Stability and its uses

Ehud Hrushovski
MIT
HEBREW UNIVERSITY IN JERUSALEM

The occasion for this paper is a proof of the Mordell-Lang conjecture for function fields, and associated results. Some of the results have been proved earlier by different methods, while others (notably in positive characteristic) are new. The main interest in both cases is the use of model-theoretic tools; indeed the proof requires no special ingenuity but follows naturally from model-theoretic considerations. This exposition attempts therefore to give an account of the proof and of the model theory that goes into it; I did not describe parts of model theory that are not relevant to the proof, but on the other hand made no effort to suppress model theoretic apparatus that is used in it.

§1 is devoted to the basic notions of model theory. The results mentioned there can be found in [4], though the point of view here is closer to that of Abraham Robinson. In §2 we describe parts of stability theory. For our purposes, it can be viewed as a foundations for certain kinds of geometry, analogous to algebraic geometry. In particular in §2.4-2.6 we consider a central notion, modularity, a dividing line between linear and non-linear geometries of this type. It is ultimately this dichotomy, applied to a certain structure lying inside an Abelian variety, that creates an equally sharp line between "arithmetic" Abelian varieties and others, and yields conclusions concerning non-isotrivial Abelian varieties, and torsion points of Abelian varieties over number fields.

In §2.6 we describe a classification of model-theoretic geometries having a dimension theory similar to that of algebraic geometry. §3.1 contains an application of this result to algebraic differential equations, in the form of a very rough classification of such equations into three classes (Theorem 3.3). One of the classes consists essentially of those equations whose solutions can be algebraically parameterized. Another involves special differential equations, discovered by Manin, inside Abelian varieties. It may seem surprising that a very basic and general model theoretic classification, applied to the case algebraic differential equations, should lead to a consideration of Abelian varieties at all.

By a quite similar scheme we obtain similar results about difference equations, as well as certain structures in characteristic $p > 0$.

In §3.2 we deduce the geometric Mordell-Lang conjecture, and mention some other corollaries involving torsion points. Here is an oversimplified description of that proof. It concerns an Abelian variety A, a certain group Γ of points of A, and a subvariety X. We will consider $(A, +, \Gamma, X)$ as an abstract geometry of the type described in §2. We will know that the axioms of these geometries are satisfied by interpreting the situation in a differential algebraic setting (following an idea of A. Buium's), or in another setting known to have the right abstract properties. The general theory of §2 will allow us to decompose the geometry into certain "minimal" or irreducible ones. The classification in §3.1 will show that each irreducible geometry is modular, or else equivalent to algebraic geometry over a certain field k. Modularity will imply that Γ meets X in a simple way. The other alternative restricts the situation considerably and in particular, following a brief investigation of the identity of the field k, will show that A is defined over a number field.

1   Model Theory

1.1   Introduction

Model theory often deals with mathematical structures familiar from algebra or geometry, but from a point of view of its own. I will try in this section to explain this point of view.

Formally, a *structure* in the sense of model theory is a set A (or perhaps a number of sets A_ν, called *sorts*), and a collection $\mathcal{D}$ of *relations*, or *definable sets*, $R \subset A^m$ (or $R \subset A_{\nu_1} \times \ldots \times A_{\nu_m}$). One makes the assumption that $\mathcal{D}$ contains the diagonals and is closed under the following operations: intersections, complements, and inverse and direct images under projections from

a product of m sorts to a partial product (such as $A^m \to A^{m-1}$.) Often $\mathcal{D}$ is generated, under these operations, by some "basic" relations $R_1, \ldots, R_k$; in this case one denotes the structure as $(A; R_1, \ldots, R_k)$, though keeping in mind the larger class of sets that can be constructed from these.

These are sometimes called *1st-order structures*. Here are some examples of structures related to number theory; (3.1-3.3) are used in the results reported on in this talk.

(1) $(\mathbb{N}; +, \cdot)$ or $(\mathbb{Z}; +, \cdot)$; $\mathcal{D}$ is generated by the graphs, in $\mathbb{Z}^3$, of addition and multiplication. This structure is the subject of Gödel's theorem (see e.g. [37]). Though one begins with the simple relations $\{+, \cdot\}$ and applies very basic geometric operations, Gödel showed that one obtains an almost arbitrary subset of $\mathbb{Z}$ within $\mathcal{D}$; the study of these sets belongs to recursion theory, or descriptive set theory, but not to geometry or number theory. Later J. Robinson showed the same is true of the field $(\mathbb{Q}; +, \cdot)$. Thus though one may ultimately be interested in $\mathbb{Z}$ or $\mathbb{Q}$, they are not directly amenable to model theoretic analysis of the geometric kind we will consider. One may note that in number theory too, one approaches these objects through a host of auxiliary structures.

(2.1) $(\mathbb{R}; +, \cdot)$. A few years after Gödel's work, Tarski showed that a quite different situation holds for the real field. The elements of $\mathcal{D}$ in this case are geometrically reasonable objects, the semi-algebraic subsets of $\mathbb{R}^n$. In particular the definable subsets of $\mathbb{R}$ itself are finite unions of intervals.

Tarski's paper was entitled "a decision procedure for elementary geometry" (and not: "for the real field"). Indeed suppose we extend this structure by adding a sort E for Euclidean 3-space, and a map $d : E^2 \to \mathbb{R}$ (distance). In this "language" we can evidently discuss triangles, circles, and general real algebraic varieties. The point is that, in the model-theoretic view of $\mathbb{R}$, this was already present. E can be interpreted as $\mathbb{R}^3$, and the relation $d(x, y) = z$ is already part of $\mathcal{D}$.

We will note below (paragraph on formulas) that if $X \subset \mathbb{R}^n$ is a definable set, the closure $cl(X)$ of X is also definable. Similarly, the derivative of a definable function is readily seen to be definable, as is distance from a point to a variety. The class of definable subsets of $\mathbb{R}^n$ thus forms a limited but useful universe; when a problem can be formulated within it, one may freely apply many geometric operations without leaving it. For some applications of this to asymptotic properties of algebraic functions, see [13], Appendix A2.

(2.2) The other (p-adic) completions of $\mathbb{Q}$, and their algebraic extensions,

enjoy similar properties. It is most natural to present these fields with three sorts: K for the field itself, k for the residue field, Γ for the value field. One has relations for the valuation ring, the valuation map, and the reduction map. In particular, every subvariety of K^n is a definable set, and so is every set of the from $\{x : v(f(x)) \geq m\}$, f a rational function. The set of m'th powers in K is also definable; it is the projection to K of the graph of $y = x^m$. A theorem of Macintyre, following work of Ax and Kochen, states that all definable sets are obtained from these (using finite unions and intersections, and pullbacks by rational functions.)

(2.3) $(\mathbb{C}; +, \cdot)$, or an algebraically closed field of another characteristic. Here the definable sets are the constructible sets; finite Boolean combinations of algebraic subvarieties of $\mathbb{C}^n$ (defined over $\mathbb{Q}$.) ("Tarski-Chevalley theorem").

The examples (2.1-2.3) considered above represent a different foundational outlook on very well known geometries. Work of Denef, Van den Dries, and others, in the past decade has extended this framework to include, for example, local analytic geometry. For each type of "geometry", one must prove the analog of the Tarski-Chevalley and Tarski-Seidenberg theorems. Introducing even one new function into the class of definable sets has a high cost; one must rewrite the foundations for each such addition. (A recent example is Wilkie's addition of the global real exponential function to the local semi-analytic universe.) Once such work is done, however, one has a geometric "paradise", in which many constructions can be carried out, with a guarantee of not encountering any pathological sets on the way.

We will now turn to examples of a different kind. Classically the geometries involved are viewed as infinite-dimensional.

(3.1) A *differential field* is a field (here, of of characteristic zero) equipped with a derivation D. (One can also consider several derivations.) Together they form a structure $(U; +, \cdot, D)$. We choose $U = \mathbb{U}$ to be the *universal domain* for differential fields of characteristic zero; this notion will be explained below. An affine differential-algebraic subvariety is a subset of $\mathbb{U}^n$ defined by the vanishing of finitely many algebraic differential equations. The algebraic theory of such varieties was developed by Ritt and by Kolchin; they are also called Kolchin - closed sets. Each Kolchin closed subset of $\mathbb{U}^n$ is definable. Conversely, by a theorem of A. Robinson's, every definable set is a finite Boolean combinations of Kolchin-closed subsets of $\mathbb{U}^n$ (defined over

$\mathbb{Q}$.)

(3.2) The analog in characteristic $p > 0$ is a separably closed field $(K, +, \cdot)$, with $[K : K^p] = p$. To bring out the analogy with (3.1), one could add a derivation of K, or a stack of Hasse derivations. These are however automatically linear over K^p or K^{p^n}, hence definable whether or not they are added explicitly. (More precisely, the family of derivations forms a definable K^p-vector space, definably dual to K; a particular derivation is definable once one picks out $a \in K \setminus K^p$.) Here we will not stress this analogy, and work directly with the field structure. The definable sets will be described in §3.1.

(3.3) An analog of a different kind is obtained if one takes an automorphism in place of a derivation (difference fields). There is again a universal domain, or rather a family of universal domains depending on the restriction of the automorphism to $\tilde{\mathbb{Q}}$. Here the family of definable sets includes "difference varieties", defined by polynomials in $X, X^\sigma, \ldots$ where σ is a formal symbol for an automorphism. But $\mathcal{D}$ consists of something more than difference varieties; the basic definable sets describe the behavior of the automorphism on etale covers of a given variety. (See §3.1).

The examples given so far were defined in terms of a fundamental structure (the complex or real field, or differential field.) We took pains to emphasize that a whole geometry (in this case, of algebraic varieties) is generated by it. Though the geometry involves high-dimensional objects, in some sense its richness emanates from the one-dimensional fundamental structure. This duality can be reversed; when presented with a geometry synthetically, one can attempt to find an associated fundamental structure that gives rise to it. (Historically this is of course the first direction.) The next examples will be directly defined as a kind of geometry. The problem of identifying an intrinsic fundamental structure or structures responsible for their nature and complexity will be a major theme of this paper.

(4.1) Let X be a compact complex manifold; let $\mathcal{D}$ be generated by the closed analytic subvarieties of X. $\mathcal{D}$ consists in this case of finite Boolean combinations of analytic subvarieties. The fact that $\mathcal{D}$ is closed under the definable operations is due to Zilber; it relies essentially on Remmert's proper mapping theorem.

(4.2) Let A_F be the set of solutions (in the universal domain) to an

algebraic differential equation $F(x, Dx, \ldots, D^n x) = 0$. If $F_1, \ldots, F_m$ are such equations, $A_i = A_{F_i}$, call a subset $W \subset A_1 \times \ldots \times A_m$ *Kolchin - closed* if it is also given as the solution set of an algebraic differential equation, in variables $X_1, \ldots, X_m$. In particular, this defines the (Kolchin)-closed subsets of $(A_F)^m$. Let $\mathcal{D}$ be the class of finite unions of differences of two Kolchin closed sets. It can be shown, using (3.1), that $(A_F, \mathcal{D})$ is a first-order structure.

If F is linear, the graph of addition on $A_F{}^3$ is a Kolchin closed set; there are many others, though they may not be immediately evident. In the non-linear case, it is less clear what $\mathcal{D}$ is, and at all events one does not expect a group or ring structure. There is no apparent intrinsic organizing principle for the class of definable sets (despite the external "upper bound" one has, in the class of all definable sets of the universal domain of differential fields). We will find such an organizing principle, arising from ideas in abstract model theory. This will in return yield information on the algebraic differential equation with which we started.

It is important not only what relations are present, but just as much so, which ones are not. For instance, compare the ring $(\mathbb{Z}, +, \cdot)$ to the group $(\mathbb{Z}, +, 1)$ in which a generator has been distinguished. Model-theoretically, they are entirely different structures. The latter has a very simple geometry of definable sets; every definable set or relation is a Boolean combination of cosets of subgroups. In particular, multiplication cannot be constructed by a finite sequence of operations of the type allowed, though it can be constructed set-theoretically since $(\mathbb{Z}, +, \cdot)$ has the same automorphism group as $(\mathbb{Z}, +, 1)$. We will see later how to establish a kind of Galois theory for this kind of question, though initially no automorphisms are present.

When seeing a copy of $\mathbb{Z}$, occurring for instance as the value group of $\mathbb{Q}_p$, the model theorist immediately asks: what kind of $\mathbb{Z}$ is it? What is the structure? (In the case of the value group of $\mathbb{Q}_p$, the answer is: precisely $(\mathbb{Z}, +, <)$; no non-obvious relations are induced on elements of $\mathbb{Z}$ as p-adic values. This is a consequence of Macintyre's theorem or Ax-Kochen-Ershov, see below.)

Formulas and sentences. Assume a structure is given in terms of some "basic" generators of the class of definable sets; for instance, the graphs of $+$ and $\cdot$ on a field. The definable sets are then formed by a sequence of geometric operations: intersection, projection, etc. A *formula* (in n variables) is a description of a definable subset (of an n-fold product, or Cartesian power),

given via the procedure from obtaining it from some "basic" definable sets by projections, intersections, etc. One thinks not so much of the definable set, but of the conditions for membership in that set. Thus the intersection of two sets corresponds to the conjunction of two formulas, denoted by &. The projection of a definable set from $U \times V$ to U corresponds to the existential quantifier: for $D \subset (U \times V)$, $u \in pr_U D$ iff $(\exists v)((u, v) \in D)$. Similarly the complement corresponds to the negation symbol $\neg$.

Formally, formulas are formed by starting with a symbol for each basic definable set, including the symbol = for the diagonal, as well as the symbols $\exists, \&, \neg$, and forming expressions according to certain (evident) syntactic rules. One can also use the universal quantifier $\forall$; $(\forall x)\phi(x)$ can be defined as $\neg \exists x \neg \phi(x)$; similarly "or" , "implies", etc. Any statement normally encountered in mathematics, concerning the objects at hand, can in fact by formulated with ease. However, there is a lot of weight put on the proviso "objects at hand"; one is forced to be quite precise about just what is part of the structure; in particular the collection of sequences, or even of finite subsets, of a structure A cannot directly be referred to, unless it is either explicitly put in as part of the structure or shown to be interpretable in another way.

Formulas can also be used to assert properties of the structure, rather than point to a particular definable set. By definition, a *sentence* is a formula with no free variables; it represents an theorem (or false assertion) about the geometry of the structure.

Examples of formulas.

1. in $(\mathbb{R}, +, \cdot)$, the inequality relation $\leq$- the subset of $\mathbb{R}^2$ lying above the diagonal- can be described as the projection of the affine variety $y = x + z^2$ in $\mathbb{R}^3$:

$$(\exists z)(y = x + z^2)$$

2. If $D \subset \mathbb{R}^n$ is a definable set, defined by a formula $\phi(x) = \phi(x_1, \dots, x_n)$, the closure of D is also definable:

$(y_1, \dots, y_n) \in$

$$cl(D) \leftrightarrow (\forall \epsilon > 0)(\exists x_1, \dots, x_n) \left(\phi(x_1, \dots, x_n) \& \sum_{i=1}^{n} (x_i - y_i)^2 < \epsilon \right)$$

3. For any field F, the following sentence is either true or false (absolutely, without reference to particular parameters):

$$(\forall a)\dots(\forall f)(\exists x_0)(\exists x_1)(\exists x_2)(\neg(x_0 = x_1 = x_2 = 0)$$
$$\&\ (ax_0^2 + bx_0x_1 + cx_0x_2 + dx_1^2 + ex_1x_2 + fx_2^2 = 0))$$

It asserts that every degree 2 hypersurface in $\mathbb{P}^2$ has a solution; similarly, one can write down a sentence $\sigma_{d,n}$ stating that every degree d hypersurface in $\mathbb{P}^n$ has a solution with coordinates from the given field. The field is algebraically closed iff every such sentence holds true in it. Similarly, with infinitely many sentences, one can express the fact that a valued field is Henselian.

The linguistic interpretation of the basic operations usefully complements the geometric one. It is sometimes harder to comprehend directly the sequence of geometric operations - pullbacks, intersections, complements and projections - leading from a set to its closure, or from the graph of a function to that of its derivatives, than to read a formula such as in (2).

There are close connections between the class of sentences true in a fundamental structure, and the nature of the definable sets in the associated "geometries". Rarely does one determine one without saying a great deal about the other. For instance, it is not very different to prove the Tarski-Seidenberg theorem mentioned above, and to prove that the sentences true in $\mathbb{R}$ are those that follow from the axioms of real closed fields.

Model-theoretic categories. We can make the above discussion more explicit by describing some categories model theory works with.

Embeddings: This is the usual category used in algebra. The objects are structures whose definable sets are generated by a given collection of symbols, the *basic* relations, such as rings by $\{+,\cdot\}$. The morphisms are maps respecting a given set of relations.

Elementary embeddings : The objects are the *models* of a given *theory*, i.e. the structures in which a given collection of sentences is true. The morphisms are those embedding that respect arbitrary formulas. In other words, they respect not only the basic relations, but the comportment of the class $\mathcal{D}$ of definable sets with respect to projections, intersections, complements.

Interpretations: This is the most characteristic model theoretic category. Here the objects are structures, that need not be in the same language. The morphisms involve the linguistic viewpoint on formulas, and do not correspond to algebraic maps. An interpretation of A in B is a partial map $f : B^n \to A$, such that for any $C \in \mathcal{D}_A$, the pullback f^*C of C under f is in $\mathcal{D}_B$. In particular, the domain of f is in $\mathcal{D}_B$, and so is the equivalence relation $E = f^{-1}(=)$ on $\mathrm{dom}(f)$.

One can therefore identify A as a set with B^n/E. The *induced structure* on B^n/E, by definition, is the class of E-invariant relations on B^n, present in $\mathcal{D}_B$. An interpretation identifies A, as a structure, with B^n/E endowed with a sub-class of the relations in the induced structure. (We will often be interested in the case where the entire induced structure is taken.)

Familiar interpretations include the "construction" of natural numbers from sets; of rational from natural numbers; of reals from the structure consisting of the rationals, and the sets of rationals; of complex from real numbers; of formulas and recursive sets in the natural numbers (Gödel); of algebraic geometric objects in the real or complex fields. Model theory undertakes to describe systematically the possible interpretations, in a given structure or class of structures.

1.2 Universal Domains

The notion of a universal domain for the class of fields was introduced by Weil in the *Foundations of Algebraic Geometry*. The idea is to have a large structure with uniform features, in which every field under consideration can be embedded; in the case of fields it will be a large algebraically closed field. Questions about a given field can sometimes be reduced to "geometric" questions about the universal domain, and "rationality" questions; for the latter one has the aid of Galois theory.

The scheme revolution put this idea into some disuse in algebraic geometry. In place of the embedding into the universal domain, one puts a structure sheaf on a variety; in this (alternate) way, one remembers the way it arose from the fundamental structure. In model theory the universal domain remains fundamental, and if a field is present, one must do some synthetic geometry to discover it. Note in particular that example (4.2) above cannot be intrinsically described as a ringed space.

Here is the general definition; it can be taken either in the embeddings or elementary embeddings category, or sometimes in intermediate ones.

Let T be a theory, i.e. a set of sentences. Let $\mathcal{C}$ be the class of structures in which T holds true: the class of fields of a given characteristic, or

differential fields of char. 0, etc. One considers $\mathcal{C}$ as a category by taking morphisms to be embeddings or elementary embeddings (or sometimes an intermediate choice), as appropriate.

Definition 1.1 A structure $\mathbb{U} \in \mathcal{C}$ is a *universal domain* for T if every countable $M \in \mathcal{C}$ embeds into $\mathbb{U}$, uniquely up to an automorphism of $\mathbb{U}$.

We have combined here two properties of the universal domain; the *existence* of an embedding, together with a certain statement of existence of extensions of a given embedding, is usually called *saturation*. The *uniqueness* is called *homogeneity*.

The restriction to countable M here is somewhat arbitrary, and can be replaced with a number of variants. We have not specified $\mathbb{U}$ completely, to avoid getting into entirely irrelevant (and easily solved) set theoretic issues. For instance if the continuum hypothesis holds, one can ask that $\mathbb{U}$ have the size of the continuum; and then $\mathbb{U}$ is determined uniquely, up to a non-canonical isomorphism.

Existence. There are precise criteria for the existence of a universal domain; the essential point is that $\mathcal{C}$ should admit *amalgamation*, i.e. given two maps $f_i : A \to B_i$ there should exist $C \in \mathcal{C}$ with and $g_i : B_i \to C$ with $g_1 f_1 = g_2 f_2$. If one uses elementary embeddings, a universal domain always exists for a given (complete) theory. If one can show amalgamation for a less restrictive category, say for embeddings, one obtains additional information on the definable sets of the universal domain. (See below). This can easily be verified in our examples.

The following general procedure has been described by Abraham Robinson:

1) Identify an aspect of a given situation that you wish to concentrate on, described by universal laws; for instance the ordering on $\mathbb{Q}$, the laws being the axioms of ordered fields. Here "universal" refers to the nature of the quantifiers in the sentences describing the laws; the class of models of these sentences is to be closed under substructures.

2) Construct a universal domain for the class of structures satisfying these laws. In the case of ordered fields, one obtains a real closed field.

3) In the universal domain, projections of definable sets are usually easy to understand. If $a \in \mathbb{U}^m$ is *not* in the projection of $R \subset \mathbb{U}^{m+n}$, this cannot be due to an accidental occurrence, but rather it must follow from one of the universal laws of the class that $(a, y) \in R$ is impossible (otherwise, there would be a solution in the universal domain). This makes the class $\mathcal{D}$

comprehensible in $\mathbb{U}$.

4) We will work with the universal domain itself. However sometimes one is given in advance a particular structure, such as $\mathbb{R}$ or $\mathbb{C}$, and wishes to determine the structure of their definable sets. This can often be done most efficiently by finding first the universal domain $\mathbb{U}$ (in these examples, of ordered fields, or fields of char. 0) and then showing that the given structure forms an elementary submodel. (The fact that the definable sets of $\mathbb{U}$ are understood makes the last task easier.) In the cases of $\mathbb{C}$ and $\mathbb{R}$, Robinson noted that this procedure yields Hilbert's Nullstellensatz and Hilbert's problem on expressing rational functions as sums of squares (proved earlier by E. Artin), respectively.

The last step expresses a connection between the topological completion process, yielding a complete ordered or valued field, and the model theoretic one. The model theoretic completion "includes" the topological one, but adds also infinite and infinitesimal elements (to $\mathbb{R}$, or to the value group.)

Note that the topological option is not present in examples (3.1-3.3). The universal domain for differential fields (say) may at first seem disturbing, being an abstract object whose elements cannot be identified with (say) functions. It is a unique object, but up to a non-canonical isomorphism. This is however not essentially different than the case of the algebraic closure of a field of characteristic $p > 0$.

The key feature of the universal domain is the compactness theorem:

Theorem 1.2 (Model theoretic compactness) *Let $\{X_i : i = 1, 2, \dots\}$ be a family of definable subsets of $\mathbb{U}^n$. If every finite subfamily has non-empty intersection, then $\cap_i X_i \neq \emptyset$*

This is the fundamental finiteness principle of model theory. The apparently infinitary nature of the construction of the universal domain really goes to bring out this finiteness.

In our application, in characteristic p, we will use this theorem as follows: when A is a definable Abelian group in the universal domain, if $p^n A \neq 0$ for any n, then $\cap_n p^n A \neq 0$; moreover this intersection reflects the properties of $p^n A$ for large n.

For another application, see the Ax-Kochen theorem, discussed below.

Galois Theory. The universal domain has a large automorphism group; sufficiently large for a good Galois theory. It is also large enough for the

"Erlangen program" to hold; different "structures" have different automorphism groups. We noted above that $(\mathbb{Z}, +, \cdot)$ and $(\mathbb{Z}, +, 1)$ have the same (trivial) automorphism group, but are quite different as structures. This can no longer happen once one goes to the universal domain. If a relation R is not definable from some others $R_1, R_2, \ldots$, there exists an automorphism of the structure with respect to $R_1, R_2, \ldots$ that does not leave R invariant.

We will later need the notion of algebraic closure. If X is a subset of the universal domain $\mathbb{U}$, let $\mathrm{Aut}(\mathbb{U}/X)$ be the group of automorphisms of $\mathbb{U}$ fixing X pointwise. We define the *algebraic closure* $\mathrm{acl}(X)$ to be the set of elements $u \in \mathbb{U}$, whose orbit under this group is finite. An equivalent definition in terms of formulas: $a \in acl(X)$ iff $\phi(a, b)$ holds for some formula ϕ and parameter b from X^k, such that $\phi(x, b)$ has a finite number of solutions. The basics of Galois correspondence are valid here; in particular there is a bijection between closed subgroups of $\mathrm{Aut}(acl(X)/X)$, and subsets of $acl(X)$ containing X and closed under definable functions. (Here one must use Shelah's imaginary elements; we will not discuss this.)

Digression: The Ax-Kochen Theorem.   Artin conjectured that a hypersurface of degree d in $\mathbb{P}^n$ has a p-adic solution, whenever $n \geq d^2$. This is not true ([39]), but is true asymptotically in p: for a fixed degree d, for all sufficiently large p, every hypersurface of degree d in $\mathbb{P}^{d^2}$ has a p-adic solution. Ax and Kochen ([1]) discovered and proved the asymptotic statement using the model theoretic ideas above. Recalling the proof in full would take us out of our way, but let us mention some aspects.

The *Ax-Kochen-Ershov principle* states that for a Henselian K of residual characteristic 0, the set of sentences true in K is determined by the the knowledge of those true in k (as a field) and Γ (as an ordered group.) To have an idea of the proof, recall the characterization of the Laurent series field $k((t))$ ($\mathrm{char}(k) = 0$) as the unique complete valued field, with residue field k value group $\mathbb{Z}$. Replacing "complete" by "Henselian and saturated" permits a generalization of this result to arbitrary value groups. At the level of first-order sentences - they do not change upon moving to a universal domain - it follows then that "Henselian" suffices. (cf. [27].)

Fix an integer d. We saw above that the statement of Artin's conjecture for this d (and say, for $n = d^2$) can be made by a first-order sentence $\sigma_{d,n}$. The property of being a Henselian valued field can also be expressed by an infinite collection of sentences.

Consider a structure with a sort P for the set of primes; an additional sort Q, and a function $f : Q \rightarrow P$. The fiber $Q_p = f^{-1}(p)$ of f above the

prime p is taken to be the valued field $\mathbb{Q}_p$.

Embed this structure in its universal domain. The set E of elements of P such that the fiber Q_p fails to satisfy σ is a definable set. So is the set E_n of all p such that, in addition, the residue characteristic of Q_p is at least n. If the asymptotic conjecture fails, then each E_n is non-empty; by the compactness theorem , there exists an element $p^* \in P$ (in the universal domain) such that $F = Q_{p^*}$ has residue characteristic 0, and σ fails. Since each Q_p is Henselian, so is F.

It suffices now to show separately, working in characteristic zero, that σ holds in F. Ax and Kochen actually do this using compactness again: σ is known to hold (Chevalley, Tsen) for the fields $F_p((T))$; repeating the construction for these, one obtains a field F', limit of the $F_p((T))$'s. But F, F' are saturated and evidently have the same residue field and value group, hence are isomorphic by the Ax-Kochen-Ershov principle mentioned above.

2 Stability

The model theory discussed so far stayed close to algebra and its foundations. We move now to another school, pursuing model theory in the abstract. These schools developed quite separately from the the 50's to the early 80's . Moreover, starting in the 70's, the abstract model theory school begun to work with hypotheses, internally motivated, that left nearly all the significant model theoretic algebra outside the door. (The only known theories of algebraic interest satisfying the stability hypotheses are theories of modules over a fixed ring, the theory of separably closed fields, and the theory of differentially closed fields.) This isolation made possible, in laboratory conditions as it were, the discovery of concepts proper to the model theoretic view. A rapprochement began in the mid-80's, partly because of the discovery within stability theory that certain classical structures play a key role in general model theoretic phenomena, and must be explicitly studied even if one is not interested in them for their own sake. A synthesis is now in its initial stages; one hopes the power of stability-theoretic analysis will eventually become available for all of the structures surrounding number theory.

The application to Mordell-Lang reported on below uses differential algebra, in any characteristic, where stability was already applicable; and difference algebra, where the appropriate theory was developed more recently, see [22],[9].

2.1   Categoricity and dimension

The motivating problem for the development of much of stability was the "spectrum problem", the question of the number of isomorphism types of models of a given theory in a given cardinality. This takes place in the category of elementary embeddings, but the question is preserved under bi-interpretations, and encourages the discovery of invariants for the latter category.

The beginning of the subject was Morley's analysis of categoricity. In an old use of the term, a theory is called categorical if it has only one model, up to isomorphism. However, an early result (Löwenheim-Skolem, cf. [4]) states that a theory with one infinite model has models of any infinite size. Thus one says that a theory is κ- *categorical* if all its models of size κ are isomorphic. Morley showed that this notion does not depend on the cardinal κ, if $\kappa > \aleph_0$. This is a typical statement of the essentially non-set-theoretic nature of model theory. (Categoricity in the cardinality of the integers is an interesting but quite different model theoretic story, which we will leave aside.)

Note that the theory of algebraically closed fields of a given characteristic is κ - categorical, for uncountable κ: an algebraically closed field is characterized by its transcendence degree over the prime field. This transcendence degree is another face of dimension theory for varieties.

Morley started with the assumption of categoricity, and recovered an analog of transcendence degree and dimension, the Morley-rank. This is only the first of a host of dimension theories occurring in model theory, usually ordinal-valued, applying in more general situations. We will however only discuss Morley rank, and only in cases where it is finite; in this case we will refer to it as (Morley) dimension. The following definition refers to definable sets in the universal domain, possibly using parameters from the universal domain.

Definition 2.1

- A definable set X in the universal domain has Morley rank ≥ 0 iff it is nonempty.

- X has Morley rank $\geq n + 1$ if for each m, it contains m disjoint definable subsets of Morley rank $\geq n$.

- If X has Morley rank $\geq n$, but not $\geq n + 1$, we say it has Morley rank n. In this case there exists a maximal m, called the *multiplicity*,

such that X is the disjoint union of m definable sets of Morley rank n.

Note that X has Morley rank 0 iff it is finite. This dimension theory can be shown in general to have some of the basic properties of dimension theory for algebraic varieties; in particular, the dimension of a Cartesian product is the sum of the dimensions of the factors. (These functorial properties are as important in practice as the definition itself.)

The dimension theory serves both to begin an analysis of arbitrary definable sets, and to point out ones of particular importance.

Definition 2.2 An infinite structure U is called *minimal* if for every definable $R \subset U \times U^m$, for some integers n, n', for any $b \in U^m$:

Either $R(a, b)$ holds for at most n values of a, *or else* it holds for all but n' values of a.

If $U = \mathbb{U}$ is a universal domain, then $\mathbb{U}$ is *minimal* iff every definable subset of $\mathbb{U}$ is finite or cofinite. This includes definable subsets allowing parameters from $\mathbb{U}$.

A minimal structure has Morley dimension 1 and multiplicity 1. Conversely, every definable subset of $\mathbb{U}$ of Morley dimension and multiplicity 1 can be viewed as a minimal structure (with the induced structure.) In seeking to identify the "fundamental structures" involved in a given one, we will begin by agreeing that a "fundamental structure" should be minimal. Later we will define a notion of equivalence (non-orthogonality) to decide when two minimal definable sets correspond to the same fundamental structure. The minimal structures interpretable within a given structure M of finite Morley dimension, throw a great deal of light about M (both in abstract categoricity theory, and in the analysis and applications to be described here.)

Examples

1. Algebraically closed fields are minimal structures.

If V is any variety over an algebraically closed field K, say for simplicity without nontrivial automorphisms (as a K-variety), one can view V as a structure by taking the basic relations on V^n to be the Zariski closed subsets. Then V interprets K (hence V is bi-interpretable with the field K.)

2. Let G be a finite group, and consider the complex representations of G. The relations are the graph of addition, of the complex scalar multipliers, and of the operations of "multiplication" by any $g \in G$. There exists a universal domain for such representations (a direct sum of infinitely many

copies of the regular representation.) The Morley dimension equals the order of the group. The minimal definable sets (up to an equivalence relation to be defined below) correspond to the irreducible representations.

Within a minimal structure D, algebraic closure induces a matroid, or combinatorial pregeometry (cf. [34].) In other words, the exchange law holds: if $a \in acl(B \cup \{c\}) \setminus acl(B)$, then $c \in acl(B \cup a)$. Thus one can define a "transcendence degree" as follows: the Morley transcendence degree of a set X is the cardinality of any maximal subset Y of X, whose elements are algebraically independent.

In example (1) above, this is the usual transcendence degree. In example (2), each minimal definable set determines a different dimension theory; the transcendence degree of a representation with respect to a given minimal set, is the number of components of the associated irreducible representation that it contains.

Digression: The Baldwin-Lachlan theorem. The Baldwin-Lachlan theorem characterizes categoricity in terms of minimal structures. One says that a model M is *prime* over a subset Y if any elementary embedding of Y into $\mathbb{U}$ extends to an embedding of M into $\mathbb{U}$. In many situations, such as when M is countable, or has Morley rank, the prime model is determined up to a (non-unique) isomorphism.

Theorem 2.3 *Let T be a complete theory in a countable language. T is $\aleph_1$-categorical iff for any model M of T,*

1. *M has a definable subset D that is minimal, with the induced structure.*

2. *M is the prime model of T over D.*

Note that in (1), D is a definable subset of M, but is considered as a structure in its own right, with the structure induced from M.

The model M is thus determined by the minimal structure D. This could happen because D interprets M; an example we say earlier is that where M is a variety over an algebraically closed field F (with the "Zariski" structure). However, the situation could be considerably more complicated; a great deal of model theoretic analysis revolves around this issue. We will not directly require this analysis, though we will use some of the theory of groups in the definable category, that was developed with these issues in

mind. It turns out that M can be constructed as a series of "fiber bundles", not locally trivial however, with structure groups interpretable in D.

Returning to the statement of the theorem, one sees that the minimal structure D controls in some way the behavior of the high-dimensional structure M. This fits with our theme of searching, when one is not initially known, for a fundamental structure responsible for a given geometry. Note the analogy to the Ax-Kochen-Ershov principle, stating that the simple objects - residue field, value group - in some sense control the more complex valued field.

2.2 Superstability and regular types

An extensive deepening and generalization of the Morley- Baldwin-Lachlan theory was achieved by Shelah ([36]). His main goal, irrelevant for our present purposes, was replacing the categoricity condition by any nontrivial constraint on the number of isomorphism types of models. One of the outcomes was a technology that applies to theories with ordinal Morley rank, and in particular to differential fields. The key notion is that of a *regular type*, generalizing the minimal structures occurring in the previous theory. These are orbits of the automorphism group of the universal domain (over some base structure), forming a pregeometry under a natural dependence relation. Thus there is an integer-valued dimension theory associated with every regular type. Shelah's theory shows that there are enough regular types to control the entire structure.

Much of the power of Shelah's theory resides in its ability to treat the regular types at higher tiers with nearly the same facility as the minimal ones; it is likely that future applications will utilize that power. For a start, one would like to see a classification of the regular types associated with PDE's (differential fields with several derivations). However, the present applications use only the finite rank theory (together with the ability to recognize it within an infinite rank ambient structure); we will limit out discussion to this case.

We will mention just one statement of Shelah's theory, simplified by being specialized to the finite Morley rank context. (Thus avoiding the problematics associated with one regular type embedded in another.) We first define the equivalence relation on minimal sets that we alluded to earlier.

Definition 2.4 Let D_1, D_2 be two minimal definable sets within a structure M. D_1, D_2 are *equivalent* if there exists a minimal D and definable, finite-to-one maps $f_1 : D \to D_1$, $f_2 : D \to D_2$, with cofinite ranges.

This notion of equivalence is very coarse from the point of view of algebraic geometry; with the structure of Zariski closed subsets, a variety is minimal iff it is a curve; and all curves are equivalent to each other! But indeed all curves are bi-interpretable; what they interpret, in effect, is algebraic geometry. In general the equivalence relation is slightly stronger than bi-interpretability, but is reasonable for this category. It can be useful if we are interested in determining whether a given structure belongs to algebraic geometry, or to some other domain.

We wish to point out the phenomenon of "orthogonality" of distinct dimension theories within stability theory. For minimal sets, it expresses itself as follows: if two (or more) minimal sets are inequivalent, they do not interact at all.

Lemma 2.5 (Orthogonality) *Let $D_1, \ldots, D_k$ be pairwise inequivalent minimal definable sets in a universal domain $\mathbb{U}$. Then the following statements hold:*

- *Every definable $R \subset D_1{}^{m_1} \times \ldots \times D_k{}^{m_k}$ is a finite union of "rectangular" sets of the form $R_1 \times \ldots \times R_k$, R_i a definable subset of $D_i{}^{m_i}$.*

- *If the D_i are defined over some algebraically closed set B, $\mathrm{Aut}(D_1 \cup \ldots \cup D_k / B) = \Pi_i \, \mathrm{Aut}(D_i / B)$.*

Both conclusions assert that the definable sets D_i are "unrelated". the 2nd statement brings out the strength of the first; it is equivalent to it, because of the homogeneity of the universal domain.

Let us note how the case $k = 2, m_1 = m_2 = 1$ of the first conclusion follows from the definitions. Remove from R the horizontal or vertical lines $a \times D_2$, $D_1 \times b$ that it may contain (or even contain up to a finite set.) Let $\pi_i : R \to D_i$ be the projections. For $a \in D_2$, $\pi_2{}^{-1}(a)$ is either finite, or a cofinite subset of D_1, by the minimality of D_1. Since we removed the lines, it must be finite. But then the maps $\pi_i : R \to D_i$ are finite, and D_1, D_2 are equivalent.

Finally, let us mention in passing how regular types enter when the object of study is the category of models and elementary embeddings. There is a notion of a finitely generated extension of a model M (prime model over M + a finite set.) There is also a notion of "tensor product" of two extensions of M. (Prime model over their independent join, see below.) This operation yields a semi-group. Shelah's theory (in the finite Morley rank context)

shows that it is a free Abelian semi-group. The generators are in one-one correspondence with equivalence classes of minimal definable sets. Two such minimal sets D,D' are equivalent iff $M(a)$, $M(a')$ are isomorphic over M when $a \in D, a' \in D'$. If they are inequivalent, then the two extensions are "free" over M, i.e. the tensor product is their only possible relation inside a larger common extension.

2.3 Stability

The notion that lends its name to the subject turns out to be a simple combinatorial condition on bipartite graphs. It is valid in all graphs interpretable in structures with ordinal Morley rank, in particular in those occurring in algebraic or differential algebra. It has consequences of surprising elegance and force: it brings the critical notion of a "generic" element into the subject; and permits the development of a smooth and very powerful calculus of dependence and independence.

We will consider bi-partite graphs $(P, Q; R)$, or more generally pairs of sets P,Q together with relations $R_1, \dots , R_k \subset P \times Q$.

A key example is L_n, the bipartite graph consisting of two copies P,Q of $\{1, \dots , n\}$, and R is the graph of the order relation $i < j$. $(P, Q; R)$ embeds L_n iff there exist $a_i \in P, b_i \in Q$ $(i = 1, \dots , n)$ such that $(a_i, b_j) \in R$ iff $i < j$.

Definition 2.6 $(P, Q; R_1, \dots , R_k)$ is *combinatorially stable* if for each i, for some n, $(P, Q; R_i)$ does not embed L_n.

Given $R \subset (P \times Q)$, we denote

$$R(b) = \{a \in P : (a, b) \in R\}$$

for $b \in Q$; we then think of the graph as a family of subsets of P, indexed by Q. Dually, $R'(a) = \{b \in Q : (a, b) \in R\}$.

A type of example that will concern us later is called a pseudo-plane. One should think of P as a rough kind of "plane", with the $R(b)$ analogous to lines on the plane.

Definition 2.7 $(P, Q; R)$ is a *pseudo-plane* if each $R(b)$ is infinite, each $R'(a)$ is infinite, but there exists m such that for distinct $b_1 \neq b_2 \in Q$, $|R(b_1) \cap R(b_2)| \leq m$.

Sometimes one adds the dual condition that "through two distinct points there pass at most finitely many lines".

Note that all pseudo-planes are combinatorially stable.

We will make a homogeneity assumption that will make the statement of the theory easier. It is not essential; in particular if the weak primitivity condition below is omitted, one shows that there exists a definable equivalence relation with finitely many classes, such that it holds on each class.

Assumption 2.8 $\Gamma = (P, Q; R_1, \dots, R_k)$ *is combinatorially stable. A group G acts on Γ by automorphisms. G is transitive on P. Moreover, there is no definable, G-invariant equivalence relation on P with finitely many classes.*

Theorem 2.9 *There exists a unique definable relation $\mathcal{F}_\Gamma \subset (P \times Q)$ such that:*

- *For $b \in Q$, $\mathcal{F}_\Gamma(b)$ is a intersection of sets of the form $R_i(b)$ and their complements.*

- *For any $b_1, \dots, b_k \in Q$, $\mathcal{F}_\Gamma(b_1) \cap \dots \cap \mathcal{F}_\Gamma(b_k) \neq \emptyset$*

Note that $\mathcal{F}_\Gamma$ is G-invariant (by the uniqueness). The relation $\mathcal{F}_\Gamma$ is called the combinatorial non-forking, or freeness relation. For instance, if $P = Q$ and R is the graph of equality, then $\mathcal{F}_\Gamma$ is the complement of R. If R is the graph of inequality, then $\mathcal{F}_\Gamma = R$. If R is a pseudo-plane, again $\mathcal{F}_\Gamma$ is the complement of R. More generally, when $k = 1$ say , $\mathcal{F}_\Gamma(b)$ is either $R(b)$ or $P \setminus R(b)$, whichever is "larger" or "more general"; $\mathcal{F}_\Gamma$ is the relation one expects between two random points. In the presence of an ordering, one expects $x < y$ or $y < x$ with equal probability, but in the stable regime the measure is 2-valued.

One has a *symmetry* phenomenon: if P, Q both satisfy 2.8, and $R^t \subset (Q \times P)$ denotes the transposed graph, $\Gamma^t = (Q, P, R^t)$, then Γ^t is stable, and $\mathcal{F}_{\Gamma^t} = \mathcal{F}_\Gamma{}^t$.

Definition 2.10 We say that $a \in P$ is *generic* over a set A, with respect to Γ, if $a \in \mathcal{F}_\Gamma(b)$ for any $b \in Q \cap A$.

In a saturated model, in particular in the universal domain, the finite intersection property in the definition of $\mathcal{F}_\Gamma$ together with the compactness theorem ensure that over any countable set A, generic elements exist. These are elements of P "in general position".

Stable structures.

Definition 2.11 A theory (or structure) is *stable* if every graph interpreted in it is combinatorially stable.

One uses this as follows. Let P and Q as two parts of a given structure, say with Q definable over a countable set A, and P forming an orbit of $Aut(\mathbb{U}/A)$. The assumption 2.8 will hold if A is *algebraically closed*. (For this one needs to take into account "imaginary elements", i.e. elements of structures interpretable in the original structure, and we will do this.) Let R_i index all the definable relations on $P \times Q$ (over A). One obtains a single "freeness" relation $\mathcal{F}$; it is the intersection of the freeness relation $\mathcal{F}_i$ corresponding to R_i. One can still use the terminology of genericity.

Definition 2.12 Let P be an orbit of $Aut(\mathbb{U}/A)$, $a \in P$, and let B be a algebraically closed substructure, $A \subset B$. Then a is *generic* over B if for any definable Q,R a is generic over B with respect to $(P, Q; R)$

It is a characteristic of stability that any two elements of P generic over B are conjugate, not only under $Aut(\mathbb{U}/A)$, but also under $Aut(\mathbb{U}/B)$. Thus given $b \in Q$, a relation $R(a, b)$ holds either for all or for no generic elements $a \in P$.

The uniqueness of the non-forking relation leads to compatibilities between the $\mathcal{F}_\Gamma$ of the various definable relations. One consequence that should be stressed in this context is Shelah's definability lemma. It has the effect that the class of definable sets is closed also under constructions using expressions such as "for a generic element $a \in P, \dots$ ". We will see an example in 2.16 below.

Lemma 2.13 *Let R be a definable relation on $P \times Q$. Then*

$$\{b \in Q : \text{ for some generic } a \in P, R(a, b)\}$$

is a definable subset of Q.

Proof. Indeed, it is just $\{b \in Q : \mathcal{F}_\Gamma(b) = R(b)\}$, where $\Gamma = (P, Q; R)$.

Independence An equivalent language, exhibiting better the symmetry, is the language of independence.

Definition 2.14 Let A be a countable subset of $\mathbb{U}$, $a, b \in \mathbb{U}$, and let P,Q denote the orbits of a, b under $Aut(\mathbb{U}/A)$. For any definable relation R, let $\Gamma_R = (P, Q, R \cap (P \times Q))$. We say that a, b are *independent over A* if for any such R, $(a, b) \in \mathcal{F}_{\Gamma_R}$.

Then a,b are independent over A if a is a generic element of P, not only over A but also over $A \cup \{b\}$.

Two of the formal properties of independence are symmetry and transitivity: a, b are independent over A iff b, a are; if a, b are independent over $A \cup \{c\}$, and a, c are independent over A, then $a, (b, c)$ are are independent over A.

Every structure of (finite) Morley rank is stable. Examples (2.1), (3.1), (3.2), (4.1) are stable, hence so are all structures interpretable therein; in particular, (4.2). In the case of subfields of $\mathbb{C}$, a,b are free over a subfield A iff the fields $A(a), A(b)$ are free over A in the sense of field theory. (One can refine to a notion generalizing linear disjointness.)

The structures of finite rank that we will consider within difference fields (analogously to 4.2) are not stable; but enough of the bipartite graphs interpretable there are stable to explain the interactions between different part of the structure, and create a good theory of independence. The 2-valuedness referred to above no longer holds however.

Stable groups and Internality.   We would like to use the above notions to say something about the possible interaction between a definable group G, and another definable set P. Generally one can show that a certain quotient of G is "internal" to P, isomorphic to a group interpretable over P; whereas the quotient is "orthogonal" to P, cannot interact with P at the level of generic elements. We will prove a special case of this here.

A definable group is a group object in the category of definable sets; i.e. a group whose underlying set and operations are all definable. We say that G is *connected* if it has no definable subgroups of finite index. The group G acts on itself by left translation; if G is connected, the hypothesis 2.8 at the beginning of this subsection holds for $P = G$, eliminating the need to work over an algebraically closed substructure.

Let P be another definable set; view it as a structure in its own right, with the induced structure.

Definition 2.15 G is *P-internal* if there exists a definable isomorphism in $\mathbb{U}$ between G and a group interpretable in P.

This stands at the opposite extreme of *orthogonality* between G and P.

We will need the following lemma, to recognize when a definable group in some part of a large structure $\mathbb{U}$, actually belongs to P.

Lemma 2.16 *Suppose G is a definable group, P a definable set, and every element of G is in the algebraic closure of some elements of P. Then there exists a finite normal subgroup K of G such that G/K is P-internal.*

Proof. We consider the Abelian case for convenience. The assumption implies the existence of a definable relation $R \subset P^n \times G$, such that $R(b)$ is finite for $b \in P^n$, and $\cup_b R(b) = G$. Let $R'(a) = \{b : a \in R(b)\}$, $K = \{b \in G :$ for generic $a \in G, R'(a) = R'(ab)\}$. Then K is definable, by lemma 2.13. It is easy to check that K forms a subgroup. For a fixed a, $\{b : R'(a) = R'(ab)\}$ is finite. Fix a generic element $(a_1, \dots, a_k)$ of G^k. One can show that if k is chosen large enough, then for any $b \in K$, some i is "sufficiently generic over b", so that $R'(a_i) = R'(a_ib)$. Hence K is finite. Moreover, mapping an element $b \in G$ to the k-tuple $R'(a_1b), \dots, R'(a_kb)$ is an injective map on G/K. The range of this map consists of definable subsets of P rather than elements; however these are uniformly definable subsets, and can be parameterized by a (definable) "moduli variety", so that we get elements. We can then construct a group H in P such that the map becomes an isomorphism. $\square$

The definition and theory of stability are due to Shelah. In the setting of groups, it is Poizat's, [33]. For a treatment along the lines described here, see [22].

2.4   Local modularity

All the structures we have considered so far possess considerable geometric complexity. We need to consider now another, much simpler class of structures, where the definable sets behave as if they were defined by linear equations. This class forced itself on the attention of model theorists in many different ways; this is reflected in its many synonymous names: 1-based, weakly normal, locally modular, linear type, (disintegrated or) module-like. There is a sharp divide between linear and nonlinear phenomena in model theory, in this sense. Moreover, certain problematics exist exclusively in the geometrically simpler "linear" case. See the survey [14] for evidence of this.

Lemma 2.17 *The following conditions on a stable structure are equivalent.*

1. *The lattice of algebraically closed subsets of $\mathbb{U}$ satisfies the modular law.*

2. *Any two algebraically closed subsets are independent over their intersection. (For this, all interpretable sorts must be taken into account.)*

> 3. *There is no psuedo-plane* $(P, Q; R)$ *with* P *an orbit of* $\mathrm{Aut}(\mathbb{U}/B)$ *for some finite or countable* B, *and* Q, R B-*definable*.

(3) asserts the non-interpretability of an even rudimentary analog of point-line geometry. (1) suggests a relation with modules; indeed there is a very close relation. Any module is stable and locally modular, if one takes the basic definable relations to be the ones defined by R-linear equations. Conversely, under certain non-degeneracy assumptions, one can show that a locally modular theory interprets an Abelian group, and behaves like a module within that group.

In the finite Morley rank context, the local modularity condition can also be stated as follows: on a definable set of Morley dimension n, there is no l-dimensional family of k - dimensional subsets, with $k + l > n$ ([6],[10]). One needs of course to define precisely the notions of a family and its dimension. Contrast this however with the case of the projective plane in algebraic geometry, of dimension 2, where the hypersurfaces of degree d form a family of 1-dimensional subsets, of dimension $(d^2 + 3d)/2$. It can be shown thatin any non-locally-modular minimal set, the plane (of dimension 2) will contain arbitrarily large famíles of 1-dimensional subsets.

We note a local-global principle ([6]), demonstrating the qualitative control of minimal types over a finite-dimensional structure.

Theorem 2.18 *A structure of finite Morley rank is locally modular iff every rank-one structure interpretable therein is locally modular.*

It can be shown ([5])that the "nondegenerate" parts of a locally modular structure are controlled by definable Abelian groups. Inside these groups, the name "module-like" is justified by the following (due to Zilber in low rank and to Pillay in the $\aleph_1$-categorical context, where we will use it.)

Theorem 2.19 *[21] Let* G *be an Abelian group interpretable in a locally modular stable theory, and let* X *be a definable subset. Then* X *is a finite Boolean combination of cosets of definable subgroups.*

The conclusion of this theorem is in clear analogy with that of the Mordell-Lang conjecture (see §3). Indeed this conjecture can be viewed, model theoretically, as stating that certain diophantine-geometric structures are locally modular. The robustness of the model theoretic notion of local modularity will allow us to transfer the question from a high-dimensional Abelian variety, where the geometry is hard, to a one-dimensional object, where we can sometimes solve it.

2.5   Zilber's Trichotomy

Let us consider three types of minimal structures D.

(1) D is called *disintegrated* if the corresponding combinatorial geometry is trivial: $\mathrm{acl}(X \cup Y) = \mathrm{acl}(X) \cup \mathrm{acl}(Y)$ for $X, Y \subset D$. Equivalently, every *family* of definable subsets of D^2 is a combination of degenerate ones, in one of the following senses: it is constant (consists of a single definable subset of D^2); it is a family of points; or it is one of the familes $(D \times \{b\} : b \in D)$ or $(\{a\} \times D : a \in D)$. There is no *moving* geometry of interest on D^n, for such a D.

There are many very interesting disintegrated minimal structures, though evidently their interest does not derive from their dimension theory or algebraic closure relation. Any infinite, locally finite simplicial complex, with transitive automorphism group on vertices, forms such a structure. The basic (n-place) relations are given by the $n - 1$-dimensional simplices; the other definable relations are all Boolean combinations of local ones, i.e. relations R such that for any x, there are only finitely many $y_1, \dots, y_k$ with $R(x, y_1, ...y_k)$. The transcendence degree of such an object is just the number of components.

It is not known if any such (infinite, connected) disintegrated structure can be interpreted using algebraic differential equations. At all events, we will see that they cannot exist inside a finite dimensional group; therefore they will play no part in the results on Abelian varieties.

(2) D non-disintegrated, but locally modular. An example of this is a vector space V over a field or division ring F, with the definable subsets of V^n being (finite Boolean combinations of) solution sets to F-linear equations $\sum_{i=1}^{n} \alpha_i x_i = 0$. Note that there are now somewhat more moving families of "lines" in D^2, but not many; every such family is one-dimensional.

Using results mentioned in 2.4, it can be shown that up to equivalence, D can be taken to carry an Abelian group structure; and then the possibilities can be completely classified, [20]. A vector space modulo a finite subgroup is a quite typical example.

(3)   An algebraically closed field, or a curve over such a field; the basic closed sets are the Zariski closed ones.

Zilber [40] conjectured that these three types are the only ones, in the

sense of non-orthogonality. This can be viewed as a foundational statement, stating that algebraic geometry is sui generis; Every "geometry" with the same general dimension- theoretic properties is thus conjectured to be either at most of linear complexity, or to include algebraic geometry. A number of results in this direction exist. The conjecture itself is not true with the original set of assumptions. ([16]). It seems that the counterexamples, though not disintegrated as in type (1), have a tree-like or hyperbolic flavor. It turns out also to be possible to combine different minimal structures (of different types), without any known mechanism for taking them apart. At the moment a satisfactory substitute conjecture is missing. Nor do we know whether any counterexamples occur "in nature". The conjecture becomes true, however, under certain stronger assumptions on the dimension theory. These stronger assumptions hold in all examples mentioned in this paper. The number theoretic application we report on in §3 is based on this result.

For *real* algebraic geometry, incidentally, an analog of the conjecture was recently proved , [32].

2.6   Zariski Geometries

In the survey paper [25], Zariski geometries are motivated ab initio. Here we will view them as a topological setting for the study of minimal structures.

Definition 2.20 A *Zariski geometry* is a minimal structure D, together with a distinguished subcollection $\mathcal{C} \subset \mathcal{D}$ (called "closed sets"), and satisfying:

1. C contains the diagonals, and is closed under intersections, finite unions, and inverse images of projections.

2. $\mathcal{D}$ is generated by $\mathcal{C}$ as a Boolean algebra (on each D^n).

3. Every closed set is the union of finitely many closed, irreducible sets.

4. If X is a proper subset of Y, both closed subsets of D^n, and Y is irreducible, then $\dim(X) < \dim(Y)$

5. If X is a closed, irreducible subset of D^n, $\dim(X) = m$, and Δ is a diagonal $x_i = x_j$, then $X \cap \Delta$ is the union of closed irreducible sets of dimension at least $m - 1$.

Here "dimension" refers to the Morley rank. The first four items assert a compatibility between the topology and the definable sets, similar to the

relation between closed and constructible sets in algebraic geometry. The fifth item is the new structural hypothesis. It holds in smooth algebraic varieties (and normal l.c.i varieties.) In [24], we make an effort to classify Zariski geometries as such (see the introduction there.) Here we are content with understanding them in the coarser category of minimal structures.

Theorem 2.21 *Zilber's conjecture holds for minimal structures admitting a Zariski geometry.*

In fact we will only use part of this theorem: a Zariski geometry is either locally modular, or interprets a field of Morley dimension 1.

To give a brief idea of the proof, recall a well known *linear* analog, the fundamental theorem of projective geometry. This theorem recognizes a division ring (and a finite dimensional linear geometry over it,say) in terms of intersection properties of points, lines and planes of that geometry. Of course these properties are much more complicated in the nonlinear case, when we are trying to recover a field from the intersection structure of Zariski closed sets. The solution is to linearize; in the tangent bundle we ought to find something like projective geometry. From non-local modularity, one can show the existence of families of "plane curves" (one dimensional subsets of D^2) of arbitrarily high dimension. We consider the sub-family of curves passing through a point $p \in D^2$, and wish to see its trace in an "infinitesimal neighborhood" of this point. The problem is to define the tangent space, or the equivalence relation of tangency, using synthetic geometry. Equivalently, one wants to define a limit of a family of definable sets C_{p_1,p_2} as the two points "approach each other". This can be done by considering the total family in an appropriate space, intersecting with the diagonal, removing uninteresting components, and recognizing the component corresponding to pairs of tangent curves. This uses the "dimension theorem" (property 5). In practice, however, it is difficult to define tangency well directly in high dimensions, and the proof stays close to dimension one, defining first the multiplicative and then the additive group by successive "synthetic differentiation".

2.7 Abelian groups of finite Morley dimension

Having gathered information about the minimal geometries involved in a given structure, one would like to globalize it to the full structure. A very large chapter of stability theory is devoted to this effort. One finds a sequence of fibrations of the given structure, with fibers either interpretable in minimal geometries, or else transitively acted on by groups interpreted

over these geometries. Because of the possible occurence of finite groups, the situation is nontrivial even when the only minimal geometriy involved is the completely degenerate one consisting of a pure set (with no relations but equality.)

When the entire structure is contained in a finite-dimensional Abelian group, however, none of this happens, and we will be able to give a complete account of the transfer of information from the minimal geometries upward.

The first lemma is an easy generalization of facts from algebraic groups. Fix a definable Abelian group A, of finite Morley dimension. We will take the minimal subsets to be *indecomposable* (as we may), meaning they cannot be cut into finitely many pieces by the cosets of a definable subgroup of A .

Lemma 2.22 *Let X be a one-dimensional indecomposable definable subset of a finite - dimensional definable group A, $0 \in X$. Then X generates a subgroup B of A, in finitely many steps; for some n, $B = \{\sum_{i=0}^{n}(-1)^i x_i : x_i \in X\}$. Moreover, if $\dim(B) = d$, a generic element of B can be expressed in finitely many ways (at least one) as a sum of d elements of X*

Proof. The first statement, due to Zilber, generalizes the "indecomposability" theorem of algebraic groups. The proof is the same, once one has learned to replace the set - theoretic stabilizer:

$$Stab_{set}(X) = \{a \in A : a + X = X\}$$

By the more robust dimension-theoretic stabilizer:

$$Stab(X) = \{a \in A : \dim((a + X \setminus X) \cup (X \setminus (a + X))) < \dim(X)\}$$

They coincide for closed irreducible sets, but here we do not have this notion.

The second statement comes from easy dimension counting. $\square$

The concern of the remaining lemmas is the possible mixture of orthogonal geometries. It thus has no analog in algebraic geometry proper (or in any *single* geometry, in the sense of non-orthogonality equivalence). Example (4.1) serves better. Suppose X is a complex torus; the structure is generated by the closed analytic subvarieties. If X is an Abelian variety, this is by-interpretable with the field $\mathbb{C}$; there is plenty of structure. If X is chosen generically, $\dim(X) > 1$, then X^n has no closed analytic subvarieties, except for translates of group subvarieties; X is then a minimal structure of locally modular type. (cf. [25]). (Away from the tori, we can also find disintegrated geometries in 4.1.) The next corollary implies (in the general

context) that X admits a composition series into the two types. One clearly cannot expect to improve this to a direct sum decomposition. But we will do almost as well, and obtain decisive information from the "bottom layer" alone, where a direct sum decomposition does occur (2.25)

Corollary 2.23

- *A minimal subset of a group A cannot be disintegrated.*

- *Suppose A is generated by finitely many minimal subsets X_i, with each X_i equivalent to a minimal set k. Then there exists a k-internal group A' and a definable map $A \to A'$ with finite kernel.*

Proof. By 2.22, a minimal subset X of A may be taken to generate a definable subgroup B. Let (b_1, b_2) be a generic element of B^2, $b_3 = b_1 + b_2$. By the lemma, each b_i has the same algebraic closure as a d-tuple from X. Putting together the three d-tuples contradicts the triviality of algebraic dependence on X.

Similarly, if A is generated by the minimal sets X_i, then any point of A is algebraic over some points of k. The conclusion in this case follows by 2.16. $\square$

Definition 2.24 The *socle* of A is the sum of all connected definable subgroups, generated by minimal subsets of A.

By the orthogonality lemma, the socle is a direct sum of subgroups, each generated by a finite union of minimal sets in the same non-orthogonality class. Using the known structure of locally modular groups and Corollary 2.23 above, we obtain a complete analysis of the socle.

For instance suppose A is interpretable in $\mathbb{U}_D$ or $\mathbb{U}_p$. The constant field k has precisely the structure of an (algebraically closed field) - there is no further induced structure. It follows from this that every definable group over k is definably isomorphic to a group $G(k)$, where G is an algebraic group over k. Thus the socle of A is isogenous to a direct sum of an algebraic group over k, with a locally modular group. (In the case of $\mathbb{U}_\sigma$, k is a pseudo-finite field, and the group may be a twisted algebraic group, e.g. a unitary group; this interesting variation is of no importance in the proof however.)

Call a definable group B *rigid* if every connected definable subgroup of B is defined over $\mathrm{acl}(B)$. This sometimes follows by consideration of torsion points; it is automatically true for locally modular groups.

Proposition 2.25 (Reduction to the socle) *Let G be an Abelian group of finite Morley dimension, X an definable subset of G. Assume :*
i)   The socle A of G is rigid.
ii) $Stab(X)$ is finite.
Then X is contained in a the union of finitely many cosets of A, and a definable set of smaller dimension.

3   Applications

3.1   Geometric translation

We return to examples (3) of §1.1, to discover the meaning in these concrete cases of the above concepts. Let $\mathbb{U} = \mathbb{U}_D, \mathbb{U}_p$, or $\mathbb{U}_\sigma$ be the universal domain of differential fields of characteristic zero, separably closed fields of imperfection index 1, and difference fields of characteristic zero (3.1-3.3, respectively.)

None of the three structures has finite Morley rank. $\mathbb{U}_D$ falls just short, with Morley rank ω; the solution set within it to any nontrivial differential polynomial does have finite Morley rank. $\mathbb{U}_p$ is stable, but has no Morley rank; nonetheless minimal types and types of finite rank exist. These "types" are countable intersections of definable sets, rather than definable sets proper; this makes little difference in the stable context, and we will ignore the distinction. $\mathbb{U}_\sigma$ is not stable at all. We will ignore this too; another dimension theory can be developed for $\mathbb{U}_\sigma$, with the same effect for our purposes ([9]).

In all three cases we are interested in the finite dimensional structures interpretable in $\mathbb{U}$. One can show ("elimination of imaginaries") that in these structures, any structure interpretable in $\mathbb{U}$ is definably isomorphic to a structure definable in $\mathbb{U}$, i.e. to a definable subset D of $\mathbb{U}^n$; the definable subsets of D^m are then among the definable subsets of $(\mathbb{U}^n)^m$. We will be interested in the maximal case, "the induced structure", where one takes the definable relations on D^m to be all subsets of D^m that are definable as subsets of $\mathbb{U}^{nm}$ (cf. §1.1).

We begin with a geometric description of the definable sets, and criteria for their finite dimensionality. Actually we will only describe a family of "basic" definable sets; but an arbitrary definable set is a finite union of differences of such basic sets. In each of the three cases, a basic definable set corresponds to an algebraic variety, together with some additional geometric data.

Basic definable sets.

$\mathbb{U}_D$ We begin by describing certain jet bundles of a variety X. We can restrict attention here to smooth affine varieties. (See [7] for a different and more detailed description. of these bundles. Note in particular that when X is not defined over the constant field $k = \{u \in \mathbb{U} : Du = 0\}$, for $m = 1$, the bundle we define is not the tangent bundle but a certain torsor thereof, determined by the structural derivation D and the Kodaira - Spencer class of X. But for our needs a more naive description suffices.)

Let $X \subset \mathbb{A}^n$ be a smooth affine variety; we will identify X with the set of points of X with coordinates in $\mathbb{U}$. When $x = (x_1, \ldots, x_m) \in X$, write Dx for $(Dx_1, \ldots, Dx_n)$. Write $D_{[m]}x$ for $(x, Dx, DDx, \ldots, D^{(m)}x)$. Let $B_m X$ be the Zariski closure of $\{D_{[m]}x : x \in X\}$. Then $B_m X$ is an algebraic variety; we have obvious algebraic projection maps $\pi_{m,m-1} : B_m \to B_{m-1}$. It can (easily) be shown that $B_m X$ does not depend on the derivation D, and that B_m glues to a functor on algebraic varieties. When X is defined over the constant field, and $m = 1$, $B_1 X$ is just the tangent bundle.

A basic definable set H corresponds to an affine variety X together with a subvariety V of some $B_m X$, projecting dominantly to X. The set is defined by:

$$H_{X,V} = D_{[m]}^{-1}(V)$$

$\mathbb{U}_p$ In the case of char. p, one can use divided powers to give a description of the definable sets strictly analogous to the char. 0 differential case. But there is no advantage in this uniformity (for our purposes), and we will give another description.

As a field, $\mathbb{U}$ is separably closed, but has a unique purely inseparable extension $\mathbb{U}_m$ of degree $q = p^m$. The map $\phi = \phi_q : x \mapsto x^q$ is an isomorphism of of $\mathbb{U}_m$ with $\mathbb{U}$. For any variety X over $\mathbb{U}$, let X' be the variety obtained by "pulling back the coefficients using ϕ". Then ϕ_q induces a bijection $\phi_{q_*} : X'(\mathbb{U}_m) \to X(\mathbb{U})$.

On the other hand, via an inseparable analog of restriction of scalars, one can interpret the field extension $\mathbb{U}_m$ over $\mathbb{U}$. More generally, for any variety X over $\mathbb{U}$, one has a variety X_m and an isomorphism $\rho : X'(\mathbb{U}_m) \simeq X_m(\mathbb{U})$.

Composing, we obtain a bijection

$$\psi_q = \phi_{q_*} \circ \rho^{-1} : X_m(\mathbb{U}) \to X(\mathbb{U})$$

A basic definable set (of level m) is determined by a smooth affine variety

X, and a subvariety $V \subset X[m]$. We let

$$H_{X,V} = \psi_q(V(\mathbb{U})) \subset X(\mathbb{U})$$

Here it is natural to consider also ∞-definable sets. They are determined by a compatible collection $V = (V_m)_m$, $V_m \subset X[m]$ (for each m); then $H_{X,V} = \cap_m H_{X,V_m}$. The saturation of $\mathbb{U}$ implies that this is nonempty, and Zariski dense if each H_{X,V_m} is.

The combination of compactness and stability make such ∞-definable sets no harder to handle than definable sets, and we will ignore the slight technical differences in the sequel. By a definable subset of $B = H_{X,V}$ we will mean the intersection of B with a definable set.

$\mathbb{U}_\sigma$ Here a basic definable subset of X arises from a subvariety $V \subset (Y \times Y^\sigma \times \ldots \times Y^{\sigma^l})$, where Y is a finite cover of X via a rational map $g : Y \to X$. We let $\sigma_l(x) = (x, \sigma(x), \ldots, \sigma^l(x))$ (in coordinates), and $H_{Y;V} = \sigma_l^{-1}(V)$, $H_{X,V;g} = g(H_{Y;V})$.

Criteria for finite dimension.

$\mathbb{U}_D$ If $\pi_{m,m-1}$ has finite fibers on V, then $H_{X,V}$ has finite Morley dimension, indeed dimension at most $m \dim(V)$. To see this consider the case: $m = 1$, and V is the graph of a section of B_1X. (Actually the general case can be reduced to this one by the standard trick, absorbing the higher jets into X.) In this case every basic subset of $H_{X,V}$ has the form $H_{X,V'}$, V' a subvariety of V; and V' must be of the form $V \cap B_1Y$, Y being a subvariety of X. Thus Zariski closure gives an injective map from Kolchin closed subsets of X to Zariski closed ones. The finiteness of the Morley dimension now follows from the finiteness of the usual (Zariski) dimension.

Note that not every Zariski closed subset Y of X must occur in this correspondence; e.g. when X is defined over k, only integral varieties of the vector field V do. Thus it is quite possible, indeed common, for $H_{X,V}$ to have Morley rank 1, though X has higher dimension.

$\mathbb{U}_p$ Let $V = (V_m)_m$ be data for a basic (∞-) definable set, $V_m \subset X[m]$. Then $\dim(V_1) \leq \dim(V_2) \leq \ldots$ as algebraic varieties. If the sequence of dimensions stabilizes, then $H_{X,V}$ has finite dimension. Again this dimension is induced from algebraic dimension theory on the V_m.

We will use this in the context of an algebraic Abelian group A. In this case A_m has the structure of an algebraic group, for each m. Let

L be an algebraic closure of $\mathbb{U}$; then $A_m(L) \simeq_\rho A'(L_m)$ where L_m is the commutative ring obtained from L by adding a p'th root to an element, equivalently $L' = L[\xi]/(\xi^p)$. The map $L' \to L$, $\xi \mapsto 0$ yields an algebraic group homomorphism $A_m \to A$, and it is easy to see that the kernel is a unipotent group, hence of exponent p^N for some N. Let V_m be the algebraic subgroup of $p^N A_m$ of A_m. Then the homomorphism $A_m \to A$ has finite kernel when restricted to V_m. Thus $\dim(V_m) = \dim(A)$, and by our criterion, $H_{A,V}$ has finite Morley dimension.

Under ρ, V_m corresponds to the group of elements of $A(\mathbb{U}_m)$ divisible by p^N. Thus $H_{A,V_m} = \psi_q(V_m)$ is the group of elements of $A(\mathbb{U})$ divisible by p^N, and $H_{A,V}$ is the group of infinitely p-divisible points. We obtain therefore:

The group of infinitely p-divisible elements of $A(\mathbb{U})$ has finite Morley dimension

$\mathbb{U}_\sigma$ The dimension is finite if the projection $(Y \times \ldots \times Y^{\sigma^l}) \to (Y \times \ldots \times Y^{\sigma^l})$ is finite on V.

Remarks

a) Suppose $X = H_{X,V}$ and $\dim(X) = 1$. It can be shown then that there is a fixed integer m such that for any definable $Y \subset X^n$, $Y = H_{X,V_Y}$ for some V_Y with $\dim_{\text{Zariski}}(V') = m \dim_{\text{Morley}}(Y)$. V' can be taken to be the Zariski closure of $D_l Y$, or $\lambda_l Y$, or $\sigma_l Y$ in the appropriate jet or product space. This integer need not equal 1.

b) The above remark is one reason it is helpful to use the Morley dimension theory directly, rather than (as one presumably could) translate to Zariski dimension theory in the corresponding jet or product space and do the work there. Inside a finite Morley rank X, one can pick a one-dimensional Y and work with it, as a structure. The corresponding jet space situation, is an m-dimensional object, with the curious property that the dimensions of all varieties constructed from it are divisible by m. It could be difficult to tell in advance that Y has this property, or to argue that it holds at each stage. (This is reminiscent of the situation, with $m = 2$, between real and complex geometry.)

Proposition 3.1 *Zilber's conjecture holds for $\mathbb{U}$: every minimal subset of $\mathbb{U}$ is locally modular, or interprets a one-dimensional field*

Proof. In the case of $\mathbb{U}_D$ and $\mathbb{U}_p$, call the basic sets defined above *closed*. After possibly removing a singular locus, the jet-space varieties they correspond to can be taken to be non-singular. The dimension theorem of

algebraic geometry, for these jet-space varieties, can be shown to yield the dimension theorem for the minimal set. ([23] for $\mathbb{U}_D$, [19] for minimal sets in $\mathbb{U}_p$ satisfying the above finite dimensionality criterion; [11] for other minimal sets.) One then uses the theorem on Zariski geometries.

In the case of $\mathbb{U}_\sigma$, it is first necessary to show that a non-locally modular is *stable*. ([9]). At this point it may be possible to prove the dimension theorem for an appropriate Zariski geometry. In fact this course is not taken; instead a version of the following principle is used ([17]): if certain automorphisms groups of a minimal structure are unimodular as locally compact groups (equal left and right Haar measure), the structure must be locally modular. (See [17] and [9] for details.)   □

Let k be the constant field of $\mathbb{U}_D$, or the fixed field of $\mathbb{U}_\sigma$, or $k = \cap_r \mathbb{U}^{p^r}$ in the case of $\mathbb{U}_p$. In all cases, k can easily be shown to be minimal; it is certainly not locally modular, since it carries a field structure.

Theorem 3.2 *Every non-locally modular minimal set D in $\mathbb{U}$ is equivalent to k (in the sense of non-orthogonality).*

Proof. By the above proposition, D is equivalent to *some* one-dimensional field k'. It is a matter of showing that all interpretable one-dimensional fields are isomorphic. This can be viewed as a special case, or at least an allied problem, to that of classifying the interpretable simple groups. Model theoretic methods are quite useful here, and for instance show effortlessly that definable groups embed in algebraic groups, with information on the embedding. One must then rely on the theory of algebraic groups. The results are due to P. Cassidy (by other methods) for $\mathbb{U}_D$, and to Ph.D. theses by Sokolovic [38] and Mesmer [29] in the cases of $\mathbb{U}_D$ and $\mathbb{U}_p$.

In the case of $\mathbb{U}_\sigma$, the definable groups and simple groups can also be classified. Twisted groups do occur; in fact the list of definable simple groups is the same as that of non-sporadic finite simple groups, and thus includes for instance the unitary groups, and in characteristic 2 analogs of the Suzuki groups. In characteristic $p > 0$, twisted *fields* occur too - such as the fixed field of $x^p \sigma$. But in characteristic 0, it can be shown that the only definable field is the fixed field k of σ, and the finite extensions thereof.   □

Interesting algebraic information is obtained upon translating the above back to language of differential equations, or difference equations. We consider equations of the form

$$f(X, DX, D^2 X, \dots, D^n X) = 0 \quad \text{or} \quad f(X, \sigma X, \dots, \sigma^n X) = 0,$$

where D and σ are symbols for a derivation, or automorphism. Let us call a differential equation *geometrically degenerate* if the set of solutions, in the universal domain, forms a minimal set of disintegrated type (in the Zil'ber classification.)

Theorem 3.3 *Any algebraic differential equation can be reduced, by a differential-algebraic change of variable, to one of the following form:*

1. *A geometrically degenerate differential equation.*

2. *An equation defining a minimal subgroup of a simple, non-isotrivial Abelian variety. (Manin's equation).*

3. *The equation $Dx = 0$ defining the field of constants.*

Remarks

1. The change-of-variable is to be performed by a differential-algebraic polynomial , that may itself involve non-constant coefficients. Thus the theorem says nothing about linear equations; there is always a differential rational function (involving a basis of solutions as coefficients) reducing an arbitrary solution to a constant. There does exist a coefficient-free theory; in this case it amounts to a slight generalization of the Lie-Kolchin Galois theory of differential equations. (Poizat, Pillay).

2. In degree one (n=1), the description of geometrically degenerate equations can be sharpened. It can be shown, using methods of Jouanolou, that such an equation is equivalent to one with no structure at all; no two solutions are algebraically dependent, and only finitely many are algebraic over any given function field of finite transcendence degree. It is possible that the same holds in higher degrees; this may have applications to rational points over function fields, of varieties of arbitrary dimension. The proof presumably requires geometric methods; but the model theory expressed in the following theorem should give a guide, if only to suggest the correct statement of the theorem in higher dimensions.

3. Any non-isotrivial Abelian variety has a unique minimal infinite subgroup definable by an algebraic differential equation. It is minimal not only as a definable subgroup, but also as a definable set. Two such groups are equivalent iff their abelian varieties are isogenous; otherwise orthogonality holds. These facts can be proved purely model-theoretically (and indeed to my knowledge no other proof exists.) The equations were found by Manin;

see Manin's paper as well as [8] for a description. The theorem shows that
they have a priori significance among all algebraic differential equations, and
must be taken into account in a general theory.

4. An analogous theorem is available for difference equations. Alterna-
tive (1) looks the same; (3) refers to the fixed field; (2) is similar but here
Abelian varieties over number fields show up in place of non-isotrivial ones.
See [9], [17].

3.2   Points on subvarieties of Abelian varieties

The Mordell-Lang conjecture grew from Mordell's 1922 hypothesis on the
finiteness of the number of rational points on a curve of high genus (Falting's
theorem, 1983.) It has had a complicated history, well told in a number of
sources; see for instance [12], [26]. Briefly, according to the Mordell-Weil
theorem, the group of rational points of an Abelian variety is finitely gener-
ated. Since a curve may be embedded in its Jacobian, Mordell's conjecture
becomes equivalent to the following statement: a curve of genus ≥ 2 over
a number field intersects a finitely generated subgroup of its Jacobian in a
finite set.

Mordell's conjecture was transposed to function fields by Lang in 1960,
and in this version was proved by Manin (in characteristic 0) in 1963 (fol-
lowed by proofs by Grauert, and in positive characteristic by Samuel.) A
related conjecture arose, by Manin and Mumford: a curve of genus ≥ 2
intersects the set of torsion points of its Jacobian in a finite set. (Proved by
Raynaud, 1982).

This led Lang (1965) to a more general conjecture, replacing the Jaco-
bian by an Abelian variety, the curve by an arbitrary subvariety, and the
finitely generated, or torsion, group by the group of division points of a
finitely generated Γ; i.e. the points a such that $ma \in \Gamma$ for some positive
integer m. This group will be denoted by $\tilde{\Gamma}$. In positive characteristic p,
however, following [2], we let $\tilde{\Gamma} = \{g : mx \in \Gamma_0, \text{ some } m \text{ prime to } p\}$.
(The conjecture remains open for other m.)

Allowing a higher dimensional variety X in the hypothesis forces a mod-
ification in the conclusion: the intersection can no longer expected to be
finite, but is rather conjectured to be contained in a finite union of group
subvarieties, contained in X.

A further generalization was made to semi-Abelian varieties. See [26],
6.3. Our methods prove immediately the strongest "relative" version of the
conjecture, for semi-Abelian varieties. For simplicity we will state only a

weaker version and only for Abelian varieties.

Theorem 3.4 (Geometric Mordell-Lang) *Let A be a Abelian variety over an algebraically closed field K. Suppose A has no nontrivial homomorphic images defined over a finite field, or a number field. Let Γ be a finitely generated subgroup of $A(K)$, and let X be a subvariety of A. Then the Zariski closure of $X \cap \tilde{\Gamma}$ is a finite union of cosets of group subvarieties of A.*

In characteristic $p > 0$, the restriction against homomorphic images defined over finite extensions of the prime field is necessary. In characteristic 0, grace to Faltings and Vojta, it is not, and rather reflects a weakness of our present tools. We are presently unable to deal with the rational field, or with number fields of small degree. For function fields of characteristic 0, a number of proofs are available, including one by Buium ([8]) using differential algebra; Buium's proof, while quite different from ours beyond the first step, uses the Manin map in the same way, and it was seeing this that inspired the model theoretic proof. In characteristic $p > 0$, the model theoretic proof is the only one known in general; but many cases are proved in [2].

While unable to handle the general case over number fields, we can prove the Manin-Mumford conjecture; the proof has an explicit and uniform character, and appears to be the only proof to give an effective bound. (Though Hindry gives explicit bounds modulo one constant.) The proof includes the semi-Abelian case, as in [28] (according to McQuillan, with some improvements.)

For instance, we obtain:

Theorem 3.5 *Let A be a commutative algebraic group, defined over some number field, and fix a very ample divisor H on A. Let X be a subvariety, defined over K^{alg}. Then there exist a finite number M of subvarieties $c_i + A_i$ of X, all translates of group subvarieties A_i of A, such that*

$$T(A) \cap X = \cup_i T(A_i) + c_i.$$

We have $M \le c \deg_H(X)^e$, with c and e depending on A but not on X.

c and e can also be written down explicitly; they are doubly exponential in some natural parameters associated with A.

The method immediately yields a proof of the following theorem. It is a restricted version of a conjecture due to Silverman, Tate and Voloch. Let $\mathbb{C}_p$

be the completion of the algebraic closure of the p-adics $\mathbb{Q}_p$. The conjecture concerns Abelian and semi-Abelian varieties over $\mathbb{C}_p$, but we are only able to handle finite extensions of $\mathbb{Q}_p$ therein.

Theorem 3.6 *Let A be a semi- Abelian variety defined over $\mathbb{Q}_p$, or a finite extension of $\mathbb{Q}_p$, with good reduction. Let X be a subvariety of A, defined over $\mathbb{C}_p$. Then there exists a bound $\lambda > 0$ such that for any prime-to-p torsion point a of $A(\mathbb{C}_p)$, either $a \in X(\mathbb{C}_p)$, or the p-adic distance from a to X is $\geq \lambda$.*

Let us now discuss the proof of the above theorems. For definiteness we will concentrate on 3.4 in characteristic $p > 0$, where a separably closed field $\mathbb{U}_p$ is used. The characteristic 0 case uses is so close that one could easily formulate the two cases in a common language; it uses a differential field $\mathbb{U}_D$. The proofs of 3.5 and 3.6 have a similar structure but use $\mathbb{U}_\sigma$ instead.

Recall that k is the derivation in characteristic 0; it is $\cap_n K^{p^n}$ in characteristic p; and it is the fixed field of the automorphism, in the case of U_σ.

In all cases, one replaces the finitely generated group Γ (or the group of torsion points, in 3.5) by a larger group B, defined using the derivation or automorphism or inseparable structure. One shows that if the conclusion fails, then B is internal to k.

In the case of 3.4, one chooses k in advance to be linearly disjoint from the field of definition of A over the algebraic closure of the prime field. Then the isotriviality of B yields isotriviality of A, i.e. A will have homomorphic images defined over the algebraic closure of the prime field, contradicting the assumption.

In the case of 3.5, we are already over a number field, and "isotriviality" takes the following sense. There one uses an automorphism σ, with A defined over the fixed field k, and $B = \ker F(\sigma)$ for a certain polynomial F. The fact that B is internal to k turns out to mean that F is cyclotomic. However, one chooses σ to be a lifting of Frobenius, and F to be the characteristic polynomial of Frobenius on A; and it follows that F has no cyclotomic factors.

Having mentioned these differences, we restrict attention to characteristic $p > 0$.

Proof of 3.4. We can put ourselves in the following situation: $\mathbb{U}$ is the universal domain for separably closed fields of characteristic p, with $[\mathbb{U} : \mathbb{U}^p] = p$. $k = \cap_n K^{p^n}$ is the largest algebraically closed subfield. A is a

semi-Abelian variety defined over K; for simplicity, suppose A is an Abelian variety, with no nontrivial homomorphic factors defined over k. Γ is a finitely generated subgroup of $A(K)$, or the group of (prime-to-p)- division points of such a group. X is a K-subvariety of A. We want to show that the Zariski closure of $\Gamma \cap X$ is a finite union of cosets of group subvarieties of A.

Step 1. Let $B = p^\infty A(K) = \cap_r p^r A(K)$. Then B has finite Morley dimension. Moreover the minimal types within it satisfy Zilber's conjecture.
Proof. This was proved in §3.1.

Step 2. The minimal definable sets within B are locally modular.
Proof. By 3.2, if a non-locally-modular minimal set C occurs in B, it must be equivalent to k in the sense of non-orthogonality. Let C' be a group generated by C, as in 2.22. By 2.16, there exists an isogeny from C' to a to a definable group $H(k)$ over k. Taking essentially the Zariski closure of the isogeny, we obtain an isogeny from the Zariski closure A' of C' in A, to $H(K)$. Since H is defined over k, this contradicts our assumption on isotrivial factors.

Step 3. $X \cap socle(B)$ is a finite Boolean combination of cosets of definable subgroups of B.
Proof. Each minimal type in the socle is locally modular, and hence the socle is locally modular (2.18). Thus by 2.19, every definable subset of the socle is a finite Boolean combination of cosets definable subgroups of B.

Step 4. $X \cap B = Y \cap B$, where Y is a finite union of group subvarieties of A, contained in X. *Proof.* We can reduce to the case that the stabilizer of X is finite. Then by 2.25, X is contained in finitely many cosets of the socle of B. By Step (3), $X \cap B$ is a finite Boolean combination of cosets of (definable) subgroups of B. Let Y be the Zariski closure of $X \cap B$. One sees easily that Y is a finite union of group subvarieties.

Step 5. For some r, $X \cap p^r A(K) = Y \cap p^r A(K)$.
Proof. This follows by compactness from step (4): the intersection of all the groups $p^r A$ and of the definable set $X \setminus Y$ is empty, by (4); hence some finite intersection is empty.

In the same way, for any translate $X + b$ of X, there exists r such that the Zariski closure of $p^r A(K) \cap (X + b)$ is a finite union of translates of group subvarieties of A. Further, by compactness (1.2), one can choose the same r to work for all translates.

Now $\Gamma/p^r\Gamma$ is a finite group; so Γ is a union of finitely many translates of $\Gamma \cap p^r A(K)$, and hence lies in a finite union of cosets of $p^r A(K)$. The Zariski closure of the intersection of X with each of these cosets is a finite union of group subvarieties; hence so is the intersection of X with their union. Thus $X \cap \Gamma$ is contained in a finite union of group subvarieties contained in X, as promised.   □

1. For a fixed Abelian variety A, and X varying within a bounded algebraic family of subvarieties of A, the bound is uniform. Further, since Γ only enters the proof in the last step, via the number of cosets of $p^r A(K)$ it meets, the bound on the number of cosets of group subvarieties required has the form $cp^{r\gamma}$, where c and r do not depend on Γ, and γ is the "rank" of Γ, i.e. the maximal integer such that Γ contains a copy of $\mathbb{Z}^\gamma$.

2. In characteristic 0, (3.4 or 3.5), B is definable by a single formula (rather than infinitely many), and so the use of compactness in Step (5) above is not needed. As a consequence, one gets explicit (doubly exponential) bounds.

Bibliography

[1] J. Ax and S. Kochen, Amer. J. Math 87 (1965) 605-630, 631-648, Diophantine Problems over local fields I,II; Annals of Math. **83**, pp. 437 -456, Diophantine Problems over local fields III.

[2] Abramovich, D. and Voloch, J.F., Toward a proof of the Mordell-Lang conjecture in characteristic p, Duke Mathematical Journal 1992 No. 5 pp. 103-115.

[3] L. Henkin et al, eds., Proceedings of the Tarski Symposium, Proceedings of symposia in pure mathematics XXV, AMS, Providence 1974

[4] Chang, and Keisler, J., *Model Theory* , 3rd Edition, North -Holland, Amsterdam - Tokyo 1990

[5] , E. Bouscaren and E. Hrushovski, Interpreting groups in one-based theories *Journal of Symbolic Logic,* to appear

[6] S. Buechler, Locally modular theories of finite rank, Annals of Pure and Applied Logic **30** (1986) no. 1, 83-94

[7] A. Buium, Differential Function Fields and Modulii of Algebraic Varieties, LNM 1226, Springer 1986

[8] A. Buium, Intersections in jet spaces and a conjecture of S. Lang, Annals of Mathematics 136(1992), pp. 557-567.

[9] Z. Chatzidakis and E. Hrushovski, The model theory of difference equations, to appear.

[10] G. Cherlin, L. Harrington, and A. Lachlan, $\aleph_0$-categorical, $\aleph_0$-stable structures, *Ann. Pure Appl. Logic* **28** (1985) 103-135.

[11] F. Delon, work to appear on minimal types in separably closed fields.

[12] M. Hindry Sur les conjectures de Mordell et Lang (d'apres Vojta, Faltings et Bombieri), *Asterisque* 209 (1992)

[13] Lars Hörmander, *The Analysis of Linear Partial Differential Operators II*, Springer-Verlag Berlin - Heidelberg - New - York Tokyo 1983

[14] E. Hrushovski, Unidimensional Theories: an introduction to geometric stability theory, in *Logic Colloquium 88*, Granada, North Holland 1989.

[15] E. Hrushovski, Locally Modular Regular Types, *Classification Theory: Chicago 1985*, editor J.T. Baldwin, Springer Verlag 1987

[16] E. Hrushovski, A New Strongly Minimal Set,*Annals of Pure and Applied Logic***62** (1993) pp. 147-166.

[17] E. Hrushovski, the Manin-Mumford conjecture and the model theory of difference fields, preprint

[18] E. Hrushovski, unimodular minimal structures, London Math. Soc. Journal

[19] E. Hrushovski, E., The Mordell-Lang Conjecture for Function Fields, to appear in *Journal of the AMS*

[20] E. Hrushovski, J. Loveys, A structure theorem for strongly minimal sets, preprint

[21] E. Hrushovski and A. Pillay, Weakly Normal Groups, in *Logic Colloquium 85*, North Holland, Amsterdam 1986

[22] Hrushovski, E. & Pillay, A., Groups Definable in Local Fields and Pseudo-Finite Fields, *Israel Journal of Mathematics* **85** (1994) pp. 203-262

[23] Hrushovski, E. & Sokolovic, Z., Minimal subsets of differentially closed fields, to appear in *Transactions of the American Mathematical Society*

[24] E. Hrushovski and B. Zilber, Zariski Geometries, to appear in *Journal of the AMS*

[25] E. Hrushovski and B. Zilber, Zariski Geometries, AMS *Bulletin***28** (1993) no. 2 pp. 315-323

[26] S. Lang, *Number Theory III: Diophantine Geometry*, EMS 60, Springer-Verlag, Berlin-Heidelberg 1991

[27] Angus Macintyre, 20 years of p-adic model theory, Logic Coll. 84, pp. 121-153, N. Holland Amsterdam - NY 1986

[28] Michael McQuillan, Division points on semi-abelian varieties, *Inventiones mathematicae* 120 143-149 (1995)

[29] M. Mesmer, Groups and fields interpretable in separably closed fields, Ph.D. thesis, University of Illinois at Chicago, 1983

[30] Anand Pillay, Introduction to Stability Theory, Clarendon Press, Oxford 1983

[31] Anand Pillay, Geometric Stability Theory (book, forthcoming).

[32] J. Peterzil, S. Starchenko, preprint

[33] B. Poizat, Sous-groupes définissables d'un groupe stable, J. Symbolic Logic 46 (1981) 137-146.

[34] Henry H. Crapo and Gian Carlo Rota, On the Foundations of Combinatorial Theory: Combinatorial Geometries, MIT Press, Cambridge 1970

[35] Jean-Pierre Serre, *Lectures on the Mordell-Weil Theorem*, Aspects of Mathematics E 15, Vieweg, Braunschweig/Wiesbaden 1990.

[36] Saharon Shelah, *Classification Theory*, revised edition, North-Holland Amsterdam-Tokyo 1990

[37] C. Szmoryński, Logical Number Theory I, An Introduction, Springer Verlag Berlin-Heidelberg 1991

[38] Z. Sokolovic, Model Theory of Differential Fields, PhD Thesis, Notre Dame, July 1992

[39] C.R. Acad. Sci. A-B 262 (1966)

[40] B.I. Zilber, Strongly minimal countably categorical theories, Siberian Mathematical Journal 21 (1980), 219-230, II-III (1984), pp. 396-412, 559-571