# Model Theory of Valued $D$-Fields
**with Applications to Diophantine Approximations in Algebraic Groups**

A thesis presented

by

Thomas Warren Scanlon

to

The Department of Mathematics

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Mathematics

Harvard University

Cambridge, Massachusetts

May 1997

**Abstract**

In this thesis we introduce a general notion of a $\mathcal{D}$-ring generalizing that of a differential or difference ring. In Chapter 3, this notion is specialized to consider valued fields $D$-fields: valued fields $K$ having an operator $D : K \to K$ and a fixed element $e \in K$ satisfying $D(x + y) = Dx + Dy$, $D(1) = 0$, $D(xy) = xDy + yDx + eDxDy$, $v(e) \geq 0$, and $v(Dx) \geq v(x)$. Upon a further specialization, namely that the residue field has characteristic zero and $v(e) > 0$, we present axioms for the model completion and prove a version of the Ax-Kochen-Ershov principle.

Using the general model theory of valued $\mathcal{D}$-fields and results of Hrushovski on groups definable in separably closed fields we prove a characteristic $p$ analog of Buium's *abc* theorem for semi-abelian varieties.

Using the same general results on estimates in valued $\mathcal{D}$-fields together with results of Chatzidakis and Hrushovski on groups definable in transformally closed fields we prove a version of a conjecture of Tate and Voloch on the $p$-adic distance from torsion points of a semi-abelian variety to a subvariety.

This thesis was written under the primary direction of Professor Ehud Hrushovski. Professor Barry Mazur provided significant guidance on Chapters 4, 5, and 6.

To Bæde

I thank my thesis advisor Ehud Hrushovski for all that he has done on behalf of this thesis and of my graduate education in general. He has been very generous – somehow, the phrase "very generous" sounds too weak – with his time and ideas. He is also owed a mathematical debt for conceiving that methods of stability theory could have number theoretic consequences. The reader will recognize that the deep theorems underlying Chapters 5 and 6 are due to him.

I thank also Barry Mazur for his time and advice on Chapters 4, 5, and 6. While at first I balked at producing more geometric versions of these arguments, I am now grateful for his insistence that I look for such arguments for the greater understanding of these problems this point of view has afforded me.

Many other people have contributed to this thesis. The list that follows is no doubt incomplete and I apologize to those people who I inadvertently omit. I thank in particular Dan Abramovic, Matthias Aschenbrenner, Brian Conrad, Keith Conrad, Wee Teck Gan, Joshua Greene, Johan de Jong, Jacob Lurie, Jessica Millar-Young, Adi Ofer, David Pollack, Gerald Sacks, Željko Sokolović, José Felipe Voloch, and Carol Wood.

I thank those people who kept the administrative side of my graduate studies running smoothly, especially Donna D'Fini.

During my first three years of graduate study I was supported by an Office of Naval Research Graduate Fellowship. My work with Professor Hrushovski was made possible through the student support provision of his NSF grant.

# Chapter 1

# Introduction

## 1.1 Problems Considered

### 1.1.1 The Elementary Theory of Valued $D$-fields

The analysis of $D$-henselian fields addresses several model theoretic questions.

- Is there a reasonable theory of valued fields with an automorphism or derivation that interacts with the valuation structure? Michaux has produced a model complete theory of valued differential fields [37], but in his theory there is very little interaction between the valuation and the derivation. Various people, including Macintyre and van den Dries, have asked whether the theory of Hardy fields considered as ordered, differential fields might admit a model companion. While I would very much like to know the answer to this question, it is not answered herein.

- Is there a theory of a non-trivial difference field which admits elimination of quantifiers in a reasonable language? In the absolute model complete case with $e \neq 0$ the theory of $D$-henselian fields provides such an example.

- Can the theories of difference rings and of differential rings be regarded as instances of one theory? The theory of valued $D$-fields with $e \neq 0$ provide evidence that the answer is yes. In these structures one has an automorphism on the valued field which specializes to a derivation on the residue field.

- What is the theory of $k((\epsilon^{\mathbb{Q}}))$ considered as a differential field where $k$ is a differentially closed field of characteristic and the derivation is extended term by term?

There are many other questions about $D$-henselian fields that we leave unaddressed. The structure of the definable sets suggests that a dimension theory lifting the stability theoretic ranks on the residue field ought to exist. Moreover, there are many unexpected definable functions which ought to be classified.

### 1.1.2 Proximity Estimates

Chapters 4, 5, and 6 concern applications of the model theory of valued $D$-fields to number theoretic questions. A model theorist may be interested in this merely as an exhibition of the use of model theoretic methods to solve concrete mathematical problems.

The Mordell-Lang conjecture in characteristic $p$ (now a theorem of Hrushovski [25]) asserts that if $G$ is a commutative algebraic group over an imperfect separably closed field $K$ and $\Gamma \subseteq G(K)$ is a subgroup with the property that $\mathrm{rk}_{\mathbb{Z}_{(p)}}\{y \in G(K) : [n]y \in \Gamma \text{ for some } n \in \mathbb{Z} \setminus (p)\} < \infty$, then for a *general* subvariety $X \subseteq G$ the Zariski closure of $\Gamma \cap X$ is a finite union of cosets of algebraic subgroups of $G$.

The Manin-Mumford conjecture (now a theorem of Raynaud [43]) asserts that if $G$ is a commutative algebraic group over a number field and $X$ is a subvariety, then the Zariski closure of the intersection of $X$ with the torsion points of $G$ is a finite union of cosets of algebraic subgroups of $G$.

If one takes the field of definition in each of these conjectures to be finite, then the results are patently false. The theorems proved in Chapters 5 and 6 show that there is some uniformity in the way that these results become false upon specialization.

More precisely, we show, for instance, that if $G$ is a semi-abelian scheme over $\mathbb{Z}_p$ and $f : G \to \mathbb{P}^1$ is a rational function defined over $\mathbb{Z}_p$, then for any torsion point $P$ in $G$ of order prime-to-$p$, either $f(P) = 0$ or $\infty$ or $|f(P)|_p > \epsilon$ for $\epsilon > 0$ depending only on $f$.

The reader may wonder why we call the main theorem of Chapter 5 an *abc* theorem. We do this because Buium called his own characteristic zero version of this theorem an *abc* theorem. Consult [5] for a discussion of how one may interpret this theorem as an instance of a general conjecture which also specializes to Mason's function field *abc* theorem.

## 1.2 Algebraic Preliminaries

Most of the algebraic facts and notation used in this thesis are either standard or explained as they are introduced.

For us, a valued field is a field $K$ given together with a homomorphism $v : K^\times \to \Gamma$ where $\Gamma$ is an ordered abelian group and $v$ satisfies the inequality $v(x + y) \geq \min\{v(x), v(y)\}$.

**Definition 1.2.1.** If $K$ is a field and $v$ is a valuation on $K$, then $\mathcal{O}_{K,v} := \{x \in K : v(x) \geq 0\}$ denotes the ring of integers of $K$ with respect to $v$. If we happen to be considering only one valuation at the time, the notation $\mathcal{O}_K$ may be used. If $\Sigma$ is a family of valuations on $K$, then $\mathcal{O}_{K,\Sigma} := \bigcap_{v \in \Sigma} \mathcal{O}_{K,v}$.

**Throughout this thesis, *ring* means *commutative ring* .**

On occasion we will need to use the existence of a henselization. Recall that a local ring $(R, \mathfrak{m})$ is *henselian* if whenever $P(X) \in R[X]$ is a polynomial and $a \in R$ reduces modulo $\mathfrak{m}$ to a simple root of the reduction of $P$, then there is a unique solution to $P$ which also reduces to the reduction of $a$. A *henselization* of a local ring $(R, \mathfrak{m}_R)$ is given by a morphism of local rings $(R, \mathfrak{m}) \to (R^h, \mathfrak{m}_{R^h})$ which induces the identity on the residue field and which has the universal property that if $R \to S$ is any local map from $R$ into a henselian ring $S$, then there is a unique map $R^h \to S$ such that the following diagram commutes.

$$\begin{array}{ccc} R & \longrightarrow & R^h \\ \downarrow & \swarrow \exists! & \\ S & & \end{array}$$

For a proof that henselizations exist, see [44].

## 1.3 Background

In the following sections, we explain some of the basic background for the non-specialist. The reader may want to skip this section and refer back to it only as needed.

### 1.3.1 Model Theory for the Algebraic Geometer

In this section we will outline some of the basics of model theory with an eye towards their use in Chapters 4, 5, and 6. We do not describe the model theory needed for Chapter 3. Though the background for Chapter 3 is fairly easy – it is usually taught in an undergraduate course on model theory – I expect that the problem addressed there will interest few non-logicians. If you happen to be one of the few, then I suggest you consult [10], [20], or [45] for the definitions of model completeness, types, saturation, etc. and the proofs of the theorems used there.

**Definition 1.3.1.** A *language* $\mathcal{L}$ is given by specifying a set of function symbols $\{F_i\}_{i \in I}$, a set of relation symbols $\{R_j\}_{j \in J}$, and a set of constant symbols $\{c_k\}_{k \in K}$. The function and relation symbols are given together with their arity. That is, $F$ is supposed to represent a function of $n_F$-variables and $R$ is supposed to represent a relation on $m_R$-tuples.

**Definition 1.3.2.** Let $\mathcal{L}$ be a language. An *$\mathcal{L}$-structure* $\mathfrak{M}$ consists of a non-empty set $M$ together with functions $F^{\mathfrak{M}} : M^{n_F} \to M^{n_F}$, sets $R^{\mathfrak{M}} \subseteq M^{m_R}$ and elements $c^{\mathfrak{M}}$ for each function symbol $F$, relation symbol $R$, and constant symbol $c$ in $\mathcal{L}$. This is also called an *interpretation* of $\mathcal{L}$ by $\mathfrak{M}$.

**Definition 1.3.3.** A *morphism of $\mathcal{L}$-structures* $f : \mathfrak{M} \to \mathfrak{N}$ is given by a function $f : M \to N$ on the universes respecting all the structure given by $\mathcal{L}$. That is, if we let $f$ also denote the induced map on the Cartesian powers of $M$, then $f(R^{\mathfrak{M}}) \subseteq R^{\mathfrak{N}}$, $F^{\mathfrak{N}}(f(\mathbf{x})) = f(F^{\mathfrak{M}}(\mathbf{x}))$, and $f(c^{\mathfrak{M}}) = c^{\mathfrak{N}}$ for relation symbols $R$, function symbols $F$, and constant symbols $c$ in $\mathcal{L}$.

*Example* 1.3.4.

- The language of groups is $\mathcal{L}_{gp} := \mathcal{L}(\cdot, {}^{-1}, e)$ where $\cdot$ is a binary function symbol, ${}^{-1}$ is a unary function symbols and $e$ is a constant symbol. A group is an $\mathcal{L}$-structure in a natural way, but not all $\mathcal{L}$-structures are groups. An $\mathcal{L}$-structure is just a pointed set $(M, e)$ given with functions $\cdot : M^2 \to M$ and ${}^{-1} : M \to M$. One usually violates the strict rules for the writing of functions by denoting $\cdot(x, y)$ by $(x \cdot y)$.
- The language of rings is $\mathcal{L}_{ring} := \mathcal{L}(+, \cdot, -, 0, 1)$ where $+$ and $\cdot$ are binary function symbols, $-$ is a unary function symbol, and 0 and 1 are constant symbols.
- The language of graphs is $\mathcal{L}_{graph} := \mathcal{L}(E)$ where $E$ is a binary relation. A graph $\Gamma = (V, E)$ may be regarded as an $\mathcal{L}$-structure by taking $V$ as the universe and setting $E^{\Gamma} := \{(v, w) \in V^2 :$ there is an edge from $v$ to $w\}$.

*Remark* 1.3.5. It is customary to notationally identify an $\mathcal{L}$-structure with its universe. This practice reduces the number of symbols that one must repeat, but in some instances one needs to use such locutions as "$M$ considered as an $\mathcal{L}$-structure." This is done in common mathematical practice as well. For example, one may wish to emphasize that some map is a morphism when considered as a map of modules but not as a map of rings.

*Remark* 1.3.6. Model theory owes its existence to the logical distinction between a language and its interpretation in a structure. While to some mathematicians this may seem to be a peculiarity of logic, it has mathematical content. The correct application of model theoretic methods depends on maintaining this distinction.

**Definition 1.3.7.** Let $\mathcal{L}$ be a language. A *term of* $\mathcal{L}$ is a formal symbol built from a set of variables $\{x_i\}_{i=1}^{\infty}$ and the function and constant symbols of $\mathcal{L}$ by composition. To be precise, the set of terms is the smallest set with the properties that

- $x_i$ is a term for each variable $x_i$,
- $c$ is a term for each constant symbol $c$ of $\mathcal{L}$, and
- $F(t_1, \ldots, t_n)$ is a term if $F$ is a function symbol on $n$ variables of the language and each $t_i$ is a term.

*Remark* 1.3.8. In general, a term can look rather complicated. One might expect that in the language of groups, a term would correspond to an element of the free group generated by the variables $\{x_i\}_{i=1}^{\infty}$. However, no relations are built into the terms. $(x_1 \cdot x_1^{-1})$ is not the same term as $e$ for instance. We will correct this problem after defining the notion of a theory.

**Definition 1.3.9.** Let $\mathcal{L}$ be a language. A *basic* (or *atomic*) *formula* is an expression of the form $t_1 = t_2$ or $R(t_1, \ldots, t_m)$ where each $t_i$ is a term and $R$ is a relation symbol of $\mathcal{L}$ on $m$ variables.

**Definition 1.3.10.** Let $\mathcal{L}$ be a language. A *formula of* $\mathcal{L}$ is built from the basic formulas by finite Boolean combinations ("not", "and", "or") and existential and universal quantification over the variables $x_i$. To be precise, the set of formulas is the smallest set with the properties that

- a basic formula is a formula,
- if $\phi$ and $\psi$ are formulas then so are $\neg(\phi)$ (read: "not $\varphi$"), $(\phi \wedge \psi)$ (read: "$\phi$ and $\psi$"), and $(\phi \vee \psi)$ (read: "$\phi$ or $\psi$"), and
- if $\phi$ is a formula, then for each variable $x_i$ so are $(\exists x_i)\phi$ and $(\forall x_i)\phi$.

*Remark* 1.3.11. If $\mathfrak{M}$ is an $\mathcal{L}$-structure, then there is a natural way to interpret any formula in $\mathfrak{M}$. If each instance of the variables $x_i$ is bound by a quantifier, then the truth or falsity of the formula in $\mathfrak{M}$ is determined. These formulas are called *sentences*. If $\phi$ is a sentence and $\mathfrak{M}$ makes $\phi$ true, then we write $\mathfrak{M} \models \phi$ and say "$\mathfrak{M}$ models $\phi$." Rather than test the reader's patience, I will not give a formal definition of bound variable or of how a formula is to be interpreted in a structure.

**Definition 1.3.12.** Let $\mathcal{L}$ be a language and $\mathfrak{M}$ be an $\mathcal{L}$-structure. Let $\phi(x_1, \ldots, x_n)$ be an $\mathcal{L}$-formula having only the variables $x_1, \ldots, x_n$ as free variables. The *solution set of* $\phi$ *in* $\mathfrak{M}$, written $\phi(\mathfrak{M})$, is the set $\{\mathbf{a} \in M^n : \mathfrak{M} \models \phi(\mathbf{a})\}$. A subset of $M^n$ is called *definable* if it is the solution set to some formula. A function is said to be definable if its graph is definable.

*Example* 1.3.13.

- Let $K$ be an algebraically closed field considered as an $\mathcal{L}_{ring}$-structure. Then the basic definable sets are affine varieties. By Chevalley's theorem on projections of constructible sets [13], the definable sets are just the constructible sets - finite Boolean combinations of affine varieties.
- If $G$ is a group, then the basic definable sets are sets of solutions to equations. The definable sets are in general much more complicated.
- If $\mathfrak{M}$ is any $\mathcal{L}$-structure and $A \subseteq M$ is any subset, then we can expand the language to $\mathcal{L}_A$ by adding new constant symbols for each $a \in A$. $\mathfrak{M}$ is naturally an $\mathcal{L}_A$-structure. We call the definable sets in $\mathfrak{M}$ with respect to $\mathcal{L}_A$ *$A$-definable* or *definable with parameters from $A$*.

**Definition 1.3.14.** A *theory* $T$ is a set of sentences from a language $\mathcal{L}$. An $\mathcal{L}$-structure $\mathfrak{M}$ is a model of $T$, written $\mathfrak{M} \models T$ if $\mathfrak{M} \models \psi$ for each sentence $\psi \in T$. $T_0$ is a set of *axioms* for $T$ if $\mathfrak{M} \models T_0 \Rightarrow \mathfrak{M} \models T$. We usually require that a theory be consistent. That is, there is some $\mathfrak{M} \models T$. If $\mathfrak{M}$ is any $\mathcal{L}$-structure, then the *theory of* $\mathfrak{M}$ is $\mathrm{Th}(\mathfrak{M}) := \{\phi : \mathfrak{M} \models \phi\}$.

With the notion of a theory, we can restrict the classes of structures considered. For instance, the usual axioms for groups, rings, and fields are axioms (in $\mathcal{L}_{gp}$ or $\mathcal{L}_{ring}$) in the above sense. It is now true that modulo the theory of groups any term in $\mathcal{L}_{gp}$ is logically equivalent to a reduced word as an element of the free group on $\{x_i\}_{i=1}^{\infty}$.

We say that a theory $T$ *eliminates quantifiers* if for every model $\mathfrak{M} \models T$ every definable set in $\mathfrak{M}$ is a boolean combination of basic definable sets As mentioned above, the theory of algebraically closed fields eliminates quantifiers, but $\mathrm{Th}(\mathbb{Q}, +, \cdot, 0, 1)$ does not. Only one of the theories we use in the number theoretic applications (the theory of transformally closed fields) fails to eliminate quantifiers, though the definable sets are fairly close to being basic in that case as well.

**Definition 1.3.15.** Let $T$ be a theory in some language $\mathcal{L}$. Let $E(\mathbf{x}, \mathbf{y})$ be an $\mathcal{L}$ formula in $2n$ variables. $E$ is a *definable equivalence relation* if $E(\mathfrak{M})$ is an equivalence relation on $M^n$ for any model $\mathfrak{M} \models T$. $T$ *eliminates imaginaries* if for each definable equivalence relation $E$ there is a definable function $f_E : M^n \to M^m$ (for some $m$) such that $\mathbf{x}E\mathbf{y} \iff f_E(\mathbf{x}) = f_E(\mathbf{y})$. To put it another way, the category of definable sets is closed under quotients.

*Example* 1.3.16.

- The theory of algebraically closed fields eliminates imaginaries (see [42]).
- The theory of valued fields in the language of rings augmented by a relation $R_v(x, y) \iff v(x) \leq v(y)$ does not eliminate imaginaries. Consider for instance the equivalence relation $v(x) = v(y)$ (see [47]).

*Remark* 1.3.17. All of the theories (differentially closed fields, transformally closed fields, and separably closed fields of finite imperfection degree) we use in Chapters 5 and 6 eliminate imaginaries (see [51], [12], and [34]).

With the basic definitions behind us, let us consider the first non-trivial theorem of model theory.

**Theorem 1.3.18.** *(Compactness) Let $\mathcal{L}$ be a language. Let $\Sigma$ be a set of $\mathcal{L}$-sentences. Assume that for each finite $\Sigma_0 \subseteq \Sigma$ that there is some $\mathfrak{N}_{\Sigma_0} \models \Sigma_0$. Then there is some $\mathfrak{M} \models \Sigma$.*

The compactness theorem implies uniformities. Let us now derive one such.

**Proposition 1.3.19.** *Let $\mathcal{L}$ be a language. Let $T$ be an $\mathcal{L}$-theory. Let $\phi(\mathbf{x})$ be an $\mathcal{L}$-formula. Suppose that for each model $\mathfrak{M} \models T$ the set $\phi(\mathfrak{M})$ is finite. Then there is a natural number $N$ such that if $\mathfrak{M} \models T$ then $|\phi(\mathfrak{M})| \leq N$.*

■ Define

$$\Sigma \quad := \quad T \cup \{(\exists \mathbf{x_1} \ldots \mathbf{x_n})[\bigwedge_{1 \leq i < j \leq N} \mathbf{x_i} \neq \mathbf{x_j} \ \wedge \bigwedge_{i=1}^{N} \phi(\mathbf{x_i})]\}_{N=1}^{\infty}$$

where the notation $\bigwedge_{i \in I} \psi_i$ is an abbreviation for the conjunction of all the $\psi_i$'s. If the proposition were false, then for any $N$ we could find some $\mathfrak{M}_N \models T$ with $|\phi(\mathfrak{M}_N)| > N$. That is, $\mathfrak{M}_N \models T \cup \{(\exists \mathbf{x_1} \ldots \mathbf{x_n})[\bigwedge_{1 \leq i < j \leq N} \mathbf{x_i} \neq \mathbf{x_j} \ \wedge \ \bigwedge_{i=1}^{N} \phi(\mathbf{x_i})]\}_{n \leq N}$. By the compactness theorem, there would be some $\mathfrak{M}_\infty \models \Sigma$. This is a contradiction since $\mathfrak{M}_\infty \models T$ so that $\phi(\mathfrak{M}_\infty)$ must be finite, but the axioms in $\Sigma$ imply that $\phi(\mathfrak{M}_\infty)$ is infinite. ■

This proposition exemplifies our main use of the compactness theorem. We will prove the existence of uniform bounds by showing finiteness for all models of some theory and then by arguing by compactness that there must be a uniform bound. Not every such argument will be a direct application of Proposition 1.3.19, but the proofs will have the same form.

The compactness theorem can be used to prove the existence of models in which families of definable sets having the finite intersection property have non-empty intersection. To be precise, an $\mathcal{L}$-structure $\mathfrak{M}$ is called *$\kappa$-saturated* (where $\kappa$ is some infinite cardinal) if whenever $\{X_i\}_{i \in I}$ is a family of definable subsets of $M$ defined over a fixed set of parameters of size strictly less than $\kappa$ and for each finite $I_0 \subseteq I$ the intersection $\bigcap_{i \in I_0} X_i$ is not empty, then $\bigcap_{i \in I} X_i \neq \emptyset$.

The compactness theorem is often proven as a consequence of Łos' Lemma on ultraproducts. For one proposition considered in Chapter 4 we revert to the ultraproduct construction rather than a direct appeal to the compactness theorem because the language involved there is somewhat complicated.

**Definition 1.3.20.** Let $I$ be a set. A *filter* $\mathcal{F}$ on $I$ is a set of subsets of $I$ satisfying

- $I \in \mathcal{F}$
- $(A \subseteq B$ and $A \in \mathcal{F}) \Rightarrow B \in \mathcal{F}$
- $A, B \in \mathcal{F} \Rightarrow A \cap B \in \mathcal{F}$
- $\emptyset \notin \mathcal{F}$

$\mathcal{F}$ is called an *ultrafilter* if in addition it satisfies

- $(\forall A \subseteq I)\ A \in \mathcal{F}$ or $(I \setminus A) \in \mathcal{F}$

Let $I$ be some set and let $\mathcal{F}$ be a filter on $I$. It follows by Zorn's Lemma that $\mathcal{F}$ may be extended to an ultrafilter $\mathcal{F}'$.

Let $\{X_i\}_{i \in I}$ be a family of sets indexed by $I$. The ultraproduct of this family with the respect to the ultrafilter $\mathcal{F}'$ is the quotient of the product $\prod_{i \in I} X_i$ by the equivalence relation $\mathbf{x} \equiv_{\mathcal{F}'} \mathbf{y} \iff \{i \in I : x_i = y_i\} \in \mathcal{F}'$. One writes the ultraproduct as $\prod_{/\mathcal{F}'} X_i$. If each $X_i$ is a $\mathcal{L}$-structure for some language $\mathcal{L}$, then so is the ultraproduct. We use square brackets to denote the equivalence classes. If $F$ is a function symbol, then define $F([\mathbf{x}]) := [F(\mathbf{x})]$. If $R$ is a relation symbol, then define $R([\mathbf{x}^1], \ldots, [\mathbf{x}]^{(n)}) \iff \{i \in I : R(x_i^{(1)}, \ldots, x_i^{(n)})\} \in \mathcal{F}'$. These are well defined. Łos' Lemma asserts that to check satisfaction for any sentence by the ultraproduct it suffices to check the satisfaction almost everywhere co-ordinatewise.

**Lemma 1.3.21 (Łos' Lemma).** *If $\{\mathfrak{M}_i\}_{i \in I}$ is a family of $\mathcal{L}$-structures for some language $\mathcal{L}$ and $\mathcal{F}$ is an ultrafilter on $I$, then for any $\mathcal{L}$ sentence $\phi$ $\prod_{/\mathcal{F}} \mathfrak{M}_i \models \phi \iff \{i \in I : \mathfrak{M}_i \models \phi\} \in \mathcal{F}$.*

We will now consider the deeper theorems lying behind Chapters 5 and 6.

**Definition 1.3.22.** Let $\mathfrak{M}$ be an $\mathcal{L}$-structure. Let $X \subseteq M^n$ be a subset. Define a language $\mathcal{L}'$ and an $\mathcal{L}'$-structure as follows. The universe of $\mathfrak{X}$ is $X$. For each $\mathcal{L}$-formula $\phi(x_1, \ldots, x_{nm})$ in $nm$ variables, there is a basic relation $R_\phi$ in $\mathcal{L}'$ and $R_\phi(\mathfrak{X}) := X^m \cap \phi(\mathfrak{M})$.

The induced structure may be more or less complicated than one might guess at first sight. For example, if $G$ is an algebraic group over an algebraically closed field, then there is more to the induced structure on $G$ than just its group structure since every subvariety of $G^n$ is a definable set. On the other hand it may happen that the structure on a definable set is much simpler than that of the ambient structure. We will meet examples of this while discussing the Zil'ber trichotomy.

Before proceeding to the Zil'ber trichotomy, let us describe the theories that we will use later.

$\mathrm{SCF}_{p^e}$ denotes the theory of separably closed fields of imperfection degree $p^e (> 1)$. That is, if $K \models \mathrm{SCF}_{p^e}$ then $K = K^{sep}$ and $[K : K^{(p)}] = p^e$.

$\mathrm{DCF}_0$ denotes the theory of differentially closed fields of characteristic zero. A model of $\mathrm{DCF}_0$ is a field $K$ given with a derivation $\partial : K \to K$ for which every reasonable differential equation has generic solutions. To be precise, if $P(X_0, \ldots, X_n)$ is an irreducible polynomial over $K$ and $Q(X_0, \ldots, X_n)$ is some other polynomial such that $\deg_{X_n} Q < \deg_{X_n} P$, then there is some $a \in K$ such that $P(a, \partial a, \ldots, \partial^n a) = 0$ but $Q(a, \partial a, \ldots, \partial^n a) \neq 0$.

$\mathrm{ACFA}_0$ denotes the theory of transformally closed fields of characteristic zero. A model of $\mathrm{ACFA}_0$ is a field $K$ given with an automorphism $\sigma : K \to K$ for which all reasonable equations involving $\sigma$ have solutions. To be precise, let $V$ be an irreducible affine variety over $K$. Let $V^\sigma$ denote the result of applying $\sigma$ to the coefficients of the equations defining $V$, or if you like, define $V^\sigma$ by the Cartesian square

$$\begin{array}{ccc} V^\sigma & \longrightarrow & V \\ \downarrow & & \downarrow \\ \mathrm{Spec} K & \xrightarrow{\sigma^*} & \mathrm{Spec} K \end{array}$$

Let $W \subseteq V \times V^\sigma$ be an irreducible subvariety projecting dominantly and generically finitely onto $V$. Then for any Zariski open $U \subseteq W$ there is some point of the form $(a, \sigma(a)) \in U(K)$.

*Remark* 1.3.23. While it is quite easy to write the axioms for $\mathrm{SCF}_{p^e}$ and $\mathrm{DCF}_0$ in the formal language, the case of $\mathrm{ACFA}_0$ is a bit tricky.

A small field $k$ is associated to each of the above examples. In the case of $\mathrm{SCF}_{p^e}$, $k := \bigcap_{n=1}^{\infty} K^{(p^n)}$. In the case of $\mathrm{DCF}_0$, $k := \ker \partial$. In the case of $\mathrm{ACFA}_0$, $k := \mathrm{Fix}(\sigma)$.

**Definition 1.3.24.** Let $T$ be a theory. A *group definable in* $T$ is a quadruple of formulas $G(\mathbf{x})$, $\mu(\mathbf{x}, \mathbf{y}, \mathbf{z})$, $\iota(\mathbf{x}, \mathbf{y})$, and $e(\mathbf{x})$ such that for any model of $\mathfrak{M} \models T$, $\mu(\mathfrak{M})$ is the graph of a group structure on $G(\mathfrak{M})$ for which $\iota(\mathfrak{M})$ is the graph of an inverse and $e(\mathfrak{M})$ is the identity element. An $\infty$-definable group $G$ is given by $G(\mathfrak{M}) = \bigcap_{i \in I} G_i(\mathfrak{M})$ where each $G_i$ is a definable group and $I$ is some index set (in our applications it is either finite - in which case $G$ is already definable - or it is countably infinite) for any model $\mathfrak{M} \models T$.

*Example* 1.3.25.

- In all of the theories $\mathrm{SCF}_{p^e}$, $\mathrm{DCF}_0$, and $\mathrm{ACFA}_0$, if $G$ is an algebraic group over $K$, then $G(K)$ is a definable group.
- In the cases of $\mathrm{DCF}_0$ and $\mathrm{ACFA}_0$, if $G$ is an algebraic group over $k$, then $G(k)$ is a definable group. However, in the case of $\mathrm{SCF}_{p^e}$, $G(k)$ is merely $\infty$-definable. We can present $G(k)$ as $\bigcap_{n=0}^{\infty} G(K^{(p^n)})$.
- $\{x \in \mathbb{G}_m(K) : \sigma(x) = x^\ell\}$ for some integer $\ell$ is a definable group in $\mathrm{ACFA}_0$. This group has the property that its induced structure sees only the group structure.

**Definition 1.3.26.** An $\infty$-definable group $G$ is *c-minimal* if any definable subgroup of $G$ of infinite index is finite. $G$ is of *finite rank* if it has a finite composition series where each successive quotient is $c$-minimal.

*Example* 1.3.27.

- In the above examples, if $G$ is an algebraic group over $k$, then $G(K)$ is not $c$-minimal (or even of finite rank), but $G(k)$ will always be of finite rank. If $G$ is a simple abelian variety, then $G(k)$ is also $c$-minimal.
- The example of $\{x \in \mathbb{G}_m(K) : \sigma(x) = x^\ell\}$ is $c$-minimal.

**Theorem 1.3.28 ([24, 23, 25, 12]).** *The Zil'ber principle for groups is true for* $\mathrm{DCF}_0$*, $SCF_{p^e}$, and* $\mathrm{ACFA}_0$*. That is, if $G$ is a c-minimal $\infty$-definable group in one of these theories, then either*

- *there is a definable homomorphism with finite kernel from $G$ to an algebraic group over $k$ or*
- *the induced structure on $G$ is locally modular – that is, the only definable subsets of $G^n$ are finite boolean combinations of cosets of definable subgroups.*

*Moreover, if $G$ is a finite rank group all of whose c-minimal subquotients in some composition series are locally modular and $H$ is a finite rank group all of whose c-minimal subquotients are non-locally modular, then $G$ and $H$ are orthogonal. That is, the only definable subsets of $G \times H$ are finite boolean combinations of sets of the form $Y \times Z$ where $Y \subseteq G$ is definable in $\mathfrak{G}$ and $Z \subseteq H$ is definable in $\mathfrak{H}$.*

*Remark* 1.3.29. This theorem is usually referred to as the Zil'ber Trichotomy, but when one specializes to the case of groups as we have done the third (so-called "trivial") case disappears. The methods of proof requires the consideration more general structures than just groups and at least in the cases of $\mathrm{DCF}_0$ and $\mathrm{SCF}_{p^e}$ make no use of the ambient algebraic structure.

### 1.3.2 Algebraic Geometry for the Model Theorist

Even though Chapters 4, 6, and 5 deal exclusively with problems in diophantine geometry, very little knowledge of that subject is required to follow the proofs. However, we do occasionally use the language of schemes and quote a couple deep theorems. So in this section we will describe schemes and these theorems as well as give some basic definitions.

Let $K$ be an algebraically closed field. Define projective $n$-space over $K$, $\mathbb{P}^n(K)$, by $\mathbb{P}^n(K) := (K^{n+1} \setminus \{\mathbf{0}\})/K^\times$. That is, $\mathbb{P}^n(K)$ consists of the set of equivalence classes for the equivalence relation on $(n+1)$-tuples $(x_0, \ldots, x_n)$ from $K$ (where not all the $x_i$'s are zero) given by $(x_0, \ldots, x_n) \equiv (\lambda x_0, \ldots, \lambda x_n)$ for any $\lambda \in K^\times$. A homogeneous polynomial over $K$ of degree $d$ is a polynomial $P(X_0, \ldots, X_n) \in K[X_0, \ldots, X_n]$ such that $P$ satisfies $P(\lambda X_0, \ldots, \lambda X_n) = \lambda^d P(X_0, \ldots, X_n)$ for any $\lambda \in K$. $V$ is a projective variety over $K$ if $V$ is of the form $\{[x_0 : \cdots : x_n] \in \mathbb{P}^n(K) : \bigwedge_{i=1}^m P_i(\mathbf{x}) = 0\}$ for some finite set of homogeneous polynomials $P_i$. The projective varieties form the closed sets of the Zariski topology on $\mathbb{P}^n(K)$. $V$ is said to be a quasi-projective variety if it is of the form $U \cap W$ where $U$ is a Zariski open subset of some $\mathbb{P}^n$ and $W$ is Zariski closed.

*Example* 1.3.30.

- Any projective variety is quasi-projective.
- Affine space $\mathbb{A}^n(K) := K^n$ is quasi-projective. Identify $\mathbb{A}^n$ with $\{[x_0 : \cdots : x_n] \in \mathbb{P}^n(K) : x_0 \neq 0\}$.
- Any affine variety is quasi-projective. Recall that an affine variety is a subset of some power of $K$ defined by the vanishing of some polynomials.

**Definition 1.3.31.** An *algebraic group* is a group $G$ whose underlying set is a variety and whose group structure is given piecewise by polynomials.

*Remark* 1.3.32. Any definable group in ACF is definably isomorphic to an algebraic group (see [3, 26, 54]).

*Example* 1.3.33.

- The additive group $\mathbb{G}_a(K) := (K, +)$ is an algebraic group.
- The multiplicative group $\mathbb{G}_m(K) := (K^\times, \cdot)$ is an algebraic group.
- The general linear group $GL_n(K)$ is an algebraic group.

**Definition 1.3.34.** An algebraic group $G$ is *connected* if $G$ has no proper algebraic subgroup of finite index.

*Remark* 1.3.35. Any algebraic group $G$ is a group of finite Morley rank. The connected component of the identity (in the sense of stability theory) $G^0$ is a connected algebraic group of finite index in $G$. An algebraic group is connected in the above sense if and only if it is connected in the Zariski topology.

**Definition 1.3.36.** A projective connected algebraic group is called an *abelian variety*.

As one might guess from the name, any abelian variety is commutative. However, the name is intended to emphasize the connections with the classical theory of abelian functions and integrals (see [56]). If $K = \mathbb{C}$, then $A(\mathbb{C})$ is a compact complex analytic group and one can invert the exponential function (modulo periods) via the theory of abelian integrals to obtain an isomorphism (of complex analytic groups) $A(\mathbb{C}) \cong \mathbb{C}^g/\Lambda$ where $g = \dim_{\mathbb{C}} A(\mathbb{C})$ and $\Lambda$ is a lattice. That is, $\Lambda \cong \mathbb{Z}^g$ and the $\mathbb{C}^g = \Lambda \otimes \mathbb{C}$. This presentation shows that for any positive integer $n$, if we denote the kernel of multiplication by $n$ on $A$ by $A[n]$, then $A(\mathbb{C})[n] \cong (\mathbb{Z}/n\mathbb{Z})^{2g}$. Since $ACF_0$ is complete, this fact is true with $\mathbb{C}$ replaced by any algebraically closed field of characteristic zero. If $\ell$ is any prime number, we define the $\ell$-Tate module of $A$ by $T_\ell A := \varprojlim A[\ell^n]$. By the above description, $T_\ell A \cong \mathbb{Z}_\ell^{2g}$ where $\mathbb{Z}_\ell$ denotes the $\ell$-adic integers defined by $\mathbb{Z}_\ell := \varprojlim \mathbb{Z}/\ell^n\mathbb{Z}$. If $\sigma$ is an automorphism of $K$ preserving $A$, then $\sigma$ acts naturally on $T_\ell A$. With respect to the profinite topology on $T_\ell A$ and the Krull topology on the Galois group, this action is continuous.

**Definition 1.3.37.** A commutative algebraic group $G$ is called *semi-abelian* if $G$ fits into an exact sequence $0 \longrightarrow T \longrightarrow G \longrightarrow A \longrightarrow 0$ where $A$ is an abelian variety and $T \cong \mathbb{G}_m^r$ for some $r \in \mathbb{N}$.

*Remark* 1.3.38. In general, a semi-abelian variety is neither affine nor projective but merely quasi-projective.

A theorem of Chevalley classifies all commutative algebraic groups.

**Theorem 1.3.39 (Chevalley [14]).** *If $G$ is a connected commutative algebraic group, then $G$ fits into an exact sequence $0 \longrightarrow V \longrightarrow G \longrightarrow S \longrightarrow 0$ where $S$ is a semi-abelian variety and $V$ has a composition series each term of which is isomorphic to $\mathbb{G}_a$.*

On occasion we need to use schemes. Scheme theory is quite deep, but we will need to use only one feature that schemes have but varieties do not: they make sense over general commutative rings rather than just over algebraically closed fields.

Let us restrict attention to affine schemes for now. Recall that an affine variety $V$ is given by the vanishing of some finite set of polynomials $f_1, \ldots, f_m \in K[X_1, \ldots, X_n]$. Certainly, the sequence $f_1, \ldots, f_n$ determines $V$, but how well does $V$ determine the $f_i$'s? That is, for what other $g_1, \ldots, g_\ell$ might we have $\mathrm{Th}(K) \vdash (\forall \mathbf{x})(\bigwedge f_i(\mathbf{x}) = 0) \iff (\bigwedge g_j(\mathbf{x}) = 0)$? A necessary and sufficient condition is that $\sqrt{(f_1, \ldots, f_m)} = \sqrt{(g_1, \ldots, g_\ell)}$.

Let now $R$ be any commutative ring. A closed subscheme of $\mathbb{A}_R^n$ is also given by some set of equations $\{f_i = 0\}_{i \in I}$ where $f_i \in R[X_1, \ldots, X_n]$. In this case we think of $X$ as a functor from the category of $R$-algebras to the category of sets by defining $X(R') := \{\mathbf{x} \in (R')^n : \bigwedge_{i \in I} f_i(\mathbf{x}) = 0\}$ for $R'$ an $R$-algebra. In this case, such a scheme determines the ideal $I(X) = (\{f_i\}_{i \in I})$. If $R$ is a noetherian ring, then $I(X)$ is finitely generated so that two presentations $f_1, \ldots, f_m$ and $g_1, \ldots, g_\ell$ define the same scheme if and only if the theory of $R$-algebras proves that $(\forall \mathbf{x})(\bigwedge f_i(\mathbf{x}) = 0) \iff (\bigwedge g_j(\mathbf{x}) = 0)$.

Presented intrinsically (ie without regard to a particular embedding in some affine space), the scheme $X$ with defining ideal $I(X) \subseteq R[x_1, \ldots, x_n]$ corresponds to the ring $R[x_1, \ldots, x_n]/I(X)$. The scheme (over $R$) corresponding to an $R$-algebra $R'$ is called $\mathrm{Spec} R'$ (the spectrum of $R'$). The correspondence $R' \mapsto \mathrm{Spec} R'$ is contravariant (and by definition) is an equivalence of categories between the category of affine schemes over $R$ and the category of $R$-algebras.

We will not need to know what $\mathrm{Spec} R'$ is, but we describe it now for the curious reader. As a topological space, $\mathrm{Spec} R' := \{\mathfrak{p} \subset R' : \mathfrak{p}$ is prime but is not the unit ideal $\}$ with basic closed sets $V(I) := \{\mathfrak{p} : I \subseteq \mathfrak{p}\}$ where $I$ runs through all the ideals of $R'$. $\mathrm{Spec} R'$ is also given with a sheaf of regular functions. That is, for each open set $U \subseteq \mathrm{Spec} R'$ there is a ring of functions $\mathcal{O}_{\mathrm{Spec} R'}(U)$ and these rings satisfy certain compatibility properties with respect to the open sets. In the case that $R = K$ is an algebraically closed field and $R' = K[x]$, we almost recover the space $\mathbb{A}^1$ considered before as a variety. The points $a \in K$ correspond to the prime ideals $(x - a) \in \mathrm{Spec} R'$. There is one more point, $\eta = (0) \in \mathrm{Spec} R'$ (called the generic point) which corresponds to the generic type in $S_1(K)$.

A general (not necessarily affine) scheme is built from affine schemes in much the same way as a manifold is built from model Euclidean spaces. The reader interested in the details should consult [19].

Our main use of schemes will come from the situation where $K$ is a valued field with ring of integers $R$. For the sake of this discussion, let us assume that the value group of $K$ is an ordered subgroup of $\mathbb{R}$. In this case, Spec$R$ consists of two points: $s = \mathfrak{m}_K$, the special point, and $\eta = (0)$, the generic point. If $\mathfrak{X}$ is a scheme over $R$, denote the reduction of $\mathfrak{X}$ mod $\mathfrak{m}_K$ by $\mathfrak{X}_s$ (called the special fibre) and $\mathfrak{X}$ considered over $K$ by $\mathfrak{X}_\eta$ (called the generic fibre). Each of these may be regarded as fibre products. In later chapters, so as to fit with notation to be introduced, we may write $\mathfrak{X}_s$ as $\mathfrak{X}_0$.

If we have a variety $X$ over $K$, we might like to find a scheme $\mathfrak{X}$ over $R$ so that $\mathfrak{X}_\eta = X$ and $\mathfrak{X}$ has "good" properties. In this case that $X$ is an abelian variety, we would like to find $\mathfrak{X}$ so that the reduction $\mathfrak{X}_s$ is also an abelian variety. Alas, this is not always possible, but one can find a projective $\mathfrak{X}$ so that after removing the singularities of $\mathfrak{X}_s$ one has a semi-abelian variety. (This is the semi-stable reduction theorem of Grothendieck [18].) If the equations for $X$ are given over some field (not algebraically closed), the equations for $\mathfrak{X}$ need not be rational over $K$, but the degree of the field extension needed to find $\mathfrak{X}$ can be bounded in terms of the geometry of $X$.

# Chapter 2

# $\mathcal{D}$-rings and Jet Spaces

## 2.1 Ring Functors and $\mathcal{D}$-structures

In this section we will describe a general framework for analyzing rings with extra structure. This general setting is not strictly necessary, but it unifies many of the constructions done with differential and difference rings and rings given with a $p$-basis.

In what follows, $k$ is some fixed commutative ring and $U$ denotes the forgetful functor $U : \mathbf{Ring}_{/k} \to \mathbf{Set}$ which associates to a $k$-algebra its underlying set.

**Definition 2.1.1.** A $\mathcal{D}$-functor over $k$ of dimension $n$ is a projective system of functors $\{\mathcal{D}_\alpha : \mathbf{Ring}_{/k} \to \mathbf{Ring}_{/k}\}_{\alpha \in \mathbb{N}^n}$ satisfying for any $k$-algebra $R$:

- $U(\mathcal{D}_\alpha(R)) = \prod_{\gamma \leq \alpha} U(R)$.
- $\mathcal{D}_0(R) = R$
- If $\beta \geq \alpha$, then $U(\pi_{\beta,\alpha}) : U(\mathcal{D}_\beta(R)) \to U(\mathcal{D}_\alpha(R))$ is the natural projection $(x_\gamma)_{\gamma \leq \beta} \mapsto (x_\gamma)_{\gamma \leq \alpha}$
- if $\psi : R \to S$ is any map of $k$-algebras, then the map $\mathcal{D}_\alpha(\psi) : \mathcal{D}_\alpha(R) \to \mathcal{D}_\alpha(S)$ is given in co-ordinates by $(x_\beta)_{\beta \leq \alpha} \mapsto (\psi(x_\beta))_{\beta \leq \alpha}$.
- The $k$-algebra structure is given in co-ordinates by $c \cdot (x_\alpha) = (cx_\alpha)$.

We let $\widehat{\mathcal{D}} := \varprojlim \mathcal{D}_\alpha$.

**Definition 2.1.2.** A $\mathcal{D}$-structure on a $k$-algebra $R$ is a section of the projection map $\widehat{\mathcal{D}}(R) \to R$. In co-ordinates, a $\mathcal{D}$-structure on $R$ is given by a family of functions $\partial_\alpha : R \to R$ such that the map $x \mapsto (\partial_\alpha x)_\alpha$ is a ring homomorphism $R \to \widehat{\mathcal{D}}(R)$.

*Example* 2.1.3.

- $n = 1$ and $\mathcal{D}_m(R) = R^{m+1}$. In this case a $\mathcal{D}$-structure is given by a family of endomorphisms $\sigma_i : R \to R$ such that $\sigma_0(x) = x$.
- $\widehat{\mathcal{D}}(R) = R[[X_1, \dots, X_n]]$. In this case a $\mathcal{D}$-structure is given by $n$ stacks of commuting Hasse derivations. When $R$ is a $\mathbb{Q}$-algebra, it is enough to specify $n$ commuting derivations $\partial_1, \dots, \partial_n : R \to R$ and then to set $\partial_\alpha := \frac{1}{\alpha_1! \cdots \alpha_n!} \partial_1^{\alpha_1} \circ \cdots \circ \partial_n^{\alpha_n}$.
- $\widehat{\mathcal{D}}(R) = W_{p^\infty}(R)$, the ring Witt vectors over $R$ with uniformizer $p$. Note that $R$ cannot have a $\mathcal{D}$-structure if $p$ is nilpotent.

*Remark* 2.1.4. Some of the restrictions in the definition of a $\mathcal{D}$-functor can be relaxed. For instance, the functors $\{\mathcal{D}_n\}$ associating to a ring $R$ its matrix ring $M_n(R)$ are ruled out by Definition 2.1.1 because these rings are non-commutative and because the dimensions are not right.

**Definition 2.1.5.** Let $\mathbf{c} = \{c_{\alpha,\beta}\}_{\alpha,\beta\in\mathbb{N}^n}$ be a sequence of elements of $k$ satisfying $c_{\beta,\gamma}\cdot c_{\beta+\gamma,\alpha} = c_{\beta,\alpha}\cdot c_{\gamma,\alpha+\beta}$. For any $k$-algebra $R$ let $\Delta_{\mathbf{c}} : \widehat{\mathcal{D}}(R) \to \widehat{\mathcal{D}}(\widehat{\mathcal{D}}(R))$ be the function given by $(x_\alpha)_{\alpha\in\mathbb{N}^n} \mapsto ((c_{\gamma,\beta}x_{\beta+\gamma})_{\beta\in\mathbb{N}^n})_{\gamma\in\mathbb{N}^n}$. $\mathcal{D}$ is called $\mathbf{c}$-*iterative* if $\Delta_{\mathbf{c}}$ is always a ring homomorphism. A $\mathcal{D}$-ring $(R,\{\partial_\alpha\})$ is $\mathbf{c}-$ *iterative* if the ring functor $\mathcal{D}$ is $\mathbf{c}$-iterative and the $\mathcal{D}$-maps satisfy the functional equations $\partial_\alpha\circ\partial_\beta = c_{\alpha,\beta}\partial_{\alpha+\beta}$.

*Remark* 2.1.6. Let $R$ be a $\mathcal{D}$-ring with the $\mathcal{D}$-structure given by $\varphi : R \to \widehat{\mathcal{D}}(R)$. Then the assertion that the $\mathcal{D}$-structure is $\mathbf{c}$-iterative is equivalent to the commutativity of the following diagram.

$$\begin{array}{ccc} R & \xrightarrow{\varphi} & \widehat{\mathcal{D}}(R) \\ {\scriptstyle\varphi}\downarrow & & \downarrow{\scriptstyle\widehat{\mathcal{D}}(\varphi)} \\ \widehat{\mathcal{D}}(R) & \xrightarrow{\Delta_{\mathbf{c}}} & \widehat{\mathcal{D}}(\widehat{\mathcal{D}}(R)) \end{array}$$

*Example* 2.1.7.

- If $\mathcal{D}_m(R) = R^{m+1}$, then we may take $c_{i,j} = 1$. In this case, a $\mathcal{D}$-ring is $\mathbf{c}$-iterative if and only if $\partial_j = \sigma^j$ where $\sigma : R \to R$ is some ring endomorphism.
- If $\widehat{\mathcal{D}}(R) = R[[X]]$, then we may take $c_{i,j} = \binom{i+j}{i}$.

**Proposition 2.1.8.** *Let $\mathcal{D}$ be $\mathbf{c}$-iterative of dimension $n$. Let $R$ be a $\mathbf{c}$-iterative $\mathcal{D}$-ring. There is a universal $\mathbf{c}$-iterative $\mathcal{D}$-ring, $R\langle X\rangle_{\mathcal{D},\mathbf{c}}$, over $R$ which as a ring is $R[\{X_\alpha\}_{\alpha\in\mathbb{N}^n}]$, the polynomial ring in infinitely many variables indexed by $\mathbb{N}^n$.*

■ Let $R' := R[\{X_\alpha\}_{\alpha\in\mathbb{N}^n}]$. Let $\varphi : R \to \widehat{\mathcal{D}}(R)$ be the map giving the $\mathcal{D}$-structure on $R$. Since $\widehat{\mathcal{D}}$ is a functor, the natural map $R \to R'$ induces a map $\widehat{\mathcal{D}}(R) \to \widehat{\mathcal{D}}(R')$. Let $\varphi' : R \to \widehat{\mathcal{D}}(R')$ be the composite of this map with $\varphi$. By the universality property of the polynomial ring, there is a unique map of rings $\tilde{\varphi} : R' \to \widehat{\mathcal{D}}(R')$ which agrees with $\varphi'$ on $R$ and which sends $X_\alpha \mapsto (c_{\beta,\alpha}X_{\alpha+\beta})_{\beta\in\mathbb{N}^n}$. This shows that $R'$ has a $\mathcal{D}$-structure extending the $\mathcal{D}$-structure on $R$. We need to check that it is $\mathbf{c}$-iterative. That is, we need to check the commutativity of

$$\begin{array}{ccc} R' & \xrightarrow{\tilde{\varphi}} & \widehat{\mathcal{D}}(R') \\ {\scriptstyle\tilde{\varphi}}\downarrow & & \downarrow{\scriptstyle\widehat{\mathcal{D}}(\tilde{\varphi})} \\ \widehat{\mathcal{D}}(R') & \xrightarrow{\Delta_{\mathbf{c}}} & \widehat{\mathcal{D}}(\widehat{\mathcal{D}}(R')) \end{array}$$

Both maps $\Delta_{\mathbf{c}}\circ\tilde{\varphi} : R' \to \widehat{\mathcal{D}}(\widehat{\mathcal{D}}(R'))$ and $\widehat{\mathcal{D}}(\tilde{\varphi})\circ\tilde{\varphi} : R' \to \widehat{\mathcal{D}}(\widehat{\mathcal{D}}(R'))$ have the property that when restricted to $R$ they are equal (because $R$ is $\mathbf{c}$-iterative).

We calculate $\widehat{\mathcal{D}}(\tilde{\varphi})\circ\tilde{\varphi}(X_\alpha)$ and $\Delta_{\mathbf{c}}\circ\tilde{\varphi}(X_\alpha)$. By definition, $\tilde{\varphi}(X_\alpha) = (c_{\beta,\alpha}X_{\alpha+\beta})_{\beta\in\mathbb{N}^n}$. Since $\widehat{\mathcal{D}}(\tilde{\varphi})$ may be computed co-ordinatewise and $\widehat{\mathcal{D}}(R')$ is a $k$-algebra,

$$\begin{aligned} \widehat{\mathcal{D}}(\tilde{\varphi})\circ\tilde{\varphi}(X_\alpha) &= \widehat{\mathcal{D}}(\tilde{\varphi})((c_{\beta,\alpha}X_{\alpha+\beta})_{\beta\in\mathbb{N}^n}) \\ &= ((c_{\beta,\alpha}c_{\gamma,\alpha+\beta}X_{\alpha+\beta+\gamma})_{\gamma\in\mathbb{N}^n})_{\beta\in\mathbb{N}^n} \end{aligned}$$

On the other hand, we compute directly

$$\begin{aligned} \Delta_{\mathbf{c}}\circ\tilde{\varphi}(X_\alpha) &= \Delta_{\mathbf{c}}((c_{\delta,\alpha}X_{\alpha+\delta})_{\delta\in\mathbb{N}^n}) \\ &= ((c_{\beta,\gamma}c_{\beta+\gamma,\alpha}X_{\alpha+\beta+\gamma})_{\gamma\in\mathbb{N}^n})_{\beta\in\mathbb{N}^n} \end{aligned}$$

Since $c_{\beta,\gamma} \cdot c_{\beta+\gamma,\alpha} = c_{\beta,\alpha} \cdot c_{\gamma,\alpha+\beta}$, the expressions for $\Delta_{\mathbf{c}} \circ \tilde{\varphi}(X_\alpha)$ and $\widehat{\mathcal{D}}(\tilde{\varphi}) \circ \tilde{\varphi}(X_\alpha)$ agree. Since these maps agree on the generators of $R'$, they agree everywhere.

As to the universality, if $R \to S$ is any map of $\mathbf{c}$-iterative $\mathcal{D}$-rings and $x \in S$, then there is a unique map of rings $\psi : R' \to S$ given by extending the map on $R$ and sending $X_\alpha$ to $\partial_\alpha(x)$. That $\psi$ is map is a map of $\mathcal{D}$-rings is equivalent to the commutativity of

$$\begin{array}{ccc} S & \longrightarrow & \widehat{\mathcal{D}}(S) \\ {\scriptstyle\psi}\Big\uparrow & & \Big\uparrow{\scriptstyle\widehat{\mathcal{D}}(\psi)} \\ R\langle X\rangle_{\mathcal{D},\mathbf{c}} & \longrightarrow & \widehat{\mathcal{D}}(R\langle X\rangle_{\mathcal{D},\mathbf{c}}) \end{array}$$

Since $R \to S$ is a map of $\mathcal{D}$-rings, this diagram is commutative when restricted to $R$. Since $S$ is $\mathbf{c}$-iterative, the diagram is commutative on the elements $\{X_\alpha\}_{\alpha\in\mathbb{N}^n}$. Hence, the diagram is commutative. ■

We will refer to $R\langle X\rangle_{\mathcal{D},\mathbf{c}}$ as the ring of $\mathbf{c}$-iterative $\mathcal{D}$-polynomials over $R$. We may omit reference to $\mathbf{c}$ and $\mathcal{D}$ if they are understood.

In general, if $R$ is a $\mathcal{D}$-ring and $I \subseteq R$ is an ideal, then $I$ is a $\mathcal{D}$*-ideal* if $\partial_\alpha(I) \subseteq I$ for each $\alpha \in \mathbb{N}^n$. In this case, the $\mathcal{D}$-structure induces such a structure on $R/I$.

If $R$ is a $\mathcal{D}$-ring and $\Sigma \subseteq R$ is a subset, then $\langle\Sigma\rangle$ is the $\mathcal{D}$*-ideal generated by* $\Sigma$. Since the intersection of a set of $\mathcal{D}$-ideals is a $\mathcal{D}$-ideal this notion is well-defined. Concretely, $\langle\Sigma\rangle = (\{\partial_\alpha x\}_{x\in\Sigma,\alpha\in\mathbb{N}^n})$.

Restrict now to the case of $\mathcal{D}$ of dimension one.

Define the *order* and *degree* of a $\mathcal{D}$-polynomial by: A constant polynomial is considered to have order $-1$ and degree $-\infty$. Otherwise, $\mathrm{ord}P := \min\{n : P \in R[X_0, \ldots, X_n]\}$. If $n$ is the order of $P$, then the degree is the degree of $P$ as a polynomial in $D^n X$.

The $\mathcal{D}$-polynomial $P$ is *simpler than* the $\mathcal{D}$-polynomial $Q$, written $P \ll Q$, if in the lexico-graphic order, the order-degree of P is less than that of $Q$.

If $P$ is a $\mathcal{D}$-polynomial of the form $P(X) = F(X, DX, \ldots, D^m X)$, then define $\frac{\partial}{\partial X_i}P$ to be the $\mathcal{D}$-polynomial $\frac{\partial}{\partial X_i}F$.

We may define a more refined degree: the *total degree*. $\mathrm{T.deg}P := (\deg_{X_i} P)_{i=0}^\infty$. Notice that the image of T. deg comprises the set $\mathbb{N}^{<\omega} := \{(n_j)_{j=0}^\infty : n_j \in \mathbb{N},\, n_j = 0 \text{ for } j \gg 0\}$. Define an ordering on $\mathbb{N}^{<\omega}$ by $(n_j)_{j=0}^\infty < (m_j)_{j=0}^\infty$ iff there is some $N$ such that $n_N < m_N$ and $n_j \le m_j$ for $j > N$. Observe that this ordering is a well-ordering of $\mathbb{N}^{<\omega}$. We define $P \prec Q$ if $\mathrm{T.deg}P < \mathrm{T.deg}Q$. The fact that the ordering on $\mathbb{N}^{<\omega}$ is a well-ordering means that we can (and will) argue by induction with respect to $\prec$.

$\prec$ has the properties

- $\frac{\partial}{\partial X_i}P \preceq P$
- If $P \prec Q$ and $P'$ and $Q'$ differ from $P$ and $Q$ respectively by linear changes of variables – that is, if $P(X) = F(\{\partial_\alpha X\})$ and $Q(X) = G(\{\partial_\alpha X\})$, then $P'(X) = F(\{a_\alpha \partial_\alpha X + b_\alpha\})$ and $Q'(X) = G(\{a'_\alpha \partial_\alpha X + b'_\alpha\})$ with $a_\alpha, a'_\alpha, b_\alpha, b'_\alpha \in k$ and $a_\alpha, a'_\alpha \neq 0$ – then $P' \prec Q'$.

## 2.2 The Ring Functor $\mathcal{D}_e$

In this section we introduce a special ring functor, $\mathcal{D}_e$, which fits into the framework of the previous section. A $\mathcal{D}$-structure relative to $\mathcal{D}_e$ generalizes differential and endomorphism structure.

$\mathcal{D}_e$ is a ring functor over $\mathbb{Z}[e]$, the polynomial ring over $\mathbb{Z}$ in the indeterminate $e$.

Let $R$ be a commutative $\mathbb{Z}[e]$-algebra. $\mathcal{D}_e(R)$ is the ring which as an abelian group is $R^2$ with multiplication defined by $(x_1, x_2) * (y_1, y_2) := (x_1 y_1, x_1 y_2 + y_1 x_2 + e x_2 y_2)$.

A $\mathcal{D}$-structure on $R$ with respect to $\mathcal{D}_e$ consists of an additive map $\partial : R \to R$ satisfying the twisted Leibniz rule $\partial(xy) = x\partial y + y\partial x + e(\partial x)(\partial y)$ and $\partial(1) = 0$.

*Remark* 2.2.1. From the standpoint of logic, the restriction to $\mathbb{Z}[e]$-algebras corresponds simply to adding a constant symbol $e$ to the language of rings.

*Remark* 2.2.2. Note that when $e = 0$ in $R$, then a $\mathcal{D}_e$-structure on $R$ is simply a derivation.

**Proposition 2.2.3.** *Let $(R, D)$ be a $\mathcal{D}_e$-ring. The function $\sigma : R \to R$ defined by $x \mapsto eDx + x$ is a ring endomorphism of $R$.*

■ Linearity is clear as is the fact that $\sigma(1) = 1$. For multiplication:

$$\begin{aligned} \sigma(xy) &= eD(xy) + xy \\ &= e(xDy + yDx + eDxDy) + xy \\ &= e^2 DxDy + exDy + eyDx + xy \\ &= (eDx + x)(eDy + y) \end{aligned}$$

■

*Remark* 2.2.4. If $e$ is a non-zero divisor, then $\sigma$ determines $\mathcal{D}_e$. So when $e$ is a unit, a $\mathcal{D}_e$-ring is just a difference ring in disguise. That is, if $e$ is a non-zero divisor, then $D$ and $\sigma$ are inter-definable. If $e$ is a unit and one includes $e^{-1}$ as a constant, then $D$ is term definable from $\sigma$ as $Dx = \frac{\sigma(x)-x}{e}$.

Of course, for fields $e$ being zero or a unit exhausts the possibilities, but for more general rings there is an intermediate case.

*Remark* 2.2.5. The Leibniz rule may also be written as $\partial(xy) = x\partial y + \sigma(y)\partial x$.

**Proposition 2.2.6.** *If $x \in R^\times$, then $D(\frac{1}{x}) = \frac{-Dx}{x\sigma(x)}$.*

■

$$\begin{aligned} 0 &= D(1) \\ &= D(xx^{-1}) \\ &= xD(x^{-1}) + Dx(x^{-1}) + eDxD(x^{-1}) \end{aligned}$$

Subtracting, we find that $(x + eDx)D(x^{-1}) = -Dx(x^{-1})$. As $\sigma(x) = x + eDx$ and $x \in R^\times$, we also have $\sigma(x) \in R^\times$. Therefore, $D(\frac{1}{x}) = \frac{-Dx}{\sigma(x)x}$. ■

**Proposition 2.2.7.** *If $R$ is a $\mathcal{D}_e$-ring and $S \subseteq R$ is a multiplicative subset of $R$ containing $1$ and is closed under $\sigma$, then there is a unique structure of a $\mathcal{D}_e$-ring on the localization $S^{-1}R$.*

■ Proposition 2.2.6 shows how $\mathcal{D}_e$ must be defined. The original $\mathcal{D}_e$-structure on $R$ corresponds to a map $R \xrightarrow{\varphi} \mathcal{D}_e(R)$. By functoriality of $\mathcal{D}_e$, there is a map $\mathcal{D}_e(R) \xrightarrow{\mathcal{D}_e(i)} \mathcal{D}_e(S^{-1}R)$. For any $s \in S$ there is an inverse to $i \circ \varphi(s) = (s, Ds)$ in $\mathcal{D}_e(R)$. Simply take $(\frac{1}{s}, \frac{-Ds}{\sigma(s)s})$. By the universal property of $S^{-1}R$, there is a unique ring homomorphism $S^{-1}R \longrightarrow \mathcal{D}_e(S^{-1}R)$ making the following diagram commute.

$$\begin{array}{ccc} R & \xrightarrow{\varphi} & \mathcal{D}_e(R) \\ {\scriptstyle i}\downarrow & & \downarrow{\scriptstyle \mathcal{D}_e(i)} \\ S^{-1}R & \xrightarrow{\exists!} & \mathcal{D}_e(S^{-1}R) \end{array}$$

■

Let us also calculate $Dx^n$.

**Proposition 2.2.8.** *If $R$ is a $\mathcal{D}_e$-ring, $x \in R$, and $n$ is a positive integer, then*

$$Dx^n = \sum_{i=1}^{n} \binom{n}{i} e^{i-1} x^{n-i} (Dx)^i$$

■ We check this by induction on $n$. For $n = 1$ the assertion is obvious. Let us now try the case of $n + 1$.

$$\begin{aligned}
Dx^{n+1} &= D(x^n x) \\
&= x^n Dx + x(Dx^n) + eDx(Dx^n) \\
&= x^n Dx + (x + eDx) \sum_{i=1}^{n} \binom{n}{i} e^{i-1} x^{n-i} (Dx)^i \\
&= \sum_{i=1}^{n} \binom{n}{i} e^i x^{n+1-i} (Dx)^i + \sum_{j=1}^{n} \binom{n}{j} e^{j-1} x^{n-j} (Dx)^{j+1} \\
&= (n+1)x^n Dx + \sum_{j=2}^{n} (\binom{n}{j} + \binom{n}{j+1}) e^{j-1} x^{n+1-j} (Dx)^i \\
&= \sum_{j=1}^{n+1} \binom{n+1}{j} e^{j-1} x^{n+1-j} (Dx)^j
\end{aligned}$$

■

**Lemma 2.2.9.** *Let $R$ be a local ring with maximal ideal $\mathfrak{m}$. Assume that $e \in \mathfrak{m}$. Then $\mathcal{D}_e(R)$ is also a local ring with maximal ideal $\pi_0^{-1}\mathfrak{m}$.*

■ Let $(x, y) \in \mathcal{D}_e(R) \setminus \pi_0^{-1}\mathfrak{m}$. That is, $x \in R^\times$. Since $e \in \mathfrak{m}$, $x + ey \in R^\times$ as well. The inverse to $(x, y)$ is then $(\frac{1}{x}, \frac{-y}{x(x+ey)})$. ■

**Proposition 2.2.10.** *If $(R, \mathfrak{m})$ is a henselian local ring and $e \in \mathfrak{m}$, then $(\mathcal{D}_e(R), \pi_0^{-1}(\mathfrak{m}))$ is also henselian.*

■ Let us denote the reduction map $R \to R/\mathfrak{m}$ by $x \mapsto \overline{x}$. Let $P(X) \in \mathcal{D}_e(R)[X]$ and let $(x, y) \in \mathcal{D}_e$ such that $\overline{x}$ is a simple root of $\overline{\pi_0(P)}(X)$. Since $R$ is henselian, there is a unique $a \in R$ such that $\pi_0(P)(a) = 0$ and $\overline{a} = \overline{x}$. Let $\epsilon := (0, 1)$. Since $\pi_0(P)(a) = 0$, there is some $Q(Y) \in R[X]$ with

$$P(a + \epsilon Y) := \epsilon Q(Y)$$

Taylor expanding $P(a + \epsilon Y)$, we compute that the linear term of $Q(Y)$ is $\pi_0(P'(a))$ and that all the higher order terms involve $e$ as a factor. Hence, $\overline{Q}$ is a linear polynomial and therefore has a unique solution in $R/\mathfrak{m}$. As $R$ is henselian, there is a unique lifting of this solution to some $b \in R$. $(a, b)$ is then the unique solution to $P(X) = 0$ with $\overline{a} = \overline{x}$. ■

## 2.3 Jet Spaces

In this section we will consider some of the properties of jet spaces associated to the ring functors considered in Section 2.1.

Let $R$ be a $\mathcal{D}$-ring via the map $\varphi : R \to \widehat{\mathcal{D}}(R)$. Let $X$ be any scheme over $R$. We define a scheme $X_{\varphi,\alpha}$ over $\mathcal{D}_\alpha(R)$ by the Cartesian square

$$\begin{array}{ccc} X_{\varphi,\alpha} & \longrightarrow & X \\ \downarrow & & \downarrow \\ \mathrm{Spec}\mathcal{D}_\alpha(R) & \xrightarrow{\varphi^*} & \mathrm{Spec}R \end{array}$$

We define a projective family of functors $\nabla_\alpha : \mathbf{Schemes}_{/R} \times \mathbf{Ring}_{/R} \to \mathbf{Sets}$ by $(X, R') \mapsto X_{\varphi,\alpha}(\mathcal{D}_\alpha(R'))$.

If we restrict to reasonable $X$ the functors $\nabla_\alpha X$ are representable (and we will denote the representing object by $\nabla_\alpha X$ as well). To use the language standard in the literature, the question of the representability of $\nabla_\alpha X$ is simply the question of whether the Weil restriction of $X_{\varphi,\alpha}$ from $\mathcal{D}_\alpha(R)$ to $R$ exists. It is known (see Theorem 4 of Section 7.6 of [2]) that it does if $X$ is quasi-projective.

So we have

**Proposition 2.3.1.** *Let $R$ be a $\mathcal{D}$-ring with $\mathcal{D}$-structure given by $\varphi : R \to \widehat{\mathcal{D}}(R)$. There is a projective family of functors $\{\nabla_\alpha\}$ which assigns to a quasi-projective $R$-scheme $X$ another quasi-projective $R$-scheme $\nabla_\alpha X$ so that there is a natural correspondence between $\nabla_\alpha X(R)$ and $X_{\varphi,\alpha}(\mathcal{D}_\alpha(R))$.*

The scheme $\nabla_\alpha X$ is called the $\alpha$-th jet space of $X$.

Before discussing some of the properties of the jet spaces, let us consider a couple of other ways of thinking about them.

For now let us consider the case that $X$ is a closed subvariety of $\mathbb{A}^n$. The characteristic feature of $\nabla_\alpha X$ is that if $x = (x_1, \ldots, x_n) \in X(R)$ then $\nabla_\alpha(x) = (\partial_\beta x_j)_{j=1,\beta=0}^{n,\alpha} \in \nabla_\alpha X(R)$. One can use this property to define the jet spaces of varieties over suitable fields (see [39]). For instance, if $K$ is a differentially closed field of characteristic zero and $X$ is closed subvariety of $\mathbb{A}^n_K$ then one can take for $\nabla_\alpha X$ the the Zariski closure of $\{(\frac{1}{i!}\partial^i x_j)_{i=0,j=1}^{m,n} \in K^{n(m+1)} : (x_1, \ldots, x_n) \in X(K)\}$. Suppose that $X = V(f_1, \ldots, f_m)$. To say that $x$ is in $X(R)$ is the same as to say that $f_i(x) = 0$ for each $i$ in the range $[1, m]$. From these equations we can start to deduce equations for $\partial_\beta x$ by looking at $\partial_\beta f_i(x) = 0$. Certainly, these equations must all be satisfied by points in $\nabla_\alpha X$.

Suppose now the $R$ is a $\mathbf{c}$-iterative $\mathcal{D}$-ring. Let $S := R[x_1, \ldots, x_n]/(f_1, \ldots, f_m)$ and $X := \mathrm{Spec}S$. Let $S_\infty := R\langle x_1, \ldots, x_n\rangle_{\mathcal{D},\mathbf{c}}/\langle f_1, \ldots, f_m\rangle$. Let $S_\alpha := S_\infty \cap R[\{x_i^{(\beta)}\}_{i=1,\beta\leq\alpha}^n]$. Then $\nabla_\alpha X = \mathrm{Spec}S_\alpha$.

This way of viewing the construction of the jet spaces suggests another construction which we will not pursue here. One can define the notion of a $\mathcal{D}$-scheme where the basic objects are of the form $\mathrm{Spec}^{\mathcal{D}} S := \{\mathfrak{p} \in \mathrm{Spec}S : \mathfrak{p} \text{ is a } \mathcal{D}-\text{ideal }\}$ for $S$ a $\mathcal{D}$-ring. The schemes $\nabla_\alpha X$ would then be approximations to the underlying scheme of the affine $\mathcal{D}$-scheme $\mathrm{Spec}^{\mathcal{D}} S$. Buium considers the case of differential schemes in depth in [7]. Some of the theory of $\mathcal{D}$-schemes associated to $\widehat{\mathcal{D}}(R) = R^\omega$ is developed by Hrushovski in [27].

*Example* 2.3.2.

- If $\widehat{\mathcal{D}}(R) = R[[\epsilon]]$ and $R$ is a ring with a trivial derivation, then for any scheme $X$ over $R$, the first jet space of $X$ is simply the tangent bundle $TX$. For more general $R$, $\nabla_1 X$ is a torsor of $TX$.

- If $\widehat{\mathcal{D}}(R) = R^\omega$ and $R$ is a $\mathcal{D}$-ring with the $\mathcal{D}$-structure determined by an endomorphisms $\sigma_i : R \to R$, then for any scheme $X$ over $R$, $\nabla_n X = X \times X^{\sigma_1} \times \cdots \times X^{\sigma_n}$.

*Remark* 2.3.3. The jet spaces have been introduced as a way to transform $\mathcal{D}$-equations into algebraic equations. However, unless we are working with a variety of $\mathcal{D}$-rings in which all $\mathcal{D}$-equations are equivalent to systems of $\mathcal{D}$-equations of the form $F(\{\partial_\alpha X\}) = 0$ for $F$ some polynomial, then the jet spaces are not fine enough to describe all $\mathcal{D}$-equations. We will use the jet spaces only for iterative $\mathcal{D}$-rings thereby avoiding this issue.

Let us list a few basic properties of the jet spaces that we will use in the sequel. In all of the following we assume that the schemes considered are quasi-projective over the $\mathcal{D}$-ring $R$.

**Lemma 2.3.4.** *If $G$ is a group scheme, then so is $\nabla_\alpha G$.*

■ Indeed, since $G$ is a group scheme the functor from $\mathbf{Ring}_{/R} \to \mathbf{Set}$ defined by $R' \mapsto G_{\varphi,\alpha}(\mathcal{D}_\alpha(R')) = \nabla_\alpha(R')$ actually is a functor into the category of groups. As $\nabla_\alpha G$ represents this functor, it a group scheme. ■

**Lemma 2.3.5.** *There is a natural map $\nabla_\alpha : X(R) \to \nabla_\alpha X(R)$.*

■ One sees this by considering

$$\begin{array}{ccccc} X & \xleftarrow{P} & \mathrm{Spec}R & \xleftarrow{\varphi^*} & \mathrm{Spec}\mathcal{D}_\alpha(R) \\ \downarrow & & & // & \\ \mathrm{Spec}R & \xleftarrow{\varphi^*} & \mathrm{Spec}\mathcal{D}_\alpha(R) & & \end{array}$$

from which one deduces the existence of a map to the fibre product $X_{\varphi,\alpha}$. ■

**Lemma 2.3.6.** *If $G$ is a group scheme, then the map $\nabla_\alpha : G(R) \to \nabla_\alpha G(R)$ is a group homomorphism.*

■ This follows from the naturality of the map $\nabla_\alpha$. ■

# Chapter 3

# A Model Complete Theory of Valued $D$-Fields

## 3.1 Notation and Conventions

The models under consideration have three sorts $(K, k, \Gamma)$.

- $K$ is the valued field given with the signature of a $\mathcal{D}_e$-ring: $(+, \cdot, -, 0, 1, e, D)$
- $k$ is the residue field also given with the signature of a $\mathcal{D}_e$-field and possibly with some extra predicates.
- $\Gamma$ is the value group given with the signature of an ordered abelian group with divisibility predicates and possibly with some extra predicates: $(+, -, 0, \leq, \{n|\}_{n=1}^{\infty})$.

For convenience, an extra symbol $\infty$ is added to the language. For instance, one defines $0^{-1} = \infty$ and $(\forall \gamma \in \Gamma)\gamma < \infty$. The sorts are connected by functions $\pi : K \to k \cup \infty$ (the residue map) and $v : K \to \Gamma \cup \infty$ (the valuation). $\mathcal{L}$ denotes the first-order language described above.

As in Chapter 2, we define $\sigma(x) := eDx + x$.

We define $K^D := \ker(D : K \to K)$.

In this chapter, we will refer to a $\mathcal{D}_e$-ring just as a $D$-ring.

Throughout this chapter, if $R$ is a $D$-ring, then $R\langle X\rangle$ denotes the ring of $D$-polynomials over $R$. If $L/K$ is an extension of $D$-fields and $a \in L$ then $K\langle a\rangle$ is the $D$-subfield of $L$ generated by $K$ and $a$.

If $M$ is an $\mathcal{L}$-structure and $P$ is a predicate then we denote the realization of $P$ in $M$ by either $P_M$ or $P(M)$. The particular choice of notion in this case is purely aesthetic. If $P$ is a particular sort, then $S_{m,P}(A)$ denotes the space of $m$-types over $A$ in the sort $P$. That is, each $p(x_1, \ldots, x_m) \in S_{m,P}(A)$ must contain the formula $\wedge_{1 \leq i \leq m} P(x_i)$. It will be proven that in the cases of $P = \Gamma$ or $k$ that $A$ may be replaced by $P_A$ when $A$ is an $\mathcal{L}$-substructure of a model of the theory described in Section 3.2.

## 3.2 Axioms

We will restrict the models considered to those with a differential field of characteristic zero as residue field. The more general cases of positive residual characteristic or a difference field as residue field present technical problems. In particular, in these cases one cannot have quantifier elimination in the language $\mathcal{L}$ described above.

Let $\mathbf{k}$ be a differential field and $\mathbf{G}$ an ordered abelian group. We assume that $\mathbf{k}$ satisfies

1. $\operatorname{char} \mathbf{k} = 0$,
2. $(\mathbf{k}^\times)^n = \mathbf{k}^\times$ for each $n \in \mathbb{Z}_+$, and
3. any non-zero linear differential operator $L \in \mathbf{k}[D]$ is surjective as a map $L : \mathbf{k} \to \mathbf{k}$. We call a differential field satisfying this condition *linearly differentially closed.*

We also assume that enough predicates have been added to $\mathcal{L}$ on the sort $k$ so that $\mathrm{Th}(\mathbf{k})$ admits elimination of quantifiers. Of course, one should take the language to be as simple as possible. It is currently unknown whether there are any non-trivial differential fields other than differentially closed fields which admit elimination of quantifiers in the language of differential rings (see [41]).

We also assume that the language for $\Gamma$ is sufficiently rich so that $\mathrm{Th}(\mathbf{G})$ admits elimination of quantifiers. In many cases of interest one may achieve this by adding divisibility predicates defined by $n|x \iff (\exists y)\,\overbrace{y + \cdots + y}^{n \text{ times}} = x$ [58]. In general, more complicated predicates may be needed.

If you would rather not worry about the relative case, then simply take $\mathbf{k} \models \mathrm{DCF}_0$ and $\mathbf{G} = \mathbb{Q}$.

The first axioms describe general valued $(\mathbf{k}, \mathbf{G})$-$D$-fields.

**Axiom 1.** *$K$ and $k$ are $D$-fields of characteristic zero and $k \models \mathrm{Th}_\forall(\mathbf{k})$.*

**Axiom 2.** *$K$ is valued field whose value group is a subgroup of $\Gamma$ via the valuation $v$ and whose residue field is a subfield of $k$ via the residue map $\pi$ and $v(e) > 0$.*

**Axiom 3.** $(\forall x \in K)\ v(Dx) \geq v(x)$

**Axiom 4.** $\Gamma \models \mathrm{Th}_\forall(\mathbf{G})$

The next six axioms together with the first four describe $(\mathbf{k}, \mathbf{G})$-$D$-henselian fields.

**Axiom 5.** $(\forall x \in K)[([\exists y \in K]\ y^n = x) \iff n|v(x)]$

**Axiom 6.** $\Gamma = v((K^D)^\times)$

**Axiom 7.** $k = \pi(\mathcal{O}_K)$

**Axiom 8 ($D$-Hensel's Lemma).** *If $P \in \mathcal{O}_K\langle X\rangle$ is a $D$-polynomial, $a \in \mathcal{O}_K$, and $v(P(a)) > 0 = v(\frac{\partial}{\partial X_i} P(a))$ for some $i$, then there is $b \in K$ with $P(b) = 0$ and $v(a - b) \geq v(P(a))$.*

If the hypotheses of the last axiom apply to $P$ and $a$, then one says that DHL applies to $P$ at $a$.

**Axiom 9.** $\Gamma \equiv \mathbf{G}$

**Axiom 10.** $k \equiv \mathbf{k}$

*Remark* 3.2.1. Axiom 2 includes the hypothesis that $D$ is a derivation on $k$.

We assumed that $\mathbf{k}$ is linearly differentially closed and is closed under roots in order to guarantee consistency of the theory of $(\mathbf{k}, \mathbf{G})$-$D$-henselian fields.

**Proposition 3.2.2.** *Axioms 1 - 10 together with $\mathbf{G} \neq 0$ imply that $\mathbf{k}$ is linearly differentially closed.*

■ Let $K$ be a $(\mathbf{k}, \mathbf{G})$-$D$-henselian field. Let $L(X) = \sum_{i=1}^n a_i D^i X$ be a non-zero linear $D$-polynomial over $k$. Let $y \in k$ be given. By Axiom 7 there are $b_i \in \mathcal{O}_K$ such that $\pi(b_i) = a_i$ and $z \in \mathcal{O}_K$ such that $\pi(z) = y$. Since $\mathbf{G} \neq 0$, by Axiom 6 there is $\epsilon \in \mathcal{O}_K$ with $D\epsilon = 0$ and $v(\epsilon) > 0$. Let $P(X) = -\epsilon \cdot z + \sum_{i=1}^n b_i D^i X$.

$v(P(0)) > 0$ and for some $i$ we have $\pi(\frac{\partial}{\partial X_i}P) = \pi(b_i) = a_i \neq 0$ so by DHL there is some $x \in \mathcal{O}_K$ such that $P(x) = 0$ and $v(x) \geq v(P(0)) = v(\epsilon)$. Let $x' = \frac{x}{\epsilon}$. We have

$$\begin{aligned} 0 &= P(x) \\ &= -\epsilon z + \sum_{i=1}^{n} b_i D^i x \\ &= \epsilon(-z + \sum_{i=1}^{n} b_i D^i x') \end{aligned}$$

Hence, $z = \sum_{i=1}^{n} b_i D^i x'$. Applying $\pi$, we find that $y = L(\pi(x'))$. ■

**Proposition 3.2.3.** *Axioms 1 - 10 imply that $(\mathbf{k}^\times)^n = \mathbf{k}^\times$ for each positive integer $n$.*

■ Let $K$ be a $(\mathbf{k}, \mathbf{G})$-$D$-henselian field. Let $x \in k^\times$. By Axiom 7 there exists $y \in \mathcal{O}_K$ such that $\pi(y) = x$. $v(y) = 0$ so $n|v(y)$ which implies by Axiom 5 that $y$ has an $n$-th root $z$. Thus, $\pi(z)$ is an $n$-th root of $x$. ■

## 3.3 Consistency and the Standard $D$-Henselian Fields

The generalized power series fields $\mathbf{k}((\epsilon^{\mathbf{G}}))$ provide canonical models for the theory of $D$-henselian fields. For the reader's convenience, we recall the definition of these fields.

As a set, $\mathbf{k}((\epsilon^{\mathbf{G}})) = \{f : \mathbf{G} \to \mathbf{k} : \mathrm{supp}(f) := \{x \in \mathbf{G} : f(x) \neq 0\}$ is well-ordered in the ordering induced by $\mathbf{G}\}$.

We think of an element $f \in \mathbf{k}((\epsilon^{\mathbf{G}}))$ as a formal power series

$$\begin{aligned} f &\leftrightarrow \sum_{\gamma \in \mathbf{G}} f(\gamma)\epsilon^\gamma \\ v(f) &:= \min \mathrm{supp}(f) \\ (f+h)(\gamma) &:= f(\gamma) + h(\gamma) \\ (fh)(\gamma) &:= \sum_{\alpha+\beta=\gamma} f(\alpha)h(\beta) \end{aligned}$$

If we wish to have $e = 0$, then define

$$(Df)(\gamma) = D(f(\gamma))$$

Otherwise, we define an endomorphism $\sigma$ with the property that $\sigma(x) = eDx + x$. Define on $\mathbf{k}$,

$$\sigma(x) := \sum_{n=0}^{\infty} \frac{D^n x}{n!} e^n$$

Extend to all of $\mathbf{k}((\epsilon^{\mathbf{G}}))$ by the rule

$$\sigma(f) = \sum \sigma(f(\gamma))\epsilon^\gamma$$

$\mathbf{k}((\epsilon^{\mathbf{G}}))$ is a maximally complete valued field [38, 46].

That this field is a $D$-henselian field is clear except perhaps for DHL. We prove DHL for $K := \mathbf{k}((\epsilon^{\mathbf{G}}))$ in a *prima facie* stronger form.

**Proposition 3.3.1.** *If $P \in \mathcal{O}_K\langle X\rangle$ is a $D$-polynomial, $a \in \mathcal{O}_K$, and $v(P(a)) > 2v(\frac{\partial}{\partial X_i}P(a))$ for some $i$, then there is $b \in K$ such that $P(b) = 0$ and $v(a-b) \geq v(P(a)) - v(\frac{\partial}{\partial X_i}P(a))$.*

■ Inductively build an ordinal indexed Cauchy sequence of approximate solutions $\{x_\alpha\}$ from $K$. If at some point $P(x_\alpha) = 0$, stop. At each point in the construction ensure that $(\forall \beta < \alpha)\ v(P(x_\alpha)) > v(x_\alpha - x_\beta) > 2v(\frac{\partial}{\partial X_i} P(a))$ and that $v(x_{\alpha+1} - x_\alpha) = v(P(x_\alpha))$. By starting with $x_0 = a$, provided that one may construct the sequence so as to be cofinal in $\mathbf{G}$, the result is proven.

At limits, simply find any $x_\lambda$ such that $v(x_\lambda - x_\alpha) < v(x_\lambda - x_\beta)$ for $\alpha < \beta < \lambda$. Such exists by completeness of $K$. Without loss of generality we may assume that $i$ was chosen so as to minimize $\gamma_0 := v(\frac{\partial}{\partial X_i} P(a))$. For each $j$, choose $c_j \in \mathbf{k}$ such that $v(c_j \epsilon^{\gamma_0} - \frac{\partial}{\partial X_j} P(a)) > \gamma_0$.

At successor stages, $\alpha + 1$, try to modify $x_\alpha$ slightly so as to increase the value of $P$. Let $\gamma = v(P(x_\alpha)) - \gamma_0$. Consider the expansion:

$$P(X\epsilon^\gamma + x_\alpha) \quad = \quad \sum_{m \geq 0} \sum_{|I|=m} \partial_I P(x_\alpha) \epsilon^m X^I$$

Every coefficient on the right hand side has value $\geq \gamma + \gamma_0$. For the constant term, this is because $\gamma + \gamma_0 = v(P(x_\alpha))$ by definition. For the linear terms, one knows that each of $v(\frac{\partial}{\partial X_j} P(x_\alpha))$ is at least $\gamma_0$. For the higher order terms, note that $n\gamma \geq 2\gamma = 2v(P(x_\alpha)) - 2\gamma_0 > v(P(x_\alpha)) + 2\gamma_0 - 2\gamma_0 = \gamma + \gamma_0$. (The strict inequality follows from the fact that $v(P(x_\alpha)) > 2\gamma_0$.) Thus, we may divide the right hand side by $\epsilon^{\gamma+\gamma_0}$ and still have a $D$- polynomial with integral coefficients. In the residue field, the equation is

$$\pi(\frac{P(X\epsilon^\gamma + x_\alpha)}{\epsilon^{\gamma+\gamma_0}}) \quad = \quad \pi(\frac{P(x_\alpha)}{\epsilon^{\gamma+\gamma_0}}) + \sum c_j X^{(j)} \tag{3.1}$$

As $c_i \neq 0$, Equation 3.1 is a non-trivial inhomogeneous linear $D$-equation over $\mathbf{k}$. As $\mathbf{k}$ is linearly differentially closed we may find $x$ which is a solution to this equation and set $x_{\alpha+1} = x\epsilon^\gamma + x_\alpha$.

Let $b = \lim x_\alpha$. ■

## 3.4 Quantifier Elimination

**Theorem 3.4.1.** *With the restrictions imposed on* $\mathbf{k}$ *and* $\mathbf{G}$ *in Section 3.2, the theory of henselian* $(\mathbf{k}, \mathbf{G})$*-$D$-fields is complete, eliminates quantifiers and is the model completion of the theory of valued* $(\mathbf{k}, \mathbf{G})$*-$D$-fields.*

■ We prove this by a standard back-and-forth. For the completeness and quantifier elimination assertions we will prove the following technical claim.

**Claim 3.4.2.** *Assume the continuum hypothesis ($2^{\aleph_0} = \aleph_1$). Let $M_1$ and $M_2$ be two saturated* $(\mathbf{k}, \mathbf{G})$*-$D$-henselian fields of cardinality $\aleph_1$. Let $A_1 \subset M_1$ and $A_2 \subset M_2$ be countable substructures. Let $f : A_1 \to A_2$ be an isomorphism of $\mathcal{L}$-structures. Then $f$ extends to an isomorphism $f : M_1 \to M_2$.*

Before proceeding, let us check that Claim 3.4.2 does in fact prove Theorem 3.4.1. A similar absoluteness argument appears in [1].

**Claim 3.4.3.** *The property of a theory $T$ in a countable language $\mathcal{L}$ eliminating quantifiers is absolute (ie if it holds in one well-founded model of ZFC containing $T$, then it holds in all such).*

✠ As $\mathcal{L}$ is countable, we may encode every $\mathcal{L}$ formula by a natural number and thereby encode all proofs via natural numbers as well. If $\mathfrak{M} \models$ ZFC is a well-founded model of set theory in which $T$ eliminates quantifiers, then for any $\mathcal{L}$-formula $\phi(\mathbf{x})$ there is some quantifier-free formula $\psi(\mathbf{x})$ such that $\mathfrak{M} \models \ulcorner T \vdash \psi(\mathbf{x}) \iff \phi(\mathbf{x}) \urcorner$. So in $\mathfrak{M}$ there is a proof that $\psi$ and $\phi$ are equivalent over $T$. This proof is described by a natural number (in $\mathfrak{M}$), but because $\mathfrak{M}$ is well-founded the proof is encoded by an actual natural number. Since the natural numbers are interpreted the same way in every well-founded model of ZFC, there is a proof that $\psi(\mathbf{x})$ is equivalent to $\phi(\mathbf{x})$ in every model of ZFC. ✠

$\mathbf{L}(T)$, the constructible universe over $T$ is one well-founded model of ZFC + CH containing $T$ (see [17]). So it suffices to work in a universe where the continuum hypothesis is true.

We use the continuum hypothesis to produce a saturated model of size $\aleph_1$. (If $2^{\aleph_0} \neq \aleph_1$, then there need not be a saturated model of size $\aleph_1$ for if $\mathbf{G} \neq 0$, there is a model of $\mathrm{Th}(\mathbf{G})$ containing a countable set over which there are continuum many 1-types.) As to quantifier elimination, we take $M_1 = M_2$ any saturated model of the theory of $(\mathbf{k}, \mathbf{G})$-$D$-henselian fields and $A_1 = A_2$ any countable substructure. In this case, the claim boils down to substructure completeness which implies quantifier elimination and thus model completeness [10]. For completeness, we use the fact that $\mathbb{Q}$ given with the trivial valuation and the trivial $D$-structure is a substructure of any valued $D$-field. So model completeness implies completeness in this case [10].

We will prove that the theory of $D$-henselian fields is the model completion of the theory of valued $D$-fields by showing that each of the constructions used in the proof of Claim 3.4.2 may actually be used to extend any valued $D$-field to $D$-henselian field.

Our strategy for the proof of Claim 3.4.2 is to enlarge $A_i$ a little bit at a time so as to exhaust all of $M_i$. We start by enlarging $A_i$ so that one has $n|\gamma \iff \gamma = n\delta$ for some $\delta \in \Gamma_{A_i}$ for each $\gamma \in \Gamma_{A_i}$. We then expand so that $A_i$ has enough constants. At this point we enlarge $A_i$ so that the residue map is surjective onto the residue field of $M_i$ (N.B.: This is $M$, not $A$.) This step will increase the cardinality of $A_i$ but not of its value group. From this point onward the induction works assuming only the countability of the value group of $A_i$.

We then expand $A_i$ to a maximal immediate extension inside $M_i$. This step will require most of the work. We then add a new element to the value group of $A_i$ and repeat the above steps. Provided that we choose the elements of the value groups so as to exhaust $\Gamma_{M_i}$, this procedure will produce an isomorphism after $\omega_1$ steps. ■

### 3.4.1 Extensions of $k$ and of $\Gamma$

**Definition 3.4.4.** The valued $D$-field $(K, \Gamma)$ *has enough constants* if it satisfies Axiom 6.

**Lemma 3.4.5.** *If $K$ is a valued $D$-field with enough constants, then for any value $\gamma \in v(K^\times)$ and any finite set of polynomials $Q_1(X), \dots, Q_n(X) \in K[X]$ with $Q_i(X) = \sum_{j=0}^{m} q_{i,j} X^j$ there is some $\epsilon \in K$ with $v(\epsilon) = \gamma$, $D\epsilon = 0$, and $v(Q_i(\epsilon)) = \min\{v(q_{i,j}) + j\gamma\}$ for each $Q_i$ from the above set.*

■ As $K$ has enough constants there is some $\eta \in K$ with $D\eta = 0$ and $v(\eta) = \gamma$. For each $i$, let $\delta_i \in K$ with $v(\delta_i) = \min\{v(q_{i,j}) + j\gamma\}$. If for each $i$ it were the case that

$$\pi(\frac{Q_i(\eta)}{\delta_i}) \quad \neq \quad 0 \tag{3.2}$$

then the desired result would be true with $\epsilon = \eta$ for Inequality 3.2 means simply that $v(Q_i(\eta)) = v(\delta_i) = \min\{v(q_{i,j}) + j\gamma\}$. Alas, it may happen that with our first choice of $\eta$, some instance of Inequality 3.2 fails. We replace $\eta$ with $c\eta$ where $v(c) = 0$ and $D\pi(c) = 0$. We need only ensure $\pi(c)$ is not a solution to any of $\sum \pi(\frac{q_{i,j}\eta^j}{\delta_i})Y^j = 0$. With finitely many exceptions, any choice from $\mathbb{Q}$ will work. ■

**Corollary 3.4.6.** *Let $K$ be an $\aleph_1$-saturated valued $D$-field with enough constants. Let $L \subset K$ be a countable subfield of $K$. Let $\gamma \in \Gamma_K$. Then there is some $\epsilon \in K$ such that $v(\epsilon) = \gamma$ and and for any polynomial $Q(X) = \sum_{j=0}^{n} q_j X^j \in L[X]$ one has $v(Q(\epsilon)) = \min\{v(q_j) + j\gamma\}$*

*Remark* 3.4.7. In the case that $\mathbf{k}$ already admits elimination of quantifiers in the natural language of differential rings, Corollary 3.4.6 may be used to give a quick proof that the isomorphism may be extended so that $A_i$ has enough constants. Unfortunately, in general the adjunction of constants may lead to an expansion of the residue field so that there may be some ambiguity as to the extension unless the types of the new elements of the residue field are controlled.

**Lemma 3.4.8.** *If $K$ is a valued field which is also a $\mathcal{D}_e$-ring and $a, b \in K^\times$ satisfy $v(Da) \geq v(a)$ and $v(Db) \geq v(b)$, then $v(D(\frac{a}{b})) \geq v(\frac{a}{b})$.*

■

$$\begin{aligned}
v(D(\frac{a}{b})) &= v(\frac{D(a)b - D(b)a}{b\sigma(b)}) \\
&\geq \min\{v(Da) - v(\sigma(b)), v(a) + v(Db) - (v(b) + v(\sigma(b)))\} \\
&= \min\{v(Da) - v(b), v(a) + v(Db) - 2v(b)\} \\
&\geq \min\{v(a) - v(b), v(a) + v(b) - 2v(b)\} \\
&= v(a) - v(b) \\
&= v(\frac{a}{b})
\end{aligned}$$

■

**Lemma 3.4.9.** *Let $K$ be a valued $D$-field. Let $p \in S_{1,\Gamma}(\Gamma_K)$ be any non-principal 1-type in the value group sort. There is a unique (up to $\mathcal{L}_K$-isomorphism) structure of a valued $D$-field on $K(X)$ such that $v(X) \models p$, $DX = 0$, and such that for any polynomial $Q(X) = \sum_{i=0}^{n} q_i X^j \in K[X]$ one has $v(\sum q_j X^j) = \min\{v(p_j) + jv(X)\}$.*

■ The hypotheses completely describe the $D$-structure and the valuation structure. Since there is no extension of the residue field, we need not consider the extra structure on it. Let us check now that this prescription actually gives a valued $D$-field. We need to check that $v(Dy) \geq v(y)$ for $y \in K(X)$. By Lemma 3.4.8, it suffices to consider $y = P(X) \in K[X]$. Write $P(X) = \sum p_i X^i$. Then

$$\begin{aligned}
DP(X) &= \sum D(p_i X^i) \\
&= \sum D(p_i)X^i + \sigma(p_i)D(X^i) \\
&= \sum D(p_i)X^i
\end{aligned}$$

Since $K$ is a valued $D$-field, $v(D(p_i)) \geq v(p_i)$. Therefore,

$$\begin{aligned}
v(DP(X)) &= \min\{v(D(p_i)) + iv(X)\} \\
&\geq \min\{v(p_i) + iv(X)\} \\
&= v(P(X))
\end{aligned}$$

■

**Lemma 3.4.10.** *If $K$ is a valued $D$-field and $L/K$ is an unramified valued field extension of $K$ given with an extension of the $D$-structure with $D(\mathcal{O}_L) \subseteq \mathcal{O}_L$, then $L$ is also a valued $D$-field.*

■ Let $x \in L^\times$. Let $y \in K$ such that $v(x) = v(y)$. Let $\alpha = \frac{y}{x}$. The hypothesis that $D(\mathcal{O}_L) \subseteq \mathcal{O}_L$ means that $v(D\alpha) \geq v(\alpha) = 0$. Since $y \in K$, $v(Dy) \geq v(y)$. As $x = \frac{y}{\alpha}$, Lemma 3.4.8 shows $v(Dx) \geq v(x)$. ■

*Remark* 3.4.11. If in Lemma 3.4.10 one drops the requirement that the extension is unramified, then the result is not true. For an example, take $K = \mathbb{Q}$ with the trivial valuation and derivation. Let $L = \mathbb{Q}((x))$ with the order of vanishing at the origin valuation and the derivation $\partial = \frac{d}{dx}$. $\frac{d}{dx}(\mathbb{Q}[[x]]) \subseteq \mathbb{Q}[[x]]$, but $\mathrm{ord}_x(\frac{d}{dx}x) = \mathrm{ord}_x(1) = 0 < 1 = \mathrm{ord}_x(x)$.

**Lemma 3.4.12.** *Let $K$ be a valued D-field. Given a type $p \in S_{1,\mathbf{k}}(k_K)$ and a D-polynomial $P \in \mathcal{O}_K\langle X\rangle$ such that*

- *if $a \models p$, then $\pi(P)$ is of minimal total degree such that $\pi(P)(a) = 0$ and*
- $\mathrm{T.\,deg}P = \mathrm{T.\,deg}\pi(P)$,

*there is a unique (up to $\mathcal{L}_K$-isomorphism) D-field $L = K\langle a\rangle$ such that $P(a) = 0$ and $\pi(a) \models p$.*

*Remark* 3.4.13. Lemma 3.4.12 applies equally well in the case that $P = 0$. Recall that in this case, $\mathrm{T.\,deg}P = \infty$.

■ We analyze $L$ as a direct limit of valued field extensions $K_n := K(a, n \ldots, D^n a)$. Let $m := \mathrm{ord}P$.

- $(n \leq m)$ Let $Q(X) \in K\langle X\rangle$ be nonzero with $Q \ll P$. Every element of $K_m$ is of the form $\frac{Q_1(a)}{Q_2(a)}$ with $Q_i \ll P$ so it suffices to compute $v(Q(a))$. Let $\alpha \in K$ such that $\alpha Q \in \mathcal{O}_K\langle X\rangle$ and $\pi(\alpha Q) \neq 0$. Since $\pi(\alpha Q)(b) \neq 0$, we must have $v(Q(a)) = -v(\alpha)$.
- $(n > m)$ If $e = 0$, then $L = K_m$ so there is nothing more to do.

  We assume now that $e \neq 0$. Let $P(X) = F(X, \ldots, D^m X)$. Let $G(X) := F(a, \ldots, D^{m-1}a, X)$. In the case $e \neq 0$, $\sigma y$ and $Dy$ are inter-definable, so it suffices to consider the extension $K(a, Da, \ldots, D^{n-1}a, \sigma^{n-m}D^m a)/K(a, Da, \ldots, D^{n-1}a)$.

  $\sigma^{n-m}D^m a$ satisfies $\sigma^{n-m}G$. So the minimal polynomial $Q_n$ of $\sigma^{n-m}D^m a$ over $K(a, \ldots, D^{n-1}a)$ divides $\sigma^{n-m}G$. Since $\sigma$ reduces to the identity automorphism, $\pi(\sigma^{n-m}D^m a) = b$ which is a simple root of $\pi(G) = \pi(\sigma^{n-m}G)$ (since $\pi(G)' \ll P$). Thus, $b$ is a simple root of $\pi(Q_n)$ so the extension is completely determined as an extension of valued fields.

We check now that the process used above to analyze such extensions may be used to produce them. Let $b \models p$ be a realization of $p$.

Let $K'$ be the field of fractions of $K[X, DX, \ldots, D^m X]/(F(X, \ldots, D^m X))$. $K'$ is given a valuation structure by setting $v(Q(X)) := \max\{-v(a) : aQ \in \mathcal{O}_K\langle X\rangle\}$ for $Q(X) \in K\langle X\rangle$ with $Q \ll P$. In the case that $e = 0$, $K'$ is already a differential field. In the case that $e \neq 0$, let $Q_1$ be the unique (up to multiplication by a unit) factor of $\sigma(G)$ (over $\mathcal{O}_{K'}[y]$) for which $\pi(Q_1)(D^m b) = 0$ and $\pi(Q_1) \neq 0$. This was proven to exist in the course of the uniqueness proof. Since $D^m b$ is a simple root of $\pi(Q_1)$, $K'[y]/(Q_1)$ is an immediate extension of $K'$ when one define $\pi(X) = D^m b$. Define $\sigma : K' \to K'[y]/(Q_1)$ as the extension of $\sigma$ on $K$ (defined by $x \mapsto eDx + x$) with $D^i X \mapsto eD^{i+1}X + X$ for $i < m$ and $D^m X \mapsto y$. We formally define $Dz := \frac{\sigma(z) - z}{e}$ as a function $K' \to K'[y]/(Q_1)$. At this point, we could continue to incrementally define $D$ as the analysis in the uniqueness part of the proof might suggest. This works and the reader is invited to finish the argument this way. We take a different tack. First we check that the required inequalities continue to hold at least for $z \in K'$.

**Claim 3.4.14.** $v(D^{m+1}X) \geq 0$

✠ Since $\sigma$ reduces to the identity on $\mathcal{O}_K[X, \ldots, D^{m-1}X]/(e)$, $D^m X$ is a root to $\sigma(G)$ modulo $e$. Thus, $v(y - D^m x) \geq v(e)$. Since $D^{m+1}x = \frac{y - D^m x}{e}$, the result is now clear. ✠

The next claim is valid in general. That is, there is no restriction on $e$.

**Claim 3.4.15.** *If $Q(X) \ll P(X)$, then $v(Q(X)) \leq v(DQ(X))$.*

✠ By Lemma 3.4.8 we may assume that $Q(X) \in \mathcal{O}_K\langle X\rangle$ and $\pi(Q) \neq 0$. This implies $v(Q(X)) = 0$. Write $Q(X) = \sum q_i(D^m X)^i$ where $q_i \in \mathcal{O}_K[X, \ldots, D^{n-1}X]$. Then $DQ(X) = \sum D(q_i)(D^m X)^i + \sigma(q_i)D(D^m X)^i \in \mathcal{O}_K\langle X\rangle$. For $j \leq m$, it is clear that $v(D^j X) \geq 0$. Claim 3.4.14 shows that $v(D^{m+1}X) \geq 0$ in the case that $e \neq 0$. In the case that $e = 0$, we observe that $D^{m+1}X = \frac{-P^\partial}{\frac{\partial}{\partial X_m}P}$ and $\frac{\partial}{\partial X_m}P \ll P$ so that its valuation is zero.

Since each of the $D^j X$ have non-negative valuation, $v(DQ(X)) \geq 0 = v(Q(X))$. ✠

Let $R$ be the henselization of $\mathcal{O}_{K'}$ and let $L$ be the field of fractions of $R$. Let $\varphi : K' \to \mathcal{D}_e(L)$ be the map $x \mapsto (x, Dx)$. By Claim 3.4.15, $\varphi(\mathcal{O}_{K'}) \subseteq \mathcal{D}_e(R)$. By Lemma 2.2.9, $\mathcal{D}_e(R)$ is a local ring with maximal ideal $\pi_0^{-1}\mathfrak{m}_R$. Since $\pi_0 \circ \varphi = \mathrm{id}_R$, $\varphi$ is a local homomorphism. By Lemma 2.2.10, $\mathcal{D}_e(R)$ is henselian. Thus, there is a unique extension of $\varphi$ to a local homomorphism $\tilde{\varphi} : R \to \mathcal{D}_e(R)$. By Lemma 2.2.6, $\tilde{\varphi}$ extends to a map $L \to \mathcal{D}_e(L)$. Let $D$ denote the function $D : L \to L$ for which $\varphi(x) = (x, Dx)$. Since $D(R) \subseteq R$, by Lemma 3.4.10 $L$ is a valued $D$-field. ■

**Proposition 3.4.16.** *Let $K$ be a valued $D$-field. Let $a \in K^\times$. There is an unramified extension $L/K$ of valued $D$-fields of the form $L = K(x)$ for which $v(x) = v(a)$ and $Dx = 0$. Moreover, the $\mathcal{L}_K$-isomorphism type of $L$ is determined by $\mathrm{tp}(\pi(\frac{a}{x})/k_K)$.*

■ We wish to find $x$ so that $Dx = 0$ and $v(x) = v(a)$. This is equivalent to finding $y =$ "$\frac{a}{x}$". Such a $y$ would have to satisfy $Dy = \frac{Da}{x} = \frac{Da}{a}\frac{a}{x} = \frac{Da}{a}y$. Conversely, if $y$ satisfies $Dy = \frac{Da}{a}y$ then

$$\begin{aligned} D(\frac{a}{y}) &= \frac{Da}{y} - \frac{\sigma(a)Dy}{y\sigma(y)} \\ &= \frac{Da}{y} - \frac{(a + eDa)\frac{Da}{a}y}{y^2(1 + e\frac{Da}{a})} \\ &= \frac{Da}{y}[1 - \frac{1 + e\frac{Da}{a}}{1 + e\frac{Da}{a}}] \\ &= 0 \end{aligned}$$

One would also need $v(y) = 0$ in order for $v(\frac{a}{y}) = v(a)$. That is, we need $y$ to be a solution to $DY = \frac{Da}{a}Y$ with $v(y) = 0$. By Lemma 3.4.12, such extensions exist and they are determined by $\mathrm{tp}(y/k_K)$. ■

**Proposition 3.4.17.** *Let $K$ be a valued $D$-field. Assume that $v(K^\times) = v((K^D)^\times)$. Let $\eta \in K^D$ such that $n|v(\eta)$. Assume that for each positive natural number $m$ dividing $n$ that $v(\eta) \notin m \cdot v(K^\times)$. Then there exists a unique (up to $\mathcal{L}_K$-isomorphism) extension of valued $D$-fields of the form $K(\sqrt[n]{\eta})$.*

■ Let $\epsilon = \sqrt[n]{\eta}$. Since the extension is totally ramified, the valuation structure is completely determined. I claim that the $D$-structure is determined by $D\epsilon = 0$. This fact would certainly fully specify the $D$-structure, the content of my claim is that one must have $D\epsilon = 0$. When $e = 0$, this follows from the fact that $0 = D(\eta) = D(\epsilon^n) = n\epsilon^{n-1}D\epsilon$. When $e \neq 0$, the assertion $D\epsilon \neq 0$ is equivalent to $\sigma(\epsilon) = \omega\epsilon$ for some nontrivial $n$-th root of unity. But then $D\epsilon = \frac{\sigma(\epsilon)-\epsilon}{e} = \epsilon\frac{\omega-1}{e}$. Since $\omega \neq 1$, $v(\frac{\omega-1}{e}) = -v(e) < 0$. So that $v(D\epsilon) < v(\epsilon)$ which violates Axiom 3.

We check that this prescription correctly defines a valued $D$-field. Let $x = \sum_{i=0}^{n-1} x_i\epsilon^i \in L$. Then $v(x) = \min\{v(x_i) + \frac{i}{n}v(\eta)\}$ and $Dx = \sum_{i=0}^{n-1} Dx_i\epsilon^i$ so that visibly, the inequality $v(Dx) \geq v(x)$ holds. ■

**Proposition 3.4.18.** *The isomorphism may be extended so that $A_i$ has enough constants.*

■ We extend the map first so that $v(A_i^\times) = v((A_i^D)^\times)$. Let $a \in A_1$ such that there is no constant in $A_1$ having the same value as that of $a$. Let $p \in S_{1,\mathbf{k}}(k_{A_1})$ be some type containing the formula $Dx = \pi(\frac{Da}{a})x$ as well as the formulas $x \neq b$ for each $b \in k_{A_1}$. By the saturation hypotheses, $p$ is realized in $k_{M_1}$ by some $b_1$ and $f(p)$ is realized in $k_{M_2}$ by some $b_2$. By the surjectivity of the residue map and DHL there is some $c_1 \in M_1$ and some $c_2 \in M_2$ such that $\pi(c_i) = b_i$, $Dc_1 = \frac{Da}{a}c_1$ and $Dc_2 = f(\frac{Da}{a})c_2$. By Proposition 3.4.16, the extension of $f$ given by $c_1 \mapsto c_2$ is an isomorphism of $\mathcal{L}$-structures and the element $\frac{a}{c_1}$ is a constant with value equal to that of $a$.

We may now assume that $v(A_i^\times) = v((A_i^D)^\times)$. We extend the map so as to have $v(A_i^\times) = \Gamma_{A_i}$. Let $\gamma \in \Gamma_{A_1} \setminus v(A_1^\times)$. In the case that $\mathrm{tp}(\gamma/\Gamma_{A_1})$ is non-principal, we find $\epsilon_i \in M_i$ with $v(\epsilon_1) = \gamma$, $v(\epsilon_2) = f(\gamma)$, and $D\epsilon_1 = D\epsilon_2 = 0$ by Axiom 6. Lemma 3.4.5 shows that the extension given by $\epsilon_1 \mapsto \epsilon_2$ is an isomorphism of $\mathcal{L}$-structures.

In the case that $n\gamma \in v(A_1^\times)$ for some $n \in \mathbb{Z}_+$, take $n$ minimal with this property and find some $\eta \in A_1^D$ with $v(\eta) = n\gamma$. By Axiom 5, we may find $\epsilon_i \in M_i$ such that $\epsilon_1^n = \eta$ and $\epsilon_2^n = f(\eta)$. By Proposition 3.4.17 this gives an isomorphism of $\mathcal{L}$-structures. ■

**Proposition 3.4.19.** *The map extends so that $M_i$ and $A_i$ have the same residue field.*

■ As the theory of the residue field of $M_1$ has quantifier elimination by assumption we may fix some isomorphism $\overline{f}$ between the residue fields of $M_1$ and $M_2$ extending the isomorphism induced by the isomorphism $f : A_1 \to A_2$.

Take $a \in \mathcal{O}_{M_1}$ so that $\pi(a)$ is a new element of the residue field. Let $p = tp(\pi(a)/k_{A_1})$. If $\pi(a)$ is differentially transcendental over $k_{A_1}$, then let $b_1 \in \mathcal{O}_{M_1}$ such that $\pi(b_1)$. Let $b_2 \in \mathcal{O}_{M_2}$ such that $\pi(b_2) = \overline{f}(a)$. Then Lemma 3.4.12 shows that we may extend the isomorphism by setting $f(b_1) = b_2$.

Otherwise, Let $P \in \mathcal{O}_{A_1}\langle X\rangle$ such that $\pi(P)(a) = 0$, $\mathrm{T.deg}P = \mathrm{T.deg}\pi(P)$, $\pi(P) \neq 0$ and $\mathrm{T.deg}P$ is minimal with this property. The minimality condition on $P$ implies that for some $i$ one has $v(\frac{\partial}{\partial X_i}P(a)) = 0$. By DHL in both $M_1$ and $M_2$ there is some $b_1 \in M_1$ and $b_2 \in M_2$ such that $P(b_1) = 0$, $f(P)(b_2) = 0$, $\pi(b_1) = \pi(a)$ and $\pi(b_2) = \overline{f}(\pi(a))$.

By Lemma 3.4.12, $A_1\langle b_1\rangle \cong A_2\langle b_2\rangle$ as $\mathcal{L}_{A_i}$-structures. ■

*Remark* 3.4.20. As the proof of Proposition 3.4.19 uses Lemma 3.4.12, the value groups of $A_1$ and $A_1\langle a\rangle$ are equal. Consequently, this construction can be iterated without changing the value group at all. To ensure that the map onto the residue fields is surjective, we should list all the elements of $k_{M_1}$ in order type $\omega_1$ and then consider them one at a time.

### 3.4.2 Immediate Extensions

The rest of this section is devoted to proving that the isomorphism may be extended to immediate extensions.

**Definition 3.4.21.**

1. A *pseudo-convergent sequence* is a limit ordinal indexed sequence $\{x_\alpha\}_{\alpha<\kappa}$ of elements of $K$ such that $(\forall \alpha < \beta < \gamma < \kappa)\ v(x_\alpha - x_\beta) < v(x_\beta - x_\gamma)$.

2. If $L/K$ is an extension of valued $D$-fields and $\{x_\alpha\}$ is a pseudo-convergent sequence from $K$, then the set of *pseudo-limits* of $\{x_\alpha\}$ in $L$ is the set of $c \in L$ such that $(\forall \alpha < \beta < \kappa)\ v(x_\alpha - c) < v(x_\beta - c)$. In this case, one writes $x_\alpha \Rightarrow c$ and says that $\{x_\alpha\}$ *pseudo-converges* to $c$.

3. The pseudo-convergent sequence $\{x_\alpha\}$ *pseudo-solves* the $D$-polynomial $P$ if either $P(x_\alpha) = 0$ for $\alpha \gg 0$ or $P(x_\alpha) \Rightarrow 0$.

4. A pseudo-convergent sequence from $K$ is *strict* if it has no pseudo-limits in $K$.

*Remark* 3.4.22. If $\{x_\alpha\}_{\alpha<\kappa}$ is a pseudo-convergent sequence and one restricts to $\{x_\alpha\}_{\alpha\in I}$ where $I$ is cofinal in $\kappa$, then the new sequence is also pseudo-convergent with the same pseudo-limits and $D$-polynomials which it pseudo-solves as the original sequence. As the value groups considered in this paper are countable, one could always assume that $\kappa = \omega$ by making such a restriction.

**Lemma 3.4.23.** *Assume that the residue field of $K$ is linearly differentially closed and that $K$ has enough constants. If DHL applies to $P \in \mathcal{O}_K\langle X\rangle$ and $a \in \mathcal{O}_K$, then either there is some $b \in \mathcal{O}_K$ with $P(b) = 0$ and $v(a-b) = v(P(a))$ or there is a strict pseudo-convergent sequence $\{x_\alpha\}_{\alpha<\lambda}$ from $K$ pseudo-solving $P$ with $v(x_\alpha - a) = v(P(a))$.*

■ This is proven exactly as in the proof of DHL for complete fields so we give only a sketch here referring the reader to the proof of Proposition 3.3.1 for the detailed computations.

We will produce the sequence with the properties that

- $x_0 = a$
- $v(x_{\alpha+1} - x_\alpha) = v(P(x_\alpha))$
- $v(P(x_\alpha))$ is strictly increasing

Start with $x_0 = a$.

At a limit stage, look for $x_\lambda$ such that $x_\alpha \Rightarrow x_\lambda$. If no $x_\lambda$ exists, then stop – the sequence $\{x_\alpha\}_{\alpha<\lambda}$ is a strict pseudo-solution of $P$.

At a successor stage, since $K$ has enough constants, there is some $\epsilon \in K^D$ with $v(\epsilon) = v(P(x_\alpha))$. We look to solve

$$P(x_\alpha + \epsilon Y) \quad = \quad 0$$

As in the proof of Proposition 3.3.1, we let $y \in \mathcal{O}_K$ be a lifting of a solution to

$$0 \quad = \quad \pi(\frac{P(x_\alpha)}{\epsilon}) + \sum_{i=0}^{\mathrm{ord}P} \pi(\frac{\partial}{\partial X_i} P(a)) D^i Y$$

Set $x_{\alpha+1} = x_\alpha + \epsilon y$. ■

**Definition 3.4.24.** Let $\mathbf{N} \in \mathbb{N}^{<\omega}$. The valued $D$-field $K$ is $\mathbf{N}$- *full* if whenever $\{x_\alpha\}$ is a pseudo-convergent sequence from $K$ and there is an immediate extension $K\langle a\rangle$ with

- $x_\alpha \Rightarrow a$
- $Q(a) = 0$ for some $Q \in K\langle X\rangle$ with $\mathrm{T.deg}Q < \mathbf{N}$
- $\{x_\alpha\}$ pseudo-satisfies $Q$

then there is some $b \in K$ with $Q(b) = 0$ and $x_\alpha \Rightarrow b$.

**Definition 3.4.25.** Let $A \in K\langle X\rangle$ be a non-zero $D$-polynomial. A *refinement* of $A$ at $a$ is a $D$-polynomial $G(Y) = \frac{A(\epsilon Y + a) - A(a)}{c}$ where $c, \epsilon \in (K^D)^\times$ and $G \in \mathcal{O}_K\langle X\rangle$ but $\pi(G) \neq 0$. $\epsilon$ is called the *internal scale* and $c$ is the *external scale*.

*Remark* 3.4.26. Notice that $G$ may be expressed as

$$G(Y) \quad = \quad \sum_{|I|>0} \frac{\epsilon^{|I|}}{c} \partial_I A(a) Y^I$$

**Definition 3.4.27.** Let $\{x_\alpha\}$ be a pseudo-convergent sequence. Let $A$ be a $D$-polynomial. A *refinement of $A$ along $\{x_\alpha\}$* is a sequence of refinements of $A$ at $x_\alpha$ having internal scale $\epsilon_\alpha$ where $v(\epsilon_\alpha) = v(x_{\alpha+1} - x_\alpha)$.

**Definition 3.4.28.** The non-zero $D$-polynomial $A(X) \in \mathcal{O}_K\langle X\rangle$ is *residually linear* if $\pi(A) \in k\langle X\rangle$ is a non-zero linear $D$-polynomial. $A$ is *potentially residually linear* if some refinement of $A$ is residually linear. We will stipulate that the zero polynomial is residually linear.

**Proposition 3.4.29.** *Let $K$ be a valued $D$-field. Let $P(X) \in K\langle X\rangle$ be an irreducible $D$-polynomial. Assume that $K$ has enough constants, has a linearly differentially closed residue field, and is* $\mathrm{T.deg}P$*-full. Let $\{x_\alpha\}_{\alpha<\kappa}$ be a strict pseudo-convergent sequence from $\mathcal{O}_K$.*

1. *If $\{x_\alpha\}$ is a pseudo-solution of $P$, then there is an immediate extension of valued $D$-fields of the form $K\langle a\rangle$ in which $P(a) = 0$ and $x_\alpha \Rightarrow a$.*
2. *If $K\langle a\rangle$ is an extension in which $x_\alpha \Rightarrow a$ and $P(a) = 0$, then $K\langle a\rangle$ is unique up to $\mathcal{L}_K$-isomorphism.*
3. *$P$ is potentially residually linear. In fact, for $\alpha \gg 0$ any refinement of $P$ along $\{x_\alpha\}$ is residually linear.*

*Remark* 3.4.30. Our proof of Proposition 3.4.29 is more complicated than one might expect it needs to be. The idea behind the proof is fairly simple, but technical problems arose for us. For the existence proof, one would like to take some sort of limit. Of course, $\{x_\alpha\}$ may be merely pseudo-convergent rather than convergent, so that there will not be a good notion of a completion. One might try to find the limit by working in some saturated extension and then specializing so as to eliminate excess infinitesimals. Instead, we employ an algebraic construction. For the uniqueness proof, one might like to argue that for $Q \prec P$ the sequence $v(Q(x_\alpha))$ is non-decreasing so that either $\{x_\alpha\}$ pseudo-solves $Q$ and hence pseudo-converges to some $a \in K$ by the inductive hypothesis and fullness (contradicting the strictness of the sequence) or the value settles down. Again some technical problems arise, notably with controlling the qualitative behavior of $\{v(Q(x_\alpha))\}$, so that our actual proof is a bit more involved.

*Remark* 3.4.31. Proposition 3.4.29 allows us to finish extending the back and forth. We will arrange that the hypotheses are true of $A_i$ by an inductive argument. We use potential residual linearity to see that after a linear change of variables, any solution to a $D$-polynomial may be analyzed as an instance of DHL so that we can find the relevant solutions on both sides.

**Definition 3.4.32.** If $\{\gamma_\alpha\}_{\alpha<\kappa}$ is a sequence of elements of $\Gamma$, then we say that the limit of the sequence exists iff the sequence is eventually constant. In that case, we write $\lim \gamma_\alpha = \gamma$ where $\gamma_\alpha = \gamma$ for $\alpha \gg 0$.

Most of the lemmas proved in what follows will be employed to prove Proposition 3.4.29 and they depend inductively on Proposition 3.4.29. We indicate this by the condition

†: The hypotheses of Proposition 3.4.29 are assumed to hold and we assume inductively on $\mathrm{T.deg}P$ that Proposition 3.4.29 is true.

**Lemma 3.4.33 (†).** *Let $Q \in \mathcal{O}_K\langle X\rangle$. Suppose that $Q \prec P$ and DHL applies to $Q$ at $a \in \mathcal{O}_K$. Then there is some $b \in K$ such that $Q(b) = 0$ and $v(b - a) = v(Q(a))$.*

■ By Lemma 3.4.23 either the lemma is true or there is a strict pseudo-convergent sequence $\{y_\beta\}_{\beta<\kappa}$ from $K$ pseudo-solving $Q$ with $y_0 = a$ and $v(y_\beta - a) = v(Q(a))$. By the inductive hypothesis (via †), there is a unique immediate extension $K\langle c\rangle$ in which $Q(c) = 0$ and $y_\beta \Rightarrow c$. By $\mathrm{T.deg}P$-fullness of $K$, $c \in K$. ■

**Lemma 3.4.34 (†).** *Let $Q \in \mathcal{O}_K\langle X\rangle$. If $Q \prec P$, then $\lim v(Q(x_\alpha))$ exists.*

■ We will prove this lemma by $\prec$-induction on $Q$. When $Q$ is a constant $D$-polynomial, the result is obvious. We may now assume that $Q$ is not constant. Thus, there is some $i$ for which $\frac{\partial}{\partial X_i}Q$ is not the zero $D$-polynomial. By the inductive hypothesis, $\lim v(\frac{\partial}{\partial X_i}Q(x_\alpha))$ exists. We finish this proof by a series of lemmas. In each of these lemmas, we assume inductively † as well as

‡: Lemma 3.4.34 is true for $\tilde{Q} \prec Q$. ■

**Lemma 3.4.35 (†, ‡).** *Let $Q \prec P$. There is a pseudo-convergent sequence $\{y_\alpha\}$ having the same pseudo-limits as $\{x_\alpha\}$ such that if $H_\alpha$ is a refinement of $Q$ at $y_\alpha$ with internal scale $\eta_\alpha$, then $v(H_\alpha(\frac{y_{\alpha+1}-y_\alpha}{\eta_\alpha})) = 0$.*

■ We may assume that no cofinal sequence in $\{x_\alpha\}$ already works. So that for each $\alpha \gg 0$ we have that $v(G_\alpha(\frac{x_{\alpha+1}-x_\alpha}{\epsilon_\alpha})) > 0$ where $G_\alpha$ is a refinement of $Q$ at $x_\alpha$ with internal scale $\epsilon_\alpha$.

We will construct the sequence $\{y_\alpha\}$ allowing repetition and later thin to get an actual pseudo-convergent sequence.

Using the inductive hypotheses we may assume that the valuations of the partials of $Q$ have stabilized (via ‡) and any refinement of $Q$ along $x_\alpha$ is residually linear (via †).

**Claim 3.4.36.** *For each $\alpha$, there is some $\beta > \alpha$ minimal with the properties that $v(G_\alpha(\frac{x_\beta-x_\alpha}{\epsilon_\alpha})) < v(\epsilon_\beta) - v(\epsilon_\alpha)$ and $v(G_\alpha(\frac{x_\gamma-x_\alpha}{\epsilon_\alpha})) \leq v(G_\alpha(\frac{x_\beta-x_\alpha}{\epsilon_\alpha}))$ for $\gamma \geq \beta$.*

✠ There cannot be a cofinal sequence of $\gamma$'s on which $v(G_\alpha(\frac{x_\gamma-x_\alpha}{\epsilon_\alpha}))$ is increasing for if this were to occur by the inductive hypothesis for existence and fullness, the sequence $\{x_\alpha\}$ would not be strict.

If the first condition were to fail, then since $G_\alpha$ is residually linear, DHL would apply at $\frac{x_\beta-x_\alpha}{\epsilon_\alpha}$ so that Lemma 3.4.33 would produce $w_\beta$ with $G_\alpha(w_\beta) = 0$ and $v(w_\beta - \frac{x_\beta-x_\alpha}{\epsilon_\alpha}) \geq v(\epsilon_\beta) - v(\epsilon_\alpha)$. Set $z_\beta := x_\alpha + \epsilon_\alpha w_\beta$. We may restrict to a subsequence of $\{z_\beta\}$ which is pseudo-convergent and has the same pseudo-limits as $\{x_\beta\}$. To do this, start with $\zeta_0$. At stage $\beta$, if $\zeta = \lim_{\gamma<\beta}\zeta_\gamma$ exists and $v(\zeta - x_{\beta+1}) \geq v(x_\beta - x_{\beta+1})$, then set $\zeta_\beta := \zeta$. Note that by the strictness of $\{x_\beta\}$ this cannot happen cofinally. In the other case, let $\zeta_\beta := z_\beta$ constructed above. Thin by including only those $\zeta_\beta$ taken from the second case.

For each $\beta$, one has $\tilde{Q}(\zeta_\beta) := Q(z_\beta) - Q(x_\alpha) = 0$ so that by the inductive hypothesis, there is some extension of $K$ in which $\zeta_\beta \Rightarrow \zeta$ and $Q(\zeta) = 0$. By fullness, we may take $\zeta \in K$, but then $\{\zeta_\beta\}$, and consequently $\{x_\alpha\}$, are not strict. ✠

We construct the sequence $\{y_\alpha\}$ now. At stage $\alpha$, let $\beta > \alpha$ be minimal so that $v(G_\alpha(\frac{x_\beta-x_\alpha}{\epsilon_\alpha})) < v(\epsilon_\beta) - v(\epsilon_\alpha)$ and has attained its maximum. As in the proof of the claim, find $z_\beta$ such that $G_\alpha(z_\beta) = 0$ and $v(z_\beta - \frac{x_\beta-x_\alpha}{\epsilon_\alpha}) = v(G_\alpha(\frac{x_\beta-x_\alpha}{\epsilon_\alpha}))$. For $\gamma$ in the range $[\alpha, \beta]$, define $y_\gamma := x_\alpha + \epsilon_\alpha z_\beta$.

Thin now to include only those $y_\beta$'s corresponding to the case the $\beta$ is the right-hand endpoint of the interval $[\alpha, \beta]$. Denote the associated $\alpha$ by $\alpha(\beta)$. I claim now that the sequence $\{y_\beta\}$ is pseudo-convergent with the same pseudo-limits as those of $\{x_\alpha\}$.

Let $\gamma > \beta$. Since $y_\gamma = x_{\alpha(\gamma)} + \epsilon_{\alpha(\gamma)}z_\gamma$, we have $v(y_\gamma - x_\beta) = v(\epsilon_\beta)$. We have also ensured that

$$\begin{aligned} v(y_\beta - x_\beta) &= v(\epsilon_{\alpha(\beta)}) + v(G_{\alpha(\beta)}(\frac{x_\beta - x_{\alpha(\beta)}}{\epsilon_{\alpha(\beta)}})) \\ &< v(\epsilon_\beta) \end{aligned}$$

This calculation itself shows – once we know that $\{y_\beta\}$ is pseudo-convergent – that the sequences $\{y_\beta\}$ and $\{x_\alpha\}$ have the same pseudo-limits.

Since,

$$v(y_\gamma - y_\beta) \quad = \quad v(y_\gamma - x_\beta + x_\beta - y_\beta)$$

we have

$$v(y_\gamma - y_\beta) \quad = \quad v(\epsilon_{\alpha(\beta)}) + v(G_{\alpha(\beta)}(\frac{x_\beta - x_{\alpha(\beta)}}{\epsilon_{\alpha(\beta)}}))$$

This last value is less than $v(\epsilon_\beta)$ but is at least $v(\epsilon_{\alpha(\beta)})$.

Thus, if $\delta > \gamma > \beta$, we have $v(y_\delta - y_\gamma) \geq v(\epsilon_{\alpha(\gamma)}) > v(\epsilon_\beta) > v(y_\delta - y_\beta)$.

To finish the calculation, let $\theta \in \Gamma$ be $\min_i\{v(\frac{\partial}{\partial X_i}Q(x_\alpha))\}$. (Recall that these values do not depend on $\alpha$.) In what follows, $H_\beta$ denotes a refinement of $Q$ at $y_\beta$ with internal scale $\eta_\beta$. Since $H_\beta$ is residually linear, the minimal valuation of a coefficient of $Q(y_\beta + \eta_\beta Y) - Q(y_\beta)$ is $\theta + v(\eta_\beta)$.

$$\begin{aligned}
v(H_\beta(\frac{y_{\beta+1} - y_\beta}{\eta_\beta})) &= v(Q(y_{\beta+1}) - Q(y_\beta)) - \theta - v(\eta_\beta) \\
&= v(Q(x_{\beta+1}) - Q(x_\alpha)) - \theta - v(\epsilon_\alpha) - v(G_\alpha(\frac{x_\beta - x_\alpha}{\epsilon_\alpha})) \\
&= v(G_\alpha(\frac{x_{\beta+1} - x_\alpha}{\epsilon_\alpha})) - v(G_\alpha(\frac{x_\beta - x_\alpha}{\epsilon_\alpha})) \\
&= 0
\end{aligned}$$

The last equality is justified by the observations that $H_\beta$ and its arguments are integral so that the final expression can be no less than zero, but it can be no more then zero because of the maximality condition on $\beta$. ■

**Lemma 3.4.37 (†, ‡).** *If $Q \prec P$, then there is some $\beta < \kappa$ such that if $L/K$ is any extension of valued $D$-fields and $c \in L$ with $v(x_{\beta+1} - c) > v(x_{\beta+1} - x_\beta)$, then $v(Q(c)) = v(Q(x_\beta))$.*

■ Let $L$ be an extension of $K$ (as a valued $D$-field) and in $L$ take $c$ with $x_\alpha \Rightarrow c$. Let $\{y_\alpha\}$ be the pseudo-convergent sequence produced by Lemma 3.4.35. Since $v(H_\beta(\frac{y_{\beta+1}-y_\beta}{\eta_\beta})) = 0$ and $y_\beta \Rightarrow c$, we have

$$\begin{aligned}
0 &= v((H_\beta(\frac{c - y_\beta}{\eta_\beta})) \\
&= v(Q(c) - Q(y_\beta)) - \theta - v(\eta_\beta)
\end{aligned}$$

So $v(Q(c) - Q(y_\beta)) = \theta + v(\eta_\beta)$. The right-hand side is growing with $\beta$ so that either $v(Q(c)) = v(Q(y_\beta))$ eventually (and we're done) or $v(Q(y_\beta))$ is increasing cofinally which would contradict strictness of $\{x_\alpha\}$.

By the compactness theorem, there is some $\beta < \kappa$ for which we have $v(c - x_{\beta+1}) > v(x_\beta - x_{\beta+1}) \Rightarrow v(Q(c)) = v(Q(x_\beta))$. ■

Lemma 3.4.37 finishes the proof of Lemma 3.4.34.

**Lemma 3.4.38 (†).** *For $\alpha \gg 0$, any refinement of $P$ along $x_\alpha$ is residually linear.*

■ By Lemma 3.4.34, $\lim v(\partial_I P(x_\alpha))$ exists for any non-zero multi-index $I$. From now on, work only with $\alpha$ large enough so that this common value has been attained.

For each $\alpha$, let $G_\alpha$ be a refinement of $P$ along $x_\alpha$ with parameters $\epsilon_\alpha$ and $c_\alpha$. Write $G_\alpha(Y) = \sum_I g_{I,\alpha} Y^I$.

**Claim 3.4.39.** *The set of multi-indices $\{I : v(g_{I,\alpha}) = 0\}$ does not depend on $\alpha$ for $\alpha$ sufficiently large. In fact, all such $I$ have the same length.*

✠ If it ever happens that for some $I$ and $J$ with $|I| < |J|$ that $v(g_{\alpha,I}) \leq v(g_{\alpha,J})$, then $v(g_{\beta,I}) < v(g_{\beta,J})$ for $\beta > \alpha$. To see this, observe that the hypothesis is that

$$\begin{aligned} v(\frac{\epsilon_\alpha^{|I|}}{c_\alpha}) + v(\partial_I Q(x_\alpha)) &= g_{I,\alpha} \\ &\leq g_{J,\alpha} \\ &= v(\frac{\epsilon_\alpha^{|J|}}{c_\alpha}) + v(\partial_J Q(x_\alpha)) \end{aligned}$$

That is,

$$v(\partial_I Q(x_\alpha)) \quad \leq \quad (|J| - |I|)v(\epsilon_\alpha) + v(\partial_J Q(x_\alpha))$$

Since the valuations of the partials are the same whether evaluated on $x_\alpha$ or $x_\beta$, $|J| - |I| > 0$, and $v(\epsilon_\beta) > v(\epsilon_\alpha)$, we conclude

$$v(\partial_I Q(x_\beta)) \quad < \quad (|J| - |I|)v(\epsilon_\beta) + v(\partial_J Q(x_\beta))$$

Reversing the above manipulations, the claim follows. ✠

If the lemma were not true, then for some non-zero multi-indexes $I$ and $J$ and cofinal sequence of $\alpha$'s we would have

- $J = (i_0, \ldots, i_{j-1}, i_j + 1, i_{j+1}, \ldots, i_n)$ where $I = (i_0, \ldots, i_n)$ and
- $v(g_{I,\alpha}) > 0 = v(g_{J,\alpha})$

Using the chain rule, we calculate

$$\partial_I G_\alpha(Y) \quad = \quad \frac{\epsilon_\alpha^{|I|}}{c_\alpha} \partial_I P(\epsilon_\alpha Y + x_\alpha) \tag{3.3}$$

By the expansion for $G_\alpha$ and the fact that $v(\partial_I P(x_\beta))$ is stable, we have

$$v(\partial_I G_\alpha(\frac{x_\beta - x_\alpha}{\epsilon_\alpha})) \quad = \quad v(g_{I,\alpha}) \tag{3.4}$$

for $\beta \geq \alpha$. So by Equation 3.3 with $J$ playing the rôle of $I$ and Equation 3.4, we have

$$\begin{aligned} v(\frac{\partial}{\partial X_j} \partial_I G(\frac{x_\beta - x_\alpha}{\epsilon_\alpha})) &= v(g_{J,\alpha}) \\ &= 0 \\ &< v(g_{I,\alpha}) \\ &= v(\partial_I G(\frac{x_\beta - x_\alpha}{\epsilon_\alpha})) \end{aligned}$$

for $\beta > \alpha$.

**Claim 3.4.40.** *There is another pseudo-convergent sequence $\{y_\delta\}$ having the same pseudo-limits as $\{x_\alpha\}$ but pseudo-satisfying $\partial_I P$.*

✠ We will build $\{y_\delta\}$ while following along the sequence $\{x_\alpha\}$. For simplicity, we will use the same index set as that used for $\{x_\alpha\}$ at the cost of repeating some of the terms in the $y$-sequence. To get an actual pseudo-convergent sequence, we thin at the end of the construction.

Start with $y_0 := x_0$. At stage $\alpha$, if $y_\beta = y$ is constant for $\alpha > \beta \gg 0$ (N.B.: If $\alpha$ is successor, this condition will always be true.) and $v(x_{\alpha+1} - y) \geq v(x_{\alpha+1} - x_\alpha)$, then set $y_\alpha := y$.

(N.B.: Since $\{x_\alpha\}$ is strict, cofinally we will not be in this case.)

Otherwise, since DHL applies to $\partial_I G_\alpha$ at $\frac{x_{\alpha+1}-x_\alpha}{\epsilon_\epsilon}$, by Lemma 3.4.33 there is some $w_\alpha \in K$ such that $\partial_I G_\alpha(w_\alpha) = 0$ and $v(w_\alpha - \frac{x_{\alpha+1}-x_\alpha}{\epsilon_\alpha}) \geq v(\partial_I G_\alpha(\frac{x_{\alpha+1}-x_\alpha}{\epsilon_\alpha})) > 0$. Define $y_\alpha := \epsilon_\alpha w_\alpha + x_\alpha$.

Let us check that this construction works. In the case where we actually change $y$, Equation 3.3 shows that $\partial_I P(y_\alpha) = 0$. By the construction, cofinally, $v(y_\alpha - x_{\alpha+1}) \geq v(\epsilon_\alpha)$ so that $\{y_\alpha\}$ and $\{x_\alpha\}$ have the same pseudo-limits. Let us check now that (up to repetition) $\{y_\alpha\}$ is pseudo-convergent. Take $\alpha < \beta < \gamma$ such that $y_\alpha, y_\beta$, and $y_\gamma$ are distinct. Take each of these ordinals so that $y_\delta = x_\delta + \epsilon_\delta w_\delta$. By the construction, $v(\epsilon_\beta) \leq v(x_\gamma - y_\beta) < v(\epsilon_\gamma)$ and $v(\epsilon_\alpha) \leq v(x_\beta - y_\alpha) < v(\epsilon_\beta)$. Since $y_\beta = x_\beta + \epsilon_\beta w_\beta$ and $y_\gamma = x_\gamma + \epsilon_\gamma w_\gamma$, these inequalities give $v(y_\gamma - y_\beta) \geq v(\epsilon_\beta) > v(y_\beta - y_\alpha)$. ✠

By the inductive hypothesis and fullness again, there is some $b \in K$ such that $y_\alpha \Rightarrow b$ and $\partial_I P(b) = 0$. As $\{y_\alpha\}$ and $\{x_\alpha\}$ have the same pseudo-limits, $x_\alpha \Rightarrow b$ as well, but this contradicts the strictness of $\{x_\alpha\}$. ■

**Lemma 3.4.41 (†).** *The structure of a valued D-field on $K\langle a\rangle$ is determined by $P(a) = 0$ and $x_\alpha \Rightarrow a$.*

■ When $e = 0$, this is an immediate consequence of Lemma 3.4.37 since every element of $K\langle a\rangle$ is of the form $\frac{Q(a)}{R(a)}$ with $Q, R \ll P$.

In any case, Lemma 3.4.37 shows that $K(a, \ldots, D^m a)$ is an immediate (and hence unramified) extension of $K$. So the extension $K(a, \ldots, D^m a)/K(a, \ldots, D^{m-1} a)$ is generated by some single element $b$ characterized by $F(b) = 0$ and $\pi(b) = c$ for some irreducible $F \in \mathcal{O}_{K(a,\ldots,D^{m-1}a)}[X]$ and $c \in k$ which is a simple root of $\pi(F)$. The extension $K(a, \ldots, D^{m+1} a)/K(a, \ldots, D^m a)$ is then generated by $\sigma(b)$ who is characterized by being the unique solution to the irreducible factor of $\sigma(F)$ for which $c$ is a simple root of the reduction. As in the proof of Lemma 3.4.12, the extension of $D$ to the henselization of $\mathcal{O}_{K(a,\ldots,D^m a)}$ is now determined. ■

At this point we would like to show that this analysis may be used to produce a valued $D$-field extending $K$ determined by the data $x_\alpha \Rightarrow X$ and $P(X) = 0$.

**Lemma 3.4.42 (†).** *Assume that $\{x_\alpha\}$ pseudo-satisfies $P$. Then there is an immediate extension of $K$ of the form $K\langle a\rangle$ such that $x_\alpha \Rightarrow a$ and $P(a) = 0$.*

■ By Lemma 3.4.38 we may assume that $P$ is residually linear.

Let $m := \mathrm{ord} P$. Let $K' := K(X, DX, \ldots, D^{m-1} X)$. We define a valuation on $K'$ by setting $v(Q) := \lim v(Q(x_\alpha))$ for $Q \in K\langle X\rangle$ with $\mathrm{ord} Q < \mathrm{ord} P$. By Proposition 3.4.37, these limits exist. One checks easily that $v$ is a valuation on $K'$ making it an immediate extensions of $K$. Let $R$ be the henselization of $\mathcal{O}_{K'}$ and let $L$ be the field of fractions of $R$. Fix an embedding (as a valued field over $K$) of the field of fractions of $K'[D^m X]/(P)$. We break the argument into two cases depending on whether $m = 0$ or not.

Let us start with the case that $m = 0$. We define a map $K[X]/(P) \xrightarrow{\varphi} \mathcal{D}_e(L)$ by extending the given map on $K'$ and sending $X \mapsto (X, a)$ where $a$ is the unique solution of $DP$ (considered now as a polynomial in $DX$ over $R$) which reduces to $D\pi(X) = D\pi(x_\alpha)$. Since $P$ is residually linear, the same is true of $DP$ so this makes sense. We write $\varphi$ as $x \mapsto (x, Dx)$. We need to check that for $Q(X) \in K'[X]$ with $\deg Q < \deg P$ that $v(DQ) \geq v(Q)$.

By Lemmas 3.4.38 and 3.4.37 the are $c, \epsilon \in (K^D)^\times$ such that

- $\widetilde{Q}_\alpha := \frac{Q(\epsilon Y + x_\alpha)}{c} \in \mathcal{O}_K[X]$
- $v(\epsilon) = v(x_{\alpha+1} - x_\alpha)$ and
- $v(Q(x_\beta)) = v(c)$ for $\beta \geq \alpha$.

Define $\widetilde{P}(Y) := \frac{P(\epsilon Y + x_\alpha)}{c}$. Set now $b := \frac{X - x_\alpha}{\epsilon}$. Then by construction $Db$ is the unique element of $R$ which is a solution to $D\widetilde{P}$ and reduces to $D\pi(\frac{x_{\alpha+1} - x_\alpha}{\epsilon})$.

We compute now

$$\begin{aligned} v(DQ(X)) &= v(DQ(\epsilon b + x_\alpha)) \\ &= v(Dc\widetilde{Q}(b)) \\ &= v(c) + v(D\widetilde{Q}(b)) \\ &\geq v(c) \\ &= v(Q) \end{aligned}$$

The last inequality follows from the fact that when calculated formally as a $D$-polynomial, each coefficient of $Q$ is integral. As we observed above, both $b$ and $Db$ are integral. Thus, the expression $D\widetilde{Q}(b)$ has value at least zero.

So we finish the existence proof by extending the map $\varphi|_{\mathcal{O}_{K[X]/(P)}}$ to $R$ (uniquely by the very definition of the henselization) and observing via Lemma 3.4.10 that the resulting structure on $L$ is that of a valued $D$-field.

Work now in the case that $m > 0$. We check that if $Q(X) \in K'$, then $v(DQ) \geq v(Q)$. By Lemma 3.4.8, it suffices to consider the case of $Q(X) \in K[X, \ldots, D^{m-1}X]$.

**Claim 3.4.43.** *If $G(X) \in K[X, \ldots, D^m X]$ then $v(G(x_\alpha))$ is eventually non-decreasing.*

✠ We prove this by induction on $\mathrm{T.deg} G$. If $G \ll P$, then the values are eventually constant by Proposition 3.4.37. In general, using the Euclidean algorithm in $K(X, \ldots, D^{m-1}X)[D^m X]$ we may write

$$HG = AP + R$$

for some $H \in K[X, \ldots, D^{m-1}X] \setminus \{0\}$, $R \ll P$, and $A \in K[X, \ldots, D^m X]$. For $\alpha \gg 0$ the values $v(H(x_\alpha))$ and $v(R(x_\alpha))$ are eventually constant. By induction, $v(AP(x_\alpha))$ is eventually increasing. If $v(G) \neq \lim v(\frac{R(x_\alpha)}{H(x_\alpha)})$, then it must be that $v(G(x_\alpha)) = v(\frac{A(x_\alpha)P(x_\alpha)}{H(x_\alpha)})$ which is increasing (and less than $\lim v(\frac{R(x_\alpha)}{H(x_\alpha)})$). ✠

By Claim 3.4.43 $v(DQ) \geq v(DQ(x_\alpha)) \geq v(Q(x_\alpha)) = v(Q)$ for $\alpha \gg 0$.

As before, we finish by extending the map $\varphi : \mathcal{O}_{K'} \to \mathcal{D}_e(R)$ by the universal property of the henselization and Lemma 3.4.10. ■

We note now that the †-Lemmas give a full proof of Proposition 3.4.29

■ Via † we proceed by induction on $\mathrm{T.deg} P$. Lemma 3.4.41 proves uniqueness. Lemma 3.4.42 proves existence. Lemma 3.4.38 proves potential residual linearity. ■

Proposition 3.4.29 is the last step needed to prove that any $(\mathbf{k}, \mathbf{G})$-$D$-field may be embedded into a $(\mathbf{k}, \mathbf{G})$-$D$-henselian field. We will give the details of this in the next subsection.

With the help of Proposition 3.4.29 we can complete the back and forth.

**Proposition 3.4.44.** *If $A_1$ has enough constants and $k_{A_1}$ is linearly differentially closed, then we may extend the isomorphism to a maximal immediate extension of $A_1$.*

■ Let $a \in \mathcal{O}_{M_1}$ be an element such that $A_1\langle a\rangle$ is an immediate extension of $A_1$. Working inductively, we may assume that $A_1$ is full with respect to $\mathrm{T.deg}P$ where $P$ is a minimal $D$-polynomial for $a$ over $A_1$. By Proposition 3.4.29, we may assume that $P$ is residually linear. Let $\{x_\alpha\}_{\alpha<\omega}$ be a pseudo-convergent sequence from $\mathcal{O}_K$ with $x_\alpha \Rightarrow a$. Using DHL and $\aleph_1$-saturation of $M_2$, we may find $b \in M_2$ such that $f(P)(b) = 0$ and $f(x_\alpha) \Rightarrow b$. By Proposition 3.4.29, $A_1\langle a\rangle \cong A_2\langle b\rangle$. ■

This completes the proof of completeness and quantifier elimination, though we recap the argument in the next subsection.

### 3.4.3 Recap

With the necessary lemmas now proven, let us go through the proof of Theorem 3.4.1 again.

Let us first prove quantifier elimination for the theory of $(\mathbf{k}, \mathbf{G})$-$D$-henselian fields.

We start by enlarging the domain of $f$ so that the substructures $A_1$ and $A_2$ have enough constants by using Proposition 3.4.18. Since the construction of Proposition 3.4.18 does not produce any new elements of the value group of $A_i$, we may iterate the procedure $\omega$ times to ensure that $A_1$ and $A_2$ have enough constants. From this point on, $A_1$ and $A_2$ will always have enough constants.

We then use Proposition 3.4.19 to extend $A_1$ and $A_2$ so that their residue fields agree with the residue fields of the ambient models. The construction of Lemma 3.4.12 does not enlarge the value groups at all so that we may iterate it freely without losing the countability of the value groups. To ensure that no mistake is made at this point, one should lay out all the elements of the residue fields (from $M_1$ and $M_2$ respectively) in order type $\omega_1$ and alternate between considering elements from $M_1$ and from $M_2$.

We then enlarge $A_i$ to a maximal immediate extension inside $M_i$. Alternating between $M_1$ and $M_2$ as above (and having all the elements listed in order type $\omega_1$) we consider an element $a \in M_i$ such that $A_i\langle a\rangle$ is an immediate extension of $A_i$ and $P(a) = 0$ with $P \in \mathcal{O}_{A_i}\langle X\rangle$ of minimal total degree. This will ensure that $A_i$ is full. Build a pseudo-convergent sequence $x_\alpha$ from $A_i$ such that $x_\alpha \Rightarrow a$. Proposition 3.4.29 shows that the extension is uniquely determined by this data. (N.B.: Proposition 3.4.29 does not require that we verify that $\{x_\alpha\}$ pseudo-solves $P$.)

Eventually, we run out of elements giving immediate extensions. We enlarge the value group either via Lemma 3.4.9 or Proposition 3.4.17. These constructions maintain the property of having enough constants. At this point we repeat the arguments involving immediate extensions. If we iterate this procedure $\omega_1$ times (again making sure to alternate between $M_1$ and $M_2$ and to list the elements of $\Gamma$ so as to exhaust everything in exactly $\omega_1$ steps), we will produce the desired isomorphism.

As to proving that the theory of $(\mathbf{k}, \mathbf{G})$-$D$-henselian fields is the model completion of the theory of $(\mathbf{k}, \mathbf{G})$-$D$-fields we use the existence parts of the lemmas cited above to show that any $(\mathbf{k}, \mathbf{G})$-$D$-field may be embedded into a henselian such field.

Let $K$ be a $(\mathbf{k}, \mathbf{G})$-$D$-field. We use the existence part of Proposition 3.4.18 to prove that $K$ may be enlarged so as to have enough constants. We use Lemma 3.4.17 to enlarge $K$ to satisfy Axiom 5 and so that its value group is a model of $\mathrm{Th}(\mathbf{G})$. We use the existence part of Lemma 3.4.12 to enlarge $K$ so that its residue field is a model of $\mathrm{Th}(\mathbf{k})$. Working inductively, we may assume that whenever we are presented with some instance of DHL to be realized that $K$ is sufficiently full and has a linearly differentially closed residue field. Using Lemma 3.4.23 we produce the necessary pseudo-convergent sequence in $K$ and then use Proposition 3.4.29 to actually find a solution.

# Chapter 4

# Distances on Varieties

## 4.1 Basic Properties of the $v$-adic Proximity Function

We begin by considering $v$-adic distances on varieties over valued fields with a single valuation specified. In this case, the proximity functions have a natural geometric interpretation. Later, it will be necessary to complicate matters so as to consider a family of valuations.

**Definition 4.1.1.** Let $K$ be a valued field with value group $\Gamma$, valuation $v : K^\times \to \Gamma$ and ring of integers $\mathcal{O}_K$.
For each $\gamma \in \Gamma_{\geq 0}$, let $I_\gamma := \{x \in \mathcal{O}_K : v(x) > \gamma\}$.
Let $\pi_\gamma : \mathcal{O}_K \to \mathcal{O}_K/I_\gamma$ be the natural quotient map.
Let $S_\gamma := \text{Spec}\mathcal{O}_K/I_\gamma$.
Let $S := \text{Spec}\mathcal{O}_K$.
Let $\eta := \text{Spec}K \hookrightarrow S$.
If $X$ is a scheme over $S$, then define $X_\gamma := X \times_S S_\gamma$ and $X_\eta := X \times_S \eta$.
Let $\pi_\gamma$ also denote the reduction map $\pi_\gamma : X(S) \to X_\gamma(S_\gamma)$. Let $\iota : X(S) \hookrightarrow X_\eta(K)$ be the natural inclusion.

For $Y \subseteq X$ a closed subscheme and $P \in X(S)$ define the $v$-adic distance from $P$ to $Y$ to be $d_v(P, Y) := \inf\{\gamma \in \Gamma_{\geq 0} : \pi_\gamma(P) \notin Y_\gamma(S_\gamma)\}$

*Remark* 4.1.2. In Definition 4.1.1, if $Y$ is not finitely presented or is not closed in $X$ then it may happen that the infimum does not exist. In that case, interpret the distance as a cut. This situation will not arise in our applications

*Remark* 4.1.3. We use scheme-theoretic language in order to give a smooth presentation of this material, but no deep properties of schemes will be used. The reader may think of $X$ as $\mathbb{P}^n$ and $Y$ as being defined by the vanishing of a finite number of equations over $\mathcal{O}_K$. $Y_\gamma$ is just the result of considering the equations for $Y$ modulo $I_\gamma$.

*Remark* 4.1.4. Since the geometric distance we have defined concerns only $S$-valued points of $X$, one may replace $X$ with $X_{red}$ and obtain the same distance. However, replacing $Y$ with $Y_{red}$ will often change the function $d_v(\cdot, Y)$.

In the next lemmas we record some easy calculations. In the following, we drop the subscript "$v$" from the notation of Definition 4.1.1.

**Lemma 4.1.5.** *Let $X_1$ and $X_2$ be schemes over $S$. Let $Y_1 \subseteq X_1$ and $Y_2 \subseteq X_2$ be subschemes. Then $d((P_1, P_2), Y_1 \times Y_2) = \min\{d(P_1, Y_1), d(P_2, Y_2)\}$ for any point $(P_1, P_2) \in X_1 \times X_2\,(S)$.*

■ By the very definition of the product, $\pi_\gamma(P_1, P_2) \in Y_1 \times_S Y_2\,(S_\gamma) \Leftrightarrow \pi_\gamma(P_1) \in Y_1(S_\gamma)$ and $\pi_\gamma(P_2) \in Y_2(S_\gamma)$. ■

**Lemma 4.1.6.** *Let $Z \subseteq Y \subseteq X$ be subschemes Then for any point $P \in X(S)$, $d(P, Z) \leq d(P, Y)$.*

■ For any $\gamma \in \Gamma_{\geq 0}$, $Z_\gamma \subseteq Y_\gamma$. So $\pi_\gamma(P) \in Z_\gamma(S_\gamma) \Rightarrow \pi_\gamma(P) \in Y_\gamma(S_\gamma)$. ■

**Lemma 4.1.7.** *Let $Y, Z \subseteq X$ be two subschemes. For any point $P \in X(S)$ one has $d(P, Y \cup Z) = \max\{d(P, Y), d(P, Z)\}$.*

■ $\pi_\gamma(P) \in (Y \cup Z)_\gamma(S_\gamma) \Leftrightarrow \pi_\gamma(P) \in Y_\gamma(S_\gamma)$ or $\pi_\gamma(P) \in Z_\gamma(S_\gamma)$. ■

**Lemma 4.1.8.** *Let $f : X' \to X$ be a map of schemes over $S$. Let $Y \subseteq X$ be a closed subscheme. Let $Y' = f^*Y := Y \times_X X'$. Then for any $P \in X'(S)$, $d(P, Y') = d(f(P), Y)$.*

■ $\pi_\gamma(P) \in Y'_\gamma(S_\gamma) \Leftrightarrow \pi_\gamma(f(P)) \in Y_\gamma(S_\gamma)$. ■

**Lemma 4.1.9.** *The function $d(P, Q)$ defines an ultra-metric on $X(S)$. That is,*

1. $d(P, Q) = \infty \Leftrightarrow P = Q$
2. $d(P, Q) = d(Q, P)$
3. $d(P, Q) \geq \min\{d(P, R), d(R, Q)\}$

■

1. If $P = Q$, then for every $\gamma \in \Gamma_{\geq 0}$ one has $P \in \{Q\}$ so that $d(P, Q) = \infty$. Conversely, if $P \neq Q$, then at least one of the equations defining $\{Q\}$ does not vanish at $P$ and hence has a non-infinite valuation (say $\gamma$). So $d(P, Q) \leq \gamma$.
2. For singletons, $\pi_\gamma(P) \in \{Q\}_\gamma$ iff $\pi_\gamma(P) = \pi_\gamma(Q)$ iff $\pi_\gamma(Q) \in \{P\}_\gamma$.
3. If $d(P, R) \geq \gamma$ and $d(R, Q) \geq \gamma$, then $\pi_\gamma(P) \in \{R\}_\gamma$ and $\pi_\gamma(R) \in \{Q\}_\gamma$. So $\pi_\gamma(P) \in \{Q\}_\gamma$ as required.

■

**Lemma 4.1.10.** *If $X = G$ a group scheme over $S$ having unit element $1 \in G(S)$, then for any $P, Q \in G(S)$ one has $d(P, Q) = d(P \cdot Q^{-1}, 1)$ and $d(P, 1) = d(P^{-1}, 1)$.*

■ As $G$ is a group scheme, the maps $\pi_\gamma : G(S) \to G_\gamma(S_\gamma)$ are group homomorphisms.
Thus,

$$\begin{aligned} \pi_\gamma(P) = \pi_\gamma(Q) \quad &\Leftrightarrow \quad \pi_\gamma(P)\pi_\gamma(Q)^{-1} = 1 \\ &\Leftrightarrow \quad \pi_\gamma(PQ^{-1}) = 1 \end{aligned}$$

Likewise, $1 = \pi_\gamma(P) \Leftrightarrow 1 = \pi_\gamma(P)^{-1} = \pi_\gamma(P^{-1})$. ■

Let $K$ be a valued field. With the notation of Definition 4.1.1, let $X$ be a scheme over $S$. Let $Y \subseteq X$ be a closed subscheme. For $P \in X(S)$, how does one compute $d(P, Y)$?

Let $U \subseteq X$ be some affine open containing the point $P$. Then $d(P, Y) = \inf\{v(f(P)) : f \in \mathcal{I}_Y(U)\}$. If $X$ is quasi-compact and the ideal sheaf $\mathcal{I}_Y$ of $Y$ in $X$ is finitely generated, then we can reduce the number of values over which the infimum is taken to be finite. Namely, let $\mathcal{U} := \{U_1, \ldots, U_n\}$ be some finite covering of $X$ by open affines and let $\mathcal{I}_Y(U_i) = (f_1^{(i)}, \ldots, f_{m_i}^{(i)})$. Then $d(P, Y) = \inf\{v(f_j^{(i)}(P)) : P \in U_i\}$.

This method of calculating the proximity functions shows that in case $X$ is quasi-compact and $Y$ is defined by a finite number of equations, then the assertion "$d(P, Y) \geq \gamma$" may be expressed in the first-order language of valued fields.

This method of calculating the $v$-adic distance can be used to define the notion of the distance to a closed subscheme of a scheme over a valued field without specifying an integral model. That is, one can simply choose an affine cover of the ambient scheme and on each affine in the cover choose equations for the subscheme and define the distance by calculating the valuations of the defining equations evaluated at the given point. In Chapter 6 this notion of $v$-adic distance is employed by the authors of the theorems we discuss. However, since this definition of the distance is sensitive to the choices made – especially when the ambient scheme is not proper – we prefer to use the definition given in Definition 4.1.1.

## 4.2 Distance Estimates over Valued $\mathcal{D}$-Fields

In this section we will prove a general proposition on approximating distances to subvarieties by the distance to other subvarieties for points inside sets defined by $\mathcal{D}$-equations.

Let us fix some notation.

**Definition 4.2.1.** Let $\mathcal{D}$ be a ring functor of the sort considered in Chapter 2. Let $\mathfrak{E}$ be a set of terms in the language of $\mathcal{D}$-rings. The *variety* of $\mathcal{D}$-rings defined by $\mathfrak{E}$, $\mathfrak{V}_{\mathfrak{E}}$, is the class of all $\mathcal{D}$-rings $(R, \{\partial_\alpha\})$ satisfying all the equations in $\mathfrak{E}$. That is, for each term $t(X_1, \ldots, X_n) \in \mathfrak{E}$ and sequence $\mathbf{a} \in R^n$ one has $t(\mathbf{a}) = 0$. A variety of $\mathcal{D}$-rings is a class of the form $\mathfrak{V}_{\mathfrak{E}}$ for some set of equations $\mathfrak{E}$.

*Remark* 4.2.2. This notion of variety of $\mathcal{D}$-rings is the standard definition from universal algebra. We might also call them equational classes or positive universal classes.

The next proposition is stated more generally than necessary (though not in as much generality as it could be). A version of this proposition for a single valuation appeared as Proposition 6.3 in [25].

**Proposition 4.2.3.** *Let $K$ be a field. Let $\Gamma$ be an ordered abelian group. Let $\Sigma$ be a set of $\Gamma$-valuations on $K$ with the property that for any $x \in K^\times$ there is some $\gamma \in \Gamma$ such that $v(x) \leq \gamma$ for every $v \in \Sigma$.*

*Let $\mathfrak{V}$ be a variety of $\mathcal{D}$-rings containing $K$.*

*Assume that for each multi-index $\alpha$, there is some $\gamma_\alpha \in \Gamma$ such that for any $x \in K$ and any $v \in \Sigma$, one has $v(\partial_\alpha x) \geq v(x) - \gamma_\alpha$.*

*Let $X$ be a scheme over* $\mathrm{Spec}\mathcal{O}_{K,\Sigma}$.

*Let $Y, Z \subseteq X$ be closed subschemes over* $\mathrm{Spec}\mathcal{O}_{K,\Sigma}$.

*Let $\mathfrak{D} \subseteq \nabla_\alpha X$ be a closed subscheme of the $\alpha$-th jet space of $X$ for some $\alpha$. Let $\Xi$ be the $\mathcal{D}$-scheme defined by $P \in \Xi \iff \nabla_\alpha(P) \in \mathfrak{D}$.*

*Suppose that for every $\mathcal{D}$-field $L \in \mathfrak{V}$ extending $K$ one has $(Y \cap \Xi)(L) = (Y \cap \Xi)(L)$.*

*Then there are constants $n \in \mathbb{N}$ and $\gamma \in \Gamma$ such that for any $v \in \Sigma$ and any point $P \in (X \cap \Xi)(\mathcal{O}_{K,v})$ one has $d_v(P, Y) \leq n \cdot d_v(P, Z) + \gamma$.*

We will prove this proposition via an ultraproduct construction. The level of generality of our hypotheses makes a direct appeal to the compactness theorem somewhat awkward as one would need to show how to describe everything involved in a first-order way.

■

If the lemma were false, then for each $\gamma \in \Gamma$ and $n \in \mathbb{Z}_+$ we could find a valuation $v = v_{(n,\gamma)} \in \Sigma$ and a point $P_{(n,\gamma)} \in \Xi(\mathcal{O}_{K,v_{(n,\gamma)}})$ such that $d_{v_{(n,\gamma)}}(P_{(n,\gamma)}, Y) > n d_{v_{(n,\gamma)}}(P_{(n,\gamma)}, Z) + \gamma$.
For $n \in \mathbb{Z}_+$ and $\gamma \in \Gamma$ define $[(n,\gamma),\infty) := \{(m,\delta) : m \geq n, \delta \geq \gamma\}$.
Let $\mathcal{F} := \{X \subseteq \mathbb{Z}_+ \times \Gamma : X \supseteq [(n,\gamma),\infty) \text{ for some } (n,\gamma) \in \mathbb{Z}_+ \times \Gamma\}$. Note that $\mathcal{F}$ is a filter on $\mathbb{Z}_+ \times \Gamma$.
Let $\mathcal{F}'$ be any ultrafilter extending $\mathcal{F}$.
Let $(\mathbf{K}, \mathbf{v}, \mathbf{\Gamma})$ be the ultraproduct $\prod_{/\mathcal{F}'}(K, v_{n,\gamma}, \Gamma)$.
Let $\mathbf{P}$ be the image of $(P_{n,\gamma})$ in $\mathcal{O}_{\mathbf{K}}$.
Let $\xi := d_{\mathbf{v}}(\mathbf{P}, Y)$.
Let $\mathfrak{p} := \{x \in R : (\forall n, \gamma \in (\mathbb{Z}_+ \times \Gamma))\ \mathbf{v}(x) \geq n\xi + \gamma\}$.

**Claim 4.2.4.** $\mathfrak{p}$ *is prime.*

✠ Let $x, y \in \mathcal{O}_{\mathbf{K}} \setminus \mathfrak{p}$. We have $\mathbf{v}(x) \leq n\xi + \gamma$ and $\mathbf{v}(y) \leq m\xi + \delta$ for some $m, n \in \mathbb{Z}_+$ and $\gamma, \delta \in \Gamma$. Thus, $v(xy) \leq (n+m)\xi + (\gamma + \delta)$ so that $xy \notin \mathfrak{p}$. ✠

**Claim 4.2.5.** $\mathfrak{p}$ *is a* $\mathcal{D}$*-ideal*

✠ Let $x \in \mathfrak{p}$. Let $n \in \mathbb{Z}_+$ and $\gamma \in \Gamma$. For any $\alpha$, we still have $\gamma + \gamma_\alpha \in \Gamma$ so that $\mathbf{v}(x) > n\xi + (\gamma + \gamma_\alpha)$. By our hypotheses and Łos' Lemma, $\mathbf{v}(\partial_\alpha x) \geq n\xi + (\gamma + \gamma_\alpha) - \gamma_\alpha = n\xi + \gamma$ so that $\partial_\alpha x \in \mathfrak{p}$ as well. ✠

The localization of $\mathcal{O}_{\mathbf{K}}$ at $\mathfrak{p}$ is $\mathcal{O}_{\mathbf{K},\mathfrak{p}} = \{x \in \mathbf{K} : (\exists (n,\gamma) \in \mathbb{Z} \times \Gamma)\ v(x) > n\xi + \gamma\}$. By the boundedness hypotheses on the valuation of $\partial_\alpha x$, $\mathcal{O}_{\mathbf{K},\mathfrak{p}}$ is a sub-$\mathcal{D}$-ring of $\mathbf{K}$. Since for each $x \in K^\times$ we have assumed that $\{v(x) : v \in \Sigma\}$ is bounded in $\Gamma$, we have that $K^\times \hookrightarrow \mathcal{O}_{\mathbf{K},\mathfrak{p}}^\times$ via the diagonal map so that composing with the quotient map we obtain a map of $\mathfrak{V}$-fields $K \to L := \mathcal{O}_{\mathbf{K},\mathfrak{p}}/\mathfrak{p}$.

Let $\mathbf{P}$ continue to denote its image in $\Xi(L)$. By construction, $\mathbf{P} \in (Z \cap \Xi)(L) \setminus (Y \cap \Xi)(L)$. This is a contradiction. ■

*Remark* 4.2.6. One would like to argue that the hypotheses imply that $\mathfrak{D} \cap \nabla_\alpha Y = \mathfrak{D} \cap \nabla_\alpha Z$ so that once one knows that $\nabla_\alpha P \in \mathfrak{D}$ the question of whether or not it lies in $Y_\gamma$ has the same answer as whether or not it lies in $Z_\gamma$. Unfortunately, our hypotheses are weaker than this for we posit only the equality $\Xi(L) \cap Y(L) = \Xi(L) \cap Z(L)$. Restricting to affine patches in $X$, this equality corresponds to the equality of the set of prime $\mathcal{D}$-ideals containing the $\mathcal{D}$-ideal if $\Xi \cap Y_\eta$ and of those containing the $\mathcal{D}$-ideal of $\Xi \cap Z_\eta$.

*Remark* 4.2.7. A direct proof is also hampered by the fact that $\mathcal{O}_K$ need not be a $\mathcal{D}$-ring. In the case that $\mathcal{O}_K$ is a $\mathcal{D}$-ring and the equalities are assumed to hold for every $R$-$\mathcal{D}$-algebra, then the straightforward proof works.

*Remark* 4.2.8.

I believe that the ultraproduct construction has another advantage over a purely algebraic proof. One would like to argue that from a sequence of counter-examples to the Proposition one could produce a limit point $P \in \Xi$ which actually lies on $Y$ but is outside of $Z$. The point $\mathbf{P}$ in the above proof may be regarded as such a limit.

# Chapter 5

# The *abc* Theorem for Algebraic Groups over Function Fields

In this chapter we will prove a characteristic $p$ analog of Buium's *abc* theorem for abelian varieties over function fields of characteristic zero [5]. While our main concern is the characteristic $p$ case, our methods give a stronger theorem in characteristic zero as well. The structure of our proof is quite similar to that of Buium's characteristic zero proof.

## 5.1 Statement of Main Theorem

In what follows, $k$ is an algebraically closed field of characteristic $p$, $C$ is a smooth projective curve over $k$, and $K = k(C)$ is the function field of $C$. Let $\mathrm{Spec}K := \eta \hookrightarrow C$ be the inclusion of the generic point into $C$. If $X \to C$ is a scheme over $C$, then $X_\eta := X \times_C \eta$. We identify each point $x \in C(k)$ with its corresponding valuation $v_x := \mathrm{ord}_x$ on $K$.

**Theorem 5.1.1.** *Let $A$ be an abelian variety over $K$. Assume that no abelian subvariety of $A$ has a non-trivial image over $k$. Let $f : A \to \mathbb{P}^1$ be a rational function. Let $r$ be a positive integer, then there is a bound $B_r \in \mathbb{Z}$ such that for any $P \in A(K)$ and any $x \in C(k)$ either there are some $a \in A(K^{sep})$ and $Q \in A(K)$ such that $f(Q) \in \{0, \infty\}$ and $P = Q + [p^r]a$ or $v_x(f(P)) < B_r$.*

Theorem 5.1.1 follows from Theorem 5.1.2 below by taking $G$ to be the Néron model of $A$, setting $\Gamma = G(C)$, and letting $X$ be the closure of $V(f)$ in $G$.

**Theorem 5.1.2.** *Let $U = C \setminus T$ where $T$ is a finite set of closed points. Let $G$ be a smooth commutative quasi-projective group scheme over $U$. Let $\Gamma \subseteq G(U)$ be a finitely generated subgroup. Consider the sections of $G$ over $U$ as $K$-rational points of $G_\eta$ via $\iota : G(U) \hookrightarrow G_\eta(K)$. Let $X \subseteq G$ be a closed* general *subscheme over $U$. Let $r$ be a positive integer.*

*Then* $\sup\{d_{v_x}(P, X) : x \in U(k), P \in \Gamma \setminus \iota^{-1}((X_\eta(K) \cap \iota(\Gamma)) + [p^r]G_\eta(K^{sep}))\} < \infty$.

The term *general* (to be defined precisely shortly) generalizes the condition that no positive dimensional subvariety of $X_\eta$ descends to $k$. In the case that $G_\eta$ is an abelian variety for which no connected subgroup variety has a non-trivial image over $k$ (in characteristic $p$ this is slightly stronger than assuming that $G_\eta$ has $K/k$-trace zero), every subvariety will be general.

We included the more general statement involving the proximity functions not only to formally strengthen the theorem but because our proof – even for the case of $X$ a hypersurface – passes through the case of $X$ of higher codimension.

*Remark* 5.1.3. The reader unhappy with the scheme-theoretic language may regard $G$ and $X$ simply as varieties over $K$ given together with specific defining equations all of whose coefficients are integral with respect to every valuation $v_x$ for $x \in U(k)$.

**Definition 5.1.4.** Let $G$ be a commutative algebraic group. Let $X \subseteq G$ be a subvariety. $X$ is said to be *general relative to* $k$ if the following condition holds.

Whenever $H \subseteq G$ is a semi-abelian variety, $H_0$ is a semi-abelian variety defined over $k$, $X_0 \subseteq H_0$ is an irreducible subvariety of $H_0$ defined over $k$, $\phi : H \to H_0$ is a map of semi-abelian varieties defined over $K^{sep}$ and $a \in G$ then $(a + \phi^* X_0) \cap X \subseteq Y \subseteq X$ for some $Y$ a coset of a group variety.

With the notation of Theorem 5.1.2, we say that $X$ is *general* if $(X_\eta)_{red}$ is general when considered as a subvariety of $G_\eta$.

*Example* 5.1.5.

- If $X$ is itself a coset of a group variety, then $X$ is general since we may take $Y = X$.
- If $G$ is itself an abelian variety of sufficiently general moduli then no connected subgroup of $G$ has a non-trivial image over $k$. Thus, every subvariety of $G$ is general.
- If $G$ is a unipotent group, then every subvariety is general.
- If $G$ and $X$ are already defined over $k$ and $X$ is irreducible, but is not a coset of a group, then $X$ is not general. While we cannot rule out Theorem 5.1.2 in this case, our methods do not apply.

*Remark* 5.1.6. The restriction on $P \in \Gamma$ in the statement of Theorem 5.1.2 is necessary. Consider for example the case of $G = \mathbb{G}_m$ and $X = 1$. Then one has $d_{v_x}([p]P, X) = p \cdot d_{v_x}(P, X)$.

## 5.2 Valuation Estimates for Hasse-Schmidt Derivations

Our goal in this section is to construct differential operators on $K$ which behave well with respect to all the $K/k$-places.

Let $R$ be a commutative ring. A stack of Hasse-Schmidt derivations (or just HS derivations) on $R$ is a $\mathbf{c}$-iterative $\mathcal{D}$-structure on $R$ relative to $\widehat{\mathcal{D}}(R) = R[[\epsilon]]$ and $c_{i,j} = \binom{i+j}{i}$. Concretely, a stack of HS derivations on $R$ is given by a sequence of functions $\{\partial_n : R \to R\}_{n=0}^{\infty}$ satisfying

- $\partial_0(x) = x$
- $\partial_n(x + y) = \partial_n(x) + \partial_n(y)$
- $\partial_n(x \cdot y) = \sum_{i+j=n} \partial_i(x) \cdot \partial_j(y)$
- $\partial_i \circ \partial_j(x) = \binom{i+j}{i} \partial_{i+j}(x)$

*Remark* 5.2.1. Iterativity is not included in the definition of HS derivations in [32]. Since we will have no use for non-iterative stacks of HS derivations, we have built it into the definition so as to avoid repeating the word "iterative."

On the field $k(t)$ there is a natural choice a stack of HS derivations given by the ring homomorphism $\sigma : k(t) \to k(t)[[\epsilon]]$ determined by $\sigma|_k = \mathrm{id}_k$ and $\sigma(t) = t + \epsilon$. Iterativity corresponds to the commutativity of

$$\begin{array}{ccc} k(t)[[\epsilon]] & \xrightarrow{\sigma} & k(t)[[\epsilon, \eta]] \\ {\scriptstyle\sigma}\uparrow & & \uparrow \\ k(t) & \xrightarrow{\sigma} & k(t)[[\epsilon + \eta]] \end{array}$$

which in our case comes down to $t + (\epsilon + \eta) = (t + \eta) + \epsilon$.

Let $a \in k$. Then $\sigma(t-a)^m = ((t-a)+\epsilon)^m = \sum_{j=0}^{\infty} \binom{m}{j}(t-a)^{m-j}\epsilon^j$ so that $\partial_j(t-a)^m = \binom{m}{j}(t-a)^{m-j}$.

Observe that $Fix(\sigma) = k$.

**Lemma 5.2.2.** *Let $k$ be an algebraically closed field. Let $\{\partial_n\}_{n=0}^{\infty}$ be the stack of HS derivations on $k(t)$ given by $t \mapsto t + \epsilon$. For any $f \in k(t)$ and $x \in \mathbb{P}^1(k)$ one has $v_x(\partial_n f) \geq v_x(f) - n$.*

■ We reduce the question to consider only $f \in k[t]$.

**Claim 5.2.3.** *If this lemma is valid for $f$ and $g$, then it is also valid for $fg$.*

✠

$$\begin{aligned} v(\partial_n(fg)) &= v(\sum_{i+j=n} \partial_i(f)\partial_j(g)) \\ &\geq \min_{i+j=n} \{v(\partial_i(f)) + v(\partial_j(g))\} \\ &\geq \min_{i+j=n} \{(v(f) - i) + (v(g) - j)\} \\ &= v(fg) - n \end{aligned}$$

✠

**Claim 5.2.4.** *If the lemma is valid for $f \neq 0$, then it is also valid for $\frac{1}{f}$.*

✠ We calculate

$$\begin{aligned} 0 &= \partial_n(1) \\ &= \partial_n(f \cdot \frac{1}{f}) \\ &= f\partial_n(\frac{1}{f}) + \sum_{i=1}^{n} \partial_i(f)\partial_{n-i}(\frac{1}{f}) \end{aligned}$$

We proceed with the proof of the claim by induction on $n$. When $n = 0$ the claim is trivial. In general,

$$\begin{aligned} v(\partial_n(\frac{1}{f})) &= v(\sum_{i=1}^{n} \frac{\partial_i f}{f} \partial_{n-i}(\frac{1}{f})) \\ &\geq \min_{1 \leq i \leq n} \{v(\frac{\partial_i f}{f}) + v(\partial_{n-i}(\frac{1}{f}))\} \\ &\geq \min_{1 \leq i \leq n} \{-i + [v(\frac{1}{f}) - (n-i)]\} \\ &= v(\frac{1}{f}) - n \end{aligned}$$

✠

By the two claims it suffices to consider $f \in k[t]$. When $f = 0$, the lemma is obvious so we take $f \neq 0$.

If $x \in \mathbb{A}^1(k)$, then we may expand $f$ as $f = \sum_{i \geq N} f_i(t-x)^i$ where $f_N \neq 0$ and each $f_i \in k$. Then by the $k$-linearity of $\partial_n$, we compute $\partial_n(f) = \sum_{i \geq N} \binom{i}{n} f_i(t-x)^{i-n}$ which visibly has $v_x$ valuation at least $N - n = v_x(f) - n$. When considering the place at $\infty$ given by $v_\infty(f) = -\mathrm{ord}(f)$ observe that each $\partial_n$ actually decreases the order so that $v_\infty(\partial_n f) \geq v(f) \geq v(f) - n$. ■

**Lemma 5.2.5.** *Let $(K, v)$ be a discretely valued field. Let $\{\partial_n\}_{n=0}^{\infty}$ be a stack of HS derivations on $K$ satisfying $\inf\{v(\partial_n x) - v(x) : x \in K^\times\} = B_n > -\infty$. Then there is a unique extension of the stack of HS derivations to completion $K_v$ also satisfying $\inf\{v(\partial_n x) - v(x) : x \in K_v^\times\} = B_n$.*

■ The hypothesis on $\partial_n$ implies that it is a continuous function on $K$. Thus, there is a unique extension of $\partial_n$ to a continuous function on the completion $K_v$. Since $K_v$ is a topological ring, each of the following functions is continuous.

$$
\begin{aligned}
Z(x) &:= \partial_0(x) - x \\
A_n(x, y) &:= \partial_n(x) + \partial_n(y) - \partial_n(x + y) \\
M_n(x, y) &:= \partial_n(xy) - \sum_{i+j=n} \partial_i(x)\partial_j(y) \\
I_{i,j}(x) &:= \partial_i \circ \partial_j(x) - \binom{i+j}{i}\partial_{i+j}(x)
\end{aligned}
$$

As $\{\partial_n\}_{n=0}^{\infty}$ is a stack of HS derivations on $K$, each of these functions is identically zero on $K$ (or $K \times K$ depending on the number of arguments). Since $K$ is dense in $K_v$, these functions must be identically zero on $K_v$ as well. That is, $\{\partial_n\}_{n=0}^{\infty}$ is a stack of HS derivations on $K_v$.

The valuation $v : K_v \to \mathbb{Z} \cup \{\infty\}$ is continuous and takes the value $\infty$ only at zero so that the functions $E_n(x) := v(\partial_n x) - v(x)$ are continuous as maps $K^\times \to \mathbb{Z}$. By the hypotheses, $E_n^{-1}\{N : N \geq B_n\} \supseteq K^\times$. Again since $K$ is dense in $K_v$, we must have $E_n^{-1}\{N : N \geq B_n\} = K_v^\times$. ■

**Lemma 5.2.6.** *If $(K, v)$ is a discretely valued field with an algebraically closed residue field and $\{\partial_n\}_{n=0}^{\infty}$ is a stack of HS derivations on $K$ satisfying $\inf\{v(\partial_n x) - v(x) : x \in K^\times\} = B_n > -\infty$, then for any finite unramified extension $L/K$, there is a unique extension of the stack of HS derivations still satisfying $\inf\{v(\partial_n x) - v(x) : x \in K^\times\} = B_n$.*

■ Since the residue field of $K$ is algebraically closed, $L$ embeds over $K$ into $K_v$ as a valued field. Since $L$ is a finite separable extension of $K$, there is a unique extension of the stack of HS derivations to $L$ ([32] Theorem 9.23). Thus the stack on $L$ must agree with the restriction of the stack on $K_v$. By Lemma 5.2.5 the stated inequalities are true on $K_v$ and hence on $L$. ■

**Lemma 5.2.7.** *Let $(K, v)$ be a complete discretely valued field with a stack of HS derivations satisfying $\inf\{v(\partial_n x) - v(x) : x \in K^\times\} = B_n > -\infty$. Let $L/K$ be a finite separable totally ramified extension of $K$. Then there is a unique extension of the stack of HS derivations to $L$. This stack satisfies $\inf\{v(\partial_n x) - v(x) : x \in L^\times\} = \tilde{B}_n > -\infty$*

*Remark* 5.2.8. In general, $\tilde{B}_n \neq B_n$. We will not need a precise calculation of $\tilde{B}_n$, but we note that it depends on $B_n$, $[L : K]$, the valuation of the the different, and linearly on $n$.

■ That there is unique extension of the stack is Theorem 9.23 of [32]. Let $e := [L : K]$. Let $\pi \in \mathcal{O}_L$ be a uniformizer. For each pair of integers $(a, i)$ with $0 \leq i < e$ and $a \in \mathbb{N}$ define $E_{a,i} := v(\partial_a \pi^i) - v(\pi^i)$. Let $n$ be given. Define $B^{(n)} := \max\{B_j : 0 \leq j \leq n\}$ and $E^{(n)} := \max\{E_{a,i} : 0 \leq a \leq n, 0 \leq i < e\}$.

**Claim 5.2.9.** *We may take $\tilde{E}_n := B^{(n)} + E^{(n)}$.*

✠ In the following calculation each $x_i \in K$ and at least one $x_i$ is not zero.

$$\begin{aligned} v(\partial_n(\sum_{i=0}^{n-1} x_i\pi^i)) &= v(\sum_{i=0}^{n-1} \partial_n(x_i\pi^i)) \\ &\geq \min_{0\leq i<n}\{v(\partial_n(x_i\pi^i))\} \\ &= \min_{0\leq i<n}\{v(\sum_{a+b=n} \partial_b(x_i)\partial_a(\pi^i))\} \\ &\geq \min_{0\leq i<n, a+b=n}\{v(x_i) - B_b + v(\pi^i) - E_{a,i}\} \\ &\geq \min_{0\leq i<n}\{v(x_i) + v(\pi^i)\} - \tilde{B}_n \\ &= v(\sum_{i=0}^{n-1} x_i\pi^i) - \tilde{B}_n \end{aligned}$$

✠

■

**Lemma 5.2.10.** *Let $k$ be an algebraically closed field. Let $K$ be a finitely generated extension of transcendence degree one. There is a stack of HS derivations $\{\partial_n\}_{n=0}^{\infty}$ satisfying*

- $K^{(p)} = \ker \partial_1$ *and*
- *there are constants $B_n \in \mathbb{Z}$ such that for any $K/k$-place $v$ and $x \in K^{\times}$ one has $v(\partial_n x) \geq v(x) - B_n$.*

■ Express $K$ as a finite separable extension of $k(t)$. Let $\{\partial_n\}_{n=0}^{\infty}$ be the stack of HS derivations on $k(t)$ corresponding to $t \mapsto t + \epsilon$. By our calculation above, $\ker \partial_1 = k(t^p)$. Let $\{\partial_n\}_{n=0}^{\infty}$ also denote the unique extension of this stack to $K$ (which exists because the extension is separable). Since $K/k(t)$ is separably algebraic, the extension on the constant fields is also separably algebraic. Thus, $\ker \partial_1 = K^{(p)}$.

As $K$ is a separable extension of $k(t)$, only finitely many places ramify. For any unramified place $v$, Lemmas 5.2.2 and 5.2.6 show that $v(\partial_n x) \geq v(x) - n$ for $x \in K^{\times}$. Use Lemmas 5.2.2 and 5.2.7 to bound the difference $v(\partial_n x) - v(x)$ for each of the finitely many ramified valuations. ■

**Lemma 5.2.11.** *Let $K$ be a field with a stack of HS derivations $\{\partial_n\}_{n=0}^{\infty}$. Let $\Gamma$ be an ordered abelian group. Let $\Sigma$ be a set of $\Gamma$-valuations on $K$. Let $\{B_n\}_{n=0}^{\infty}$ be a sequence of elements of $\Gamma$ with the property that for each $x \in K^{\times}$ and valuation $v \in \Sigma$ one has $v(\partial_n x) \geq v(x) - B_n$.*

*Then for any $\mathcal{D}$-polynomial $F(X_1, \ldots, X_n) \in K\langle X_1, \ldots, X_n\rangle_{\mathcal{D}}$ with $F(0, \ldots, 0) = 0$ there is some $B_F \in \Gamma$ such that for any tuple $\mathbf{a} := (a_1, \ldots, a_n) \in K^n$ and valuation $v \in \Sigma$ if $v(\mathbf{a}) := \min_i\{v(a_i)\} \geq 0$, then $v(F(\mathbf{a})) \geq v(\mathbf{a}) - B_F$.*

■ We proceed by induction on the construction of $F$. If $F = \partial_j X_i$, then the result is already true by hypothesis with $B_F = B_j$.

Suppose $F = GH$ and the result is true for $G$ and $H$. Let $\mathbf{a}$ and $v$ be given with $v(\mathbf{a}) \geq 0$.

$$\begin{aligned} v(F(\mathbf{a})) &= v(G(\mathbf{a})) + v(H(\mathbf{a})) \\ &\geq v(\mathbf{a}) - B_G + v(\mathbf{a}) - B_H \\ &\geq v(\mathbf{a}) - (B_G + B_H) \end{aligned}$$

So we may take $B_F = B_G + B_H$. (N.B.: We used $v(\mathbf{a}) \geq 0$ to obtain the last inequality.)

Suppose now $F = G + H$ and the result is true for $G$ and $H$. Again take $\mathbf{a} \in K^n$ and $v \in \Sigma$ with $v(\mathbf{a}) \geq 0$.

$$\begin{aligned} v(F(\mathbf{a})) &= v(G(\mathbf{a}) + H(\mathbf{a})) \\ &\geq \min\{v(G(\mathbf{a}), v(H(\mathbf{a})\} \\ &\geq \min\{v(\mathbf{a}) - B_G, v(\mathbf{a}) - B_H\} \\ &= v(\mathbf{a}) - \max\{B_G, B_H\} \end{aligned}$$

So we may take $B_F = \max\{B_G, B_H\}$. ■

## 5.3 Manin Maps

In this section we will construct homomorphism $\psi_r : G(K^{sep}) \to U_r(K^{sep})$ which in co-ordinates are polynomials in $\{\partial_m X_i\}$ and have $\ker \psi_r = [p^r]G(K^{sep})$. It was observed in the introduction to [25] that the existence of these maps follows from elimination of quantifiers and imaginaries for the the theory of separably closed fields of imperfection degree one in the differential language. Buium and Voloch have constructed such maps for ordinary abelian varieties in the case of $r = 1$ using jet space and explicit $p$-descent methods [9]. We will construct these maps using a strictly model theoretic argument and also by using jet spaces. Of course, these methods come to the same thing. One could also construct these maps via cohomology (fppf or crystalline) or by an analysis of the formal groups. We will leave these points of view to another paper.

### 5.3.1 Model Theoretic Construction of Manin Maps

All of what is said in this section can be done with only the hypothesis $1 < [K : K^p] < \infty$ by merely changing the notation slightly, but to avoid the use of multi-indices and because we only need the case of $[K : K^p] = p$, we will work only in this case.

Let $t \in K \setminus K^p$. For any $r$, $\{t^i\}_{i=0}^{p^j-1}$ is a basis for $K$ over $K^{p^j}$. In fact, for any $L/K$ a separable extension satisfying $[L : L^p] = p$ this set is still a basis for $L$ over $L^p$. Define co-ordinate functions by the formula

$$x = \sum_{i=0}^{p^j-1} \xi_i^j(x)^{p^j} t^i$$

To use the language of Chapter 2, the $\xi$-functions correspond to a $\mathcal{D}$-structure on $K$ relative to the ring functor $\mathcal{D}_n(K) = K(\sqrt[p^n]{t})$.

The stability of the theory of separably closed fields was first proved by Wood and Shelah [59]. Quantifier elimination in the language with the co-ordinate functions was proved by Delon [15]. Elimination of imaginaries was proved by Messmer [34]. See [35] for a more complete discussion of the model theory of separably closed fields. As noted in the introduction to [25], one can deduce the existence of the Manin maps from these model theoretic properties of separably closed fields.

**Lemma 5.3.1.** *Let $G$ be a commutative algebraic group over $L$ a separably closed field with $[L : L^p] = p$ and fixed a $p$-basis given by $t \in L \setminus L^p$. For any positive integer $r$, there are a unipotent algebraic group $H_r$ and a group homomorphism $\psi_r : G(L) \to H_r(L)$ where $\psi_r$ is given piecewise as a rational function in the $\xi$ co-ordinate functions such that* $\ker \psi_r = [p^r]G(L)$.

■ Without loss of generality, we may assume $G$ is connected. Working in co-ordinates, we may take $G$ to be a definable (in the field language) group. $[p^r]G(L)$ is definable by the formula $x \in [p^r]G \iff (\exists y \in G)\ [p^r]y = x$. By elimination of imaginaries, there is a definable function $\psi : G(L) \to L^m$ for some $m$ such that the fibres of $\psi$ are the cosets of $[p^r]G(L)$. By elimination of quantifiers, $\psi$ may be given piecewise as a rational function in the $\xi$ co-ordinate functions. By the Weil-Hrushovski group chunk theorem [3], the image of $\psi$ embeds into a definable group $H$ such that the generic type of $G(L)$ maps to the generic type of $H$. By [34], there is an embedding $\phi : H \to W$ of $H$ into an algebraic group. Again, $\phi$ may be given by rational functions in the $\xi$ functions. Replacing $W$ with the Zariski closure of the image of $\phi$, we may assume that that the generic type of $H$ maps to the field theoretic generic type of $W$. In the statement of the theorem, $\psi_r = \phi \circ \psi$ and $H_r = W$. Since the exponent of $G(K)/[p^r]G(K)$ is $\leq p^r$, the same is true of $W$. This implies that $W$ is unipotent. ■

### 5.3.2 Jet Space Construction of Manin Maps

We now change the language slightly so that this map may be understood as a differential rational map. We choose an iterative stack of Hasse derivations having the property that $\ker \partial_1 = K^{(p)}$. To be explicit, Let $t \in K \setminus K^{(p)}$ and take the stack determined by the equations $\partial_n t^m = \binom{m}{n} t^{m-n}$. Since $\partial_n$ is linear over $K^{(p^{\lceil \log_p(n)+1 \rceil})}$, these equations do fully determine the stack of HS derivations and also show how to calculate the functions $\partial_n(x)$ in terms of the functions $\xi_i^j(x)$. If $t$ is fixed as a parameter, then one can calculate the $\xi$-functions in terms of the HS derivations as well.

It is shown in [36] that the theory of $K^{sep}$ given with the stack of derivations (but without necessarily fixing $t$) still admits quantifier elimination and elimination of imaginaries.

Recall from Chapter 2 that corresponding to any quasi-projective scheme $X$ over $K$ there is a projective system of jet schemes $\{\nabla_n X\}_{n=0}^{\infty}$.

Let us now use the language jet schemes to re-interpret the construction of the Manin maps.

We use the model theoretic results at only one point.

**Lemma 5.3.2.** *Let $K$ be a separably closed field of imperfection degree one. Let $\{\partial_n\}_{n=0}^{\infty}$ be a stack of HS derivations on $K$ with $\ker \partial_1 = K^{(p)}$. Let $G$ be a commutative algebraic group over $K$. For any $r \in \mathbb{N}$ there is some $N_r \in \mathbb{N}$ and a constructible set $Y_r \subseteq \nabla_{N_r} G$ such that $[p^r]G(K) = \nabla_{N_r}^{-1} Y_r(K)$.*

■ Theorem 3.12 of [36] shows that the theory of $K$ in the differential language admits elimination of quantifiers. This implies that the solution set to the formula $\phi(x) = \ulcorner x \in G(K) \& (\exists y \in G(K))[p^r]y = x \urcorner$ is equivalent to a quantifier-free formula in $x$. Such formulas correspond to Boolean combinations of differential equations. A differential equation is simply an algebraic equation on $\nabla_N(x)$ for $N \gg 0$. $Y_r$ is the constructible subset of $\nabla_N X$ describing these equations. ■

Let us fix some notation. If $f : X \to Y$ is a map of schemes, then $f_* X$ will denote the scheme theoretic image of $X$ in $Y$. In general, $f_* X$ need not be closed in $Y$, but in case $X$ and $Y$ are both group schemes and $f$ is a homomorphism, then $f_* X$ is itself a closed subgroup scheme of $Y$.

**Proposition 5.3.3.** *Let $K$ be a separably closed field with $[K : K^{(p)}] = p$. Let $\{\partial_n\}_{n=0}^{\infty}$ be a stack of HS derivations on $K$ with $\ker \partial_1 = K^{(p)}$. Let $G$ be a commutative algebraic group over $K$. Let $r \in \mathbb{N}$. Then there is a unipotent group $W_r$ and a function $\psi_r : G(K) \to W_r(K)$ which is locally a $\mathcal{D}$-polynomial – in fact, $\psi_r$ is of the form $\phi_r \circ \nabla_{N_r}$ for $\phi_r : \nabla_{N_r} X \to W_r$ a regular function – such that $\ker \psi_r = [p^r]G(K)$.*

■ Before proceeding, we need a little more information about $Y_r$.

**Claim 5.3.4.** *$Y_r$ in Lemma 5.3.2 may be taken to be a group variety. In fact, we may take $Y_r = [p^r]_* \nabla_{N_r} G$.*

✠ Any Zariski closed subvariety of $Y_r$ containing the image of $[p^r]G(K)$ under $\nabla_{N_r}$ will work. So we might as well take the Zariski closure of this image. By the very definition of the jet space, $\nabla_{N_r}(G(K))$ is Zariski dense in $\nabla_{N_r}G$. Hence, $[p^r] \circ \nabla_{N_r}(G(K)) = \nabla_{N_r}([p^r]G(K))$ is Zariski dense in the algebraic group $[p^r]_* \nabla_{N_r} G$. ✠

Let $W_r$ be the quotient $\nabla_{N_r}G/[p^r]_*\nabla_{N_r}G$ and let $\phi_r : \nabla_{N_r}G \to W_r$ be the quotient map. The map $\psi_r$ is then $\phi_r \circ \nabla_{N_r}$. Since $W_r(K^{alg})$ has exponent at most $p^r$, it must be unipotent. ■

## 5.4 Uniformities in the Function Field Mordell-Lang Conjecture

We will need to make concrete some of the uniformities inherent in Hrushovski's Theorem [25].

Before we proceed, recall that a model $L$ is $\aleph_1$-compact if whenever $\{X_i\}_{i=0}^{\infty}$ is a countable collection of definable subsets of $L$ and $\bigcap_{i=0}^{N} X_i \neq 0$ for each $N$, then $\bigcap_{i=0}^{\infty} X_i \neq 0$. Recall that if $L$ is a separably closed field of finite imperfection degree given with a stack of HS derivations, then each of the $X_i$'s may be taken to be a finite Boolean combination of solutions to differential equations.

Let us first restate the main theorem of [25].

**Theorem 5.4.1 (Hrushovski).** *Let $K$ be a separably closed field of characteristic $p$. Assume the $1 < [K : K^{(p)}] < \infty$. Let $G$ be a semi-abelian variety over $K$. Let $L$ be an $\aleph_1$-compact separable separably closed extension of $K$ with the property that $L = L^{(p)}K$ and $[L : L^{(p)}] = [K : K^{(p)}]$. Define $[p^\infty]G(L) := \bigcap_{n=1}^{\infty}[p^n]G(L)$. Let $X \subseteq G$ be a subvariety of $G$ defined over $L$. Assume that $X$ is general relative to $\bigcap_{n=1}^{\infty} L^{(p^n)}$. Then there finitely many group subvarieties $G_1, \ldots, G_n$ of $G$ and points $a_1, \ldots, a_n \in G(L)$ such that $X \cap [p^\infty]G(L) = (\bigcup_{i=1}^{n} a_i + G_i) \cap [p^\infty]G(L)$.*

We observe that one may replace "semi-abelian variety" by "commutative algebraic group" in the hypotheses on $G$ in Hrushovski's theorem since for any commutative algebraic group $G$ one has that $[p^r]_*G$ is a semi-abelian variety for $r \gg 0$.

In the next proposition we will use the compactness theorem of first order logic to re-interpret the above statement to give a uniformity result over $K$.

**Proposition 5.4.2.** *Let $K$ be a separably closed field of characteristic $p$ with $1 < [K : K^{(p)}] < \infty$. Let $k = \cap K^{p^n}$. Let $G$ be a commutative algebraic group over $K$. Let $X \subseteq G$ be a subvariety of $G$ defined over some $L$ a separably closed separable extension of $K$ having the same $p$-basis. Assume that $X$ is general with respect to $\bigcap_{n=0}^{\infty} L^{(p^n)}$. Then there is a finite set $\Xi$ of semi-abelian subvarieties of $G$ and integers $N$ and $M$ such that for any point $a \in G(L)$ one has*

$$(X + a) \cap [p^N]G(L) \quad = \quad \bigcup_{i=1}^{m} (a_i + H_i) \cap [p^N]G(L)$$

*with $0 \leq m \leq M$ and $H_i \in \Xi$.*

■ If this proposition were false, then for each natural number $N$ and finite set of semi-abelian subvarieties $\Xi = \{G_1, \ldots, G_n\}$ of $G$ (possibly listed with multiplicity) defined over $K$, it would be consistent with the theory of $L$ that $(\exists b)(\forall a_1, \ldots, a_n)(X + b) \cap [p^N]G \neq (\bigcup_{i=1}^{n} a_i + G_i) \cap [p^N]G$. By the compactness theorem, the following set of formulas has a model:

1. the elementary diagram of $L$

2. for each natural number $N$ and finite sequence of semi-abelian subvarieties of $G$ defined over $K$, $G_1, \ldots, G_n$, the formula

$$(\forall a_1, \ldots, a_n \in G)(\exists y)(\exists z)[y = [p^N]z \text{ and } (y - \mathbf{c} \in X \setminus \bigcup_{i=1}^{n} a_i + G_i \text{ or } y - \mathbf{c} \in [\bigcup_{i=1}^{n} a_i + G_i) \setminus X])]$$

Let $M$ be such a model which is $\aleph_1$-compact and let $\mathbf{c} \in G(M)$ be the point interpreting the formal symbol $\mathbf{c}$.

**Claim 5.4.3.** *1. $M$ is a separable separably closed extension of $K$ with $M = M^{(p)}K$ and $[M : M^{(p)}] = [M : M^{(p)}]$.*

*2. There is no finite sequence $G_1, \ldots, G_n$ of semi-abelian subvarieties of $G$ defined over $K$ and points $a_1, \ldots, a_n \in G(M)$ such that $\mathbf{c} + X \cap [p^\infty]G(M) = (\bigcup_{i=1}^{n} a_i + G_i) \cap [p^\infty]G(M)$.*

✠

1. Since $M$ is a model of the elementary diagram of $L$, the extension $M/L$ is elementary and hence $M/K$ is elementary. The property of being separably closed is first-order so $M$ is separably closed. The other property may be expressed by fixing a basis $B$ for $K$ over $K^{(p)}$ and insisting that $B$ also be a basis for $M$ over $M^{(p)}$.
2. Suppose that $G_1, \ldots, G_n$ are semi-abelian subvarieties of $G$ defined over $K$ and $a_1, \ldots, a_n \in G(M)$ such that $(X + \mathbf{c}) \cap [p^\infty]G(M) = (\bigcup_{i=1}^{m} a_i + G_i) \cap [p^\infty]G(M)$. Since $M$ is $\aleph_1$-compact, for $N$ sufficiently large we must have $(X + \mathbf{c}) \cap [p^N]G(M) = (\bigcup_{i=1}^{m} a_i + G_i) \cap [p^N]G(M)$. This violates a condition on $\mathbf{c}$.

✠

Since every semi-abelian subvariety $G$ is defined over the separable closure of the field of definition of $G$, Hrushovski's theorem implies that in fact $(X + \mathbf{c}) \cap [p^\infty]G(M) = \bigcup_{i=1}^{m} (a_i + G_i) \cap [p^\infty]G(M)$ for some $a_i \in G(M)$ and $G_i$ semi-abelian subvarieties defined over $K$. This gives the contradiction. ■

## 5.5 Main Theorem

For the next two lemmas, $K$ and $C$ continue to have the meaning assigned before the statement of Theorem 5.1.1, but $G$ will denote a commutative algebraic group over $K$ rather than a group scheme over $U$.

**Lemma 5.5.1.** *Let $G$ be a commutative algebraic group over $K$. Let $\Lambda \subseteq G(K)$ be a finitely generated subgroup. Let $a \in G(K)$. Let $r \in \mathbb{Z}_+$. Let $V \subseteq G$ be an affine open containing $a$. Take affine co-ordinates on $V$ so that $\mathcal{I}_{\{a\}}(V) = (y_1, \ldots, y_n)$. Then there is a bound $B_r$ such that for any $\lambda \in \Lambda \setminus (a + [p^r]G(K^{sep}))$ and any $x \in C(k)$ one has $\min\{v(y_i(\lambda))\} < B_r\}$.*

■ Let $\psi_r : G(K) \to U_r(K)$ be the map of Proposition 5.3.1 with kernel $G_\eta(K) \cap [p^r]G_\eta(K^{sep})$. The underlying variety of $U_r$ is an affine space. Let $\tau : U_r \to U_r$ be the translation (with respect to the usual additive group structure on affine space)which takes $\psi_r(a)$ to the origin. Set $\phi_r := \tau \circ \psi_r$.

By our hypothesis on $\Lambda$, $\Lambda/(\Lambda \cap [p^r]G(K^{sep}))$ is finite. Thus $\phi_r(\Lambda)$ is a finite set.

For any particular $\mathbf{b} \in \mathbb{A}^n(K) \setminus \{(0, \ldots, 0)\}$, there is a constant $C$ such that $-C \leq v_x(\mathbf{a}) \leq C$ for $x \in U(k)$. Let $C$ be the maximum over these constants for the non-zero elements of $\phi_r(\Lambda)$. By Lemma 5.2.11, for $P \in V(K) \setminus [p^r]G(K^{sep})$ and $x \in U(k)$ either $\min\{v(y_i(P))\} < 0$ or we have $C \geq v_x(\phi_r(P)) \geq v_x(P) - B_{\phi_r}$. The bound is then $C + B_r$. ■

**Lemma 5.5.2.** *Let $G$ be a commutative algebraic group over $K$. Let $H \subseteq G$ be an algebraic subgroup. Let $a \in G(K)$. Let $V \subseteq G$ be an affine open. Let $(f_1, \ldots, f_n) = \mathcal{I}_{a+H}(V)$. Let $\Lambda \subseteq G(K)$ be a finitely generated subgroup. Let $r \in \mathbb{Z}_+$. Then there is a bound $B_r$ such that for any $x \in C(k)$ and any $\lambda \in \Lambda \setminus [(a + H(K)) + [p^r]G(K^{sep})]$, one has $\min\{v(f_i(\lambda))\} < B_r$.*

■ Apply Lemma 5.5.1 to the algebraic group $G/H$. ■

We return now to the notation of Theorem 5.1.2.

**Lemma 5.5.3.** *Let $X \subseteq G$ be a closed subscheme whose generic fibre $X_\eta$ is a coset of an algebraic subgroup of $G$. Let $r \in \mathbb{Z}_+$. Then $d_{v_x}(P, X)$ is bounded independently of $x \in U(k)$ and $P \in \Gamma \setminus \iota^{-1}(X_\eta(K) + [p^r]G_\eta(K^{sep}))$.*

■ Take a finite affine cover $\mathcal{V} := \{V_i\}_{i=1}^N$ of $G$ and equations $(f_1^{(i)}, \ldots, f_{m_i}^{(i)}) = \mathcal{I}_X(V_i)$ for $X$ over $U$ relative to this cover. The distance from $P$ to $X$ may now be computed in terms of these equations as $d_{v_x}(P, X) = \min\{v_x(f_j^{(i)}(P)) : P \in V_i(U)\}$. By Lemma 5.5.2 these values are bounded on $\Gamma \setminus \iota^{-1}(X_\eta(K) + [p^r]G_\eta(K^{sep}))$. ■

We can turn now to the proof of the main theorem.

■ Let $N$ be large enough so that each translate of $[p^N]G_\eta(K^{sep})$ meets $X_\eta$ as does a finite union of translates of group subvarieties. Let $\Xi$ be the finite set of group varieties of Proposition 5.4.2.

Since the statement is stronger for $r$ larger, we may assume that $r \geq N$.

Let $\Sigma$ be a set of coset representatives for $\iota^{-1}(\iota(\Gamma) \cap [p^N]G(K^{sep}))$ in $\Gamma$. For $\sigma \in \Sigma$ let $G_1^\sigma, \ldots, G_{n_\sigma}^\sigma \in \Xi$ and $a_1^\sigma, \ldots, a_{n_\sigma}^\sigma \in G(U)$ such that

$$X_\eta(K^{sep}) \cap ([p^N]G_\eta(K^{sep}) + \iota(\sigma)) \quad = \quad [\bigcup_{i=1}^{n_\sigma} (\iota(a_i^\sigma) + G_i^\sigma)(K^{sep})] \cap [p^N]G_\eta(K^{sep})$$

Since $\#\Gamma/\iota^{-1}(\iota(\Gamma) \cap [p^N]G_\eta(K^{sep}))$ is finite, it suffices to show for each $\sigma \in \Sigma$ that there is some $C_r^\sigma \in \mathbb{N}$ such that for any $x \in U(k)$ and any $P \in \iota^{-1}[([p^N]G_\eta(K^{sep}) \cap \iota(\Gamma)) + \sigma] \setminus \iota^{-1}[(X_\eta(K) \cap \iota(\Gamma)) + [p^r]G_\eta(K^{sep})]$ one has $d_{v_x}(P, X) \leq C_r^\sigma$. We can then set $C_r = \max_{\sigma \in \Sigma} C_r^\sigma$.

By Lemma 4.2.3, for $P \in \Gamma$ with $\iota(P) \in [p^N]G(K^{sep}) + \iota(\sigma)$, the distance to $X$ is uniformly (in $P$ and in $v_x$) comparable to the distance to $\bigcup_{i=1}^{n_\sigma} \overline{a_i^\sigma + G_i^\sigma}$.

By Lemma 4.1.7, it suffices to bound the distance to $\overline{a_i^\sigma + G_i^\sigma}$.

By Lemma 5.5.3, for $P \in \Gamma$ with $\iota(P) \notin (\iota(a_i^\sigma) + G_i)(K) + [p^r]G(K^{sep})$ $\quad d_{v_x}(P, \overline{a_i^\sigma + G_i^\sigma})$ is bounded independently of $P$ and $x$.

Putting this all together, the result follows. ■

# Chapter 6

# A conjecture of Tate and Voloch

## Notation

$\mathbb{C}_p$ denotes the completion of the algebraic closure of $\mathbb{Q}_p$, the $p$-adic numbers. $\mathbb{C}_p$ is given together with the $p$-adic valuation normalized to have $v_p(p) = 1$. Unless otherwise noted, distances will be computed with respect to $v_p$.

If $G$ is a commutative group scheme and $n \in \mathbb{Z}_+$, then $G[n] := \ker([n] : G \to G)$. Let $\ell$ be any rational prime. We denote by $G[\ell^\infty]$ the direct limit $\varinjlim G[\ell^m]$.

For $G$ any abelian group and $p$ any prime number, the torsion group of $G$ is $G_{tor} := \{x \in G : [m]x = 0 \text{ for some } m \in \mathbb{Z}_+\}$ and the prime to $p$ torsion group is $G_{p'-tor} := \{x \in G : [m]x = 0 \text{ for some } m \in \mathbb{Z}_+ \text{ with } (m,p) = 1\}$.

If $L$ is a field and $\sigma$ is an automorphism of $L$, then $\mathrm{Fix}(\sigma)$ is the subfield of $L$ fixed by $\sigma$.

## 6.1 Statements of Conjecture, Known Results, and Theorem

Tate and Voloch proved the following approximation theorem on linear forms in $p$-adic roots of unity in [53]:

**Theorem 6.1.1 (Tate, Voloch).** *Let $(a_1, \ldots, a_n) \in \mathbb{A}^n(\mathbb{C}_p)$. Then there is a constant $\epsilon > 0$ such that for any $n$-tuple $(\zeta_1, \ldots, \zeta_n)$ of roots of unity, either $\sum_{i=1}^n a_i \zeta_i = 0$ or $|\sum_{i=1}^n a_i \zeta_i|_p \geq \epsilon$.*

They observed that this theorem may be interpreted as a special case of the following conjecture.

**Conjecture 6.1.2 (Tate, Voloch).** *Let $G$ be a semi-abelian variety over $\mathbb{C}_p$. Let $X \subseteq G$ be a closed subscheme defined over $\mathbb{C}_p$. There is a constant $N \in \mathbb{Q}$ such that for any torsion point $P \in G(\mathbb{C}_p)_{tor}$ either $P \in X$ or $d(P, X) \leq N$.*

Several people have made progress on Conjecture 6.1.2. The first result in this direction would be the theorem of Nigel and Lutz showing that the distance from non-zero torsion points to zero on an elliptic curve is bounded (in this case $G$ is an elliptic curve and $X = 0$). This theorem was generalized by Mattuck to the case of $G$ any commutative algebraic group. Theorem 6.1.1 may be interpreted as an instance of this conjecture by taking $G = \mathbb{G}_m^n$ and $X$ the hypersurface defined by $\sum_{i=1}^n a_i X_i = 0$. For $G = \mathbb{G}_m^n$ one can reduce the case of $X$ arbitrary to the case of $X$ a hyperplane.

Buium proved a version of this conjecture under a restrictive hypothesis on $G$.

**Theorem 6.1.3 (Buium).** *Let $G$ be a semi-abelian variety over $\mathbb{C}_p$. Suppose that $G$ has good reduction and that there is an algebraic endomorphism $F \in \mathrm{End}(G)$ which reduces to a Frobenius. Let $T := \{P \in G(\mathbb{C}_p)_{tor} : (\exists n \in \mathbb{Z}_+)\ F^n(P) = P\}$. Then for any subvariety $X \subseteq G$ defined over $G$ there is some $r \in \mathbb{Q}$ such that for any $P \in T$ either $P \in X$ or $d(P, X) \leq r$.*

Voloch proved a version of this conjecture with $P$ restricted to be a point in the formal group.

**Theorem 6.1.4 (Voloch).** *Let $K$ be a complete subfield of $\mathbb{C}_p$ finitely ramified over $\mathbb{Q}_p$. Let $G$ be a semi-abelian variety over $K$. Let $X \subseteq G$ be a subvariety defined over $K$. There are constants $r \in \mathbb{Q}$ and $N \in \mathbb{N}$ depending only on the degree of $X$ and a set of points $\Sigma$ of size at most $N$ such that for any torsion point $P$ in the formal group of $G$ either $P \in X \cup \Sigma$ or $d(P, X) \leq r$.*

Voloch observes that Theorems 6.1.3 and 6.1.4 may be combined to give

**Theorem 6.1.5 (Voloch).** *Let $K$ be a complete subfield of $\mathbb{C}_p$ finitely ramified over $\mathbb{Q}_p$. Let $G$ be a semi-abelian variety over $K$ having good reduction and an algebraic endomorphism $F$ lifting a Frobenius. Let $X \subseteq G$ be a subvariety defined over $K$. Then there is a constant $r \in \mathbb{Q}$ such that for any torsion point $P \in G(\mathbb{C}_p)_{tor}$ either $P \in X$ or $d(P, X) \leq r$.*

Using methods similar to those we will be using, Hrushovski obtained

**Theorem 6.1.6 (Hrushovski).** *Let $K$ be a complete subfield of $\mathbb{C}_p$ having a finite residue field. Let $G$ be a semi-abelian variety over $K$ having good reduction. Let $X \subseteq G$ be a subvariety of $G$ defined over $\mathbb{C}_p$. There is a constant $r \in \mathbb{Q}$ such that for any prime to $p$ torsion point $P \in G(\mathbb{C}_p)_{p'-tor}$ either $P \in X$ or $d(P, X) \leq r$.*

The proof of Theorem 6.1.6 actually gives a stronger result. Without any extra effort, one may take $P$ to be in the group generated by the prime to $p$ torsion points and the points of the étale part of the $p$-divisible group of $G$ (or to be more accurate, of a semi-abelian model of $G$ over $\mathrm{Spec}\mathcal{O}_{\mathbb{C}_p}$).

We prove a strengthening of Hrushovski's theorem to the case of bad reduction. We revert to scheme-theoretic language with the notation following that introduced in Chapter 4 with the exception that we do not explicitly refer to the valuation used to define the distance functions.

**Theorem 6.1.7.** *Let $K$ be a finite extension of $\mathbb{Q}_p$. Let $G$ be a semi-abelian scheme over $\mathcal{O}_K$. Let $X \subseteq G$ be a closed subscheme of $G$ defined over $\mathcal{O}_{\mathbb{C}_p}$. There is a constant $r \in \mathbb{Q}$ such that for any prime to $p$ torsion point $P \in G(\mathcal{O}_{\mathbb{C}_p})_{p'-tor}$ either $\iota(P) \in X_\eta(\mathbb{C}_p)$ or $d(P, X) \leq r$.*

We prove this theorem as a special case of a theorem on groups defined by difference equations.

**Theorem 6.1.8.** *Let $(L, v, \Gamma)$ be an algebraically closed valued field of characteristic zero. Let $K$ be a discretely valued subfield of $L$. Let $G$ be a semi-abelian scheme defined over $\mathcal{O}_K$. Let $P(X) \in \mathbb{Z}[X]$ be a polynomial having no cyclotomic factors. Let $\daleth = G(\mathcal{O}_L)_{tor}, G(\mathcal{O}_L)_{p'-tor},$ or $G[p^\infty](\mathcal{O}_L)$. Call an automorphism $\sigma$ of $L$* good *if $\sigma$ satisfies*

1. *$v(\sigma(x)) = v(x)$ for every $x \in L$,*
2. *$\sigma(G(L)) = G(L)$, and*
3. *$\{x \in G(L) : P(\sigma) \circ (\sigma - 1)(x) = 0\} \supseteq \daleth$.*

*Assume that there are good automorphisms. Assume that the common fixed field of the good automorphisms is $K$ and for any algebraic subgroup $H$ of $G_\eta$, there are only finitely many torsion points in $(G_\eta/H)(K)$.*

*Then for any closed subscheme $X \subseteq G$ defined over $\mathcal{O}_L$ there is some constant $\gamma \in \Gamma$ such that for any point $x \in \daleth$ either $\iota(x) \in X_\eta$ or $d(x, X) \leq \gamma$.*

The arguments that go into the proof of Theorem 6.1.1 yield information for groups of rank greater than zero in many cases. We list now one easy case.

**Proposition 6.1.9.** *Let $K$, $L$, $v$, $\Gamma$, $G$, and $P$ be as in Theorem 6.1.8. Let $\daleth \subseteq G(\mathcal{O}_L)$ be a subgroup. Suppose that there is some automorphism $\sigma \in \mathrm{Gal}(L/K)$ such that $v(x) = v(\sigma(x))$ for any $x \in L$ and $\daleth \subseteq \ker(P(\sigma) : G(L) \to G(L))$.*

*Let $X \subseteq G$ be a closed subscheme defined over $\mathcal{O}_L$. Then there is some $\gamma \in \Gamma$ such that for any $x \in \daleth$ either $\iota(x) \in X_\eta$ or $d(x, X) \leq \gamma$.*

*Remark* 6.1.10. Theorem 6.1.1 may be seen as a consequence of Theorem 6.1.8. Take $(L, v, \Gamma) = (\mathbb{C}_p, v_p, \mathbb{Q})$, $G = \mathbb{G}_m^r$ and $P(X) = (X - p)(X - \ell)$ where $\ell$ is some prime different from $p$.

*Remark* 6.1.11. By Grothendieck's semi-stable reduction theorem (Theorem 3.6 of [18]), if $G$ is a semi-abelian variety over $K$ a finite extension of $\mathbb{Q}_p$, then there is a semi-abelian scheme $\mathfrak{G}$ over $\mathcal{O}_{K'}$ where $K'$ is a finite extension of $K$ for which $\mathfrak{G}_\eta = G \times_{\mathrm{Spec}K} \mathrm{Spec}K'$. So the requirement in Theorem 6.1.7 that $G_\eta$ actually has a semi-abelian integral model does not restrict the class of semi-abelian varieties considered.

*Remark* 6.1.12. Tate and Voloch interpret their work in [53] as an instance of differential algebra with respect to the "$p$-derivation" given by "differentiating with respect to $p$." With this language, a point $P \in G(\mathcal{O}_K)$ is thought of as a "curve" and the $p$-adic distance from $P$ to $X$ may be interpreted as the order of contact of the "curve" $P$ with $X$.

It will take a little work to show that Theorem 6.1.7 follows from Theorem 6.1.8. We do this in Section 6.3. We will discuss in Section 6.4 some other cases of Conjecture 6.1.2 which follow from Theorem 6.1.8.

## 6.2 General Theorem

In this section we will give a proof of Theorem 6.1.8. Throughout this section $L$, $K$, $v$, $\Gamma$, $G$, and $X$ will have the meaning assigned to them in the statement of Theorem 6.1.8.

### 6.2.1 Easy Reductions

**Lemma 6.2.1.** *Let $H$ be an algebraic subgroup of $G_\eta$ defined over $K'$, a finite extension of $K$ of degree $m$. Let $\sigma$ be a good automorphism of $L$ relative to $G$ and $P(X)$. Let $\mathcal{Q}$ be a semi-abelian scheme over* $\mathrm{Spec}\mathcal{O}_{K'}$ *having $\mathcal{Q}_\eta = (G_\eta/H)$. Then $\sigma$ is good relative to $\mathcal{Q}$ and $(P(X))^{m!}$.*

■

**Claim 6.2.2.** *The map $G_\eta(L)_{tor} \xrightarrow{\ \pi\ } (G_\eta/H)(L)_{tor}$ is surjective.*

✠ Let $m := [H : H^0]$ be the index of the connected component of $H$ in $H$. Let $x \in (G_\eta/H)[n](L)$. Let $y \in \pi^{-1}\{x\}$. Then $[n]y = a \in H$. So that $[mn]y = [m]a \in H^0$. $H^0$ is divisible so we may find $a' \in H^0$ such that $[mn]a' = -[m]a$. Let $y' = y + a'$. So $\pi(y') = \pi(y) = x$ and $[mn]y' = [mn](y + a') = [mn]y - [mn]a' = [m]a - [m]a = 0$. ✠

**Claim 6.2.3.** *If $\varphi : A \twoheadrightarrow B$ is a surjective map between torsion abelian groups, then for any prime $\ell$, the restriction of $\varphi$ to the $\ell$-primary component of $A$ is surjective onto the $\ell$-primary component of $B$.*

✠ Let $x \in B$ with $[\ell^n]x = 0$. As $\varphi$ is surjective, there is some $y \in A$ with $\varphi(y) = x$. Since $A$ is torsion, $[m]y = 0$ for some $m \in \mathbb{Z}_+$. Write $m = m'\ell^j$ where $(m', \ell) = 1$. Let $a, b \in \mathbb{Z}$ such that $am' + b\ell = 1$. Then, $x = \varphi(y) = \varphi((am' + b\ell)y) = \varphi([am']y)$. Moreover, $[\ell^j][am']y = [a][m]y = 0$. ✠

Since $H$ is defined over $K'$ for any $\sigma \in \mathrm{Gal}(L/K)$ there can be at most $m$ elements of the set $\{\sigma^j(H)\}_{j=1}^\infty$. Thus, $\sigma^{m!}$ acts on $G/H$. The following diagram is commutative.

$$\begin{array}{ccccccc}
G(L) & \xrightarrow{P(\sigma)} & G(L) & \xrightarrow{P(\sigma)} & \cdots & \xrightarrow{P(\sigma)} & G(L) \\
{\scriptstyle\pi}\big\downarrow & & \big\downarrow{\scriptstyle\pi} & & \cdots & & \big\downarrow{\scriptstyle\pi} \\
(G/H)(L) & \xrightarrow{P(\sigma)} & (G/H)^\sigma(L) & \xrightarrow{P(\sigma)} & \cdots & \xrightarrow{P(\sigma)} & (G/H)^{\sigma^{m!}}(L)
\end{array}$$

Restricting to the $\ell$ power torsion points (for appropriate $\ell$ depending on the choice of $\urcorner$), since $\sigma$ is good on $G$, we have

$$\begin{aligned} 0 &= \pi \circ P(\sigma)^{m!}|_{G[\ell^\infty](L)} \\ &= P(\sigma)|^{m!}_{(G/H)[\ell^\infty](L)} \circ \pi_{G[\ell^\infty](L)} \end{aligned}$$

By Claim 6.2.2, $\pi_{G[\ell^\infty](L)}$ is surjective so that it must be that $P(\sigma)|_{(G/H)[\ell^\infty](L)} = 0$. ■

*Remark* 6.2.4. The reader may have noticed that $\mathcal{Q}$ did not appear at all in the last proof. This is not an accident. The existence of integral models for the semi-abelian varieties under consideration facilitates our proof, but we don't need to use much beyond their existence.

**Lemma 6.2.5.** *Let $S$ be the stabilizer of $X_\eta$ in $G_\eta$. $\tilde{G}$ be a semi-abelian integral model for $G_\eta/S$ and let $\tilde{X}$ be the closure of $X_\eta/S$ in $\tilde{G}$. If Theorem 6.1.8 holds with $\tilde{G}$ and $\tilde{X}$ replacing $G$ and $X$, then it is true of $G$ and $X$.*

■ By Proposition 2.7 of Chapter I of [16], The map $\pi_\eta : G_\eta \to (G_\eta/S)$ extends to a morphism $\pi : G \to \tilde{G}$. Let $\gamma \in \Gamma$ be the bound provided by the hypothesis that Theorem 6.1.8 is true of $\tilde{G}$ and $\tilde{X}$. Since $\pi$ is a group scheme homomorphism, if $P \in G(\mathcal{O}_L)_{tor}$, then $\pi(P) \in \tilde{G}(\mathcal{O}_L)_{tor}$. Since $S$ stabilizes $X_\eta$, $(\pi^*\tilde{X})_\eta = X_\eta$. By Lemma 4.1.8, $d(P, \pi^*\tilde{X}) = d(\pi(P), \tilde{X})$.

So if $\iota(P) \notin X_\eta$, then $\iota \circ \pi(P) \notin \tilde{X}_\eta$. By Lemma 4.2.3, there is an integer $n$ and element $\delta \in \Gamma$ such that $d(P, \pi^*X) = n \cdot d(\pi(P), \tilde{X}) + \delta$. By hypothesis, if $P \in G(\mathcal{O}_L)_{tor} \setminus \iota^{-1}X_\eta(L)$, then $d(\pi(P), \tilde{X}) \leq \gamma$.

Putting this all together: if $P$ is a torsion point whose image in the generic fibre does not lie on the generic fibre of $X$, then the distance from $P$ to $X$ is comparable to the distance to $\pi^*\tilde{X}$ which is equal to the distance from $\pi(P)$ to $\tilde{X}$ which is bounded by $\gamma$. ■

**Lemma 6.2.6 (Mattuck).** *There is a constant $\gamma \in \Gamma$ such that for any torsion point $P \in G(\mathcal{O}_L)_{tor}$ either $P = 0$ or $d(P, 0) \leq \gamma$.*

■ This is a theorem of Mattuck [33] at least in the case that $K$ is a $\mathfrak{p}$-adic field. The proof goes through in general.

For the reader's convenience we sketch the proof.

Replace $L$ with $\mathbf{L}$ a maximal completion. See [46] chapter 2 for a proof that $\mathbf{L}$ exists and is algebraically closed.

If $P$ does not reduce to zero, then $d(P, 0) = 0$ so we need not worry about $P$.

There is a natural isomorphism $\{x \in G(\mathcal{O}_{\mathbf{L}}) : \pi_0(x) = 0\} \cong \widehat{G}(\mathfrak{m}_{\mathbf{L}})$. There is a neighborhood of the origin in $\widehat{G}(\mathfrak{m}_{\mathbf{L}})$ on which the formal logarithm of $\widehat{G}$ converges to define a homomorphism $\log_{\widehat{G}} : \widehat{G}(\mathfrak{m}_{\mathbf{L}}) \to \widehat{\mathbb{G}}_a^M(\mathfrak{m}_{\mathbf{L}})$ for some $M$. Moreover, there is a neighborhood of the identity in $\widehat{\mathbb{G}}_a^M(\mathfrak{m}_{\mathbf{L}})$ on which the formal exponential of $\widehat{G}$ is defined and gives an inverse to the logarithm. Thus, a neighborhood of the identity in $G$ is isomorphic to a neighborhood of the identity in $\mathbb{G}_a^M$ so that near the origin of $G$ there can be no other torsion points. ■

**Lemma 6.2.7.** *If $a \in G(\mathcal{O}_L)$ is any point then there is a constant $\gamma \in \Gamma$ such that for any torsion point $P \in G(L)_{tor}$ either $P = a$ or $d(P, a) \leq \gamma$.*

■ Let $\gamma$ be the bound computed in Lemma 6.2.6. By Lemma 4.1.9, the distance function is a metric. So if $P$ and $Q$ are distinct torsion points and $d(P, a) > \gamma$ and $d(Q, a) > \gamma$, then $d(P, Q) > \gamma$. By Lemma 4.1.10, this implies $d(P - Q, 0) > \gamma$ which contradicts Lemma 6.2.6. ■

**Lemma 6.2.8.** *If $H$ is an algebraic subgroup of $G_\eta$ defined over $L$ and $a \in G(\mathcal{O}_L)$ is any point, then there is some $\gamma \in \Gamma$ such that for any torsion point $P \in G(\mathcal{O}_L)_{tor}$ either $\iota(P) \in \iota(a) + H$ or $d(P, \overline{a + H}) \leq \gamma$.*

■ Let $\mathcal{Q}$ be a semi-abelian model of $G_\eta/H$ over $\mathrm{Spec}\mathcal{O}_L$. As above, let $\pi : G \to \mathcal{Q}$ be a map which on the generic fibre is the quotient map. Let $\gamma$ be the bound for the distance from torsion points to $\pi(a)$ in $\mathcal{Q}$ provided by Lemma 6.2.7. By Lemma 4.1.8, $d(P, \pi^*\pi(a)) = d(\pi(P), \pi(a))$. By construction $(\pi^*\pi(a))_\eta = \iota(a) + H$ so that by Lemma 4.2.3 there is some $n \in \mathbb{Z}_+$ and $\delta \in \Gamma$ such that $d(P, \overline{\iota(a) + H}) \leq n \cdot d(P, \pi^*\pi(a)) + \delta$. So for $P \in G(\mathcal{O}_L)_{tor}$ with $\iota(P) \notin \iota(a) + H$, we have $d(P, \overline{\iota a + H}) \leq n\gamma + \delta$. ■

### 6.2.2 Model Theory of Difference Fields

A difference ring is a ring $R$ given together with a ring endomorphism $\sigma : R \to R$. In the language of Chapter 2, a difference ring is a $\mathcal{D}$-ring with respect to the ring functor $\mathcal{D}_1(R) = R^2$.

The theory of difference fields has a model companion which in the literature is known as ACFA. ACFA is axiomatized by saying that a model $\mathcal{K}$ is an algebraically closed field and $\sigma$ is an automorphism of $\mathcal{K}$ together with the axiom schema:

Let $V$ be an irreducible variety over $\mathcal{K}$. Let $V^\sigma$ denote the variety obtained by applying $\sigma$ to the equations defining $V$. If you like, $V^\sigma$ fits into the Cartesian square

$$\begin{array}{ccc} V^\sigma & \longrightarrow & V \\ \downarrow & & \downarrow \\ \mathrm{Spec}\mathcal{K} & \xrightarrow{\sigma^*} & \mathrm{Spec}\mathcal{K} \end{array}$$

Let $W \subseteq V \times V^\sigma$ be a closed irreducible subvariety with $\dim W = \dim V$. Let $V \times V^\sigma \xrightarrow{\pi} V$ denote the projection onto the first co-ordinate. Assume that $\pi(W)$ is Zariski dense in $V$. Let $W' \subseteq W$ be a proper closed subvariety. Then there is a point of the form $(x, \sigma(x)) \in (W \setminus W')(\mathcal{K})$.

Models of ACFA are called *transformally closed fields*.

ACFA is a simple theory (in the sense of Shelah) [29] and Chatzidakis and Hrushovski have shown in characteristic zero that the Zil'ber trichotomy holds for minimal types [12]. The main consequences of these facts for groups definable in ACFA are described in [21] and [11].

Let us now spell out what these facts mean in our case.

**Definition 6.2.9.** If $(R, \sigma)$ is a difference ring extending $(\mathcal{O}_K, \mathrm{id})$, then let

$$\begin{aligned} \Omega(R,\sigma) &:= \ker G(R) \xrightarrow{P(\sigma)\circ(\sigma-1)} G(R) \\ \Lambda(R,\sigma) &:= \ker G(R) \xrightarrow{P(\sigma)} G(R) \\ \Phi(R,\sigma) &:= \ker G(R) \xrightarrow{\sigma-1} G(R) \end{aligned}$$

**Lemma 6.2.10 (Hrushovski).** *If $(\mathcal{K}, \sigma)$ is a transformally closed field extending $(K, \mathrm{id})$ and $V$ is any subvariety of $G_\eta \times G_\eta$, then there are varieties $Y_1, \ldots, Y_n$ and $Z_1, \ldots, Z_n$ such that*

1. *$V \cap (\Lambda \times \Phi)(\mathcal{K}, \sigma) = (\bigcup_{i=1}^n Y_i \times Z_i) \cap (\Lambda \times \Phi)(\mathcal{K}, \sigma)$*
2. *each $Y_i$ is a translate of a group subvariety of $G$.*

■ By Lemma 3.60 of [21] the group $\Lambda(\mathcal{K}, \sigma)$ is locally modular and stably embedded and hence every definable subset is a Boolean combination of cosets of definable subgroups. By Theorem 5.5 of [12] $\Lambda(\mathcal{K}, \sigma)$ is orthogonal to the fixed field. This implies by Lemma 3.24 of [21] every definable subset of the product $(\Lambda \times \Phi)(\mathcal{K}, \sigma)$ is a Boolean combination of products of definable subsets of $\Lambda(\mathcal{K}, \sigma)$ with definable subsets of $\Phi(\mathcal{K}, \sigma)$. In the statement above $\bigcup_{i=1}^n Y_i \times Z_i$ is the Zariski closure of $V \cap (\Lambda \times \Phi)(\mathcal{K}, \sigma)$. ■

**Lemma 6.2.11.** *In Lemma 6.2.10, if $K'$ is a subfield of $\mathcal{K}$ closed under $\sigma$ and $\sigma^{-1}$ and $V$ is defined over $K'$, then each of the $Y_i$'s and $Z_i$'s are defined over ${K'}^{alg}$.*

■ If $\Sigma \subseteq \mathcal{K}$ is any subset, then the algebraic closure of $\Sigma$ in the model theoretic sense (ie the set of elements of $\mathcal{K}$ which satisfy a formula with parameters from $\Sigma$ having only finitely many solutions) is equal to the algebraic closure in the sense of field theory of the field generated by $\{\sigma^n(x) : x \in \Sigma, n \in \mathbb{Z}\}$ (see [12] Proposition 1.7). If $\Sigma = K'$ is a field which is already closed under $\sigma$ and $\sigma^{-1}$, then the model theoretic algebraic closure is just ${K'}^{alg}$.

Each $Y_i$ and $Z_i$ is algebraic over $K'$ model theoretically and therefore algebraically. ■

**Lemma 6.2.12.** *Let $V \subseteq G_\eta \times G_\eta$ be a subvariety defined over $L$. There is a finite set $\Xi$ of subvarieties of $V$ defined over $L$ such that*

1. *If $Y \in \Xi$, then each irreducible component of $Y$ is of the form $W \times Z$ where $W$ is a translate of a group subvariety of $G$.*
2. *If $(\mathcal{K}, \sigma)$ is a transformally closed field extending $(K, \mathrm{id})$ given with a fixed embedding of $L$, then for some $Y \in \Xi$ one has $V \cap (\Lambda \times \Phi)(\mathcal{K}, \sigma) = Y \cap (\Lambda \times \Phi)(\mathcal{K}, \sigma)$.*

■ This follows by the compactness theorem of first order logic from Lemmas 6.2.10 and 6.2.11.

That is, if there were no uniform choice for $\Xi$, then the following set of sentences would be consistent.

1. $(\mathcal{K}, \sigma)$ is a transformally closed field extending $(K, \mathrm{id})$ and is given with a field embedding of $L$.
2. $\{\bigwedge_{Y \in \Xi} X \cap (\Lambda \times \Phi)(\mathcal{K}, \sigma) \neq Y \cap (\Lambda \times \Phi)(\mathcal{K}, \sigma) : \Xi$ a finite set of varieties as in the statement of the Lemma $\}$

By the compactness theorem, all of these sentences can be realized simultaneously. This contradicts Lemmas 6.2.10 and 6.2.11. ■

*Remark* 6.2.13. If $V \subseteq G \times G$ is a closed subscheme over $\mathrm{Spec}\mathcal{O}_L$, then there is a finite set $\Xi$ of closed subschemes of $V$ having properties analogous to $\Xi$ of Lemma 6.2.12. One sees this by observing that the varieties in $\Xi$ come from components of projections back to $G_\eta \times G_\eta$ of intersections in jet spaces. We will not need to use this fact.

**Lemma 6.2.14.** *If $\sigma \in \mathrm{Gal}(L/K)$ is good, then $\Lambda(L, \sigma) + \Phi(L, \sigma) \supseteq ⅂$*

■ If $(\mathcal{K}, \sigma)$ is any transformally closed field extending $(L, \sigma)$, then $|\Omega(\mathcal{K}, \sigma)/[\Lambda(\mathcal{K}, \sigma) + \Phi(\mathcal{K}, \sigma)]|$ is finite (see the proof of Theorem 5.4 of [11]).

**Claim 6.2.15.**

$$\Lambda(L, \sigma) + \Phi(L, \sigma) \quad = \quad (\Lambda(\mathcal{K}, \sigma) + \Phi(\mathcal{K}, \sigma)) \cap \Omega(L, \sigma)$$

✠ Since $\Lambda = \ker P(\sigma)$ and $\Phi = \ker(\sigma - 1)$, $\Lambda(\mathcal{K}, \sigma) \cap \Phi(\mathcal{K}, \sigma) \subseteq G[P(1)](\mathrm{Fix}(\sigma))$ which is finite because $P(1) \neq 0$. Thus if $x = a + b$ with $(a, b) \in \Lambda(\mathcal{K}, \sigma) \times \Phi(\mathcal{K}, \sigma))$, then $a$ and $b$ are model-theoretically algebraic over $x$. By the criterion for algebraicity in transformally closed fields, if $x \in G(L)$, then $a, b \in G(L)$. ✠

Thus the map $\Omega(L, \sigma)/[\Lambda(L, \sigma) + \Phi(L, \sigma)] \longrightarrow \Omega(\mathcal{K}, \sigma)/[\Lambda(\mathcal{K}, \sigma) + \Phi(\mathcal{K}, \sigma)]$ is injective. So $m := |\Omega(L, \sigma)/[\Lambda(L, \sigma) + \Phi(L, \sigma)]| \leq |\Omega(\mathcal{K}, \sigma)/\Lambda(\mathcal{K}, \sigma)| < \infty$.

As $\sigma$ is good, $\Omega(L, \sigma) \supseteq ⅂$. So $\Lambda(L, \sigma) + \Phi(L, \sigma) \supseteq [m]\Omega(L, \sigma) \supseteq [m]⅂ = ⅂$. ■

### 6.2.3 Proof of Theorems

Before giving a proof of Theorem 6.1.8 we give an easy proof of Proposition 6.1.9. This proposition and proof, if not exactly stated this way, appears in a letter from Hrushovski to Voloch [28].

■ Let $\sigma$ be as in the statement of the Proposition. By Lemma 6.2.10, there is subvariety $Y$ of $X_\eta$ defined over $L$ all of whose irreducible component are translates of group varieties having the property that $(X_\eta \cap \Lambda)(L, \sigma) = (Y \cap \Lambda)(L, \sigma)$. Since $\daleth \subseteq \Lambda(\mathcal{O}_L, \sigma)$, by Lemma 4.2.3 it suffices to bound the distance to $\overline{Y}$. By Lemma 4.1.7, it suffices to bound the distance to each irreducible component of $\overline{Y}$. Since each component is generically a coset, Lemma 6.2.8 says that the desired bound exists. ■

We now give the proof of Theorem 6.1.8

■ By Lemma 6.2.5, we may assume that $X_\eta$ has a trivial stabilizer. By Lemma 4.1.7, we may assume that $X$ is irreducible. If $X_\eta$ is empty, then by Lemma 4.2.3 the distance from any point in $G(\mathcal{O}_L)$ to $X$ is uniformly bounded. Thus, we may assume that $X_\eta$ is scheme theoretically dense in $X$.

We prove the theorem by induction on the dimension of $X_\eta$. If $\dim X_\eta = 0$, then we are in the situation of Lemma 6.2.7.

In general, let $\tilde{X} := +^*X := \{(x, y) \in G \times G : x + y \in X\}$. By Lemma 4.1.8, $d((x, y), \tilde{X}) = d(x + y, X)$. Let $\Xi$ be the set of subvarieties of $\tilde{X}_\eta$ produced by Lemma 6.2.12 with $V = \tilde{X}_\eta$. Let $\Pi$ and $\Upsilon$ be the sets of irreducible varieties with the property that the irreducible components of elements of $\Xi$ are of the form $W \times Z$ with $W \in \Pi$ and $Z \in \Upsilon$.

Each $W \in \Pi$ is a translate of a group so that by Lemma 6.2.8 there is some constant $\gamma_W \in \Gamma$ such that for any torsion point $P \in G(\mathcal{O}_L)_{tor}$ either $\iota(P) \in W$ or $d(P, \overline{W}) \leq \gamma_W$.

By induction, if $Z \in \Upsilon$ and $\dim Z < \dim X_\eta$, then there is some constant $\gamma_Z \in \Gamma$ so that any point $P \in \daleth$ not on $Z$ in the generic fibre satisfies $d(P, \overline{Z}) \leq \gamma_Z$. Let $C_1$ and $C_2$ be the constants produced by Lemma 4.2.3 so that for any transformally closed field $(\mathcal{K}, \sigma)$ extending $(K, \mathrm{id})$ for which we have $\tilde{X}_\eta \cap (\Lambda \times \Phi)(\mathcal{K}, \sigma) = Y \cap (\Lambda \times \Phi)(\mathcal{K}, \sigma)$ for some $Y \in \Xi$, then for $x \in (\Lambda \times \Phi)(L, \sigma)$ the inequality $d(x, \tilde{X}) \leq C_1 d(x, \overline{Y}) + C_2$.

Let $\gamma' := \min\{C_1\gamma_Y + C_2 : Y \in \Pi \text{ or } (Y \in \Upsilon \text{ and } \dim Y < \dim X_\eta)\}$.

Let $\sigma$ be a good automorphism.

By Lemma 6.2.14, if $\sigma$ is good, then $(\Lambda(L, \sigma) \cap \daleth) + (\Psi(L, \sigma) \cap \daleth) \supseteq \daleth$. So if $P \in \daleth$ and $d(P, X)$ is large, there are $a \in \Lambda(L, \sigma) \cap \daleth$ and $b \in \Psi(L, \sigma) \cap \daleth$ with $a + b = P$ and $d((a, b), \tilde{X})$ large.

Let $Y \in \Xi$ such that $\tilde{X}_\eta \cap (\Lambda \times \Phi)(L, \sigma) = Y \cap (\Lambda \times \Phi)(L, \sigma)$. Write $Y = \bigcup_{i=1}^N W_i \times Z_i$ with $W_i \in \Pi$ and $Z_i \in \Upsilon$.

By Lemma 4.1.7, $d((y, z), \bigcup_{i=1}^N \overline{W_i \times Z_i}) = \max d((y, z), \overline{W_i \times Z_i})$.

By Lemma 4.1.5, $d((y, z), \overline{W_i \times Z_i}) = \min\{d(y, \overline{W_i}), d(z, \overline{Z_i})\}$.

Thus, if $P \in \daleth$ and $d(P, X) > \gamma'$, it must be that for any good $\sigma$ one can write $P = a + b$ with $(a, b) \in (\Lambda \times \Phi)(L, \sigma)$ and $d(a, \overline{W}) > \frac{\gamma' - C_2}{C_1}$ and $d(b, \overline{Z}) > \frac{\gamma' - C_2}{C_1}$ for some $W \in \Pi$ and $Z \in \Upsilon$.

By the definition of $\gamma'$, the only way we can have $d(a, \overline{W}) > \frac{\gamma' - C_2}{C_1}$ for $a$ a torsion point is to have $\iota(a) \in W$. If $\dim Z < \dim X_\eta$, then having $d(b, \overline{Z}) > \frac{\gamma' - C_2}{C_1}$ would also violate the definition of $\gamma'$ unless $\iota(b) \in Z$. Since $W \times Z \subseteq \tilde{X}_\eta$, if $\iota((a, b)) \in W \times Z$, then $P = a + b \in X_\eta$.

So we must have $\iota(a) \in W$ and $\dim X_\eta = \dim Z$. Since $W \times Z \subseteq \tilde{X}_\eta$, we have $W + Z \subseteq X$. Since $X$ is irreducible, has trivial stabilizer, and has dense generic fibre, it must be that $\overline{W} = \{a\}$ and $\overline{Z} = X - a$. Let $m$ be the least common multiple of the orders of $y$ such that $\{y\} \in \Pi$.

Then we have that $[m]P = [m](a + b) = [m]b \in \Phi(L, \sigma)$. That is, for any choice of a good $\sigma$, $[m]P$ is fixed by $\sigma$. Since there are only finitely many torsion points fixed by every good $\sigma$, there are only finitely many choices for $[m]P$ and hence for $P$.

Let $\gamma := \max\{d(Q, X) : \iota(Q) \notin X_\eta, [m]Q \in \urcorner \cap G(K)\} \cup \{\gamma'\}$.
This completes the induction. ■

## 6.3 Prime to $p$ Torsion

In this section we prove that Theorem 6.1.7 follows from Theorem 6.1.8.

Let $\mathbb{F}_q$ be the residue field of $K$. Let $\mathrm{Frob}_q \in \mathrm{Gal}(\mathbb{F}_q^{alg}/\mathbb{F}_q)$ be the $q$-power Frobenius $x \mapsto x^q$. Let $\mathrm{Frob}_q$ also denote the endomorphism of $G_0$ induced by $\mathrm{Frob}_q$. Let $P(X) \in \mathbb{Z}[X]$ be the minimal polynomial over $\mathbb{Z}$ of $\mathrm{Frob}_q$ considered as an element of $\mathrm{End}(G_0)$.

**Lemma 6.3.1.** *$P(X)$ has no cyclotomic factors.*

■ By the Riemann hypothesis for semi-abelian varieties over finite fields, all the roots of $P$ in $\mathbb{C}$ have size $q$ or $\sqrt{q}$.

For the reader unhappy with the Weil conjectures, we sketch a proof given by Pillay in [40].

Since $P(X)$ is the minimal polynomial of $\mathrm{Frob}_q$ over $\mathbb{Z}$, $P$ is irreducible over $\mathbb{Q}$. So if $P$ has a cyclotomic factor, $P$ itself is a cyclotomic polynomial and $\mathrm{Frob}_q$ acts as a root of unity on $\mathcal{G}_0$. Thus, for some $n \in \mathbb{Z}_+$, $\mathrm{Frob}_q^n = \mathrm{id}_{G_0}$. Since $\mathrm{Frob}_q^n$ topologically generates $\mathrm{Gal}(\mathbb{F}_q^{alg}/\mathbb{F}_{q^n})$, this would imply that $G_0(\mathbb{F}_q^{alg}) = G_0(\mathbb{F}_{q^n})$. This is absurd as $G_0(\mathbb{F}_{q^n})$ is finite but $\dim G_0 = \dim G > 0$ so $G_0(\mathbb{F}_q^{alg})$ is infinite. ■

**Definition 6.3.2.** A continuous automorphism $\sigma \in \mathrm{Gal}(\mathbb{C}_p/K)$ is called a Frobenius if for every $x \in \mathcal{O}_{\mathbb{C}_p}$ one has $\sigma(x) = x^q \mod \mathfrak{m}_{\mathbb{C}_p}$.

**Lemma 6.3.3.**

1. *Frobenii exist.*
2. *If $\sigma \in \mathrm{Gal}(\mathbb{C}_p/K)$ is continuous, then $\sigma$ is a Frobenius iff $\sigma$ satisfies the conditions of a Frobenius on the strict henselization of $\mathcal{O}_K$ (ie the ring of integers of the maximal algebraic unramified extension of $K$).*
3. *If $\sigma$ is a Frobenius, then $\sigma|_{K^{unr}}$ topologically generates $\mathrm{Gal}(K^{unr}/K)$.*

■

**Claim 6.3.4.** $K^{unr} = K(\{\zeta : \zeta^n = 1$ *for some* $n \in \mathbb{Z}_+\})$.

✠ If $L$ and $M$ are two unramified extensions of $K$, then so is the compositum $LM$. Thus, there is at most one unramified extension of $K$ having residue field $\mathbb{F}_{q^n}$ for each positive integer $n$. By Hensel's Lemma, if $p \nmid n$, then the extension $K(\zeta_n)/K$ is unramified for any $n$-th root of unity $\zeta$. As every element of $\mathbb{F}_q^{alg}$ is a root of unity of order prime to $p$, the claim is proved. ✠

The map $x \mapsto x^q$ on $\mu_n$ generates $\mathrm{Gal}(K(\mu_n)/K)$ for each $n$ prime to $p$. So Frobenii exist and they topologically generate $\mathrm{Gal}(K^{unr}/K)$.

Let $x \in \mathcal{O}_{\mathbb{C}_p}$ be arbitrary. Let $x_0 \in \{\zeta \in \mathbb{C}_p^\times : \zeta^n = 1 \text{ for some } n \in \mathbb{Z}_+, (n,p) = 1\} \cup \{0\}$ such that $x = x_0 \mod \mathfrak{m}_{\mathbb{C}_p}$. If $\sigma$ is any automorphism of $\mathbb{C}_p$, then $\sigma(x) = \sigma(x_0) + \sigma(x - x_0)$. If $\sigma$ is continuous, then $\sigma(x - x_0) \in \mathfrak{m}_{\mathbb{C}_p}$ (Elements of $\mathrm{Gal}(\mathbb{Q}_p^{alg}/\mathbb{Q}_p)$ preserve the valuation. Since $\mathbb{Q}_p^{alg}$ is dense in $\mathbb{C}_p$, the same must be true of continuous automorphisms of $\mathbb{C}_p$.) Thus, $\sigma(x) = x^q \mod \mathfrak{m}_{\mathbb{C}_p} \iff \sigma(x_0) = x_0^q \mod \mathfrak{m}_{\mathbb{C}_p}$. ■

**Lemma 6.3.5.** *The common fixed field of the Frobenii is $K$.*

■ By Lemma 6.3.3, the Frobenii topologically generate $\mathrm{Gal}(K^{unr}/K)$. So if $x \in K^{unr}$ and some Frobenius fixes $x$, then $x \in K$. Suppose now that $x \notin K^{unr}$. Let $\sigma$ be a Frobenius. If $\sigma(x) \neq x$, then we're done. Otherwise, let $\tau \in \mathrm{Gal}(K^{alg}/K^{unr})$ such that $\tau(x) \neq x$. Then by Lemma 6.3.3 $\tau \circ \sigma$ is a Frobenius. Visibly, $\tau \circ \sigma(x) \neq x$. ■

**Lemma 6.3.6.** *If $H$ is any semi-abelian variety over $K$, then $H(K)_{tor}$ is finite.*

■ Replacing $K$ with a finite extension can only increase the number of torsion points, so we may assume (via semi-stable reduction) that $H$ is the generic fibre of a semi-abelian scheme $\mathfrak{H}$ over $\mathcal{O}_K$. By Lemmas 6.2.7 and 4.1.9, there is some $\gamma$ such that for any two distinct torsion points $P$ and $Q$ one has $d(P, Q) > \gamma$. Thus, on the torsion points the reduction map $\pi_\gamma : \mathfrak{H}(\mathcal{O}_K) \to \mathfrak{H}_\gamma(\mathcal{O}_K/I_\gamma)$ is injective. Since $K$ is finite over $\mathbb{Q}_p$, the ring $\mathcal{O}_K/I_\gamma$ and hence the set $\mathfrak{H}_\gamma(\mathcal{O}_K/I_\gamma)$ is finite. ■

**Lemma 6.3.7 (Grothendieck).** *Let $n := \mathrm{rk}_{\mathbb{Z}_\ell} T_\ell G_\eta - \mathrm{rk}_{\mathbb{Z}_\ell} T_\ell G_0$ for any prime $\ell \neq p$. Let $Q \in \mathbb{Z}[X]$ be the minimal polynomial of $\mathrm{Frob}_q^{n!}$ on $G_0$. If $\sigma \in \mathrm{Gal}(\mathbb{C}_p/K)$ is a Frobenius, then $Q(\sigma^{n!}) \circ (\sigma^{n!} - 1)$ vanishes on $G(\mathcal{O}_{\mathbb{C}_p})_{p'-tor}$.*

■ By Corollary 4.4 of [18] for any prime $\ell \neq p$, there is a Galois invariant submodule $U$ of the $\ell$-Tate module of $G$ which is isomorphic to $T_\ell G_0$ as a Galois module (if one identifies the Galois group of the residue field with the group $\mathrm{Gal}(K^{unr}/K)$). Moreover, the Galois group has as eigenvalues $n!$-th roots of unity on $T_\ell G/U$. The result should now be clear. ■

At this point we replace $q$ with $q^{n!}$ and $P$ with $Q$.
These lemmas give a proof of Theorem 6.1.7.

■ Lemmas 6.3.3, 6.3.5, 6.3.6, and 6.3.7 ensure that the hypotheses of Theorem 6.1.8 hold. ■

## 6.4 Other Cases

The proof of Theorem 6.1.8 requires very little information about the arithmetic of $G$. The number theoretic cost of applying Theorem 6.1.8 comes from verifying the hypotheses. In this section we will discuss a few other cases of Conjecture 6.1.2 that may be amenable to our method.

The main challenge in applying Theorem 6.1.8 to the case of $\daleth = G[p^\infty](\mathcal{O}_{\mathbb{C}_p})$ is finding good automorphisms. If one restricts for the moment to the considerations of points in the formal group, then one might try to find the equations as a characteristic polynomial of some element of the inertia group acting on the Tate module of the formal group. Theorems of Serre, Tate, and Sen describe the image of this Galois representation (see [52], [50], and [48]) so one might hope to deduce from their results the existence of good automorphisms. One would then like to proceed by an analysis of the monodromy of the Galois representation on the full Tate module at $p$ to argue that $P(\sigma) \circ (\sigma - 1)$ vanishes on $G[p^\infty]$ for some $P$ with no cyclotomic factors and good $\sigma$.

This approach may work in general, but it seems that the current state of knowledge about the Galois representation on the Tate module of the formal group is insufficient. There are cases, however, where this will work. For instance, if the formal group has rank at most one, then it is quite easy to choose a Frobenius which acts on the formal group by multiplication by a rational integer distinct from $\pm 1$. In general, if the maximal abelian quotient of the reduction of $G$ is ordinary, then one can find elements of the inertia group which act on the formal group with characteristic polynomials over $\mathbb{Z}$ having no cyclotomic factors. However, once the $p$-rank of the reduction surpasses one, it is not clear that the automorphism may be chosen to be a Frobenius.

This brings us to another case which is closer to Conjecture 6.1.2. To handle the case of $\daleth = G(\mathcal{O}_{\mathbb{C}_p})_{tor}$ one needs to treat the $p$-power torsion together with the prime-to-$p$ torsion. Ideally, we would work with Frobenii that also behaved well on the formal group. As mentioned above, even in the cases where one can find automorphisms behaving well on the formal group, it is not so easy to find Frobenii with this property. One could work with the theory of transformally closed fields with respect to several automorphisms instead. Our methods do work in this case, though the translation is not immediate.

One might also apply these methods to the case of the *abc* theorem where $G$ and $X$ are defined over $k$. At present, we have to specialize to the case of $\operatorname{char} k = 0$. To apply Theorem 6.1.8, we need to find $\sigma$ which preserve $k$ but act on $\Gamma$ by an integral characteristic polynomial having no cyclotomic factors. For some $\Gamma$, it is easy to find such $\sigma$. For instance, if $\Gamma$ is generated by $n$ generic elements of $G$, then one can take $P(X) = X - 2$. Restrict now to the case that $G$ is an abelian variety. If we could find a finitely generated field $L$ containing $k(\Gamma)$ and having an endomorphism $\sigma : L \to L$ with $\operatorname{Fix}(\sigma) = k$, then the characteristic polynomial of $\sigma$ acting on $G(L)/G(k)$ would have integer co-efficients and no cyclotomic factors. If $V$ is a variety with $k(V) = k(\Gamma)$ then to get $L$ we need to find another variety $W$, a rational dominant map $W \dashrightarrow V$ and an endomorphism of $W$ with a Zariski dense orbit.

# Bibliography


[1] J. Ax and S. Kochen, Diophantine problems over local fields II: a complete set of axioms for $p$-adic number theory, *Amer. J. Math.* **87** (1965), pp 631-48.

[2] S. Bosch, W. Lütkebohmert, and M. Raynaud, Néron Models, Springer-Verlag, Berlin, 1990.

[3] E. Bouscaren, Model theoretic versions of Weil's theorem on pregroups, in Model Theory of Groups (A. Pillay and A. Nesin, editors), University of Notre Dame Press, South Bend, IN, 1989.

[4] A. Buium, An approximation property for Teichmüller points, *Math. Res. Lett.* **3** (1996), no. 4, 453 – 457.

[5] A. Buium, The *abc* theorem for abelian varieties, *Internat. Math. Res. Notices* (1994), no 5, 219 ff.

[6] A. Buium, Geometry of $p$-jets, *Duke Math. J.* **82** (1996), no. 2, 349 - 367.

[7] A. Buium, The geometry of differential polynomial functions I: Algebraic groups, *Am. J. Math.* **115** (1993), no. 6, 1385 – 1444.

[8] A. Buium and J. F. Voloch, Lang's conjecture in characteristic $p$: an explicit bound, *Compositio Math.* **103** (1996), no. 1, 1 – 6.

[9] A. Buium and J. F. Voloch, Reduction of the Manin map modulo $p$, *J. Reine Angew. Math.* **460** (1995), 117-126.

[10] Chang and J. Keisler, *Model Theory*, North-Holland, 3rd ed, 1990.

[11] Z. Chatzidakis, Groups definable in ACFA, *Proc. NATO-ASI Conf. Model Theory of Fields* (Fields Institute, Toronto), August 1996, (to appear). `http://boole.logique.jussieu.fr/www.zoe/`

[12] Z. Chatzidakis and E. Hrushovski, The model theory of difference fields, preprint Université Paris VII, July 1996. `http://boole.logique.jussieu.fr/www.zoe/`

[13] C. Chevalley, Fondements de la géométrie algébique, *Sec. Math.*, Paris, 1958.

[14] C. Chevalley, La théorie des groupes algébriques, Proc. Int. Cong. Math. 1958, Cambridge University Press, Cambridge, pp 53 – 68.

[15] F. Delon Ideaux et types sur les corps separablement clos, *Mém. Soc. Math. France (N. S.)*, No. 33 (1988), 76 pp.

[16] G. Faltings and C.-L. Chai, Degeneration of Abelian Varieties, Ergebenisse der Mathematik und ihrer Grenzgebiete **22**, Springer-Verlag, Berlin Heidelberg, 1990.

[17] K. GÖDEL The Consistency of the Continuum hypothesis, Annals of Mathematics Studies, no. 3. Princeton University Press, Princeton, N. J., 1940. 60 pp.

[18] A. GROTHENDIECK, Modeles de Néron et monodromie, SGA $7_I$, exposeé IX.

[19] R. HARTSHORNE Algebraic Geometry, GTM **52**, Springer-Verlag, New York-Heidelberg, 1977.

[20] W. HODGES, Model Theory, Encyclopedia of Mathematics and its Applications, **42**, Cambridge University Press, Cambridge, 1993.

[21] E. HRUSHOVSKI, The model theory of difference fields and the Manin-Mumford conjecture, preprint, 1996.

[22] E. HRUSHOVSKI and A. PILLAY, Weakly normal groups, Logic Colloquium '85, North-Holland, Amsterdam, 1987.

[23] E. HRUSHOVSKI and Ž. SOKOLOVIĆ, Minimal subsets of differentially closed fields, *Trans. Amer. Math. Soc.* (to appear).

[24] E. HRUSHOVSKI and B. ZIL'BER, Zariski geometries, *J. Amer. Math. Soc.* **9** (1996), 1 – 56.

[25] E. HRUSHOVSKI, The Mordell-Lang conjecture for function fields, *J. Amer. Math. Soc.* **9** (1996), no 3, 667-90.

[26] E. HRUSHOVSKI, Contributions to stable model theory, Ph. D. thesis, UC Berkeley, 1986.

[27] E. HRUSHOVSKI, The first order theory of the Frobenius, preprint, 1996.

[28] E. HRUSHOVSKI, e-mail to J. F. Voloch, November 1996.

[29] B. KIM and A. PILLAY, The independence theorem for simple theories, preprint, 1996.

[30] A. MACINTYRE, On definable subsets of $p$-adic fields, *JSL* **41** (1976), 605-610.

[31] D. MARKER, Model theory of differential fields, in Model Theory of Groups, LNL **5**, Springer-Verlag, New York.

[32] H. MATSUMURA Commutative Ring Theory, Cambridge Studies in Advanced Mathematics **8**, Cambridge University Press, Cambridge, 1986.

[33] A. MATTUCK, Abelian varieties over $p$-adic ground fields, *Ann. of Math. (2)* **62**, (1955), 92 – 119.

[34] M. MESSMER, Groups and fields interpretable in separably closed fields, *Trans. Amer. Math. Soc.* **344** (1994), no 1, 361-377.

[35] M. MESSMER Some model theory of separably closed fields, in Model Theory of Fields, LNL **5**, Springer, Berlin, 1996.

[36] M. MESSMER and C. WOOD, Separably closed fields with higher derivations, *JSL* **60** (1995), no 3, 898-910.

[37] C. MICHAUX, Ph. D. thesis, Université Paris 7 (1991).

[38] NEUMANN, On ordered division rings, *Trans. Amer. Math. Soc.* **66** (1949), pp 202 - 252.

[39] A. PILLAY, Differential algebraic groups and the number of countable differentially closed fields, in Model Theory of Fields, LNL **5**, Springer, Berlin, 1996.

[40] A. Pillay, ACFA and the Manin-Mumford conjecture, *Proc. NATO-ASI Conf. Model Theory of Fields* (Fields Institute, Toronto), August 1996, (to appear).

[41] A. Pillay and Ž. Sokolović, Superstable differential fields, *JSL* **57** (1992), no. 1, 97 – 108.

[42] B. Poizat, Une theorie de Galois imaginaire. *JSL* **48** (1983), no. 4, 1151 – 1170.

[43] M. Raynaud, Sous-varietes d'une variete abelienne et points de torsion, Arithmetic and geometry, Vol. I, Birkhäuser Boston, Boston, MA, 1983, 327–352.

[44] M. Raynaud, Anneaux locaux henséliens, LNM **169**, Springer-Verlag, Berlin-New York 1970.

[45] G. Sacks, Saturated Model Theory, W. A. Benjamin, Reading, MA, 1972.

[46] Schilling, The Theory of Valuations, AMS Mathematical Survey # 4, 1950.

[47] P. Scowcroft On the elimination of imaginaries from certain valued fields, *Ann. Pure Appl. Logic* **61** (1993), no. 3, 241 – 276.

[48] S. Sen Lie algebras of Galois groups arising from Hodge-Tate modules, *Ann. of Math. (2)* **97** (1973), 160 – 170.

[49] J. P. Serre Lie algebras and Lie groups, LNM **1500**, Springer-Verlag, Berlin-New York, 1992.

[50] J. P. Serre Groupes algébriques associés aux modules de Hodge-Tate, *Journées de Géométrie Algébrique de Rennes* (Rennes, 1978), Vol III, pp. 155 – 188, Astérisque 65, Soc. Math. France, Paris, 1979.

[51] Ž. Sokolović, Model theory of differential fields, Ph. D. thesis, Notre Dame, July, 1992.

[52] J. Tate $p$-divisible groups, *1967 Proc. Conf. Local Fields* (Driebergen, 1966), pp. 158 – 183, Springer, Berlin.

[53] J. Tate and J. F. Voloch, Linear forms in $p$-adic roots of unity, *Internat. Math. Res. Notices* 1996, no. 12, 589 – 601.

[54] L. van den Dries, Weil's group chunk theorem: a topological setting, *Illinois J. Math.* **34** (1990), no. 1, 127 – 139.

[55] J. F. Voloch, Diophantine approximation on abelian varieties in characteristic $p$, *Amer. J. Math.* **117** (1995), no. 4, 321-5.

[56] A. Weil, Variétés abéliens et courbes algébriques, Hermann, Paris, 1948.

[57] V. Weispfenning, Existential equivalence of ordered abelian groups with parameters, *Arch. Math. Logic* **29** (1990), no. 4, 237 – 248.

[58] V. Weispfenning, Quantifier eliminable ordered abelian groups, Algebra and order (Luminy-Marseille, 1984), 113 – 126, *R & E Res. Exp. Math.*, 14, Heldermann, Berlin, 1986.

[59] C. Wood, Notes on the stability of separably closed fields, *JSL* **44** (1979), 412 – 416.