Cambridge University Press
978-1-107-05775-3 - A Guide to NIP Theories
Pierre Simon
Excerpt
More information

Chapter 1

INTRODUCTION

This text is an introduction to the study of NIP (or dependent) theories. It is meant to serve two purposes. The first is to present various aspects of NIP theories and give the reader sufficient background material to understand the current research in the area. The second is to advertise the use of honest definitions, in particular in establishing basic results, such as the so-called *shrinking of indiscernibles*. Thus although we claim no originality for the theorems presented here, a few proofs are new, mainly in chapters 3, 4 and 9.

We have tried to give a *horizontal* exposition, covering different, sometimes unrelated topics at the expense of exhaustivity. Thus no particular subject is dealt with in depth and mainly *low-level* results are included. The choices made reflect our own interests and are certainly very subjective. In particular, we say very little about algebraic structures and concentrate on combinatorial aspects. Overall, the style is concise, but hopefully all details of the proofs are given. A small number of facts are left to the reader as exercises, but only once or twice are they used later in the text.

The material included is based on the work of a number of model theorists. Credits are usually not given alongside each theorem, but are recorded at the end of the chapter along with pointers to additional topics.

We have included almost no preliminaries about model theory, thus we assume some familiarity with basic notions, in particular concerning compactness, indiscernible sequences and ordinary imaginaries. Those prerequisites are exposed in various books such as that of Poizat [95], Marker [81], Hodges [56] or the recent book [115] by Tent and Ziegler. The material covered in a one-semester course on model theory should suffice. No familiarity with stability theory is required.

History of the subject. In his early works on classification theory, Shelah structured the landscape of first order theories by drawing dividing lines defined by the presence or absence of different combinatorial configurations. The most important one is that of stability. In fact, for some twenty years, pure model theory did not venture much outside of stable theories. Shelah discovered the independence property when studying the possible behaviors for the function relating the size of a subset to the number of types over it. The class of theories lacking the independence property, or NIP theories, was studied very little in the early days. However some basic results were established, mainly by Shelah and Poizat (see [95, Chapter 12] for an account of those works).

As the years passed, various structures were identified as being NIP: most notably, Henselian valued fields of characteristic 0 with NIP residue field and ordered group (Delon [32]), the field $\mathbb{Q}_p$ of p-adics (see Bélair [15]) and ordered abelian groups (Gurevich and Schmitt [50]). However, NIP theories were not studied *per se*. In [91], Pillay and Steinhorn, building on work of van den Dries, defined o-minimal theories as a framework for tame geometry. This has been a very active area of research ever since. Although it was noticed from the start that o-minimal theories lacked the independence property, very little use of this fact was made until recently. Nevertheless, o-minimal theories provide a wealth of interesting examples of NIP structures.

In the years since 2000, the interest in NIP theories has been rekindled and the subject has been expanding ever since. First Shelah initiated a systematic study which lead to a series of papers: [107], [109], [102], [111], [110]. Amongst other things, he established the basic properties of forking, generalized a theorem of Baisalov and Poizat on externally definable sets, defined some subclasses, so called "strongly dependent" and "strongly$^+$ dependent". This work culminates in [110] with the proof that NIP theories have few types up to automorphism (over saturated models). Parallel to this work, Hrushovski, Peterzil and Pillay developed the theory of *measures* (a notion introduced by Keisler in [71]) in order to solve Pillay's conjecture on definably compact groups in o-minimal theories.

A third line of research starts with the work of Hrushovski, Haskell and Macpherson on algebraically closed valued fields (ACVF) and in particular on *metastability* ([52]). This lead to Hrushovski and Loeser giving a model theoretic construction of Berkovich spaces in rigid geometry as spaces of stably-dominated types, which made an explicit use of the NIP property along with the work on metastability.

Motivated by those results, a number of model theorists became interested in the subject and investigated NIP theories in various directions. We will present some in the course of this text and mention others at the end of each chapter. It is not completely clear at this point how the subject will develop and what topics will turn out to be the most fruitful.

Let us end this general introduction by mentioning where NIP sits with respect to other classes of theories. First, all stable theories are NIP, as are o-minimal and C-minimal theories. Another well-studied extension of stability is that of simple theories (see Wagner [122]), however it is in a sense orthogonal to NIP: a theory is both simple and NIP if and only if it is stable. Simple and NIP theories both belong to the wider class of NTP$_2$ theories (defined in Chapter 5).

Organization of this text. Aside from the introduction and appendices, the text is divided into 8 chapters, each one focussing on a specific topic. In Chapter 2, we present the classical theory as it was established by Shelah and Poizat. We first work formula-by-formula giving some equivalent definitions of NIP. We then move to invariant types and Morley sequences. Starting then, and throughout most of the text, we assume that our ambient theory T is NIP. That assumption will be dropped only for the first three sections of Chapter 5. Many results could be established for an NIP *formula* (or type) inside a (possibly) independent theory, but for the sake of clarity we will not work at this level of generality.

The end of Chapter 2 is a collection of appendices on different subjects. We study dense trees in some detail as one can obtain from them a lot of intuition on NIP theories, we recall basic facts on stable theories, discuss the strict order property and give the original characterization of NIP theories by counting types.

In Chapter 3 we define honest definitions. They serve as a substitute to definability of types in NIP theories. We use them to prove Shelah's theorem on expanding a model by externally definable sets and the very important results about shrinking of indiscernibles.

Chapter 4 deals with dp-rank and strong dependence. In the literature, one can find up to three different definitions of dp-rank, based on how one handles the problem of *almost finite*, non-finite rank. None of them is perfect, but we have decided to use the same convention as in Adler's paper [2] on burden, although we refrain from duplicating limit cardinals into κ_- and κ. Instead, we define when dp-rk$(p) < \kappa$, and it may happen that the dp-rank of a type is not defined (for example it can be $< \aleph_0$ but greater than all integers).

In Chapter 5, we study forking and dividing. The main results are $bdd(A)$-invariance of non-forking extensions (Hrushovski and Pillay [61]) and equality of forking and dividing over models (Chernikov and Kaplan [25]). The right context for this latter result is NTP_2 theories, but here again we assume NIP which slightly simplifies some proofs.

The next three chapters have a different flavor. In Chapter 6, we change the framework to that of finite combinatorics. We are concerned with families of finite VC-dimension over finite sets. The finite and infinite approaches come together to prove uniformity of honest definitions. In Chapter 7, the two frameworks are combined with the introduction of Keisler measures. The most important class of examples of such measures is that of translation-invariant measures on definable groups. Those are investigated in Chapter 8. We also discuss there connected components of groups.

The last chapter addresses the problem of characterizing NIP structures which are in some sense completely unstable. They are called *distal* structures.

Finally, two appendices are included. The first one gives some algebraic examples and in particular records some facts about valued fields. Most of the proofs are omitted, but we explain how to show that those structures are NIP. The other appendix is very short and collects results about probability theory for reference in the text.

Acknowledgments. Part of the material presented here was exposed in 2011 during a series of lectures in Paris. I would like to thank all the participants of this seminar: Élisabeth Bouscaren, Zoé Chatzidakis, Pablo Cubides Kovacsics, Françoise Delon, Martin Hils, Samaria Montenegro, Françoise Point, Silvain Rideau and Patrick Simonetta. I have received very helpful advice from a number of people: David Bradley-Williams, Rafel Farré, Martin Hils, Udi Hrushovski, Itay Kaplan, Dugald Macpherson, Dave Marker and the two anonymous referees. Special thanks to Alex Kruckman for his extensive list of comments and corrections.

1.1. Preliminaries

We work with a complete, usually one-sorted, theory T in a language L. We have a monster model $\mathcal{U}$ which is $\bar{\kappa}$-saturated and homogeneous. A subset $A \subset \mathcal{U}$ is *small* if it is of size less than $\bar{\kappa}$. For $A \subseteq \mathcal{U}$, $L(A)$ denotes the set of formulas with parameters in A. In particular, $\phi(x) \in L$ means that ϕ is without parameters.

We do not usually distinguish between points and tuples. If a is a tuple of size $|a|$, we will write $a \in A$ to mean $a \in A^{|a|}$. Similarly, letters such as $x, y, z, \ldots$ are used to denote tuples of variables.

We often work with *partitioned formulas*, namely formulas $\phi(x; y)$ with a separation of variables into object and parameters variables. The intended partition is indicated with a semicolon.

If $A \subset \mathcal{U}$ is any set, and $\phi(x)$ is a formula, then $\phi(A) = \{a \in A^{|x|} : \mathcal{U} \models \phi(a)\}$. The set of types over A in the variable x is denoted by $S_x(A)$. We will often drop the x. If $p \in S_x(A)$, we might write p_x or $p(x)$ to emphasize that p is a type in the variable x. We say that a type p *concentrates* on a definable set $\phi(x)$ if $p \vdash \phi(x)$. If $A \subseteq B$ and p is a type over B, we denote by $p \upharpoonright A$, or by $p|_A$ the *restriction* of p to A.

We will often write either $\models \phi(a)$ or $a \models \phi(x)$ to mean $\mathcal{U} \models \phi(a)$, and similarly for types.

We use the notation ϕ^0 to mean $\neg\phi$ and ϕ^1 to mean ϕ. If $\phi(x; y)$ is a partitioned formula, a ϕ-*type* over A is a maximal consistent set of formulas of the form $\phi(x; a)^{\varepsilon}$, for $a \in A$ and $\varepsilon \in \{0, 1\}$. The set of ϕ-types over A is denoted by $S_\phi(A)$.

A *global type*, is a type over $\mathcal{U}$.

The group of *automorphisms* of $\mathcal{U}$ is denoted by $\mathrm{Aut}(\mathcal{U})$, whereas $\mathrm{Aut}(\mathcal{U}/A)$ refers to the subgroup of $\mathrm{Aut}(\mathcal{U})$ of automorphisms fixing A pointwise.

1.1.1. Indiscernible sequences. We will typically denote sequences of tuples by $I = (a_i : i \in \mathcal{I})$ where $\mathcal{I}$ is some linearly ordered set. The order on $\mathcal{I}$ will be denoted by $<_{\mathcal{I}}$ or simply $<$ if no confusion arises. If $I = (a_i : i \in \mathcal{I})$ and $J = (b_j : j \in \mathcal{J})$, then we write the concatenation of I and J as $I + J$. It has I as initial segment and J as the complementary final segment. We use the notation (a) to denote the sequence which has a as unique element.

We say that the sequence I is *endless* if the indexing order $\mathcal{I}$ has no last element.

Let Δ be a finite set of formulas and A a set of parameters. A (possibly finite) sequence $I = (a_i : i \in \mathcal{I})$ is Δ-indiscernible over A, if for every integer k and two increasing tuples $i_1 <_{\mathcal{I}} \cdots <_{\mathcal{I}} i_k$ and $j_1 <_{\mathcal{I}} \cdots <_{\mathcal{I}} j_k$, $b \in A$ and formula $\phi(x_1, \ldots, x_k; y) \in \Delta$, we have $\phi(a_{i_1}, \ldots, a_{i_k}; b) \leftrightarrow \phi(a_{j_1}, \ldots, a_{j_k}; b)$. An *indiscernible sequence* is an *infinite* sequence which is Δ-indiscernible for all Δ.

Let $I = (a_i : i \in \mathcal{I})$ be any sequence. We define the Ehrenfeucht-Mostowski type (or *EM-type*) of I over A to be the set of $L(A)$-formulas $\phi(x_1, \ldots, x_n)$ such that $\mathcal{U} \models \phi(a_{i_1}, \ldots, a_{i_n})$ for all $i_1 < \cdots < i_n \in \mathcal{I}$, $n < \omega$. If I is an indiscernible sequence, then for every n, the restriction of the EM-type of I to formulas in n variables is a complete type over A. If $A = \emptyset$, then we can omit it. We will write $I \equiv_A^{EM} J$ to mean that I and J are two A-indiscernible sequences having the same EM-type over A. If I is any sequence and $\mathcal{J}$ is any infinite linear order, then using Ramsey's theorem and compactness, we can find an indiscernible sequence J indexed by $\mathcal{J}$ and realizing the EM-type of I (see e.g., [115, Lemma 5.1.3]).

A sequence I is *totally indiscernible* (or set indiscernible) if every permutation of it is indiscernible. If a sequence $(a_i : i \in \mathcal{I})$ is not totally indiscernible, then there is some formula $\phi(x, y)$, possibly with parameters, which orders it, that is such that $\phi(a_i, a_j)$ holds if and only if $i \leq j$.

Most of the time, Ramsey and compactness will be sufficient for us to construct indiscernible sequences. However, we will need once or twice a more powerful result which is an easy application of the Erdős-Rado theorem.

Proposition 1.1. *Let A be a set of parameters, $\kappa > |T| + |A|$ and $\lambda = \beth_{(2^\kappa)^+}$. Let $(a_i : i < \lambda)$ be a sequence of tuples all of the same size $\leq \kappa$. Then there is an indiscernible sequence $(b_i : i < \omega)$ such that for any $i_1 < \cdots < i_n < \omega$, there are some $j_1 < \cdots < j_n < \lambda$ with*

$$a_{i_1} \ldots a_{i_n} \equiv_A b_{j_1} \ldots b_{j_n}.$$

See e.g. [21, Proposition 1.6] for a proof.

Cambridge University Press
978-1-107-05775-3 - A Guide to NIP Theories
Pierre Simon
Excerpt
More information

Chapter 2

THE NIP PROPERTY AND INVARIANT TYPES

In this chapter, we introduce the basic objects of our study. We first define the notion of an NIP formula. The combinatorial definition is not very handy, and we give an equivalent characterization involving indiscernible sequences which is the one we will most often use. We then define NIP theories as theories in which all formulas are NIP and give some examples. We discuss invariant types and their relation to indiscernible sequences. In particular, we define generically stable types, which share some characteristics with types in stable theories.

To illustrate the notions considered, we prove some results on definable groups in NIP theories: the Baldwin-Saxl theorem, and Shelah's theorem on existence of definable envelopes for abelian subgroups.

In the "additional topics" section we introduce trees, which serve as a paradigm for NIP theories. Many examples of NIP theories are either explicitly constructed as a tree with additional structure, or have an underlining tree-structure (valued fields for example). We discuss in more details the theory of dense meet-trees, and in particular describe indiscernible sequences in it. The next subsection collects some facts about stable formulas and theories. We then present the strict order property and finally give yet another characterization of NIP in terms of counting types.

2.1. NIP formulas

Let $\phi(x; y)$ be a partitioned formula. We say that a set A of $|x|$-tuples is *shattered* by $\phi(x; y)$ if we can find a family $(b_I : I \subseteq A)$ of $|y|$-tuples such that

$$\mathcal{U} \models \phi(a; b_I) \iff a \in I, \quad \text{for all } a \in A.$$

By compactness, this is equivalent to saying that every finite subset of A is shattered by $\phi(x; y)$.

Definition 2.1. A partitioned formula $\phi(x; y)$ is NIP (or dependent) if no infinite set of $|x|$-tuples is shattered by $\phi(x; y)$.

If a formula is not NIP, we say that it has IP.

Remark 2.2. The acronym IP stands for the *Independence Property* and NIP is its negation. Some authors (notably Shelah) use the terminology dependent/independent instead of NIP/IP.

Remark 2.3. If $\phi(x; y)$ is NIP, then by compactness, there is some integer n such that no set of size n is shattered by $\phi(x; y)$.

The maximal integer n for which there is some A of size n shattered by $\phi(x; y)$ is called the *VC-dimension* of ϕ. If there is no such integer, that is if the formula ϕ has IP, then we say that its VC-dimension is infinite.

Example 2.4.

- Let T be DLO: the theory of dense linear orders with no endpoints. Then the formula $\phi(x; y) = (x \leq y)$ is NIP of VC-dimension 1. Indeed, if we have $a_1 < a_2$, then we cannot find some $b_{\{2\}}$ such that
$$\mathcal{U} \models \neg\phi(a_1; b_{\{2\}}) \wedge \phi(a_2; b_{\{2\}}).$$
- If $\phi(x; y)$ is a stable formula, then it is NIP (see Section 2.3.2 if needed).
- If T is the theory of arithmetic, then the formula $\phi(x; y) =$ "x divides y" has IP. To see this, take any $N \in \mathbb{N}$ and $A = \{p_0, \ldots, p_{N-1}\}$ a set of distinct prime numbers. For any $I \subseteq N$, set b_I to be $\prod_{i \in I} p_i$. We have $\models \phi(p_i, b_I) \iff i \in I$. Thus the set A is shattered and $\phi(x; y)$ has infinite VC-dimension.
- If T is the random graph in the language $L = \{R\}$, then the formula $\phi(x; y) = xRy$ has IP. In fact any set of elements is shattered by ϕ.
- If T is a theory of an infinite Boolean algebra, in the natural language $\{0, 1, \neg, \vee, \wedge\}$, then the formula $x \leq y$ (defined as $x \wedge y = x$) has IP. Indeed, it shatters any set A with $a \wedge b = 0$ for $a \neq b \in A$.

If $\phi(x; y)$ is a partitioned formula, we let $\phi^{opp}(y; x) = \phi(x; y)$. Hence ϕ^{opp} is the same formula as ϕ, but we have exchanged the role of variables and parameters. The following fact will be used throughout this text, often with no explicit mention.

LEMMA 2.5. *The formula $\phi(x; y)$ is NIP if and only if $\phi^{opp}(y; x)$ is NIP.*

PROOF. Assume that $\phi(x; y)$ has IP. Then by compactness, we can find some $A = \{a_i : i \in \mathfrak{P}(\omega)\}$ which is shattered by $\phi(x; y)$ as witnessed by tuples b_I, $I \subseteq \mathfrak{P}(\omega)$. Let $B = \{b_j : j \in \omega\}$ where $b_j := b_{I_j}$ and $I_j := \{X \subseteq \omega : j \in X\}$. Then for any $J_0 \subseteq \omega$, we have
$$\models \phi(a_{J_0}, b_j) \iff j \in J_0.$$
This shows that B is shattered by ϕ^{opp}. Therefore ϕ^{opp} has IP. ⊣

Remark 2.6. The VC-dimension of a formula ϕ need not be equal to the VC-dimension of the opposite formula ϕ^{opp}. For example, let T be the theory of equality, then the formula $\phi(x; y_1y_2y_3) = (x = y_1 \vee x = y_2 \vee x = y_3)$ has VC-dimension 3, but the opposite formula only has VC-dimension 2.

See Lemma 6.3 for inequalities linking the VC-dimensions of ϕ and ϕ^{opp}. For now, our only concern is whether they are finite or not.

We now give an equivalent characterization of NIP which is often the most convenient one to use.

LEMMA 2.7. *The formula $\phi(x; y)$ has IP if and only if there is an indiscernible sequence $(a_i : i < \omega)$ and a tuple b such that*

$$\models \phi(a_i; b) \iff i \text{ is even.}$$

PROOF. $(\Leftarrow)$: Assume that there is a sequence $(a_i : i < \omega)$ and a tuple b as above. Let $I \subseteq \omega$. We show that there is some b_I such that $\phi(a_i; b_I)$ holds if and only if $i \in I$. We can find an increasing one-to-one map $\tau : \omega \to \omega$ such that for all $i \in \omega$, $\tau(i)$ is even if and only if i is in I. Then by indiscernibility the map sending a_i to $a_{\tau(i)}$ for all $i < \omega$ is a partial isomorphism. It extends to a global automorphism σ. Then take $b_I = \sigma^{-1}(b)$.

$(\Rightarrow)$: Assume that $\phi(x; y)$ has IP. Let $A = (a_i : i < \omega)$ be a sequence of $|x|$-tuples which is shattered by $\phi(x; y)$. By Ramsey and compactness, we can find some indiscernible sequence $I = (c_i : i < \omega)$ of $|x|$-tuples realizing the EM-type of A. It follows that for any two disjoint finite sets I_0 and I_1 of I, the partial type $\{\phi(x; c) : c \in I_0\} \cup \{\neg\phi(x; c) : c \in I_1\}$ is consistent. Then by compactness, I is shattered by $\phi(x; y)$. In particular, there is b such that $\phi(c_i; b)$ holds if and only if i is even. $\dashv$

Let $\phi(x; y)$ be an NIP formula, then there is a finite set Δ of formulas and an integer $n_{\phi,\Delta}$ such that the following do not exist:

- $(a_i : i < n_{\phi,\Delta})$ a Δ-indiscernible sequence of $|x|$-tuples;
- b a $|y|$-tuple, such that $\neg(\phi(a_i; b) \leftrightarrow \phi(a_{i+1}; b))$ holds for $i < n_{\phi,\Delta} - 1$.

Indeed, if we could not find such Δ and $n_{\phi,\Delta}$, then the partial type in variables $(x_i : i < \omega)\hat{\ }y$ stating that $(x_i : i < \omega)$ is an indiscernible sequence and $\neg(\phi(x_i; y) \leftrightarrow \phi(x_{i+1}; y))$ holds for all $i < \omega$ would be consistent, contradicting the previous lemma.

Let $I = (a_i : i \in \mathcal{I})$ be an indiscernible sequence and take an NIP formula $\phi(x; y) \in L$ and a tuple of parameters $b \in \mathcal{U}$. Then there is a maximal integer n such that we can find $i_0 < \cdots <_{\mathcal{I}} i_n$ with $\neg(\phi(a_{i_k}; b) \leftrightarrow \phi(a_{i_{k+1}}; b))$ for all $k < n$. We call such an n the number of alternations of $\phi(x; b)$ on the sequence I and write it as $\text{alt}(\phi(x; b), I)$. We let $\text{alt}(\phi(x; y))$ denote the maximum value of $\text{alt}(\phi(x; b), I)$ for b ranging in $\mathcal{U}$ and I ranging over all indiscernible sequences. Note that this maximum exists and is bounded by the number $n_{\phi,\Delta}$ of the previous paragraph. We sometimes call $\text{alt}(\phi(x; y))$ the *alternation rank* (or *number*) of $\phi(x; y)$.

PROPOSITION 2.8. *The formula $\phi(x; y)$ is NIP if and only if for any indiscernible sequence $(a_i : i \in \mathcal{I})$ and tuple b, there is some end segment $\mathcal{I}_0 \subseteq \mathcal{I}$ and $\varepsilon \in \{0, 1\}$ such that $\phi(a_i; b)^\varepsilon$ holds for any $i \in \mathcal{I}_0$.*

PROOF. If $\mathcal{I}$ has a last element i_0, let $\mathcal{I}_0 = \{i_0\}$. Otherwise, this follows immediately from the discussion above. ⊣

Of course the equivalence also holds if we restrict to sequences indexed by ω, or in fact by any given linear order with no last element.

LEMMA 2.9. *A Boolean combination of NIP formulas is NIP.*

PROOF. It is clear from the definition that the negation of an NIP formula is NIP.

Let $\phi(x; y)$ and $\psi(x; y)$ be two NIP formulas and we want to show that $\theta(x; y) = \phi(x; y) \wedge \psi(x; y)$ is NIP. We use the criterion from Proposition 2.8. Let $(a_i : i \in \mathcal{I})$ be an indiscernible sequence of $|x|$-tuples and let b be a $|y|$-tuple. Let $\mathcal{I}_\phi \subseteq \mathcal{I}$ be an end segment such that $\phi(a_i; b) \leftrightarrow \phi(a_j; b)$ holds for $i, j \in \mathcal{I}_\phi$. Define $\mathcal{I}_\psi$ similarly and let $\mathcal{I}_0 = \mathcal{I}_\phi \cap \mathcal{I}_\psi$. Then $\mathcal{I}_0$ is an end segment of $\mathcal{I}$ and we have $\theta(a_i; b) \leftrightarrow \theta(a_j; b)$ for $i, j \in \mathcal{I}_0$. This shows that $\theta(x; y)$ is NIP. ⊣

2.1.1. NIP theories.

DEFINITION 2.10. The theory T is NIP if all formulas $\phi(x; y) \in L$ are NIP.

Note that if T is NIP, then also all formulas $\phi(x; y)$ with parameters are NIP, since if $\phi(x; y, d)$ has IP, then so does $\phi(x; y\hat{\ }z)$.

PROPOSITION 2.11. *Assume that all formulas $\phi(x; y) \in L$ with $|y| = 1$ are NIP, then T is NIP.*

PROOF. Assume that all formulas $\phi(x; y)$ with $|y| = 1$ are NIP.

CLAIM. *Let $(a_i : i < |T|^+)$ be an indiscernible sequence of tuples, and let $b \in \mathcal{U}$, $|b| = 1$. Then there is some $\alpha < |T|^+$ such that the sequence $(a_i : \alpha < i < |T|^+)$ is indiscernible over b.*

PROOF OF CLAIM. If this does not hold, then for every $\alpha < |T|^+$, for some formula $\delta_\alpha(x_1, \dots, x_{k(\alpha)}; y)$, we can find

$$\alpha < i_1 < \cdots < i_{k(\alpha)} < |T|^+ \quad \text{and} \quad \alpha < j_1 < \cdots < j_{k(\alpha)} < |T|^+$$

such that $\models \delta_\alpha(a_{i_1}, \dots, a_{i_{k(\alpha)}}; b) \wedge \neg\delta_\alpha(a_{j_1}, \dots, a_{j_{k(\alpha)}}; b)$. There is some formula $\delta(x_1, \dots, x_k; y)$ such that $\delta_\alpha = \delta$ for cofinally many values of α. Then we can construct inductively a sequence $I = (i_1^l\hat{\ }\dots\hat{\ }i_k^l : l < \omega)$ such that $i_1^l < \cdots < i_k^l < i_1^{l+1}$ for all $l < \omega$ and $\delta(a_{i_1^l}, \dots, a_{i_k^l}, b)$ holds if and only if l is even. As the sequence $(a_{i_1^l} \dots a_{i_k^l} : l < \omega)$ is indiscernible, this contradicts the assumption that $\delta(x_1, \dots, x_k; y)$ is NIP. ⊣ (Claim)

Now let $\phi(x; y)$ be any formula, where $y = y_1\hat{\ }\dots\hat{\ }y_n$ is an n-tuple. Let $(a_i : i < |T|^+)$ be any indiscernible sequence of $|x|$-tuples and let $b = b_1\hat{\ }\dots\hat{\ }b_n$ be an n-tuple. By the claim, there is some $\alpha_1 < |T|^+$ such that the sequence