Model theory and stability theory, with applications in differential algebra and algebraic geometry

Anand Pillay†
University of Leeds

This article is based around parts of the tutorial given by E. Bouscaren and A. Pillay at the training workshop at the Isaac Newton Institute, March 29 - April 8, 2005. The material is treated in an informal and free-ranging manner. We begin at an elementary level with an introduction to model theory for the non logician, but the level increases throughout, and towards the end of the article some familiarity with algebraic geometry is assumed. We will give some general references now rather than in the body of the article. For model theory, the beginnings of stability theory, and even material on differential fields, we recommend [5] and [8]. For more advanced stability theory, we recommend [6]. For the elements of algebraic geometry see [10], and for differential algebra see [2] and [9]. The material in section 5 is in the style of [7]. The volume [1] also has a self-contained exhaustive treatment of many of the topics discussed in the present article, such as stability, ω-stable groups, differential fields in all characteristics, algebraic geometry, and abelian varieties.

1 Model theory

From one point of view model theory operates at a somewhat naive level: that of point-sets, namely (definable) subsets X of a fixed universe M and its Cartesian powers $M \times \cdots \times M$. But some subtlety is introduced by the fact that the universe M is "movable", namely can be replaced by an elementary extension M', so a definable set should be thought of more as a functor.

† This work is supported by a Marie Curie Chair as well as NSF grants DMS-0300639 and the FRG DMS-0100979

Subtlety or no subtlety, model theory operates at a quite high level of generality.

A (1-*sorted*) structure M is simply a set (also called M) together with a fixed collection of distinguished relations (subsets of $M \times \cdots \times M$) and distinguished functions (from $M \times \cdots \times M$ to M). We always include the diagonal $\{(x,x) : x \in M\} \subset M \times M$ among the distinguished relations. (Example: Any group, ring, lattice,... is a structure under the natural choices for the distinguished relations/functions.) These distinguished relations/functions are sometimes called the *primitives* of the structure M. From the collection of primitives, one constructs using the operations composition, finite unions and intersections, complementation, Cartesian product, and projection, the class of $\emptyset$-*definable* sets and functions of the structure M. Let us call this class $Def_0(M)$, which should be seen as a natural "category" associated to the structure M: the objects of $Def_0(M)$ are the $\emptyset$-definable sets (certain subsets of $M \times \cdots \times M$) and the morphisms are $\emptyset$-definable functions $f : X \to Y$ (i.e. $graph(f)$ is $\emptyset$-definable). The category $Def(M)$ of definable (with parameters) sets in M is obtained from $Def_0(M)$ by allowing also fibres of $\emptyset$-definable functions as objects: if $f : X \to Y$ is in $Def_0(M)$ and $b \in Y$ then $f^{-1}(b)$ is a definable set (defined with parameter b). For A a subset of the (underlying set of M) $Def_A(M)$ denotes the category of definable sets in M which are *defined over* A, namely defined with parameter which is a tuple of elements of A. By convention, by a definable set we mean a set definably possibly with parameters. By a uniformly definable family of definable sets we mean the family of fibres of a definable map $f : X \to Y$.

We give a couple of examples.

The reals.
Consider the structure consisting of $\mathbb{R}$ with primitives $0, 1, +, -, \cdot$. Then the natural total ordering on $\mathbb{R}$ is a 0-definable set, being the projection on the first two coordinates of $\{(x, y, z) \in \mathbb{R}^3 : y - x = z^2 \text{ and } x \neq y\}$. Tarski's "quantifier elimination" theorem says that the definable sets in $(\mathbb{R}, 0, 1, +, -, \cdot)$ are precisely the *semialgebraic sets*, namely finite unions of subsets of $\mathbb{R}^n$ of the form

$$\{x \in \mathbb{R}^n : f(x) = 0 \text{ and } g_1(x) > 0 \text{ and } \ldots\ g_r(x) = 0\}$$

where f and the g_i are polynomials over $\mathbb{R}$.

Algebraically closed fields

Consider the field $\mathbb{C}$ of complex numbers. An (affine) algebraic variety is a subset $X \subseteq \mathbb{C}^n$ defined by a finite system of polynomial equations in n-variables and with coefficients from $\mathbb{C}$. If the equations have coefficients from $\mathbb{Q}$ we say that X is defined over $\mathbb{Q}$. A morphism between algebraic varieties X and Y is a map from X to Y given by a tuple of polynomial functions. Such a morphism is over $\mathbb{Q}$ if the polynomial functions have coefficients from $\mathbb{Q}$. View $\mathbb{C}$ as a structure with primitives $0, 1, +, -, \cdot$. Then it is a theorem (quantifier-elimination) that the category $Def_0(\mathbb{C})$ consists, up to Boolean combination, of the affine algebraic varieties defined over $\mathbb{Q}$ with morphisms defined over $\mathbb{Q}$. Likewise $Def(\mathbb{C})$ is (up to Boolean combination) just the category of algebraic varieties and morphisms. Everything we have said applies with any algebraically closed field K in place of $\mathbb{C}$ and with the prime field in place of $\mathbb{Q}$.

Given a structure M, the *language* or *signature* $L = L(M)$ of M is an indexing of the primitives, or rather a collection of (relation/ function) symbols corresponding to the primitives of M. We call M an L-structure or structure for the signature L. There is a natural notion of an L-structure M being a substructure or extension of an L-structure N (generalizing the notions subgroup, subring, ...). But somewhat more crucial notions for model theory are those of *elementary substructure* and *elementary extension*. We may take the Tarski-Vaught criterion as a definition: So assume that M, N are L-structures and M is a substructure of N (notationally $M \subseteq N$). Then M is an elementary substructure of N if whenever $X \subseteq N^n$, $X \in Def_M(N)$, and $X \neq \emptyset$, then $X \cap M^n \neq \emptyset$.

It is usual to begin by introducing first order formulas and sentences of L, define the notion of their satisfaction/truth in L-structures, and develop the rest of the theory afterwards. So the first order formulas of L are built up in a syntactically correct way, with the aid of parentheses, from primitive formulas $R(x_1, \ldots, x_n)$, $f(x_1, \ldots, x_n) = y$ (where R, f are relation/function symbols of L and $x_1, \ldots, x_n, y$ are "variables" or "indeterminates") using $\neg, \wedge, \vee$, and quantifiers $\exists x$, $\forall x$. Among the L-formulas are those with no unquantified variables. These are called L-sentences. An L-formula with unquantified variables $x_1, \ldots, x_n$ is often written as $\phi(x_1, \ldots, x_n)$. Given an L-structure M and L-sentence σ there is a natural notion of "σ is true in M" which is written $M \models \sigma$, and for $\phi(x_1, \ldots, x_n)$ an L-formula, and $b = (b_1, \ldots, b_n) \in M^n$, a natural notion of "$\phi$ is true of b in M", written $M \models \phi(b)$.

As the logical operations $\neg, \vee, \ldots$ correspond to complementation, union, ... we see that for M an L-structure the 0-definable sets of M come from L-formulas: if $\phi(x_1, \ldots, x_n)$ is an L-formula, then $\{b \in M^n : M \models \phi(b)\}$ is a 0-definable set in M and all 0-definable sets of M occur this way. Depending on one's taste, the syntactic approach may be more easily understandable. For example if M is a group $(G, \cdot)$, then the centre of G is a 0-definable set, defined by the formula $\forall y(x \cdot y = y \cdot x)$.

Likewise the definable sets in M are given by L-formulas with parameters from M.

With this formalism M is an elementary substructure of N (N is an elementary extension of M) if $M \subseteq N$ are both L-structures and for each L-formula $\phi(x_1, \ldots, x_n)$ and tuple $b = (b_1, \ldots, b_n) \in M^n$, $M \models \phi(b)$ iff $N \models \phi(b)$.

The compactness theorem of first order logic gives rise to elementary extensions of M of arbitrarily large cardinality, as long as (the underlying set of) M is infinite. Such an elementary extension N could be considered as some kind of "nonstandard" extension of M, in which all things true in M remain true. If X is a definable set in M then X has a canonical extension, say $X(N)$ to a definable set in N (in fact $X(N)$ is just defined in the structure N by the same formula which defines X in M). The usefulness of passing to an elementary extension N of a structure M is that we can find such elementary extensions with lots of symmetries (automorphisms) and "homogeneity" properties. Such models play the role of Weil's universal domains in algebraic geometry (and Kolchin's universal differential fields in differential algebraic geometry). The relative unfashionability of such objects in modern algebraic geometry is sometimes an obstacle to the grasp of what is otherwise the considerably naive point of view of model theory. Another advantage of such nonstandard models is that uniformly definable families of definable sets have explicit "generic fibres".

Given a cardinal κ, a structure M is called *κ-compact* if whenever $\{X_i : i \in I\}$ is a collection of definable subsets of M with the finite intersection property, and $|I| < \kappa$, then $\cap_{i \in I} X_i \neq \emptyset$.

Under some mild set-theoretic assumptions, any structure M has κ-compact elementary extensions of cardinality κ for sufficiently large cardinals κ. There is a related notion, *κ-saturation*: M is said to be κ-saturated if for any subset of M of cardinality $< \kappa$ any collection of A-definable subsets of M, which has the finite intersection property, has nonempty intersection. For κ strictly greater than the cardinality of L (number of L-formulas), κ-saturation coincides with κ-compactness.

Let κ be any uncountable cardinal. Then any algebraically closed field K of cardinality κ is κ-compact. Moreover let $\lambda < \kappa$, let $K[x_i : i < \lambda]$ be the polynomial ring in λ unknowns over K, and let S be any proper ideal of this ring, there is a common zero $(a_i)_{i<\lambda}$ of S whose coordinates lie in K.

We finish this section with some additional notation, conventions, and examples (aimed at the nonlogician).

Fix a language L. An L-*theory* is a set Σ of L-sentences which has a model. If σ is an L-sentence we write $\Sigma \models \sigma$ to mean that every model of Σ is a model of σ. The L-theory Σ is said to be *complete* if for every L-sentence σ either $\Sigma \models \sigma$ or $\Sigma \models \neg\sigma$.

A complete theory is often denoted by T. If T is such then we are interested in models of T and definable sets in such models. It has been a convention to choose a κ-compact model $\bar{M}$ of T of cardinality κ for some large κ. Then every model of T of cardinality $< \kappa$ is (up to isomorphism) an elementary substructure of $\bar{M}$. So when we speak of a model of T we refer to a small (cardinality $< \kappa$) elementary substructure of $\bar{M}$. We use $A, B, \ldots$ to denote small subsets of (the underlying set of) $\bar{M}$.

Complete types play an important role in model theory, especially in stability theory. If a is a tuple (usually finite) from $\bar{M}$ and A a subset of $\bar{M}$ then $tp(a/A)$ denotes the set of formulas $\phi(x)$ with parameters from A which are true of a in $\bar{M}$. Working rather with definable sets, $tp(a/A)$ can be identified with the collection of A-definable subsets of $\bar{M}^n$ which contain the point a, and is an ultrafilter on the set of A-definable subsets of $\bar{M}^n$. Then tuples a and b have the same type over A if there is an automorphism of $\bar{M}$ fixing A pointwise and taking a to b.

A basic example of a complete theory is ACF_p the theory of algebraically closed fields of characteristic p (where p is a prime, or is 0). The language here is the language of rings $(0, 1, +, -, \cdot)$. (Note that it is easy to write down first order sentences giving the axioms.)

A theory Σ (complete or not) in language L is said to have *quantifier-elimination* if every L-formula $\phi(x_1, \ldots, x_n)$ is equivalent (in models of Σ) to a quantifier-free L-formula. For example the (incomplete) theory ACF has QE. For specific theories it is important to have some kind of quantifier-elimination or relative quantifier elimination theorem so as to understand to some extent definable sets. But as far as the general theory of definability goes one can always assume quantifier-elimination by expanding the language by new relation symbols $R_\phi(x_1, \ldots, x_n)$ for each formula $\phi(x_1, \ldots, x_n)$.

The origin of stability theory was the (abstract) study of theories T which are *uncountably categorical*. The convention here is that the language L is countable, and then T is said to be uncountably categorical if for every (equivalently some, by Morley's theorem) uncountable λ, T has exactly one model of cardinality λ up to isomorphism. ACF_p is uncountably categorical.

Differentially closed fields.
An important example in this series of talks will be *differentially closed fields*. The relevant complete theory is DCF_0. The language here is the language of differential rings, namely the language of rings together with a new unary function symbol ∂. The axioms are the axioms for fields of characteristic 0 with a derivation ∂ (∂ is an additive homomorphism, and $\partial(x \cdot y) = \partial(x) \cdot y + x \cdot \partial(y)$), together with axioms which state that any finite system of differential polynomial equations and inequations with parameters (in finitely many indeterminates) which has a solution in differential field extension already has a solution in the model in question. It is a nontrivial fact that one can find such axioms, and in fact there are much simpler axioms (referring just to single differential equations in one indeterminate) which suffice, as shown by Blum. The theorem is that DCF_0 is complete and has quantifier-elimination.

Other important complete theories with interest are the theory of separably closed fields of characteristic p with Ershov invariant e (meaning that the dimension of K over K^p is p^e), and the theory of nontrivially valued algebraically closed fields of a given pair of characteristics. The latter is an important first order context for dealing with "infinitesimals".

Many-sorted structures
It is natural to consider many-sorted structures and theories in place of one-sorted ones. In this more general context, a structure M will be a family $(M_s : s \in S)$ of universes. The primitive relations and functions will be be on and between Cartesian products of universes. The language L of the structure will include the set of sorts S. Moreover any variable x comes equipped with a specific sort, and thus quantifiers will range over designated sorts. The whole machinery of first order logic (elementary

extensions, saturation, formulas, theories) generalizes without difficulty to the many-sorted context. In fact this fits in very well with the above-mentioned notion of the category of definable sets in a given structure. For example, given a one-sorted structure M, form a new many-sorted structure $Def_0(M)$ whose sorts are the $\emptyset$-definable sets of M and whose relations are those induced by $\emptyset$-definable sets in M. For example, if X and Y are $\emptyset$-definable sets of M, and $R \subseteq X \times Y$ is a $\emptyset$-definable relation of M, then we have a corresponding basic relation in $Def_0(M)$ between the sort of X and that sort of Y. (Note that $Def_0(M)$ will automatically have quantifier-elimination in this presentation.) In fact one can go further: we can consider not only $\emptyset$-definable sets in M but also quotients of such by $\emptyset$-definable equivalence relations. So we take as sorts all sets of the form X/E where X is $\emptyset$-definable in M and E is an $\emptyset$-definable equivalence relation on X. Again we take as relations things induced by $\emptyset$-definable relations on M. Note that among the new basic functions will be the canonical surjections $X \to X/E$. We call this new many-sorted structure M^{eq}. The point is that M^{eq} is the "same" as M. (The technical term for sameness here is "bi-interpretable")

A typical example is obtained when we start with ACF_0 say and form the category of algebraic varieties defined over $\mathbb{Q}$ (again with the induced structure). We call this many-sorted structure AG_0 (algebraic geometry in characteristic 0).

A somewhat richer structure is the many-sorted structure $\mathcal{A}$ whose sorts are compact complex analytic spaces (up to biholomorphism) and whose relations are analytic subvarieties of (finite) cartesian products of sorts. More details will be given in a subsequent paper in the volume. The structure $\mathcal{A}$ is NOT κ-saturated for any cardinal κ. This is because every element of every sort is essentially named by a constant. We let CCM denote the first order theory of $\mathcal{A}$. Among the sorts in $\mathcal{A}$ are the projective algebraic varieties, and in this way AG_0 can be seen as a "subcategory" of CCM. CCM has quantifier elimination.

In many cases, it is not necessary to pass to M^{eq} in that the quotient sets are already present in M. This is when M (or $Th(M)$) has so-called elimination of imaginaries. So the structure M is said to have *elimination of imaginaries* if whenever X and E are $\emptyset$-definable sets in M ($X \subset M^n$ say and E an equivalence relation on X) then there is another $\emptyset$-definable $Y \subset M^k$ and a $\emptyset$-definable surjective function from X to Y such that $f(x_1) = f(x_2)$ iff $E(x_1, x_2)$.

ACF_p, DCF_0, CCM and $SCF_{p,e}$ all eliminate imaginaries (the latter after naming a p-basis).

The notions of algebraic and definable closure are important in what follows: Given a possibly many-sorted structure which eliminates imaginaries (for example M^{eq}), and a subset A of M, $dcl(A) = \{f(a) : a$ a finite tuple from A and f a $\emptyset$-definable function$\}$. We let $acl(A)$ denote the union of all finite A-definable sets. (For a structure which does not necessarily eliminate imaginaries, we have described what are usually called $dcl^{eq}(A), acl^{eq}(A)$.)

2 Stability

We will fix a many-sorted structure $M = (M_s : s \in S)$, which we assume to be saturated (κ-saturated of cardinality κ for some large κ) for convenience. We also assume that M has 'elimination of imaginaries. Let $X, Y, \ldots$ denote definable sets, and $A, B, \ldots$ sets of parameters.

What kind of relationships between definable sets can be formulated at this general level?

Definition 2.1 (i) X and Y are *fully orthogonal* if every definable $Z \subseteq X^n \times Y^m$ is up to Boolean combination of the form $Z_1 \times Z_2$.
(ii) At the opposite extreme: X is *internal* to Y if there is a definable surjective map f from Y^n to X (for some n).

A naive example of full orthogonality is the case where M consists of two infinite sorts M_0, M_1 with no additional relations. Put $X = M_0$ and $Y = M_1$. A rather trivial example of X being internal to Y is when $X = Y^n$ for some n. A more interesting example is when Y is equipped with a definable group structure, and there is a definable strictly transitive action of Y on X. Then the choice of a point $x \in X$ yields a definable bijection between Y and X.

A slight weakening of internality is *almost internality* where the map f above is replaced by a definable relation $R \subset Y^n \times X$ such that for any $x \in X$, there are only finitely many, but at least one, $y \in Y^n$ such that $R(y, x)$. In any case internality is a fundamental model-theoretic notion. The subtlety is that X and Y may be $\emptyset$-definable, and X may be internal to Y but only witnessed by a definable function defined with additional parameters. In such a situation there will be an associated nontrivial Galois group arising: a definable group naturally isomorphic to the group of permutations of X induced by automorphisms of M which fix Y pointwise.

Note that if X is finite then X is fully orthogonal to any Y. In AG

orthogonality is vacuous. In fact in AG if X, Y are infinite definable sets, then each is almost internal in the other.

Stability is an assumption on M (or on $Th(M)$) which played a very large role in Shelah's classification theory program, but also has many consequences for the structure of definable sets. For stable structures the study of *complete types* plays an important role.

Definition 2.2 M is *stable* if there is no definable relation $R(x, y)$ and a_i, b_i in M for $i < \omega$ such that for all $i, j < \omega$, $R(a_i, b_j)$ iff $i < j$.

As the ordering on $\mathbb{R}$ in the structure $(\mathbb{R}, <, +, \cdot)$ is definable we see that this structure is unstable. On the other hand AG is stable.

Independence (also called nondividing, nonforking).
Under the assumption of stability, a notion of *freeness* can be developed, giving meaning to "a is independent (free) from B over A" where $A \subset B$ are sets of parameters and a is a finite tuple of elements of M. In the case of AG (as a 1-sorted structure), assuming $A \subset B$ are subfields $F_1 < F_2$ of K this will mean precisely that $tr.deg(F_1(a)/F_1) = tr.deg(F_2(a)/F_2)$.

The precise definition depends on the notion of *indiscernibility*: a sequence $(b_i : i \in \omega)$ of tuples b_i of the same length is said to be indiscernible over a set A if for all n, $tp(b_{i_1}, \ldots, b_{i_n}/A) = tp(b_{j_1}, \ldots, b_{j_n}/A)$ whenever $i_1 < \cdots < i_n$ and $j_1 < \cdots < j_n$.

Definition 2.3 Let a, b be possibly infinite tuples, and A a set of parameters. We say that $p(x, b) = tp(a/A, b)$ *divides* over A if there is an A-indiscernible sequence $(b_i : i < \omega)$ with $b_0 = b$ such that $\{p(x, b_i) : i < \omega\}$ is inconsistent (not realized in M).

For T stable (or more generally "simple") nondividing is our notion of freeness and it has good properties: so a is free from b over A if $tp(a/A, b)$ does not divide over A, and we have properties such as
symmetry: a is free from b over A iff b is free from a over A;
free extensions: for every a, A and b there is a' with $tp(a'/A) = tp(a/A)$ and a' is free from b over A;
small bases: for any finite tuple a and set A, there is $A_0 \subseteq A$ of cardinality $\leq |L|$ such that a is free from A over A_0.

Stationarity. A characteristic property of independence in stable theories is "uniqueness of generic types" or "uniqueness of free extensions"

Fact 2.4 *Assume M stable. Let $A \subseteq B \subset M$ be sets of parameters. Assume A is algebraically closed. Let a_1, a_2 be tuples such that $tp(a_1/A) = tp(a_2/A)$ and each of a_1, a_2 is independent from B over A. Then $tp(a_1/B) = tp(a_2/B)$.*

We express the above fact by saying that complete types over algebraically closed sets are *stationary*.

In the case of AG, any stationary type is the "generic" type of an (absolutely) irreducible variety: Fact 2.4 says that if V is an irreducible variety over F, and a_1, a_2 are generic points of V over F then there is an automorphism of K fixing F pointwise and taking a_1 to a_2.

The notion of a "general" or "generic" point of a definable set may make sense in many contexts, especially where there is a notion of "dimension" for definable sets. However in the case of stable theories Fact 2.4 leads to an independence-theoretic characterization of full orthogonality:

Lemma 2.5 *(T stable.) Let X, Y be $\emptyset$-definable sets. Then X is fully orthogonal to Y iff for any set A of parameters, and $a \in X$ and $b \in Y$, a is independent from b over A.*

There is a notion of orthogonality for stationary types: p and q are orthogonal iff for any set A of parameters including the domains of p and q, and a realizing p independent from A over $dom(p)$ and and b realizing q independent from A over $dom(q)$ then a is independent from b over A. So the lemma above can be restated as: X and Y are fully orthogonal if and only if for all complete stationary types $p(x)$ containing $x \in X$ and $q(y)$ containing $y \in Y$, p is orthogonal to q.

Generalizing the notion of smallest field of definition of an algebraic variety, is the notion of the canonical base of a stationary type:

Fact 2.6 *(T stable.) Assume $tp(a/A)$ is stationary. Then there is smallest $A_0 \subset A$ such that a is independent from A over A_0 and $tp(a/A_0)$ is stationary. A_0 is called the canonical base of p.*

Morley rank and t.t. theories.

The notions of stability theory are a little clearer for so-called totally transcendental (t.t.) theories. T is said to be *t.t* if every definable set has an ordinal valued Morley rank.

Again work in a possibly many sorted saturated structure M. Let X