

# Jordan decomposition for formal $G$ -connections



MASOUD KAMGARPOUR, SAMUEL WEATHERHOG

## Abstract

A theorem of Hukuhara, Levelt, and Turrittin states that every formal differential operator has a Jordan decomposition. This theorem was generalised by Babbit and Varadarajan to the case of formal  $G$ -connections where  $G$  is a semisimple group. In this paper, we provide straightforward proofs of these facts, highlighting the analogy between the linear and differential settings.

MSC 2010. Primary: 13N10.

## 1 Introduction

### 1.1 Overview

The Hukuhara–Levelt–Turrittin Theorem on Jordan decomposition for formal connections is a fundamental result with numerous applications in the theory of differential operators and related areas, cf. [8, 9, 20, 12, 2, 7]. The existence result was first proved by Turrittin [22], building on earlier work of Hukuhara [4]. Turrittin’s argument was rather complicated, involving nine different cases. Subsequently, Levelt gave a more conceptual (albeit still not straightforward) proof. A different (rather involved) argument was provided by Babbit and Varadarajan [1], who also considered the generalisation to semi-simple groups. Levelt noted that Jordan decomposition implies that every (formal) differential operator has an eigenvalue. He asked if one can give a direct proof of this fact, noting that this would significantly simplify the results. In this paper, we answer Levelt’s question by proving directly that every differential polynomial over  $\mathbb{C}((t))$  has a root over  $\mathbb{C}((t^{1/m}))$  for some  $m$ .

Our proof is elementary and consists of two parts: first we use Hensel’s lemma to prove that every differential polynomial over  $\mathbb{C}[[t]]$  has a root. Then we use Newton polygons to extend the result to  $\mathbb{C}((t))$ . This will immediately imply that every differential operator has an eigenvalue and leads to an easy proof of Jordan decomposition. In the last section, we establish Jordan decomposition for formal  $G$ -connections using the fact that, after a finite base change, every semisimple  $G$ -connection can be represented by an element of the Cartan subalgebra. As far as we know, this is the first time this natural result appears in the literature.

For the proof, we use some facts about differential Galois groups of formal connections (this is the only non-elementary part of our paper).

### 1.2 Differential operators and gauge transformations

Let  $\mathcal{K} := \mathbb{C}((t))$  be the field of formal Laurent series and consider the derivation  $d : \mathcal{K} \rightarrow \mathcal{K}$  defined by  $d := t \frac{d}{dt}$ . Let  $V$  be a finite dimensional vector space over  $\mathcal{K}$ . A *formal differential operator* is a  $\mathbb{C}$ -linear map  $D : V \rightarrow V$  satisfying the Leibniz rule

$$D(av) = aD(v) + d(a)v, \quad a \in \mathcal{K}, \quad v \in V. \quad (1.1)$$

It is well-known that linear operators encode linear equations. Similarly, differential operators encode (ordinary) differential equations. Thus, the study of formal differential operators is indispensable in the theory of meromorphic differential equations; see [23] for an extensive review.

In analogy with linear operators, differential operators have matrix presentations and it will be convenient to have these at our disposal. Indeed, choosing a basis for  $V$ , we can represent  $D$  as an operator  $d + A$  where  $A$  is an  $n \times n$  matrix with values in  $\mathcal{K}$ . Changing the basis by an element  $g \in GL_n(\mathcal{K})$  amounts to changing the operator  $d + A$  to  $d + g^{-1}Ag + g^{-1}dg$ . Here,  $dg$  denotes the matrix obtained by applying the derivation  $d$  to each entry of the matrix  $g$ . The map

$$A \mapsto g^{-1}Ag + g^{-1}dg$$

is called *gauge transformation* and plays an important role in the theory.

### 1.3 Semisimple Connections

To formulate a Jordan decomposition, we need a notion of semisimplicity. We start with a definition for formal differential operators.

**Definition 1.1.** *Let  $D : V \rightarrow V$  be a formal differential operator. Then  $D$  is*

- (i) *simple if  $V$  has no  $D$ -invariant subspace*

- (ii) semisimple if every  $D$ -invariant subspace has a  $D$ -invariant complement
- (iii) diagonalisable if it has a presentation of the form  $d + A$  where  $A$  is a diagonal matrix
- (iv) potentially diagonalisable if it is diagonalisable after a finite base change.

It is easy to show that an operator is semisimple if and only if it is a direct sum of simple ones. The following theorem gives an explicit description of semisimple operators.

**Theorem 1.2** (Levelt). *A formal differential operator is semisimple if and only if it is potentially diagonalisable.*

For future use, we will need the following functorial property. Let  $D : V \rightarrow V$  be a differential operator and write  $D = d + A$  with  $A \in \mathfrak{gl}(V)$ . Consider the adjoint map

$$\text{ad} : \mathfrak{gl}(V) \rightarrow \mathfrak{gl}(\mathfrak{gl}(V))$$

Then  $\text{ad}(A)$  is a linear operator on  $\mathfrak{gl}(V)$ ; therefore,  $d + \text{ad}(A)$  is a differential operator on  $\mathfrak{gl}(V)$ . The following observation will be useful.

**Proposition 1.3.** *The differential operator  $d + A$  is semisimple if and only if the differential operator  $d + \text{ad}(A) : \mathfrak{gl}(V) \rightarrow \mathfrak{gl}(V)$  is semisimple.*

## 1.4 Jordan Decomposition

We are now ready to discuss the notion of Jordan decomposition.

**Theorem 1.4** (Hukuhara-Levelt-Turrittin). *Every formal differential operator  $D$  can be written as a sum  $D = S + N$  of a semisimple differential operator  $S$  together with a nilpotent  $\mathcal{K}$ -linear operator  $N$  such that  $S$  and  $N$  commute (as  $\mathbb{C}$ -linear maps). Moreover, the pair  $(S, N)$  is unique.*

As a corollary, he concluded:

**Corollary 1.5.** *Every formal differential operator has, after an appropriate finite base change, an eigenvalue.*

Levelt asked for a direct proof of this corollary, noting that this would considerably simplify the proofs of the above theorems. Subsequently, several authors provided alternative approaches to these theorems cf. [24, 14, 18, 1, 16, 21, 10]. One of our main goals is to provide an elementary proof of the fact that every differential operator has an eigenvalue and use it to provide a simple proof of the existence of Jordan decomposition, thus fulfilling Levelt's vision.

We now provide a brief summary of our approach. Let  $\mathcal{K}\{x\}$  denote the non-commutative ring of *differential polynomials*. As an abelian group  $\mathcal{K}\{x\} = \mathcal{K}[x]$  but multiplication is modified by the rule  $xa = ax + da$  for all  $a \in \mathcal{K}$ . Using a version of Hensel's lemma and Newton polygons, we prove:

**Theorem 1.6.** *Every non-constant differential polynomial in  $\mathcal{K}\{x\}$  has a linear factorisation over a finite extension of  $\mathcal{K}$ .*

The above result is established in §2. Note that Malgrange [14] and Robba [18] also use Newton polygons and differential Hensel's lemma in their treatment of the Hukuhara-Levelt-Turrittin Theorem; however, our formulation and proof of Jordan decomposition is different from theirs; for instance, we do not use the cyclic vector lemma<sup>1</sup>.

In §3, we show that Theorem 1.2 and Corollary 1.5 follow easily from Theorem 1.6, thus illustrating the analogies between linear and differential setting. Using these results, we obtain a generalised eigenspace decomposition for differential operators. In other words, we obtain that every differential operator has a representation  $d + X$  where  $X$  is a block-upper triangular matrix and each block has a unique (up to similarity) eigenvalue. At this point, we encounter a subtle difference between the linear and differential setting. Let us write  $X = Y + Z$  where  $Y$  is diagonal and  $Z$  is strictly upper triangular. If we were considering linear operators, then  $Y$  would be the semisimple and  $Z$  the nilpotent part of  $X$  and these two commute. In the differential setting, however, the situation is more subtle because the operators  $d + Y$  and  $Z$  do not necessarily commute. In fact, these two operators commute if and only if the entries of  $Z$  are complex numbers (i.e. have no powers of  $t$ ). We prove that indeed we can arrange so that the entries of  $Z$  are complex numbers by using Katz's classification of unipotent differential operators [9].

## 1.5 Formal $G$ -connections

The above considerations have a natural generalisation to the setting of algebraic groups. Let  $G$  be a connected, semisimple, linear algebraic group over  $\mathbb{C}$  and let  $\mathfrak{g}$  denote its Lie algebra. A *formal  $G$ -connection* is an expression of the form

$$\nabla = d + A, \quad A \in \mathfrak{g}(\mathcal{K}) := \mathfrak{g} \otimes \mathcal{K}.$$

The group  $G(\mathcal{K})$  acts on the space of connections by gauge transformation

$$g \cdot (d + A) = d + \text{Ad}_g(A) + (dg)g^{-1}, \quad g \in G(\mathcal{K}), \quad A \in \mathfrak{g}(\mathcal{K}).$$

One way to make sense of the expression  $(dg)g^{-1}$  is to choose a faithful representation  $\rho : G \rightarrow \text{GL}_n$  (e.g. the adjoint representation) and show that  $d(\rho(g)) \cdot \rho(g)^{-1}$ , a priori in  $\mathfrak{gl}_n(\mathcal{K})$ , actually belongs to  $\mathfrak{g}(\mathcal{K})$ , and is independent of the chosen representation; see [1, §1.6], [3, §1.2.4], [17, §1.12].

### 1.5.1 Semisimple $G$ -connections

To discuss Jordan decomposition, we first need a notion of semisimplicity for formal  $G$ -connections. Proposition 1.3 allows us to define such a notion:

<sup>1</sup> For the advantages and disadvantages of the cyclic vector lemma, cf. [10, §5.7].

**Definition 1.7.** A  $G$ -connection  $\nabla = d + A$ ,  $A \in \mathfrak{g}(\mathcal{K})$ , is called semisimple if  $d + \text{ad}(A)$  is semisimple (as a formal  $\text{GL}(\mathfrak{g})$ -connection).

The above is analogous to the definition of ad-semisimplicity for elements in a semisimple Lie algebra, cf. [5, §5.4]. Next, let  $H \subseteq G$  be a maximal (complex) torus and  $\mathfrak{h} := \text{Lie}(H)$  the corresponding Cartan subalgebra. We then have an analogue of Theorem 1.2:

**Theorem 1.8.** A  $G$ -connection  $\nabla = d + A$  is semisimple if and only if, after a finite base change  $\mathcal{K}'/\mathcal{K}$ ,  $\nabla$  is gauge equivalent to a connection of the form  $d + X$  where  $X \in \mathfrak{h}(\mathcal{K}')$ .

As far as we know this is the first time the above natural theorem has been formulated in the literature. While the above theorem has an elementary formulation, for its proof, we use properties of the differential Galois group; see §4. This is the only part of the paper which is not “elementary”.

### 1.5.2 Jordan decomposition

We are ready to state Jordan decomposition for formal  $G$ -connections.

**Theorem 1.9** (Jordan decomposition). *Every  $G$ -connection  $\nabla = d + A$  can be written as a sum  $\nabla = S + N$ , where  $S$  is a semisimple  $G$ -connection,  $N \in \mathfrak{g}(\mathcal{K})$  is a nilpotent element and  $S$  and  $N$  commute. Moreover, the pair  $(S, N)$  is unique.*

When we say  $S$  and  $N$  commute, we mean they commute as elements of the extended loop algebra  $\hat{\mathfrak{g}} = \mathfrak{g}(\mathcal{K}) \oplus \mathbb{C}d$ , where the bracket is defined by

$$[(x \otimes p(t), \alpha.d), (y \otimes q(t), \beta.d)] := [x \otimes p(t), y \otimes q(t)] + \alpha y \otimes d(q(t)) - \beta x \otimes d(p(t)),$$

with  $x, y \in \mathfrak{g}$ ,  $p(t), q(t) \in \mathcal{K}$ ,  $\alpha, \beta \in \mathbb{C}$ .

Following a suggestion of Deligne, Babbitt and Varadarajan proved an equivalent form of the above theorem in [1]. Their proof, which uses intrinsic properties of algebraic groups, is the only proof of this fundamental result available in the literature. In this note, we give an alternative proof which uses the adjoint representation and reduces the problem to the  $\text{GL}_n$ -case. Our approach is thus similar to the standard proofs of (usual) Jordan decomposition for semisimple Lie algebras, cf. [5]. Moreover, the functoriality property of Jordan decomposition becomes transparent in our approach. We refer the reader to §4 for details.

**ACKNOWLEDGMENT:** We thank Philip Boalch, Peter McNamara, Daniel Sage, Ole Warnaar, and Sinead Wilson for helpful conversations. We are grateful to Claude Sabbah for sending copies of [14] and [19]. The material in this paper forms a part of SW’s Master’s Thesis. MK was supported by an ARC DECRA Fellowship.

## 2 Factorisation of differential polynomials

The goal of this section is to prove Theorem 1.6. This theorem should be thought of as a differential analogue of a classical theorem of Puiseux. As in Section 1, we consider the differential field  $\mathcal{K} := \mathbb{C}((t))$  with derivation  $d$ . An important implication of Puiseux’s theorem is that for every positive integer  $b$ ,  $\mathcal{K}_b := \mathbb{C}((t^{1/b}))$  is the unique extension of  $\mathcal{K}$  of degree  $b$ . The derivation  $d$  extends canonically to a derivation  $d_b$  on  $\mathcal{K}_b$ .

Let  $R$  be a  $\mathbb{C}$ -algebra and  $d : R \rightarrow R$  a derivation. We denote by  $R\{x, d\}$  the ring of differential polynomials over  $(R, d)$ . We will generally be interested in the cases  $R = \mathcal{O} := \mathbb{C}[[t]]$  and  $R = \mathcal{K} := \mathbb{C}((t))$  with derivation of the form  $\delta_m := t^m \frac{d}{dt}$ , for some positive integer  $m$ . According to [15], the ring  $\mathcal{K}\{x, \delta_m\}$  is a left and right principal ideal domain.

### 2.1 Differential Hensel’s Lemma

Let  $f \in \mathcal{O}\{x, \delta_m\}$  be a differential polynomial. We write  $f \pmod{t^n}$  for the polynomial obtained by first moving all factors of  $t$  to the left and then reducing the coefficients modulo  $t^n$ . We denote  $f \pmod{t}$  by  $\bar{f}$ . Note that this is a polynomial in  $\mathbb{C}[x]$ . Without loss of generality, we assume throughout that  $\bar{f} \neq 0$ .

Now suppose we have a factorisation of the form

$$\bar{f} = g_0 h_0, \quad g_0, h_0 \in \mathbb{C}[x].$$

Our aim is to lift this to a factorisation of  $f$  in  $\mathcal{K}\{x, \delta_m\}$ . We think of the following result as a differential analogue of Hensel’s lemma.

**Proposition 2.1.** *Let  $f \in \mathcal{O}\{x, \delta_m\}$  and  $\bar{f} = g_0 h_0$  as above. Suppose that*

$$\begin{cases} \gcd(g_0(x+n), h_0(x)) = 1, & \forall n \in \mathbb{Z}_{>0} & \text{if } m = 1 \\ \gcd(g_0(x), h_0(x)) = 1 & & \text{if } m > 1. \end{cases}$$

*Then we have a factorisation  $f = gh$  with  $g, h \in \mathcal{O}\{x, \delta_m\}$ ,  $\deg(g) = \deg(g_0)$ ,  $\bar{g} = g_0$  and  $\bar{h} = h_0$ .*

We note that a version of this proposition appeared in [16]\*Lemma 1.

*Proof.* First of all, in the differential polynomial ring  $\mathcal{K}\{x, \delta_m\}$ , easy induction arguments show that

$$h(x)t^i = t^i h(x + it^{m-1}), \quad \forall h(x) \in \mathcal{K}\{x, \delta_m\}, \quad \forall i \in \mathbb{Z}, \tag{2.1}$$

and

$$(t^d x)^k = \sum_{j=0}^{k-1} a_j t^{kd+(m-1)j} x^{k-j}, \quad \forall d \in \mathbb{Z} - \{0\}, \quad \forall k \in \mathbb{N}, \tag{2.2}$$

for some constants  $a_j \in \mathbb{C}$ ,  $a_0 = 1$ .

Our goal is to inductively build a sequence of polynomials

$$g_n(x) = g_0 + tp_1 + t^2p_2 + \cdots + t^{n-1}p_{n-1} + t^np_n, \\ p_i \in \mathbb{C}[x] \quad (2.3)$$

$$h_n(x) = h_0 + tq_1 + t^2q_2 + \cdots + t^{n-1}q_{n-1} + t^qn_n, \\ q_i \in \mathbb{C}[x], \quad (2.4)$$

which satisfy:

$$f \equiv g_n(x)h_n(x) \pmod{t^{n+1}}.$$

If we can do this, then by letting  $n \rightarrow \infty$  we will obtain elements  $g, h \in \mathcal{O}\{x, \delta_m\}$  such that  $f = gh$ .

Suppose that we know the  $p_i$  and  $q_i$  for  $1 \leq i \leq n-1$ . In view of (2.3) and (2.4) we have:

$$g_n = g_{n-1} + t^n p_n, \quad h_n = h_{n-1} + t^n q_n.$$

Requiring that  $f \equiv g_n(x)h_n(x) \pmod{t^{n+1}}$  then gives us the following condition:

$$f \equiv g_n(x)h_n(x) \pmod{t^{n+1}} \\ \equiv (g_{n-1}(x) + t^n p_n(x))(h_{n-1}(x) + t^n q_n(x)) \pmod{t^{n+1}} \\ \equiv g_{n-1}(x)h_{n-1}(x) + g_{n-1}(x)t^n q_n(x) + t^n p_n(x)h_{n-1}(x) + t^n p_n(x)t^n q_n(x) \pmod{t^{n+1}}.$$

We need to shift the powers of  $t$  to the left. By (2.1),  $g_{n-1}(x)t^n = t^n g_{n-1}(x + nt^{m-1})$ , so we have:

$$f - g_{n-1}(x)h_{n-1}(x) \equiv t^n g_{n-1}(x + nt^{m-1})q_n(x) + t^n p_n(x)h_{n-1}(x) \pmod{t^{n+1}} \\ \equiv t^n (g_{n-1}(x + nt^{m-1})q_n(x) + p_n(x)h_{n-1}(x)) \pmod{t^{n+1}},$$

and thus

$$\frac{f - g_{n-1}(x)h_{n-1}(x)}{t^n} \equiv g_{n-1}(x + nt^{m-1})q_n(x) + p_n(x)h_{n-1}(x) \pmod{t} \\ \equiv g_0(x + nt^{m-1})q_n(x) + p_n(x)h_0(x) \pmod{t}.$$

For notational convenience, we set:

$$f_n = \frac{f - g_{n-1}(x)h_{n-1}(x)}{t^n}.$$

so that we have

$$f_n \equiv g_0(x + nt^{m-1})q_n(x) + p_n(x)h_0(x) \pmod{t}. \quad (2.5)$$

Now if  $m > 1$ , then (2.5) reduces to

$$f_n \equiv g_0(x)q_n(x) + p_n(x)h_0(x) \pmod{t}.$$

Since  $\mathbb{C}[x]$  is a Euclidean domain, we will be able to solve this for  $p_n$  and  $q_n$  provided that  $g_0$  and  $h_0$  are coprime. On the other hand, if  $m = 1$ , then (2.5) becomes

$$f_n \equiv g_0(x+n)q_n(x) + p_n(x)h_0(x) \pmod{t}.$$

In this case, we will only be able to generate the entire sequence if  $g_0(x+n)$  and  $h_0(x)$  are coprime for all  $n \in \mathbb{Z}_{>0}$ .

All that remains to show is that we can control the degree of the  $g_n$ 's. We will show this in the case  $m = 1$ . The proof in the case  $m > 1$  is similar (replace  $g_0(x+n)$  with  $g_0(x)$  everywhere). Since  $g_0(x+n)$  and  $h_0(x)$  are coprime, we can find  $a, b \in \mathbb{C}[x]$  such that

$$g_0(x+n)a(x) + h_0(x)b(x) = 1.$$

Multiplying through by  $f_n$  yields

$$g_0(x+n)a(x)f_n(x) + h_0(x)b(x)f_n(x) = f_n(x). \quad (2.6)$$

Using the division algorithm we can find unique  $p_n$  and  $q_n$  such that  $\deg(p_n) < \deg(g_0)$ . Write:

$$b(x)f_n(x) = Q(x)g_0(x) + R(x)$$

with  $\deg(R) < \deg(g_0)$ . Equation (2.6) then becomes:

$$g_0(x+n)(a(x)f_n(x) + Q(x)h_0(x)) + h_0(x)R(x) \equiv f_n(x) \pmod{t}.$$

Setting  $p_n = R$  and  $q_n = af_n + Qh_0$  gives us the required  $g_n$  and  $h_n$ .  $\square$

**Corollary 2.2.** *Let  $f \in \mathcal{O}\{x, \delta_1\}$  be a monic differential polynomial. Then  $f$  admits a factorisation of the form*

$$(x - \Lambda)h,$$

with  $\Lambda \in \mathcal{O}$  and  $h \in \mathcal{O}\{x, \delta_1\}$ .

*Proof.* Let  $\bar{f} \in \mathbb{C}[x]$  be the reduction of  $f \pmod{t}$ . Since  $f$  is monic,  $\bar{f}$  is non-constant and hence factors over  $\mathbb{C}$  into linear factors:

$$\bar{f} = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_n), \quad \lambda_i \in \mathbb{C}.$$

Without loss of generality, we can order these factors so that  $\text{Re}(\lambda_1) \leq \text{Re}(\lambda_2) \leq \cdots \leq \text{Re}(\lambda_n)$ . With this ordering we then have

$$\bar{f} = g_0 h_0,$$

where

$$g_0 = x - \lambda_1, \quad h_0 = (x - \lambda_2) \cdots (x - \lambda_n).$$

By our choice of ordering,  $g_0(x+n)$  has no common factor with  $h_0$  for all  $n \in \mathbb{Z}_{>0}$ . Hence we can apply Proposition 2.1 to obtain a factorisation of the form

$$f = (x - \Lambda)h, \quad \Lambda \in \mathcal{O}, h \in \mathcal{O}\{x, \delta_1\},$$

as required.  $\square$

*Remark 2.3.* Note that the above result is false for the usual polynomial ring  $\mathcal{O}[x]$ . Indeed,  $x^2 + t - t^2$  does not have a linear factorisation over this ring, but if we consider it as an element of  $\mathcal{O}\{x, \delta_1\}$ , then  $x^2 + t - t^2 = (x - t)(x + t)$ .

## 2.2 From power series to Laurent series via Newton polygons

In the previous section, we settled linear factorisation for differential polynomials in  $\mathcal{O}\{x, \delta_1\}$ . In this section, we explain how, by a change of variable, we can transform polynomials with coefficients in  $\mathcal{K}$  to those with power series coefficients. The price is that we have to go to a finite extension  $\mathcal{K}_q$  of  $\mathcal{K}$  and, more seriously, the derivation is not simply the canonical extension of  $\delta_1$  to  $\mathcal{K}_q$ . Nevertheless, we shall see that this change of variable allows us to factor elements of  $\mathcal{K}\{x, \delta_1\}$ . Throughout we let  $v_t(\cdot)$  denote the  $t$ -adic valuation on  $\mathcal{K}$ .

**Lemma 2.4.** *Consider the monic differential polynomial  $f(x) = x^n + \sum_{i=1}^n a_i x^{n-i} \in \mathcal{K}\{x, \delta_1\}$ . Let  $r := \min\left\{\frac{v_t(a_i)}{i}\right\}$ . Then  $g(X) = t^{-nr} f(t^r X)$  is a monic differential polynomial with power series coefficients.*

*Proof.* To be more precise, write  $r = \frac{p}{q}$  with  $\gcd(p, q) = 1$  and  $q > 0$ . If  $r \geq 0$ , then each  $v_t(a_i) \geq 0$  and so  $f \in \mathcal{O}\{x, \delta_1\}$ . Since we have already dealt with this case in Corollary 2.2, we may assume that  $r < 0$ . In order to make the change of variables  $x = t^r X$ , we require a field extension to  $\mathcal{K}_q = \mathbb{C}((t^{1/q}))$ . Let  $s := t^{1/q}$  so that our change of variables becomes  $x = s^p X$ . Note that this change of variables means that the relation  $xt = tx + t$  becomes  $Xs^q = s^q X + s^{q-p}$ . Hence differential polynomials in  $X$  lie in the ring  $\mathcal{K}_q\{X, \frac{1}{q}s^{1-p}\frac{d}{ds}\}$  (note this new derivation sends  $s^q$  to  $s^{q-p}$ ).

Applying (2.2) to  $f(s^p X)$  yields  $f(s^p X) = s^{np}g(X)$  where

$$g(X) = s^{-np}a_n + \sum_{k=0}^{n-1} a_k \sum_{j=0}^{n-1-k} m_{n-k,j} s^{(-j-k)p} X^{n-1-k-j},$$

$$a_0 = 1.$$

Let  $v_s(\cdot)$  denote the  $s$ -adic valuation on  $\mathcal{K}_q$ . Since  $v_s(a_i) = qv_t(a_i)$ ,  $v_t(a_i) \geq \frac{ip}{q}$  implies that  $v_s(a_i) \geq ip$ . Thus, for  $0 \leq l \leq n-1$ , the coefficient,  $b_l$ , of  $X^{n-l}$  in  $g$  satisfies

$$v_s(b_l) = \min_{0 \leq k \leq l} \{v_s(a_k s^{-lp})\} \geq \min_{0 \leq k \leq l} \{kp - lp\} = 0,$$

where the last equality follows since  $p < 0$ .

It is clear that  $v_s(b_l)$  will be 0 exactly when  $v_s(a_l) = lp$ , that is, if, and only if,  $v_t(a_l) = lr$ . For the ‘‘constant’’ term of  $g$  we have

$$v_s(b_n) = v_s(a_n s^{-np}) \geq np - np = 0,$$

again with equality exactly when  $v_t(a_n) = nr$ . Thus

$$g(X) = X^n + b_1 X^{n-1} + \dots + b_n, \quad b_i \in \mathbb{C}[[s]],$$

with  $\min\{v_s(b_i)\} = 0$ . Furthermore,  $v_s(b_i) = 0$  if, and only if,  $v_t(a_i) = ir$ . This shows that  $g(X) \in \mathbb{C}[[s]]\{X, \frac{1}{q}s^{1-p}\frac{d}{ds}\}$ .  $\square$

Consider  $g(x)$  from the above lemma. If  $\bar{g}(x)$  has two distinct roots, then Hensel’s lemma allows us to factor it. We now study the opposite extreme, i.e., when all roots of  $\bar{g}(x)$  are equal. It will be helpful to use the notion of Newton polygons for differential polynomials, cf. [10, §6.4]. Throughout the rest of this section, we will assume that  $r < 0$ , unless explicitly stated otherwise.

**Definition 2.5** (Newton Polygon). *Let  $f \in \mathcal{K}\{x, \delta_m\}$  be a differential polynomial and write*

$$f(x) = \sum_{i=0}^n a_i x^{n-i}, \quad a_i \in \mathcal{K}.$$

*Consider the lower boundary of the convex hull of the points*

$$\{(n-i), v_t(a_i) : 0 \leq i \leq n\} \subset \mathbb{R}^2.$$

*The Newton polygon of  $f$ , denoted  $NP(f)$ , is obtained from this boundary by replacing all line segments of slope less than  $1-m$  with a single line segment of slope exactly  $1-m$ .*

Lemma 2.4 now has the following corollary.

**Corollary 2.6.** *Let  $f$  and  $g$  be as in Lemma 2.4 and suppose that  $\bar{g} := g \pmod{s} = (X + \lambda)^n$ ,  $\lambda \in \mathbb{C}$ . Then  $\lambda$  is non-zero and the Newton polygon of  $f$  has a single integral slope.*

*Proof.* As in Lemma 2.4, write

$$g(X) = X^n + b_1 X^{n-1} + \dots + b_n, \quad b_i \in \mathbb{C}[[s]].$$

Since  $\min\{v_s(b_i)\} = 0$ ,  $\lambda \neq 0$ . Now since,  $\lambda \neq 0$ , expanding the bracket  $(X + \lambda)^n$  shows that  $v_s(b_i) = 0$  for all  $i$  and hence  $v_t(a_i) = ir$ . Thus, the Newton polygon of  $f$  has a single slope of  $-r$  and since  $v_t(a_1) = r$ ,  $r$  is an integer.  $\square$

For future use, we also record the following lemma.

**Lemma 2.7.** *Let  $f$  and  $g$  be as in Lemma 2.4 and suppose that  $\bar{g} = (X + \lambda)^n$ ,  $\lambda \in \mathbb{C}$ . Then the slopes of the Newton polygon of  $f(x - \lambda t^r)$  are all strictly smaller than the slope of the Newton polygon of  $f(x)$ .*

*Proof.* By Corollary 2.6,  $r$  is an integer and hence no extension of  $\mathcal{K}$  is necessary. Since  $\bar{g} = (X + \lambda)^n$ , we can write  $g$  as

$$g = (X + \lambda)^n + e_1(X + \lambda)^{n-1} + \dots + e_n, \quad e_i \in \mathcal{O},$$

with  $v_t(e_i) > 0$  for all  $i$ . Now

$$f(t^r X) = t^{nr} \left( (X + \lambda)^n + e_1(X + \lambda)^{n-1} + \dots + e_n \right)$$

and so

$$f(x) = t^{nr} \left( (t^{-r}x + \lambda)^n + e_1(t^{-r}x + \lambda)^{n-1} + \dots + e_n \right).$$

Hence

$$f(x - \lambda t^r) = t^{nr} \left( (t^{-r}x)^n + e_1(t^{-r}x)^{n-1} + \dots + e_n \right).$$

Applying (2.2), we have, for  $m_{k,l} \in \mathbb{C}$ ,

$$\begin{aligned} f(x - \lambda t^r) &= t^{nr} \left( t^{-nr} \sum_{j=0}^{n-1} m_{n,j} x^{n-j} \right. \\ &\quad \left. + e_1 t^{-(n-1)r} \sum_{j=0}^{n-2} m_{n-1,j} x^{n-1-j} + \dots + e_n \right) \\ &= \sum_{j=0}^{n-1} m_{n,j} x^{n-j} \\ &\quad + e_1 t^r \sum_{j=0}^{n-2} m_{n-1,j} x^{n-1-j} + \dots + t^{nr} e_n. \end{aligned}$$

Since  $v(e_i) > 0$ , the valuation of the coefficient of  $x^{n-j}$  in  $f(x - \lambda t^r)$  is strictly greater than the corresponding coefficient in  $f(x)$ . This means that the slopes of the Newton polygon for  $f(x - \lambda t^r)$  are strictly less than the slope of the Newton polygon for  $f(x)$ .  $\square$

*Example 2.8.* In order to illustrate Corollary 2.6 and Lemma 2.7, consider the differential polynomial

$$f_1(x) = x^2 + (4t^{-2} + 2t^{-1} + 2)x + (4t^{-4} + 4t^{-3} + t^{-2} + t^{-1} + 1).$$

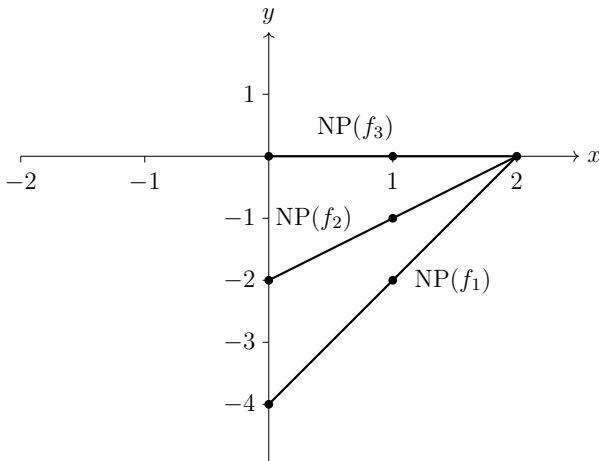
In this case,  $r = -2$  and the change of variables  $x = t^{-2}X$  yields

$$g_1(X) = X^2 + (4 + 2t)X + (4 + 4t + t^2 + t^3 + t^4),$$

and so  $\bar{g}_1(X) = (X + 2)^2$ . The figure below shows that the Newton polygon of  $f_1$  has only a single slope of  $-2$  (cf. Cor 2.6). Making the translation  $x \mapsto x + 2t^{-2}$  as in Lemma 2.7 yields the new polynomial

$$f_2(x) = x^2 + (2t^{-1} + 2)x + (t^{-2} + t^{-1} + 1).$$

This has a single slope  $r = -1$  and a final translation  $x \mapsto x - t^{-1}$  yields  $f_3(x) = x^2 + 2x + 1$ . This can easily be factorised and reversing the change of variables yields the full factorisation  $f_1(x) = (x + 2t^{-2} + t^{-1} + 1)^2$ .



### 2.3 Proof of Theorem 1.6

Write  $f(x) = x^n + a_1x^{n-1} + \dots + a_n \in \mathcal{K}\{x, \delta_1\}$  and let  $r := \min \left\{ \frac{v_t(a_i)}{i} \right\} \in \mathbb{Q}$ . If  $r \geq 0$ , then the result follows from the differential Hensel's Lemma (see Corollary 2.2) so we may assume  $r < 0$ . Let us write

$$r = \frac{p}{q}, \quad q > 0, \quad \gcd(p, q) = 1.$$

Consider the transformation  $x \mapsto t^r X$ . Under this transformation the differential field  $(\mathcal{K}, \delta_1)$  changes to  $(\mathcal{K}_q, \frac{1}{q}s^{1-p} \frac{d}{ds})$  where  $s := t^{1/q}$ . Moreover, we obtain a monic differential polynomial  $g(X) \in \mathbb{C}[[s]]\{y, \frac{1}{q}s^{1-p} \frac{d}{ds}\}$ . Let  $\bar{g}(X)$  denote the reduction of  $g(X)$  modulo the maximal ideal of  $\mathbb{C}[[s]]$ . If  $\bar{g}(X)$  has two distinct roots, then we can again apply Proposition 2.1 to reduce the problem to a polynomial of degree strictly less than  $f$ . Thus, we are reduced to the case that  $\bar{g}(X)$  has a unique repeated root  $\lambda$ . For inductive purposes, we rename  $f$  to  $f_1$ . In this case, by Corollary 2.6,  $\lambda \neq 0$  and the Newton polygon of  $f_1$  has a single *integral* slope. Now we make the transformation  $x \mapsto x - \lambda t^r$ . As shown in Lemma 2.7, under this transformation  $f_1$  is mapped to a polynomial  $f_2$  whose Newton polygon has slopes strictly less than that of  $f_1$ . Note that this transformation does not change the differential field.

Now we start the process with the polynomial  $f_2(x) := x^n + b_1x^{n-1} + \dots + b_n \in \mathcal{K}\{x, \delta_1\}$ ; i.e., we let  $r_2 := \min \left\{ \frac{v(b_i)}{i} \right\}$ . If  $r_2 \geq 0$  we are done. Otherwise, we make the change of variable  $x \mapsto t^{r_2}X$  to obtain a new polynomial  $g_2(X)$ . If  $\bar{g}_2(X)$  has distinct roots, then we are done; otherwise, applying Corollary 2.6 again, we conclude that the Newton polygon of  $f_2$  has a single *integral* slope. Since the slope of  $f_2$  is a nonnegative integer strictly less than slope of  $f_1$ , this process must stop in finitely many steps at which point we have a factorisation of our polynomial.  $\square$

## 3 Proof of Jordan decomposition for differential operators

Recall that for each positive integer  $b$ ,  $\mathcal{K}_b$  denotes the unique finite extension of  $\mathcal{K}$  of degree  $b$ . Given a differential operator  $D$ , one has a canonical differential operator

$$D \otimes_{\mathcal{K}} \mathcal{K}_b : V \otimes_{\mathcal{K}} \mathcal{K}_b \rightarrow V \otimes_{\mathcal{K}} \mathcal{K}_b$$

called the base change of  $D$  to  $\mathcal{K}_b$ . All base changes considered in this article are of this form. Henceforth, we will use the notation  $V_b := V \otimes_{\mathcal{K}} \mathcal{K}_b$  and  $D_b = D \otimes_{\mathcal{K}} \mathcal{K}_b$ .

### 3.1 Proof of Corollary 1.5 (Every differential operator has an eigenvalue)

The argument proceeds exactly as in the linear setting. Let  $D : V \rightarrow V$  be a differential operator and

$v \in V$  be a non-zero vector. Consider the sequence  $v, D(v), D^2(v), \dots$ . As  $V$  has finite dimension over  $\mathcal{K}$ , we must have that

$$D^n(v) + a_1 D^{n-1}(v) + \dots + a_{n-1} D(v) + a_n v = 0, \quad a_i \in \mathcal{K},$$

where  $n = \dim_{\mathcal{K}}(V)$ . Now consider the polynomial  $f(x) = x^n + a_1 x^{n-1} + \dots + a_n$  in the twisted polynomial ring  $\mathcal{K}\{x\}$ . After a finite extension, we can write

$$f(x) = (x - \Lambda_1) \cdots (x - \Lambda_n) \in \mathcal{K}_b\{x\}, \quad \Lambda_i \in \mathcal{K}_b, b \in \mathbb{Z}_{>0}.$$

Thus,

$$(D_b - \Lambda_1) \cdots (D_b - \Lambda_n)v = 0.$$

Let  $i \in \{1, 2, \dots, n\}$  be the largest number such that  $(D_b - \Lambda_i) \cdots (D_b - \Lambda_n)v = 0$ . If  $i = n$ , then  $v$  is an eigenvector of  $D_b$  with eigenvalue  $\Lambda_n$ . Otherwise  $(D_b - \Lambda_{i+1}) \cdots (D_b - \Lambda_n)v$  is an eigenvector of  $D_b$  with eigenvalue  $\Lambda_i$ .  $\square$

### 3.2 Proof of Theorem 1.2 (Semisimple operators are diagonalisable)

We need the following lemma. The proof is an easy argument using the Galois group  $\text{Gal}(\mathcal{K}_b/\mathcal{K})$ ; see [11, §1(e)] for details.

**Lemma 3.1.**  *$D$  is semisimple if and only if  $D_b$  is.*

Now we are ready to prove Theorem 1.2. Suppose  $D$  is semisimple. We prove by induction on  $\dim(V)$  that, after an appropriate base change, it is diagonalisable. If  $\dim(V) = 1$  the result is obvious, so assume  $\dim(V) > 1$ . Without loss of generality, assume  $D$  has an eigenvector  $v$  (if not, do an appropriate base change; by the previous lemma, the operator remains semisimple). Let  $U = \text{span}_{\mathcal{K}}\{v\}$ . Then  $U$  is a one-dimensional,  $D$ -invariant subspace<sup>2</sup> of  $V$ ; thus, there exists a  $D$ -invariant complement  $W$ . Now  $D : W \rightarrow W$  is semisimple so by our induction hypothesis (after an appropriate base change), we can write  $W$  as a direct sum of one-dimensional subspaces. Thus, after an appropriate base change, we have a decomposition of our vector space into one-dimensional, invariant subspaces and so  $D$  is diagonalisable.

Conversely, suppose  $D_b$  is a diagonalisable operator. Then clearly  $D_b$  is semisimple and thus, by Lemma 3.1, so is  $D$ .  $\square$

### 3.3 Invariant Properties of Differential Operators

The goal of this section is to prove Proposition 1.3. To this end, we need to establish some properties of differential operators.

#### 3.3.1 Invariant Subspaces

**Lemma 3.2.** *Let  $D : V \rightarrow V$  be a differential operator with Jordan decomposition  $D = S + N$ . Suppose that  $W \subset V$  is a  $D$ -invariant subspace. Then  $W$  is also  $S$ -invariant.*

*Proof.* Note that  $V$  decomposes into generalised eigenspaces and that these generalised eigenspaces are  $D$ ,  $S$  and  $N$  invariant [11]\*§4. Hence we need only consider the case where  $V$  itself is a generalised eigenspace. In this case, there exists a finite extension,  $\mathcal{K}_b$ , of  $\mathcal{K}$  such that  $S = d + \lambda I$  for some  $\lambda \in \mathcal{K}_b$ . We first prove the result in the case of unipotent differential operators (i.e. in the case  $\lambda = 0$ ). As in Section 3.5, we denote by  $\mathcal{U}$  the category of unipotent differential operators. Recall this category is equivalent to the category Nilp whose objects are pairs  $(V_0, N)$  where  $V_0$  is a  $\mathbb{C}$ -vector space and  $N$  is a nilpotent linear operator.

The restriction  $D|_W : W \rightarrow W$  gives us a monomorphism in the category  $\mathcal{U}$ . Under the equivalence  $F$  we obtain a monomorphism in Nilp. Hence, there is a basis of  $V$  for which we can write  $D = d + N$ . Since  $(W_0, N') \hookrightarrow (V_0, N)$ , in this basis we have  $dW = W$ . That is,  $W$  is  $S$ -invariant.

This result clearly extends to differential operators with a unique (up to similarity) eigenvalue.

For the general case, recall that after a finite extension to  $\mathcal{K}_b$ , we can write  $D = S + N$  where  $S$  is diagonalisable. Now  $W_b$  is a  $D_b$ -invariant subspace of  $V_b$  and so by the above,  $W_b$  is also  $S_b$ -invariant. If  $W$  were not  $S$ -invariant, then  $W_b$  would not be  $S_b$ -invariant, hence  $W$  must be  $S$ -invariant.  $\square$

#### 3.3.2 Adjoint differential operator

Let  $V$  be a finite dimensional vector space over  $\mathcal{K}$ . Given a differential operator  $d + A : V \rightarrow V$ , we write  $d + S + N$  for its Jordan decomposition. Note that  $S$  is not necessarily a semisimple linear operator on  $V$ ; rather,  $d + S$  is a semisimple differential operator on  $V$ .

**Lemma 3.3.** *Let  $d + A : V \rightarrow V$  be a differential operator, where  $A \in \mathfrak{gl}(V)$ . Then  $d + \text{ad } S + \text{ad } N$  is the Jordan decomposition of  $d + \text{ad } A$ .*

*Proof.* There exists a finite extension  $\mathcal{K}_b$  of  $\mathcal{K}$  such that we can pick a basis for  $V \otimes \mathcal{K}_b$  to put  $d + A$  in Jordan normal form. In this case,  $S$  is diagonal and  $N$  is a constant nilpotent matrix with 1's or 0's on the super-diagonal, and  $S$  and  $N$  commute. Thus,  $d + \text{ad}(S)$  is a semisimple differential operator on  $\mathfrak{gl}(V \otimes \mathcal{K}_b)$ . We claim that it commutes with  $\text{ad}(N)$ . Indeed,

$$[d + \text{ad}(S), \text{ad}(N)] = [d, \text{ad}(N)] + [\text{ad}(S), \text{ad}(N)],$$

where the bracket is for the extended Lie algebra  $\widehat{\mathfrak{gl}(\mathfrak{g})}$ . Now  $\text{ad}(N)$  is constant, so the first bracket is zero. Since  $S$  and  $N$  commute, the second bracket is also zero.  $\square$

<sup>2</sup> Indeed, if  $Dv = \lambda v$ ,  $\lambda \in \mathcal{K}$ , and  $a \in \mathcal{K}$  then  $D(av) = (a\lambda + d(a))v \in U$ .

### 3.3.3 Proof of Proposition 1.3

If  $d + A$  is semisimple, then we have seen that so is  $d + \text{ad}(A)$ . If  $d + A$  is not semisimple, then suppose  $d + S + N$  is its Jordan decomposition. By assumption,  $N \neq 0$ . This implies that  $\text{ad} N$  is not trivial. Thus,  $d + \text{ad} A$  is not semisimple.

### 3.4 Generalised eigenspace decomposition

Let  $D : V \rightarrow V$  be a formal differential operator and let  $a \in \mathcal{K}$ .

**Definition 3.4.** The generalised eigenspace  $V(a)$  of  $D$  is defined as

$$V(a) := \text{span}_{\mathcal{K}}\{v \in V \mid \exists n \geq 1, (D - a)^n v = 0\}$$

The goal of this section is to prove the following theorem.

**Theorem 3.5** (Generalised eigenspace decomposition). *For some finite extension  $\mathcal{K}_b$  of  $\mathcal{K}$  there exists a canonical decomposition  $V_b = \bigoplus_i V_b(a_i)$ . Moreover,*

$$\begin{aligned} V_b(a_i) \cap V_b(a_j) \neq \{0\} &\iff a_i \text{ is similar to } a_j \\ &\iff V_b(a_i) = V_b(a_j). \end{aligned}$$

Before proving this theorem, we need to recall some facts about differential operators. Let  $D : V \rightarrow V$  be a differential operator. Define

$$H^0(V) := \ker(D),$$

$$H^1(V) := V/D(V).$$

Note that these are vector spaces over  $\mathbb{C}$  (not over  $\mathcal{K}$ ). The following proposition due to Malgrange [13, Theorem 3.3] is an analogue of the rank-nullity theorem for formal differential operators.

**Proposition 3.6.** *Let  $D : V \rightarrow V$  be a formal differential operator. Then*

$$\dim_{\mathbb{C}} H^0(V) = \dim_{\mathbb{C}} H^1(V).$$

Next, recall that the dual differential operator  $D : V \rightarrow V$  is the operator  $D^*$  on the vector space  $V^* = \text{Hom}_{\mathcal{K}}(V, \mathcal{K})$  defined by

$$D^* : V^* \rightarrow V^*, \quad D^*(f) = d \circ f - f \circ D, \quad f \in V^*.$$

Let  $D : V \rightarrow V$  and  $D' : V' \rightarrow V'$  be differential operators. Then, we can define a differential operator  $D \otimes D'$  on  $V \otimes V'$  by

$$(D \otimes D')(v \otimes v') := D(v) \otimes v' + v \otimes D'(v').$$

The set of all of  $\mathcal{K}\{x\}$ -linear maps from  $V$  to  $V'$  is denoted  $\text{Hom}_{\mathcal{K}\{x\}}(V, V')$ . This is a  $\mathbb{C}$ -vector space.

The Yoneda extension group  $\text{Ext}_{\mathcal{K}\{x\}}^1(V, V')$  consists of equivalence classes of extensions of  $\mathcal{K}\{x\}$ -modules

$$0 \rightarrow V \rightarrow V'' \rightarrow V' \rightarrow 0$$

As usual, two extensions are equivalent if there exists a  $\mathcal{K}\{x\}$ -linear isomorphism between them inducing the identity on  $V$  and  $V'$ .

**Proposition 3.7.** *Let  $D : V \rightarrow V$  and  $D' : V' \rightarrow V'$  be two formal differential operators. Then, we have*

$$(i) \dim_{\mathbb{C}} \text{Ext}_{\mathcal{K}\{x\}}^1(V, V') = \dim_{\mathbb{C}} H^0(V^* \otimes V').$$

(ii) *If no eigenvalue of  $D$  is similar to an eigenvalue of  $D'$ , then  $\text{Ext}_{\mathcal{K}\{x\}}^1(V, V') = 0$ .*

*Proof.* One can show (see [10]\*Lemma 5.3.3) that there is a canonical isomorphism of  $\mathbb{C}$ -vector spaces:

$$\text{Ext}_{\mathcal{K}\{x\}}^1(V, V') \simeq H^1(V^* \otimes V').$$

This fact together with Proposition 3.6 implies (i).

The eigenvalues of  $D^* \otimes D'$  are of the form  $-a + a'$  where  $a$  and  $a'$  are eigenvalues of  $D$  and  $D'$ , respectively. By assumption,  $-a + a'$  is never similar to zero; thus, kernel of  $D^* \otimes D'$  is trivial. Part (ii) now follows from Part (i).  $\square$

*Proof of Theorem 3.5.* We may assume, without the loss of generality, that all eigenvalues of  $D$  are already in  $\mathcal{K}$  (if not, do an appropriate base change). We use induction on  $\dim(V)$  to prove the theorem. If  $\dim(V) = 1$  then the claim is trivial. Suppose  $\dim(V) > 1$ . Then by assumption  $D$  has an eigenvector. Hence, we have a one-dimensional invariant subspace  $U \subset V$ . Let  $W := V/U$ . Then  $D$  defines a differential operator on  $W$ . Moreover,  $V \in \text{Ext}_{\mathcal{K}\{x\}}^1(U, W)$ . By induction we may assume that  $W$  decomposes as

$$W = \bigoplus_i W(a_i), \quad a_i \in \mathcal{K},$$

for non-similar  $a_i$ . Now

$$V \in \text{Ext}_{\mathcal{K}\{x\}}^1\left(U, \bigoplus_i W(a_i)\right) \simeq \bigoplus_i \text{Ext}_{\mathcal{K}\{x\}}^1(U, W(a_i)).$$

If the eigenvalue  $a$  of  $D|_U$  is not similar to any  $a_i$  then by the above proposition all the extension groups are zero, and so  $V = W \oplus U$  and the theorem is established. If  $a$  is similar to  $a_j$ , for some  $j$ , then the only non-trivial component in the above direct sum is  $\text{Ext}_{\mathcal{K}\{x\}}^1(U, W(a_j))$ . But it is easy to see that all differential operators in  $\text{Ext}_{\mathcal{K}\{x\}}^1(U, W(a_j))$  have only a single eigenvalue  $a_j$  (up to similarity). Hence  $V$  has the required decomposition.  $\square$



### 3.5 Unipotent differential operators

Theorem 3.5 implies that we only need to prove Jordan decomposition for differential operators with a unique eigenvalue. By translating if necessary, we can assume this eigenvalue is zero. Thus, we arrive at the following:

**Definition 3.8** (Unipotent Operators). *A differential operator is unipotent if all of its eigenvalues are similar to zero.*

We now give a complete description of unipotent differential operators. Let  $\text{Nilp}_{\mathbb{C}}$  denote the category whose objects are pairs  $(V, N)$  where  $V$  is a  $\mathbb{C}$ -vector space and  $N$  is a nilpotent endomorphism. The morphisms of  $\text{Nilp}_{\mathbb{C}}$  are linear maps which commute with  $N$ . Let  $\mathcal{U}$  be the category of pairs  $(V, D)$  consisting of a vector space  $V/\mathcal{K}$  and a unipotent differential operator  $D : V \rightarrow V$ . Define a functor

$$F : \text{Nilp}_{\mathbb{C}} \rightarrow \mathcal{U}, \quad (V, N) \mapsto (\mathcal{K} \otimes_{\mathbb{C}} V, d + N).$$

The following result appears (without proof) in [9, §2].

**Lemma 3.9.** *The functor  $F$  defines an equivalence of categories with inverse given by*

$$G : \mathcal{U} \rightarrow \text{Nilp}_{\mathbb{C}}, \quad (V, D) \mapsto \left( \ker(D^{\dim_{\mathcal{K}}(V)}), D \right).$$

*Proof.* We first show that the composition  $G \circ F$  equals the identity. Let  $(V, N) \in \text{Nilp}_{\mathbb{C}}$  with  $n := \dim(V)$  and consider  $F(V, N) = (V \otimes \mathcal{K}, d + N)$ . The kernel of the operator  $(d + N)^n$  acting on  $V \otimes \mathcal{K}$  is the set of all constant vectors. This is an  $n$ -dimensional  $\mathbb{C}$ -vector space. Since  $d$  acts as 0 on this space, applying  $G$  to  $(\mathcal{K} \otimes V, d + N)$  recovers the pair  $(V, N)$ .

Next, let  $D : V \rightarrow V$  be a unipotent differential operator and let  $n := \dim_{\mathcal{K}}(V)$ . We first show by induction that  $\ker(D^n)$  contains  $n$ ,  $\mathcal{K}$ -linearly independent vectors. If  $n = 1$  this is obvious. If  $n > 1$ , then there exists  $v \in V$  such that  $Dv = 0$ . Set  $U := \text{span}_{\mathcal{K}}\{v\}$  and consider the differential module  $V/U$ . This has dimension  $n - 1$  so we may assume there exist  $\{v_1, \dots, v_n\}$   $\mathcal{K}$ -linearly independent vectors in  $\ker(D^{n-1})$ . For each  $v_i$  we have  $D^{n-1}v_i + U = U$  and hence  $D^{n-1}v_i = a_i v$  for some  $a_i \in \mathcal{K}$ . Now observe that we can choose  $b_i$  such that  $d^{n-1}(b_i) = a_i - a_{i,0}$  where  $a_{i,0}$  is the constant term of  $a_i$ ; since we can always “integrate” elements with no constant term. Now we have

$$\begin{aligned} D^{n-1}(v_i - b_i v) &= D^{n-1}v_i - D^{n-1}(b_i v) \\ &= a_i v - \sum_{j=0}^{n-1} \binom{n-1}{j} d^j(b_i) D^{n-1-j}(v) \\ &= a_i v - d^{n-1}(b_i)v = a_{i,0}v. \end{aligned}$$

Hence

$$D^n(v_i - b_i v) = D(a_{i,0}v) = 0$$

so  $\{v, v_1 - b_1 v, \dots, v_{n-1} - b_{n-1} v\}$  is a set of  $\mathcal{K}$ -linearly independent vectors in  $\ker(D^n)$ .

Note the functor  $G$  sends  $V$  to the  $\mathbb{C}$ -vector space  $W := \ker(D^n) = \text{span}_{\mathbb{C}}\{v, v_1 - b_1 v, \dots, v_{n-1} - b_{n-1} v\}$ . Moreover,  $D$  induces a  $\mathbb{C}$ -linear operator  $N$  on  $W$ . By construction, this operator is nilpotent and for this basis, the matrix of  $N$  is constant (i.e., its entries belong to  $\mathbb{C}$ ). Applying the functor  $F$  to  $(W, N)$  now recovers the differential module  $(V, D)$ .  $\square$

*Remark 3.10.* A formal differential operator  $D$  is said to be *regular singular* if it has a matrix representation of the form

$$A_0 + A_1 t + \dots, \quad A_i \in \mathfrak{gl}_n(\mathbb{C}).$$

It is known that, in this case,  $D$  can actually be represented by a constant matrix; i.e., by a matrix  $A \in \mathfrak{gl}_n(\mathbb{C})$ . The conjugacy class of  $A$  is uniquely determined by  $D$  and is called the *monodromy* [1, §3]. The above lemma implies that a unipotent differential operator is the same as a regular singular differential operator with unipotent monodromy.

### 3.6 Proof of Theorem 1.4 (Jordan Decomposition)

The uniqueness part of the theorem is relatively easy. Since we don't have anything new to add to Levelt's original proof, we refer the reader to [11] for the details. It remains to prove existence.

Let  $D : V \rightarrow V$  be a formal differential operator. By Theorem 3.5, there exists a positive integer  $b$  such that  $D_b : V_b \rightarrow V_b$  admits a generalised eigenspace decomposition. Thus,  $D_b$  can be represented by a block diagonal matrix where each block is upper triangular with a unique (up to similarity) eigenvalue. Thus, we may assume without the loss of generality that  $D_b$  has a unique, up to similarity, eigenvalue  $a$ . Replacing  $D_b$  by  $D_b - a$ , we may assume that  $D_b$  is unipotent in which case the result follows from Lemma 3.9. This proves the existence of Jordan decomposition for  $D_b$ .

We now show that the Jordan decomposition of  $D_b$  descends to a decomposition of  $D$ . The proof is similar to the linear setting. Picking a  $\mathcal{K}$ -basis of  $V$  and extending it to a basis of  $V_b$  allows us to write  $D_b = d + A$  where  $A$  is a matrix with entries in  $\mathcal{K}$ . Let  $S_b = d + B$  and  $N_b = C$  for matrices  $B$  and  $C$  with respect to this basis. Then, for any  $\sigma \in \text{Gal}(\mathcal{K}_b/\mathcal{K})$ , it is clear that  $d + A = d + \sigma(B) + \sigma(C)$  is a second Jordan decomposition of  $D_b$ . Thus, we must have  $C = \sigma(C)$  and  $\sigma(B) = B$ . Hence,  $d + B$  and  $C$  are defined over  $\mathcal{K}$ .  $\square$

## 4 Proof of Jordan decomposition for $G$ -connections

### 4.1 How to think about $G$ -connections?

There are (at least) two ways of thinking about formal  $G$ -connections. The first approach, discussed in the introduction, involves considering gauge equivalence classes

of expressions of the form  $d + A$ ,  $A \in \mathfrak{g}(\mathcal{K})$ . The second approach involves considering the differential Galois group  $I_{\mathcal{K}}$ . For the benefit of the reader, we include a brief summary of this approach which is crucial in the proof of Theorem 1.8 (but not used elsewhere in the article).

We follow the Tannakian approach of [9, §2.5] for defining  $I_{\mathcal{K}}$ . Let  $\mathcal{C}$  denote the category of pairs  $(V, \nabla)$  consisting of a finite dimensional vector space  $V$  over  $\mathcal{K}$  equipped with a connection  $\nabla$ . The operation  $(V, \nabla) \otimes (V', \nabla') := (V \otimes V', \nabla \otimes \nabla')$  defines a tensor category structure on  $\mathcal{C}$ . Moreover, Katz used Theorem 1.4 to construct a canonical fibre functor  $F : \mathcal{C} \rightarrow \text{Vect}_{\mathbb{C}}$ . Thus,  $\mathcal{C}$  is a Tannakian category. The group  $I_{\mathcal{K}}$  is defined to be the automorphism of the fibre functor  $F$ .

With the differential Galois group  $I_{\mathcal{K}}$  at hand, specifying a pair  $(V, \nabla)$  amounts to specifying a representation  $I_{\mathcal{K}} \rightarrow \text{GL}(V)$ . In this language, specifying a formal  $G$ -connection  $\nabla$  amounts to specifying a homomorphism  $\rho_{\nabla} : I_{\mathcal{K}} \rightarrow G$ . The Zariski closure of the image of  $\rho_{\nabla}$  is an algebraic subgroup of  $G$  called the *differential Galois group* of  $\nabla$  and denoted by  $G_{\nabla}$ . For an alternative point of view on  $G_{\nabla}$ , cf. [21, §1.4].

## 4.2 Proof of Theorem 1.8

Suppose the differential operator  $d + A$ ,  $A \in \mathfrak{g}(\mathcal{K})$ , is gauge equivalent to  $d + X$  with  $X \in \mathfrak{h}(\mathcal{K}')$  for some finite extension  $\mathcal{K}'$  of  $\mathcal{K}$ . Then  $\text{ad}(X) \in \text{ad}(\mathfrak{h})(\mathcal{K}')$  and  $\text{ad}(\mathfrak{h})$  is contained in some Cartan subalgebra of  $\mathfrak{gl}(\mathfrak{g})$ . Then there exists  $g \in \text{GL}(\mathfrak{g})(\mathbb{C})$  such that  $g^{-1} \text{ad}(\mathfrak{h})g \in \mathfrak{d}$  where  $\mathfrak{d}$  consists of diagonal matrices in  $\mathfrak{gl}(\mathfrak{g})$ . As  $dg = 0$ , the gauge action of  $g$  on  $d + \text{ad}(X)$  yields  $d + g^{-1} \text{ad}(X)g$ . Thus  $d + \text{ad}(X)$  is gauge equivalent to a diagonal differential operator and is therefore semisimple by Theorem 1.2. As  $d + \text{ad}(X)$  is gauge equivalent to  $d + \text{ad}(A)$ , this implies that  $d + \text{ad}(A)$  is semisimple. By definition, then  $d + A$  is semisimple.

Conversely, suppose that  $d + A$  is semisimple, i.e.  $d + \text{ad}(A)$  is semisimple. By Theorem 1.2,  $d + \text{ad}(A)$  is diagonalisable after a finite extension  $\mathcal{K}'$  of  $\mathcal{K}$ . This implies that the image of the composition

$$I_{\mathcal{K}'} \rightarrow G \rightarrow \text{GL}(\mathfrak{g})$$

is a subgroup of a torus in  $\text{GL}(\mathfrak{g})$ . This then implies that the image of  $I_{\mathcal{K}'} \rightarrow G$  is a subgroup of a maximal torus  $H \subset G$ ; that is, the above map factors through a map  $I_{\mathcal{K}'} \rightarrow H$ . Thus,  $d + A$  is equivalent to a connection of the form  $d + X$  for some  $X \in \mathfrak{h}(\mathcal{K}')$ .  $\square$

*Remark 4.1.* Let  $\nabla = d + X$  be a semisimple formal  $G$ -connection. By Theorem 1.8, we may assume (after a finite base change) that  $X \in \mathfrak{h}(\mathcal{K})$  where  $\mathfrak{h}$  is a Cartan subalgebra of  $\mathfrak{g}$ . Write  $X = \sum_i X_i t^i$  where  $X_i \in \mathfrak{h}(\mathbb{C})$  and set

$$X^+ = \sum_{i \geq 1} X_i t^i \in \mathfrak{h}[[t]], \quad X^- = \sum_{i \leq 0} X_i t^i \in \mathfrak{h}[t^{-1}].$$

Let

$$A := \exp\left(-\int \frac{X^+}{t} dt\right) \in H(\mathbb{C}[[t]]).$$

Then gauge transformation of  $\nabla$  by  $A$  yields  $d + X^-$ . This is the canonical form of  $\nabla$  in the sense of [1].

## 4.3 Jordan decomposition for $G$ -connections

We start with a lemma, which is an analogue of a standard result in Lie theory, c.f. [5]\*§6.4.

**Lemma 4.2.** *Let  $\mathfrak{g} \subset \mathfrak{gl}(V)$  be a Lie subalgebra. Let  $d + A : V \rightarrow V$ ,  $A \in \mathfrak{g}(\mathcal{K})$ , be a differential operator with Jordan decomposition (as a  $\text{GL}(V)$ -connection)  $D = d + X + N$ . Then  $X \in \mathfrak{g}(\mathcal{K})$ ; moreover,  $d + X$  is a semisimple  $G$ -connection.*

*Proof.* By definition  $\mathfrak{g}(\mathcal{K})$  is a  $(d + \text{ad } A)$ -invariant subspace of  $\mathfrak{gl}(V) \otimes \mathcal{K}$ . Thus, by Lemma 3.2, it is also  $(d + \text{ad } X)$ -invariant. By definition,  $\mathfrak{g}(\mathcal{K})$  is  $d$  invariant. Thus,  $\mathfrak{g}(\mathcal{K})$  is  $\text{ad } X$ -invariant. This implies that  $(d + \text{ad } X) - d : \mathfrak{g}(\mathcal{K}) \rightarrow \mathfrak{g}(\mathcal{K})$  is a  $\mathcal{K}$ -linear derivation on  $\mathfrak{g}(\mathcal{K})$  and hence  $X \in \mathfrak{g}(\mathcal{K})$  (since every  $\mathcal{K}$ -linear derivation is inner).  $\square$

*Proof of Theorem 1.9.* We are now ready to prove the theorem. Let  $\nabla = d + A$  be a  $G$ -connection. Note that the adjoint action gives an embedding  $\mathfrak{g}(\mathcal{K}) \subset \mathfrak{gl}(\mathfrak{g}(\mathcal{K}))$ . Let  $d + X + N$  denote the Jordan decomposition of  $d + A$  as a differential operator  $\mathfrak{g}(\mathcal{K}) \rightarrow \mathfrak{g}(\mathcal{K})$ . Then by the previous lemma,  $X \in \mathfrak{g}(\mathcal{K})$  and  $d + X$  is a semisimple  $G$ -connection. It follows that  $N = d + A - (d + X)$  is a nilpotent element of  $\mathfrak{g}(\mathcal{K})$ . Now  $d + X$  and  $N$  commute in the extended loop algebra of  $\mathfrak{gl}(\mathfrak{g}(\mathcal{K}))$ . This implies that they commute in the extended loop algebra of  $\mathfrak{g}(\mathcal{K})$ , this establishes the existence of Jordan decomposition.

For uniqueness, suppose  $d + X_1 + N_1$  and  $d + X_2 + N_2$  are Jordan decompositions for  $\nabla$ . Then  $d + \text{ad}(X_1) + \text{ad}(N_1)$  and  $d + \text{ad}(X_2) + \text{ad}(N_2)$  are Jordan decompositions for  $\text{ad}(\nabla)$ . By uniqueness of Jordan decomposition for differential operators (Theorem 1.4), we obtain  $\text{ad}(X_1) = \text{ad}(X_2)$  and  $\text{ad}(N_1) = \text{ad}(N_2)$ . As the adjoint representation is faithful, we conclude  $X_1 = X_2$  and  $N_1 = N_2$ .  $\square$

As a corollary of our proof we obtain:

**Corollary 4.3.** *Let  $f : G \rightarrow H$  be a homomorphism of algebraic groups. Let  $\nabla$  be a formal  $G$ -connection with Jordan decomposition  $d + S + N$ . Then  $d + f(S) + f(N)$  is the Jordan decomposition of  $f_* \nabla$ , where the latter is the  $H$ -connection defined by the composition*

$$I_{\mathcal{K}} \xrightarrow{\rho_{\nabla}} G \xrightarrow{f} H.$$

*Remark 4.4.* A formal  $G$ -connection  $\nabla$  is called *unipotent* if its semisimple part is trivial. One can show that  $\nabla$  is unipotent if and only if its differential Galois group  $G_{\nabla}$  is unipotent. According to [21, thm. 11.2],  $G_{\nabla}$  is

then a one-parameter subgroup of  $G$  generated by a unipotent element. This implies that the map

$$Y \mapsto d + Y$$

from nilpotent elements of  $\mathfrak{g}(\mathbb{C})$  to formal unipotent  $G$ -connections defines a bijection between nilpotent orbits in  $\mathfrak{g}(\mathbb{C})$  and equivalence classes of unipotent connections. Thus, one obtains a generalisation of Katz's results (Lemma 3.9) to the setting of  $G$ -connections.

## References

- [1] B. BABBITT, P. VARADARAJAN, *formal reduction theory of formal differential equations: a group theoretic view*, Pacific J. Math. **109** (1) (1983), p. 1–80.
- [2] P. BOALCH, D. YAMAKAWA *Twisted wild character varieties*, arXiv:1512.08091 (2015).
- [3] E. FRENKEL, *Langlands Correspondence for Loop Groups*, Cambridge Studies in Advanced Mathematics **103**, Cambridge University Press, Cambridge (2007).
- [4] M. HUKUHARA, *Théorèmes fondamentaux de la théorie des équations différentielles ordinaires. II*, Mem. Fac. Sci. Kyūsyū Imp. Univ. A. **2** (1941), 1–25.
- [5] J.E. HUMPHREYS, *Introduction to Lie Algebras and Representation Theory*, Graduate Texts in Mathematics **9**, Second printing, revised, Springer-Verlag, New York-Berlin (1978).
- [6] V. Kac, *Infinite-dimensional Lie algebras*, Edition 3, Cambridge University Press, Cambridge (1990).
- [7] M. KAMGARPOUR, D. SAGE, *A geometric analogue of a conjecture of Gross and Reeder*, Amer. J. of Math. **141** (2019), 1457–1476.
- [8] N. M. KATZ, *Nilpotent connections and the monodromy theorem: Applications of a result of Turrittin*, Inst. Hautes Études Sci. Publ. Math. **39** (1970), 175–232.
- [9] N.M. KATZ, *On the calculation of some differential Galois groups*, Invent. Math. (1) **87** (1987), 13–61.
- [10] K. S. KEDLAYA,  *$p$ -adic differential equations*, Cambridge Studies in Advanced Mathematics **125**, Cambridge University Press, Cambridge (2010).
- [11] G. LEVELT, *Jordan decomposition for a class of singular differential operators*, Arkiv för Matematik **13** (1975), 1–27.
- [12] M. LUU, *Local Langlands duality and a duality of conformal field theories*, arXiv:1506.00663 (2015).
- [13] B. MALGRANGE, *Sur les points singuliers des équations différentielles*, Enseignement Math. (2) **20** (1974), 147–176.
- [14] B. MALGRANGE, *Sur la réduction formelle des équations différentielles à singularités irrégulières* (1979).
- [15] O. ORE, *Theory of non-commutative polynomials*, Ann. of Math. (2) **34** (1933), no. 3, 480–508.
- [16] C. PRAAGMAN, *The formal classification of linear difference operators*, Indagationes Mathematicae (Proceedings), **86** (1983), no. 2, 249 - 261.
- [17] S. RASKIN, *On the notion of spectral decomposition in local geometric Langlands*, arXiv:1511.01378 (2015).
- [18] P. ROBBA, *Lemmes de Hensel pour les opérateurs différentiels. Application à la réduction formelle des équations différentielles*, Enseign. Math. (2) **26** (1980), no.3-4, 279–311 (1981).
- [19] C. SABBABH, *Introduction to algebraic theory of linear systems of differential equations*, Unpublished lecture notes.
- [20] C. SABBABH, *An explicit stationary phase formula for the local formal Fourier-Laplace transform*, Singularities I, Contemp. Math. **474** (2008), 309–330.
- [21] M. VAN DER PUT, M. F. SINGER, *Galois theory of linear differential equations*, Grundlehren der Mathematischen Wissenschaften **328**, Springer-Verlag, Berlin (2003).
- [22] H. L. TURRITTIN, *Convergent solutions of ordinary linear homogeneous differential equations in the neighborhood of an irregular singular point*, ACTA MATH. **93** (1955), 27–66.
- [23] V. S. VARADARJAN, *Linear formal differential equations: a modern point of view*, BULLETIN OF AMS **33** (1996), NO.1.
- [24] W. WASOW, *Asymptotic expansions for ordinary differential equations*, Pure and Applied Mathematics, Vol. XIV, Interscience Publishers John Wiley & Sons (1965).

M. Kamgarpour  
SCHOOL OF MATHEMATICS AND PHYSICS,  
THE UNIVERSITY OF QUEENSLAND,  
ST. LUCIA, QUEENSLAND, AUSTRALIA  
E-mail address: masoud@uq.edu.au

S. Weatherhog  
SCHOOL OF MATHEMATICS AND PHYSICS,  
THE UNIVERSITY OF QUEENSLAND,  
ST. LUCIA, QUEENSLAND, AUSTRALIA  
E-mail address: samuel.weatherhog@gmail.com