NOTES ON DIFFERENTIAL ALGEBRA

REID DALE

Abstract.

Contents

Date: July, 2016.

This work was supported by an NSF Graduate Research Fellowship.

The author has benefited tremendously from the patient guidance of Tom Scanlon and Silvain Rideau.

1. Introduction

2. Basic Differential Algebra

2.1. **Derivations and Dual Numbers.** The fundamental object that we will be working with throughout this section is that of a (commutative) *differential* ring (R, ∂), which is a ring R equipped with a *derivation*

$$\partial : R \to R,$$

which is a function satisfying

$$\partial(x+y) = \partial(x) + \partial(y)$$
$$\partial(xy) = x\partial(y) + y\partial(x).$$

While this ∂ is not a ring homomorphism, it is possible to view a derivation as a *component* of a certain homomorphism from R to the dual numbers $R[\epsilon]/(\epsilon^2)$ as follows:

Proposition 2.1. *There is a bijective correspondence between derivations ∂ on R and sections $s : R \to R[\epsilon]/(\epsilon^2)$ of the canonical projection $\pi : R[\epsilon]/(\epsilon^2) \to R$.*

Proof. We first define a map from derivations to sections $\partial \mapsto s_\partial$ given by

$$s_\partial(x) = x + \epsilon\partial(x).$$

The function s_∂ is a ring homomorphism as it is certainly additive, and it is multiplicative as

$$\begin{aligned} s_\partial(x)s_\partial(y) = (x+\epsilon\partial(x))(y+\epsilon\partial(y)) &= xy + \epsilon(x\partial(y) + y\partial(x)) + \epsilon^2(\partial(x)\partial(y)) \\ &= xy + \epsilon\partial(xy) = s_\partial(xy). \end{aligned}$$

Conversely, we define a map $s \mapsto \partial_s$ given by mapping a section $s = x + \epsilon f(x)$, where $f : R \to R$ is a function, to $\partial_s = f$. Note that the f described here is a well-defined function as $\{1, \epsilon\}$ is a basis for $R[\epsilon]/(\epsilon^2)$ as an R-module. This map is a derivation as the fact that s is a section forces

$$(x+\epsilon f(x))(y+\epsilon f(y)) = xy + \epsilon(xf(y) + yf(x)) = s(xy) = xy + \epsilon f(xy)$$

and so $f(xy) = xf(y) + yf(x)$. Additivity is clear, so that ∂_s is indeed a derivation on R.

Finally, the maps $\partial \mapsto s_\partial$ and $s \mapsto \partial_s$ are inverse to each other and so these sets are in bijective correspondence. $\square$

This identification of derivations with certain sections has the advantage of simplifying a lot of arguments about the existence and uniqueness of certain extensions of derivations; instead of doing explicit computations once can often exploit functoriality to cook up the desired extension since *functors always preserve sections.*

Proposition 2.2. *Let (R, ∂) be a differential ring with and $S \subset R$ a multiplicatively closed set. Then there exists an extension of ∂ to $S^{-1}R$.*

Proof. Working in the category of R-algebras, consider $R[\epsilon]/(\epsilon^2)$ as an R-algebra by equipping it with the structure map $s_\partial : R \to R[\epsilon]/(\epsilon^2)$. This is a section of the natural projection $R[\epsilon]/(\epsilon^2) \to R$ in the category of R-algebras. Moreover, by composing with the natural homomorphism $R[\epsilon]/(\epsilon^2) \to (S^{-1}R)[\epsilon]/(\epsilon^2)$ we get a map

$$s_\partial : R \to (S^{-1}R)[\epsilon]/(\epsilon^2).$$

By the universal property of the localization, if s_∂ maps S to the units of $(S^{-1}R)[\epsilon]/(\epsilon^2)$ then s_∂ extends to a unique map on $S^{-1}R$. We check that for all $x \in S$, $s_\partial(x)$ is a unit. Indeed

$$(s_\partial(x)) \cdot \frac{x - \epsilon\partial x}{x^2} = \frac{(x + \epsilon\partial(x))(x - \epsilon\partial(x))}{x^2} = \frac{x^2}{x^2} = 1$$

and so the map extends. It is clear that the homomorphism is a section of the canonical projection since $s_\partial(\frac{1}{x}) = \frac{x - \epsilon\partial(x)}{x^2} = \frac{1}{x} + \epsilon\frac{\partial(x)}{x^2}$. □

Another useful application of the section point of view is in constructing the ring of differential polynomials over R.

Proposition 2.3. *Let (R, ∂) be a differential ring. Then the ring of differential polynomials*

$$R\{x\} := R[x, x', x^{(2)}, \cdots]$$

is a differential ring with $\tilde{\partial} : R\{x\} \to R\{x\}$ given by $\tilde{\partial}|_R := \partial$, $\tilde{\partial}(x^{(n)}) = x^{(n+1)}$, and extended to the whole domain by Leibniz's rule.

Proof. To construct a derivation on $R\{x\}$ it suffices to construct a section

$$s : R\{x\} \to R\{x\}[\epsilon]/(\epsilon^2).$$

Since, as a ring, $R\{x\}$ is the free commutative R-algebra on the elements $\{x^{(n)}\}_{n\in\omega}$, to define a map $s : R\{x\} \to R\{x\}[\epsilon]/(\epsilon^2)$ it suffices to specify a homomorphism $s_0 : R \to R\{x\}[\epsilon]/(\epsilon^2)$ as well as the elements $s(x^{(n)})$ for all $n \in \omega$. We set

- $s_0 : R \to R\{x\}[\epsilon]/(\epsilon^2)$ by $s_0(r) = r + \epsilon\partial(r)$
- $s(x^{(n)}) = x^{(n)} + \epsilon x^{(n+1)}$

Then s is a section of the dual numbers for $R\{x\}$ and $\partial_s = \tilde{\partial}$, the desired derivation on $R\{x\}$. □

Finally, one can show that derivations extend uniquely to separable field extensions in this manner using more universal properties in commutative algebra.

Proposition 2.4. *Let (K, ∂) be a differential field and let L/K be a separable algebraic extension. Then ∂ extends uniquely to L.*

Proof. We first show the result for a finite separable extension L/K and then conclude by noting that if L/K is separable then L is the union of an ascending chain of separable extensions of K. As L is finite separable we may write $L = K[x]/(f(x)) = K(a)$ with $f(x) = \sum b_i x^i$ the minimal polynomial of a such that the *formal derivative* $f'(x)$ given by $f'(x) = \sum i b_i x^{i-1}$ satisfies $f'(a) \neq 0$. We first show *uniqueness* of the extension of ∂ to L, assuming that it exists.

If $\tilde{\partial}$ is an extension of ∂ to L then

$$\tilde{\partial}(\sum c_j a^j) = \sum \left(\partial(c_j)a^j + \tilde{\partial}(a) j c_j a^{j-1}\right)$$

so that knowing $\tilde{\partial}(a)$ and ∂ determines $\tilde{\partial}$ uniquely. Thus, to show that ∂ extends to *at most* one derivation on L we show that the value of $\partial(a)$ is uniquely determined. First we compute what ∂ *should* be. If $L = K(a)$ then $f(a) = 0$ with $f(x) = \sum b_i x^i$, so that

$$\partial[f(a)] = 0.$$

Expanding this expression gives us

$$\partial\left(\sum b_i x^i\right) = \sum \partial(b_i)x^i + i b_i x^{i-1}\partial(x) = \left(\sum \partial(b_i)x^i\right) + \partial(x)f'(x).$$

The polynomial $\left(\sum \partial(b_i)x^i\right)$ is just the polynomial obtained by applying ∂ to the coefficients of f and we set

$$f^{\partial}(x) = \left(\sum \partial(b_i)x^i\right).$$

Applying the above formula to a,

$$0 = \partial[f(a)] = f^{\partial}(a) + \partial(a)f'(a)$$

so that

$$\partial(a) = -\frac{f^{\partial(a)}}{f'(a)}$$

showing that ∂ extends in at most one way to a derivation on L.

To show the *existence* of the derivation extending ∂ to L we use the characterization of derivations as certain sections of the dual numbers. Since ∂ is a derivation, we obtain a canonical section $s_{\partial} : K \to K[\epsilon]/(\epsilon^2)$. We may compose s_{∂} with the natural injection $K[\epsilon]/(\epsilon^2) \to L[\epsilon]/(\epsilon^2)$ to get a map we abusively name $s_{\partial} : K \to L[\epsilon]/(\epsilon^2)$. Our goal is to extend this to a section $s_{\tilde{\partial}} : L \to L[\epsilon]/(\epsilon^2)$. Towards this, we first extend s_{∂} to $K[x]$ via the map

$$x \mapsto a - \epsilon\frac{f^{\partial(a)}}{f'(a)}$$

which, as we showed above, is the only possible option for $\tilde{\partial}(a)$. To show that this map descends to L we need to show that *in the dual numbers* $L[\epsilon]/(\epsilon^2)$,

$$f\left(a - \epsilon\frac{f^{\partial(a)}}{f'(a)}\right) = 0.$$

To evaluate polynomials in the dual numbers we use the formula

$$f(c + \epsilon d) = \sum (b_i + \epsilon\partial(b_i))(c + \epsilon d)^i$$

since we are thinking of $K[\epsilon]/(\epsilon^2)$ as having the K-algebra structure given by s_{∂}. Now note that

$$(c + \epsilon d)^i = \sum \binom{i}{j} c^j(\epsilon d)^{i-j}$$

which is 0 for all $i - j \geqslant 2$ as $(\epsilon d)^2 = 0$. Thus $(c + \epsilon d)^i = c^i + i\epsilon c^{i-1}d$ so that

$$f(c + \epsilon d) = \sum (b_i + \epsilon\partial(b_i))(c^i + i\epsilon c^{i-1}d).$$

Now we need to check that $f\left(a - \epsilon\frac{f^{\partial(a)}}{f'(a)}\right) = 0$. Expanding, we have that

$$\begin{aligned} f\left(a - \epsilon\frac{f^{\partial(a)}}{f'(a)}\right) &= \sum (b_i + \epsilon\partial(b_i))\left(a^i - i\epsilon a^{i-1}\frac{f^{\partial(a)}}{f'(a)}\right) \\ &= \sum\left(b_i a^i + \epsilon\partial(b_i)a^i - ib_i\epsilon a^{i-1}\frac{f^{\partial(a)}}{f'(a)} + \epsilon^2(i\partial(b_i)a^{i-1}\frac{f^{\partial(a)}}{f'(a)}\right) \end{aligned} \tag{1}$$

□

2.2. **Differential Ideals and Ritt Noetherianity.** As in usual ring theory, the notion of a *differential* ideal plays a fundamental role in differential algebra and differential algebraic geometry. The motivation is nearly identical as in the case of algebraic geometry: if we know that a differential equation $f = 0$ holds, then differentiating both sides yields that $\partial(f) = 0$ as well.

Throughout these notes we adopt the convention that if $X \subseteq R$, then the ideal generated by X is denoted $\langle X \rangle$.

Definition 2.5. Let (R, ∂) be a differential ring. An ideal $I \subset R$ is a *differential* ideal just in case $\partial(I) \subseteq I$.

Given a family $\mathcal{F} = \{r_\alpha\}_{\alpha<\lambda} \subseteq R$ of elements of R, we can consider the differential ideal $J(\mathcal{F})$ generated by $\mathcal{F}$, which is the smallest differential ideal containing $\mathcal{F}$. It has a straightforward explicit presentation:

Proposition 2.6. *Let (R, ∂) be a differential ring and $\mathcal{F} \subseteq R$ be a family of elements. If we enumerate $\mathcal{F} = \{r_\alpha\}_{\alpha<\lambda} \subseteq R$, then*

$$J(\mathcal{F}) = \left\langle \{\partial^i(r_\alpha)\}_{\alpha<\lambda; i<\omega} \right\rangle,$$

i.e. $J(\mathcal{F})$ is generated as an ideal by $\mathcal{F}$ and all of the higher derivatives of its elements.

Proof. Write $\mathcal{F} = \{r_\alpha\}_{\alpha<\lambda} \subseteq R$. As $\mathcal{F} \subseteq J(\mathcal{F})$ and $J(\mathcal{F})$ is a differential ideal, we immediately have that

$$\left\langle \{\partial^i(r_\alpha)\}_{\alpha<\lambda; i<\omega} \right\rangle \subseteq J(\mathcal{F})$$

To show that

$$J(\mathcal{F}) \subseteq \left\langle \{\partial^i(r_\alpha)\}_{\alpha<\lambda; i<\omega} \right\rangle$$

we need only show that $\left\langle \{\partial^i(r_\alpha)\}_{\alpha<\lambda; i<\omega} \right\rangle$ is a differential ideal since $J(\mathcal{F})$ is minimal amongst all differential ideals. Suppose that

$$g \in \left\langle \{\partial^i(r_\alpha)\}_{\alpha<\lambda; i<\omega} \right\rangle$$

so that $g = \sum g_{\alpha,i} \partial^i(r_\alpha)$. Then

$$\partial(g) = \sum \left(\partial(g_{\alpha,i}) \partial^i(r_\alpha) + g_{\alpha,i} \partial^{i+1}(r_\alpha) \right) \in \left\langle \{\partial^i(r_\alpha)\}_{\alpha<\lambda; i<\omega} \right\rangle$$

and so $\left\langle \{\partial^i(r_\alpha)\}_{\alpha<\lambda; i<\omega} \right\rangle$ is a differential idea. Therefore $J(\mathcal{F}) = \left\langle \{\partial^i(r_\alpha)\}_{\alpha<\lambda; i<\omega} \right\rangle$, as desired. □

A useful geometric fact in algebraic geometry is that the Zariski topology is a noetherian topology, which can be seen as a corollary of the Hilbert Basis Theorem. In the context of *differential* algebraic geometry, the Kolchin topology is *also* noetherian, but the straightforward analogue of the Hilbert Basis theorem is false: there exist strictly increasing ascending chains of differential ideals in differential polynomial rings. However, in the case of *radical* differential ideals, the Ritt-Raudenbush theorem tells us that all *radical* differential ideals in a differential polynomial ring over a field K containing $\mathbb{Q}$ are finitely differentially-radically generated.

Example 2.7. We give an example of an ascending chain of differential ideals that does not terminate. Consider the chain of ideals $I_n \subseteq \mathbb{Q}\{x\}$ given by

$$I_n = J(x^2, (x')^2 \cdots, (x^{(n-1)})^2)$$

as well as the ideal $I = J(\{(x^{(i)})^2\}_{i\in\omega})$

Clearly whenever $i \leqslant j$ we have $I_i \subseteq I_j$, so our goal is to show that for all $n \in \omega$, $I_n \subsetneq I_{n+1}$ to yield a strictly ascending chain. To do this, we will often work with the auxiliary ideal I defined above, because it admits a set of especially nice looking generators.

Claim 2.8. $I = \left\langle x^{(i)}x^{(j)} \,|\, i \leqslant j \right\rangle$

To show that $\left\langle x^{(i)}x^{(j)} \,|\, i \leqslant j \right\rangle \subseteq I$, we go by induction on $k = j - i$. The case $k = 0$ is immediate by the definition of I, and the $k = 1$ case follows since $\partial((x^{(i)})^2) = 2x^{(i)}x^{(i+1)}$ and since in $\mathbb{Q}$ we can divide by 2.

Suppose now that all $x^{(i)}x^{(j)} \in I$ for (i,j) with $j - i \leqslant k$ for $k \geqslant 1$; we wish to show that $x^{(i)}x^{(j)} \in I$ for all (i,j) with $j - i = k + 1$. Let (i,j) be such that $j - i = k + 1$. Then $(i,j) = (i_0, j_0 + 1)$ with $j_0 - i_0 = k$. Then $\partial(x^{(i_0)}x^{(j_0)}) = x^{(i_0+1)}x^{(j_0)} + x^{(i_0)}x^{(j_0+1)}$. Since $j_0 - (i_0 + 1) = k - 1$ we have that $x^{(i_0+1)}x^{(j_0)} \in I$ by induction hypothesis and so $x^{(i_0)}x^{(j_0+1)} = x^{(i)}x^{(j)} \in I$, as desired.

To show that $I \subseteq \left\langle x^{(i)}x^{(j)} \,|\, i \leqslant j \right\rangle$ it suffices by 2.6 to show that the higher derivatives of its differential generators $(x^{(i)})^2$ are all expressible as sums of products of elements of $\mathbb{Q}\{x\}$ with elements of the form $x^{(i)}x^{(j)}$, which is a very straightforward computation (in fact, this computation is nested in the induction step of the $\left\langle x^{(i)}x^{(j)} \,|\, i \leqslant j \right\rangle \subseteq I$ direction of the argument). **I think that this direction is the only one needed for the argument, actually...**

Now that we understand the ideal I we proceed with the argument. We wish to show that $I_n \subsetneq I_{n+1}$ by arguing that $(x^{(n)})^2 \notin I_n$. We make two slight simplifications:

- We may work in $\mathbb{Q}\{x\}/J(x^{(n+1)}) \cong \mathbb{Q}[x, x', \cdots, x^n]$ since if we can show that $(x^{(n)})^2 \bmod J(x^{(n+1)}) \notin I_n \bmod J(x^{(n+1)})$ then $(x^{(n)})^2 \notin I_n$
- We instead show that

$$(x^{(n)})^2 \bmod J(x^{(n+1)}) \notin \left\langle x^{(i)}x^{(j)} \,|\, 0 \leqslant i < j \leqslant n \text{ or } 0 \leqslant i = j \leqslant n-1 \right\rangle \bmod J(x^{(n+1)})$$

which suffices since

$$I_n \bmod J(x^{(n+1)}) \subseteq \left\langle x^{(i)}x^{(j)} \,|\, 0 \leqslant i < j \leqslant n \right\rangle$$

by one direction of the containment argument for I.

Now suppose that inside $\mathbb{Q}[x, x', \cdots, x^{(n)}]$ we have that

$$(x^{(n)})^2 \in \left\langle x^{(i)}x^{(j)} \,|\, 0 \leqslant i < j \leqslant n \text{ or } 0 \leqslant i = j \leqslant n-1 \right\rangle.$$

Then writing out a witnessing expression to $(x^{(n)})^2$ being in the above ideal we find that

$$(x^{(n)})^2 = \sum_{\ell} \sum_{0 \leqslant i \leqslant j \leqslant n-1} f_{i,j,\ell} \cdot (x^{(n)})^{\ell} x^{(i)}x^{(j)} + \sum_{m} \sum_{1 \leqslant k \leqslant n-1} g_{k,m} \cdot (x^{(n)})^{m+1} x^{(k)}$$

with $f_{i,j,\ell}$ and $g_{k,m}$ inside $\mathbb{Q}[x, x', \cdots, x^{(n)}]$. But then since it's a polynomial in the variable $x^{(n)}$ we can simplify this to

$$(x^{(n)})^2 = \sum_{0 \leqslant i \leqslant j \leqslant n-1} \tilde{f}_{i,j} \cdot (x^{(n)})^2 x^{(i)}x^{(j)} + \sum_{1 \leqslant k \leqslant n-1} \tilde{g}_k \cdot (x^{(n)})^2 x^{(k)}.$$

But then

$$1 = \sum_{0 \leqslant i \leqslant j \leqslant n-1} \tilde{f}_{i,j} x^{(i)}x^{(j)} + \sum_{1 \leqslant k \leqslant n-1} \tilde{g}_k x^{(k)}$$

which is impossible as $1 \notin \left\langle x^{(i)} \,|\, 0 \leqslant i \leqslant n \right\rangle \subseteq K[x, x', \cdots, x^{(n)}]$.

Therefore $I_n \subsetneq I_{n+1}$ and so we get a strictly increasing chain of differential ideals

$$I_1 \subsetneq I_2 \subsetneq \cdots \subsetneq I_n \subsetneq \cdots$$

□

One way to circumvent pathologies of this sort is to look at ideals of geometric significance from the Kolchin viewpoint: the *radical* differential ideals. The motivation for considering this class of ideals comes from the same motivation as in algebraic geometry: if (K, ∂) is a field of functions and $x \in K$ is a point, then if $(f^n)(x) = 0$ then $f(x) = 0$ as well, so that the ideal of differential polynomials vanishing on x is radical.

Definition 2.9. A differential ring (R, ∂) is called *Ritt-noetherian* provided every properly ascending chain of *radical* differential ideals is finite.

Our aim now is to prove the Ritt-Raudenbush theorem:

Theorem 2.10. *Let $R \supseteq \mathbb{Q}$ be a Ritt-noetherian differential ring. Then $R\{x\}$ is Ritt-noetherian.*

Before proving this, we first establish some basic properties of radical differential ideals and Ritt-noetherianity.

Proposition 2.11. *Let $\mathcal{F} \subset R$ be a family of elements with $R \supset \mathbb{Q}$. Then the minimal radical ideal containing $\mathcal{F}$, $\{\mathcal{F}\}$, can be characterized by the equation*

$$\{\mathcal{F}\} = \sqrt{J(\mathcal{F})}$$

i.e. the minimal radical differential ideal containing $\mathcal{F}$ is the radical of the minimal differential ideal containing $\mathcal{F}$.

Proof. Since intersections of radical differential ideals are radical differential, $\{\mathcal{F}\}$ exists (here allowing the possibility that $\{F\} = R$). It's immediate that $\sqrt{J(\mathcal{F})} \subseteq \{\mathcal{F}\}$, so it suffices to show that $\sqrt{J(\mathcal{F})}$ is itself a radical differential ideal. It's certainly a radical ideal, so we just check that $\partial(\sqrt{J(\mathcal{F})}) \subseteq \sqrt{J(\mathcal{F})}$.

Suppose that $a \in \sqrt{J(\mathcal{F})}$, so that $a^n \in J(\mathcal{F})$. We want to show that there is an m such that $\partial(a)^m \in J(F)$. We do this by differentiating a^n and seeing what we get. Since $a^n \in J(\mathcal{F})$, $\partial(a^n) \in J(\mathcal{F})$. Expanding we get

$$\partial(a^n) = \partial(a) n a^{n-1} \in J(\mathcal{F}).$$

Differentiating again, we see that

$$\partial^2(a^n) = \partial^2(a) n a^{n-1} + \partial(a)^2 n(n-1) a^{n-2} \in J(\mathcal{F})$$

While this looks ugly, multiplying by $\partial(a)$ on both sides yields

$$\partial^2(a) \left(\partial(a) n a^{n-1}\right) + \partial(a)^3 n(n-1) a^{n-2} \in J(\mathcal{F})$$

so that since $\partial^2(a) \left(\partial(a) n a^{n-1}\right) \in J(\mathcal{F})$ we can conclude that

$$\partial(a)^3 n(n-1) a^{n-2} \in J(\mathcal{F}).$$

Repeating this exact process repeatedly we can conclude that

$$\partial(a)^{2n-1} n! \in J(\mathcal{F}).$$

Since $R \supset \mathbb{Q}$ we can conclude that

$$\partial(a)^{2n-1} \in J(\mathcal{F})$$

so that $\partial(a) \in \sqrt{J(\mathcal{F})}$. $\square$

Ritt-noetherianity can, like usual noetherianity, be expressed in terms of a kind of finite generation of ideals:

Proposition 2.12. *(R, ∂) is Ritt-noetherian if and only if all radical differential ideals $I \subseteq R$ are finitely generated; that is, that there exists a finite set $I_0 \subseteq I$ such that $I = \{I_0\}$.*

Proof. Suppose that there is an infinite ascending chain of radical differential ideals $I_0 \subset I_1 \subset \cdots$. Let $I = \bigcup I_n$; this is a proper radical differential ideal since if $1 \in I$ then $1 \in I_m$ for some $m < \omega$. If I were finitely generated, then $I = \{f_1, \cdots, f_n\}$ for some finite collection of f_i's. But then there is m such that $f_1, \cdots, f_n \in I_m$, so that $I_m = \{f_1, \cdots, f_n\}$. But then the ideal chain stabilizes at m, a contradiction.

Conversely, suppose that there is a radical differential ideal I which is not finitely generated. We build a chain inductively as follows:

- Pick $r_0 \in I$ and set $I_0 = \{r_0\}$.
- Given $I_n = \{r_0, \cdots r_n\} \subseteq I$ finitely generated, select $r_{n+1} \in I \backslash I_n$, which exists as otherwise I would be finitely generated. Then let $I_{n+1} = \{r_0, \cdots, r_{n+1}\}$.

This yields a properly increasing chain of radical differential ideals. $\square$

Proposition 2.13. *Let $X, Y \subseteq R$ be two sets. Then*

$$\{X\}\{Y\} \subseteq \{xy \,|\, x \in X, y \in Y\}.$$

Proof. To prove this we prove some slightly more general lemmas.

Suppose that I is a radical differential ideal and $S \subset R$ is closed under multiplication (e.g. S is an ideal). Then I claim that $T_S = \{x \in R \,|\, xS \subset I\}$ is a radical differential ideal. Note that if $x \in T$ then $\partial(x) \in T$ since if $ab \in I$ then $\partial(a)b \in I$ by Leibniz' rule together with the differential radicality of I (the full argument is written up later in the proof of 2.40). Hence T is a differential ideal, and moreover if $x^n \in T$ then $x^n S \subset I$. Since S is multiplicatively closed, $x^n S^n \subset I$ so that since I is radical, $xS \subset I$ and x is in T.

Now, we prove the proposition in the case that X is a single element. If $X = \{a\}$ and Y is any set, then

$$a\{Y\} \subset \{aY\}$$

since the set

$$T_a = \{x \in R \,|\, xa^n \in \{aY\}\}$$

is a radical ideal containing $a\{Y\}$. Then if X is larger, this shows that

$$T_X = \{x \in R \,|\, x\{Y\} \subset \{Y\}\}$$

contains $\{X\}$, so that

$$\{X\}\{Y\} \subset \{xy \,|\, x \in X, y \in Y\}$$

$\square$

A crucial part of the usual proof of Hilbert's basis theorem is the division lemma for polynomial rings; we will rely on an analogue of it for differential rings to prove the Ritt-Raudenbush theorem. To state the division lemma we will need to define a convenient quantity associated to differential polynomials.

Definition 2.14. Let (R, ∂) be a differential ring and $f \in R\{x\} \backslash R$. The *order* of f, $\mathrm{ord}(f)$ is the largest n such that $x^{(n)}$ appears in f; if $f \in R$, then its order is -1. Given f of order n we can write

$$f = \sum_{i=0}^{d} g_i \cdot (x^{(n)})^i$$

with all $g_i \in R[x, x', \cdots, x^{(n-1)}]$ and $g_d \neq 0$. In this case we say that f has degree d.

We write $f \ll g$ in case $\mathrm{ord}(f) < \mathrm{ord}(g)$ or if $\mathrm{ord}(f) = \mathrm{ord}(g)$ and $\deg(f) < \deg(g)$.

Recall the usual division algorithm lemma for polynomial rings over fields:

Lemma 2.15. *Let $f, g \in K[x]$ be a polynomial. Then there exists a polynomial $\tilde{g} \in K[x]$ with $\deg(\tilde{g}) < \deg(f)$ such that*

$$g \equiv \tilde{g} \operatorname{mod} \langle f \rangle$$

While we cannot achieve a differential division algorithm as clean as this since, as a pure ring, $R\{x\}$ is the polynomial ring on countably many variables, the differential structure on $R\{x\}$ allows us to simplify differential polynomials.

Two crucial quantities associated to a differential polynomial f, the *initial* I_f and the *separant* s_f occur naturally in the course of devising the division algorithm. Let $\mathrm{ord}(f) = n$ and $\deg_{x^{(n)}}(f) = d$. The initial I_f is the leading coefficient of f considered as a polynomial in $(R[x, x^{(1)}, \ldots, x^{(n-1)}])[x^{(n)}]$. In other words, I_f is the unique element such that

$$f = f = I_f \cdot (x^{(n)})^d + \sum_{0 \leqslant i \leqslant d-1} h_i \cdot (x^{(n)})^i$$

with each $h_i \in R[x, x^{(1)}, \ldots, x^{(n-1)}]$. The *separant* is the *initial* of $\partial(f)$: $s_f = I_{\partial(f)}$. Its importance stems from the fact (to be proven shortly) that, in fact, $s_f = I_{\partial^k(f)}$ for all $k > 0$, which is a key observation for carrying out the division algorithm.

Lemma 2.16. *Let R be a commutative differential ring containing $\mathbb{Q}$, $f \in R\{x\}$ be of order $n > 0$ and degree d. Writing*

$$f = I_f \cdot (x^{(n)})^d + \sum_{0 \leqslant i \leqslant d-1} h_i \cdot (x^{(n)})^i$$

with initial I_f and coefficients h_i all inside $R[x, x', \ldots, x^{(n-1)}]$. Then for all $g \in R\{x\}$ there exist $\tilde{g} \in R\{x\}$ such that $\tilde{g} \ll f$ (in the order-degree ordering), an element $r \in R$, and integers ℓ and t with

$$r(I_f)^{\ell}(s_f)^t g \equiv \tilde{g} \operatorname{mod} J(f)$$

Proof. By induction we see that

$$f^{(k)} = s_f x^{(n+k)} + f_k$$

with $\mathrm{ord}(f_k) \leqslant n + k - 1$.

- $(k = 1)$ Writing $f = \sum_{i=0}^{\ell} h_i (x^{(n)})^i$ with $\mathrm{ord}(h_i) \leqslant n - 1$ we have that

$$f' = \sum_{i=0}^{\ell} \left(i h_i (x^{(n)})^{i-1} x^{(n+1)} + (h_i)'(x^{(n)})^i \right) = s_f x^{(n+1)} + f_1$$

with $\mathrm{ord}(f_1) \leqslant n = n+1-1$.

- $(k \geqslant 1)$ Suppose that $f^{(k)} = s_f x^{(n+k)} + f_k$ with $\mathrm{ord}(f_k) \leqslant n+k-1$. Then
$$f^{(k+1)} = s_f(x^{n+k+1}) + (s_f)' x^{(n+k)} + (f_k)'.$$
But then as $\mathrm{ord}(s_f) \leqslant n$ and $\mathrm{ord}(f_k) \leqslant n+k-1$ we have that $f_{k+1} := (s_f)' x^{(n+k)} + (f_k)'$ has order
$$\mathrm{ord}(f_{k+1}) \leqslant n+k-1+1 = n+k = n+(k+1)-1$$
as desired.

Now let $g \in R\{x\}$. If g has order $n+k$, writing $g = \sum_{i=0}^{d} v_i (x^{(n+k)})^i$ with $\mathrm{ord}(g_i) \leqslant n+k-1$ we have that $rs_f^d g - v_d f^{(k)} \ll g$. Iterating this process we can replace g by $\tilde{g} = \sum_{j=0}^{\tilde{d}} \tilde{v}_j (x^{(n)})^j$ equivalent mod $J(f)$ with $\mathrm{ord}(\tilde{g}) = q$. If $\tilde{d} := \deg(\tilde{g}) \geqslant \deg(f)$ then we may reduce the degree $\tilde{d}$ of $\tilde{g}$ by multiplying by some power of I_f and $r \in R$ to get $\deg(rI_f\tilde{g} - \tilde{v}_{\tilde{d}}(x^{(n)})^{\tilde{d}-\ell} f) < \tilde{d}$.[1] Iterating this process we can push the degree of $\tilde{g}$ below $\ell = \deg(f)$.

Thus, by collecting all of these steps, we only multiplied g by powers of I_f and s_f and so the result holds. □

noticing the appearance of both the *initial* I_f and the *separant* s_f in the differential division lemma. This adds a step of complication in our proof of the Ritt-Raudenbush theorem. We now prove the Ritt-Raudenbush theorem (following the proof from Marker's notes, but organized in a different way):

Proof. Suppose that (R, ∂) is a commutative Ritt-noetherian differential ring containing $\mathbb{Q}$. We wish to show that $R\{x\}$ is as well. By 2.12 this is equivalent to showing that every radical differential ideal $I \subseteq R\{x\}$ is finitely generated.

Step 1: Find a maximal counterexample. Suppose for contradiction that there is a radical differential ideal $I \subseteq R\{x\}$ which is not finitely generated. I claim that we can take I to be a *maximal* ideal amongst the family of radical, non-finitely-generated differential ideals by Zorn's lemma. Consider the family
$$\mathcal{I} = \{I \subseteq R\{x\} \mid I \text{ is a proper, radical, non-finitely-generated differential ideal}\}$$
ordered by inclusion. By assumption, $\mathcal{I}$ is nonempty so it suffices to show that every chain in $\mathcal{I}$ has an upper bound in $\mathcal{I}$. Let $\{I_\alpha\}_{\alpha<\lambda}$ be a chain of elements of $\mathcal{I}$. Their union $\tilde{I} = \bigcup_{\alpha<\lambda} I_\alpha$ is a radical differential ideal. It is a proper ideal since if $1 \in \tilde{I}$ then $1 = \sum r_i f_i$ with $r_i \in R$ and each f_i in *some* I_α. But as only finitely many f_i occur in this expression and since $\{I_\alpha\}_{\alpha<\lambda}$ forms an ascending chain, we must have that $1 \in I_{\alpha_0}$ for *some* $\alpha_0 < \lambda$. But then I_{α_0} is not a proper ideal, a contradiction. Similarly, $\tilde{I}$ is not finitely generated, for if $I = \{f_1, \cdots, f_n\}$ then there would exist an I_{α_0} with $f_1, \cdots, f_n \in I_{\alpha_0}$ and so as
$$\{f_1, \cdots, f_n\} \subseteq I_{\alpha_0} \subseteq \tilde{I} = \{f_1, \cdots, f_n\},$$
we would have that I_{α_0} is a finitely generated radical differential ideal.

Thus, by Zorn's lemma, we may assume that the radical, non-finitely-generated differential ideal I that we take is maximal amongst that family.

Step 2: Intersect I with R, find a minimal element f outside the radical-differential ideal generated by $I \cap R$ Since I is a radical differential

[1]Unlike the case $R = K$ a field, we may need to multiply by some $r \in R$, e.g. in the case $R = \mathbb{Z}$

ideal, so is $I \cap R$. Now, $I \cap R \subseteq R$ is finitely generated; say $I \cap R = \{r_1, \cdots r_m\} \subseteq R$. Set $I_0 = \sqrt{J(I \cap R)} = \{r_1, \cdots r_m\} \subseteq R\{x\}$; I_0 does not depend on choice of generators for $I \cap R$. Now, as $I \neq I_0$ we may pick $f \in I \backslash I_0$ of minimal order-degree.

Our goal is to reduce every element of I modulo f using the differential division lemma 2.16. [2].

Writing

$$f = I_f(x^{(n)})^d + \sum_{0 \leqslant i \leqslant d-1} h_i(x^{(n)})^i = I_f(x^{(n)})^d + f_0$$

we see that $I_f \notin I$ for if $I_f \in I$ then since $I_f \ll f$ we have that $I_f \in I_0$ and so $\sum_{0 \leqslant i \leqslant d-1} h_i(x^{(n)})^i \in I \backslash I_0$ with lower order-degree than that of f. Similarly, the separant $s_f \notin I$, for if $s_f \in I$ then $s_f \in I_0$ and so $f_0 = f - \frac{1}{d} s_f x^{(n)} \in I \backslash I_0$[3], again contradicting minimality. Since coefficients of the form $I_f^\ell s_f^k$ occur in 2.16, we wish to show that $I_f s_f \notin I$. One way to accomplish this is to show that I is in fact a *prime* ideal[4]. Therefore $\{I, I_f s_f\} \supset I$ is a radical differential ideal properly containing I, so that $\{I, I_f s_f\} = \{g_1, \cdots, g_\ell, I_f s_f\}$ with each $g_i \in I$.

Step 3: Divide modulo $J(f)$ and apply radicality. The way we intend to use the radicality of I is to use the following immediate fact: if I_0 is a radical ideal and $I_1^k \subseteq I_0$ for some k, then $I_1 \subseteq I_0$. Given this, our goal is to contain some *power* of our maximal counterexample I inside a finitely generated radical ideal and then show that they are, in fact, equal.

To construct a candidate finitely generated ideal we first reduce every element of I modulo $J(f)$ using 2.16. Pick $g \in I$. There is some $\tilde{g} \in R\{x\}$ with

$$r(I_f)^k (s_f)^m g \equiv \tilde{g} \bmod J(f)$$

with $\tilde{g} \ll f$. But since $g, f \in I$, $\tilde{g} \in I$ and hence in I_0. Thus

$$(I_f)^k (s_f)^m g \in \{I_0, f\};$$

by multiplying $(I_f)^k (s_f)^m g$ by $(I_f)^{t-k} (s_f)^{t-m} g^{t-1}$ for $t = \max\{k, m, 1\}$ and applying the radicality of I we have that

$$I_f s_f g \in J(I_0, f).$$

Since g was arbitrary we have that

$$(I_f s_f) I \subseteq \{I_0, f\}$$

But then

$$I^2 \subseteq I\{I, I_f s_f\} \subseteq I\{g_1, \cdots, g_\ell, I_f s_f\} \subseteq \{I \cdot g_1, I \cdot g_2, \cdots, I \cdot g_\ell, I_f s_f I\} \subseteq \{I_0, f, g_1, \cdots, g_\ell\} \subseteq I.$$

But then if $g \in I$ then $g^2 \in I^2$ and hence in the finitely-generated radical differential ideal $\{I_0, g_1, \cdots, g_\ell\}$, so that $g \in \{I_0, g_1, \cdots, g_\ell\}$. Thus $I = \{I_0, g_1, \cdots, g_\ell\}$ and is finitely generated, contradicting our original assumption, giving us that I is finitely generated to start with.

Appendix: the primality of I We claim that I is a prime differential ideal. Suppose that I is not prime, so that there is some $a, b \in R$ with $ab \in I$ but $a, b \notin I$. Consider the radical differential ideals $\{I, a\}$ and $\{I, b\}$; (note that here we are *not* assuming on the outset that these ideals are necessarily properly contained inside

[2]Marker does not note this, but his argument uses a version of 2.16 that assumes that f is irreducible. Because I is prime, one may however reduce to this case.

[3]Here using that $R \supseteq \mathbb{Q}$

[4]We prove the primality of I at the very end of the argument

$R\{x\}$). These ideals properly contain I and so they are finitely generated[5]. But then we may write

$$\{I, a\} = \{f_1, \cdots, f_n\} \text{ and } \{I, b\} = \{g_1, \cdots, g_m\}$$

with all $f_i^{n_i} \in J(I, a)$ and $g_j^{n_j} \in J(I, b)$. In fact, we may write

$$\{I, a\} = \{\tilde{f}_1, \cdots, \tilde{f}_\ell, a\} \text{ and } \{I, b\} = \{\tilde{g}_1, \cdots, \tilde{g}_k, b\}$$

with all $\tilde{f}_i, \tilde{g}_j \in I$ by rewriting the original $(f_i)^{n_i}$'s in terms of finitely many elements $\tilde{f}_1, \cdots, \tilde{f}_\ell$ and a (and likewise for the $\tilde{g}_j$'s). But then by 2.13

$$\{I, a\}\{I, b\} \subseteq \{ab, \{\prod \tilde{g}_j \tilde{f}_i\}_{ij}, \{\tilde{f}_i a\}_i, \{\tilde{g}_j b\}_j\} \subseteq I$$

But then if $z \in I$, then $z^2 \in I \cap (\{I, a\}\{I, b\})$ and so as $\{ab, \{\prod \tilde{g}_j \tilde{f}_i\}_{ij}, \{\tilde{f}_i a\}_i, \{\tilde{g}_j b\}_j\}$ is radical, $z \in \{ab, \{\prod \tilde{g}_j \tilde{f}_i\}_{ij}, \{\tilde{f}_i a\}_i, \{\tilde{g}_j b\}_j\}$. Thus $I = \{ab, \{\prod \tilde{g}_j \tilde{f}_i\}_{ij}, \{\tilde{f}_i a\}_i, \{\tilde{g}_j b\}_j\}$ and so I would finitely generated, a contradiction. Thus I is prime. $\square$

2.3. **Characteristic Sets and the Partial Ritt-Raudenbush.** The framework for differential algebra that we've considered corresponds to the study of a certain class of *ordinary* differential equations, but can be extended to study algebraic properties of *partial* differential equations as well.

Definition 2.17. A *partial differential ring* (R, Δ) is a ring R equipped with a (finite) family $\Delta = \{\partial_1, \cdots, \partial_n\}$ of **commuting** R-derivations.[6]

The analogue of the Ritt-Raudenbush theorem is true in this setting, although the proof is more involved: to perform an analogue of the reduction step of the ordinary case of Ritt-Raudenbush, we must consider not a *single* differential polynomial f of a specific type in the ideal I but rather a finite family $\mathcal{C}$ of Δ-polynomials called a *characteristic set*. To define and motivate characteristic sets, we need the concept of a *ranking* on the set of "Δ-variables" (**Better name for this?**) in the ring of Δ-polynomials $R_\Delta\{x\}$.

Definition 2.18. Let (R, Δ) be a partial differential ring with $\Delta = \{\partial_1, \cdots \partial_n\}$. The *ring of* Δ*-polynomials in* m *variables* $R_\Delta\{x_1, \cdots, x_m\}$ is the ring

$$R\left[\left(\partial_1^{\ell_1} \cdots \partial_n^{\ell_n} x_j\right)_{(\ell_1, \cdots, \ell_n) \in \omega^n,\, j \in \{1, \cdots, m\}}\right]$$

with each $k_i \in \{1, \ldots, n\}$ equipped with an extension of the elements of Δ from R to $R_\Delta\{x_1, \cdots, x_m\}$ given by setting

$$\partial_j(\partial_{k_\ell} \cdots \partial_{k_1} x_q) = \partial_j \partial_{k_\ell} \cdots \partial_{k_1} x_q.$$

Let $\mathcal{M}_\Delta = \{\theta x_i \mid \theta = \partial_{k_m} \cdots \partial_{k_1}\} \cup \{x_1, \cdots, x_m\}$ be the set of Δ-*variables.* A *ranking* $\prec$ on $\mathcal{M}_\Delta$ is a well-ordering satisfying two further conditions:

- For all $u, v \in \mathcal{M}_\Delta$ and $\theta = \partial_{k_m} \cdots \partial_{k_1}$, $u < v$ implies $\theta u < \theta v$.
- For all $u \in \mathcal{M}_\Delta$ and $\theta = \partial_{k_m} \cdots \partial_{k_1} \neq \mathrm{id}$, $u < \theta u$.

Remark 2.19.
- The proof that $R_\Delta\{x_1, \cdots, x_m\}$ gives us a well-defined Δ-ring is essentially the exact same as the argument that $R\{x\}$ is a ∂-ring.

[5]If either ideal were not a proper ideal then they would still be finitely generated as they are then the unit ideal.

[6]Recall that ∂_i and ∂_j commute provided $\partial_i(\partial_j r) = \partial_j(\partial_i r)$ for all $r \in R$.

- An ordering on $\mathcal{M}_\Delta$ is essentially an ordering on the variables of $R_\Delta\{x\}$, thought of as $R[\theta x_i]_{\theta x_i \in \mathcal{M}_\Delta}$ compatible with the application of derivations.

Fixing a ranking $\prec$ on $\mathcal{M}_\Delta$ we can define, given a Δ-polynomial f, the auxiliary Δ-polynomials of *initial* and *separant* as in the ordinary case.

Definition 2.20. Fix a ranking $\prec$ on $\mathcal{M}_\Delta$ and $f \in R_\Delta\{x_1, \cdots, x_m\}\backslash R$. The variable of highest $\prec$-rank, θx_i, is called the *leader* u_f of f.

Writing

$$f = g_d(u_f)^d + \cdots + g_1 u_f + g_0$$

with each $g_i \in R[\theta x_i \,|\, \theta x_i \prec u_f]$ and $g_d \neq 0$, we call g_d the *initial* I_f of f.

Let $\partial \in \Delta$ and $f \in R_\Delta\{x_1, \cdots, x_m\}\backslash R$. Then

$$\partial f = \partial(\sum g_i(u_f)^i) = \sum \partial(g_i(u_f)^i) = \sum \big(\partial(g_i)(u_f)^i + i g_i(u_f)^{i-1}\partial(u_f)\big)$$

By inspection, $\partial(u_f)$ is the leader of $\partial(f)$ by combining the two compatibility conditions necessary of $\prec$ and the fact that u_f is the leader of f. But then we may write

$$\partial f = s_f \partial(u_f) + \tilde{g}$$

where $\tilde{g} \in R[\theta x_i \,|\, \theta x_i \prec \partial(u_f)]$. The coefficient of ∂_{u_f} is the *separant* of f and, by the above computation, is independent of choice of ∂ and is equal to

$$s_f = \sum_{i=1}^{d} i g_i(u_f)^{i-1}$$

The notions of ranking and of leaders of differential polynomials give us a way to measure the *complexity* of a differential polynomial, allowing us to perform reduction and division procedures in an algorithmic fashion.

Definition 2.21. Let $f, g \in R_\Delta\{x_1, \cdots, x_m\}\backslash R$. We say that f is *reduced* with respect to g provided that

- f is *partially reduced* with respect to g: no term in f contains an instance of a *proper* derivative of u_g.
- If $u_f = u_g =: u$, then $\deg_u(f) < \deg_u(g)$.

Using this notion of reduction we can compare Δ-polynomials: we define for differential polynomials f and g the relation $f < g$ just in case $u_f \prec u_g$ or $u_f = u_g$ and $\deg_{u_f}(f) < \deg_{u_g}(g)$. Write $f \sim g$ if $f \leqslant g$ and $g \leqslant f$, i.e. if f and g have the same leading term and degree.

Remark 2.22. The definition of reduction makes no essential use or mention of the underlying arithmetic of the coefficient ring R. Because of this, when we prove the partial case of Ritt-Raudenbush we will have to simultaneously study the arithmetic of R in conjunction with the bare structure of differential polynomials.

Example 2.23. In $\mathbb{Q}\{x\}$ (equipped with the *unique* ranking $\prec$), x is partially reduced with respect to x' since x contains no proper derivatives of x', but x' is *not* partially reduced with respect to x.

Now we come to the key technical notion underlying the definition of characteristic sets, that of an *autoreduced set* of differential polynomials.

Definition 2.24. A subset $\mathcal{A} \subseteq R_\Delta\{x_1, \cdots, x_m\}$ is *autoreduced* provided that for all $f \neq g \in \mathcal{A}$, f is reduced with respect to g.

Example 2.25. If $f \in R_\Delta\{x_1, \cdots, x_m\}\backslash\{0\}$ then the singleton $\{f\}$ is autoreduced; thus there are always autoreduced sets.

In $K\{x, y, z\}$ for $\Delta = \{\partial\}$, with monomial ordering

$$x < y < z < \partial x < \partial y < \partial z < \cdots$$

the set $\{\partial^2 x, \partial^2 y - \partial x, \partial x \partial y \partial z\}$ is autoreduced.

A useful property of autoreduced sets is that they are necessarily finite:

Proposition 2.26. *Let $\mathcal{A} \subseteq R_\Delta\{x_1, \cdots, x_m\}$ be a set autoreduced with respect to some ranking $\prec$. Then $\mathcal{A}$ is finite.*

Proof. We first recast what it means for $\mathcal{A}$ to be autoreduced in terms of the leading monomials u_f. Since being autoreduced demands that for *any* ordered pair $f, g \in \mathcal{A}$ that f is reduced with respect to g, the condition that $u_f = u_g$ implies that $\deg_{u_f}(f) < \deg_{u_g}(g)$ cannot ever hold and so for all $f \neq g \in \mathcal{A}$, $u_f \neq u_g$. Thus, if we show that only finitely many leading terms u_f occur in $\mathcal{A}$ then we will have shown that $\mathcal{A}$ is finite. Moreover, the condition that no term of f contains an instance of some proper derivative u_g implies that u_f is not some proper derivative of u_g (in fact, u_f contains a proper derivative of u_g if and only if it *is* a proper derivative of u_g).

Moreover, if $\mathcal{A}$ is infinite then within $\mathcal{A}$ there is an infinite subset $\mathcal{A}_x \subseteq \mathcal{A}$ such that for all $f \in \mathcal{A}_x$, the variable of u_f is x, and certainly $\mathcal{A}_x$ is autoreduced.

With these observations in mind we can translate this problem to a combinatorial problem. Since each leading term u_f for $f \in \mathcal{A}_x$ can be rewritten uniquely as $\partial_1^{\ell_1} \cdots \partial_n^{\ell_n} x$, and the statement that $u_f = \partial_1^{\ell_1} \cdots \partial_n^{\ell_n} x$ is a proper derivative of $u_g = \partial_1^{k_1} \cdots \partial_n^{k_n} x$ is equivalent to saying that $(k_1, \cdots, k_n) < (\ell_1, \cdots, \ell_n)$ where $\leqslant$ is the partial order given by

$$(k_1, \cdots, k_n) < (\ell_1, \cdots, \ell_n) \iff k_i < \ell_i \text{ for } i \leqslant n.$$

If $\mathcal{A}_x$ were an infinite autoreduced set, then the set

$$\{(\ell_1, \cdots, \ell_n) \in \mathbb{N}^n \mid (\exists f \in \mathcal{A}_x)\, u_f = \partial_1^{\ell_1} \cdots \partial_n^{\ell_n} x\}$$

must form an infinite antichain with respect to the pointwise partial ordering $\leqslant$ on $\mathbb{N}^n$. Thus we reduce the problem to the following claim: the pointwise order on $\mathbb{N}^n$ has no infinite antichains. To do this it suffices to show that given an infinite $X \subseteq \mathbb{N}^n$ there is exist a comparable pair of tuples $(x_1, \ldots, x_n), (y_1, \cdots, y_n) \in X$. We prove this by induction on n:

- $(n = 1)$ $\mathbb{N}$ is a linear order, so this is automatically satisfied.
- $(n+1)$ Suppose that the result holds for $\mathbb{N}^n$ and that $X \subseteq \mathbb{N}^{n+1}$ is an infinite set. Consider the projection $\pi : X \to \mathbb{N}$ which maps $(x_1, \cdots, x_{n+1}) \mapsto x_{n+1}$. Then one of two things can happen: either $\pi(X)$ contains an element m with infinite fiber or every fiber $\pi^{-1}(m)$ is finite and $\pi(S)$ is infinite.

  Suppose that $\pi(X)$ contains an element m with infinite fiber, consisting of elements of the form $(x_1, \cdots, x_n, m)$. The set

  $$\{(x_1, \cdots, x_n) \mid (x_1, \cdots, x_n, m) \in \pi^{-1}(m)\} \subset \mathbb{N}^n$$

  is infinite and thus contains a comparable pair $(u_1, \cdots, u_n) \leqslant (v_1, \cdots, v_n)$. But then $(u_1, \cdots, u_n, m) \leqslant (v_1, \cdots, v_n, m)$ is a comparable pair in S, as desired.

Conversely, suppose that every element $m \in \mathbb{N}$ has $\pi^{-1}(m)$ finite. Then $\pi(X)$ is necessarily infinite and so there exists an ascending sequence $m_1 < m_2 < \cdots$. By the proof of the infinite fiber case above, we may assume that for all $j \leqslant n+1$ we have an infinite ascending chain in each coordinate: consider the coordinate projections $\rho_i : \mathbb{N}^{n+1} \to \mathbb{N}$. Then $\rho_1(X) \subseteq \mathbb{N}$ is an infinite linearly ordered set. Pick for each $m \in \rho_i(X)$ a tuple $\overline{x}_m$ with $\rho_1(\overline{x}_m) = m$ and set

$$X_1 = \{\overline{x}_m \mid m \in \rho_1(X)\}$$

Now, the image $\rho_2(X_1)$ is infinite by assumption as $X_1 \subseteq X$ and no fiber of any number under any projection is infinite by assumption. Constructing sets $X_2, \cdots, X_{n+1}$ in the following way yields at the end an infinite set X_{n+1} of comparable elements under $\leqslant$ since, by construction, at stage ℓ we ensured that the projection $X_\ell \to \mathbb{N}^\ell$ dropping the last $n+1-\ell$ coordinates consists only of comparable elements.

$\square$

Remark 2.27. In the analogous case of several *noncommuting* derivations, autoreduced sets need not be finite. For instance, in the case of the free noncommuting derivations $\Delta = \{\partial, \theta\}$ on $\mathbb{Q}_\Delta\{x\}$ we have that the set $\{\partial\theta^n x\}_{n\in\omega}$ is infinite and autoreduced. One way to account for this is the existence of infinite antichains in the tree $2^{<\omega} \cong \{\theta, \partial\}^{<\omega}$, where each sequence $s \in \{\partial, \theta\}^{<\omega}$ can be thought of as a term in $\mathcal{M}_\Delta$.

The proposition also fails in the case of admitting infinitely many variables: consider the ring $K\{(x_i)_{i\in\omega}\}$ with $\Delta = \{\partial\}$ and ordering the monomials in any way. Then the set $\mathcal{A} = \{(x_i)_{i\in\omega}\}$ is autoreduced but infinite.

Autoreduced sets give admit a convenient division algorithm similar to that of dividing by a single differential polynomial.

Proposition 2.28. *Let $\mathcal{A} = \{a_1 < \cdots, a_k\}$ be an autoreduced set with respect to an ordering $\prec$ and $f \in R_\Delta\{x_1, \cdots, x_n\}$. Then there exists an $\tilde{f} \in R_\Delta\{x_1, \cdots, x_n\}$*

- $r\left(\prod_{q=1}^{k} I_{a_q}^{\ell_q} s_{a_q}^{t_q}\right) f \equiv \tilde{f} \bmod J(\mathcal{A})$ *for some tuple $(\ell_1, t_1, \cdots, \ell_k, t_k) \in \mathbb{N}^{2k}$ and $r \in R$,*
- *$\tilde{f}$ reduced with respect to all elements of $\mathcal{A}$,*
- $\tilde{f} \leqslant f$

We call $\tilde{f}$ a remainder of f and say that f reduces to $\tilde{f}$ via $\mathcal{A}$, written $f \to_{\mathcal{A}} \tilde{f}$[7].

Proof. The process is very similar to that for dividing by a single differential polynomial. If f is already reduced with respect to all elements in $\mathcal{A}$, then taking $\tilde{f} = f$ satisfies the conclusions of the proposition.

Suppose that f is not reduced with respect to all elements in $\mathcal{A}$. If f is not reduced with respect to $\mathcal{A}$, we first ensure that f is *partially* reduced with respect to $\mathcal{A}$. We've ordered $\mathcal{A}$ by the differential polynomial ordering $<$. By applying the

[7] *a priori* f may reduce to many $\tilde{f}$ with respect to $\mathcal{A}$ depending on how one performed a reduction procedure. This is fine for us.

single differential polynomial division algorithm[8] to divide f by a_k we may find a differential polynomial f_k such that f_k is reduced with respect to a_k and

$$I_{a_j}^{\ell_j} s_{a_j}^{t_j} f \equiv f_k \operatorname{mod} J(\mathcal{A})^9$$

Note that if f is already reduced with respect to a_k then $f_k = f$. Repeating this process to f_k, then f_{k-1}, all the way through f_1 and noting that by construction f_j is reduced with respect to all a_m for $m \geqslant j$ by construction, after a finite number of steps we will find an $\tilde{f} := f_1$ which is reduced with respect to *all* elements of $\mathcal{A}$, obtained only by multiplying f by powers of I_a and s_a for $a \in \mathcal{A}$. □

We now define a partial ordering between autoreduced sets; the minimal elements in this partial ordering will be our *characteristic sets*.

Definition 2.29. Given two autoreduced sets $\mathcal{A} = \{a_1, \cdots, a_n\}$ and $\mathcal{B} = \{b_1, \cdots, b_m\}$ with $a_1 < \cdots < a_n$ and $b_1 < \cdots < b_n$[10] we write $\mathcal{A} < \mathcal{B}$ if either:

- There is some $1 \leqslant i \leqslant n$ so that for all $1 \leqslant j \leqslant i-1$, $a_j \sim b_j$ but $a_i < b_i$
- $m < n$ and for all $j \leqslant m$, $a_j \sim b_j$.

Remark 2.30. In this partial ordering, for all autoreduced sets $\mathcal{A}$ we have that $\mathcal{A} \prec \varnothing$.

As Tracey McGrail says in [7], "the order is lexicographic in nature, but 'humane' comes before 'human.'"

Corollary 2.31. *Let $\mathcal{A}_1 \succ \mathcal{A}_2 \succ \cdots$ be a descending chain of autoreduced sets. Then the chain eventually stabilizes; that is, there exists an n such that for all $m > n$, $\mathcal{A}_n = \mathcal{A}_m$.*

Proof. Suppose that the chain $\mathcal{A}_1 \succ \mathcal{A}_2 \succ \cdots$ does not stabilize. Then by the definition of $\prec$ we have a descending chain of

$$a_{1,1} > a_{2,1} > \cdots$$

where $a_{i,1} \in \mathcal{A}_i$ is the first element. But this is an infinite descending chain in $\mathbb{N}^{n+1}$ (counting degree), which is therefore finite and stabilizes: the $<$-class of $a_{1,j}$ is the same for all sufficiently large $j > 0$. Now, amongst those $\mathcal{A}_i$ with $a_{1,i}$ stabilized, we may compare the second elements $a_{2,i}$, which must stabilize as well. Proceeding in this fashion, we find that for all n we have a differential monomial $\tilde{u}_n$ in all $\mathcal{A}_\ell$ for all ℓ sufficiently large. Picking a representative a_j from some $\mathcal{A}_\ell$ of the stabilized class of the j^{th} elements, this yields for all n the autoreduced set $\{a_j\}_{j \leqslant n}$. But as being reduced is a property of pairs of pairs, we must have that $\{f_j\}_{j \in \omega}$ is an infinite autoreduced set. By 2.26, such sets do not exist and so $\mathcal{A}_1 \succ \mathcal{A}_2 \succ \cdots$ must stabilize. □

Definition 2.32. Let $I \subseteq R_\Delta\{x_1, \cdots, x_n\}$ be a Δ-ideal. A *characteristic set* for I is a $\prec$-minimal autoreduced subset of Δ-polynomials in I.

[8]We did not consider the multivariate case in our original account of it, but the proof works word-for-word for a choice of a differential ranking on the variables of $R\{x_1, \cdots, x_n\}$ in the ordinary case

[9]since $J(\{a_j\}) \subseteq J(\mathcal{A})$.

[10]note that, as argued above, the case $u_f = u_g$ with $f \neq g$ does not occur for autoreduced sets

Note that for *any* differential ideal I (or, really, any ideal) characteristic sets exist. The two crucial properties we need for proving the Ritt-Raudenbush theorem are their *finiteness* and their *minimality*; their minimality allows us to control the reduction process in much the same way that we were able to control the reduction process in the proof of the ordinary case of Ritt-Raudenbush. The main lemma is this:

Lemma 2.33. *Let $I \subseteq R_\Delta\{x_1, \cdots x_m\}$ be a differential ideal and $\mathcal{C}$ a characteristic set for I. If $f \in I$ is with respect to $\mathcal{C}$ then $f \in I \cap R$. Moreover, $s_c, I_c \notin I$ for all $c \in \mathcal{C}$.*

In the special case when $R = K$ is a field, if $f \in I$ is reduced with respect to $\mathcal{C}$ then $f \in I \cap K = \{0\}$.

Proof. Suppose that $f \in I$ with $f \notin I \cap R$ is reduced with respect to $\mathcal{C}$. Let

$$\mathcal{C}_f = \{c \in \mathcal{C} \mid c < f\} \cup \{f\}.$$

Then $\mathcal{C}_f$ is autoreduced and $\mathcal{C}_f \prec \mathcal{C}$ by definition of $\mathcal{C}_f$. But this violates the minimality of $\mathcal{C}$ amongst autoreduced subsets of I, and so f must have been in $I \cap R$.

Now suppose that either I_c or s_c were in I. If $I_c \in I$ for some $c \in \mathcal{C}$ then since $I_c \ll c$ we can form the new autoreduced set $\mathcal{C}' = \{g \in \mathcal{C} \mid g < c\} \cup \{c - I_c u_c\}$ where u_c is the leader of c. But then $\mathcal{C}' \prec \mathcal{C}$, contradicting minimality.

Likewise, if $s_c \in I$ then set

$$\mathcal{C}' = \{g \in \mathcal{C} \mid g < c\} \cup \{c - \frac{1}{d} s_c u_c\}$$

where $d = \deg_{u_c}(c)$ would also satisfy $\mathcal{C}' \prec \mathcal{C}$ strictly, again contradicting the assumption that $\mathcal{C}$ is a characteristic set. □

This reduction algorithm can actually be used to characterize which autoreduced sets are characteristic sets:

Proposition 2.34. *An autoreduced set $\mathcal{A}$ is a characteristic set for a differential ideal I if and only if for all $f \in I$, f reduces to some element in $I \cap R$ with respect to $\mathcal{A}$.*

Proof. The above argument yields the left-to-right direction.

For the other direction, suppose that $\mathcal{A}$ is autoreduced and for all $f \in I$, f reduces to some element of $I \cap R$ with respect to $\mathcal{A}$. Then no matter which $f \in I \backslash \mathcal{A}$, we have that

$$\mathcal{A} \prec \mathcal{A}_f = \{a \in \mathcal{A} \mid a < f\} \cup \{f\} = \{a_1, \ldots, a_n, f\}$$

since "humane comes before human:" as f is reduced with respect to $\mathcal{A}$ we have that $a_{n+1} \sim f$. □

Example 2.35. Consider the case of $\mathbb{Q}\{x, y\}$ with a single derivation and monomial ordering given by $x < y < \partial x < \partial y < \cdots$, and the ideal $I = J(\partial^2 x + y, \partial^2 y + x)$. Then I is a prime differential ideal since it is linear, and the set $\mathcal{A} = \{\partial^2 x + y, \partial^2 y + x\}$ is autoreduced. Since every element of I reduces to 0.

This process can be generalized: note that in the above example that $J(\mathcal{A}) = I$. If $I \subset K_\Delta\{x_1, \ldots, x_n\}$ is a differential ideal such that $I = J(f_1, \ldots, f_n)$ with each $f_1, \ldots, f_n$ having unit initial and separant then a characteristic set is just any autoreduced generating set for I by the above criterion. Since the theoretical interest in characteristic sets arises in *radical* differential ideals, which *may or may not* be finitely generated as differential ideals, the assumption that $I = J(f_1, \ldots, f_n)$ is a very special one to make. In the case of finitely generated linear differential ideals like the one above, which are always prime, characteristic sets can be found by repeatedly reducing the elements of a given generating set with respect to each other, which must terminate.

Theorem 2.36. *Suppose that* (R, Δ) *is a commutative Ritt-Noetherian ring containing* $\mathbb{Q}$*. Then so is* $R_\Delta\{x_1, \cdots, x_n\}$*.*

Proof. Suppose that (R, Δ) is a commutative Ritt-Noetherian ring containing $\mathbb{Q}$. We need to show that every radical differential ideal $I \subseteq R_\Delta\{x_1, \cdots, x_n\}$ is finitely generated.

Step 1: Find a maximal counterexample. We wish to show the existence of a radical differential ideal maximal amongst those that are not finitely generated. The proof is word-for-word the same as the one given in the proof of 2.10.

Step 2: Find and use a characteristic set for I. By 2.31 we may extract from I a characteristic set $\mathcal{C}$. Now we have that for all $c \in \mathcal{C}$, $I_c, s_c \notin I$ by 2.33. By primality of I, this implies that $\prod_{c\in\mathcal{C}} I_c s_c \notin I$. This product makes sense as $\mathcal{C}$, being autoreduced, is finite by 2.26. Thus $\{I, \prod_{c\in\mathcal{C}} I_c s_c\}$ is a finitely generated radical differential ideal, which we may write as

$$\{I, \prod_{c\in\mathcal{C}} I_c s_c\} = \{g_1, \cdots, g_m, \prod_{c\in\mathcal{C}} I_c s_c\}$$

Now let $f \in I$. By applying the division algorithm 2.28 to f we find an $\tilde{f}$ such that $\tilde{f}$ is reduced with respect to $\mathcal{C}$ and

$$r\left(\prod_{c\in\mathcal{C}} I_c^{\ell_c} s_c^{p_c}\right) f - \tilde{f} \in J(\mathcal{C}, I \cap R) \subseteq I$$

But then $\tilde{f} \in I$ and is reduced with respect to $\mathcal{C}$ and therefore by 2.33 $\tilde{f} \in I \cap R$[11]. Therefore

$$r\left(\prod_{c\in\mathcal{C}} I_c^{\ell_c} s_c^{p_c}\right) f \in J(\mathcal{C}, I \cap R) \subseteq I$$

and so by multiplying by appropriate powers of f, I_c, and s_c for $c \in \mathcal{C}$ we can conclude that

$$r\left(\prod_{c\in\mathcal{C}} I_c s_c\right) f \in \{C, I \cap R\} \subseteq I$$

Since $f \in I$ was arbitrary we have that

$$\{\left(\prod_{c\in\mathcal{C}} I_c s_c\right) I\} \subseteq \{C, I \cap R\}$$

so that

$$I^2 \subseteq I\{I, \prod_{c\in\mathcal{C}} I_c^{\ell_c} s_c^{p_c}\} \subseteq \{I \prod_{c\in\mathcal{C}} I_c^{\ell_c} s_c^{p_c}, I g_1, \cdots, I g_m\} \subseteq \{\mathcal{C}, I \cap R, g_1, \cdots, g_m\} \subseteq I.$$

[11] In proofs of the Ritt-Raudenbush theorem in the case of differential polynomial rings over fields, this step is usually omitted: in that case 2.33 yields that f is 0. Since we are not assuming that R is a field, we must take care to ensure that the arithmetic of R is accounted for in the proof.

Since $I \cap R \subseteq R$ is a radical differential ideal, it is finitely generated by the assumption of Ritt-noetherianity for R. But then by radicality of $\{\mathcal{C}, g_1, \cdots, g_m\}$, entails that $I = \{\mathcal{C}, r_1, \cdots, r_k, g_1, \cdots, g_m\}$ is finitely generated.

Appendix: The primality of I. Same proof as in the proof of 2.10. $\square$

2.4. **Basic Differential Algebraic Geometry: Properties of the Kolchin Topology.** To study the geometric properties of algebraic differential equations we define an analogue of the Zariski topology for algebraic varieties– called the *Kolchin topology*– that shares many of the same fundamental properties as the Zariski topology. Our focus will be primarily on fields equipped with many commuting derivations.

Definition 2.37. Let (K, Δ) be a differential field. A subset $X \subset K^n$ is said to be *Kolchin closed* provided X is the zero set of finitely many elements $K_\Delta\{x_1, \ldots, x_n\}$; namely,

$$X = Z(f_1, \ldots, f_m) = \left\{x \in K^n \mid \bigwedge_{1 \leqslant i \leqslant m} f_i(x) = 0\right\}$$

A priori this particular set of subsets of K^n may not form a topology since it only mentions zero sets of *finite* collections of differential polynomials.

Proposition 2.38. *The set of Kolchin closed subsets of K^n forms a topology on K^n.*

Proof. To verify that the Kolchin closed subsets of K^n form a topology, we need to check that $\varnothing$ and K^n are both Kolchin closed and that the family of closed sets is closed under finite unions and arbitrary intersections. Clearly $Z(0) = K^n$ and $Z(1) = \varnothing$ and so both are Kolchin closed.

To prove the more nontrivial properties, note that it suffices to show that the union of any two Kolchin closed sets is Kolchin closed. Given $X = Z(f_1, \ldots, f_m)$ and $Y = Z(g_1, \ldots, g_\ell)$, I claim that

$$X \cup Y = Z\left(\left\{f_i g_j\right\}_{i \leqslant m, j \leqslant \ell}\right)$$

We first show that $X \cup Y \subset Z\left(\left\{f_i g_j\right\}_{i \leqslant m, j \leqslant \ell}\right)$. If $x \in X \cup Y$ then either $x \in X = Z(f_1, \ldots, f_m)$ or $x \in Y = Z(g_1, \ldots, g_\ell)$, so that either $f_i(x) = 0$ for all $i \leqslant m$ or $g_j(x) = 0$ for all $j \leqslant \ell$. In either case, $f_i g_j(x) = 0$ for all $i \leqslant m$ and $j \leqslant \ell$.

Conversely, if $x \notin X \cup Y$ then there is some i_0 and j_0 such that $f_{i_0}(x) \neq 0$ and $g_{j_0}(x) \neq 0$. Since K is a field, $f_{i_0} g_{j_0}(x) \neq 0$ and so $x \notin Z\left(\left\{f_i g_j\right\}_{i \leqslant m, j \leqslant \ell}\right)$.

Finally, we wish to show that the intersection of an arbitrary family of Kolchin closed sets is Kolchin closed. To do so, suppose that $X_i = Z(\mathcal{F}_i)$ is Kolchin-closed with $\mathcal{F}_i$ a finite family of elements of $K_\Delta\{x_1, \ldots, x_n\}$. To prove the results, we define the locus of a differential ideal I as follows:

$$Z(I) = \{x \in K^n \mid f(x) = 0 \text{ for all } f \in I\}$$

Note that since K is a field, $Z(I) = Z(\sqrt{I})$ and, moreover, for an arbitrary set of differential polynomials $\mathcal{F}$, $Z(\mathcal{F}) = Z(\{\mathcal{F}\})$. Finally,

$$\bigcap Z(\mathcal{F}_i) = Z\left(\bigcup \mathcal{F}_i\right)$$

so that
$$\bigcap Z(\mathcal{F}_i) = Z\left(\bigcup \mathcal{F}_i\right) = Z\left(\left\{\bigcup \mathcal{F}_i\right\}\right) = Z(g_1, \ldots, g_m)$$
for some finite list of $g_1, \ldots, g_m \in K_\Delta\{x_1, \ldots, x_n\}$ since every radical differential ideal is finitely generated. □

Moreover, the proof that the Zariski topology is closed under intersection immediately implies that the Zariski topology is Noetherian: since
$$\bigcap Z(\mathcal{F}_i) = Z\left(\bigcup \mathcal{F}_i\right) = Z(g_1, \ldots, g_m)$$
with each $g_i \in \mathcal{F}_{d(i)}$ for some function $d : \mathbb{N} \to \mathbb{N}$ we have that an arbitrary intersection of closed sets is equal to an intersection of finitely many of them. Thus

Proposition 2.39. *The Kolchin topology on K^n is Noetherian.*

Ritt-Noetherianity moreover implies an analogue of primary decomposition which allows us to talk about *irreducible components* in the context of the Kolchin topology

Proposition 2.40. *Let R be a Ritt-Noetherian ring. Then any non-unit radical differential ideal I is the intersection of a finite set of prime differential ideals. Moreover, the set of prime differential ideals occurring in any such decomposition is unique provided the decomposition is irredundant in the sense that if $\mathcal{P}$ is the set of primes then $\bigcap \mathcal{P} = I$ but for all $\mathcal{P}' \subset \mathcal{P}$ proper then $\bigcap \mathcal{P}' \neq I$.*

Proof. Suppose otherwise, so that there exists an ideal I maximal amongst those that are not expressible as the intersection of a finite number of prime ideals. To apply Zorn's lemma to show that there is such a maximal element, let $\{I_\alpha\}_{\alpha<\lambda}$ be a chain. By Ritt-Noetherianity, there are no infinite chains of radical differential ideals and so the maximal element of the chain is an upper bound amongst those radical differential ideals that are not the intersections of finitely many prime ideals.

By construction, I itself is not prime, so that there is some $r = ab \in I$ with $a, b \notin I$. But then $\{I, a\}$ and $\{I, b\}$ contain I properly and so are themselves the intersection of finitely many prime differential ideals, provided that they are non-unit ideals. To see this, suppose that $\{I, a\} = (1)$. Then $J(I, a) = 1$ since $1^m = 1$ for all m. Then we may write
$$1 = c + \sum_\theta r_\theta \theta(a)$$
so that by multiplying by b we obtain
$$b = cb + \sum_\theta r_\theta \theta(a) b$$
I now claim that since I is radical, for all $\theta = \partial_1^{m_1} \cdots \partial_n^{m_n}$ we have that $b\theta(a) \in I$. It suffices to show that this is the case for $\theta = \partial$ for some $\partial \in \Delta$. Indeed, since I is a radical differential ideal
$$\partial(ab) = a\partial(b) + b\partial(a) \in I$$
and so
$$ab\partial(b) + b^2\partial(a) \in I$$
but then since $ab \in I$ we have that, by subtracting $ab\partial(b)$ and then multiplying by $\partial(a)$
$$b^2\partial(a)^2 \in I$$

so that by radicality

$$b\partial(a) \in I$$

as desired.

But

$$\{I, a\}\{I, b\} \subseteq \{I^2, Ia, Ib, ab\} \subseteq I$$

and if $c \in \{I, a\} \cap \{I, b\}$ then $c^2 \in I$ by above, so that $c \in I$ by radicality. Hence

$$\{I, a\} \cap \{I, b\} = I$$

is the intersection of finitely many prime ideals.

Now, for uniqueness, suppose that

$$I = \bigcap \mathcal{P} = \bigcap \mathcal{Q}$$

for irredundant families of prime differential ideals $\mathcal{P}$ and $\mathcal{Q}$. We first claim that for all $q \in \mathcal{Q}$ there exists a $p \in \mathcal{P}$ such that $q \supseteq p$. Suppose otherwise; if q contains none of the $p \in \mathcal{P}$ then we may find $a_p \in p \backslash q$ for all $p \in \mathcal{P}$. Then $\prod_{p \in \mathcal{P}} a_p \in \bigcap \mathcal{P} = I \subset q$ but none of the $a_p \in q$, contradicting primality of q, proving the result. Likewise for every $p \in \mathcal{P}$ there is a $q \in \mathcal{Q}$ such that $p \supset q$. Similarly, every $p \in \mathcal{P}$ is contained in some $q \in \mathcal{Q}$.

We go by induction on the size of $\mathcal{Q}$. The case that $|\mathcal{Q}| = 1$ is trivial. Suppose $|\mathcal{Q}| \geqslant 1$. Pick $q \in \mathcal{Q}$ and find $p \subset q$ inside $\mathcal{P}$. Then either $p = q$ or $p \neq q$. If $p = q$ then applying induction to $\mathcal{Q}\backslash\{q\}$ and $\mathcal{P}\backslash\{p\}$ finishes the job. If $p \neq q$ then by the above argument we may find some $q' \in \mathcal{Q}$ such that $p \supset q'$. But then

$$q \supset p \supset q'$$

so that $q \supset q'$ properly, contradicting the irredundancy of q. Thus the induction goes through and $\mathcal{P} = \mathcal{Q}$. □

Remark 2.41. The natural notion of Noetherian *dimension* is not as well-behaved in the context of differential algebraic geometry as it is in usual algebraic geometry. Recall that a closed set X in a Noetherian space has dimension defined as follows:

$$\dim(X) = \sup\{n \mid \exists X = X_0 \supset X_2 \supset \cdots \supset X_n \text{ with each } X_n \text{ closed irreducible } \}$$

Now note that in the case of Kolchin-closed sets the supremum may not exist in $\mathbb{N}$. For instance, we have an ascending chain $\{X_i\}$ with each $X_i = Z(\partial^i(x) = 0)$ which is properly ascending over a sufficiently rich differential field K (such as a differentially closed field, which will be discussed shortly). Then for each n we have the chain

$$K = X_0 \supset X_n \supset X_{n-1} \supset \cdots \supset X_1 = Z(\partial(x) = 0)$$

Later on we will discuss some notions of dimension that are amenable to studying differential algebraic geometry.

Primary decomposition gives some insight into the structure of maximal *differential* ideals:

Proposition 2.42. *Let R be a Ritt-Noetherian ring and $I \subset R$ a maximal proper differential ideal. Then I is prime.*

Proof. Since I is proper, $\sqrt{I}$ is also proper as $1 \notin I$ means that $1^n \notin I$ for all $n \in \omega$. Then $I \subset \sqrt{I}$ and so, by maximality, $I = \sqrt{I}$. But then I is a radical differential ideal and so $I = \bigcap P_i$ for some prime differential ideals P_i by primary decomposition. But then $I = P_i$ for one (any) of the P_i's occuring in the decomposition, so that I is prime. □

Remark 2.43. While every maximal differential ideal is a prime ideal, it is not the case that every maximal differential ideal is a *maximal* ideal. For example, consider the ring $R = \mathbb{C}[x, y]$ equipped with the derivation that is trivial on $\mathbb{C}$ and satisfying $\partial(x) = x$ and $\partial(y) = -y$. Then the prime ideal $I = \langle xy - 1\rangle$ is a differential ideal as

$$\partial(xy - 1) = \partial(x)y + \partial(y)x = xy - yx = 0.$$

Since I is a curve and $\mathbb{C}$ is algebraically closed, the only prime ideals properly containing I are maximal ideals of the form $m_{a,b} = \langle x - a, y - b\rangle$ with $a, b \in \mathbb{C}$ such that $ab = 1$. In particular, $a, b \neq 0$. Now, the ideals $m_{a,b}$ are not differential ideals: if they were, then $\partial(x - a) = x$ so that

$$a = (x - a) - \partial(x - a) \in m_{a,b}$$

so that $m_{a,b} = (1)$, which is clearly false.

Since I is a prime differential ideal and the only prime *ideals* containing it are maximal, non-differential ideals, I is a maximal differential ideal which is *not* a maximal ideal.

2.5. **Differentially Closed Fields.** At this point we introduce the fields that play an analogous role to algebraically closed fields in the context of studying differential equations: differentially closed fields. For simplicity we give only the definitions for the single-derivation case:

Definition 2.44. A differentially closed field (F, ∂) is an *differentially closed field*, i.e. if for every finite system of equations and inequations involving differential polynomials in $F\{x_1, \ldots, x_m\}$ for some m with a solution in some $L \supset F$ then there exists a solution in F, i.e. if F is an existentially closed differential field.

A differential field (K, ∂) is a model of DCF_0 provided that for any nonconstant differential polynomials $f, g \in K\{x\}$ with $\mathrm{ord}(g) < \mathrm{ord}(f)$ there is some $x \in K$ such that $f(x) = 0$ and $g(x) \neq 0$.

The above axiomatization of DCF_0 can be easily translated into the first-order language of differential rings

$$\mathcal{L}_{\partial-\text{rings}} = \{0, 1, +, \times, \partial\}.$$

We will show that models of DCF_0 are differentially closed. The following proposition implies the converse: that differentially closed fields model DCF_0.

Proposition 2.45. *Let (K, ∂) be a differential field. Then there exists a $(\hat{K}, \hat{\partial}) \supseteq (K, \partial)$ modeling* DCF_0*.*

Proof. Suppose that $f \in K\{x\}$ and $\mathrm{ord}(g) < \mathrm{ord}(f)$. Then there is an irreducible factor $\tilde{f}$ of f which has order $\mathrm{ord}(f)$. As $\tilde{f}$ is irreducible, the division algorithm guarantees that $g \notin J(f)$. But then setting $K' = \mathrm{Frac}(K\{x\}/J(f))$ yields a field such that the image of x under the canonical projection satisfies $f(x) = 0$ but $g(x) \neq 0$.

To construct a differentially closed field using this procedure, first enumerate the set of pairs

$$\{(f,g) \mid \operatorname{ord}(g) < \operatorname{ord}(f) \text{ and there is no } x \in K \text{ with } f(x) = 0 \text{ and } g(x) \neq 0\}$$

This set has size $\leqslant |K|$. Sequentially construct differential field extensions adding points witnessing differential closure as above. The new field K_1 obtained this way may not be differentially closed; but repeating this process to construct K_2, K_3, and so on yields an ascending chain

$$K \subset K_1 \subset K_2 \subset \cdots$$

such that each pair (f,g) in K_i with no $x \in K_i$ witnessing $f(x) = 0$ and $g(x) \neq 0$ has such a witness in K_{i+1}. Thus the field $\hat{K} = \bigcup K_i$ is differentially closed. By this construction, we see that we can take $|\hat{K}| = |K|$ since it is obtained by adding only $|K|$-many points at each stage and that there are only countably many stages. $\square$

We now prove that DCF_0 admits elimination of quantifiers in this language:

Theorem 2.46. DCF_0 *eliminates quantifiers in the language* $\mathcal{L}_{\partial-rings}$.

Proof. We use a standard model-theoretical test for quantifier-elimination proven in the appendix (A.1): if for any $K, L \models \mathrm{DCF}_0$ with $k \subset K, L$ a differential field and $\overline{a} \in k^n$ and $\phi(\overline{v}, w)$ a quantifier free formula in $\mathcal{L}_{\partial-\mathrm{rings}}$ with $K \models \phi(\overline{a}, b)$ then $L \models \exists w \phi(\overline{a}, w)$.

Replacing K, L with sufficiently saturated elementarily equivalent models and that $k = \operatorname{dcl}(\overline{a})$. We simply need to show that $k\langle b\rangle_\partial \cong k\langle b'\rangle_\partial$ for some $b' \in L$ so that $L \models \phi(\overline{a}, b')$. We argue this as follows:

Let $K, L \models \mathrm{DCF}_0$ be sufficiently saturated and suppose that $k \subset K$, $l \subset L$ are small isomorphic ∂-fields, isomorphic via σ. Then for all $a \in K$ there is $b \in L$ such that σ extends to an isomorphism

$$\hat{\sigma} : k\langle a\rangle_\partial \to l\langle b\rangle_\partial \,.$$

We break into two cases: if a is satisfies a nontrivial differential polynomial over k or not[12]. If a is differentially transcendental then by ω-saturation we may find a realization of $\sigma_*(\operatorname{tp}(a/k)) = p \in S(l)$ inside L, so that $k\langle a\rangle_\partial \cong l\langle b\rangle_\partial$.

If a is differential algebraic over k then let f be a minimal polynomial of $I_\partial(a/k)$. Let $g = \sigma_*(f) \in l\{x\}$. Then the partial type

$$\{g(x) = 0\} \cup \{h(x) \neq 0 \mid h(x) \in l\{x\} \text{ and } h \ll g\}.$$

This type is finitely satisfiable, and so it is satisfiable and by saturation we find a realization b of it in L. But then

$$k\langle a\rangle_\partial \cong k\{x\}/I_\partial(a/k) \cong l\{x\}/I_\partial(b/l) \cong l\langle b\rangle_\partial \,.$$

Thus, in either case we may extend the isomorphism and so we have quantifier elimination. $\square$

Quantifier elimination immediately implies the following characterization of definable sets in DCF_0.

[12] If a does satisfy a nontrivial differential polynomial over k then we say it's differentially algebraic over k; otherwise we say that a is differentially transcendental over k

Proposition 2.47. *Every formula in $\mathcal{L}_{\partial-rings}$ is equivalent modulo* DCF_0 *to a formula of the form*

$$\bigvee_{0\leqslant i\leqslant m}\left[\bigwedge_{1\leqslant j\leqslant n_i} f_{ij}(\overline{v})=0\wedge g_i(\overline{v})\neq 0\right].$$

Proof. Quantifier elimination immediately tells us that every formula is equivalent to one of the form

$$\bigvee_{0\leqslant i\leqslant m}\left[\bigwedge_{1\leqslant j\leqslant n_i} f_{ij}(\overline{v})=0\wedge \bigwedge_{1\leqslant k\leqslant \ell_i} g_{ki}(\overline{v})\neq 0\right].$$

But since $\bigwedge_{1\leqslant k\leqslant \ell_i} g_{ki}(\overline{v})\neq 0$ is equivalent to $\prod_{1\leqslant k\leqslant \ell_i} g_{ki}\neq 0$ since we're in a field, we can set $g_i=\prod_{1\leqslant k\leqslant \ell_i} g_{ki}$ to yield the desired result. □

This characterization allows us to classify and count the types in DCF_0.

Proposition 2.48. *Given a type $p\in S_n(k)$, let*

$$I_p=\{f\in k\{x_1,\dots,x_n\}\mid \text{“}f(\overline{x})=0\text{”}\in p\}.$$

The map $p\mapsto I_p$ is a bijection from $S_n(k)$ to the set of prime differential ideals over $k\{x_1,\dots,x_n\}$.

Proof. The map is clearly well-defined, and I_p is prime since for all f,g if $fg(x)=0\in p$ then either $f(x)=0\in p$ or $g(x)=0\in p$ since p is a complete type. Thus we need to show that it is both injective and surjective.

For injectivity, suppose that $p\neq q\in S_n(k)$. Then there is a formula $\phi(x,a)\in p\backslash q$ equivalent to a formula of the form

$$\bigvee_{0\leqslant i\leqslant m}\left[\bigwedge_{1\leqslant j\leqslant n_i} f_{ij}(\overline{v})=0\wedge g_i(\overline{v})\neq 0\right]$$

with f_{ij} and g_i all in $k\{x_1,\dots,x_n\}$. But then $\phi(x,a)\in p$ iff $f_{ij}\in I_p$ and $g_i\notin I_p$ for all i,j. Since $p\neq q$, this means that either some $f_{ij}\notin I_q$ or some $g_i\in I_q$; in either case $I_p\neq I_q$.

For surjectivity, let I be a prime differential ideal, so that $k\{x_1,\dots,x_n\}/I$ is a differential domain. Then $\mathrm{Frac}(k\{x_1,\dots,x_n\}/I)$ is a field and $\mathrm{tp}(x/k)\mapsto I$ under the above function. Hence $p\mapsto I_p$ is surjective. □

Corollary 2.49. *Over any base field k, $|S_n(k)|=|k|$ and so* DCF_0 *is ω-stable and, in particular, totally transcendental.*

Proof. Over any differential field K, there are at most $|K^{<\omega}|=|K|$ prime differential ideals in K by the Ritt-Raudenbush theorem. Thus there are at most $|K|$ types, so that DCF_0 is ω-stable. □

Using quantifier elimination we can prove that models of DCF_0 are differentially closed as well as the differential analogue of the Nullstellensatz:

Theorem 2.50. (1) *(Models of* DCF_0 *are existentially closed) Let k be a ∂-field and Σ a finite collection of equations and inequations over k with a solution in some differential field $l\supset k$, then Σ has a solution in any $K\supset k$ with $K\models$* DCF_0.

(2) *(Algebra-geometry correspondence)Let* $K \models \mathrm{DCF}_0$ *and* $\Sigma \subset K\{x_1, \ldots, x_n\}$ *and* $V \subset K^n$. *Set*

$$V(\Sigma) = \{x \in K^n \mid (\forall f \in \Sigma)\, f(x) = 0\}$$

and

$$I(V) = \{f \in K\{x_1, \ldots, x_n\} \mid (\forall x \in V) f(x) = 0$$

Then

$$I(V(\Sigma)) = \{\Sigma\}$$

Proof. (1) Suppose that there is a solution to Σ in some extension $l \supset k$. Then l is contained inside a differentially closed $\hat{l}$, and any point in l solving Σ remains a solution to Σ in $\hat{l}$. But then by quantifier elimination, there being a solution to Σ is equivalent to a quantifier-free sentence ϕ_Σ over k, so that

$$\hat{l} \models \phi_\Sigma \iff l \models \phi_\Sigma \iff k \models \phi_\Sigma$$

But then if $K \supset k$ is any differentially closed field,

$$k \models \phi_\Sigma \iff K \models \phi_\Sigma$$

so that K has a solution of Σ.

(2) Note first of all that $\{\Sigma\} \subseteq I(V(\Sigma))$ since $I(V(\Sigma))$ contains Σ and is a radical differential ideal as K is a field so that if $f^n(x) = 0$ for all $x \in V(\Sigma)$ then $f(x) = 0$.

Conversely we show that $I(V(\Sigma)) = \{\Sigma\}$. Suppose that $I(V(\Sigma)) \neq \{\Sigma\}$. Since $\{\Sigma\} \subset I(V(\Sigma))$ this means that there is some $g \in I(V(\Sigma)) \backslash \{\Sigma\}$. But then $g \notin \{\Sigma\}$ and so by the decomposition theorem we may find a prime ideal $P \supset \{\Sigma\}$ with $g \notin P$. Then the field $\mathrm{Frac}(K_\partial\{x_1, \ldots, x_n\}/P) \supset K$ has a point $z \in V(\Sigma)$ such that $g \notin I_\partial(z/k)$, as does its differential closure. But then by the above argument there must be such a point in K, contradicting the assumption that $g \in I(V(\Sigma))$.

□

While we showed above that any differential field F is contained inside *some* differentially closed field, DCF_0 being ω-stable actually gives us much more: there is a unique-up-to-differential-isomorphism differentially closed $\hat{F} \supset F$ with the property that if $L \models \mathrm{DCF}_0$ contains F then $\hat{F}$ embeds into L.

Corollary 2.51. *Let* (F, ∂) *be a differential field. Then there exists a differential field* $\hat{F}$, *unique up to differential field isomorphism, such that if* $K \models \mathrm{DCF}_0$ *and contains* F, *then there is an embedding* $\hat{F} \to K$.

Proof. By ω-stability of DCF_0 (2.49) and the results on existence (A.5) and uniqueness (A.6) of prime models for ω-stable theories in the appendix on model theory we have prime models $\hat{F} \models \mathrm{DCF}_0$ for any differential field F.

Note that by quantifier elimination (2.46) if $K \models \mathrm{DCF}_0$ then any differential field embedding $\hat{F} \to K$ is an elementary embedding. □

Remark 2.52. We note here that we can give an *a priori* non-first-order axiomatization of what it means for a differential field with m commuting derivations to be differentially closed:

A Δ-field (F, Δ) is differentially closed provided every finite system of differential polynomial equations and inequations in $F_\Delta\{x_1, \ldots, x_m\}$ for any m that has a solution in some $L \supset F$ has a solution in F.

It turns out that there is a first-order axiomatization for this theory, but we will not discuss it here.

A crucial fact that we will use when we come to differential Galois theory is the precise relationship between the fields of constants of F and $\hat{F}$.

Proposition 2.53. *Let F be a differential field and $\hat{F}$ its differential closure. Then*

$$C_{\hat{F}} = C_F^{alg.}$$

Proof. First note that $C_{\hat{F}} \supset C_F^{alg}$ since every order-zero (i.e. algebraic) differential polynomial over C_F has a solution inside $C_{\hat{F}}$.

We now wish to show that $C_{\hat{F}} \subset C_F^{alg}$, which means that we need to show that every $a \in C_{\hat{F}}$ is algebraic over F. It suffices to show that $\mathrm{trdeg}_F(F_\partial \langle a \rangle) = 0$. Since $\partial a = 0$, $\mathrm{trdeg}_F(F_\partial \langle a \rangle) \leqslant 1$. Moreover, since $a \in F$, its type $\mathrm{tp}(a/F)$ is isolated by a quantifier free formula ϕ of the form

$$\bigwedge f_i(x) = 0 \wedge g \neq 0.$$

If $\mathrm{trdeg}(a/F) = 1$ then, since C is a pure algebraically closed field it is strongly minimal ϕ is of the form

$$\partial(x) = 0 \wedge g(x) \neq 0$$

for $g \in F[x]$ a *polynomial.* But this cannot be an isolating formula since there exists $a \in C_F$ satisfying this formula. Thus $\mathrm{trdeg}(a/F) = 0$ and so $a \in C_F^{alg}$. □

Similarly, every element of $\hat{F}$ is *differentially algebraic* over F:

Proposition 2.54. *Let $a \in \hat{F}$. Then a satisfies a nontrivial differential polynomial $f \in F\{x\} \backslash \{0\}$.*

Proof. Suppose a satisfies no nonzero differential polynomial $f \in F\{x\} \backslash \{0\}$. Since $\mathrm{tp}(a/F)$ is isolated and a, we may pick an isolating formula ϕ of the form $f(x) \neq 0$. Suppose that f is of order n; then inside $\hat{F}$ there is a solution b to

$$\partial^{n+1}(x) = 0 \wedge f(x) \neq 0$$

so that $a \neq b$. But then $\mathrm{tp}(a/F)$ is not isolated by ϕ, a contradiction. □

Finally, we end this section by showing that the class of definable sets in DCF_0 represents quotients. In other words, that DCF_0 eliminates imaginaries.

Proposition 2.55. *Let T be a theory that has at least two constant symbols and eliminates imaginaries. Then for all definable equivalence relations E on M^n there exists a definable function $f_E : M^n \to M^m$ such that*

$$T \models (xEy \iff f_E(x) = f_E(y))$$

Remark 2.56. Note that if T eliminates imaginaries then, given a definable set X and equivalence relation on X we may identify the quotient X/E with the image $f_E(X)$, which is another definable set.

Theorem 2.57. DCF_0 *eliminates imaginaries.*

Proof. The proof of the theorem is in three steps:

Step 1: Reduce to coding conjugacy classes of differential ideals. First of all note that every definable equivalence relation E is of the following form:

$$E_\phi(y,z) \iff \forall x\,(\phi(x,z) \leftrightarrow \phi(x,y))\,.$$

and that, for all $\phi(x,y)$, E_ϕ is an equivalence relation[13]. Then an automorphism of a model $K \models \mathrm{DCF}_0$ fixes $\phi(x,a)$ if and only if it fixes the E_ϕ-class of a. Let $p_1,\dots,p_n$ be the finitely many types over U of maximal Morley rank containing $E_\phi(y,a)$ and partition them into their G-conjugacy classes $P_1 \cup \dots \cup P_k$ where G is the group of global automorphisms fixing $E_\phi(y,a)$ setwise. If we can find for each conjugacy class P_j a finite tuple b_j depending on a such that P_j is fixed setwise if and only if b_j is fixed pointwise then by compactness we can find formulas ψ_j such that b_j is the unique element such that

$$E_\phi(y,a) \iff \bigwedge_j \psi_j(y,b_j)$$

so that we get a definable map $a \mapsto b = (b_1,\dots,b_k)$.

We reduce to the case of looking at a single conjugacy class of $\{p_1,\dots,p_k\}$; by concatenating tuples we get the result that we want. So assume that $I_{p_1},\dots,I_{p_k}$ are conjugate prime differential ideals. Now an automorphism σ permutes $\{p_1,\dots,p_k\}$ if and only if it permutes the corresponding differential prime ideals $I_{p_1},\dots,I_{p_k}$. Our goal is therefore to find a finite tuple b so that the p_i are permuted if and only if b is fixed pointwise.

Step 2: Reduce to the algebraic case. Let $I = \bigcap I_{p_j}$. Then I_{p_j} is a radical differential ideal, and σ fixes I setwise if and only if it permutes $I_{p_1},\dots,I_{p_k}$. By 2.10 we know that that $I = \{f_1,\dots,f_m\}$, so that there is an m such that each

$$f_\ell \in K[x_i^{(j)} \mid 1 \leqslant i \leqslant m \text{ and } 1 \leqslant j \leqslant n] = R_0$$

Set $I_0 = I \cap R_0$. Then σ fixes I setwise if and only if σ fixes I_0 setwise. We now find a field $k \subset K$ finitely differentially generated such that I is fixed setwise if and only if k is fixed pointwise if and only if some choice of finite tuple of generators b for k is fixed pointwise. We can resort to looking at fields of definition for polynomial ideals by our reduction to looking at the ideal $I_0 \subset R_0$.

Step 3: Construct fields of definition. Consider the K vector space R_0/I_0. Let B be a basis of monomials for this vector space. Then every monomial $u \in R_0$ can be written as $u = (\sum a_{u,\ell} b_\ell) + g_u$ with $a_{u,\ell} \in K$ and $b_\ell \in B$ and $g_u \in I_0$. Note that the $a_{u,\ell}$ are uniquely determined by our choice of B. Then $u - (\sum a_{u,\ell} b_\ell)$ is in I_0 and in fact generate I_0 since the B are a basis. Let

$$k = \mathbb{Q}\langle a_{u,\ell}\rangle$$

Then every element of I_0 has coefficients in k. Moreover this is a finitely generated field extension since the ideal generated by the $a_{u,\ell}$ is finitely generated, so that $k = \mathbb{Q}\langle a_1,\dots,a_n\rangle$. ☐

[13]If E is a definable equivalence relation defined by $\phi(x,y)$, then $E_\phi(y,z)$ holds iff $E(y,z)$ holds by a very easy computation.

2.6. **Differential Dimension Polynomials.** We argue here the polynomial growth of the transcendence degree of a differential field extension. On the face of it, however, "growth" has to measure something changing, and a differential field extension on its own is not something that changes. However, given a differential field extension $K = F\left(\eta_1, \cdots, \eta_n\right)_\Delta)$ over F, we may write K as the union of the fields

$$K_{\eta,q} := F\left(\theta\eta_i \,|\, 1 \leqslant i \leqslant n \text{ and } \theta = \partial_1^{j_1} \cdots \partial_\ell^{j_\ell} \text{ with } \sum_{1 \leqslant k \leqslant \ell} j_k \leqslant q\right).$$

The K_q are simply the fields generated by applications of operators in Θ of order less than or equal to q to the generators $\eta_1, \cdots, \eta_n$. The differential dimension polynomial $\omega_{\eta/F}$ measures the transcendence degree $\operatorname{trdeg}_F K_q$ for $q \gg 0$ sufficiently large.

Theorem 2.58. *Let $\eta = \{\eta_1, \cdots, \eta_n\}$ be a finite tuple of elements of some $\Delta = \{\partial_1, \ldots, \partial_m\}$-field extending F and let $K = F\left(\eta\right)_\Delta$. Then there exists a numerical polynomial $\omega_{\eta,F}$ such that*

(1) *For $q \gg 0$,*
$$\operatorname{trdeg}_F(K_{\eta,q}) = \omega_{\eta,F}(q)$$
(2) $\deg(\omega_{\eta,F}) \leqslant |\Delta| = m$

Proof. The proof of the theorem goes by reducing the algebraic problem– counting the size of a transcendence basis for K_q over F– to a combinatorial problem obtained by looking at a characteristic set of Δ-locus $\operatorname{loc}_\Delta(\eta/F)$.

Step 1: Find dependencies.

Since $\operatorname{loc}_\Delta(\eta/F)$ is a radical differential ideal of $F_\Delta\{x_1, \cdots, x_n\}$ we can extract from it a characteristic set $\mathcal{C}_\eta$. As

$$f \in \operatorname{loc}_\Delta(\eta/F) \iff f(\eta) = 0$$

we have that for all $c \in \mathcal{C}_\eta$, $c(\eta) = 0$ but that $S_c(\eta) \neq 0 \neq I_c(\eta)$. As $I_c(\eta) \neq 0$ we have that $u_c(\eta)$ is algebraic over the field extension

$$F\left(\theta\eta_i \,|\, 1 \leqslant \ell \leqslant n \text{ and } \theta\eta_i < u_a\right)$$

and so if $v = \theta u_c$ then similarly $v(\eta)$ is algebraic over

$$F\left(\theta\eta_i \,|\, 1 \leqslant \ell \leqslant n \text{ and } \theta\eta_i < v\right)$$

by differentiating the polynomial witnessing the algebraicity of u_a and using the fact that $I_v(\eta) = S_c(\eta) \neq 0$.

Step 2: Reduce to a combinatorial problem Set

$$V = \{\theta x_i \,|\, \theta x_i \neq \theta' u_c \text{ for any } \theta' \text{ and } c \in \mathcal{C}_\eta \text{ with } \operatorname{ord}(\theta') \geqslant 1\} = \{\theta x_i \,|\, (\forall c \in \mathcal{C}_\eta)\theta x_i \not\geqslant u_c\}$$

and let $V(t) = \{\theta x_i \,|\, \theta x_i \in V \text{ and } \operatorname{ord}(\theta) \leqslant t\}$

By construction $K_{\eta,t}$ is algebraic over $K_{V,t} := F(v(\eta) \,|\, v \in V(t))$ and, moreover, $\operatorname{trdeg}_F K_{V,t} = |V(t)|$ since otherwise we would find some nonzero Δ-polynomial f in $\operatorname{loc}_\Delta(\eta/F)$ such that $f(v_1(\eta), \cdots, v_m(\eta)) = 0$ for some enumeration of $V(t)$. Reducing $(f(v_1, \cdots, v_m))(u)$ with respect to $\mathcal{C}$ yields another polynomial $\tilde{f}$ equivalent to $f(v_1, \cdots, v_m)$ modulo $\mathcal{C}$ and thus identical as functions on $K_{V,t}$. But then $\tilde{f} = 0$ since it is reduced with respect to $\mathcal{C}$, so that $f(v_1, \cdots, v_n)$ is the zero function, contradicting our assumption that it was a nontrivial relation between the $v(\eta)$'s for $v \in V(t)$.

We put V in correspondence with a subset of $\{1,\dots,n\}\times\mathbb{N}^m$ as follows. Define a map L from differential variables θx_i to $\{1,\dots,n\}\times\mathbb{N}^m$ by mapping

$$\theta x_i = \partial_1^{\ell_1}\cdots\partial_m^{\ell_m} x_i \mapsto (i,\ell_1,\dots,\ell_m).$$

Then let $\mathcal{C}_\eta \to \{1,\dots,n\}\times\mathbb{N}^m$ given by mapping $c\mapsto L(u_c)$ and call its image $L(\mathcal{C}_\eta)$ the *lattice* of $\mathcal{C}_\eta$. Then L is a bijection between V and $L(V)$ and, moreover, $L(V)$ is the complement of the set of elements greater than or equal to the elements of $L(\mathcal{C}_\eta)$.

Step 3: Count. It therefore suffices to show that for $V\subseteq\mathbb{N}$ as above, $|V(t)|$ has polynomial growth in t. We construct ω_η by induction on m and the quantity

$$S(\mathcal{C}_\eta)=\sum_{c\in\mathcal{C}_\eta}\sum_{i=1}^{|\Delta|} n_{i,c}$$

where $n_{i,c}$ is the order of u_c relative to ∂_i. In other words, $n_{i,c}$ is the unique natural number such that $u_c=\partial_1^{n_{1,c}}\cdots\partial_i^{n_{i,c}}\cdots\partial_k^{n_{m,c}}x_j$.

If $S(\mathcal{C}_\eta)=0$ then either $L(\mathcal{C}_\eta)=\varnothing$ or $L(\mathcal{C}_\eta)=(0,\dots,0)$. If it's the former case then

$$|V(t)|=\sum_{i=1}^{n}|\{(\ell_1,\dots,\ell_m)\mid \sum \ell_i\leqslant t\}|=\binom{t+m}{m}$$

If it's the latter case then $|V(t)|=0$, so the case $S(\mathcal{C}_\eta)=0$ is finished.

Suppose now that $S(\mathcal{C}_\eta)>0$ but that for all $n<S(\mathcal{C}_\eta)$ and $p<m=|\Delta|$ we have the result. If $S(\mathcal{C}_\eta)>0$ then there is some point $(i,\ell_1,\dots,\ell_m)\in L(\mathcal{C}_\eta)$ with not all ℓ_j equal to 0. We may assume that $\ell_m\neq 0$. We partition $L(\mathcal{C}_\eta)$ into the two sets: L_0 and L_1 as follows:

$$L_0=\{v=v(i,\ell_1,\dots,\ell_{m-1})\in\{1,\dots,k\}\times\mathbb{N}^{m-1}\mid (v,0)\in L(\mathcal{C}_\eta)\}$$

$$L_1=\{(i,\ell_1,\dots,\ell_m)\mid \ell_m\neq 0 \text{ and } (i,\ell_1,\dots,\ell_m+1)\in L(\mathcal{C}_\eta) \text{ or } \ell_m=0 \text{ and } (i,\ell_1,\dots,\ell_m)\in L(\mathcal{C}_\eta)\}$$

By induction there is a polynomial $\omega_0(t)$ that is asymptotically equal to the size of the complement of L_0 of size $\leqslant t$, while the size of the complement of L_1 of size less than t is also asymptotically a polynomial $\omega_1(t-1)$, so that

$$|V(t)|=\omega_0(t)+\omega_1(t-1)=\omega_\eta(t)$$

is a polynomial. □

Thus, associated to a tuple $\eta=(\eta_1,\cdots,\eta_m)$ in some Δ-field extension K/F we may associate to it a numerical polynomial $\omega_{\eta/F}$. A natural question arises: is $\omega_{\eta/F}$ an invariant of the *field extension* $F_\Delta(\eta)/F$? The answer to this is no, as can be seen by how taking prolongation sequences affects the behavior of the Kolchin polynomial.

Definition 2.59. Let $\eta=(\eta_1,\cdots,\eta_n)\in K$ be an element in a Δ-field. Fix an ordering $\Delta=\{\partial_1,\cdots,\partial_k\}$. The prolongation of η, $\nabla(\eta)$, is the element

$$\nabla(\eta)=(\eta,\partial_1(\eta),\cdots,\partial_k(\eta))\in K^{(m+1)n}.$$

For $\ell>1$ The ℓ^{th} prolongation of η, $\nabla^\ell(\eta)$ is defined recursively as follows:

$$\nabla^\ell(\eta)=\nabla(\nabla^{\ell-1}(\eta))\in K^{(m+1)^\ell n}$$

Remark 2.60. The fact that $\nabla^{\ell+1}(\eta)$ extends $\nabla^{\ell}(\eta)$ can be restated by saying that applying the natural projection

$$\rho_{\ell}^{\ell+1} : K^{(m+1)^{\ell+1}n} \to K^{(m+1)^{\ell}n}$$

given by projecting the first $(m+1)^{\ell}n$ coordinates to $\nabla^{\ell+1}$ behaves as follows:

$$\rho_{\ell}^{\ell+1}(\nabla^{\ell+1})(\eta) = \nabla^{\ell}(\eta)$$

Note that, as sequences, $\nabla^{\ell+1}(\eta)$ extends $\nabla^{\ell}(\eta$ for all $\ell \in \omega$. The *full prolongation sequence* of η, $\nabla^{\infty}(\eta) = \{\nabla^{\ell}(\eta)\}_{\ell\in\omega}$

In this context, we note that the Δ-field generated by η over F is precisely the *pure field* extension $F(\nabla^{\infty}(\eta)$, and that moreover for all $\ell \in \omega$,

$$F_{\Delta}(\nabla^{\ell}(\eta)) = F(\nabla^{\infty}(\eta))$$

Thus, if $\omega_{\eta/F}$ were an invariant of the Δ-field generated by η, then in particular it would have to satisfy $\omega_{\eta/F} = \omega_{\nabla(\eta)/F}$. However, ∇ acts as the shift operator at the level of Kolchin polynomials.

Proposition 2.61. *Let $\eta \in K^n$ with $F \subseteq K$. Then*

$$\omega_{\nabla(\eta)/F}(n) = \omega_{\eta/F}(n+1)$$

Proof. Restating the definition of the Kolchin dimension-counting function in terms of prolongations we have that for sufficiently large m that

$$\omega_{\nabla(\eta)/F}(m) = \mathrm{trdeg}_F(F(\nabla^m(\nabla(\eta))) = \mathrm{trdeg}_F(F(\nabla^{m+1}(\eta)) = \omega_{\eta/F}(m+1).$$

Thus $\omega_{\nabla(\eta)/F}(m) - \omega_{\eta/F}(m+1) = 0$ for $m \gg 0$ and so, as they are univariate polynomials in m we have that $\omega_{\nabla(\eta)/F}(m) \equiv \omega_{\eta/F}(m+1)$ on the nose as polynomials in m. □

This proposition allows us to cook up many examples witnessing the fact that the Kolchin polynomial is *not* an invariant of the Δ-field extension K/F.

Example 2.62. Let $F = \mathbb{Q}$, $\Delta = \{\partial_1\}$, and let η be a differential transcendental element. For instance, take $K = \mathrm{Frac}(\mathbb{Q}_{\Delta}\{x\})$ and $\eta = x$. Then

$$\omega_{x/\mathbb{Q}}(n) = n$$

but

$$\omega_{\nabla(x)/\mathbb{Q}}(n) = n+1.$$

Note that if η is comprised solely of elements in $K^{\Delta=0}$ or elements in F then $\omega_{\eta/F}$ is a constant number and so $\omega_{\nabla(\eta)/F} = \omega_{\eta/F}$.

While the Kolchin polynomial is not a *birational* invariant of the point η, it is a *generic* property of η in the following sense:

Proposition 2.63. *Suppose that $\eta_1, \eta_2 \in K^n$ are such that*

$$I_{\Delta}(\eta_1/F) = I_{\Delta}(\eta_2/F) = \mathfrak{p}.$$

Then

$$\omega_{\eta_1/F} = \omega_{\eta_2/F}$$

Proof. This follows from the proof that $\omega_{\eta/F}$ is a numerical polynomial; namely, $\omega_{\eta/F}$ can be computed solely in terms of a characteristic set for the ideal $\mathfrak{p} = I_{\Delta}(\eta/F)$. Thus, if two points η_1 and η_2 have the same associated differential ideals over F then they have the same characteristic sets and therefore the same Kolchin polynomials. □

3. Differential Galois Theory

3.1. **Binding Groups and Internality.** In this section we give a naïve approach to the construction of model theoretic Galois groups called *binding groups* in a similar manner as that developed by Poizat in *Stable Groups* [9]. Loosely speaking, Poizat's approach takes as input two definable sets X and Y, with X *internal* to Y. Intuitively, internality is basically a condition that says that Y parametrizes X in a strong definable way. Using this parametrizing function, one builds up a *definable*[14] group of automorphisms of X fixing Y. This theory allows one to provide a nice, coherent generalization of Kolchin's theory of strongly normal extensions and the associated Galois theory as well as a slick, conceptual proof that the differential Galois group is an algebraic group.

Definition 3.1. Let T be a theory and let X and Y be definable sets. We say that X is internal to Y provided there exists a point $\overline{c} \in X^m$ and function $u : X^n \times Y^m \to X$ such that for all $x \in X$

$$x = u(\overline{c}, \overline{y})$$

for some $\overline{y} \in Y^n$. Such a u is called an *internality function* and a choice of $\overline{c}$ is called a *fundamental system* of solutions of X relative to Y via u.

Remark 3.2.
- We often represent the data implicit in the statement "X is internal to Y via u" as a triple (X, Y, u). We call this an *internal triple*. A *fundamental system* for (X, Y, u) is a tuple $\overline{c} \in X^n$ such that the function $u(\overline{c}, \overline{y}) : Y^m \to X$ is surjective.
- An example of internality that we've already seen is the case of linear differential equations: In a model of DCF_0, a linear differential operator $\mathcal{L}$ of order n in a single variable has a solution space $Z(\mathcal{L})$ which has dimension n over the constants C. Let $\overline{c} = (c_1, \cdots, c_n) \in K^n$ be a basis for $Z(\mathcal{L})$. Then $Z(\mathcal{L})$ is internal to C via the function $u : Z(\mathcal{L})^n \times C^n$ given by

$$u(\overline{x}, \overline{y}) = \sum_{1 \leqslant i \leqslant n} x_i y_i$$

  and taking as our fundamental set of solutions $\overline{x} = \overline{c}$.
- Note that, in general, there is no unique choice of u to witness the internality of X to Y. For instance, in the above case of linear differential equations we could replace the function $u(\overline{x}, \overline{y}) = \sum_{1 \leqslant i \leqslant n} x_i y_i$ with the function $\tilde{u}(\overline{x}, \overline{y}) = \sum_{1 \leqslant i \leqslant n} 2x_i y_i$ still provides a witness to internality. **We will see later on that the binding groups constructed depend only on the pair (X, Y); in other words, the choice of u does not affect the binding group even though the explicit presentation does invoke u.**

Given an internal triple (X, Y, u) and choice of fundamental system $\overline{c}$ we can construct the binding group $\mathrm{Bind}(X, Y, u, \overline{c})$, which is an explicitly presented interpretable group in T.

[14] Really, it's a group interpretable in Y together with all induced structure

Construction 3.3. Let T be a (complete) totally transcendental theory **MT**, $M \models T$ the prime model, and let (X, Y, u) be an internal triple and $\overline{c} \in X(M)^n$ a fundamental system[15].

Step 1: Find a definable set in natural correspondence with the group of automorphisms in question. The binding group $\operatorname{Bind} X, Y, u, \overline{c}$ is a definable group of permutations of $X(M)$ fixing $Y(M)$ pointwise, defined as follows. We first note that given $\overline{c} \in X(M)^n$ a fundamental set of solutions for X and $\sigma \in \operatorname{Aut}(M/Y(M)))$, $\sigma|_X$ is determined *uniquely* by where σ maps $\overline{c}$, since for all x the equation

$$\sigma(x) = \sigma(u(\overline{c}, \overline{y})) = u(\sigma(\overline{c}), \overline{y})$$

is determined by the fact that σ is an automorphism fixing $Y(M)$. Now, the collection of all fundamental systems $\overline{c}$ for the internal triple (X, Y, u) is a definable subset of X^n:

$$\operatorname{Fund}(X, Y, u) = \{\overline{z} \in X^n \mid (\forall x \in X)\,(\exists \overline{y} \in Y^m)\, x = u(\overline{z}, \overline{y})\} \subseteq X^n$$

Within the set $\operatorname{Fund}(X, Y, u)$ is the subset $\operatorname{tp}(\overline{c}/\varnothing) \subseteq \operatorname{Fund}(X, Y, u)$. This type is isolated by a formula $\phi_c \in \mathcal{L}(M)$.[16] Using ϕ_c we can set up a bijective correspondence

$$\operatorname{Aut}_X(M/Y(M)) = \{\sigma|_{X(M)} \mid \sigma \in \operatorname{Aut}(M/Y(M))\} \cong \phi_{\overline{c}}(M)$$

given by mapping $\sigma \mapsto \sigma(\overline{c}) \in \phi_{\overline{c}}(M)$ and $\overline{z} \in \phi_{\overline{c}}(M)$ mapping to the unique $\sigma \in \operatorname{Aut}(M/Y(M))$ mapping $\overline{c}$ to $\overline{z}$. The map $\phi_{\overline{c}}(M) \to \operatorname{Aut}_X(M/Y(M))$ is well-defined because prime models are homogeneous **Model theory fact; include** and because the and these maps are clearly inverses. This therefore identifies $\operatorname{Aut}(M/Y(M))$ *as a set* with a definable *set* $\phi_{\overline{c}}$.

Step 2: Lift the problem from X to Y and use u to define the group law. We now wish to show that we can endow $\phi_{\overline{c}}$ with the structure of a definable group in such a way that $\operatorname{Aut}_X(M/Y(M)) \cong \phi_{\overline{c}}$ *as groups*. This definable group will be called the *binding group* $\operatorname{Bind}(X, Y, u, \overline{c})$. The approach we take for doing this is to use the internality function u, together with $\overline{c}$, to lift the problem to a problem in Y.

Since elements of $\operatorname{Aut}_X(M/Y(M))$ are in natural correspondence with elements of $\phi_{\overline{c}}(M)$, which is a certain subset of fundamental systems living inside X^n, we slightly tweak u in order to obtain another, related, internal triple that allows us to represent the composition of automorphisms. The function

$$\hat{u} : Y^{nm} \times X^n \to X^n$$

given by the equation

$$\hat{u}(\overline{y}_1, \ldots, \overline{y}_n, \overline{x}) = (u(\overline{y_1}, \overline{x}), \ldots, u(\overline{y_n}, \overline{x}))$$

is surjective, as is the function $\hat{u}_{\overline{c}} = \hat{u}(-, \overline{c}) : Y^{mn} \to X^n$.[17] With this in mind, the set $\phi_{\overline{c}}$ is internal to the subset

$$\hat{Y}_{\overline{c}} = \hat{u}_{\overline{c}}^{-1}(\phi_{\overline{c}}) \subseteq Y^{mn}$$

[15] We can find such a $\overline{c}$ in M because saying that there exists a fundamental system for the internal triple (X, Y, u) can be straightforwardly expressed as a single first order sentence

[16] *A priori* this depends on choice of choice of function u witnessing internality as well as the choice of fundamental system $\overline{c}$. In the end we will argue that the binding group is independent of this choice up to definable isomorphism.

[17] In fact, it is surjective for *any* fundamental system $\overline{z} \in \operatorname{Fund}(X, Y, u)$, not just $\overline{c}$.

via $\hat{u}$.[18]

Thus we may regard $\phi_{\overline{c}}$ as a set interpretable in the induced structure on Y as follows:

$$\phi_{\overline{c}}(M) \cong \hat{Y}_{\overline{c}}^{mn} / \sim_{\overline{c}}$$

where

$$(\overline{y}_1, \cdots, \overline{y}_n) \sim_{\overline{c}} (\overline{y}'_1, \cdots, \overline{y}'_n) \iff \hat{u}(\overline{y}_1, \cdots, \overline{y}_n, \overline{c}) = \hat{u}(\overline{y}'_1, \cdots, \overline{y}'_n, \overline{c})$$

We may now represent the group law on $\phi_{\overline{c}}$ by coding a $\sigma|_X \in \mathrm{Aut}_X(M/Y(M))$ by a equivalence class of $[(\overline{y_1}, \ldots, \overline{y_m})]_{\sim_{\overline{c}}} := [y] \in \hat{Y}_{\overline{c}}$ as follows: let $[y], [w] \in \hat{Y}_{\overline{c}}$. Then

$$[y] * [w] := \text{ the unique } [v] \text{ such that } \hat{u}([v], \overline{c}) = \hat{u}([y], \hat{u}([w], \overline{c}))$$

The identification $\hat{Y}_{\overline{c}} / \sim_{\overline{c}}$ with $\mathrm{Aut}_X(M/Y(M))$ via the function

$$[y] \mapsto \sigma_{[y]} \text{ the unique } \sigma|_X \in \mathrm{Aut}_X(M/Y(M)) \text{ such that } \sigma(\overline{c}) = u([y], \overline{c})$$

satisfies

$$\sigma_{[y]*[w]} = \sigma_{[y]} \circ \sigma_{[w]}$$

by construction and therefore induces a group structure on $\hat{Y}_{\overline{c}}$ isomorphic to that of $\mathrm{Aut}_X(M/Y(M))$. We define the *binding group* of $(X, Y, u, \overline{c})$ to be

$$\mathrm{Bind}(X, Y, u, \overline{c}) := (\hat{Y}_{\overline{c}}, *)$$

equipped with all structure induced by T^{eq}.

Remark 3.4.
- Note that the choice of $\overline{c}$ and u are immaterial at the level of classifying $\mathrm{Bind}(X, Y, u, \overline{c})(M)$ as an *abstract* group: no matter what choice of internality function u and fundamental system $\overline{c}$ we choose, the identification of $\mathrm{Aut}_X(M/Y(M))$ with $\phi_{\overline{c}}$ goes through and, at the level of group structure, we have that if $(X, Y, u_1, \overline{c}_1)$ and $(X, Y, u_2, \overline{c}_2)$ are two quadruples witnessing the internality of X to Y then the above argument yields
$$\mathrm{Bind}(X, Y, u_1, \overline{c}_1)(M) \cong \mathrm{Aut}_X(M/Y(M)) \cong \mathrm{Bind}(X, Y, u_2, \overline{c}_2)(M)$$
- In fact, more than being isomorphic as *abstract* groups, these groups are *definably* isomorphic. If $(X, Y, u_1, \overline{c}_1)$ and $(X, Y, u_2, \overline{c}_2)$ are as above with isomorphisms
$$\theta_i : \mathrm{Bind}(X, Y, u_i, \overline{c}_i) \to \mathrm{Aut}_X(M/Y(M))$$
given by the identification
$$[y]_{\overline{c}_i} \mapsto \text{ the unique } \sigma \text{ such that } \sigma(\overline{c}_i) = u_i([y]_{\overline{c}_i}, \overline{c}_i)$$
Then
$$\theta_2^{-1} \circ \theta_1 : \mathrm{Bind}(X, Y, u_1, \overline{c}_1)(M) \to \mathrm{Bind}(X, Y, u_2, \overline{c}_2)(M)$$
is a *definable* automorphism, mapping
$$[y_1]_{\overline{c}_1} \mapsto \text{ the unique } [y_2]_{\overline{c}_2} \text{ such that } u_1([\overline{y}_1]_{\overline{c}_1}) = u_2([\overline{y}_2]_{\overline{c}_2})$$
which is a definable relation. It is an isomorphism as it is the composition of two isomorphisms.
- The group $\mathrm{Bind}(X, Y)$ is *interpretable* in the induced structure on Y.

[18] Again, this set facially depends on the choice of internality function u and fundamental system $\overline{c}$.

So far we've constructed a definable group of automorphisms of the *prime* model M of T, but what is the significance of the definable group $\mathrm{Bind}(X,Y)$ as we move to other models of T?

Proposition 3.5. *Let T be totally transcendental, (X,Y,u) an internal triple, and $N \models T$. Then for all elements $a,b \in X(N)$, a and b are conjugate by an element of* $\mathrm{Bind}(X,Y)$ *if and only if* $\mathrm{tp}(a/Y(N)) = \mathrm{tp}(b/Y(N))$.

Proof. Suppose that $\mathrm{tp}(a/Y(N)) = \mathrm{tp}(b/Y(N))$. Then since $a = u(y,\overline{c})$ for some class in $Y(N)$, we have that $b = u(y,\overline{c}')$ since $(\exists \overline{z} \in \mathrm{Fund}(X,Y,u))\, b = u(y,\overline{z})$ is a formula in $\mathrm{tp}(a/Y(N)) = \mathrm{tp}(b/Y(N))$. But then the automorphism $\sigma|_X$ mapping $\overline{c} \mapsto \overline{c}'$ exists and is unique, and takes $a \mapsto b$.

On the other hand, suppose that a and b are conjugate via an element of the group $\mathrm{Bind}(X,Y,u,\overline{c})(N)$. Now, over the prime model M of T, $\mathrm{tp}(\overline{c}/Y(M))$ is isolated by $\phi_{\overline{c}}$. This exactly says that there are no tuples of $Y(M)$ witnessing $\phi_{\overline{c}}(z) \wedge \phi_{\overline{c}}(z') \wedge \psi(z,y) \wedge \psi(z',y)$; in other words, for *all* formulas $\phi(z,y)$ we have that

$$T \models (\forall y \in Y)\left[(\phi_{\overline{c}}(z) \wedge \phi_{\overline{c}}(z)) \to (\psi(z,y) \leftrightarrow \psi(z',y))\right]$$

which means that $\phi_{\overline{c}}$ isolates the type of $\overline{c}$ over $Y(N)$ as well. But since any element of $\mathrm{Bind}(X,Y,u,\overline{c})$ preserves $\phi_{\overline{c}}$, which isolates the type of $\overline{c}$, it preserves the types of any element of $X(N)$ as, over any base $X(N) \subseteq \mathrm{dcl}(Y(N) \cup \{\overline{c}\})$. □

3.2. **Pillay's X-strongly-normal theory.** Using the machinery of binding groups, Pillay is able to generalize Kolchin's Galois theory of so-called *strongly normal* extensions of differential fields, which themselves generalize the Picard-Vessiot theory of linear differential equations. Pillay's definition guarantees that the automorphism groups in question have the structure of binding groups and that there is a Galois correspondence.

Throughout this section we fix a large model $U \models \mathrm{DCF}_0$ that everything we consider embeds into.

Definition 3.6. Let F be a differential field, X a set definable from parameters in F in the language of ∂-rings, and K a differential field such that $\hat{F} \supset K \supset F$. We say that K is an X*-strongly-normal* extension of F provided

(1) $X(F) = X(\hat{K})$ for some differential closure $\hat{K}$ of K.
(2) K is finitely generated over F as a differential field.
(3) For any embedding $\sigma : K \to U$ fixing F,

$$\sigma(K) \subseteq K\left\langle X(U)\right\rangle_\partial$$

Remark 3.7.
- Condition (1) above is the analogue of the *algebraically closed constants and no new constants* condition in the Picard-Vessiot theory.
- Condition (2) guarantees that $K = \mathrm{dcl}(F,a)$ for some finite tuple a; the finiteness of this tuple is a technical assumption that will let us move to the framework of binding groups.

We now show how to go from an X-strongly-normal extension to a binding group, which has the structure of an interpretable group in the induced structure on X.

Construction 3.8. Given K an X-normal extension of F, we wish to construct a *definable* group G with isomorphism $\theta : \mathrm{Aut}_X(K\left\langle X(U)\right\rangle/F) \to G(U)$ and that $\theta(\mathrm{Aut}(K/F)) = G(\hat{F})$ under this same identification.

To do this, we construct a binding group out of information from the X-strongly normal extension. By assumption, we may pick a tuple $a \in U$ such that $K = F\langle a\rangle = \mathrm{dcl}(F \cup \{a\})$. Since $\mathrm{DCF}_0(F)$ is totally transcendental, the type $\mathrm{tp}(a/F)$ is isolated by a formula ϕ_a. Now, $b \in \phi_a(U)$ if and only if there is some $\sigma \in \mathrm{Aut}(U/F)$ such that $\sigma(a) = b$ by the homogeneity of U. But then $\sigma(a) \in \sigma(K) \subseteq K\langle X(U)\rangle$ and so, given a, $b \in \mathrm{dcl}(F \cup X(U) \cup \{a\})$. Thus there is an F-definable function $f_b(a,x)$ such that $f_b(a,x) = b$ for *some* choice of tuple $x \in X(U)^m$. Because we cover ϕ_a as the ranges of such functions, the compactness theorem allows us to find a *single* function $u : X_0 \to \phi_a$ witnessing the internality of ϕ_a to a subset of $X_0 \subseteq X^m$. Let

$$G := \mathrm{Bind}(\phi_a, X_0)$$

We now wish to relate G to the automorphism groups in question.

First of all, since $\hat{F}$ is the prime model of $\mathrm{DCF}_0(F)$, we claim that

$$G(\hat{F}) \cong \mathrm{Aut}_{\phi_c}(\hat{F}/(F \cup X_0(\hat{F}))) \cong \mathrm{Aut}(K/F).$$

The first isomorphism is given by the usual identification of the binding group with the group of automorphisms of the prime model of a totally transcendental theory T fixing pointwise the parameters F and $X_0(F)$ with the prime-model-points of the group. The second identification occurs since an automorphism of K fixing F is determined uniquely by the restriction of σ to a as $K = \mathrm{dcl}(F \cup \{a\})$, and since any such automorphism must map a to some element $\sigma(a) \in \phi_a$ that generates K.

We can identify

$$G(U) \cong \mathrm{Aut}_X(K\langle X(U)\rangle/F)$$

by the homogeneity of U and the property that ϕ_a isolates the type of a over F.

This construction moreover admits a very general Galois correspondence:

Theorem 3.9. *Let K be an X-strongly-normal extension of F with Galois group G. For L ($F \subset L \subset K$) an intermediate differential field set*

$$G_L = \{g \in G \mid (\forall c \in L)\, g(c) = c\}.$$

Then

(1) *K is an X-strongly-normal extension of L,*
(2) *G_L is an F-definable subgroup of G and G_L is the Galois group of K over L*
(3) *The assignment $L \mapsto G_L$ is a bijective correspondence between intermediate finitely-differentially-generated differential subfields of K containing F and the F-definable subgroups of G*
(4) *L is an X-strongly-normal extension of F if and only if $G_L \subset G$ is a normal subgroup.*

Proof. (1) We first check that K is an X-strongly-normal extension of L. This follows straightforwardly from the definition since
- $X(F) = X(\hat{K}) = X(\hat{L}) = X(L)$ since $F \subset L \subset K$ and $\hat{F} = \hat{L} = \hat{K}$.
- K is finitely differentially generated over L as

$$K = F_\partial\langle a\rangle \subset L_\partial\langle a\rangle \subset K$$

- if $\sigma : K \to U$ is an L-embedding then σ is also an F-embedding, so that $\sigma(K) \subset K_\partial\langle X(U)\rangle$.

Note that in this step of the proof we did not use that L is finitely differentially generated over F.

(2) We first show that G_L is an F-definable subgroup of G. Since $F \subset L \subset K = F_\partial(a)$ with $\hat{F} = \hat{K}$, we have that $\hat{L} = \hat{F}$. Moreover, $L = F_\partial\langle b\rangle$ for some finite tuple $b \in L$. Since $b \in K$, $b = h(a)$ for some F-definable function h. Let $u(-,-) : \mathrm{Fund}(K, X_0) \times G \to K$ be the function mapping a fundamental system for K to its image under application by G: $u(a, g) = g(a)$. Then for $g \in G$,

$$g(b) = b \iff h(a) = g(h(a)) = h(g(a)) = h(u(a, g)) \iff (\forall c \in \mathrm{Fund}(K, X_0))h(c) = h(u(c, g))$$

and so G_L is an F-definable[19] subgroup of G.

We now wish to show that G_L is the Galois group of K over L. Let $\sigma \in \mathrm{Aut}_X(K_\partial\langle X(U)\rangle/L)$ and let $g_\sigma = \theta(\sigma) \in G$. Then

$$h(a) = h(\sigma(a)) = h(g(a))$$

so that $g \in G_L$ by the above characterization of G_L. Likewise if $\sigma \in \theta^{-1}(G_L)$ then $\sigma \in \mathrm{Aut}_X(K_\partial\langle X(U)\rangle/L)$ since it stabilizes L and fixes X pointwise.

(3) We now check that F-definable subgroups of G correspond to intermediate finitely-differentially-generated extensions of F. Let $H \subset G$ be an F-definable subgroup and consider the K-definable coset of a under H:

$$W_H := \{g(a) \mid g \in H\}$$

By elimination of imaginaries in DCF_0 there is a code c_H for W_H living in K. Set

$$L_H = F_\partial\langle c_H\rangle.$$

As argued above K is an X-strongly-normal extension of L^H with Galois group G_{L^H}. We claim that

$$G_{L^H} = H.$$

To see this we show that $H \subset G_{L^H}$ and that $G_{L^H} \subset H$.

- If $g \in H$ then for all $h \in H$ we have that
$$g(h(a)) = (gh)(a) \in W$$
so that $\theta^{-1}(g)(W) = W$ and thus $\theta^{-1}(g)(c_H) = c_H$, so that $g \in G_{L^H}$ by construction.
- If $g \in G_{L^H}$ then $g(W) = W$ so that $g(a) \in W$. But then $g(a) = h(a)$ for some $h \in H$, hence $g = h \in H$.

This shows that the map $H \mapsto L^H \mapsto G_{L^H}$ is the identity; we now check that the map $L \mapsto G_L \mapsto L^{G_L}$ is the identity. Indeed, if $L = F_\partial\langle b\rangle$ then b is a code for G_L, so that

$$L_{G_L} = F_\partial\langle c_{G_L}\rangle = F_\partial\langle b\rangle = L$$

so that these maps are mutual inverses.

(4) We finally wish to check that normal F-definable subgroups of G correspond to intermediate X-normal subextensions L/F of K.

First we check that if $H \lhd G$ then L^H is X strongly normal over F Note that, by assumption, L is finitely generated over F and that $X(F) = X(\hat{L})$,

[19] The function h may have introduced parameters from F not present in the original data.

so that we need only check the condition that for any $\sigma : L^H \to U$ fixing F,

$$\sigma(L^H) \subset (L^H)_\partial \langle X(U) \rangle$$

Indeed, we know that $\sigma(L^H) \subset K_\partial \langle X(U) \rangle$ since K is X strongly normal. Now we wish to show that $\sigma(L^H) \subset (L^H)_\partial \langle X(U) \rangle$. This follows immediately from normality of H.

□

We can sharpen this result by showing that the hypothesis that L is a *finitely differentially generated* subextension of K over F is redundant:

Proposition 3.10. *Suppose that K is an X strongly normal extension of F. Then any intermediate differential field $F \subset L \subset K$ is also finitely differentially generated.*

Proof. **SKETCH: Requires more thorough treatment of the Galois theory of types and type-definable Galois groups** The above Galois correspondence goes through verbatim for type-definable subgroups of G and (*a priori*) infinitely differentially generated subextensions L. But the descending chain condition on differential algebraic groups would yield for any infinitely differentially generated subextensions L an F-*definable* group G_L, whose code is a *finite* tuple which also generates L. □

3.3. **Galois Theory of Linear Differential Equations.** The theory of binding groups and X strongly normal extensions outlined above has conceptual elegance and give a very general and widely applicable account of Galois theory. In this section, however, we study the classical Picard-Vessiot theory of linear differential equations which Kolchin's strongly normal and Pillay's X strongly normal theories generalize. We first review some of the basic theory of ordinary linear differential equations and then reconstruct the Picard-Vessiot approach to Galois theory using the model-theoretical machinery that we have developed.

Definition 3.11. Fix (F, ∂) an ordinary differential field. A *linear differential operator* of order n over F is a differential polynomial $\mathcal{L} \in F\{x\}$ of the form

$$\mathcal{L}(x) = x^{(n)} + \sum_{i=0}^{n-1} a_i x^{(i)}$$

A *homogeneous linear differential equation* of order n over F is an equation of the form

$$\mathcal{L}(x) = 0$$

for $\mathcal{L}$ an order n linear differential operator.

Our goal is to understand, over various differential field extensions $K \supset F$, the structure of the space of solutions $Z(\mathcal{L})(K) \subset K$. We first establish that $Z(\mathcal{L})$ admits a natural C_K vector space structure and that its dimension is bounded above by $\mathrm{ord}(\mathcal{L})$.

Proposition 3.12. *Let $\mathcal{L}$ be a linear differential operator over F and $K \supset F$. Then $(Z(\mathcal{L}), +)$ forms an additive subgroup of K which is a vector space of dimension $\leqslant \mathrm{ord}(\mathcal{L})$.*

The typical proof of the dimension bound of this proposition uses the notion of the *wronskian* of a tuple of elements in K.

Definition 3.13. Let $x_1, \ldots, x_n \in K$. The *wronskian* $w(x_1, \ldots, x_n)$ of this collection of elements is the determinant of the matrix

$$Wr(x_1, \ldots, x_n) = \left(x_i^{(j)}\right)_{1 \leqslant i \leqslant n; 0 \leqslant j \leqslant n-1}$$

The wronskian gives a way of measuring linear independence over the constants.

Proposition 3.14. *Let $x_1, \ldots, x_n \in K$. Then $w(x_1, \ldots, x_n) = 0$ if and only if $x_1, \ldots, x_n$ are linearly dependent over C_K.*

Proof. Suppose that $x_1, \ldots, x_n$ are linearly dependent over C_K, so that there exist $c_1, \ldots, c_n \in C_K$ not all zero with

$$\sum c_i x_i = 0.$$

Then $\sum c_i x_i^{(j)} = 0$ for all $j \in \omega$, so that

$$\sum_{i=1}^{n} c_i [x_i^{(j)}]^T_{0 \leqslant j \leqslant n-1} = 0$$

yielding linear dependence of the column vectors of $Wr(x_1, \ldots, x_n)$, so that

$$w(x_1, \ldots, x_n) = 0.$$

Conversely, suppose that $w(x_1, \ldots, x_n) = 0$. Then there are $a_1, \ldots, a_n \in K$ such that

$$\sum_{j=1}^{n} a_j [x_i^{(j)}]^T_{0 \leqslant j \leqslant n-1} = 0$$

By dividing and reordering we may assume that $a_1 = 1$ and that $w(x_2, \ldots, x_n) \neq 0$. But then

$$x_1^{(j)} + \sum_{j=2}^{n} a_j x_i^{(j)} = 0$$

for all j. Differentiating we have that

$$x_i^{(j+1)} + \sum_{j=2}^{n} a_j x_i^{(j+1)} + \sum_{i=2}^{n} (a_j)' x_i^{(j)} = 0.$$

But since

$$x_1^{(j+1)} + \sum_{j=2}^{n} a_j x_i^{(j+1)} = 0$$

we have that

$$\sum_{j=2}^{n} (a_j)' x_i^{(j)} = 0$$

for all j. But then if $(a_1)', \ldots, (a_n)' \in K$ are not all 0, then

$$w(x_2, \ldots, x_n) = 0$$

a contradiction. Thus $(a_1)', \ldots, (a_n)'$ are all 0 so that $a_1, \ldots, a_n \in C_K$. ∎

Remark 3.15. The real utility of the wronskian is that gives a *field-independent* way to measure the linear dependence of a set of elements living in some differential field. In other words, for any differential field $K \supset F$, we have the following equivalences

$$x_1, \ldots, x_n \text{ are linearly independent over } C_F \iff w(x_1, \ldots, x_n) \neq 0$$
$$\iff x_1, \ldots, x_n \text{ are linearly independent over } C_K.$$

This characterization of linear dependence over the constants of a differential field provides us with our upper bound on the C_K-dimension on $Z(\mathcal{L})(K)$.

Proof. We first check that $Z(\mathcal{L})(K)$ is a C_K vector space. If $s_1, s_2 \in Z(\mathcal{L})(K)$ and $c_1, c_2 \in C_K$ then

$$\mathcal{L}(c_1 s_1 + c_2 s_2) = c_1 \mathcal{L}(s_1) + c_2 \mathcal{L}(s_2) = 0 + 0 = 0$$

so that $c_1 s_1 + c_2 s_2 \in Z(\mathcal{L})(K)$.

We now argue that $\dim_{C_K}(Z(\mathcal{L})(K)) \leqslant \operatorname{ord}(\mathcal{L})$. Let $\operatorname{ord}(\mathcal{L}) = n$ and let $\mathcal{L} = x^{(n)} + \sum a_i x^{(i)}$. If $x_1, \ldots, x_{n+1} \in Z(\mathcal{L})(K)$ then the first j rows for $1 \leqslant j \leqslant n$ are of the form $(x_1^{(j-1)}, \ldots, x_{n+1}^{(j-1)})$ while the last row can be rewritten as

$$\left(\sum -a_i x_1^{(i)}, \ldots, \sum -a_i x_{n+1}^{(i)}\right)$$

so that the rows are dependent and $w(x_1, \ldots, x_{n+1}) = 0$, so that any set of $n+1$ elements of $Z(\mathcal{L})(K)$ is dependent. Thus $\dim_{C_K}(Z(\mathcal{L})(K)) \leqslant \operatorname{ord}(\mathcal{L})$. □

In general, a differential field (F, ∂) may have no nontrivial solutions to a homogeneous linear differential equation. For instance, if $(F, \partial) = (\mathbb{C}, 0)$ then the differential equation $x' + x = 0$ has no nonzero solutions. That being said, given any linear differential operator $\mathcal{L}$ it is possible to find a $K \supset F$ with maximum possible dimension.

Proposition 3.16. *Let F be a differential field and $\mathcal{L}$ a linear differential operator of order n, then inside any differentially closed $F \subset K \models \mathrm{DCF}_0$ we have that $\dim_{C_K}(Z(\mathcal{L})(K)) = n$.*

Proof. We build up solutions to $\mathcal{L}$ inside K by induction using the axiomatization of DCF_0. A first solution exists inside K since

$$(\mathcal{L}(x) = 0) \wedge (1 \neq 0)$$

satisfies the criteria for having a solution inside K.

Suppose that $a_1, \ldots, a_\ell$ are C_K-linearly independent solutions to $\mathcal{L}$ for $\ell < n$. We wish to find an $\ell+1^{\text{st}}$ linearly independent solution. Since $\ell < n$, the differential polynomial $[w(a_1, \ldots, a_\ell)](x) \in K\{x\}$ has order $\ell < n$, so that by the axioms of differentially closed fields, we know that the system

$$\mathcal{L}(x) = 0 \wedge [w(a_1, \ldots, a_\ell)](x) \neq 0$$

is has a solution $a_{\ell+1}$ in K. Since $w(a_1, \ldots, a_\ell, a_{\ell+1}) \neq 0$, $a_1, \ldots, a_{\ell+1}$ are C_K linearly independent.

This process terminates after adjoining an n^{th} solution since then the polynomial $[w(a_1, \ldots, a_n)](x)$ has order n and we can no longer adjoin new linearly independent elements. □

We call a C_K-basis for $Z(\mathcal{L})(K)$ a *fundamental system of solutions* for $\mathcal{L}$. Our the Picard-Vessiot theory centers on *Picard-Vessiot* extensions, which are highly related to Pillay's X strongly normal extensions.

Definition 3.17. Let K/F be differential fields. K is a Picard-Vessiot extension of F for the linear operator $\mathcal{L}$ provided

- $K = F_\partial \langle a_1, \dots, a_n \rangle$ for $\{a_1, \dots, a_n\}$ a fundamental system of solutions for $\mathcal{L}$.
- $C_K = C_F$.

Many results in the Picard-Vessiot extension are proven under the assumption that $C_K = C_K^{alg}$, and in this case a Picard-Vessiot extension is in fact an example of a C normal extension in the Pillay theory.

Proposition 3.18. *Assume that K/F is Picard-Vessiot with $C_K = C_K^{alg}$. Then K is C strongly normal.*

Proof. K is a finitely differentially generated extension by construction, and since $C_K = C_K^{alg}$ we have that
$$C_{\hat{K}} = C_K = C_F.$$
It remains to show that for all embeddings $\sigma : K \to U$ fixing F that $\sigma(K) \subset K \langle C_U \rangle$ for our universe U. If $\sigma : K \to U$ is an F-embedding, then it is determined by its image $\sigma(a_1), \dots, \sigma(a_n)$. But each $\sigma(a_i)$ must also lie in $Z(\mathcal{L})(U)$ as the coefficients of $\mathcal{L}$ are all in F. But then as $\{a_1, \dots, a_n\}$ are linearly independent over C_K they are also linearly independent over C_U by the wronskian condition, and therefore form a basis for $Z(\mathcal{L})(U)$. This means that there exist constants $c_1, \dots, c_n \in C_U$ such that
$$\sigma(a_i) = \sum c_i a_i$$
so that $\sigma(a_i) \in K_\partial \langle C_U \rangle$. □

Therefore whenever we have a Picard-Vessiot extension K/F with $C_K = C_K^{alg}$ we are free to use any of the results from Pillay's X strongly normal theory, including the Galois correspondence.

Our main theoretical goals at this point are twofold:

(1) Under the assumption that $C_F = C_F^{alg}$ we will show that Picard-Vessiot extensions always exist by using the model-theoretic machinery we've built up.
(2) Since Picard-Vessiot extensions are C strongly normal, their Galois groups are algebraic groups over C. We will show that, in fact, their Galois groups are linear algebraic by giving a definable representation of $Gal(K/F)$ into $GL_n(C)$ for $n = \text{ord}(\mathcal{L})$. We will see this using the explicit function witnessing the internality of fundamental sets of solutions to $\mathcal{L}$ to C.

Proposition 3.19. *Let F be a differential field such that $C_F = C_F^{alg}$ and $\mathcal{L}$ over F a linear differential operator. Then there exists a Picard-Vessiot extension K/F contained inside the differential closure of F, $\hat{F}$.*

Proof. Model-theoretic proof. Since $C_F = C_F^{alg}$, $C_F = C_{\hat{F}}$. Since we can always find a fundamental system of solutions to $\mathcal{L} = 0$ inside $\hat{F}$ pick $a_1, \dots, a_n \in \hat{F}$ a C_F-basis for $\mathcal{L} = 0$ and set
$$K = F_\partial \langle a_1, \dots, a_n \rangle .$$

Then K is generated by a fundamental system of solutions to $\mathcal{L} = 0$ and

$$C_F \subset C_K \subset C_{\hat{F}} = C_F$$

so that $C_K = C_F$.

□

We now show that given a Picard-Vessiot extension K/F for $\mathcal{L}$ of order n, the Galois group is not just an algebraic group over C, but in fact a subgroup of $GL_n(C)$.

Proposition 3.20. *Let K/F be a Picard-Vessiot extension K/F for $\mathcal{L}$ of order n with binding group G. Then there is a faithful definable representation $G \to GL_n(C)$, so that G is a linear algebraic group.*

Proof. We identify G as a set with the set of realizations of $p = \mathrm{tp}(a_1, \dots, a_n/F)$ of some fundamental set of solutions of $\mathcal{L}$ as in the binding group construction. If $g \in G(F)$ then knowing the formulas $g(a_i) = \sum c_{ij} a_i$ for all i uniquely determines g, so that the map

$$g \mapsto A_g := (c_{ij}) \in GL_n(C)$$

is injective, and is well-defined since the a_i being a basis for $Z(\mathcal{L})(\hat{F})$ means that the (c_{ij}) are unique. Moreover, the set

$$\tilde{G} \subset GL_n(C)$$

given by

$$\tilde{G} = \{(c_{ij}) \in GL_n(C) \mid (\exists g \in G) \wedge g(a_i) = \sum c_{ij} a_i\}$$

is definable. The map $g \mapsto A_g$ is a group homomorphism since

$$A_{gh}(a_1, \dots, a_n)^T = gh(a_1, \dots, a_n) = g(h(a_1, \dots, a_n)) = A_g(A_h(a_1, \dots, a_n)^T)$$

for all $g, h \in G$. □

To finish off our study of Picard-Vessiot extensions we compute a few differential Galois groups.

We first address the relationship between Picard-Vessiot theory and the usual algebraic Galois theory: algebraic Galois extensions of differential fields with algebraically closed constants are Picard-Vessiot? The answer is yes, following Cormier, Singer, and Ulmer [11].

Theorem 3.21. *Let K be a differential field such that $C_K = C_K^{alg}$. Let $f \in K[y]$ be an irreducible polynomial. Then its splitting field K_f is a Picard-Vessiot extension of K.*

Proof. We assume that f is irreducible of degree m and write f as

$$f(y) = y^m + \sum_{i=0}^{m-1} a_i y^i \in \mathbb{C}(x)[y].$$

We now construct a linear operator $\mathcal{L}_f$ such that $Z(\mathcal{L}_f)$ is spanned by the solutions of f.

Let $z_1, \dots, z_m$ be the solutions of f. Then for each z_i the (unique) derivation on $K(z_i)$ extending ∂ is given by

$$\partial(z_i) = -\frac{\sum\limits_{j=0}^{m-1} a_j' z_i^j}{m z_i^{m-1} + \sum\limits_{j=1}^{m-1} j a_j z_i^{j-1}}$$

given by differentiating the formula $f(z_i) = 0$. Note that this equation is implied by the equation $f = 0$.

We claim that there is *some* n such that the solutions of f satisfy a nontrivial order n homogeneous linear differential equation. Since $K(z_i)/K$ is a dimension m vector space over K we have that

$$z_i, (z_i)', \dots, z_i^{(m)}$$

must be linearly dependent over K: there exists $b_0, \dots, b_m$ not all zero such that

$$\sum_{j=0}^{m} b_j z_i^{(j)} = 0.$$

The same b_j work for all z_i satisfying f. Pick $n \leqslant m$ minimal such that the z_i satisfy a linear differential equation of order n and call it $\mathcal{L}_f(y)$.

Then any root of f solves $\mathcal{L}_f$. To show that K_f is itself Picard-Vessiot we must show that inside K_f we can find a fundamental system of solutions and that K_f has no new constants.

- (No new constants) Since K_f is an algebraic extension of K, $K_f \subset \hat{K}$ which has constants C_K, so that
$$C_K \subset C_{K_f} \subset C_{\hat{K}} = C_K$$
so K_f has no new constants.
- (Fundamental system of solutions) Let L be the Picard-Vessiot extension $\mathcal{L}_f$ and let G be its Galois group. Then G acts on $Z(f)$ and so the vector space V generated by $\{z_1, \dots, z_m\}$ is invariant under the action of G. and, therefore, satisfies a linear differential equation of order $\leqslant \mathrm{ord}(\mathcal{L}_f)$ which is a factor of $\mathcal{L}_f$:
$$\mathcal{L}_V = w(y, z_1, \dots, z_m)/w(z_1, \dots, z_m)$$
has coefficients fixed by G and is therefore an equation over K, of order $\leqslant \mathrm{ord}(\mathcal{L}_f)$. But $\mathcal{L}_f$ has minimal order, so that $\mathcal{L}_f$ and $\mathcal{L}_V$ differ only by a multiple factor and so V generates L as well as K_f. Thus K_f is a Picard-Vessiot extension.

Finally we claim that $G(C_K) = Gal(K_f/K)$. Since any $\sigma \in Gal(K_f/K)$ fixes C_K by construction, and since $\partial(z_i)$ is a rational function in K, any $\sigma \in Gal(K_f/K)$ is an element of $G(C_K)$ and visa versa. $\square$

Example 3.22. Consider $K = \mathbb{C}(x)$ and consider the equation

$$y^3 - x = 0.$$

Then differentiating the equation on both sides yields

$$3y^2 y' - 1 = 0$$

so that

$$y' = \frac{1}{3y^2} = \frac{y}{3y^3} = \frac{1}{3x}y$$

But then

$$\mathcal{L}_f = y' - \frac{1}{3x}y$$

is our associated linear differential operator. The solutions of f are

$$Z_f = \{x^{1/3}, \xi x^{1/3}, \xi^2 x^{1/3}\}$$

for ξ a primitive cube root of unity. But then the $\mathbb{C}$ vector space generated by Z_f is dimension 1 and the Galois group is cyclic of order 3.

3.4. **Algebraic D-Groups and Logarithmic Derivatives.** With the tools developed in the previous section applied to the single-derivation case we can show that every X-strongly normal field extension K over an algebraically closed base field $F = F^{alg}$ can be written as

$$K = F\langle\alpha\rangle_\partial$$

where α is a tuple satisfying a certain equation called a *logarithmic differential equation* over some algebraic group.

FINISH

3.5. **Constrained Cohomology.** In analogy with algebraic geometry, one can study an analogue of Galois cohomology called Kolchin's *constrained* cohomology in the context of differential algebraic geometry. Given the model-theoretic tools that we have developed so far, we opt to follow the approach of Pillay, who showed that constrained cohomology is a special case of his so-called *definable* cohomology. To give some substance to the theory we will discuss how one may use definable cohomology in classifying certain special extensions of structures, including how to use constrained cohomology to classify generalized strongly normal extensions of a given differential field.

The general setup of Pillay's theory of definable cohomology is to work in a first-order structure M and subset $A \subset M$ such that M is atomic and (**strongly?**) homogeneous over A. $G = G(M)$ will be an A-definable group[20] and Gal will be the group $Aut(M/A)$ automorphisms of M fixing A pointwise.

Note that Gal acts on G since G is A-definable: if $\psi(a, x)$ is the formula defining G, then $\sigma(G)$ is defined by the $\psi(\sigma(a), x) = \psi(a, x)$, so that $\sigma(G) = G$ and so any $\sigma \in Gal$ induces an automorphism of G.

Definition 3.23. A *cocycle* from Gal to G is a set-theoretic function $f : Gal \to G$ such that for all $\sigma, \tau \in Gal$,

$$f(\sigma \circ \tau) = f(\sigma) \cdot \sigma[f(\tau)] \in G.$$

We say that f is a *definable* cocycle provided that it is represented by a definable function $h(x, y)$ in the following sense: there exists a finite tuple c such that for all $\sigma \in Gal$,

$$f(\sigma) = h(a, \sigma(a))$$

Cocycles f and g are *cohomologous*, written $f \sim g$ provided there is some $b \in G$ such that for all $\sigma \in Gal$

$$g(\sigma) = b^{-1} f(\sigma)\sigma(b).$$

[20] There's no need to take G to be abelian for the general construction

Note that $\sim$ is an equivalence relation as G is a group. A *trivial cocycle* is one that is cohomologous to the cocycle $e : Gal \to G$ given by the function $e(\sigma) = e_G$ for all $\sigma \in Gal$. Namely, a trivial cocycle is a cocycle of the form $f_b(\sigma) = b^{-1}\sigma(b)$ for some given $b \in G$.

The first definable cohomology *set* $H^1_{def}(Gal, G)$ is the set of cocycles modulo the relation $\sim$ of being cohomologous.[21]

We now give two geometric interpretations of $H^1_{def}(Gal, G)$: one corresponding to classifying principal homogeneous spaces of the group G up to G-equivariant definable isomorphism, and one corresponding to classifying the A-forms of an A-definable set X.

To interpret definable cohomology in the context of definable principal homogeneous spaces, we fix a few definitions.

Definition 3.24. A definable principal homogeneous space X over A for a definable group G consists of the following data:

- A definable set X definable over A
- A definable regular (right) action of G on X; that is, a right action $G \circlearrowright X$ definable over A such that for all $x_1, x_2 \in X$ there is a unique $g \in G$ such that $x_1 \cdot g = x_2$

An A-isomorphism of definable G-principal homogeneous spaces X and Y over A (with actions $\cdot_X$ and $\cdot_Y$ of G on X and Y respectively) is a definable isomorphism $f : X \to Y$ over A such that for all $g \in G$ and $x \in X$,

$$f(x \cdot_X g) = f(x) \cdot_Y g,$$

i.e. f is a G-equivariant definable isomorphism between X and Y.

The set of A-definable principal homogeneous spaces for G up to isomorphism is denoted $PHS_A(G)$.

Definable cohomology (over A) classifies definable principal homogeneous spaces for G up to A-isomorphism.

Proposition 3.25. *There is a correspondence between $PHS_A(G)$ and classes of cocycles in $H^1_{def}(Gal, G)$.*

Proof. The main idea is to find a canonical way to associate to an $X \in PHS_A(G)$ an element of $c_X \in H^1_{def}(Gal, G)$ and visa versa.

First suppose that $X \in PHS_A(G)$. Then pick $x_0 \in X$. For any $\sigma \in Gal$ there is a unique $g_\sigma \in G$ such that

$$\sigma(x_0) = x_0 \cdot g_\sigma.$$

Define $c_X(\sigma) = g_\sigma$. This map is a cocycle as

$$x_0 \cdot g_{\sigma\tau} = \sigma\tau(x_0) = (x_0 \cdot g_\tau) \cdot \tau(g_\sigma)$$

by construction. Note that this is a definable cocycle as it is represented by the map

$$h(x, y) = \text{ the unique } g \in G \text{ such that } x \cdot g = y$$

[21]If G is a group, then $H^1_{def}(Gal, G)$ is in fact a group. The set of cocycles is a group under pointwise multiplication in G: $f * g(\sigma) = f(\sigma)g(\sigma)$ is a cocycle as

$$f * g(\sigma\tau) = f(\sigma\tau)g(\sigma\tau) = f(\sigma)\sigma[f(\tau)]g(\sigma)\sigma[g(\tau)] = f * g(\sigma)\sigma[f * g(\tau)].$$

It is easy to see that trivial cocycles are a normal subgroup of this group.

by taking $x = x_0$.

The cohomology class of c_X is independent of choice of x_0, since if $x_1 = x_0 \cdot g$ then the resulting cocycle $\tilde{c}$ is cohomologous to c_X via

$$\tilde{c}(\sigma) = g^{-1} \cdot c_X(\sigma) \cdot \sigma(g)$$

which comes from "untwisting" the action of $\tilde{c}(\sigma)$ by first moving x_1 to x_0 via g^{-1} and computing everything from there.

Conversely, we wish to construct an A-definable principal homogeneous space X_c out of a given cocycle $c \in H^1_{def}(Gal, G)$. By representing c as $h(x_0, \sigma(x_0))$, let X_0 be a formula isolating $\mathrm{tp}(x_0/A)$ (which exists by our assumption that M is atomic over A) and for all $x, y \in X_0$, $h(x, y) \in G$ by choice of X_0. □

Forms

THE ARITHMETIC PICTURE

Following the yoga of using binding groups to glean information about differential Galois theory, the geometric interpretations of H^1_{def} can be used to give results about the existence and uniqueness of strongly normal extensions of a differential field k.

To motivate how definable cohomology could show up in this context, consider the case of a linear differential operator $\mathcal{L}$ over a field K with $C_K = C_K^{alg}$. In this setting we have the existence and uniqueness of Picard-Vessiot extensions of K for the equation $\mathcal{L} = 0$. In this case any extension of the form $K(a)$ for any element $a \in \mathrm{Fund}(Z(\mathcal{L}))$ is a Picard-Vessiot extension of K, and if two elements $a, b \in \mathrm{Fund}(Z(\mathcal{L}))$ have the same type over K then the resulting extensions $K(a)$ and $K(b)$ are K-isomorphic differential fields for trivial reasons. On the other hand, if $\mathrm{tp}(a/A) \neq \mathrm{tp}(b/A)$, when do we know that $K(a) \cong_K K(b)$?

For instance, consider the differential operator $\mathcal{L} = \partial^2$ over the field $K = \mathbb{C}$. Then the tuples $(1, t)$ and $(t, 1)$ in $\mathbb{C}(t)^2$ have different types over K but still yield isomorphic Picard-Vessiot extensions of K, as can be witnessed by the isomorphism $f : \mathrm{tp}((1, t)/\mathbb{C}) \to \mathrm{tp}((t, 1)/\mathbb{C})$ given by $f(x, y) = (y, x)$, which is defined over $\mathbb{C}$. In other words, there exists a $\mathbb{C}$-definable isomorphism between the types of these two elements which guarantees that these extensions are isomorphic. More geometrically, the types $\mathrm{tp}((1, t)/\mathbb{C})$ and $\mathrm{tp}((t, 1)/\mathbb{C})$ are both principal homogeneous spaces of the differential Galois group of $\mathcal{L}$ over K, and they are isomorphic by a G-equivariant definable action. It is this perspective that allows us to use definable cohomology to give a precise answer to questions like: how many non-isomorphic Picard-Vessiot extensions of k are there for the equation $\mathcal{L} = 0$?

3.6. **The Galois Groupoid.** Intrinsic Galois Group; the action induced from internality

Connected components

4. Differential Algebraic Groups

For me, a differential algebraic group will simply be a Kolchin-closed set G equipped with a differential morphism $m : G \times G \to G$ satisfying the usual group multiplication laws and $i : G \to G$ for the inversion map $g \mapsto g^{-1}$ compatible with m.

Appendix A. Preliminaries from Model Theory

In this section we prove a few model-theoretic results used in the main text that would have taken us too far afield. We go roughly in order of appearance. We start with the proof of the quantifier elimination test we used to show that DCF_0 eliminates quantifiers.

Proposition A.1. *Suppose that L is a language, T an L-theory, and $\phi(v)$ an L-formula. Then the following are equivalent:*

(1) *There is a quantifier-free $\psi(v)$ equivalent to $\phi(v)$ modulo T*
(2) *For all models $M, N \models T$ and common substructure $A \subset M, N$, then $M \models \phi(a)$ if and only if $N \models \phi(a)$ for all tuples a from A.*

Proof. (1 implies 2) If ϕ is equivalent to a quantifier-free ψ then for all tuples a from A we have that

$$M \models \phi(a) \iff M \models \psi(a) \iff A \models \psi(a) \iff N \models \psi(a) \iff N \models \phi(a).$$

(2 implies 1) We first handle two degenerate cases: if $T \models \forall v \phi(v)$ or $T \models \forall v \neg\phi(v)$ then $\phi(v)$ is equivalent modulo T to the formulas $v = v$ and $v \neq v$ respectively. Thus we may assume that both $T \cup \{\exists v \phi(v)\}$ and $T \cup \exists v \neg\phi(v)$ are consistent.

Let

$$\Gamma_+(v) = \{\psi(v) \,|\, \psi(v) \text{ quantifier-free such that } T \models \forall v(\phi(v) \to \psi(v)\}$$

i.e. $\Gamma_+(v)$ is the set of quantifier-free consequences of ϕ. If we can show that a realization of $\Gamma_+(v)$ realizes ϕ then by compactness there exists a finite set $\{\psi_1, \ldots, \psi_k\}$ such that

$$\left(\bigwedge \psi_k(v)\right) \to \phi(v)$$

so that, since each ψ_k was a consequence of ϕ,

$$\left(\bigwedge \psi_k(v)\right) \leftrightarrow \phi(v)$$

modulo T.

For contradiction suppose that there were a realization a of Γ_+ that $\neg(\phi)(a)$. Let $M \models T$ contain a and let $A = \langle x \rangle$ be the substructure generated by a. Then the type

$$\Sigma = T \cup \mathrm{diag}_{qf}(A) \cup \{\phi(a)\}$$

is satisfiable since, if unsatisfiable, it it because there exists $\psi_1, \ldots, \psi_\ell \in \Gamma_+(v)$ such that

$$T \models \forall v \left(\bigwedge \psi_k(v) \to \neg\phi(v)\right)$$

so that

$$T \models \forall v \left(\phi(v) \to \bigvee \neg\psi_k(v)\right)$$

contradicting the fact that the $\psi_k(v)$ are all consequences of $\phi(v)$.

Pick $N \models T$ containing A such that $N \models \phi(a)$. Then $M \models \neg\phi(a)$ but $N \models \phi(a)$ and $A \subset M, N$, a contradiction. □

Remark A.2. In applying the above quantifier elimination test we may replace M and N with saturated elementary extensions $\tilde{M} \succ M$, $\tilde{N} \succ N$ and it does not affect either direction of the proof. Thus it suffices to show the result for sufficiently saturated models of T.

Moreover, it suffices to apply the result to existential formulas since one can perform quantifier elimination one quantifier at a time.

We also used a result in stability theory known as the *stable embeddedness of definable sets*:

Definition A.3. Let X be a definable set in some theory T. Then X is stably embedded provided that for all definable subsets $Y = \phi(x, m) \subset X^n$ defined with a parameter in some $M \models T$, we may find $m' \in X(M)$ such that $Y = \phi(x, m')$.

Every definable set in a stable theory is stably embedded.

Proposition A.4. *Let T be stable and X be definable. Then X is stably embedded.*

Proof. Let $Y = \phi(m, x) \subset X^n$. Then since $p = \mathrm{tp}(m/X(M))$ is definable, we have that

$$\phi(m, X(M)) = (d_p x)\phi(x, X(M))$$

by definability of types. But $(d_p x)\phi(x, y)$ is defined over $X(M)$. □

Theorem A.5. *Let T be an ω-stable theory. Then over every set A of parameters $T(A)$ has a prime model.*

Theorem A.6. *Let T be an ω-stable theory, A a set of parameters, and $M, N \supset A$ be two prime models. Then $M \cong N$.*

Proof. We break the proof into two parts: first showing that every prime model is constructible and then showing that any two constructible models over A are isomorphic.

Step 1: Prime models are constructible.

Step 2: Constructible Models are pairwise isomorphic. Let M and N be constructible models of $T(A)$. The goal is to perform a (somewhat subtle) back-and-forth argument using an explicit construction of each model. □

References

[1] M. Artin. Algebra.
[2] E. Hrushovski and Z. Sokolović. Minimal types in differentially closed fields. Preprint, 1992.
[3] I. Kaplansky. An Introduction to Differential Algebra.
[4] E. Kolchin. Differential Algebra and Algebraic Groups
[5] D. Marker. Model Theory: An Introduction.
[6] D. Marker. Model Theory of Differential Fields.
[7] T. McGrail. The Model Theory of Differential Fields with Finitely Many Commuting Derivations.
[8] A. Ovchinnikov. Differential Algebra.
[9] B. Poizat. Stable Groups.
[10] D. Pierce and A. Pillay. A note on the axioms for differentially closed fields of characteristic zero. *Journal of Algebra*, 204(1):108–115, 1998.
[11] O. Cormier, M. Singer, and F. Ulmer. Computing the Galois Group of a Polynomial Using Linear Differential Equations.
[12] A. Pillay and M. Ziegler. Jet spaces of varieties over differential and difference fields. *Selecta Math. (N.S.)*, 9(4):579–599, 2003.

Reid Dale, University of California at Berkeley, Group in Logic and the Methodology of Science, Evans Hall, Berkeley, CA 94720, USA

E-mail address: `reiddale@math.berkeley.edu`