THE UNIVERSITY OF QUEENSLAND

AUSTRALIA

Jordan Decomposition for Differential Operators

Samuel Weatherhog

BSc Hons I, BE Hons IIA

A thesis submitted for the degree of Master of Philosophy at The University of Queensland in 2017

School of Mathematics and Physics

Abstract

One of the most well-known theorems of linear algebra states that every linear operator on a complex vector space has a Jordan decomposition. There are now numerous ways to prove this theorem, however a standard method of proof relies on the existence of an eigenvector. Given a finite-dimensional, complex vector space V, every linear operator $T : V \to V$ has an eigenvector (i.e. a $v \in V$ such that $(T - \lambda I)v = 0$ for some $\lambda \in \mathbb{C}$). If we are lucky, V may have a basis consisting of eigenvectors of T, in which case, T is *diagonalisable*. Unfortunately this is not always the case. However, by relaxing the condition $(T - \lambda I)v = 0$ to the weaker condition $(T - \lambda I)^n v = 0$ for some $n \in \mathbb{N}$, we can *always* obtain a basis of *generalised eigenvectors*. In fact, there is a canonical decomposition of V into *generalised eigenspaces* and this is essentially the Jordan decomposition.

The topic of this thesis is an analogous theorem for *differential operators*. The existence of a Jordan decomposition in this setting was first proved by Turrittin following work of Hukuhara in the one-dimensional case. Subsequently, Levelt proved uniqueness and provided a more conceptual proof of the original result. As a corollary, Levelt showed that every differential operator has an eigenvector. He also noted that this was a strange chain of logic: in the linear setting, the existence of an eigenvector is a much easier result and is in fact used to obtain the Jordan decomposition. Levelt remarked that a direct proof of his corollary would provide a much simpler proof of the Jordan decomposition for differential operators. It is this remark that stimulated the work of this thesis. Although there have been numerous alternative proofs and applications of the Hukuhara-Levelt-Turrittin theorem, it appears that Levelt's suggestion has not been carried out in the literature. Our goal is to provide a proof of the Hukuhara-Levelt-Turrittin theorem that mimics the proof of the usual Jordan decomposition theorem.

Declaration by author

This thesis is composed of my original work, and contains no material previously published or written by another person except where due reference has been made in the text. I have clearly stated the contribution by others to jointly-authored works that I have included in my thesis.

I have clearly stated the contribution of others to my thesis as a whole, including statistical assistance, survey design, data analysis, significant technical procedures, professional editorial advice, and any other original research work used or reported in my thesis. The content of my thesis is the result of work I have carried out since the commencement of my research higher degree candidature and does not include a substantial part of work that has been submitted to qualify for the award of any other degree or diploma in any university or other tertiary institution. I have clearly stated which parts of my thesis, if any, have been submitted to qualify for another award.

I acknowledge that an electronic copy of my thesis must be lodged with the University Library and, subject to the policy and procedures of The University of Queensland, the thesis be made available for research and study in accordance with the Copyright Act 1968 unless a period of embargo has been approved by the Dean of the Graduate School.

I acknowledge that copyright of all material contained in my thesis resides with the copyright holder(s) of that material. Where appropriate, I have obtained copyright permission from the copyright holder to reproduce material in this thesis.

Publications during candidature

No publications.

Publications included in the thesis

No publications included.

Contributions by others to the thesis

My principal advisor Masoud Kamgarpour suggested the possibility of a new approach to proving this theorem. Throughout the research, he provided invaluable advice and guidance both mathematically and editorially. My co-advisor Ole Warnaar also provided valuable suggestions and helped with the editing of the final thesis.

Statement of parts of the thesis submitted to qualify for the award of another degree

None.

Acknowledgements

I would like to thank my principal supervisor, Dr. Masoud Kamgarpour, for introducing me to the beautiful subject of geometric representation theory. In addition to the many hours of advice, he also gave me numerous opportunities to broaden my mathematical knowledge through conferences both locally and internationally for which I am truly grateful. I would also like to thank my co-advisor, Prof. Ole Warnaar, for many useful discussions and especially for his patience while I learned to write mathematics well. His attention to detail is something to aspire to.

More generally, I would like to thank the algebra group at UQ for the many conversations that have helped me become a more rounded mathematician as well as inspired me to continue my journey in this field. I'm also very grateful to the School of Mathematics and Physics at UQ for their support throughout my candidature.

Keywords

differential operators, formal connections, Jordan decomposition, Kac-Moody algebras, differential polynomials

Australian and New Zealand Standard Research Classifications (ANZSRC)

ANZSRC code: 010109 Ordinary Differential Equations, Difference Equations and Dynamical Systems 80%

ANZSRC code: 010101 Algebra and Number Theory 20%

Fields of Research (FoR) Classification

FoR code: 0101 Pure Mathematics 100%

Contents

1. Introduction

One of the most well-known theorems of linear algebra states that every linear operator on a complex vector space has a Jordan decomposition. There are now numerous ways to prove this theorem, however a standard method of proof relies on the existence of an eigenvector. Given a finite-dimensional, complex vector space V, every linear operator $T : V \to V$ has an eigenvector (i.e. a $v \in V$ such that $(T - \lambda I)v = 0$ for some $\lambda \in \mathbb{C}$). If we are lucky, V may have a basis consisting of eigenvectors of T, in which case, T is *diagonalisable*. Unfortunately this is not always the case. However, by relaxing the condition $(T - \lambda I)v = 0$ to the weaker condition $(T - \lambda I)^n v = 0$ for some $n \in \mathbb{N}$, we can *always* obtain a basis of *generalised eigenvectors*. In fact, there is a canonical decomposition of V into *generalised eigenspaces* and this is essentially the Jordan decomposition.

The topic of this thesis is an analogous theorem for *differential operators*. The existence of a Jordan decomposition in this setting was first proved by Turrittin [Tur55] following work of Hukuhara [Huk41] in the one-dimensional case. Subsequently, Levelt [Lev75] proved uniqueness and provided a more conceptual proof of the original result. As a corollary, Levelt showed that every differential operator has an eigenvector. He also noted that this was a strange chain of logic: in the linear setting, the existence of an eigenvector is a much easier result and is in fact used to obtain the Jordan decomposition. Levelt remarked that a direct proof of his corollary would provide a much simpler proof of the Jordan decomposition for differential operators. It is this remark that stimulated the work of this thesis. Although there have been numerous alternative proofs and applications of the Hukuhara-Levelt-Turrittin theorem (see [BV83, BBDE05, KS16, Kat87, Kat70, Ked10a, Luu15, Mal79, Pra83, Ras15, Rob80, vdPS03, Was65]), it appears that Levelt's suggestion has not been carried out in the literature. Our goal is to provide a proof of the Hukuhara-Levelt-Turrittin theorem that mimics the proof of the usual Jordan decomposition theorem.

In the linear setting, the existence of an eigenvector is a consequence of the fundamental theorem of algebra. This theorem guarantees that every polynomial splits completely into linear factors over an algebraically closed field. In the differential setting, there is also a ring of polynomials however this ring is not commutative. Despite this, there is a factorisation result for polynomials in this ring (see e.g. [Pra83]). In section 7 we give a proof of this result and use this to prove Levelt's corollary directly in Section 8.1. We then give a new proof of the Hukuhara-Levelt-Turrittin theorem using this result.

Sometime after Levelt proved his result, Babbitt and Varadarajan [BV83] gave a group theoretic proof of the Hukuhara-Levelt-Turrittin theorem. They also showed that the decomposition could be extended to reductive algebraic groups other than $\mathrm{GL}(n)$. This is perhaps not surprising given that this can also be done in the linear setting (usually known as the Jordan-Chevalley decomposition). Again, though, the proof in [BV83] is very different to the proof in the linear setting. We have therefore given an alternative proof of their result in Section 9 with the intent of emphasising the analogy between the linear and differential settings.

2. Background

The purpose of this section is to informally introduce the main objects of this thesis: differential operators. We will do this through various points of view, each increasing in sophistication. Hopefully this provides the reader with some insight into why we would be interested in a decomposition theorem such as the Hukuhara-Levelt-Turrittin theorem. A more formal treatment of differential operators is given in sections 6.1 and 6.2.

2.1. **Linear Differential Equations.** Consider a linear differential equation of the form:

$$y^{(n)} + a_1 y^{(n-1)} + \cdots + a_n y = 0, \quad y = y(t), a_i \in \mathbb{C}((t)). \tag{2.1}$$

Given such an equation, one can always obtain a system of first-order linear differential equations by introducing n variables $y_1, \ldots, y_n$ with $y_1 := y$ and $y_i = y_{i-1}'$ for $2 \leq i \leq n$. We can write the resulting system of equations in matrix form:

$$\frac{d}{dt}\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \\ y_n \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & & & 1 \\ -a_n & -a_{n-1} & \ldots & & -a_1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \\ y_n \end{pmatrix}, \tag{2.2}$$

which can be compactly written as $\boldsymbol{y}' = -A\boldsymbol{y}$ (the reason for the negative sign is so that we can write this instead as $\boldsymbol{y}' + A\boldsymbol{y} = 0$). Note that $\boldsymbol{y}$ is now a vector which lives in the vector space $V := (\mathbb{C}((t)))^n$. Our point of view is that the equation (2.2) defines an operator on this vector space which we denote $d + A : V \to V$. Here $d : V \to V$ is the operator which sends $\boldsymbol{y}$ to $\boldsymbol{y}'$. Note that whilst $d + A$ is a $\mathbb{C}$-linear operator, it is *not* a $\mathbb{C}((t))$-linear operator (i.e. it is *not* a linear operator on V). In fact, $d + A$ satisfies the Leibniz rule:

$$(d + A)(av) = a(d + A)(v) + a'v, \quad \forall a \in \mathbb{C}((t)).$$

We call such operators *differential operators* on V. From this point of view, we need not restrict ourselves to considering only matrices arising from linear differential equations (so-called companion matrices), but can allow A to be any $n \times n$ matrix with $\mathbb{C}((t))$ entries[1]. It turns out, however, that this does not actually introduce anything new; i.e. we can always find a basis of V such that A is the companion matrix of some linear differential equation. This amazing fact is called the *cyclic vector theorem* (which will be discussed further in Section 6) and is a major point of difference between differential and linear operators. In addition, whereas the rational canonical form of a linear operator is essentially unique, this is not the case in the differential setting.

Another point of difference between the two settings is the action of $\mathrm{GL}(V)$ on the matrix A. In both settings, this action arises as a change of basis of V, but in the differential setting, such a change of basis does not correspond to a conjugation of A by some element in $\mathrm{GL}(V)$. In fact, if we make a change of basis by $g \in \mathrm{GL}(V)$, then, in the new basis, our matrix equation (2.2) becomes:

$$\begin{aligned} (g\boldsymbol{y})' = -Ag\boldsymbol{y} &\implies g'\boldsymbol{y} + g\boldsymbol{y}' = -Ag\boldsymbol{y} \\ &\implies \boldsymbol{y}' = -(g^{-1}Ag + g^{-1}g')\boldsymbol{y}, \end{aligned}$$

where g' is the matrix obtained by differentiating each entry of g. We call the action $g \cdot A \mapsto g^{-1}Ag + g^{-1}g'$ the *gauge action* of $\mathrm{GL}(V)$ on A and say that the matrices A and $g^{-1}Ag + g^{-1}g'$ are *gauge equivalent*. From this point of view, the Hukuhara-Levelt-Turrittin theorem roughly tells us that there exists a gauge transformation that will put the matrix A in Jordan canonical form.

[1] In fact, we can allow the entries of these matrices to be from any differential field.

2.2. **Regular Singular Differential Equations.** In the setting of linear differential equations, there is an 'easy' case in which the Hukuhara-Levelt-Turrittin decomposition can be proved directly: the case of *regular singular* differential equations. In Section 2.1, we allowed our system of differential equations to have entries in $\mathbb{C}(\!(t)\!)$. This means that there are points at which the matrix A can have singularities. If A is gauge equivalent to a matrix that contains at most a pole of order one at a point, the equation is said to be *regular singular* at that point; otherwise, the equation is said to have an irregular singularity at the point. This dichotomy plays an important role in all proofs of the Hukuhara-Levelt-Turrittin theorem and the themes introduced here will appear in our discussion of a differential Hensel's Lemma (see Section 7.3). There are also many equivalent characterisations of regular singularities which will be discussed throughout.

Suppose that our system of differential equations $\boldsymbol{y}' = -A\boldsymbol{y}$ has a regular singularity at $t = 0$. We can re-write[2] this system as $t\boldsymbol{y}' = -B\boldsymbol{y}$, where the entries of B are now holomorphic at $t = 0$. It is now possible to naively search for a gauge transformation that will simplify the matrix B (c.f. [Was65, §5], [vdPS03, §3.1.1]). Let

$$B = B_0 + B_1 t + B_2 t^2 + \cdots, \quad P = I + P_1 t + P_2 t^2 + \cdots,$$

so that P is an invertible matrix. We claim that P can be chosen so that gauge transformation of B by P yields B_0. That is, we can choose P such that $PB_0 = BP + tP'$. Expanding both sides of this equation we have:

$$B_0 \sum_{i \geq 0} P_i t^i = \sum_{i \geq 0} \Big(iP_i + \sum_{j=0}^{i} B_{i-j} P_j \Big) t^i, \quad P_0 := I.$$

Comparing powers of t leads to the following recursive definition of the P_i:

$$B_0 P_i - P_i(B_0 - iI) = -(B_i + B_{i-1}P_1 + \cdots + B_1 P_{i-1}). \tag{2.3}$$

Here we come across an inherent technical difficulty in the differential setting: if (and only if) B_0 and $B_0 - iI$ share an eigenvalue, it is possible that the left-hand side of (2.3) is 0 without P_i necessarily being 0 (see [Was65, Theorem 4.1]). Thus, if the eigenvalues of B_0 do not differ by integers, then we can solve (2.3) for P. Luckily, if the eigenvalues of B_0 do differ by integers, there is a gauge transformation that will yield a matrix whose constant term has repeated eigenvalues (see [Was65, §5.3]). Such transformations are classically called *shearing* transformations and will be discussed further in section 7.3. Note that having made a gauge transformation to B_0, the Hukuhara-Levelt-Turrittin decomposition can be realised by taking the usual (linear) Jordan decomposition of B_0.

We briefly mention what happens for *irregular singular* equations. One can attempt the same naive approach as above, however the recursive definition obtained (c.f. equation (2.3)) is always singular. In fact, in the irregular singular case, the left-hand side of (2.3) is just $B_0 P_i - P_i B_0$. It is possible for this to be 0 without P_i necessarily being 0. Thus, a more sophisticated approach is required (see [Was65, §11]).

2.3. **Flat Connections on Vector Bundles.** In this section we introduce the geometric setting of our work. Many of the applications of the Hukuhara-Levelt-Turrittin theorem are in the realm of geometric representation theory. In this setting one is interested in *flat connections* on vector bundles (and

[2] This introduces a subtle difference in the meaning of gauge transformation. Now g' is the matrix obtained by applying the derivation $t\frac{d}{dt}$ to the entries of g. This will be discussed in detail in §6.

more generally, flat connections on principal G-bundles for reductive groups G other than $\mathrm{GL}(n, \mathbb{C})$). In general, one can think of connections as a coordinate-free way to describe differential equations on a real manifold. For a proper treatment of the material presented here, the reader should refer to [GH94, Chapter 0, §5] and [Fre07, §1.2].

Let X be a smooth manifold and $\pi : E \to X$ a complex vector bundle of rank n over X. That is, E is a smooth manifold and for each point $x \in X$, there is neighbourhood $U \subset X$ of x such that $\pi^{-1}(U) \cong U \times \mathbb{C}^n$. Moreover, for any two overlapping trivializations, say U_α, U_β, we obtain a transition map $g_{\alpha\beta} : U_\alpha \cap U_\beta \to \mathrm{GL}(n, \mathbb{C})$. Denote by $\Omega^0(E)$ the sheaf of sections of the vector bundle E and by $\Omega^1(E)$ the sheaf of E-valued one-forms. We define a connection on E as a $\mathbb{C}$-linear mapping $D : \Omega^0(E) \to \Omega^1(E)$ that satisfies the Leibniz rule:

$$D(f\xi) = f \cdot D(\xi) + df \cdot \xi, \quad f \in \mathcal{O}_X, \xi \in \Omega^0(E),$$

where $\mathcal{O}_X$ the sheaf of smooth functions on X and df is the differential of the function f.

Recall that a local frame for E over an open subset $U \subset X$ is a collection of n smooth sections $\{f_1, \ldots, f_n\}$ such that $\{f_1(x), \ldots, f_n(x)\}$ forms a basis of the fibre $\pi^{-1}(x)$ for each $x \in U$. The existence of a local frame is equivalent to the existence of a local trivialization. Given a local frame $\{f_1, \ldots, f_n\}$, we can represent the action of D by a matrix of one-forms whose columns are the $D(f_i)$; that is, locally we can write $D = d + A$ where A is a matrix of one-forms. If we make a change of frame $f_i' = \sum_j g_{ij} f_j$, then the connection matrix becomes

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

This is precisely the gauge transformation that was discussed in Section 2.1.

We now discuss the adjective *flat*. Using the connection D, we can define operators $D : \Omega^p(E) \to \Omega^{p+1}(E)$ by simply forcing Leibniz's rule, i.e. by requiring that $D(f\xi) = df \cdot \xi + f \wedge D(\xi)$ for $f \in \Omega^p(E)$ and $\xi \in \Omega^0(E)$. By abuse of notation we also call these operators D and in particular, we denote by D^2 the composition $D \circ D : \Omega^0(E) \to \Omega^2(E)$. This operator is known as the curvature tensor of D. If $D^2 = 0$, then the connection is said to be *flat*. In this setting, the Hukuhara-Levelt-Turrittin theorem can be viewed as a classification of meromorphic flat connections on the open disc with its center removed.

2.4. **Extended Loop Algebras.** The relationship between differential operators and linear operators can be pushed further: recall that the collection of all linear operators on a finite-dimensional vector space forms a Lie algebra $\mathfrak{gl}(V)$. The set of all differential operators on a vector space also has a natural Lie algebra structure and we can reformulate the Hukuhara-Levelt-Turrittin theorem in this setting.

Let $\mathfrak{g}$ be a simple Lie algebra over $\mathbb{C}$ and set $\mathcal{K} := \mathbb{C}((t))$. The *loop algebra*[3] is defined to be the Lie algebra $\mathfrak{g}_\mathcal{K} := \mathfrak{g} \otimes_\mathbb{C} \mathcal{K}$ with Lie bracket $[x \otimes f, y \otimes g] := [x, y] \otimes fg$ (where $[x, y]$ denotes the Lie bracket of x and y in $\mathfrak{g}$). Let $d : \mathcal{K} \to \mathcal{K}$ be the map $f \mapsto t\frac{df}{dt}$. This map is additive and satisfies the Leibniz rule (i.e. d is a *derivation* on $\mathcal{K}$ — see Section 6.1). Note that d extends to a derivation of the Lie algebra $\mathfrak{g}_\mathcal{K}$ via $d(x \otimes f) = x \otimes d(f)$ (where, by abuse of notation, we have used d for both

[3]This algebra can be viewed as the completion of the algebra of polynomial maps from S^1 to $\mathfrak{g}$.

the derivation on $\mathcal{K}$ and its extension to $\mathfrak{g}_{\mathcal{K}}$). Using d, we can construct the *extended loop algebra* $\hat{\mathfrak{g}} := \mathbb{C}d \oplus \mathfrak{g}_{\mathcal{K}}$ which is a Lie algebra with Lie bracket:

$$[\alpha d + A, \beta d + B] := [A, B] + \alpha d(B) - \beta d(A), \tag{2.4}$$

for $A, B \in \mathfrak{g}_{\mathcal{K}}$ and $\alpha, \beta \in \mathbb{C}$.

The extended loop algebra "exponentiates" to an algebraic group $\hat{G}$. We briefly describe the construction of this group following [Kum02, §13.2]. It is well known that there is a unique (up to isomorphism) connected, simply-connected, complex algebraic group G whose Lie algebra is $\mathfrak{g}$ (see [Bou98, Chapter III, §6.3]). We define the *loop group* to be the $\mathcal{K}$-points of G, denoted $G(\mathcal{K})$. One can consider $G(\mathcal{K})$ to be the "exponentiation" of the loop algebra $\mathfrak{g}_{\mathcal{K}}$. We can extend $G(\mathcal{K})$ by "exponentiating" the derivation d in $\hat{\mathfrak{g}}$ to obtain the *extended loop group* $\hat{G}$. The adjoint action of $\hat{G}$ on $\hat{\mathfrak{g}}$ restricts to give the following action of $G(\mathcal{K})$ on $\hat{\mathfrak{g}}$:

$$\mathrm{Ad}(g) \cdot (x + \alpha d) = \mathrm{Ad}(g)(x) + \alpha g^{-1} t \frac{dg}{dt} + \alpha d, \quad g \in \mathrm{G}(\mathcal{K}).$$

Note that $\mathrm{Ad}(g)(x)$ on the right-hand side is the extension of the usual adjoint action of G on $\mathfrak{g}$ to an action of $G(\mathcal{K})$ on $\mathfrak{g}_{\mathcal{K}}$. It is clear that this action of $G(\mathcal{K})$ on $\hat{\mathfrak{g}}$ coincides with the gauge action described in Section 2.1.

We can reformulate the Hukuhara-Levelt-Turrittin theorem in this context. The adjoint action of $\hat{G}$ partitions $\hat{\mathfrak{g}}$ into orbits. The Hukuhara-Levelt-Turrittin theorem implies that (after a finite extension) each orbit contains a unique element of the form $\lambda d + S + N$ where $\lambda \in \mathbb{C}$, $\lambda d + S$ is semisimple, N is a nilpotent linear operator and $[\lambda d + S, N] = 0$ (where the bracket is the one defined by (2.4)). From this point of view, it would be desirable to seek a classification of conjugacy classes of affine Kac-Moody algebras (including the twisted Kac-Moody algebras). For some results in this direction, cf. [BV83]. If instead of conjugacy classes in $\hat{\mathfrak{g}}$, we consider the conjugacy classes in $\hat{G}$, then we obtain the notion of *q-difference operators*. There is also a version of Jordan decomposition for these objects, cf. [Pra83].

3. Jordan Decomposition for Linear Operators

The purpose of this section is to remind the reader of the general theory of linear operators on finite-dimensional vector spaces. Many of the results presented here have analogues in the differential setting and the intent is that the development of the classical theory here will be mimicked in the differential setting.

Let k be a field and $\bar{k}$ the algebraic closure of k. Throughout this section V will denote a *finite-dimensional* vector space. We refer the reader to Dummit and Foote [DF04, chapter 12] for the general theory of modules over a principal ideal domain (in particular for the definitions of principal ideal domain, Euclidean domain, etc.) and to [Axl15] for a good introduction to linear algebra.

3.1. **Jordan Decomposition over Algebraically Closed Fields.** It is well known that over $\bar{k}$ any matrix can be put into Jordan normal form:

Theorem 3.1 (Jordan Normal Form). *Let V be a vector space over $\bar{k}$ and $T : V \to V$ a linear operator. Then there exists a basis of V in which the operator T is represented by a Jordan matrix, i.e. a matrix of the form:*

$$\begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_n \end{pmatrix}$$

where each J_i is a Jordan block:

$$J_i = \begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & 1 & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}$$

and $\lambda_i \in \bar{k}$. Moreover, this matrix representation is unique up to reordering of the Jordan blocks.

One way to prove Theorem 3.1 is to consider V as a $\bar{k}[x]$-module where x acts as a linear operator on V. Because $\bar{k}$ is algebraically closed, the non-zero prime ideals of $\bar{k}[x]$ are generated by $x - \lambda$ where $\lambda \in \bar{k}$. The fundamental theorem of finitely-generated modules over a principal ideal domain (PID) then implies that V decomposes as

$$V \cong \bar{k}[x]/\langle x - \lambda_1 \rangle^{r_1} \oplus \cdots \oplus \bar{k}[x]/\langle x - \lambda_n \rangle^{r_n}.$$

Moreover, this decomposition is unique up to reordering of the summands. For the summand $\bar{k}[x]/\langle x - \lambda_i \rangle^{r_i}$, choosing the basis $\{(x-\lambda_i)^{r_i-1}, (x-\lambda_i)^{r_i-2}, \dots, 1\}$ yields a Jordan block matrix for that summand.

There is a second formulation of Theorem 3.1 which we will use frequently. It relies on the following types of linear operators:

Definition 3.2 (Nilpotent Linear Operator). *Let V be a vector space. A linear operator $T : V \to V$ satisfying $T^n = 0$ for some positive integer n is called* nilpotent.

Definition 3.3 (Diagonalizable Operator)**.** *Let V be a vector space over $\bar{k}$ and $T : V \to V$ be a linear operator. Then T is* diagonalizable *if the following equivalent conditions hold:*

(i) There exists a basis of V for which the matrix of T is diagonal;

(ii) There is a basis of V consisting of eigenvectors for T;

(iii) The minimal polynomial of T splits in $\bar{k}$ with distinct roots.

(iv) Every T-invariant subspace $W \subset V$ has a T-invariant complement W', i.e. $V = W \oplus W'$.

Theorem 3.1 implies the following:

Theorem 3.4 (Jordan Decomposition over Algebraically Closed Fields)**.** *Let V be a vector space over $\bar{k}$ and let $T : V \to V$ be a linear operator. Then there exists a diagonalizable operator, S, and a nilpotent operator, N, such that:*

(i) $T = S + N$;

(ii) S and N commute, i.e. $SN = NS$.

Moreover, (S, N) is the unique pair of operators satisfying (i) and (ii).

3.2. **Jordan Decomposition over Arbitrary Fields.** We now want to discuss the notion of Jordan decomposition over the field k. This introduces a problem in that some operators that "should" be diagonalizable will not be (in the sense of Definition 3.3). The main issue is that k may no longer contain the eigenvalues of an operator. To remedy this, we make the following definitions:

Definition 3.5 (Simple Linear Operator)**.** *Let V be a vector space over k and let $T : V \to V$ be a linear operator. We call T* simple *if the following equivalent conditions hold:*

(i) The corresponding $k[x]$-module is simple (i.e. it contains no non-zero proper submodules);

(ii) V contains no non-trivial T-invariant subspaces.

(iii) The minimal polynomial of T is irreducible (over k) of degree $\dim(V)$.

Definition 3.6 (Semisimple Linear Operator)**.** *Let V be a vector space over k and let $T : V \to V$ be a linear operator. We call T* semisimple *if the following equivalent conditions hold:*

(i) The corresponding $k[x]$-module is a direct sum of simple modules

(ii) Every T-invariant subspace, W, of V has a complement, W' (i.e. $V = W \oplus W'$)

(iii) The minimal polynomial of T is square-free.

Note that if $k = \bar{k}$, then T is semisimple if and only if T is diagonalizable. At this point we need to distinguish between two very similar concepts: semisimplicity and diagonalizability. In this context we extend the notion of diagonalizability to include operators that are diagonalizable after a possible field extension.

Definition 3.7 (Potential Diagonalizability)**.** *Let V be a vector space over a field k and $T : V \to V$ be a linear operator. We call T* potentially diagonalizable *if the following equivalent conditions hold:*

(i) *After some finite extension of* k, T *can be represented by a diagonal matrix.*

(ii) *The minimal polynomial of* T *is separable (i.e. it splits with distinct roots in some field extension of* k*).*

We would like to know the extent to which the notions of diagonalizability and semisimplicity coincide. From the definitions it is clear that diagonalizable operators are potentially diagonalizable and that potentially diagonalizable operators are semisimple. Over an algebraically closed field all three notions are equivalent. We will see in the following subsections that semisimplicity and potential diagonalizability are equivalent when k is a perfect field.

3.2.1. *Jordan Decomposition over Perfect Fields.* Recall that k is said to be perfect if every finite extension of k is separable. This condition means that every irreducible polynomial over k of degree n has n distinct roots in $\bar{k}$. All finite fields and fields of characteristic zerro are perfect. A good example to keep in mind for what follows is the field $\mathbb{C}(\!(t)\!)$ of formal Laurent series with complex coefficients. For perfect fields, potential diagonalizability and semisimplicity are equivalent conditions.

Lemma 3.8. *Let* V *be a vector space over a perfect field* k *and let* $T : V \to V$ *be a linear operator. Then* T *is semisimple if and only if* T *is potentially diagonalizable.*

Proof. As we have previously observed, one direction is clear from the definition. For the other direction, suppose that T is semisimple so that its minimal polynomial is square-free. Since every extension of k is separable, the minimal polynomial of T must have n distinct roots in $\bar{k}$. Hence there is an extension of k for which the minimal polynomial of T splits with distinct roots. Thus T is potentially diagonalizable. □

Remark 3.9. The proof of Lemma 3.8 also shows that if T acts semisimply on V then it also acts semisimply on $V \otimes_k k'$ where k' is a finite extension of k.

The notion of Jordan decomposition can be extended to linear operators $T : V \to V$, where the underlying field, k, is an arbitrary perfect field (c.f. [HP99, §2]).

Theorem 3.10 (Jordan Decomposition over Perfect Fields)**.** *Let* V *be a vector space over a perfect field* k *and let* $T : V \to V$ *be a linear operator. Then there exists a semisimple operator* S *and a nilpotent operator* N *such that:*

(i) $T = S + N$;

(ii) $[S, N] := SN - NS = 0$.

Moreover, (S, N) *is the unique pair of operators satisfying (i) and (ii).*

Proof. In order to prove Theorem 3.10, we consider the operator $\bar{T} : V \otimes_k \bar{k} \to V \otimes_k \bar{k}$ and denote by $\mathrm{Gal}(\bar{k}/k)$ the Galois group of $\bar{k}$ over k. By Theorem 3.4, there exists a Jordan decomposition $\bar{T} = \bar{S} + \bar{N}$ for unique semisimple $\bar{S}$ and nilpotent $\bar{N}$. The issue now is that $\bar{S}$ and $\bar{N}$ may only be defined over $\bar{k}$. However, choosing a non-trivial element $\sigma \in \mathrm{Gal}(\bar{k}/k)$ we have

$$\sigma\bar{T} = \sigma\bar{S} + \sigma\bar{N} \tag{3.5}$$

and since $\bar{T}$ is defined over k, $\sigma\bar{T} = \bar{T}$. This gives us a second Jordan decomposition of $\bar{T}$. Uniqueness of the decomposition then implies that $\sigma\bar{S} = \bar{S}$ and $\sigma\bar{N} = \bar{N}$. Hence $\bar{S}$ and $\bar{N}$ are, in fact, defined over k. □

Note that perfectness is a necessary condition here as the above argument fails if we allow non-separable extensions. In this case, there may be no $\sigma \in \mathrm{Gal}(\bar{k}/k)$ that yields a second Jordan decomposition of $\bar{T}$.

3.2.2. *Failure over Non-Perfect Fields.* If we allow non-perfect fields then the correspondence between semisimplicity and potential diagonalizability starts to fail. In particular, for k non-perfect:

1. semisimplicity does not behave well under field extensions;
2. semisimplicity does not imply potentially diagonalizable;
3. a Jordan decomposition may not exist for a linear operator.

The following lemma establishes the failures 1 and 2 explicitly for the (non-perfect) field $k = \mathbb{F}_p(t^p)$ where t is an indeterminate and not a p^{th} root. We can extend k to the field $V = \mathbb{F}_p(t)$ (note that this extension is not separable). Then V is a vector space over k of dimension p and we can consider the linear operator, $T : V \to V$, which acts as multiplication by t. Extending scalars of V to $V' := V \otimes_k V = \mathbb{F}_p(t)^p$ and considering the action of T on V' will yield the required counter-example.

Lemma 3.11. *Let $k = \mathbb{F}_p(t^p)$, $V = \mathbb{F}_p(t)$ and $V' = V \otimes_k V$. The linear operator T acts simply on V but does not act semisimply on V'.*

Proof. Using the basis $\{1, t, \ldots, t^{p-1}\}$, the matrix for T is:

$$T = \begin{pmatrix} 0 & 0 & \ldots & 0 & t^p \\ 1 & 0 & & 0 & 0 \\ 0 & 1 & \ddots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 \end{pmatrix}.$$

This has characteristic polynomial $X^p - t^p$ which is irreducible over k and hence T acts simply on V (hence, semisimply). However, if we extend k to V and let T act on V' then the characteristic polynomial of T is $X^p - t^p = (X - t)^p$ (in characteristic p). Clearly, $T - tI \not\equiv 0$ so the minimal polynomial of T is not square-free. Hence T does not act semisimply on V'. □

This establishes the failure 1 above since T acts semisimply on V but not on the extension to V' (cf. Remark 3.9). It is also clear from the above that T is not potentially diagonalizable (the minimal polynomial of T does not have distinct roots in $\bar{k}$). Hence T acts semisimply on V but T is not potentially diagonalizable, establishing the failure 2.

Finally we give an example showing that the Jordan decomposition of an operator may not exist when k is non-perfect.

Lemma 3.12. *Let $k = \mathbb{F}_2(t)$ and $V = k[x]/\langle (x^2 - t)^2 \rangle$. The linear operator $T : V \to V$ which acts as multiplication by x does not have a Jordan decomposition.*

Proof. The vector space V has a basis $\{1, x, x^2, x^3\}$. In this basis T is represented by the matrix:

$$T = \begin{pmatrix} 0 & 0 & 0 & t^2 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

This has characteristic and minimal polynomial $(X^2 - t)^2 \in k[X]$. Since the minimal polynomial is not square-free, T is not semisimple. It is also clear that T is not nilpotent since the characteristic polynomial of a nilpotent operator would be X^4. Hence we can try to find a Jordan decomposition $T = S + N$ for S semisimple and N nilpotent with $[S, N] = 0$. Suppose that such a decomposition exists. We will show that $N = 0$ which contradicts the fact that T is not semisimple.

Let $N \in \mathcal{M}_4(\mathbb{F}_2(t))$. Since S and N commute, they both commute with $T = S + N$ and so $TN = NT$. Computing NT and TN shows that N must be of the form:

$$N = \begin{pmatrix} a & t^2 d & t^2 c & t^2 b \\ b & a & t^2 d & t^2 c \\ c & b & a & t^2 d \\ d & c & b & a \end{pmatrix}, \qquad a, b, c, d \in \mathbb{F}_2(t).$$

Now N has characteristic polynomial $x^4 + a^4 + b^4 t^2 + c^4 t^4 + d^4 t^6$ but since N is nilpotent, its characteristic polynomial should be just x^4. Hence:

$$a^4 + b^4 t^2 + c^4 t^4 + d^4 t^6 = 0$$
$$\iff a^4 = (b^4 + c^4 t^2 + d^4 t^4) t^2$$

and so $a \in \mathbb{F}_2(\sqrt{t})$ (but not in $\mathbb{F}_2(t)^\times$) which implies that $a = 0$. Then we have $b^4 = (c^4 + d^4 t^2) t^2$ and so now $b = 0$ by the same argument. Finally, $c^4 = d^4 t^2$ and so $c = d = 0$ and hence $N = 0$. Thus $T = S$ is semisimple which is a contradiction. Hence no such decomposition exists. $\square$

4. Jordan-Chevalley Decomposition

4.1. **Semisimple and Reductive Lie Algebras.** Theorem 3.10 provides a Jordan decomposition for elements of the Lie algebra $\mathrm{End}(V) = \mathfrak{gl}(V)$. It is possible to extend the notion of Jordan decomposition to other Lie algebras. In particular, the next section will show that this can be done for reductive Lie algebras. We first discuss semisimple Lie algebras. We refer the reader to Humphreys and Borel [Hum78, Bor91] for the basic facts of Lie algebras and linear algebraic groups.

The adjoint representation of a Lie algebra $\mathfrak{g}$ is the linear map $\phi : \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$ that sends $x \in \mathfrak{g}$ to the endomorphism $\mathrm{ad}_x : \mathfrak{g} \to \mathfrak{g}$ defined by $\mathrm{ad}_x(y) = [x, y]$ for all $y \in \mathfrak{g}$.

Recall that the radical of a Lie algebra is its maximal solvable ideal. We will denote this by $\mathrm{rad}(\mathfrak{g})$. The following theorem characterizes semisimple Lie algebras.

Theorem 4.1 (Semisimple Lie Algebras)**.** *Let $\mathfrak{g}$ be a Lie algebra. We call $\mathfrak{g}$ semisimple if it satisfies any of the following equivalent conditions.*

i) $\mathrm{rad}(\mathfrak{g}) = 0$.

ii) The Killing form $\kappa_{\mathfrak{g}} : \mathfrak{g} \times \mathfrak{g} \to \mathbb{C}$ *defined by* $\kappa_{\mathfrak{g}}(x, y) = \mathrm{Tr}(\mathrm{ad}_x \cdot \mathrm{ad}_y)$ *is non-degenerate.*

iii) $\mathfrak{g}$ *is a finite direct sum of non-abelian simple Lie algebras*

The content of Theorem 4.1 is the equivalence of the three conditions. We refer the reader to [Hum78, Sec. 5.1] for the equivalence of (i) and (ii). The equivalence of (i) and (iii) follows from Weyl's theorem:

Theorem 4.2 (Weyl's Theorem)**.** *Let* $\mathfrak{g}$ *be a complex semisimple Lie algebra and* $\phi : \mathfrak{g} \to \mathfrak{gl}(V)$ *a finite-dimensional representation. Then* ϕ *is completely reducible.*

For a proof see for instance [Hum78, FH91].

Let $x \in \mathfrak{g} \subset \mathfrak{gl}(V)$ and $x = s + n$ be the Jordan decomposition of x (in $\mathfrak{gl}(V)$). A priori, it is not clear that $s, n \in \mathfrak{g}$. The following lemma shows that we at least have $[s, \mathfrak{g}] \subset [\mathfrak{g}, \mathfrak{g}]$ and $[n, \mathfrak{g}] \subset [\mathfrak{g}, \mathfrak{g}]$.

Lemma 4.3. *Let* $x \in \mathfrak{g} \subset \mathfrak{gl}(V)$ *and* $x = s + n$ *be the Jordan decomposition of* x*. Then* $\mathrm{ad}_s(\mathfrak{g}) \subset \mathfrak{g}$ *and* $\mathrm{ad}_n(\mathfrak{g}) \subset \mathfrak{g}$.

Proof. First observe that $\mathrm{ad}_x = \mathrm{ad}_s + \mathrm{ad}_n$ is the Jordan decomposition of ad_x. This follows from the fact that if x is semisimple (respectively, nilpotent) then ad_x is semisimple (respectively, nilpotent). The fact that ad_s and ad_n commute follows from the fact that s and n commute:

$$[\mathrm{ad}_s, \mathrm{ad}_n] = \mathrm{ad}_{[s,n]} = 0.$$

Now a result from linear algebra (see [Hum78, Sec. 4.2] for instance) tells us that ad_s and ad_n can be expressed as polynomials in ad_x. Since $\mathrm{ad}_x(\mathfrak{g}) \subset \mathfrak{g}$ we must therefore have $\mathrm{ad}_s(\mathfrak{g}) \subset \mathfrak{g}$ and $\mathrm{ad}_n(\mathfrak{g}) \subset \mathfrak{g}$. □

Remark 4.4. Note that the proof of Lemma 4.3 implies that if $x(W) \subset W$ for some $\mathfrak{g}$-submodule W of V, then $s(W) \subset W$ and $n(W) \subset W$.

Weyl's theorem has the following important consequence:

Theorem 4.5. *Let* $\mathfrak{g} \subset \mathfrak{gl}(V)$ *be a semisimple Lie algebra and* $x \in \mathfrak{g}$*. Then* $\mathfrak{g}$ *contains the semisimple and nilpotent parts of* x*.*

Proof. Let $x = s + n$ be the Jordan decomposition of x in $\mathfrak{gl}(V)$. The idea of the proof is to construct $\mathfrak{g}$ in a way that makes it clear that it contains s and n. There are two subalgebras we need to consider.

Firstly, set $N_{\mathfrak{gl}(V)}(\mathfrak{g}) = \{x \in \mathfrak{gl}(V) | [x, \mathfrak{g}] \subset \mathfrak{g}\}$. This is the normalizer of $\mathfrak{g}$ in $\mathfrak{gl}(V)$. By Lemma 4.3, $s, n \in N_{\mathfrak{gl}(V)}(\mathfrak{g})$. It is clear that $\mathfrak{g} \subset N_{\mathfrak{gl}(V)}(\mathfrak{g})$ however $\mathfrak{g} \neq N_{\mathfrak{gl}(V)}(\mathfrak{g})$ as $N_{\mathfrak{gl}(V)}(\mathfrak{g})$ contains the scalar matrices, for example.

Secondly, we consider the subalgebras $\mathfrak{g}_W = \{y \in \mathfrak{gl}(V) | y(W) \subset W \text{ and } \mathrm{Tr}(y|_W) = 0\}$ for a $\mathfrak{g}$-submodule W of V. Since $\mathfrak{g}$ is semisimple, $\mathfrak{g} \subset \mathfrak{sl}(V)$ and since W is a $\mathfrak{g}$-submodule, every element of $\mathfrak{g}$ stabilizes W. Thus $\mathfrak{g} \subset \mathfrak{g}_W$. Remark 4.4 shows that $s(W) \subset W$ and $n(W) \subset W$ so $s, n \in \mathfrak{g}_W$ for all submodules W.

We have now constructed a family of subalgebras of $\mathfrak{gl}(V)$ each of which contains s, n and $\mathfrak{g}$. We now show that

$$\mathfrak{g}^* := N_{\mathfrak{gl}(V)}(\mathfrak{g}) \bigcap \left(\bigcap_W \mathfrak{g}_W \right),$$

is in fact equal to $\mathfrak{g}$. By our discussion above, s and n lie in the intersection so this will prove the result. Clearly $\mathfrak{g}$ is an ideal of $\mathfrak{g}^*$ and since $\mathfrak{g}^*$ is a finite-dimensional $\mathfrak{g}$-module, Weyl's Theorem allows us to write

$$\mathfrak{g}^* = \mathfrak{g} \oplus M$$

for some complementary $\mathfrak{g}$-submodule M. Since $\mathfrak{g}^* \subset N_{\mathfrak{gl}(V)}(\mathfrak{g})$ we have $[\mathfrak{g}, \mathfrak{g}^*] = \mathfrak{g}$ which implies that $[\mathfrak{g}, M] = 0$. If W is an irreducible $\mathfrak{g}$-submodule of V then Schur's lemma implies that any $y \in M$ acts as a scalar on W. In fact, y acts as 0 on W since we have $\mathrm{Tr}(y|_W) = 0$. Writing V as a direct sum of irreducible $\mathfrak{g}$-submodules (which is possible by Weyl's Theorem) we find that y acts as 0 on V and so $y = 0$. Since y was arbitrary in M this shows that $M = 0$. Hence $\mathfrak{g}^* = \mathfrak{g}$. □

The importance of Theorem 4.5 is that it will allow us to unambiguously define semisimple and nilpotent elements of any reductive Lie algebra. The following corollary is the key to doing this:

Corollary 4.6. *Let $\mathfrak{g} \subset \mathfrak{gl}(V)$ be a semisimple Lie algebra. Then $x \in \mathfrak{g}$ is semisimple (respectively, nilpotent) if, and only if, ad_x is semisimple (respectively, nilpotent).*

Proof. One direction of this proof is easy and we have already used it in Lemma 4.3 (see [Hum78, Sec. 3.2] for details). For the other direction, suppose $\mathrm{ad}_x = \mathrm{ad}_s + \mathrm{ad}_n$ is the Jordan decomposition of ad_x. We know by Lemma 4.3 and Theorem 4.5 that $x = s + n$ is the Jordan decomposition of x in $\mathfrak{g}$. If ad_x is semisimple then $\mathrm{ad}_n = 0$ and since ad is a faithful representation, this implies that $n = 0$. Hence x is semisimple. Similarly, if ad_x is nilpotent then $s = 0$ and so x is also nilpotent. □

Definition 4.7 (Reductive Lie Algebra)**.** *A Lie algebra, $\mathfrak{g}$, is called* reductive *if* $\mathrm{rad}(\mathfrak{g}) = \mathfrak{z}$ *where $\mathfrak{z}$ is the center of* $\mathfrak{g}$.

A reductive Lie algebra, $\mathfrak{g}$, can always be decomposed as:

$$\mathfrak{g} = \mathfrak{z} \oplus \mathfrak{s}$$

where $\mathfrak{s}$ is a semisimple subalgebra of $\mathfrak{g}$.

4.2. **Jordan-Chevalley Decomposition for Reductive Lie Algebras.** We now explain how to extend the Jordan decomposition to elements of an arbitrary reductive Lie algebra by using the adjoint representation. We assume throughout that k is perfect.

We use the adjoint representation to define semisimple and nilpotent elements of $\mathfrak{g}$:

Definition 4.8 (Semisimple and Nilpotent Elements)**.** *Let $\mathfrak{g}$ be a Lie algebra over k. We call $x \in \mathfrak{g}$* semisimple *(respectively,* nilpotent*) if ad_x is semisimple (respectively, nilpotent).*

Remark 4.9. Note that a priori, this definition may not be consistent with Definitions 3.2 and 3.6 in the special case that $\mathfrak{g} \subset \mathfrak{gl}(V)$. In this case we rely on Corollary 4.6 which tells us that these definitions are consistent.

One can also show that $x \in \mathfrak{g}$ is semisimple if it is conjugate to an element of $\mathfrak{h} \otimes_k \bar{k}$ where $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}$.

In order to extend the Jordan decomposition to arbitrary reductive Lie algebras we first need to discuss derivations of a Lie algebra.

Definition 4.10 (Derivation). *A map $\delta \in \mathrm{End}(\mathfrak{g})$ is a* derivation *if it satisfies the Leibniz rule:*

$$\delta([a, b]) = [a, \delta(b)] + [\delta(a), b], \quad \forall a, b \in \mathfrak{g}.$$

The set of all derivations on $\mathfrak{g}$ is denoted $\mathrm{Der}(\mathfrak{g})$ and is a Lie subalgebra of $\mathfrak{gl}(\mathfrak{g})$.

Thanks to the Jacobi identity we have the following lemma:

Lemma 4.11. *Let $\mathfrak{g}$ be a Lie algebra and $x \in \mathfrak{g}$. The map $ad_x \in \mathrm{End}(\mathfrak{g})$ is a derivation.*

Proof. Let $y, z \in \mathfrak{g}$. The Jacobi identity and skew-symmetry give us:

$$\begin{aligned} &[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0 \\ \implies &[x, [y, z]] = [[x, y], z] + [y, [x, z]] \\ \implies &\mathrm{ad}_x([y, z]) = [\mathrm{ad}_x(y), z] + [y, \mathrm{ad}_x(z)]. \end{aligned}$$

Hence ad_x is a derivation. □

We call a derivation of the form ad_x, for some $x \in \mathfrak{g}$, an *inner derivation* and denote by $\mathrm{ad}(\mathfrak{g})$ the set of all inner derivations. The above lemma shows that $\mathrm{ad}(\mathfrak{g}) \subseteq \mathrm{Der}(\mathfrak{g})$. In the case of semisimple Lie algebras, $\mathrm{ad}(\mathfrak{g}) = \mathrm{Der}(\mathfrak{g})$.

Lemma 4.12. *Let $\mathfrak{g}$ be a semisimple Lie algebra. Every $\delta \in \mathrm{Der}(\mathfrak{g})$ is of the form ad_x for some $x \in \mathfrak{g}$.*

Proof. Since $\mathfrak{g}$ is semisimple, the map $\mathfrak{g} \to \mathrm{ad}(\mathfrak{g})$ is a Lie algebra isomorphism. This implies that $\mathrm{ad}(\mathfrak{g})$ is semisimple and so it has non-degenerate Killing form by (ii) of Definition 4.8. By Lemma 4.11, $\mathrm{ad}(\mathfrak{g}) \subseteq \mathrm{Der}(\mathfrak{g})$. In fact, $\mathrm{ad}(\mathfrak{g})$ is an ideal of $\mathrm{Der}(\mathfrak{g})$. This can be seen as follows. Let $\delta \in \mathrm{Der}(\mathfrak{g})$ and $\mathrm{ad}_x \in \mathrm{ad}(\mathfrak{g})$. Then for $y \in \mathfrak{g}$,

$$\begin{aligned} &[\delta, \mathrm{ad}_x](y) \\ =&(\delta \circ \mathrm{ad}_x)(y) - (\mathrm{ad}_x \circ \delta)(y) \\ =&\delta([x, y]) - [x, \delta(y)] \\ =&[\delta(x), y] + [x, \delta(y)] - [x, \delta(y)] \\ =&[\delta(x), y] \\ =&\, \mathrm{ad}_{\delta(x)}(y). \end{aligned}$$

So $[\delta, \mathrm{ad}_x] = \mathrm{ad}_{\delta(x)} \in \mathrm{ad}(\mathfrak{g})$.

Since $\mathrm{ad}(\mathfrak{g})$ is an ideal of $\mathrm{Der}(\mathfrak{g})$, the Killing form on $\mathrm{ad}(\mathfrak{g})$ is just the restriction of the Killing form on $\mathrm{Der}(\mathfrak{g})$. Hence we can consider the orthogonal subspace $\mathrm{ad}(\mathfrak{g})^\perp$ in $\mathrm{Der}(\mathfrak{g})$. Note $\mathrm{ad}(\mathfrak{g})^\perp$ consists of those $\delta \in \mathrm{Der}(\mathfrak{g})$ satisfying $\kappa(\delta, \mathrm{ad}_x) = 0$ for all $\mathrm{ad}_x \in \mathrm{ad}(\mathfrak{g})$. Since the Killing form is non-degenerate on $\mathrm{ad}(\mathfrak{g})$, we have $\mathrm{ad}(\mathfrak{g}) \cap \mathrm{ad}(\mathfrak{g})^\perp = 0$ and since both $\mathrm{ad}(\mathfrak{g})$ and $\mathrm{ad}(\mathfrak{g})^\perp$ are ideals of $\mathrm{Der}(\mathfrak{g})$, $[\mathrm{ad}(\mathfrak{g}), \mathrm{ad}(\mathfrak{g})^\perp] = 0$. Now given any derivation $\delta \in \mathrm{ad}(\mathfrak{g})^\perp$ we have for all $x \in \mathfrak{g}$:

$$0 = [\delta, \mathrm{ad}_x] = \mathrm{ad}_{\delta(x)}\,.$$

Since $\mathfrak{g} \to \mathrm{ad}(\mathfrak{g})$ is an isomorphism, this implies that $\delta(x) = 0$ for all $x \in \mathfrak{g}$, so $\delta = 0$. Thus $\mathrm{ad}(\mathfrak{g})^\perp = 0$ and so $\mathrm{ad}(\mathfrak{g}) = \mathrm{Der}(\mathfrak{g})$. □

Theorem 4.13 (Jordan-Chevalley Decomposition). *Let $\mathfrak{g}$ be a reductive Lie algebra over k. Then for any $x \in \mathfrak{g}$ there exists a semisimple element $s \in \mathfrak{g}$ and a nilpotent element $n \in \mathfrak{g}$ such that*

(i) $x = s + n$;

(ii) $[s, n] = 0$.

Moreover, (s, n) is the unique pair of elements in $\mathfrak{g}$ satisfying (i) and (ii).

Proof. Using Theorem 3.10, we can write $\mathrm{ad}_x = S + N$ where $S, N \in \mathfrak{gl}(\mathfrak{g})$, S is semisimple and N is nilpotent. The key problem is to show that $S = \mathrm{ad}_s$ and $N = \mathrm{ad}_n$ for some elements $s, n \in \mathfrak{g}$. Given that ad_x is a derivation, it is enough to show that the set $\mathrm{Der}(\mathfrak{g})$ contains the semisimple and nilpotent parts of its elements (in $\mathrm{End}(\mathfrak{g})$). This will imply that $S, N \in \mathrm{Der}(\mathfrak{g}) = \mathrm{ad}(\mathfrak{g})$ (by Lemma 4.12) so that $S = \mathrm{ad}_s$ and $N = \mathrm{ad}_n$ for some $s, n \in \mathfrak{g}$.

Given $\mathrm{ad}_x = S + N$ as above, we will show that $S \in \mathrm{Der}(\mathfrak{g})$. Firstly, note that ad_x and S have the same eigenvalues. Hence the generalized eigenspaces $\mathfrak{g}_\lambda = \{y \in \mathfrak{g} | (\mathrm{ad}_x - \lambda)^n) y = 0, n \in \mathbb{Z}\}$ of ad_x coincide with the generalized eigenspaces of S. We can write $\mathfrak{g}$ as a direct sum of generalized eigenspaces:

$$\mathfrak{g} = \bigoplus_{\lambda \in \bar{k}} \mathfrak{g}_\lambda.$$

Given two generalized eigenspaces $\mathfrak{g}_\alpha, \mathfrak{g}_\beta$ we have

$$\mathfrak{g}_\alpha \mathfrak{g}_\beta \subseteq \mathfrak{g}_{\alpha+\beta},$$

which follows from the formula $(\mathrm{ad}_x - (\alpha + \beta))^n (yz) = \sum_{i=0}^n \binom{n}{i} ((\mathrm{ad}_x - a)^{n-i} y) \cdot ((\mathrm{ad}_x - b)^i z)$ for $y, z \in \mathfrak{g}$ (this formula is proved in the following Lemma 4.14). Since S acts on $\mathfrak{g}_\lambda$ by multiplication by λ, we have for $x \in \mathfrak{g}_\alpha, y \in \mathfrak{g}_\beta$

$$S(xy) = (\alpha + \beta) xy.$$

However we also have:

$$\begin{aligned} S(x)y + xS(y) &= \alpha xy + x\beta y \\ &= (\alpha + \beta) xy. \end{aligned}$$

Hence $S(xy) = S(x)y + xS(y)$. Since $\mathfrak{g} = \oplus \mathfrak{g}_\lambda$ this shows that S is a derivation on $\mathfrak{g}$. Hence $S \in \mathrm{Der}(\mathfrak{g})$ and therefore $N \in \mathrm{Der}(\mathfrak{g})$.

Since $\mathfrak{g}$ is semisimple, Lemma 4.12 implies that $S = \mathrm{ad}_s$ and $N = \mathrm{ad}_n$ for some $s, n \in \mathfrak{g}$. Moreover, these must be unique since ad is a faithful representation when $\mathfrak{g}$ is semisimple. □

Lemma 4.14. *Let $\mathfrak{g}$ be a semisimple Lie algebra over a perfect field k. Given $\alpha, \beta \in k$, $x, y \in \mathfrak{g}$ and $\delta \in \mathrm{Der}(\mathfrak{g})$ we have:*

$$\big(\delta - (\alpha+\beta)\big)^n (xy) = \sum_{i=0}^{n} \binom{n}{i} \big((\delta - a)^{n-i} x\big) \cdot \big((\delta - b)^i y\big).$$

Proof. This formula is essentially an equivalent of the binomial theorem for derivations. We prove it by induction on n. It is easy to verify that the formula holds in the cases $n = 0, 1$. Suppose that the formula holds for all $0 \le k \le n$. Then for $n+1$ we have:

$$\begin{aligned}
&\big(\delta - (\alpha+\beta)\big)\big(\delta - (\alpha+\beta)\big)^n (xy) \\
&= \big(\delta - (\alpha+\beta)\big) \left(\sum_{i=0}^{n} \binom{n}{i} \big((\delta-\alpha)^{n-i} x\big)\big((\delta-\beta)^i y\big) \right) \\
&= \delta \left(\sum_{i=0}^{n} \binom{n}{i} \big((\delta-\alpha)^{n-i} x\big)\big((\delta-\beta)^i y\big) \right) - (\alpha+\beta) \left(\sum_{i=0}^{n} \binom{n}{i} \big((\delta-\alpha)^{n-i} x\big)\big((\delta-\beta)^i y\big) \right) \\
&= \sum_{i=0}^{n} \binom{n}{i} \delta\big((\delta-\alpha)^{n-i} x\big)\big((\delta-\beta)^i y\big) + \sum_{i=0}^{n} \binom{n}{i} \big((\delta-\alpha)^{n-i} x\big)\delta\big((\delta-\beta)^i y\big) \\
&\qquad - (\alpha+\beta) \left(\sum_{i=0}^{n} \binom{n}{i} \big((\delta-\alpha)^{n-i} x\big)\big((\delta-\beta)^i y\big) \right) \\
&= \sum_{i=0}^{n} \binom{n}{i} (\delta-\alpha)\big((\delta-\alpha)^{n-i} x\big)\big((\delta-\beta)^i y\big) + \sum_{i=0}^{n} \binom{n}{i} \big((\delta-\alpha)^{n-i} x\big)(\delta-\beta)\big((\delta-\beta)^i y\big) \\
&= \sum_{i=0}^{n} \binom{n}{i} \big((\delta-\alpha)^{n+1-i} x\big)\big((\delta-\beta)^i y\big) + \sum_{i=0}^{n} \binom{n}{i} \big((\delta-\alpha)^{n-i} x\big)\big((\delta-\beta)^{i+1} y\big) \\
&= (\delta-\alpha)^{n+1}(x)y + \sum_{i=1}^{n} \binom{n}{i} \big((\delta-\alpha)^{n+1-i} x\big)\big((\delta-\beta)^i y\big) \\
&\qquad + \sum_{i=0}^{n-1} \binom{n}{i} \big((\delta-\alpha)^{n-i} x\big)\big((\delta-\beta)^{i+1} y\big) + x(\delta-\beta)^{n+1} y \\
&= (\delta-\alpha)^{n+1}(x)y + x(\delta-\beta)^{n+1}(y) \\
&\qquad + \sum_{i=1}^{n} \binom{n}{i} \big((\delta-\alpha)^{n+1-i} x\big)\big((\delta-\beta)^i y\big) + \sum_{i=1}^{n} \binom{n}{i-1} \big((\delta-\alpha)^{n+1-i} x\big)\big((\delta-\beta)^i y\big) \\
&= (\delta-\alpha)^{n+1}(x)y + x(\delta-\beta)^{n+1}(y) + \sum_{i=1}^{n} \left[\binom{n}{i} + \binom{n}{i-1} \right] \big((\delta-\alpha)^{n+1-i} x\big)\big((\delta-\beta)^i y\big) \\
&= (\delta-\alpha)^{n+1}(x)y + x(\delta-\beta)^{n+1}(y) + \sum_{i=1}^{n} \binom{n+1}{i} \big((\delta-\alpha)^{n+1-i} x\big)\big((\delta-\beta)^i y\big) \\
&= \sum_{i=0}^{n+1} \binom{n+1}{i} \big((\delta-\alpha)^{n+1-i} x\big)\big((\delta-\beta)^i y\big)
\end{aligned}$$

$\square$

5. Factorisation of Commutative Polynomials

The proof of Theorem 3.1 relies on the fact that polynomials in $\bar{k}[x]$ can be completely factorized into linear factors. The extension of this to the differential setting requires a differential analogue of Hensel's lemma. In this section we discuss Hensel's lemma in the commutative polynomial ring with coefficients in the field $\mathbb{C}(\!(t)\!)$.

5.1. **Hensel's Lemma.** The field $\mathbb{C}(\!(t)\!)$ is the field of formal Laurent series in t with coefficients in $\mathbb{C}$. That is, for non-zero $f \in \mathbb{C}(\!(t)\!)$ we can write:

$$f(t) = \sum_{i=-m}^{\infty} a_i t^i, \qquad m \in \mathbb{Z}, a_i \in \mathbb{C}, a_{-m} \neq 0. \tag{5.6}$$

We can equip $\mathbb{C}(\!(t)\!)$ with the t-adic valuation v_t. This is a function $v_t : \mathbb{C}(\!(t)\!) \to \mathbb{Z} \cup \{\infty\}$ which for f given by (5.6) is defined by:

$$v_t(f) = -m,$$

and

$$v_t(0) := \infty.$$

With respect to this valuation, $\mathbb{C}(\!(t)\!)$ is complete. We define the valuation ring of $\mathbb{C}(\!(t)\!)$ as:

$$\mathbb{C}[\![t]\!] := \{f \in \mathbb{C}(\!(t)\!) | v_t(f) \geq 0\}.$$

The ring $\mathbb{C}[\![t]\!]$ consists of formal power series in t with coefficients in $\mathbb{C}$. This ring is a local ring which means that it has a unique maximal ideal:

$$t\mathbb{C}[\![t]\!] := \{f \in \mathbb{C}(\!(t)\!) | v_t(f) \geq 1\}.$$

Since $t\mathbb{C}[\![t]\!]$ is a maximal ideal, the quotient:

$$\mathbb{C}[\![t]\!]/t\mathbb{C}[\![t]\!] \cong \mathbb{C}$$

is a field. This is the residue field of $\mathbb{C}(\!(t)\!)$. The above discussion establishes that $\mathbb{C}(\!(t)\!)$ is a local field. That is,

1. The field $\mathbb{C}(\!(t)\!)$ is complete with respect to a discrete valuation.
2. The residue field of $\mathbb{C}(\!(t)\!)$ is perfect.

We now introduce the classical version of Hensel's lemma:

Lemma 5.1 (Hensel's Lemma)**.** *Let $f \in \mathbb{C}[\![t]\!][x]$ and $\bar{f} \in \mathbb{C}[x]$ be the reduction of f mod $t\mathbb{C}[\![t]\!]$. If there exists a factorization:*

$$\bar{f} = \bar{g}\bar{h}, \qquad \bar{g}, \bar{h} \in \mathbb{C}[x],$$

such that $\gcd(\bar{g}, \bar{h}) = 1$*, then this lifts to a factorization:*

$$f = gh, \qquad g, h \in \mathbb{C}[\![t]\!][x],$$

with $\deg(g) = \deg(\bar{g})$ *and*

$$g \equiv \bar{g} \pmod{t\mathbb{C}[\![t]\!]}, \quad h \equiv \bar{h} \pmod{t\mathbb{C}[\![t]\!]}.$$

Proof. We follow the proof given in Neukirch [Neu99, Sec. 4]. Let $g_0 = \bar{g}$ and $h_0 = \bar{h}$. We would like to build a sequence of functions:

$$g_n = g_0 + p_1 t + p_2 t^2 + \cdots + p_{n-1} t^{n-1} + p_n t^n, \qquad p_i \in \mathbb{C}[\![t]\!][x] \tag{5.7}$$

$$h_n = h_0 + q_1 t + q_2 t^2 + \cdots + q_{n-1} t^{n-1} + q_n t^n, \qquad q_i \in \mathbb{C}[\![t]\!][x] \tag{5.8}$$

for all $n \geq 0$, which satisfy:

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

If we can do this, then by letting $n \to \infty$ we will obtain functions $g, h \in \mathbb{C}[\![t]\!][x]$ such that $f = gh$.

Suppose that we know the p_i and q_i for $1 \leq i \leq n-1$ (i.e. we have g_{n-1}, h_{n-1} such that $f \equiv g_{n-1}h_{n-1} \pmod{t^n}$), we will give a procedure for finding p_n and q_n. Firstly, observe from (5.7) and (5.8) that:

$$g_n = g_{n-1} + p_n t^n, \qquad h_n = h_{n-1} + q_n t^n. \tag{5.9}$$

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

$$\begin{aligned}
& f \equiv g_n h_n && \pmod{t^{n+1}} \\
\implies\; & f \equiv (g_{n-1} + p_n t^n)(h_{n-1} + q_n t^n) && \pmod{t^{n+1}} \\
\implies\; & f \equiv g_{n-1}h_{n-1} + g_{n-1}q_n t^n + h_{n-1}p_n t^n + p_n q_n t^{2n} && \pmod{t^{n+1}} \\
\implies\; & f \equiv g_{n-1}h_{n-1} + g_{n-1}q_n t^n + h_{n-1}p_n t^n && \pmod{t^{n+1}}.
\end{aligned}$$

Rearranging we see that:

$$\begin{aligned}
& f - g_{n-1}h_{n-1} \equiv (g_{n-1}q_n + h_{n-1}p_n)t^n && \pmod{t^{n+1}} \\
\implies\; & \frac{f - g_{n-1}h_{n-1}}{t^n} \equiv g_{n-1}q_n + h_{n-1}p_n && \pmod{t}.
\end{aligned}$$

For notational convenience, we set:

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

Now from (5.7) and (5.8) we have $g_{n-1} \equiv g_0 \pmod{t}$ and $h_{n-1} \equiv h_0 \pmod{t}$ so the condition on p_n and q_n becomes:

$$f_n \equiv g_0 q_n + h_0 p_n.$$

Since $\gcd(g_0, h_0) = 1$, there are unique p_n and q_n satisfying this with $\deg(p_n) < \deg(g_0)$ and $\deg(q_n) < \deg(h_0)$. To show this explicitly let:

$$a g_0 + b h_0 \equiv 1 \pmod{t}$$

for $a, b \in \mathbb{C}[\![t]\!][x]$. Then we have:

$$g_0(a f_n) + h_0(b f_n) \equiv f_n \pmod{t}. \tag{5.10}$$

At this point it would be nice if we could set $q_n = a f_n$, $p_n = b f_n$ but we require $\deg(p_n) < \deg(g_0)$ and $\deg(q_n) < \deg(h_0)$. In order to meet these conditions, we use the division algorithm to write:

$$b f_n \equiv Q g_0 + R \pmod{t}$$

with $\deg(R) < \deg(g_0)$. Equation (5.10) then becomes:

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

Setting

$$p_n \equiv R \qquad \pmod t,$$
$$q_n \equiv af_n + Qh_0 \qquad \pmod t$$

then gives us the required g_n and h_n. $\square$

5.2. **Newton-Puiseux Theorem.** We can use Hensel's lemma to prove the following result about finite extensions of $\mathbb{C}(\!(t)\!)$.

Theorem 5.2. *Every finite extension of* $\mathbb{C}(\!(t)\!)$ *is a field of the form* $\mathbb{C}(\!(t^{1/m})\!)$ *for some* $m \in \mathbb{N}$.

Proof. The proof follows [vdPS03, Sec. 3.1]. We first remark that the extension $\mathbb{C}(\!(t^{1/m})\!)$ is a Galois extension of $\mathbb{C}(\!(t)\!)$. This is because $\mathbb{C}(\!(t)\!)$ is perfect and $\mathbb{C}(\!(t^{1/m})\!)$ is a splitting field for the polynomial $x^n - t$. This implies that if a polynomial in $\mathbb{C}(\!(t)\!)[x]$ has a root in $\mathbb{C}(\!(t^{1/m})\!)$ then it has all its roots in $\mathbb{C}(\!(t^{1/m})\!)$. Thus, in order to prove Theorem 5.2, it suffices to show that any polynomial:

$$x^d + a_1x^{d-1} + \cdots + a_{d-1}x + a_d \in \mathbb{C}(\!(t)\!)[x]$$

has a root in $\mathbb{C}(\!(t^{1/m})\!)$ for some $m \in \mathbb{N}$.

In order to apply Hensel's lemma, we first need to clear the coefficients, a_i, of any negative powers of t. Making the substitution

$$x = t^{-\lambda}X$$

where

$$\lambda := \min\left\{ \frac{v(a_i)}{i} \,\middle|\, 1 \leq i \leq d \right\}$$

achieves this. We obtain a polynomial whose coefficients are now formal power series in $t^{1/m}$, where m is the denominator of λ:

$$f(X) = X^d + b_1X^{d-1} + \cdots + b_{d-1}X + b_d \in \mathbb{C}[\![t^{1/m}]\!][X].$$

Let $\bar{f}$ be the reduction of $f \pmod{t^{1/m}}$. We are now able to apply Hensel's lemma to find a root of f. We use induction on the degree d.

If $\bar{f}$ has two distinct roots in $\mathbb{C}$ then we have a factorization $\bar{f} = \bar{g}\bar{h}$ with $\deg(\bar{g}), \deg(\bar{h}) < d$. This lifts to a factorization $f = gh$ with $\deg(g), \deg(h) < d$ and we can apply the inductive hypothesis to g, h to conclude that f has a root in $\mathbb{C}(\!(t^{1/m})\!)$.

If $\bar{f}$ does not have two distinct roots then we have $\bar{f} = (X - c_0)^d$ for some $c_0 \in \mathbb{C}$. By our choice of λ, we have $\min\{v(b_i)\} = 0$ (i.e. at least one b_i has non-zero constant term). Hence $\bar{f} \neq X^d$ so $c_0 \neq 0$. But now $\bar{f}$ has $d+1$ non-zero terms and so each of the b_i has non-zero constant term. In particular, $v(b_1) = 0$ which implies that $v(a_1) = \lambda$ and so λ is an integer and $m = 1$ (so we have not actually extended our field). In this case, we write

$$f = (X - c_0)^d + e_1(X - c_0)^{d-1} + \cdots + e_{d-1}(X - c_0) + e_d. \tag{5.11}$$

Now since $\bar{f} = (X - c_0)^d$ we must have $v(e_i) > 0$ for $1 \leq i \leq d$. Let $\lambda_1 = \min\left\{ \frac{v(e_i)}{i} \,\middle|\, 1 \leq i \leq d \right\}$, m_1 be the denominator of λ_1, and make the substitution

$$X = c_0 + t^{\lambda_1}X_1.$$

This yields a polynomial

$$f(X_1) = (t^{\lambda_1}X_1)^d + e_1(t^{\lambda_1}X_1)^{d-1} + \cdots + e_{d-1}(t^{\lambda_1}X_1) + e_d.$$

As above, there are two cases: either $f \pmod{t^{1/m_1}}$ has two distinct roots in $\mathbb{C}$ or λ_1 is an integer and $m_1 = 1$. In the first case we are finished by induction. In the second case we make the substitution:

$$X_1 = c_1 + t^{\lambda_2}X_2.$$

This is equivalent to the substitution:

$$X = c_0 + c_1t^{\lambda_1} + t^{\lambda_1+\lambda_2}X_2$$

in (5.11). Continuing in this fashion we will either find some substitution for which $\bar{f}$ has two distinct roots or we will generate an infinite expression:

$$c_0 + c_1t^{\lambda_1} + c_2t^{\lambda_1+\lambda_2} + \cdots$$

with $\lambda_i \in \mathbb{Z}$ such that

$$f = (X - (c_0 + c_1t^{\lambda_1} + c_2t^{\lambda_1+\lambda_2} + \cdots))^d.$$

In either case, f has a root in $\mathbb{C}((t^{1/m}))$ for some $m \in \mathbb{N}$. □

6. The Differential Setting

We now turn to the main topic of this thesis which is differential operators and differential modules. The pinnacle of this section will be a new proof of the Hukuhara-Levelt-Turrittin theorem. This theorem gives a Jordan decomposition for differential operators and should be considered an analogue of Theorem 3.10. Despite the strong resemblance in the statement of the theorems, the current methods of proof are vastly different. The goal of this section is to rectify this by providing a new proof analogous to the proof of Jordan decomposition in the linear setting. Along the way we will introduce differential operators, differential modules, and differential polynomials and point out the similarities and differences to the linear setting.

6.1. **Differential Fields.** Our discussion of Jordan decomposition in the linear setting began by studying linear operators on vector spaces. In this new setting, the role of linear operators will be played by differential operators on a vector space. In order to define differential operators, the vector space itself must have an underlying differential structure. This differential structure is provided by the field over which the vector space is defined.

Definition 6.1 (Differential Field)**.** *A* differential field*, k, is a field equipped with a derivation, i.e. an additive map $d : k \to k$ satisfying the Leibniz rule:*

$$d(ab) = ad(b) + d(a)b, \qquad \forall a, b \in k.$$

One can generalize this to define a differential ring (see [vdPS03, §1.1]) however we will not need this level of generality here. Before we provide some interesting examples of differential fields, it is worth noting that every field can be trivially made a differential field by letting d be the zero map.

Example 6.2. We can take the field of rational functions $\mathbb{C}(t)$ or the field of formal Laurent series $\mathbb{C}((t))$ with derivation $d = p\frac{d}{dt}$ for some polynomial $p \in \mathbb{C}[t]$. In what follows, we generally work over the field $\mathcal{K} = \mathbb{C}((t))$ with derivation $d = t\frac{d}{dt}$.

The kernel of d forms a special subfield of k.

Definition 6.3 (Constants)**.** *The set $C := \{c \in k : d(c) = 0\}$ is a subfield of k called the* field of constants *of k.*

It is easy to verify that C is a subfield. If k has characteristic 0, then the next lemma shows that any element algebraic over C is also a constant.

Lemma 6.4. *Let k be a differential field of characteristic 0. If c is algebraic over C, then $d(c) = 0$.*

Proof. Since c is algebraic over C, there exists a minimal polynomial P:

$$P(x) = x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n, \qquad a_i \in C,$$

such that $P(c) = 0$. Applying the derivation to both sides we see that:

$$nc^{n-1}d(c) + d(a_1)c^{n-1} + (n-1)a_1c^{n-2}d(c) + \cdots + d(a_{n-1})c + a_{n-1}d(c) + d(a_n) = 0.$$

Since $a_i \in C$, $d(a_i) = 0$ for $1 \leq i \leq n$. So we have:

$$d(c)\big(nc^{n-1} + (n-1)a_1c^{n-2} + \cdots + a_{n-1}\big) = 0.$$

So either $d(c) = 0$ or $nc^{n-1} + (n-1)a_1c^{n-2} + \cdots + a_{n-1} = 0$. In the latter case, P is not the minimal polynomial of c over C (note that this is true because we are in characteristic 0: in characteristic $p > 0$, it may be possible that the second factor is 0). Hence we must have $d(c) = 0$. □

Corollary 6.5. *If $d \not\equiv 0$ on k, then k is infinite-dimensional over C.*

Proof. If k was algebraic over C then, by the previous lemma, we would have $d(a) = 0$ for all $a \in k$. Hence, if $k \neq C$, then k contains at least one element which is transcendental over C. □

In what follows, we will frequently make use of field extensions. Luckily, derivations behave nicely under field extensions.

Lemma 6.6. *Let (k, d) be a differential field and $k(t)$ be a transcendental extension of k. Then, given $a \in k(t)$, there is a unique extension, d', of d to $k(t)$ such that $d'(t) = a$.*

Proof. Once a choice has been made for $d'(t)$, the value of d' for every other element of $k(t)$ is uniquely determined by the additive property and the Leibniz rule. □

Example 6.7. Consider the complex numbers as a (trivial) differential field $(\mathbb{C}, \frac{d}{dt})$. When we extend to the field, $\mathbb{C}(t)$, of rational functions in t, the derivation can be extended so that it maps t to any rational function $a \in \mathbb{C}(t)$. This is just extending the derivation $\frac{d}{dt}$ to the derivation $a\frac{d}{dt}$ for some $a \in \mathbb{C}(t)$.

In the case of finite field extensions, there is no longer any choice for the new derivation.

Lemma 6.8. *Let (k, d) be a differential field and $k(a)$ be a finite extension of k. Then there is a unique extension, d', of d to $k(a)$.*

Proof. The proof is almost identical to Lemma 6.4. Since a is algebraic over k, it satisfies some minimal polynomial:

$$0 = a^n + c_1 a^{n-1} + \cdots + c_n, \qquad c_i \in k.$$

Applying d' to both sides of this we have

$$0 = na^{n-1}d'(a) + d(c_1)a^{n-1} + (n-1)c_1 a^{n-2}d'(a) + \cdots + d(c_n)$$

which implies that

$$d'(a)\big(na^{n-1} + (n-1)c_1 a^{n-2} + \cdots + c_{n-1}\big) = d(c_1)a^{n-1} + \cdots + d(c_n),$$

and thus

$$d'(a) = \frac{d(c_1)a^{n-1} + \cdots + d(c_n)}{na^{n-1} + (n-1)c_1 a^{n-2} + \cdots + c_{n-1}}.$$

Hence there is only one possible value for $d'(a)$ and all other values of d' are determined by the additive property and the Leibniz rule. □

Example 6.9. Consider the differential field $(\mathcal{K}, t\frac{d}{dt})$. We have seen that the finite extensions of $\mathcal{K}$ are all of the form $\mathcal{K}_b = \mathbb{C}(\!(t^{1/b})\!)$ for some positive integer b. If we let $s = t^{1/b}$ then the derivation $t\frac{d}{dt}$ extends uniquely to the derivation

$$t\frac{d}{dt} = s^b \frac{ds}{dt} \cdot \frac{d}{ds} = \frac{1}{b} s \frac{d}{ds}.$$

6.2. **Differential Operators.** We now introduce the objects that will play the role of linear operators in the differential setting.

Definition 6.10 (Differential Operator)**.** *Let V be a vector space defined over a differential field, (k, d). A* differential operator *is an additive map $D : V \to V$, satisfying*

$$D(av) = aD(v) + d(a)v$$

for all $a \in k, v \in V$.

Note that in the case d is trivial, the above definition reduces to that of a k-linear operator. In this sense, differential operators are a generalization of linear operators. We call the pair (V, D) a *differential module* over k.

Example 6.11. Let (k, d) be a differential field with field of constants k_0. Let V_0 be an n-dimensional k_0-vector space and $V := k \otimes_{k_0} V_0$. Given a k-linear operator $T : V \to V$, the operator $D = T + d$ is a differential operator. Here $d : V \to V$ is defined by $d(a \otimes v_0) = d(a) \otimes v_0$ for $v_0 \in V_0$ and $a \in k$ (this is well-defined since the tensor product is over k_0). Since T is additive, D is also additive. For $a \in k, v \in V$ we have

$$D(av) = T(av) + d(av) = aT(v) + ad(v) + d(a)v = aD(v) + d(a)v.$$

As in the linear setting, a choice of basis for V allows us to represent a differential operator by a matrix. Let $\{e_1, \ldots, e_n\}$ be a basis for V. The action of a differential operator, $D : V \to V$, on some $y = y_1 e_1 + \cdots + y_n e_n \in V$ is:

$$D(y_i e_i) = y_i D(e_i) + d(y_i)e_i.$$

This can be put into matrix form:

$$Dy = Ay + d(y). \tag{6.12}$$

We will generally write $D = A + d$ for the differential operator.

Now suppose $\{e'_1, \ldots, e'_n\}$ is a second basis of V and let U be the change of basis matrix, i.e.

$$\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = U \begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix}.$$

In this new basis, the action of D on $v \in V$ is given by:

$$U^{-1}DU(v) = U^{-1}\big(AUv+d(Uv)\big) = U^{-1}\big(AUv+d(U)v+Ud(v)\big) = U^{-1}AU(v)+U^{-1}d(U)v+d(v),$$

where $d(U)$ is the matrix obtained by applying d element-wise to U. Thus, the change of basis gives a new matrix of action $A \mapsto U^{-1}AU+U^{-1}d(U)$. Later we will think of this transformation as an action of the group $\mathrm{GL}(V)$ on the Lie algebra $\mathfrak{gl}(V)$. In this context, the group action is usually called gauge action and one refers to the transformation $A \mapsto U^{-1}AU+U^{-1}d(U)$ as a gauge transformation.

6.3. **Differential Polynomials.** The relationship between linear operators and modules over the polynomial ring $k[x]$ is an extremely useful tool in the linear setting. Fortunately, there exists a similar relationship in the differential setting, however the polynomial ring is no longer commutative.

Notation 6.12. Given a differential field (k, d), we will denote by $k\{x\}$ (or by $k\{x, d\}$ if we wish to emphasise the derivation) the ring of differential polynomials.

As an abelian group, $k\{x\} = k[x]$, however, multiplication in $k\{x\}$ is subject to the relation

$$xa = ax + d(a), \quad \forall a \in k.$$

For non-trivial d, this means that $k\{x\}$ is non-commutative. The ring $k\{x\}$ is a special example of the non-commutative rings studied by Ore[4] [Ore33]. The following lemma shows that $k\{x\}$ retains some of the nice properties of the usual (commutative) polynomial ring.

Lemma 6.13. *The ring $k\{x\}$ is both a left and right Euclidean domain.*

The proof of this is a simple adaptation of the usual proof for commutative polynomials. An easy way to exchange left and right multiplication is given by the involution sending x to $-x$, i.e.

$$\sum_i a_i x^i \mapsto \sum_i (-x)^i a_i, \quad a_i \in k.$$

This involution will become important when we discuss differential modules. The new polynomial is called the formal adjoint[5] of the original.

As usual, Lemma 6.13 implies that $k\{x\}$ is both a left and right principal ideal domain. One can also construct a left and right least common multiple and greatest common divisor (see [Ore33, §3] — note that Ore uses the terminology union and cross-cut respectively). Due to the non-commutativity

[4] The generalization studied by Ore allows one to treat differential operators and difference operators simultaneously.

[5] See [Ked10a, §5.5.4] and [vdPS03, §2.1] for more detailed discussions of this involution.

of $k\{x\}$, it is not obvious how one can do this (simply multiplying two polynomials together will not work). The following example illustrates this difficulty for the differential field $(\mathcal{K}, t\frac{d}{dt})$.

Example 6.14. Note that for $x+t$, $x-t$, their product $(x+t)(x-t) = x^2 - t - t^2$ is *not* a right common multiple because it is not right divisible by $x+t$. Instead, following [Ore33], we obtain the "union"

$$\begin{aligned} 2t(x+t)\frac{1}{2}t^{-1}(x-t) &= t(xt^{-1}+1)(x-t) \\ &= t(t^{-1}x - t^{-1} + 1)(x-t) \\ &= (x-1+t)(x-t) \\ &= x^2 - x - t^2. \end{aligned}$$

Right division by $x+t$ shows that $(x-1+t)(x-t) = (x-1-t)(x+t)$ is divisible on the right by both $x-t$ and $x+t$.

Remark 6.15. In analogy to the case of linear operators, a differential module (V, D) is equivalent to a finite-dimensional left $k\{x\}$-module. That is, $k\{x\}$ acts on V by letting x act as D. The cyclic vector theorem guarantees the existence of a vector $v \in V$ such that $v, D(v), \ldots, D^{n-1}(v)$ form a basis of V. This implies that $D^n(v) + a_1 D^{n-1}(v) + \cdots + a_0 v = 0$ for some $a_i \in k$. In this basis, we can write $D = d + A$ where A is the companion matrix of the differential polynomial $P = x^n + a_1 x^{n-1} + \cdots + a_0 \in k\{x\}$. This gives us an isomorphism $V \cong k\{x\}/Pk\{x\}$.

6.4. **Semisimple Differential Operators.**

We can now introduce the notions of simplicity and semisimplicity for differential operators, as defined in [Lev75] (c.f. Definitions 3.5 and 3.6).

Definition 6.16 (Simple Differential Operator)**.** *Let V be a k-vector space and let $D : V \to V$ be a differential operator. Then D is called* simple *if the following equivalent conditions hold:*

(i) The corresponding $k\{x\}$-module is simple (i.e. it contains no non-zero proper submodules);

(ii) V contains no non-trivial D-invariant subspaces.

Definition 6.17 (Semisimple Differential Operator)**.** *Let V be a vector space over k and let $D : V \to V$ be a differential operator. Then D is called* semisimple *if the following equivalent conditions hold:*

(i) The corresponding $k\{x\}$-module is a direct sum of simple modules;

(ii) Every D-invariant subspace of V has a D-invariant complement.

There is also a notion of a diagonalizable differential operator:

Definition 6.18 (Diagonalizable Differential Operator)**.** *Let $D : V \to V$ be a differential operator and write $D = d + A$. Then D is* diagonalizable *over k if the following equivalent conditions hold:*

(i) A is gauge equivalent to a diagonal matrix;

(ii) V is a direct sum of one-dimensional $k\{x\}$-modules.

We now specialise to the differential field $\mathcal{K} := \mathbb{C}((t))$ with $d = t\frac{d}{dt}$ and its finite extensions $\mathcal{K}_b := \mathbb{C}((t^{1/b}))$. We have seen that both the t-adic valuation and the derivation d extend uniquely to $\mathcal{K}_b$.

Given a $\mathcal{K}$-vector space V and a finite extension $\mathcal{K} \subset \mathcal{K}_b$ we can extend scalars to form the vector space $V_b := \mathcal{K}_b \otimes_{\mathcal{K}} V$. A differential operator $D : V \to V$ then extends to $D_b : V_b \to V_b$ via:

$$D_b : a \otimes v \mapsto \bar{d}(a) \otimes v + a \otimes Dv,$$

where $\bar{d}$ is the unique extension of d to $\mathcal{K}_b$.

In this setting, semisimplicity behaves well with respect to taking extensions:

Lemma 6.19 (Levelt)**.** *Let V be a vector space over K and $D : V \to V$ a differential operator. Then D is semisimple if, and only if, $D_{K_m} : V_{K_m} \to V_{K_m}$ is semisimple for any finite field extension K_m of K.*

Proof. See [Lev75, §1(e)]. □

There is also a nice relationship between semisimple operators and diagonalizability. A one-dimensional $\mathcal{K}\{x\}$-module is simple; thus diagonalizable differential operators are semisimple. In Section 8.1 we give a new proof of a converse to this (after a possible finite extension of $\mathcal{K}$). Thus a differential operator is semisimple if and only if, after an appropriate finite extension, it is diagonalisable. Note that this is an analogue of Lemma 3.8 in the linear setting.

6.4.1. *Nilpotent Differential Operators.* We make one brief remark about the solutions of systems of linear differential equations. Given a system $d(v) = Av$ of differential equations, there will be, in general, a vector space of solutions:

$$V = \{v \in k^n \,|d(v) = A(v)\}.$$

This solution space is the kernel of the differential operator $D = A - d$. The following lemma says that this vector space is relatively small:

Lemma 6.20. *Let k be a differential field of characteristic 0 and C its field of constants. Let $d(v) = Av$ be a system of differential equations with $v \in k^n, A \in M_n(k)$ and let V be the solution space of the system. Then V is a finite-dimensional vector space over C with dimension at most n.*

Proof. The proof follows Singer and van der Put [vdPS03, §1.2]. It is easy to verify that V is a C-vector space. We will show that linear dependence over k implies linear dependence over C. Since any $n+1$ vectors in $V \subset k^n$ must be linearly dependent over k, this will imply that V has dimension at most n over C.

Suppose that $v_1, \ldots, v_m \in V$ are linearly dependent over k but that any proper subset is linearly independent. Then we can write

$$v_1 = \sum_{i=2}^{m} a_i v_i, \qquad a_i \in k.$$

Since $v_i \in V$ we have:

$$\begin{aligned} 0 = v_1' - Av_1 &= \left(\sum_{i=2}^m a_i v_i\right)' - A\left(\sum_{i=2}^m a_i v_i\right) \\ &= \sum_{i=2}^m a_i' v_i + \sum_{i=2}^m a_i v_i' - \sum_{i=2}^m a_i A v_i \\ &= \sum_{i=2}^m a_i' v_i + \underbrace{\sum_{i=2}^m a_i (v_i' - Av_i)}_{=0} \\ &= \sum_{i=2}^m a_i' v_i. \end{aligned}$$

Since $v_2, \ldots, v_m$ are linearly independent over k we must have $a_i' = 0$ for all $i = 2, \ldots, m$. Hence $a_i \in C$ and so $v_1, \ldots, v_m$ are linearly dependent over C. □

One implication of Lemma 6.20 is that there are no non-trivial nilpotent differential operators:

Lemma 6.21. *Let V be a vector space over the differential field (k, d) and $D : V \to V$ be a differential operator. If there exists an integer, n, such that:*

$$D^n v = 0$$

for all $v \in V$, then $d = 0$, and D is a nilpotent k-linear operator.

Proof. Suppose $D^n v = 0$ for some $n \in \mathbb{Z}$ and $d \neq 0$. Then we have the system:

$$(A + d)^n v = 0.$$

Expanding the left hand side will yield a system of m differential equations of order n where $m = \dim(A)$. This is equivalent to a system of mn first-order equations by creating new variables for $d(v), d^2(v), \ldots, d^n(v)$. Thus the system is equivalent to a matrix differential equation:

$$\boldsymbol{y}' = B\boldsymbol{y},$$

where $\boldsymbol{y} \in k^{nm}, \dim(B) = mn$. By Lemma 6.20, the solution space of this system is finite-dimensional over the field of constants. Since k itself is infinite-dimensional over its field of constants (see Corollary 6.5) it certainly can not be the case that $D^n v = 0$ for all $v \in k^n$. Hence d must be trivial. □

6.5. **The Category of Differential Modules.** In this section we denote by $\mathrm{Diff}_{\mathcal{K}}$ the category of finite-dimensional differential modules over $\mathcal{K} := \mathbb{C}((t))$. The objects of this category are differential modules (i.e. pairs (V, D), with V a finite-dimensional $\mathcal{K}$-vector space and $D : V \to V$ a differential operator) and the morphisms are the $\mathcal{K}$-linear maps that commute with the differential operators. If (V_1, D_1), (V_2, D_2) are two objects of $\mathrm{Diff}_{\mathcal{K}}$, then we will usually write

$$\mathrm{Hom}_{\mathcal{K}\{x\}}(V_1, V_2) := \{f \in \mathrm{Hom}_{\mathcal{K}}(V_1, V_2) : f \circ D_1 = D_2 \circ f\}$$

for the set of morphisms from (V_1, D_1) to (V_2, D_2). Note that $\mathrm{Hom}_{\mathcal{K}\{x\}}(V_1, V_2)$ is an abelian group but not a differential module.

Most of the usual constructions from linear algebra (e.g. duals, tensor products, direct sums, etc.) have analogues in $\mathrm{Diff}_{\mathcal{K}}$. We will now define these formally and give some results that we will need in later sections.

Definition 6.22 (Direct Sum). *Let* (V_1, D_1), (V_2, D_2) *be differential modules over* $\mathcal{K}$. *The* direct sum $V_1 \oplus V_2$ *is a differential module with differential operator* $D := D_1 \oplus D_2$ *given by*

$$D(v_1 \oplus v_2) := D_1(v_1) \oplus D_2(v_2).$$

With the direct sum as both the (categorical) product and coproduct, $\mathrm{Diff}_{\mathcal{K}}$ becomes an abelian category.

Definition 6.23 (Tensor Product of Modules). *Let* (V_1, D_1) *and* (V_2, D_2) *be differential modules over* $\mathcal{K}$. *The* tensor product $V_1 \otimes_{\mathcal{K}} V_2$ *is a differential module with differential operator* $D := D_1 \otimes D_2$ *given by:*

$$(D_1 \otimes_{\mathcal{K}} D_2)(v_1 \otimes_{\mathcal{K}} v_2) := D_1(v_1) \otimes_{\mathcal{K}} v_2 + v_1 \otimes_{\mathcal{K}} D_2(v_2).$$

We will generally write $\otimes$ for $\otimes_{\mathcal{K}}$ where there is no confusion. It is straight-forward to verify that $D_1 \otimes D_2$ is $\mathbb{C}$-linear and

$$\begin{aligned}(D_1 \otimes D_2)(av_1 \otimes v_2) = D_1(av_1) \otimes v_2 + av_1 \otimes D_2(v_2) &= aD_1(v_1) \otimes v_2 + d(a)v_1 \otimes v_2 + av_1 \otimes D_2(v_2)\\ &= a(D_1 \otimes D_2)(v_1 \otimes v_2) + d(a)v_1 \otimes v_2,\end{aligned}$$

shows that the Leibniz rule is satisfied for all $a \in \mathcal{K}$.

We have already noted that $\mathrm{Hom}_{\mathcal{K}\{x\}}(V_1, V_2)$ is not a differential module. The set of $\mathcal{K}$-linear maps from V_1 to V_2 is, however, a differential module.

Definition 6.24 (Internal Homs). *Let* (V_1, D_1) *and* (V_2, D_2) *be differential modules over* $\mathcal{K}$. *The set* $\mathrm{Hom}_{\mathcal{K}}(V_1, V_2)$ *is a differential module with differential operator* D, *given by:*

$$D(f) := D_2 \circ f - f \circ D_1.$$

Again, we verify the Leibniz rule. For $a \in \mathcal{K}$ we have

$$D(af) = D_2 \circ (af) - af \circ D_1 = aD_2 \circ f + d(a)f - af \circ D_1 = a(D_2 \circ f - f \circ D_1) + d(a)f.$$

In the category of R-modules, the two functors $- \otimes X$ and $\mathrm{Hom}(X, -)$ are *adjoint functors* (see [Rot09, Theorem 2.75]). There is an analogue of this result in $\mathrm{Diff}_{\mathcal{K}}$:

Proposition 6.25. *Let* (T, D_T), (X, D_X), *and* (Y, D_Y) *be differential modules over* $\mathcal{K}$. *We have the following isomorphism of abelian groups:*

$$\mathrm{Hom}_{\mathcal{K}\{x\}}(T \otimes X, Y) \cong \mathrm{Hom}_{\mathcal{K}\{x\}}(T, \mathrm{Hom}_{\mathcal{K}}(X, Y)).$$

Proof. The construction of the isomorphism here is motivated by the corresponding construction for the category of R-modules. Define

$$\tau : \mathrm{Hom}_{\mathcal{K}\{x\}}(T \otimes X, Y) \to \mathrm{Hom}_{\mathcal{K}\{x\}}(T, \mathrm{Hom}_{\mathcal{K}}(X, Y))$$

by $f \mapsto \tau(f)$ where $\tau(f) : T \to \mathrm{Hom}_{\mathcal{K}}(X, Y)$ is defined by

$$\tau(f)_t : x \mapsto f(t \otimes x).$$

It is easy to check that τ is an isomorphism of abelian groups. Note that the map $\tau(f)_t$ is $\mathcal{K}$-linear but in general does not commute with the differential operators; i.e.

$$\tau(f)_t(D_X(x)) = f(t \otimes D_X(x)),$$

but

$$\begin{aligned} D_Y(\tau(f)_t(x)) &= D_Y(f(t \otimes x)) \\ &= f(D_{T\otimes X}(t \otimes x)) \\ &= f(D_T(t) \otimes x) + f(t \otimes D_X(x)). \end{aligned}$$

□

Remark 6.26. Proposition 6.25 shows that the functor $\mathcal{F} : \mathrm{Diff}_{\mathcal{K}} \to \mathrm{Set}$ given by

$$\mathcal{F}(T) = \mathrm{Hom}_{\mathcal{K}\{x\}}(T \otimes X, Y), \quad X, Y \in \mathrm{obj}(\mathrm{Diff}_{\mathcal{K}}),$$

is representable and represented by $\mathrm{Hom}_{\mathcal{K}}(X, Y)$. Thus, in the language of Tannakian categories (see [DM82, §1]) the differential modules $\mathrm{Hom}_{\mathcal{K}}(X, Y)$ are the internal homs.

Definition 6.27 (Dual Module)**.** *Let* $V^* := \mathrm{Hom}_{\mathcal{K}}(V, \mathcal{K})$ *denote the vector space dual to* V*. We can define a differential operator* $D^* : V^* \to V^*$ *by*

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

The differential module (V^*, D^*) *is called the* dual *of* (V, D).

Note that the dual differential module is really just $\mathrm{Hom}_{\mathcal{K}}(V, \mathbb{1})$ where $\mathbb{1}$ is the trivial differential module $(\mathcal{K}, d)$.

Remark 6.28. It is clear that $\mathbb{1}$ is a unit object for the tensor product defined above. Note that any $\mathcal{K}$-linear map $f : \mathcal{K} \to \mathcal{K}$ is completely determined by $f(1)$ and thus the only such maps are the "multiplication by a" maps for some $a \in \mathcal{K}$. In order for f to be a $\mathcal{K}\{x\}$-map, we also require $d(a) = 0$, i.e. we need $a \in \mathbb{C}$. Thus $\mathrm{Hom}_{\mathcal{K}\{x\}}(\mathbb{1}, \mathbb{1}) \cong \mathbb{C}$. These facts, together with the constructions defined in this section are almost enough to show that $\mathrm{Diff}_{\mathcal{K}}$ is a neutral Tannakian category. We only require the existence of a fibre functor $\omega : \mathrm{Diff}_{\mathcal{K}} \to \mathrm{Vect}_{\mathbb{C}}$. Interestingly, the Hukuhara-Levelt-Turritin decomposition implies the existence of such a fibre functor (see [Kat87, §2.4]) and hence $\mathrm{Diff}_{\mathcal{K}}$ is a neutral Tannakian category.

7. Factorisation of Differential Polynomials

In this section we prove a differential analogue of the Newton-Puiseux theorem (see Section 5.2). Most of the proofs given in this section have the same structure as those given in Section 5. As before, we fix $\mathcal{K} := \mathbb{C}((t))$ and $\mathcal{O} := \mathbb{C}[[t]]$.

7.1. **Newton Polygon of a Differential Polynomial.**

We now introduce an object which will be an extremely useful tool for factorising differential polynomials: the Newton polygon. There is an analogous object in the linear setting which we introduce first.

Definition 7.1 (The Newton Polygon of a Commutative Polynomial). *Let $f \in \mathcal{K}[x]$ and write*

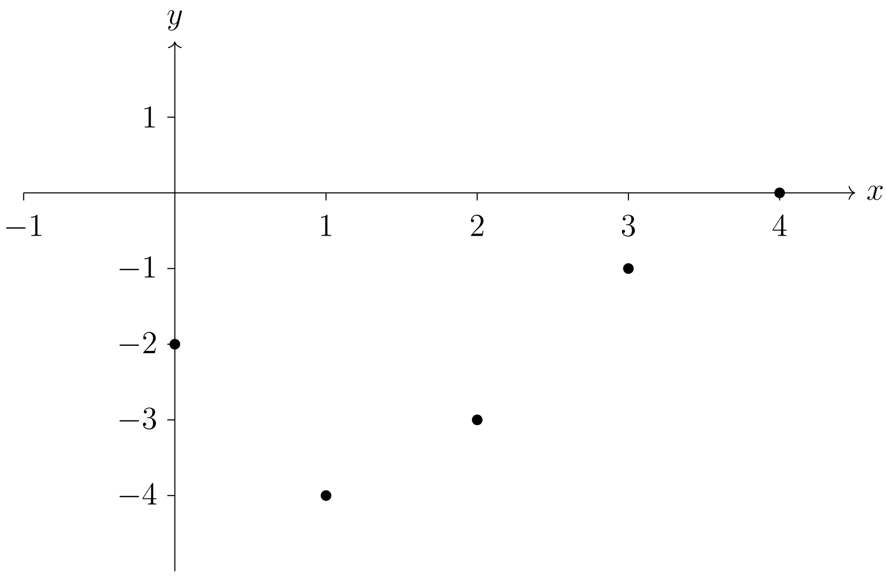
$$f(x) = \sum_{i=0}^{n} a_i x^{n-i}, \qquad a_i \in \mathcal{K}.$$

The Newton polygon *of f, denoted $\mathrm{NP}(f)$, is the lower boundary of the convex hull of the points*

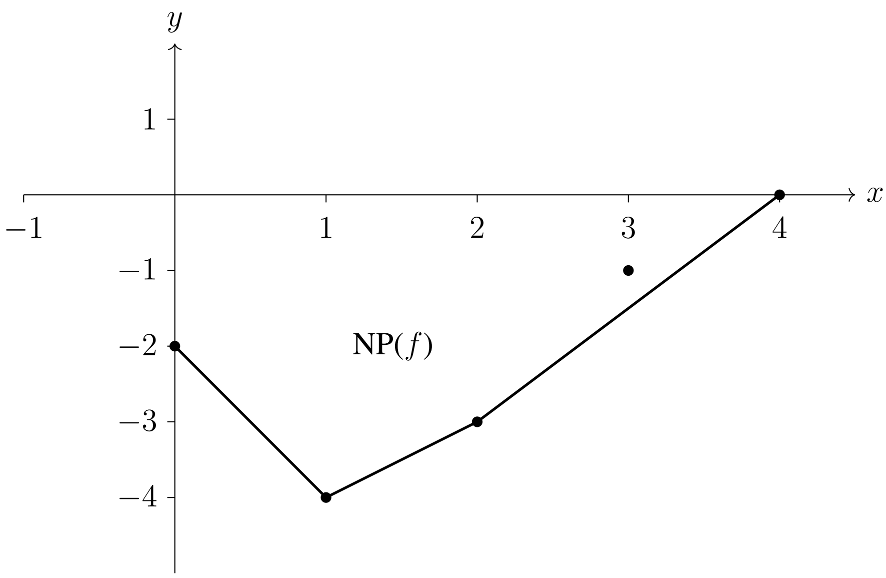
$$\{(n-i), v_t(a_i) : 0 \le i \le n\} \subset \mathbb{R}^2.$$

The easiest way to understand the above definition is by an example.

Example 7.2. To construct the Newton polygon for $f = x^4 + t^{-1}x^3 - (t^{-3} + 1 + t^2)x^2 + t^{-4}x - t^{-2}$, first plot the points whose x-coordinate is given by the power of x and whose y-coordinate is given by the valuation of the corresponding coefficient:



The Newton polygon is then the lower boundary of the convex hull of these points:



Newton originally introduced these polygons in order to factorise polynomials in the commutative ring $\mathcal{K}[x]$ (see [BK86, §8.3] for an excellent account of the applications and history of Newton polygons in this setting).

In the differential setting, one must modify the above construction so that the polygon is gauge invariant. The required modification will depend on the derivation. We follow the definition in [Ked10b]. Firstly, recall the operator norm:

Definition 7.3 (Operator Norm)**.** *Let T be a bounded linear operator on a normed vector space V. The* operator norm*, $|T|$, of T is*

$$|T| := \sup_{v \in V} \left\{ \frac{|T(v)|}{|v|} \right\}.$$

In our situation, we will take T to be a derivation, d, on $V = \mathcal{K}$ viewed as a vector space over $\mathbb{C}$. The norm on $\mathcal{K}$ is the norm coming from the t-adic valuation: $|\cdot| = \exp(-v_t(\cdot))$. In order to define a gauge invariant Newton polygon we introduce the quantity

$$r_0 := \log(|d|).$$

Our main examples will be the following.

Example 7.4. Let $v = \sum_{i \geq n}^{\infty} a_i t^i \in \mathcal{K}$ and $\delta_m = t^m \frac{d}{dt}$. We have $\delta_m(v) = \sum_{i \geq n}^{\infty} a_i i t^{i+m-1}$ and so

$$|\delta_m| = \sup_{v \in \mathcal{K}} \left\{ \frac{|\delta_m(v)|}{|v|} \right\} = \sup_{v \in \mathcal{K}} \left\{ \frac{\exp(1-n-m)}{\exp(-n)} \right\} = \exp(1-m).$$

In this case we have $r_0 = 1 - m$.

Example 7.5. We can generalise the above calculation to an arbitrary derivation $d = a\frac{d}{dt}$, for some $a \in \mathcal{K}$. Again, set $v = \sum_{i \geq n}^{\infty} a_i t^i \in \mathcal{K}$. Then:

$$|d| = \sup_{v \in \mathcal{K}} \left\{ \frac{|d(v)|}{|v|} \right\} = \sup_{v \in \mathcal{K}} \left\{ \frac{\exp\left(-v_t(\sum_{i \geq n}^{\infty} a i a_i t^{i-1})\right)}{\exp(-n)} \right\} = \sup_{v \in \mathcal{K}} \left\{ \frac{\exp(v_t(a) - n + 1)}{\exp(-n)} \right\} = \exp(1 - v_t(a)),$$

and so $r_0 = 1 - v_t(a)$.

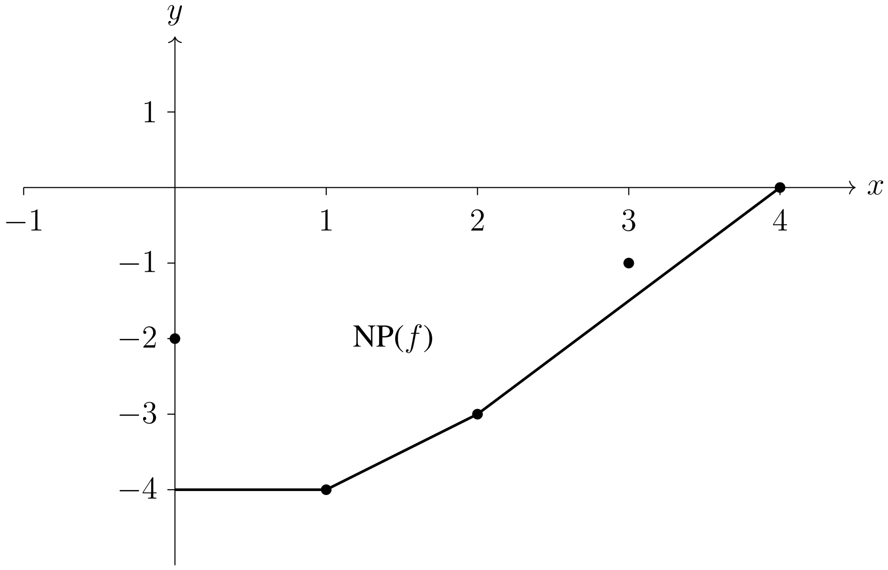
Now, given a polynomial $f \in \mathcal{K}\{x, d\}$, one first constructs the corresponding commutative Newton polygon and then replaces all slopes less than r_0 with a single slope of exactly r_0. To make this more precise, define a partial order on $\mathbb{R}^2$ by $(x_1, y_1) \geq (x_2, y_2)$ if $x_1 \leq x_2$ and $y_1 \geq y_2$. Using this partial order we have the following definition:

Definition 7.6. *Let $f \in \mathcal{K}\{x, d\}$ and set $r_0 = \log(|d|)$, as above. The* Newton polygon *of f is the lower boundary of the convex hull of the set*

$$\{(a, b) \in \mathbb{R}^2 : f \textit{ has a monomial } t^m x^n \textit{ with } (a, b) \geq (n, m + (x - n) r_0)\}.$$

Note that the condition on the y-coordinate in Definition 7.6 defines a line above which the points (a, b) must lie. Again, an example is the most efficient way to understand this definition.

Example 7.7. As in Example 7.2, we take $f = x^4 + t^{-1}x^3 - (t^{-3} + 1 + t^2)x^2 + t^{-4}x - t^{-2}$ but now view f as a polynomial in $\mathcal{K}\{x, t\frac{d}{dt}\}$. In this case $r_0 = 0$ and so we must replace any negative slopes appearing in the commutative Newton polygon with a single slope of 0. This yields the following differential Newton polgyon:



7.2. **Irregularity.** The most important characteristic of the differential Newton polygon is its set of slopes. In fact, one can reformulate most of the following factorisation results in terms of the slopes of a differential polynomials Newton polygon. In this section, we give a characterisation of regular singularity using the slopes of the differential Newton polygon. This notion will coincide with the one introduced in Section 2.2. We first define the irregularity of a differential polynomial.

Definition 7.8 (Irregularity)**.** *Let $f \in \mathcal{K}\{x, d\}$ be a differential polynomial and let $s_1, \ldots, s_n$ be the slopes of its Newton polygon. The* irregularity *of f is defined as*

$$\operatorname{Irr}(f) := \sum_{i=1}^{n}(s_i + r_0).$$

We now make the following obvious definition.

Definition 7.9 (Regular Singular Differential Polynomial)**.** *Let $f \in \mathcal{K}\{x, d\}$ be a differential polynomial. We say that f is* regular singular *if* $\operatorname{Irr}(f) = 0$*. Equivalently, f is regular singular if its Newton polygon has a single slope of* r_0*.*

For the derivation δ_1, the regular singular differential polynomials are precisely those in $\mathcal{O}\{x, \delta_1\}$ (since, for the irregularity to be 0, the Newton polygon must have a single slope of 0). Note also that in this case, the irregularity is exactly the sum of the slopes of the corresponding Newton polygon. We now show that Definition 7.9 coincides with the one given in Section 2.2.

Proposition 7.10. *Let (V, D) be a differential module and suppose that $V \cong \mathcal{K}\{x\}/f\mathcal{K}\{x\}$ for some differential polynomial $f \in \mathcal{K}\{x, \delta_1\}$. Then V is a regular singular differential module if and only if* $\operatorname{Irr}(f) = 0$*.*

Proof. By Remark 6.15, the isomorphism $V \cong \mathcal{K}\{x\}/f\mathcal{K}\{x\}$ allows us to write $D = d + A$ where A is the companion matrix of f. It is now clear that the coefficients of f are in $\mathcal{O}$ if and only if the entries of A are in $\mathcal{O}$. Thus both definitions of regular singular are equivalent. $\square$

7.3. **Hensel's Lemma for Differential Polynomials.** In this section we restrict our attention to the differential field $\mathcal{K} = \mathbb{C}(\!(t)\!)$ and its finite extensions, $\mathcal{K}_b = \mathbb{C}(\!(t^{1/b})\!)$, b a positive integer (see Section 5). Our derivations will be of the form $\delta_m := t^m \frac{d}{dt}$. Since we will need to emphasise the derivation, we denote by $\mathcal{K}\{x, \delta_m\}$ the ring of differential polynomials subject to the relation

$$xa = ax + \delta_m(a), \qquad \forall a \in \mathcal{K}. \tag{7.13}$$

As in the linear setting, Hensel's lemma will allow us to factorize differential polynomials with coefficients in $\mathcal{O} = \mathbb{C}[\![t]\!]$. In the special case of $\mathcal{O}\{x, \delta_1\}$ we find a surprising difference between the differential and linear settings: every polynomial in $\mathcal{O}\{x, \delta_1\}$ has a linear factorization (i.e. we do not require a field extension in order to obtain a linear factorization). Unfortunately, this result does not hold for the other δ_m.

Since $\mathcal{O}\{x, \delta_m\}$ is a non-commutative ring, we must be careful when reducing coefficients mod t. For $f \in \mathcal{O}\{x, \delta_m\}$, we will 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. Formally, what we mean by $f \pmod{t^n}$ is the image of f in the quotient $\mathcal{O}\{x, \delta_m\}/t^n\mathcal{O}\{x, \delta_m\}$. There is no particular significance in the choice of left: one can instead move all factors of t to the right and consider the image of f in the quotient $\mathcal{O}\{x, \delta_m\}/\mathcal{O}\{x, \delta_m\}t^n$. For notational convenience, let $\bar{f} := f \pmod{t}$. Note that $\bar{f} \in \mathbb{C}[x]$. In order to shift factors of t to the left, we will frequently make use of the following identities.

Lemma 7.11. *For all $h(x) \in \mathcal{K}\{x, \delta_m\}$ and $i \in \mathbb{Z}$*

$$h(x)t^i = t^i h(x + it^{m-1}). \tag{7.14}$$

Proof. Firstly, taking $a = t^i$ in (7.13) yields $xt^i = t^i(x + it^{m-1})$. An easy induction argument then yields the identity

$$x^n t^i = t^i (x + it^{m-1})^n, \qquad \forall m \in \mathbb{N}.$$

The result now follows by considering each term in $h(x)$. ∎

Remark 7.12. In order to factor all polynomials in $\mathcal{K}\{x\}$ we will need to allow field extensions to $\mathbb{C}(\!(s)\!)$ where $s^q = t, q \in \mathbb{N}$. Under such a field extension, the derivation δ_m extends uniquely to the derivation

$$\delta'_m = t^n \frac{d}{dt} = s^{nq} \frac{ds}{dt} \cdot \frac{d}{ds} = \frac{1}{q} s^{nq-q+1} \frac{d}{ds}.$$

Lemma 7.11 now applies to $\mathbb{C}(\!(s)\!)\{x, \delta'_m\}$ by replacing t with s. This will allow us to deal with fractional powers of t when they arise.

Lemma 7.13. *For the derivation δ_m*

$$(t^d x)^n = \sum_{j=0}^{n-1} a_j t^{nd+(m-1)j} x^{n-j}, \qquad \forall d \in \mathbb{Z} - \{0\}, n \in \mathbb{N} \tag{7.15}$$

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

Proof. The proof uses induction on n. The case $n = 1$ is clear. We have

$$\begin{aligned}
(t^d x)^{n+1} &= (t^d x) \sum_{j=0}^{n-1} a_j t^{nd+(m-1)j} x^{n-j} = \sum_{j=0}^{n-1} a_j t^d (x t^{nd+(m-1)j}) x^{n-j} \\
&= \sum_{j=0}^{n-1} a_j t^d \Big(t^{nd+(m-1)j} x + \big(nd + (m-1)j\big) t^{nd+(m-1)(j+1)} \Big) x^{n-j} \\
&= \sum_{j=0}^{n-1} a_j t^{(n+1)d+(m-1)j} x^{n+1-j} + a_j \big(nd + (m-1)j\big) t^{(n+1)d+(m-1)(j+1)} x^{n-j} \\
&= a_0 t^{(n+1)d} x^{n+1} + \sum_{j=1}^{n} \Big(a_j + a_{j-1} \big(nd + (m-1)(j-1)\big)\Big) t^{(n+1)d+(m-1)j} x^{n+1-j} \\
&= \sum_{j=0}^{n+1} b_j t^{(n+1)d+(m-1)j} x^{n+1-j}, \qquad b_j \in \mathbb{Z},
\end{aligned}$$

where $b_0 = a_0 = 1$. This proves the identity. □

Now suppose we have a factorization of $\bar{f}$:

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

As in the linear setting, our goal is to lift this to a factorization of f in $\mathcal{K}\{x, \delta_m\}$. The following result should be thought of as a differential analogue of Hensel's lemma.

Proposition 7.14. *Let $f \in \mathcal{O}\{x, \delta_m\}$ and $\bar{f} = g_0 h_0$ as above. Suppose $\bar{f} \not\equiv 0$ and*

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

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

A version of the above result appears in [Pra83, Lemma 1]. The proof has the outline of the proof of the usual Hensel's lemma.

Proof. Our goal is to inductively build a sequence of functions

$$g_n(x) = g_0 + t p_1 + t^2 p_2 + \cdots + t^{n-1} p_{n-1} + t^n p_n, \qquad p_i \in \mathbb{C}[x] \tag{7.16}$$

$$h_n(x) = h_0 + t q_1 + t^2 q_2 + \cdots + t^{n-1} q_{n-1} + t^n q_n, \qquad q_i \in \mathbb{C}[x] \tag{7.17}$$

for all $n \geq 0$, which satisfy:

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

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

Suppose that we know the p_i and q_i for $1 \leq i \leq n-1$. In view of (7.16) and (7.17) we have:

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

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

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

We need to shift the powers of t to the left. By (7.14), $g_{n-1}(x)t^n = t^n g_{n-1}\left(x + nt^{m-1}\right)$, so we have:

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

and thus

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

For notational convenience, we set:

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

so that we have

$$f_n \equiv g_0(x + nt^{m-1})q_n(x) + p_n(x)h_0(x) \pmod{t}. \tag{7.18}$$

Now if $m > 1$ then (7.18) 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 (7.18) 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

$$g_0(x + n)a(x)f_n(x) + h_0(x)b(x)f_n(x) = f_n(x). \tag{7.19}$$

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 (7.19) then becomes:

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

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

In the differential setting, Hensel's lemma immediately implies the following.

Corollary 7.15. *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 mod 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 $\mathrm{Re}(\lambda_1) \leq \mathrm{Re}(\lambda_2) \leq \cdots \leq \mathrm{Re}(\lambda_n)$. 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 7.14 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. □

Remark 7.16. 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 $x^2 + t - t^2 = (x - t)(x + t)$.

We now give some examples of this algorithm.

Example 7.17. Consider the differential polynomial $f = x^2 - (1 - t + t^2)x + (t + t^2 - t^3) \in \mathcal{K}\{x, \delta_1\}$. We have $\bar{f} = x^2 - x = x(x - 1)$ so we set $g_0 = x$, $h_0 = x - 1$. We now look for p_1 and q_1 satisfying:

$$f \equiv (g_0 + p_1 t)(h_0 + q_1 t) \pmod{t^2}.$$

This gives us the following condition on p_1 and q_1:

$$\begin{aligned} x^2 - (1 - t)x + t &\equiv (x + p_1 t)(x - 1 + q_1 t) \pmod{t^2} \\ \implies tx + t &\equiv q_1(tx + t) + p_1 tx - p_1 t \pmod{t^2} \\ \implies x + 1 &\equiv (q_1 + p_1)x + (q_1 - p_1) \pmod{t}. \end{aligned}$$

Comparing coefficients, we find $p_1 = 0, q_1 = 1$ and so $g_1 = x$ and $h_1 = x - 1 + t$. We then look for p_2 and q_2 satisfying:

$$f \equiv (x + p_2 t^2)(x - 1 + t + q_2 t^2) \pmod{t^3}.$$

Repeating the calculation above, we find that $p_2 = 1, q_2 = 0$. At this point, we in fact have a full factorisation of $f = (x + t^2)(x - 1 + t)$.

Remark 7.18. Example 7.17 is somewhat contrived in that the algorithm terminated in finitely many steps. In general, this process could continue indefinitely (as in the next example) but it will always yield a factorisation.

The second example shows that this algorithm may not always give the "nicest" factorisation of a polynomial.

Example 7.19. Consider the polynomial $f = x^2 + tx - 1 \in \mathcal{K}\{x, \delta_1\}$. In this case, $\bar{f} = x^2 - 1 = (x-1)(x+1)$ and so we set $g_0 = x+1$, $h_0 = x-1$. Applying the first two iterations of the algorithm we find $g_2 = x + 1 + \frac{2}{3}t + \frac{1}{18}t^2$ and $h_2 = x - 1 + \frac{1}{3}t - \frac{1}{18}t^2$, i.e.

$$f \equiv \Big(x + 1 + \frac{2}{3}t + \frac{1}{18}t^2\Big)\Big(x - 1 + \frac{1}{3}t - \frac{1}{18}t^2\Big) \pmod{t^3}.$$

This is still not a perfect factorisation of f. In fact, expanding out we find $g_2h_2 = x^2 + tx - 1 - \frac{1}{54}t^3 - \frac{1}{324}t^4$. We can keep computing terms indefinitely to obtain better factorisations, however we would need to continue to an infinite number of terms to get the complete factorisation. The polynomial, f, does have a simple factorisation however. One can check that $f = (x-1)(x+1+t)$. The reason we did not find this factorisation is because the ordering of the roots given in Corollary 7.15 does not match the ordering of the roots in this simple factorisation.

7.4. **Change of Variables.** In order to factorise a general polynomial in the linear setting we made a change of variables of the form $x = t^r Y$. As we are trying to mimic this proof, we would also like to do this here. There is, however, an important implication of this change of variables in the differential setting: the derivation changes. To see this, recall that multiplication in the ring $\mathcal{K}\{x, d\}$ is defined by the relation

$$xa = ax + d(a), \qquad \forall a \in \mathcal{K}.$$

If we make the change of variables $x = t^r Y$ then this relation tells us that

$$t^r Y a = at^r Y + d(a) \implies Ya = aY + t^{-r}d(a).$$

Hence making this change of variable will yield a differential polynomial in the ring $\mathcal{K}\{Y, t^{-r}d\}$. Much of this section is devoted to technical lemmas which we will use to factorise differential polynomials.

Given a differential polynomial, $f = \sum a_i x^{n-i} \in \mathcal{K}\{x, \delta_1\}$, set $r := \min\left\{\frac{v(a_i)}{i}\right\} = \frac{p}{q}$, with $\gcd(p, q) = 1$. The change of variables we would like to make is $x = t^r Y$ (as in the linear setting). The following lemma shows that despite changing the derivation, this still gives us a polynomial with power series coefficients.

Lemma 7.20. *Consider the differential polynomial $f = \sum a_i x^{n-i} \in \mathcal{K}\{x, \delta_1\}$ with r as above. Let $g(Y) = s^{-np} f(s^p Y)$, where $s = t^{1/q}$. Then, $g(Y) \in \mathbb{C}[\![s]\!]\{Y, \frac{1}{q}s^{1-p}\frac{d}{ds}\}$.*

Proof. Applying (7.15) to $f(s^p Y)$ yields $f(s^p Y) = s^{np} g(Y)$ where

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

Note that $v_s(a_i) = qv_t(a_i)$, $v_t(a_i) \geq \frac{ip}{q}$ implies that $v_s(a_i) \geq ip$. Thus, for $0 \leq l \leq n-1$, the coefficient, b_l, of Y^{n-l} in g satisfies

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

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

$$g(Y) = Y^n + b_1 Y^{n-1} + \cdots + b_n, \qquad b_i \in \mathbb{C}[\![s]\!]$$

with $\min(v_s(b_i)) = 0$. Furthermore, $v_s(b_i) = 0$ if, and only if, $v_t(a_i) = ir$. $\square$

Lemma 7.20 provides us with a clear direction to complete our factorisation proof. The next logical step is to try to apply Hensel's Lemma to $g(Y)$. Since the derivation may have changed, there is one case in which we will not be able to apply Proposition 7.14: the case where $\bar{g}(Y) := g(Y) \pmod{s}$ has a single repeated root. We now investigate this case specifically. The first point of interest is that the Newton polygon of f is very special in this case.

Lemma 7.21. *Let $f(x)$ and $g(Y)$ be as in Lemma 7.20 and suppose that $\bar{g}(Y) = (Y + \lambda)^n$ for some $\lambda \in \mathbb{C}$. Then λ is non-zero and the Newton polygon of f has a single* integral *slope.*

Proof. As in Lemma 7.20, write

$$g = Y^n + b_1 Y^{n-1} + \cdots + b_n, \qquad b_i \in \mathbb{C}[\![s]\!].$$

Since $\min\{v_s(b_i)\} = 0$, $\lambda \neq 0$. Now since, $\lambda \neq 0$, expanding $(X + \lambda)^n$ 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 equal to r and since $v_t(a_1) = r$, r is an integer. $\square$

In order to deal with this case completely, we will make a change of variables of the form $x \mapsto x - \lambda t^r$. This has the effect of reducing the slope of the Newton polygon:

Lemma 7.22. *Let f and g be as in Lemma 7.20 and suppose that $\bar{g} = (Y + \lambda)^n$, $\lambda \in \mathbb{C}$. Then the Newton polygon of $f(x - \lambda t^r)$ has a single slope strictly smaller than the slope of the Newton polygon of $f(x)$.*

Proof. By Lemma 7.21, r is an integer and hence no extension of $\mathcal{K}$ is necessary. Since $\bar{g} = (Y + \lambda)^n$, we can write g as

$$g = (Y + \lambda)^n + e_1 (Y + \lambda)^{n-1} + \cdots + e_n, \qquad e_i \in \mathcal{O},$$

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

$$\begin{aligned} f(t^r Y) &= t^{nr}\big((Y + \lambda)^n + e_1(Y + \lambda)^{n-1} + \cdots + e_n\big) \\ \implies f(x) &= t^{nr}\big((t^{-r}x + \lambda)^n + e_1(t^{-r}x + \lambda)^{n-1} + \cdots + e_n\big) \end{aligned}$$

and hence

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

Applying (7.15), we have, for $c_{k,l} \in \mathbb{C}$,

$$\begin{aligned} f(x-\lambda t^r) &= t^{nr}\Big(t^{-nr}\sum_{j=0}^{n-1} c_{n,j}x^{n-j} + e_1 t^{-(n-1)r}\sum_{j=0}^{n-2} c_{n-1,j}x^{n-1-j} + \cdots + e_n\Big) \\ &= \sum_{j=0}^{n-1} c_{n,j}x^{n-j} + e_1 t^r \sum_{j=0}^{n-2} c_{n-1,j}x^{n-1-j} + \cdots + t^{nr}e_n. \end{aligned}$$

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 slope of the Newton polygon for $f(x-\lambda t^r)$ is strictly less than the slope of the Newton polygon for $f(x)$. $\square$

Before leaving this section we mention an interesting aside. Note that even though the change of variables in Lemma 7.20 gives us a differential polynomial with power series coefficients, it is not necessarily regular singular since, in general, the derivation is no longer δ_1. In fact, the property of being regular singular is preserved under this change of variables:

Lemma 7.23. *Let $f \in \mathcal{K}\{x, \delta_m\}$ be a regular singular polynomial. Under the change of variable $x = t^p Y$, $p \in \mathbb{Z}$, the resulting polynomial $g(Y)$ is regular singular for the derivation δ_{m-p}.*

Proof. Write

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

Suppose

$$f(x) = x^n + a_1 x^{n-1} + \cdots + a_n, \qquad a_i \in \mathcal{K}$$

is regular singular for δ_m. By definition the Newton polygon of f has single slope $1-m$. Since the right-most point $(n, v_t(a_0))$ is a vertex of the Newton polygon, the valuations of the coefficients a_i must lie above the line $y = (1-m)x + v_t(a_0) - n(1-m)$. Hence for $0 \le k \le n$ we must have $v_t(a_k) \ge v_t(a_0) - k(1-m)$.

By Lemma 7.15, making the change of variable $x = t^p Y$ yields

$$\begin{aligned} g(Y) &= a_0(t^pY)^n + a_1(t^pY)^{n-1} + \cdots + a_n \\ &= a_0 \sum_{i=0}^{n-1} b_{0,i} t^{(n-i)p+i(m-1)} Y^{n-i} + a_1 \sum_{i=0}^{n-2} b_{1,i} t^{(n-1-i)p+i(m-1)} Y^{n-1-i} + \cdots + a_n, \qquad b_{j,i} \in \mathbb{Z}. \end{aligned}$$

For g to be regular singular with respect to δ_{m-p}, the corresponding Newton polygon must have single slope $1-m+p$. Note that the coefficient of the Y^n term is $a_0 t^{np}$ so that the point $(n, np + v_t(a_0))$ is the right-most point of the new Newton polygon of g. In order for the Newton polygon to have a single slope $1-m+p$, any point arising from a monomial t^jY^k must lie above the point $\big(k, np + (k-n)(1-m+p) + v_t(a_0)\big)$. We now look at the minimum valuation of the Y^k term. The coefficient of the Y^k term is

$$c_k = \sum_{j=0}^{n-k} a_j b_{j,n-j-k} t^{kp+(n-j-k)(m-1)}, \qquad b_{j,i} \in \mathbb{Z}.$$

Now the valuation of c_k must be achieved for some $0 \le j \le m-k$. But we have:

$$\begin{aligned} & v_t(a_j) + v_t\big(t^{kp+(n-j-k)(m-1)}\big) \\ & \ge v_t(a_0) - j(1-m) + kp + (n-j-k)(m-1) \\ & = kp + (n-k)(m-1) \\ & = np + (k-n)(1-m+p). \end{aligned}$$

Hence $v(c_k) \ge np + (k-n)(1-m+p)$ and so all the points (a,b) corresponding to monomials t^bY^a of $g(Y)$ lie above the line of slope $1-m+p$. Thus the Newton polygon of g with respect to δ_{m-p} has single slope $1-m+p$ and so g is regular singular with respect to δ_{m-p}. □

Remark 7.24. Firstly, in light of Remark 7.12 we do not lose generality by only considering changes of the form $x = t^pY$. If the change made involves a rational power of t then one simply extends the field $\mathcal{K}$ and relabels variables. Secondly, if we consider $f \in \mathcal{K}\{x,d\}$ for an arbitrary derivation $d = a\frac{d}{dt}$, $a \in \mathcal{K}$, then the only term of a that will affect the valuation of the coefficients of $g(Y)$ is the term involving the lowest power of t. Hence it is enough to consider only derivations of the form δ_m. That is, Lemma 7.23 shows that the property of being regular singular is preserved under a change of variables $x = t^pY$ for all derivations.

7.5. **Factorisation of Differential Polynomials.** We are now in a position to generalise our factorisation result to the full ring $\mathcal{K}\{x,\delta_1\}$.

Proposition 7.25. *Every $f \in \mathcal{K}\{x,\delta_1\}$ has a linear factorisation.*

Proof. We use induction on the degree, n, of the polynomial. The case $n=1$ is obvious, so let $f_1 = \sum_i a_i x^i \in \mathcal{K}\{x,\delta_1\}$ be monic and set $r_1 := \min\left\{\frac{v_t(a_i)}{i}\right\} = \frac{p}{q}$ with $\gcd(p,q)=1$. We may assume that $r_1 < 0$ otherwise we can apply Proposition 7.14 to factor f_1. Making the change of variables $x = t^{r_1}Y$ and applying Lemma 7.20 yields a monic polynomial, $g_1(Y) \in \mathcal{K}_q\left\{Y, \frac{1}{q}s^{1-p}\frac{d}{ds}\right\}$, with power series coefficients. Hence, we can reduce the coefficients of $g_1(Y)$ to obtain a polynomial, $\bar{g}_1(Y) \in \mathbb{C}[x]$. If $\bar{g}_1(Y)$ has at least two distinct roots, then we may apply Proposition 7.14 (with the derivation δ_{1-p}) to obtain a factorisation of $g_1(Y)$ and we are done by induction. If this is not the case, then $\bar{g}_1(Y)$ has a single repeated root and thus satisfies the hypotheses of Lemma 7.21. Hence, the Newton polygon of $f_1(x)$ has a single *integral* slope. By Lemma 7.22, we can apply the change of variables $x \mapsto x - \lambda t^r$ to f_1 to yield a new polynomial, $f_2(x) := x^n + b_1x^{n-1} + \cdots + b_n \in \mathcal{K}\{x,\delta_1\}$. The Newton polygon of f_2 will have a single slope strictly less than the slope of the Newton polygon of f_1.

Now we start the process with the polynomial $f_2(x)$; i.e., we let $r_2 = \min\left\{\frac{v(b_i)}{i}\right\}$. If $r_2 \ge 0$ we are done. Otherwise, we make the change of variable $x \mapsto t^{r_2}y$ to obtain a new polynomial $g_2(y)$. If $\bar{g}_2(y)$ has distinct roots, then we are done; otherwise, applying Lemma 7.21 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. □

Remark 7.26. We can extend this result to the case of an arbitrary derivation $d = a\frac{d}{dt}$, $a \in \mathcal{K}$. Given a polynomial, $f \in \mathcal{K}\{x,d\}$, the change of variables $x = t^{-1}aY$ will yield a new polynomial in the ring $\mathcal{K}\{Y, \delta_1\}$. We have shown that this new polynomial has a factorisation (after a possible finite extension) and hence the original polynomial also has a factorisation.

We now provide a number of examples to illustrate Proposition 7.25.

Example 7.27. The first example requires only a simple change of variables, after which we can immediately apply Hensel's lemma. Let $f = x^2 + (2 - t^{-1})x - t^{-1} \in \mathcal{K}\{x, \delta_1\}$. We have $r = \min\left\{\frac{-1}{1}, \frac{-1}{2}\right\} = -1$ so we make the substitution $x = t^{-1}Y$. This yields:

$$\begin{aligned} f(t^{-1}Y) &= (t^{-1}Y)^2 + (2 - t^{-1})t^{-1}Y - t^{-1} \\ &= t^{-1}Yt^{-1}Y + 2t^{-1}Y - t^{-2}Y - t^{-1} \\ &= t^{-1}(t^{-1}Y - 1)Y + 2t^{-1}Y - t^{-2}Y - t^{-1} \\ &= t^{-2}(Y^2 - (1-t)Y - t). \end{aligned}$$

Hence, in this case, $g(Y) = Y^2 - (1-t)Y - t$ and $\bar{g}(Y) = Y^2 - Y = Y(Y-1)$ and so we may apply Proposition 7.14. One iteration yields the full factorisation $g(Y) = (Y+t)(Y-1)$. Working backwards, we obtain a factorisation of f:

$$\begin{aligned} f &= t^{-2}(Y+t)(Y-1) = t^{-2}(tx+t)(tx-1) \\ &= (t^{-1}x + t^{-1})(tx-1) = (xt^{-1} + t^{-1} + t^{-1})(tx-1) \\ &= (x+2)t^{-1}(tx-1) = (x+2)(x - t^{-1}). \end{aligned}$$

8. Jordan Decomposition for Differential Operators

In the previous section we established a differential analogue of the classical Newton-Puiseux theorem. We will now show that this provides a direct proof of Levelt's corollary (c.f. [Lev75, §1]), i.e. we show that every differential operator has an eigenvector. In the linear setting, the existence of an eigenvector does not guarantee an eigenspace decomposition. Instead, one must extend the notion of eigenspace to that of generalised eigenspaces. The Jordan decomposition theorem follows fairly easily once the generalised eigenspace decomposition has been established. In this section we provide a proof of the Hukuhara-Levelt-Turrittin theorem in analogy to the proofs given in Section 4.

8.1. **Eigenvalues, Semisimplicity, and Diagonalisability.**

As in the linear setting, Proposition 7.25 gives us a simple proof that every differential operator has an eigenvector (after a possible finite extension). Note that this provides a direct proof of Levelt's corollary (see [Lev75, §1]).

Proposition 8.1. *Let $D : V \to V$ be a differential operator. There exists a finite extension, $\mathcal{K}_b$, of $\mathcal{K}$ such that D_b has an eigenvector.*

Proof. The argument proceeds exactly as in the linear setting. Let $D : V \to V$ be a differential operator and $v \in V$ be a non-zero vector. Consider the sequence $v, D(v), D^2(v), \cdots$. As V has finite dimension over $\mathcal{K}$, we must have that

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

where $n = \dim_{\mathcal{K}}(V)$. Now consider the polynomial $f(x) = x^n + a_1 x^{n-1} + \cdots + a_n \in \mathcal{K}\{x, \delta_1\}$. By Proposition 7.25, we can, after a finite extension, write

$$f(x) = (x - \Lambda_1) \cdots (x - \Lambda_n) \in \mathcal{K}_b\{x, \delta_1'\}, \quad b \in \mathbb{Z}_{>0}, \Lambda_i \in \mathcal{K}_b.$$

Thus,

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

Let $i \in \{1, 2, \ldots, n\}$ 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 Λ_n. Otherwise $(D_b - \Lambda_{i+1}) \cdots (D_b - \Lambda_n) v$ is an eigenvector of D_b with eigenvalue Λ_i. □

There are two important corollaries of Proposition 8.1 which are analogues of well-known results in the linear setting. The rest of this section will be devoted to proving these corollaries. Note that the proof of Corollary 8.5 provides a simple proof of Theorem II in [Lev75, §1].

Corollary 8.2. *Let $D : V \to V$ be a differential operator. After a possible finite extension to $\mathcal{K}_b$, there exists a basis of V_b such that the operator $D_b : V_b \to V_b$ can be represented by a matrix in upper-triangular form.*

Proof. By Proposition 8.1, there exists an extension, $\mathcal{K}_b$, of $\mathcal{K}$ and an eigenvector $v_1 \in V_b$. We now induct on $\dim_{\mathcal{K}_b}(V_b)$. If $\dim(V_b) = 1$ then the result is trivial. Suppose $n := \dim(V_b) > 1$ and the result holds for all differential operators on vector spaces of dimension $n - 1$. Set $W = \operatorname{span}_{\mathcal{K}_b}\{v_1\}$. Since W is D_b-invariant we can form the quotient differential module V_b/W with the induced differential operator, D_b', given by

$$D_b'(v + W) = D_b(v) + W.$$

Clearly this quotient has dimension $n - 1$ over $\mathcal{K}_b$. By the inductive hypothesis, there exists a $\mathcal{K}_b$ basis $\{v_2 + W, \ldots, v_n + W\}$ of V_b/W such that D_b' has an upper-triangular matrix with respect to this basis. This implies that

$$D_b'(v_j + W) \in \operatorname{span}\{v_2 + W, \ldots, v_j + W\}$$

for each $j = 2, \ldots, n$. Hence $D_b(v_j) \in \operatorname{span}\{v_1, \ldots, v_j\}$ for each $j = 1, \ldots, n$ and so D_b has an upper-triangular matrix with respect to the basis $\{v_1, \ldots, v_n\}$. □

As one would expect, given a differential operator in upper-triangular form, the diagonal entries are precisely the eigenvalues of the differential operator. This shows that, up to similarity[6], a differential operator has only finitely many eigenvalues. The following lemma and its corollary are valid for arbitrary derivations on $\mathcal{K}$. Note that $a \in \mathcal{K}$ being an eigenvalue of D is equivalent to the statement $D - a$ is not injective.

Lemma 8.3. *Let $D : V \to V$ be a differential operator and suppose that there exists a basis, $\{v_1, \ldots, v_n\}$ of V such that $D = d + A$ where A is upper-triangular. Then D is injective if, and only if, none of the entries on the diagonal of A are similar to zero.*

[6] Note that there may be infinitely many *similar* eigenvalues.

Proof. Firstly, we will prove that if one of the diagonal entries of A is similar to zero, then D is not injective. Let $a_1, \ldots, a_n$ be the diagonal entries of A. If a_1 is similar to 0 then $a_1 = c^{-1}d(c)$ for some $c \in \mathcal{K}$. We have $Dv_1 = a_1 v_1$ and so $D(c^{-1}v_1) = (a_1 - c^{-1}d(c))(c^{-1}v_1) = 0$. Hence D is not injective. Now suppose a_k is similar to 0 for some $1 < k \leq n$, i.e. $a_k = c^{-1}d(c)$ for some $c \in \mathcal{K}$. Since $D(c^{-1}v_k) \in \operatorname{span}_{\mathcal{K}}\{v_1, \ldots, v_{k-1}\}$ we can write:

$$D(c^{-1}v_k) = c_1 v_1 + \cdots + c_{k-1}v_{k-1}, \qquad c_i \in \mathcal{K}.$$

Now consider the $\mathbb{C}$-vector space $W_1 = \operatorname{span}_{\mathbb{C}}\{c_1 v_1, \ldots, c_{k-1}v_{k-1}\}$. Let $W_2 = \operatorname{span}_{\mathbb{C}}\{c_1 v_1, \ldots, c^{-1}v_k\}$ and consider the $\mathbb{C}$-linear operator:

$$D|_{W_2} : W_2 \to W_1.$$

Since $\dim_{\mathbb{C}}(W_2) > \dim_{\mathbb{C}}(W_1)$ this map is not injective. Hence D itself is not injective.

In the other direction, suppose D is not injective. Then there exists $v \in V$ such that $Dv = 0$. Write $v = c_1 v_1 + \cdots + c_k v_k$ with $c_k \neq 0$. Then

$$0 = Dv = D(c_1 v_1 + \cdots + c_k v_k) = c_1 Dv_1 + d(c_1)v_1 + \cdots + c_k Dv_k + d(c_k)v_k.$$

Hence $Dv_k + c_k^{-1}d(c_k) \in \operatorname{span}_{\mathcal{K}}\{v_1, \ldots, v_{k-1}\}$ since A is upper-triangular. Noting that $c_k^{-1}d(c_k)v_k \notin \operatorname{span}_{\mathcal{K}}\{v_1, \ldots, v_{k-1}\}$, we must have

$$Dv_k = b_1 v_1 + \cdots + b_{k-1}v_{k-1} - c_k^{-1}d(c_k)v_k, \qquad b_i \in \mathcal{K},$$

and so the matrix for D in the basis $\{v_1, \ldots, v_n\}$ has value $-c_k^{-1}d(c_k)$ as the k^{th} diagonal entry. That is, a_k must be similar to 0. $\square$

Corollary 8.4. *Let $D : V \to V$ be a differential operator and suppose that there exists a basis, $\{v_1, \ldots, v_n\}$ of V such that $D = d + A$ where A is upper-triangular. Then the diagonal entries of A are precisely the eigenvalues of D (up to similarity).*

Proof. Let $a \in \mathcal{K}$ and consider the matrix of $D - a$ with respect to the basis $\{v_1, \ldots, v_n\}$:

$$D - a = d + \begin{pmatrix} a_1 - a & * & * & * \\ & a_2 - a & * & * \\ & & \ddots & * \\ & & & a_n - a \end{pmatrix}.$$

By Lemma 8.3, $D - a$ is not injective if, and only if, $a_k - a$ is similar to 0 for some $1 \leq k \leq n$. That is, a is an eigenvalue if, and only if, a is similar to a_k for some $1 \leq k \leq n$. $\square$

We now prove one of the main results in [Lev75]. We will make use of Lemma 6.19 which appears in [Lev75, §1(e)].

Corollary 8.5. *Let $D : V \to V$ be a differential operator. Then D is semisimple if, and only if, for some finite extension, D_b is diagonalisable.*

Proof. Suppose $D : V \to V$ is a semisimple differential operator. By Proposition 8.1, there exists a finite extension, $\mathcal{K}_b$, of $\mathcal{K}$ such that $D_b : V_b \to V_b$ has an eigenvetor, v. We prove by induction on the dimension of V_b that D_b can be represented by a diagonal matrix. If $\dim_{\mathcal{K}_b}(V_b) = 1$ the result is obvious. Suppose $\dim_{\mathcal{K}_b}(V_b) > 1$ and let $U = \operatorname{span}_{\mathcal{K}_b}\{v\}$. Then U is a one-dimensional, D_b-invariant subspace of V_b. By Lemma 6.19, D_b is also semisimple so there exists a D_b-invariant complement,

W, of U. Now $D_b : W \to W$ is semisimple so by our induction hypothesis, we can write W as a direct sum of one-dimensional D_b-invariant subspaces. Thus we have a decomposition of V_b into one-dimensional, D_b-invariant subspaces and hence D_b is diagonalisable.

Conversely, suppose D_b is a diagonalisable operator. Then clearly D_b is semisimple. By the previous proposition, D is also semisimple. $\square$

8.2. **Generalised Eigenspace Decomposition.** In the linear setting, the existence of an eigenvector does not guarantee an eigenspace decomposition. Instead, one must extend the notion of eigenspace to that of generalised eigenspaces. Once a generalised eigenspace decomposition has been established, proving the existence of a Jordan decomposition reduces to the case that the linear operator is nilpotent. In this section we introduce the notion of generalised eigenspaces for differential operators and prove an analogue of the generalised eigenspace decomposition. Firstly, we give an analogue of the rank-nullity theorem for formal differential operators due to Malgrange [Mal74, Theorem 3.3].

Let (V, D) be a differential module over $\mathcal{K}$. Define

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

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

Note that these are vector spaces over $\mathbb{C}$ (not over $\mathcal{K}$).

Proposition 8.6. *Let $D : V \to V$ be a formal differential operator. Then*

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

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

$$0 \to V \to V'' \to V' \to 0$$

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 8.7. *Let $D : V \to V$ and $D' : V' \to V'$ be two formal differential operators. Then, we have*

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

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

Proof. One can show (see [Ked10a, Lemma 5.3.3]) that there is a canonical isomorphism of $\mathbb{C}$-vector spaces:

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

This fact together with Proposition 8.6 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). $\square$

Definition 8.8. *Let $D : V \to V$ be a formal differential operator and let $a \in \mathcal{K}$. The* generalised eigenspace $V(a)$ *of D is defined as*

$$V(a) := \operatorname{span}_{\mathcal{K}}\{v \in V \mid (D-a)^n v = 0, \quad \textit{for some positive integer } n.\}$$

Theorem 8.9 (Generalised Eigenspace Decomposition)**.** *There exists a finite extension $\mathcal{K}_b/\mathcal{K}$ such that we have a canonical decomposition $V_b = \bigoplus_i V_b(a_i)$, $a_i \in \mathcal{K}_b$. Moreover,*

$$V_b(a_i) \cap V_b(a_j) \neq \{0\} \iff a_i \textit{ is similar to } a_j \iff V_b(a_i) = V_b(a_j).$$

Proof. We may assume, without the loss of generality, that all eigenvalues of D are already in $\mathcal{K}$ (if not, carry out 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, $V \in \operatorname{Ext}^1_{\mathcal{K}\{x\}}(U, W)$. 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

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

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 $\operatorname{Ext}^1_{\mathcal{K}\{x\}}(U, W(a_j))$. But it is easy to see that all differential operators in $\operatorname{Ext}^1_{\mathcal{K}\{x\}}(U, W(a_j))$ have only a single eigenvalue a_j (up to similarity). Hence V has the required decomposition. $\square$

8.3. **Unipotent Differential Operators.** Theorem 8.9 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 8.10. *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 $\mathrm{Nilp}_{\mathbb{C}}$ 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 $\mathrm{Nilp}_{\mathbb{C}}$ are linear maps which commute with N. Let $\mathcal{U}$ be the category of pairs (V, D) consisting of a $\mathcal{K}$-vector space V and a unipotent differential operator $D : V \to V$. Define a functor

$$F : \mathrm{Nilp}_{\mathbb{C}} \to \mathcal{U}, \qquad (V, N) \mapsto (\mathcal{K} \otimes_{\mathbb{C}} V, d + N).$$

The following result appears (without proof) in [Kat87, §2].

Lemma 8.11. *The functor F defines an equivalence of categories with inverse given by*

$$G : \mathcal{U} \to \mathrm{Nilp}_{\mathbb{C}}, \qquad (V, D) \mapsto \big(\ker(D^{\dim_{\mathcal{K}}(V)}), D\big).$$

Proof. We first show that the composition $G \circ F$ equals the identity. Let $(V, N) \in \mathrm{Nilp}_{\mathbb{C}}$ 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 $D : V \to V$ 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 $U := \mathrm{span}_{\mathcal{K}}\{v\}$ and consider the differential module V/U. This has dimension $n - 1$ so we may assume there exist a set $\{v_1, \ldots, v_n\}$ of $\mathcal{K}$-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

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

Hence $D^n(v_i - b_i v) = D(a_{i,0} v) = 0$ so $\{v, v_1 - b_1 v, \ldots, v_{n-1} - b_{n-1} v\}$ 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 $W := \ker(D^n) = \mathrm{span}_{\mathbb{C}}\{v, v_1 - b_1 v, \ldots, v_{n-1} - b_{n-1} v\}$. 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 8.12. A formal differential operator D is said to be *regular singular* if it has a matrix representation of the form

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

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* [BV83, §3]. The above lemma implies that a unipotent differential operator is the same as a regular singular differential operator with unipotent monodromy.

8.4. **Jordan Decomposition.** We are now in a position to prove the Hukuhara-Levelt-Turrittin theorem. The uniqueness part of the theorem is relatively easy. Since we do not have anything new to add to Levelt's original proof, we refer the reader to [Lev75] for the details. It remains to prove existence. For convenience, we first restate the theorem.

Theorem 8.13 (Hukuhara-Levelt-Turrittin). *Let $D : V \to V$ be a differential operator on a $\mathcal{K}$-vector space V. Then D can be written as a sum $D = S + N$ of a semisimple differential operator S and a nilpotent $\mathcal{K}$-linear operator N such that S and N commute. Moreover, the pair (S, N) is unique.*

Proof. Let $D : V \to V$ be a formal differential operator. By Theorem 8.9, there exists a positive integer b such that $D_b : V_b \to V_b$ 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 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 8.11. This proves the existence of a 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 $\sigma \in \mathrm{Gal}(\mathcal{K}_b/\mathcal{K})$, 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 $\mathcal{K}$. □

9. Jordan Decomposition for G-Connections

9.1. G**-Connections.** In this section we consider differential operators in a Lie-theoretic framework. The starting point is to define an analogue of the gauge action for reductive algebraic groups. As in Section 6.2, gauge equivalence classes will then be what we call differential operators. In this context, however, differential operators are usually called connections.

For simplicity, we work over the field $\mathcal{K} := \mathbb{C}(\!(t)\!)$. Let G be a connected, reductive algebraic group over $\mathbb{C}$ and $\mathfrak{g} = \mathrm{Lie}(\mathrm{G})$. The Maurer-Cartan one-form allows us to define the gauge action of $\mathrm{G}(\mathcal{K})$ (the $\mathcal{K}$-points of G) on $\mathfrak{g}_{\mathcal{K}} := \mathfrak{g} \otimes_{\mathbb{C}} \mathcal{K}$.

Lemma 9.1. *Given $g \in \mathrm{G}(\mathcal{K})$ there exists a unique element $\delta_{\mathrm{G}}(g) \in \mathfrak{g}_{\mathcal{K}}$ such that for all rational representations $\rho : \mathrm{G} \to \mathrm{GL}(V)$,*

$$d\rho(\delta_{\mathrm{G}}(g)) = (\rho(g))^{-1} d(\rho(g)),$$

where $d(\rho(g))$ is the matrix obtained by applying the derivation element-wise to the matrix $\rho(g)$.

Remark 9.2. There are a number of ways one can make sense of the element $\delta_{\mathrm{G}}(g)$ (see [BV83, §1.6], [Fre07, §1.2.4], [Ras15, §1.12])

Note that Lemma 9.1 has two important consequences:

1. Taking $\rho = \mathrm{Ad} : \mathrm{G} \to \mathrm{GL}(\mathfrak{g})$ we have $\mathrm{ad}(\delta_{\mathrm{G}}(g)) = (\mathrm{Ad}_g)^{-1} d(\mathrm{Ad}_g)$;
2. Taking ρ to be the identity morphism we see that $\delta_{\mathrm{GL}(V)}(g) = g^{-1} d(g)$.

Hence we have:

$$\mathrm{ad}(\delta_{\mathrm{G}}(g)) = \delta_{\mathrm{GL}(\mathfrak{g})}(\mathrm{Ad}_g), \qquad \forall g \in \mathrm{G}\,.$$

We now define an action of $\mathrm{G}(\mathcal{K})$ on its Lie algebra $\mathfrak{g}_{\mathcal{K}}$:

Lemma 9.3 (Gauge Action)**.** *There exists an action of* $\mathrm{G}(\mathcal{K})$ *on* $\mathfrak{g}_{\mathcal{K}}$*, called the gauge action, given by:*

$$g \cdot A = \mathrm{Ad}_g(A) + \delta_{\mathrm{G}}(g), \quad A \in \mathfrak{g}_{\mathcal{K}}, g \in \mathrm{G}(\mathcal{K}).$$

The gauge action induces an equivalence relation on $\mathfrak{g}_{\mathcal{K}}$. This gives us the following definition:

Definition 9.4 (G-Connection)**.** *A* G-connection *is a gauge equivalence class of elements of* $\mathfrak{g}_{\mathcal{K}}$.

We will denote the gauge equivalence class containing A by both $[A]$ and $d + A$. As a consequence of Lemma 9.3 we can define an adjoint connection:

Lemma 9.5 (Adjoint Connection)**.** *Let $A \in \mathfrak{g}_{\mathcal{K}}$ and let $[A]$ be the corresponding* G*-connection. The map:*

$$[A] \mapsto [\mathrm{ad}_A] \tag{9.20}$$

is a well-defined map from G*-connections to* $\mathrm{GL}(\mathfrak{g}_{\mathcal{K}})$*-connections.*

Proof. Note that $g \cdot A \in \mathfrak{g}_{\mathcal{K}}$ so we can consider the adjoint representation of $g \cdot A$:

$$\begin{aligned}\mathrm{ad}(g \cdot A) &= \mathrm{Ad}_{\mathrm{Ad}_g}(\mathrm{ad}(A)) + \mathrm{ad}(\delta_G(g))\\ &= (\mathrm{Ad}_g)^{-1}\,\mathrm{ad}(A)\,\mathrm{Ad}_g + \delta_{GL(\mathfrak{g})}(\mathrm{Ad}_g)\\ &= (\mathrm{Ad}_g)^{-1}\,\mathrm{ad}(A)\,\mathrm{Ad}_g + (\mathrm{Ad}_g)^{-1} d(\mathrm{Ad}_g).\end{aligned}$$

It is clear that $\mathrm{ad}(g \cdot A)$ is gauge equivalent to $\mathrm{ad}(A)$ so the adjoint representation respects gauge equivalence. Thus the map

$$d + A \mapsto d + \mathrm{ad}(A)$$

is a well-defined map from G-connections to $\mathrm{GL}(\mathfrak{g})$-connections. □

9.2. **Semisimple** G**-Connections.** Lemma 9.5 will allow us to define semisimple G-connections in analogy to the definition of semisimple elements of a Lie algebra in the linear setting. Unfortunately, we will have to deal with the same problem that arose in the linear setting: when $\mathfrak{g}_{\mathcal{K}} \subset \mathfrak{gl}(V)$, we will have two possibly different definitions of semisimplicity. The results of this section will allow us to show that both definitions are equivalent.

Definition 9.6 (Semisimple G-Connection)**.** *Let $d + A$ be a* G*-connection. We call $d + A$* semisimple *if $d + \mathrm{ad}(A) : \mathfrak{g}_{\mathcal{K}} \to \mathfrak{g}_{\mathcal{K}}$ is a semisimple differential operator.*

We now show that this definition coincides with the usual definition of semisimplicity in the case $\mathrm{G} \subseteq \mathrm{GL}(V)$.

Lemma 9.7. *Let $d + A : V \to V$ be a differential operator, where $A \in \mathfrak{gl}(V)$ and V is a $\mathcal{K}$-vector space. Then $d + \mathrm{ad}\, A_S + \mathrm{ad}\, A_N$ is the Jordan decomposition of $d + \mathrm{ad}\, A$.*

Proof. Note that $\mathrm{ad}\, A$ is a $\mathcal{K}$-linear operator on $\mathfrak{gl}(V)$. Thus, $d+\mathrm{ad}\, A : \mathfrak{gl}(V) \to \mathfrak{gl}(V)$ is a differential operator. We claim that $d+\mathrm{ad}\, A_S+\mathrm{ad}\, A_N$ is its Jordan decomposition. To see this, we pick a basis for V to put $d + A$ in Jordan normal form (after a possible finite extension). In this case, A_S is diagonal and A_N is constant nilpotent matrix with 1's on the sub-diagonal, and A_S and A_N commute. Thus, $d + \mathrm{ad}(A_S)$ is a semisimple differential operator on $\mathfrak{gl}(V \otimes_{\mathcal{K}} \mathcal{K}_b)$. We claim that it commutes with $\mathrm{ad}(A_N)$. Indeed,

$$[d + \mathrm{ad}(A_S), \mathrm{ad}(A_N)] = [d, \mathrm{ad}(A_N)] + [\mathrm{ad}(A_S), \mathrm{ad}(A_N)].$$

Now $\mathrm{ad}(A_N)$ is constant, so the first bracket is zero. Since A_S and A_N commute, the second bracket is also zero. As in the proof of Theorem 8.13, this decomposition descends to a decomposition over $\mathcal{K}$. □

Corollary 9.8. *Let $d + A$ be as above. Then $d + A$ is semisimple if and only if $d + \mathrm{ad}\, A$ is semisimple.*

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

Note that Corollary 9.8 guarantees that both definitions of semisimple are equivalent in the case $\mathrm{G} \subseteq \mathrm{GL}(V)$.

We give one further characterisation of semisimplicity for G-connections. This should be viewed as a generalisation of Corollary 8.5. In the setting of G-connections an element $A \in \mathfrak{g}_{\mathcal{K}}$ is "diagonalisable" if it lies in a Cartan subalgebra $\mathfrak{h} \otimes K \subset \mathfrak{g} \otimes K$. A semisimple G-connection will be "diagonalisable" after a possible finite extension of $\mathcal{K}$. That is:

Lemma 9.9. *Let $d + A$ be a* G*-connection. Then $d + A$ is semisimple if, and only if, A is gauge equivalent to an element of $\mathfrak{h} \otimes_{\mathcal{K}} \mathcal{K}_b$ for some finite extension $\mathcal{K}_b \supset \mathcal{K}$.*

In order to prove this result, we require some properties of differential Galois groups. We briefly define and state these properties now. For a proper treatment of this topic see [vdPS03, §1.4] and [Kat87, §2].

Definition 9.10. *Let $\omega : \mathrm{Diff}_{\mathcal{K}} \to \mathrm{Vect}_{\mathbb{C}}$ be a fibre functor*[7]*. The* local differential Galois group*, I, is defined to be the tensor automorphisms of this fibre functor, i.e.*

$$I := \mathrm{Aut}^{\otimes}(\omega).$$

In the language of Tannakian categories, I is the affine group scheme such that the category $\mathrm{Diff}_{\mathcal{K}}$ is equivalent to the category of representations of I. Thus, given a non-zero differential module (V, D) over $\mathcal{K}$, we obtain a representation

$$\rho_V : I \to \mathrm{GL}(\omega(V)).$$

We claim that the image $\rho_V(I)$, is potentially diagonalisable.

Lemma 9.11. *Let (V, D) be a differential module. Then D is semisimple if and only if there exists a finite extension $\mathcal{K}_b$ of $\mathcal{K}$ such that $\rho_V(I_{\mathcal{K}_b})$ is diagonalisable.*

Proof. We have already shown that D is semisimple if and only if its Jordan form is diagonalisable after some finite extension of $\mathcal{K}$. Hence, we need only show that having diagonalisable Jordan form is equivalent to $\rho(I_{\mathcal{K}_b})$ being diagonalisable.

One direction follows from a basic fact about differential Galois groups: if $H \subset \mathrm{GL}(n, \mathbb{C})$ is a closed subgroup and $D = d + A$ with $A \in \mathfrak{h}$ (where $\mathfrak{h} = \mathrm{Lie}(H)$), then $\rho(I_{\mathcal{K}})$ is contained in a conjugate of H (see [vdPS03, Prop. 1.31]). If the Jordan form of D is diagonalisable, then we can write $D = d + A$ with A diagonal (i.e. A is contained in $\mathrm{Lie}(T)$ for a maximal torus $T \subset \mathrm{GL}(n, \mathbb{C})$). Hence $\rho(I_{\mathcal{K}_b})$ is conjugate to a maximal torus in $\mathrm{GL}(n, \mathbb{C})$ and is therefore diagonalisable.

For the other direction, suppose that the Jordan form of D is given by

$$D = d + (D_{-r}t^{-r} + \cdots + D_{-1}t^{-1} + D_0 + N), \quad N \neq 0.$$

[7]See the discussion in Remark 6.28 or [Kat87, §2.5]

In this case, $\exp(D_0 + N) \in \rho(I_{\mathcal{K}_b})$[8] (see [vdPS03, §3.2]). Since D_0 and N commute, we have the (multiplicative) Jordan form of $\exp(D_0 + N)$ given by $\exp(D_0)\exp(N)$. Since $N \neq 0$, $\exp(N) \neq \mathrm{Id}$ and so $\exp(D_0 + N)$ is not diagonalisable. Hence $\rho(I_{\mathcal{K}_b})$ is not diagonalisable. □

We are now able to prove Lemma 9.9.

Proof of Lemma 9.9. The 'if' direction being obvious we consider the opposite implication. Let $\rho : I_{\mathcal{K}} \to \mathrm{G}$ denote the corresponding homomorphism. By assumption $d + A$ is semisimple; that is to say, $d + \mathrm{ad}(A)$ is semisimple. By Lemma 9.9, there exists a finite extension $\mathcal{K}_b$ of $\mathcal{K}$ such that the image of the composition

$$I_{\mathcal{K}_b} \xrightarrow{\rho} \mathrm{G} \xrightarrow{\mathrm{Ad}} \mathrm{GL}(\mathfrak{g}_{\mathcal{K}})$$

is diagonalisable. This implies that the image $\rho(I_{\mathcal{K}_b})$ is contained in a maximal torus $H \subset \mathrm{G}$. It follows that $d + A$ is gauge equivalent, under $\mathrm{G}(\mathcal{K}_b)$, to a connection of the form $d + X$ where $X \in \mathfrak{h} \otimes_{\mathcal{K}} \mathcal{K}_b$. □

9.3. **Invariant Properties of Differential Operators.** We would like to show that, given a differential operator $D : V \to V$, any D-invariant subspace, $W \subset V$, is also S-invariant, where $D = S + N$ is the Jordan decomposition of D. We have already seen that V decomposes into generalised eigenspaces and that these generalised eigenspaces are D, S and N invariant. Hence we need only consider the case where V itself is a generalised eigenspace. In this case, there exists a finite extension of $\mathcal{K}$ such that $S = \lambda I + d$.

We first prove the result in the case of unipotent differential operators (i.e., the case $\lambda = 0$ — see Section 8.3). We denote by $\mathcal{U}$ the category of unipotent differential operators. The objects of this category are pairs (V, D) with $D : V \to V$ unipotent. There is an equivalence of categories between $\mathcal{U}$ and Nilp given by the functor

$$F : \mathcal{U} \to \mathrm{Nilp}, \qquad (V, D) \mapsto (\ker(D^n), D),$$

and its inverse

$$G : \mathrm{Nilp} \to \mathcal{U}, \quad (V, N) \mapsto (V \otimes_{\mathbb{C}} \mathcal{K}, N \otimes_{\mathbb{C}} d).$$

Proposition 9.12. *Let $D : V \to V$ be a unipotent differential operator and suppose that $W \subset V$ is a D-invariant subspace. Then W is also S-invariant.*

Proof. The restriction $D|_W : W \to W$ gives us an inclusion in the category $\mathcal{U}$. Under the equivalence F we obtain an inclusion in Nilp:

$$\begin{tikzcd} (W, D|_W) \arrow[r, hook] \arrow[d, leftrightarrow] & (V, D) \arrow[d, leftrightarrow] \\ (W_0, N') \arrow[r, hook] & (V_0, N) \end{tikzcd}$$

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.

[8] In fact, this element topologically generates the formal monodromy part of the differential Galois group.

Proposition 9.13. *Let $D : V \to V$ be a differential operator and suppose that $W \subset V$ is a D-invariant subspace. Then W is also S-invariant.*

Proof. 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 Proposition 9.12, 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. □

The results of this section allow us to prove the following important fact:

Lemma 9.14. *Let $\mathfrak{g} \subset \mathfrak{gl}(V)$. Let $d + A : V \to V$ be a differential operator with $A \in \mathfrak{g}_{\mathcal{K}}$ and suppose that $d + A_S + A_N$ is its Jordan decomposition (in $\mathfrak{gl}(V) \otimes_{\mathbb{C}} \mathcal{K}$). Then $A_S \in \mathfrak{g}$.*

Proof. Note that $\mathfrak{g}_{\mathcal{K}} \subset \mathfrak{gl}(V) \otimes_{\mathbb{C}} \mathcal{K}$ is a $d + \operatorname{ad} A$-invariant subspace of $\mathfrak{gl}(V) \otimes_{\mathbb{C}} \mathcal{K}$ and hence, by Proposition 9.13, it is also $d + \operatorname{ad} A_S$-invariant. By definition, $\mathfrak{g}_{\mathcal{K}}$ is d invariant. Thus, $\mathfrak{g}_{\mathcal{K}}$ is $\operatorname{ad} A_S$-invariant. This implies that $\operatorname{ad} A_S : \mathfrak{g}_{\mathcal{K}} \to \mathfrak{g}_{\mathcal{K}}$ is a $\mathcal{K}$-linear derivation on $\mathfrak{g}_{\mathcal{K}}$ and hence $A_S \in \mathfrak{g}_{\mathcal{K}}$. □

9.4. **Jordan Decomposition for G-Connections.** As in the case of differential operators, there is a notion of a Jordan decomposition for G-connections.

Theorem 9.15. *Every operator $d + A$, $A \in \mathfrak{g}_{\mathcal{K}}$ can be written as $d + A_S + A_N$, where $A_S, A_N \in \mathfrak{g}_{\mathcal{K}}$, $d + A_S$ is semisimple, A_N is nilpotent, and $d + A_S$ and A_N commute (in the extended loop algebra $\hat{\mathfrak{g}}$). Moreover, this decomposition is unique.*

Proof of Theorem 9.15. Note that the adjoint action gives an embedding $\mathfrak{g}_{\mathcal{K}} \subset \mathfrak{gl}(\mathfrak{g}_{\mathcal{K}})$. Using this embedding we get a Jordan decomposition $d + A = d + A_S + A_N$ of $d + A$ as an operator on $\mathfrak{g}_{\mathcal{K}}$. By Lemma 9.14, $A_S \in \mathfrak{g}_{\mathcal{K}}$ and it follows that $A_N \in \mathfrak{g}_{\mathcal{K}}$ also. Hence we have a Jordan decomposition of $d + A$ in $\mathfrak{g}_{\mathcal{K}}$. It is clear that this decomposition is unique. □

10. Outlook

It is hoped that the material in this thesis gives the reader a more intuitive understanding of the Hukuhara-Levelt-Turrittin theorem and of differential operators in general. This theorem has already found many applications most notably in the geometric Langlands correspondence (see [Ras15, KS16]). We now briefly discuss some possible extensions of the work presented here. As usual, we fix $\mathcal{K} = \mathbb{C}(\!(t)\!)$ and $\mathcal{K}_b = \mathbb{C}(\!(t^{1/b})\!)$.

The first, and perhaps most obvious, direction for further research is the question of rationality: given a differential operator $D = d + A$ over $\mathcal{K}_b$ in Jordan form, does there exist a differential operator over $\mathcal{K}$ whose Jordan form is $d + A$? This question has been answered in [BV83, §7] however it would be desirable to have a characterisation similar to the one available in the linear setting. For example, in the linear setting, we have the following equivalences:

Proposition 10.1. *Let $A \in \mathfrak{gl}_n(\mathcal{K}')$ with $\mathcal{K} \subset \mathcal{K}'$ a finite extension. The following are equivalent:*

(i) A is similar to an element of $\mathfrak{gl}_n(\mathcal{K})$.

(ii) The orbit $\mathrm{GL}_n(\mathcal{K}') \cdot A$ is fixed by $\mathrm{Gal}(\mathcal{K}'/\mathcal{K})$.

Babbitt and Varadarajan have given a proof of this result using their reduction theory (see [BV83]). It would be interesting to obtain an analogue of Proposition 10.1 in the differential setting using the same proof techniques as in the linear setting.

A second direction of interest is to extend the work here to the setting of p-adic differential operators. In this setting, we replace the field $\mathbb{C}(\!(t)\!)$ by $\mathbb{Q}_p(\!(t)\!)$. There are many subtleties that then arise due to the positive characteristic, however one can still define notions such as a Newton polygon in this setting (c.f. [Ked10a, §8.2]). There are also decomposition theorems for p-adic differential modules similar to the Hukuhara-Levelt-Turrittin decomposition. We hope that the work presented here might lead to a more unified approach to proving these decomposition theorems.

References

[Axl15] S. Axler, *Linear Algebra Done Right*, 3rd ed., Undergraduate Texts in Mathematics, Springer, 2015.

[BV83] D. G. Babbitt and V. S. Varadarajan, *Formal reduction theory of meromorphic differential equations: a group theoretic view*, Pacific J. Math. **109** (1983), no. 1, 1–80.

[BBDE05] A. Beilinson, S. Bloch, P. Deligne, and H. Esnault, *Periods for irregular connections on curves* (2005).

[Bor91] A. Borel, *Linear Algebraic Groups*, 2nd ed., Graduate Texts in Mathematics, vol. 126, Springer-Verlag, New York, 1991.

[Bou98] Nicolas Bourbaki, *Lie Groups and Lie Algebras. Chapters 1–3*, Elements of Mathematics (Berlin), Springer-Verlag, Berlin, 1998. Translated from the French; Reprint of the 1989 English translation. MR1728312

[BK86] Egbert Brieskorn and Horst Knörrer, *Plane Algebraic Curves*, Modern Birkhäuser Classics, Birkhäuser/Springer Basel AG, Basel, 1986. Translated from the German original by John Stillwell; [2012] reprint of the 1986 edition.

[DM82] Pierre Deligne and James S. Milne, *Hodge Cycles, Motives, and Shimura Varieties*, Lecture Notes in Mathematics, vol. 900, Springer-Verlag, Berlin-New York, 1982.

[DF04] D. S. Dummit and R. M. Foote, *Abstract Algebra*, 3rd ed., John Wiley & Sons, Inc., Hoboken, NJ, 2004.

[Fre07] Edward Frenkel, *Langlands Correspondence for Loop Groups*, Cambridge Studies in Advanced Mathematics, vol. 103, Cambridge University Press, Cambridge, 2007.

[FH91] W. Fulton and J. Harris, *Representation Theory*, Graduate Texts in Mathematics, vol. 129, Springer-Verlag, New York, 1991. A first course; Readings in Mathematics.

[GH94] Phillip Griffiths and Joseph Harris, *Principles of Algebraic Geometry*, Wiley Classics Library, John Wiley & Sons, Inc., New York, 1994. Reprint of the 1978 original.

[HP99] A. W. Hales and I. B. S. Passi, *Jordan decomposition*, Algebra, Trends Math., Birkhäuser, Basel, 1999, pp. 75–87.

[Huk41] M. Hukuhara, *Théorèmes fondamentaux de la théorie des équations différentielles ordinaires. II*, Mem. Fac. Sci. Kyūsyū Imp. Univ. A. **2** (1941), 1–25.

[Hum78] J. E. Humphreys, *Introduction to Lie Algebras and Representation Theory*, Graduate Texts in Mathematics, vol. 9, Springer-Verlag, New York-Berlin, 1978. Second printing, revised.

[KS16] M. Kamgarpour and D. Sage, *A geometric analogue of a conjecture of Gross and Reeder*, arXiv:1606.00943 (2016).

[Kat70] Nicholas M. Katz, *Nilpotent connections and the monodromy theorem: Applications of a result of Turrittin*, Inst. Hautes Études Sci. Publ. Math. **39** (1970), 175–232.

[Kat87] ———, *On the calculation of some differential Galois groups*, Invent. Math. **87** (1987), no. 1, 13–61.

[Ked10a] K. S. Kedlaya, *p-adic Differential Equations*, Cambridge Studies in Advanced Mathematics, vol. 125, Cambridge University Press, Cambridge, 2010.

[Ked10b] Kiran S. Kedlaya, *Good formal structures for flat meromorphic connections, I: surfaces*, Duke Math. J. **154** (2010), no. 2, 343–418.

[Kum02] Shrawan Kumar, *Kac-Moody Groups, Their Flag Varieties and Representation Theory*, Progress in Mathematics, vol. 204, Birkhäuser Boston, Inc., Boston, MA, 2002.

[Lev75] A. H. M. Levelt, *Jordan decomposition for a class of singular differential operators*, Ark. Mat. **13** (1975), 1–27.

[Luu15] M. Luu, *Local Langlands duality and a duality of conformal field theories*, arXiv:1506.00663 (2015).

[Mal74] B. Malgrange, *Sur les points singuliers des équations différentielles*, Enseignement Math. (2) **20** (1974), 147–176.

[Mal79] ———, *Sur la réduction formelle des équations différentielles à singularités irrégulières* (1979).

[Neu99] J. Neukirch, *Algebraic Number Theory*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 322, Springer-Verlag, Berlin, 1999. Translated from the 1992 German original and with a note by Norbert Schappacher; With a foreword by G. Harder.

[Ore33] O. Ore, *Theory of non-commutative polynomials*, Ann. of Math. (2) **34** (1933), no. 3, 480–508.

[Pra83] C. Praagman, *The formal classification of linear difference operators*, Indagationes Mathematicae (Proceedings) **86** (1983), no. 2, 249–261.

[Ras15] S. Raskin, *On the notion of spectral decomposition in local geometric Langlands*, arXiv:1511.01378 (2015).

[Rob80] 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.

[Rot09] Joseph J. Rotman, *An Introduction to Homological Algebra*, 2nd ed., Universitext, Springer, New York, 2009.

[vdPS03] M. van der Put and M. F. Singer, *Galois Theory of Linear Differential Equations*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 328, Springer-Verlag, Berlin, 2003.

[Tur55] 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.

[Var96] V. S. Varadarajan, *Linear meromorphic differential equations: a modern point of view*, Bull. Amer. Math. Soc. (N.S.) **33** (1996), no. 1, 1–42.

[Was65] W. Wasow, *Asymptotic Expansions for Ordinary Differential Equations*, Pure and Applied Mathematics, Vol. XIV, Interscience Publishers John Wiley & Sons, Inc., New York-London-Sydney, 1965.