Model theory, VC classes, and Henselian rings

Will Johnson

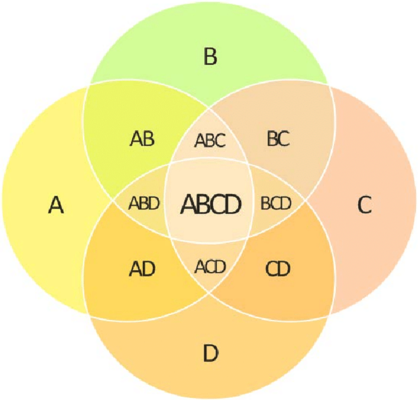
Fudan University

November 18, 2024

Section 1

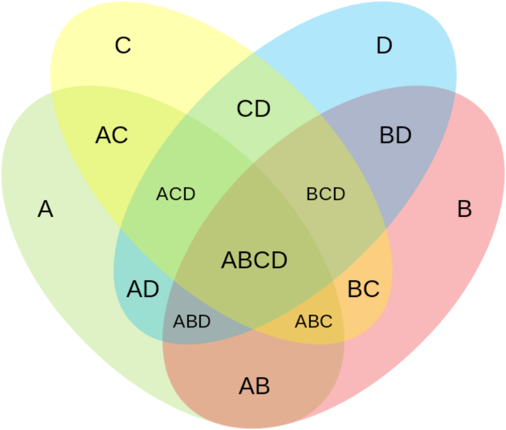
NIP

A Venn diagram with four sets

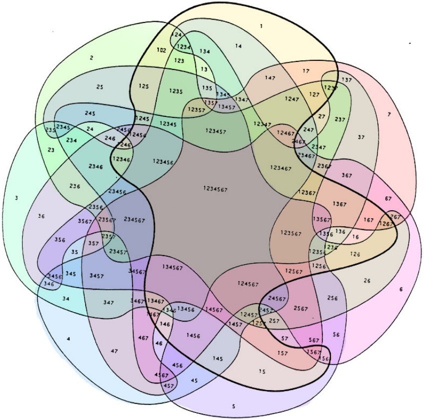


What's wrong with this picture?

A Venn diagram with four sets



A Venn diagram with seven sets



Slogan

In a Venn diagram of n sets, the complexity of the sets increases with n

Slogan

Slogan

In a Venn diagram of n sets, the complexity of the sets increases with n

- What is a "Venn diagram"?
- How is "complexity" measured?

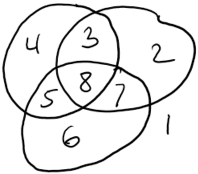
Independent sets

Fix a set U.

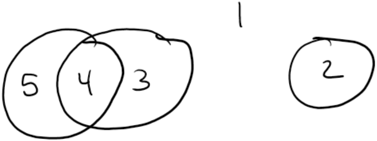
Definition

Sets $X_1, \ldots, X_n \subseteq U$ are *independent* if they partition U into 2^n subsets. More precisely, for every $S \subseteq \{1, \ldots, n\}$,

$$\bigcap_{\substack{1 \le i \le n \\ i \in S}} X_i \setminus \bigcup_{\substack{1 \le i \le n \\ i \notin S}} X_i \neq \varnothing.$$



Three independent sets.



Three non-independent sets.

Slogan

Slogan

In a Venn diagram of n sets, the complexity of the sets increases with n

Slogan

If $X_1, \ldots, X_n \subseteq \mathbb{R}^2$ are independent sets, then the complexity of the sets $X_1, \ldots, X_n$ increases with n.

- How is "complexity" measured?

Definable sets

For the structure $(\mathbb{R}, +, \cdot, \leq)\ldots$

Definition

$D \subseteq \mathbb{R}^n$ is a *definable set* if

$$D = \{\vec{x} \in \mathbb{R}^n : \varphi(\vec{x})\}$$

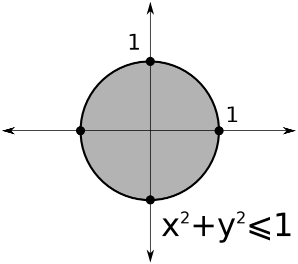
for some first-order formula $\varphi(\vec{x})$.

"First-order":

- $+$, $\cdot$, $\leq$, $=$.
- $\wedge$ (and), $\vee$ (or), $\neg$ (not)
- $\forall x \in \mathbb{R}$, $\exists x \in \mathbb{R}$
- ~~$\forall S \subseteq \mathbb{R}$, $\exists S \subseteq \mathbb{R}$~~

Definable sets

The unit disk is definable in $(\mathbb{R}, +, \cdot, \leq)$:



$$\{(x, y) \in \mathbb{R}^2 \mid \exists z : x \cdot x + y \cdot y + z \cdot z = 1\}$$
$$= \{(x, y) \in \mathbb{R}^2 \mid x \cdot x + y \cdot y \leq 1\}.$$

Definable families

Definition

$\mathcal{F} \subseteq \mathrm{Pow}(\mathbb{R}^n)$ is a *definable family* if

$$\mathcal{F} = \{\{\vec{x} \in \mathbb{R}^n : \varphi(\vec{x}, \vec{a})\} : \vec{a} \in \mathbb{R}^n,\ \psi(\vec{a})\}$$

for some first-order formulas φ, ψ.

Example

The family of open disks in $\mathbb{R}^2$ is a definable family:

$$\{\{(x, y) \in \mathbb{R}^2 : (x - a)^2 + (y - b)^2 < r^2\} : (a, b, r) \in \mathbb{R}^3,\ r > 0\}$$

Idea

In a definable family, the sets have bounded complexity.

A precise statement

Slogan

In a Venn diagram of n sets, the complexity of the sets increases with n

Theorem

If $\mathcal{F}$ is a definable family in the structure $(\mathbb{R}, +, \cdot)$, then

$$\sup\{n : \textit{there are independent } X_1, \ldots, X_n \in \mathcal{F}\} < \infty.$$

Example

If $X_1, \ldots, X_n$ are independent disks, then $n \leq 3$.

IP and NIP

Let M be a structure.

Definition

M has the *independence property* (*IP*) if there is a definable family $\mathcal{F}$ such that

$$\sup\{n : \text{there are independent } X_1, \ldots, X_n \in \mathcal{F}\} = \infty.$$

Definition

M is *NIP* if for every definable family $\mathcal{F}$,

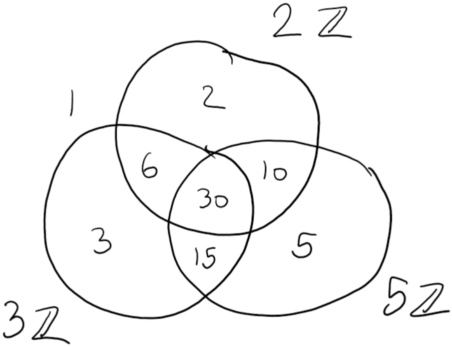
$$\sup\{n : \text{there are independent } X_1, \ldots, X_n \in \mathcal{F}\} < \infty.$$

IP and NIP

Example

$(\mathbb{Z}, +, \cdot)$ has the IP:

- $2\mathbb{Z}, 3\mathbb{Z}, 5\mathbb{Z}, 7\mathbb{Z}, \ldots$ are independent.
- $\{n\mathbb{Z} : n \in \mathbb{Z}\}$ is a definable family.



Theorem (Wilkie)

The real exponential field $(\mathbb{R}, +, \cdot, \leq, \exp)$ *is NIP.*

Motivating questions

Question

Which fields are NIP?

Question

Which rings are NIP?

From now on, “ring” means “commutative unital ring”.

Section 2

A detour to statistics

VC classes

Definition

A family of sets $\mathcal{F}$ is a *VC class* if

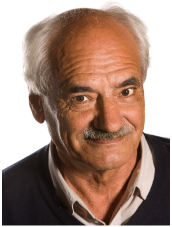
$$\sup\{n : \text{there exist independent } X_1, \ldots, X_n \in \mathcal{F}\} < \infty.$$

A structure is NIP iff every definable family is a VC class.

VC classes

What does “VC” stand for?

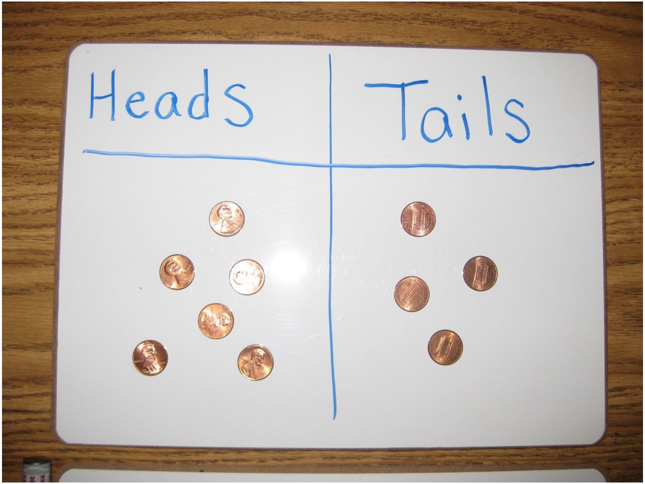
Vapnik



Chervonenkis

These are statisticians, not logicians!

The law of large numbers



The law of large numbers

Let (Ω, μ) be a probability space. Fix $\epsilon > 0$.

Theorem

If $X_1, \ldots, X_n$ are independently distributed according to μ, and $E \subseteq \Omega$, then

$$\lim_{n \to \infty} \text{Prob}\left(\left| \mu(E) - \frac{\#\{i : X_i \in E\}}{n} \right| < \epsilon \right) = 1$$

Idea

If $X_1, \ldots, X_n$ are random samples from (Ω, μ), and $n \gg 0$, then with high probability,

$$\mu(E) \approx \frac{\#\{i : X_i \in E\}}{n}.$$

Uniform law of large numbers

Let (Ω, μ) be a probability space and $\mathcal{F}$ be a VC class on Ω. Fix $\epsilon > 0$.

Theorem (Vapnik-Chervonenkis)

If $X_1, \ldots, X_n$ are independently distributed according to μ, then

$$\lim_{n \to \infty} \text{Prob}\left(\sup_{E \in \mathcal{F}} \left| \mu(E) - \frac{\#\{i : X_i \in E\}}{n} \right| < \epsilon \right) = 1.$$

Idea

If $X_1, \ldots, X_n$ are random samples from (Ω, μ), and $n \gg 0$, then

$$\mu(E) \approx \frac{\#\{i : X_i \in E\}}{n}$$

for *every* $E \in \mathcal{F}$, with high probability.

Uniform law of large numbers

Idea

If $X_1, \ldots, X_n$ are random samples from (Ω, μ), and $n \gg 0$, then

$$\mu(E) \approx \frac{\#\{i : X_i \in E\}}{n}$$

for *every* $E \in \mathcal{F}$, with high probability.

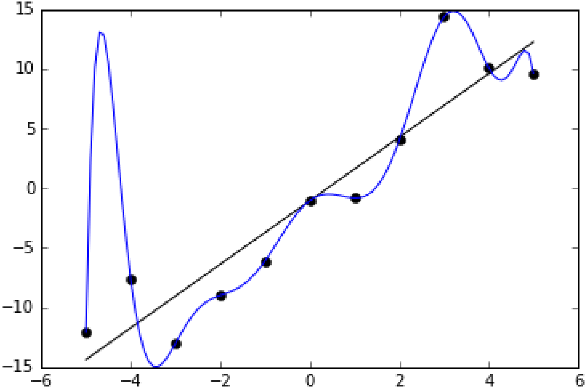
Why is this non-trivial?

- A random set of numbers between 0 and 99:

$$82, 26, 83, 31, 9, 29, 89, 5, 8, 91$$

- If $E = \{82, 26, 83, \ldots, 91\}$, then $\mu(E) = 0.1$, but the sample suggests $\mu(E) \approx 1$

Overfitting

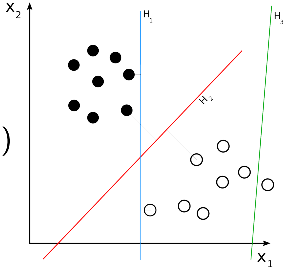


Overfitting and VC classes

- Suppose we are training a classifier $f : \Omega \to \{0, 1\}$.
- Let $\mathcal{F}$ be the set of possibilities for f.
- If $\mathcal{F}$ is a VC class, then

  $$\text{sample accuracy}(f) \approx \text{population accuracy}(f)$$

  for *all* $f \in \mathcal{F}$
- ...so maximizing sample accuracy *nearly* maximizes population accuracy.
  - No overfitting!



Overfitting and NIP

Theorem (Wilkie)

The real exponential field $(\mathbb{R}, +, \cdot, \exp)$ *is NIP.*

Corollary

In a sigmoidal neural network, the set of possible classifiers $\mathbb{R}^n \to \{0, 1\}$ *is a VC class.*

Karpinski and Macintyre calculate more precise bounds for the VC theorem, in this case.

Section 3

Model theory

Motivating questions

Question

Which fields are NIP?

Question

Which rings are NIP?

These questions belong to *model theory*.

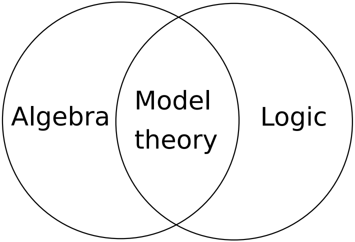
Model theory

Model theory is the study of *algebraic structures*...

- groups, rings, fields,...

...using tools from mathematical logic:

- definable sets, elementary equivalence,...



Structures

Definition

A *structure* is a set with some functions and relations.

Examples:

- $(\mathbb{C}, +, \cdot)$
- $(\mathbb{Z}, +, \leq)$
- $(\mathbb{R}, +, \cdot, \exp)$

Elementary equivalence

Definition

Two structures M and N are *elementarily equivalent* ($M \equiv N$) if M and N satisfy the same first-order sentences.

$$\mathbb{R} \text{ satisfies } \forall x\ \exists y : y \cdot y \cdot y = x$$
$$\mathbb{Q} \text{ doesn't satisfy } \forall x\ \exists y : y \cdot y \cdot y = x$$

So $\mathbb{R} \not\equiv \mathbb{Q}$.

Elementary equivalence

Any structure M has an *elementary equivalence class*

$$\{N : N \equiv M\}.$$

Theorem

$(K, +, \cdot) \equiv (\mathbb{C}, +, \cdot)$ *if and only if* $\mathrm{char}(K) = 0$ *and* $K = K^{\mathrm{alg}}$.

A form of the Lefschetz principle in algebraic geometry(?)

Example

$\mathbb{Q}^{\mathrm{alg}} \equiv \mathbb{C}$.

Definable sets and functions

Fix a structure M.

Definition

$D \subseteq M^n$ is *definable* if

$$D = \{\vec{x} : \varphi(\vec{x})\}$$

for some first-order formula φ.

Definition

If X, Y are definable, then $f : X \to Y$ is a *definable function* if the graph $\Gamma(f)$ is definable:

$$\Gamma(f) = \{(x, y) \in X \times Y : y = f(x)\}.$$

Definable sets and functions form a category.

Definable sets in $(\mathbb{C}, +, \cdot)$

Theorem

$D \subseteq \mathbb{C}^n$ is definable if and only if D is constructible in the sense of algebraic geometry:

$$D = \bigcup_{i=1}^{m} C_i \setminus C_i'$$

for Zariski closed sets C_i, C_i'.

A form of Chevalley's theorem in algebraic geometry(?)

Theorem

In $\mathbb{C}$, a function $f : X \to Y$ is definable iff f is piecewise rational.

What can model theory do?

Categoricity

Let M be an infinite structure and κ be an infinite cardinal.

Theorem (Löwenheim-Skolem)

M is elementarily equivalent to a structure of size κ.

Definition

M is *κ-categorical* if there's a unique $N \equiv M$ of size κ, up to isomorphism.

Example

$\mathbb{C}$ is κ-categorical for any $\kappa > \aleph_1$.

(There's only one K with $K = K^{\text{alg}}$, $\text{char}(K) = 0$, and $\text{tr.deg}(K/\mathbb{Q}) = \kappa$.)

Uncountable categoricity

Suppose M is κ-categorical for some $\kappa > \aleph_0$.

- Every definable set $D \subseteq M^n$ has a "dimension" $\dim(D) \in \mathbb{N}$, satisfying things like $\dim(X \times Y) = \dim(X) + \dim(Y)$.
- If K is a definable field, then K is finite or $K = K^{\mathrm{alg}}$.
- If G is a definable group...
  - G has a "connected component" G^0.
  - There is a subnormal series of definable subgroups

    $$1 = H_0 \lhd H_1 \lhd H_2 \lhd \cdots \lhd H_n = G$$

    such that the quotients H_{i+1}/H_i are simple or abelian.
  - Conjecturally, the simple H_{i+1}/H_i are algebraic groups over algebraically closed fields (Cherlin-Zilber).

Section 4

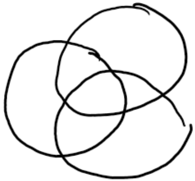
Stability theory and Neostability theory

NIP and NSOP

Let M be a structure.

Definition

M has the *IP* (*independence property*) if there's a definable family $\mathcal{F}$ such that for every $n < \infty$, there are independent $X_1, \ldots, X_n \in \mathcal{F}$.
Otherwise, M is *NIP*.

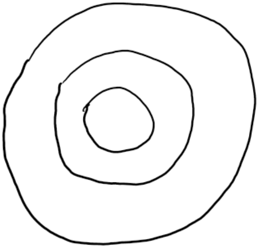


Definition

M has the *SOP* (*strict order property*) if there's a definable family $\mathcal{F}$ such that for every $n < \infty$, there are $X_1, \ldots, X_n \in \mathcal{F}$ with

$$X_1 \subsetneq X_2 \subsetneq X_3 \subsetneq \cdots$$

Otherwise, M is *NSOP*.



NIP and NSOP

Theorem

$(\mathbb{R}, +, \cdot, \leq)$ *is NIP.*

On the other hand,

Remark

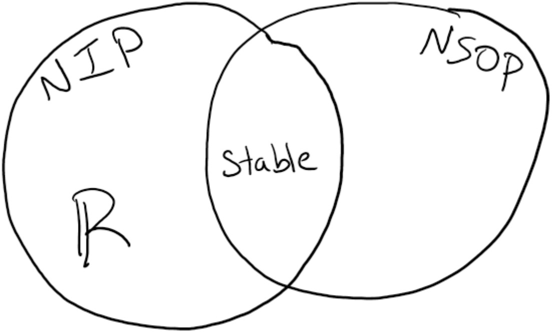
$(\mathbb{R}, +, \cdot, \leq)$ has the SOP:

$$(-\infty, 1) \subsetneq (-\infty, 2) \subsetneq (-\infty, 3) \subsetneq \cdots$$

Stability

Definition

A structure M is *stable* if M is NIP and NSOP.



Stability: what good is it?

These structures are stable:

- Algebraically closed fields.
- Free groups (Sela).
- Abelian groups, modules.
- Differentially closed fields $(K, +, \cdot, \partial)$.
- Uncountably categorical structures.

In fact,

Theorem (Shelah)

Unless M is stable, for every $\kappa > \aleph_0$,

$$\#\{N : N \equiv M,\ \#N = \kappa\}/\cong \quad \textit{is} \quad 2^\kappa.$$

Stability: what good is it?

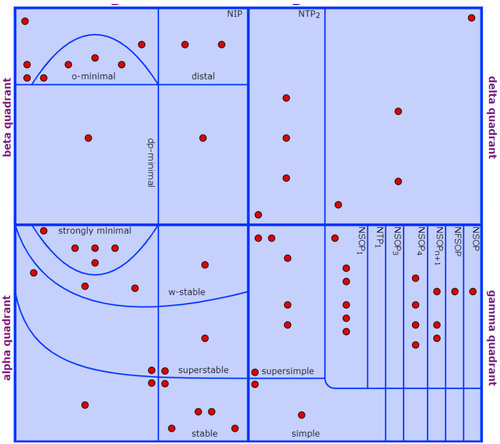
In a stable structure, we can define. . .

- The "dimension" dim(D) of a definable set.*
- "Independence" of elements $a_1, a_2, a_3, \ldots \in M$.
- The "connected component" G^0 of a definable group.
- "Prime models"

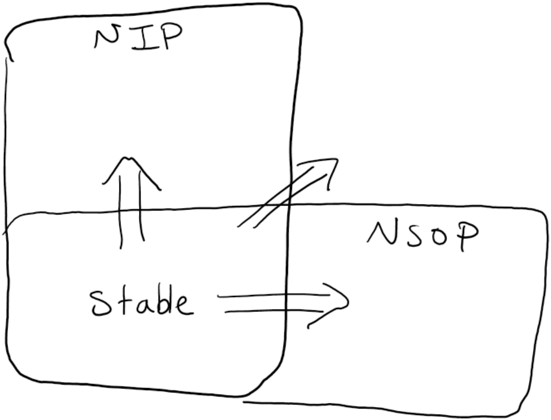
Applications of stability theory:

- Classification theory (Shelah)
- Differential algebra (Poizat, many others)
- Function-field Mordell-Lang (Hrushovski)
- Approximate subgroups (also Hrushovski)

The neostability universe



The neostability universe (simplified)



Neostability theory: generalizing stability theory to bigger classes, like NIP.

Why NIP?

These things are NIP but not stable:

- The fields $\mathbb{R}, \mathbb{Q}_p$.
- Ordered abelian groups like $(\mathbb{Z}, +, \leq)$.
- Algebraic groups like $SU(n)$, $SO(n)$.
- O-minimal structures like $(\mathbb{R}, +, \cdot, \exp)$.

NSOP is too restrictive.

Section 5

NIP rings

Reminder

"Ring" means "commutative unital ring".

Typical examples

These fields and rings are NIP:

- The real numbers $\mathbb{R}$.
- The p-adic integers
$$\mathbb{Z}_p = \varprojlim_{n\to\infty} \mathbb{Z}/p^n.$$
- The p-adic numbers $\mathbb{Q}_p = \mathrm{Frac}(\mathbb{Z}_p)$.
- The formal power series ring
$$k[[X]] = \{a_0 + a_1 X + a_2 X^2 + \cdots : a_0, a_1, a_2, \ldots \in k\},$$
for $k = \mathbb{C}, \mathbb{Q}_p, \mathbb{R}$.
- The formal Laurent series field $k((X)) = \mathrm{Frac}\, k[[X]]$ for $k = \mathbb{C}, \mathbb{Q}_p, \mathbb{R}$.

Warning

$\mathbb{F}_p[[X]]$ and $\mathbb{F}_p((X))$ are not NIP.

Local rings and henselianity

Let R be a ring.

Definition

- R is a *local ring* if there is a unique maximal ideal $\mathfrak{m}$.

Suppose R is local.

- The *residue field* is the quotient $k = R/\mathfrak{m}$.
- R is *Henselian* if: For any monic polynomial $P(X) \in K[x]$, if $\overline{P}(X) \in k[X]$ is the reduction mod $\mathfrak{m}$, and $\alpha \in k$ is a simple root of $\overline{P}(X)$, then α lifts to a root of $P(X)$.

Fact

$\mathbb{Z}_p$ is a henselian local ring with residue field $\mathbb{F}_p = \mathbb{Z}/p$.
$k[[X]]$ is a henselian local ring with residue field k.
The localization $\mathbb{Z}_{(2)} = \{a/(2b+1) : a, b \in \mathbb{Z}\}$ is a non-henselian local ring.

Prime ideals in an NIP ring

Let R be an NIP ring.

Theorem (Simon)

There is some n such that if $\mathfrak{p}_1, \ldots, \mathfrak{p}_k$ are pairwise incomparable prime ideals, then $k \leq n$.

Proof idea: $\mathfrak{p}_1, \ldots, \mathfrak{p}_k$ are independent sets.

Corollary

- *R is semilocal: only finitely many maximal ideals.*
- *If R is Noetherian, then $\dim(R) \leq 1$.*

Generalized henseliany conjecture

Conjecture (Generalized henselianity conjecture)

If R is an NIP ring, then $R = A_1 \times A_2 \times \cdots \times A_n$ for some henselian local rings $A_1, \ldots, A_n$.

Equivalent conjectures:

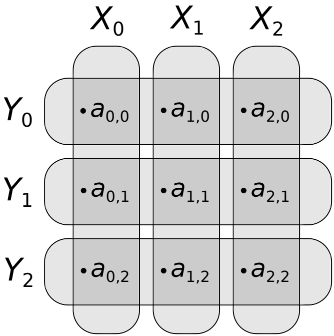
1. NIP integral domains are local.
2. NIP local rings are Henselian.
3. If $(K, +, \cdot, \ldots)$ is an NIP field with a definable field topology, then the inverse/implicit function theorem holds for polynomials.

Theorem (J)

The generalized henselianity conjecture holds when $\mathrm{char}(R) > 0$ *or when R is "finite-dimensional."*

Dimension in NIP structures

- In NIP structures, any definable set D has a “dimension” $\dim(D)$ called its “dependence rank”.
- $\dim(D)$ is a cardinal number, possibly infinite.
- $\dim(X \times Y) = \dim(X) + \dim(Y)$.
- Many NIP structures M satisfy $\dim(M) \leq 1$.



$\dim(D) \geq 2$ iff D contains a pattern of definable sets like this, ROUGHLY.

Positive characteristic: ideas in the proof

Let K be an infinite NIP field of characteristic $p > 0$.

Theorem (Kaplan-Scanlon-Wagner)

The Artin-Schreier map $\alpha(x) = x^p - x$ is onto.

Proof idea: the family of subgroups $\{b \cdot \alpha(K) : b \in K^\times\}$ isn't a VC class otherwise.

Corollary

If L/K is a finite separable extension, $[L : K]$ is prime to p.

Positive characteristic: ideas in the proof

Setting:

- R an NIP integral domain
- $\text{char}(R) = p > 0$.
- R has finitely many maximal ideals $\mathfrak{m}_1, \ldots, \mathfrak{m}_n$, $\underline{\text{and } n > 1}$.
- $J = \bigcap_i \mathfrak{m}_i$.

Strategy:

- Show that Artin-Schreier image $\alpha(J)$ has index $p^{n-1} > 1$ in J.
- Show that $\{b \cdot \alpha(J) : b \in K^\times\}$ isn't a VC class.

Section 6

NIP fields

Overview

For NIP fields...

- We *have* a <u>conjectural</u> classification.
- The classification is known in the finite-dimensional case.

Valued fields

Definition

A *valued field* is a pair (K, R) where K is a field and R is a subring such that for every $x \in K$, at least one of x or x^{-1} is in R.

R is called the *valuation ring*; it is always local, and has an associated *residue field* $R/\mathfrak{m}$.

K	R	$R/\mathfrak{m}$	Henselian?
$\mathbb{Q}_p$	$\mathbb{Z}_p$	$\mathbb{Z}/p$	Yes
$k((t))$	$k[[t]]$	k	Yes
$\mathbb{Q}$	$\mathbb{Z}_{(p)}$	$\mathbb{Z}/p$	No

Theorem (Anscombe-Jahnke)

A henselian valued field (K, R) is NIP iff the residue field $R/\mathfrak{m}$ is NIP plus a bunch of other conditions when $\mathrm{char}(R/\mathfrak{m}) > 0$*, henceforth abbreviated as "henselian*".*

The Shelah conjecture

Conjecture (Shelah)

If K is an NIP field, then one of the following holds:

- *K is finite.*
- *K is separably closed.*
- *K is real closed ($K \equiv \mathbb{R}$).*
- *K admits a henselian valuation ring $R \subsetneq K$.*

Theorem (Anscombe-Jahnke)

Assuming the Shelah conjecture, a field K is NIP iff there is a henselian* valuation ring R on K with $R/\mathfrak{m}$ being finite, real closed, or algebraically closed.

The Shelah conjecture

Everything holds in the "finite dimensional" case:

Theorem (J)

If K is a finite-dimensional NIP field, then K is finite, real closed, separably closed, or henselian.

From this, we get a classification of finite-dimensional NIP fields.

The Shelah conjecture

Theorem (J)

If K is a finite-dimensional NIP field, then K is finite, real closed, separably closed, or henselian.

Proof strategy:

Theorem

If K is a finite-dimensional NIP field, then there is usually a unique definable field topology on K of "valuation type".

ROUGHLY,

- Uniqueness implies "any definable valuation ring is henselian."
- Existence implies "some definable valuation ring exists".

The Shelah conjecture

Theorem

If K is a finite-dimensional NIP field, then there is usually *a unique definable field topology on K of "valuation type".*

Proof strategy: show that

$$\{D - D : D \subseteq K \text{ is definable and } \dim(D) = \dim(K)\}$$

is a neighborhood basis of 0 for a suitable field topology on K, where $D - D = \{x - y : x, y \in D\}$.

Idea

We expect $\dim(D) = \dim(K)$ to imply "D has non-empty interior". This forces the topology to be as above.

Prospects for classifying NIP rings

Theorem (d'Elbée, Halevi, Johnson)

- *If R is a 1-dimensional NIP valuation ring, and I is an ideal, and A is a finite subring of R/I, then the pullback $R \times_{R/I} A$ is a 1-dimensional NIP integral domain.*
- *All 1-dimensional NIP integral domains arise this way.*

Summary:

- 1-dimensional NIP integral domains are close to valuation rings.
- This fails for 2-dimensional NIP integral domains, strange new behavior occurs...

TO BE CONTINUED...

...hopefully.

Questions?