Annals of Pure and Applied Logic 164 (2013) 294–318

Contents lists available at SciVerse ScienceDirect

Annals of Pure and Applied Logic

www.elsevier.com/locate/apal

Distal and non-distal NIP theories

Pierre Simon

École Normale Supérieure, DMA, 45, rue d'Ulm, 75005 Paris, France

ARTICLE INFO

Article history:
Received 9 May 2011
Received in revised form 16 May 2012
Accepted 17 October 2012
Available online 1 November 2012
Communicated by T. Scanlon

MSC:
03C45
03C68

Keywords:
NIP
Distal
Stable
Keisler measures
Invariant types
VC-dimension

ABSTRACT

We study one way in which stable phenomena can exist in an NIP theory. We start by defining a notion of 'pure instability' that we call 'distality' in which no such phenomenon occurs. O-minimal theories and the p-adics for example are distal. Next, we try to understand what happens when distality fails. Given a type p over a sufficiently saturated model, we extract, in some sense, the stable part of p and define a notion of stable independence which is implied by non-forking and has bounded weight.

1. Introduction

We study one way in which stability and order can interact in an NIP theory. More precisely, we are interested in the situation where stability and order are intertwined. We start by giving some very simple examples illustrating what we mean.

Consider $M_0 \models$ DLO. A type of $S_1(M_0)$ is determined by a cut in M_0 and two types corresponding to different cuts are orthogonal. If we take now M_1 a model of some o-minimal theory, still a 1-type is determined by a cut, but in general, types that correspond to different cuts are not orthogonal. However this is true over indiscernible sequences in the following sense: assume $\langle a_t\colon t < \omega + \omega \rangle \subset M_1$ is an indiscernible sequence. By NIP, the sequences of types $\langle \mathrm{tp}(a_t/M_1)\colon t < \omega \rangle$ and $\langle \mathrm{tp}(a_{\omega+t}/M_1)\colon t < \omega \rangle$ converge in $S(M_1)$. Then the two limit types are orthogonal (this follows from dp-minimality, see Corollary 2.30). An indiscernible sequence with that property will be called *distal*.[1] A theory is distal if all indiscernible sequences are distal. So any o-minimal theory is distal.

Distality for an indiscernible sequence can be considered as an opposite notion to that of total indiscernibility.

Let now M_2 be a model of ACVF (or any other C-minimal structure) and consider an indiscernible sequence $(a_i)_{i<\omega}$ of elements from the valued field sort. Two different behaviors are possible: either the sequence is totally indiscernible, this happens if and only if $\mathrm{val}(a_i - a_j) = \mathrm{val}(a_{i'} - a_{j'})$ for all $i \neq j$, $i' \neq j'$, or the sequence is distal. Again, this will follow from the results in Section 2, but could be proved directly. So M_2 is neither stable nor distal; the two phenomena exist but do not interact in a single indiscernible sequence of points.

E-mail address: pierre.simon@normalesup.org.

[1] Thanks to Itay Kaplan for suggesting the name.

0168-0072/$ – see front matter
http://dx.doi.org/10.1016/j.apal.2012.10.015

Consider now a fourth structure (a 'colored order') M_3 in the language $L_3 = \{\leqslant, E\}$: M_3 is totally ordered by $\leqslant$ and E defines an equivalence relation, each E class being dense co-dense with respect to $\leqslant$. Now an indiscernible sequence of elements from different E classes is neither totally indiscernible nor distal. Given two limit types p_x and q_y of different cuts in such a sequence, the type $p_x \cup q_y$ is consistent with xEy and with $\neg xEy$. Here it is clear that the 'stable part' of a type should be its E-class.

The idea behind the work in this paper is that every ordered indiscernible sequence in an NIP theory should look like a colored order: there is an order for which different cuts are orthogonal and something stable on top of it which does not see the order (see Section 3).

1.0.1. A word about measures

Keisler measures will be used a little in this work, however the reader not familiar with them can skip all parts referring to measures without harm. For this reason, we will be very brief in recalling some facts about them and refer the reader to [8] and [9]. They however give some understanding of the intuition behind some definitions and results. We explain this now.

A Keisler measure (or simply a measure) is a Borel probability measure on a type space $S_x(A)$. Basic definitions for types (non-forking, invariance, coheir, Morley sequence etc.) generalize naturally to measures (see [8] and [9]). Of interest to us is the notion of *generically stable measure*. A measure is generically stable if it is both definable and finitely satisfiable over some small set. Equivalently, its Morley sequence is totally indiscernible. Such measures are defined and studied by Hrushovski, Pillay and the author in [9]. Furthermore, it is shown in [17] that some general constructions give rise to them, and in this sense they are better behaved than the more natural notion of generically stable type.

This paper can be considered as an attempt to understand where generically stable measures come from. What stable phenomena do generically stable measures detect? What does the existence of generically stable measures in some particular theory tell us about types? The first test question was: Can we characterize theories which have non-trivial generically stable measures? Here "non-trivial" means "non-smooth": a measure is smooth if it has a unique extension to any bigger set of parameters. This question is answered in Section 2: a theory has a non-smooth generically stable measure if and only if it is not distal.

The main tool at our disposal to link measures to indiscernible sequences is the construction of an average measure of an indiscernible segment (see [9, Lemma 3.4] or [17, Section 3] for a more elaborate construction). Such a measure is always generically stable. The intuition we suggest is that the 'order' component of the sequence is evened out in the average measure and only the 'stable' component remains.

1.0.2. Organization of the paper and main results

The paper is organized as follows. The first section contains some basic facts about NIP theories and Keisler measures. We give a number of definitions concerning indiscernible sequences and some basic results illustrating how we can manipulate them. Section 2 studies distal theories. They are defined as theories in which every indiscernible sequence is distal, as explained above. We show that this condition can also be seen through invariant types and generically stable measures. The main results can be summarized by the following theorem.

Theorem 1.1. *Let T be NIP. Then the following are equivalent*:

- *T is distal,*
- *Any two invariant types that commute are orthogonal,*
- *All generically stable measures are smooth.*

Furthermore, it is enough to check any one of those conditions in dimension 1.

As a consequence, o-minimal theories and the p-adics are distal as are more generally any dp-minimal theory with no generically stable type.

Section 3 can be read almost independently of the previous one: it contains a study of the intermediate case of an NIP theory that is neither stable nor distal. We deal with the problem of understanding to what extend non-distality is witnessed by stable-like interactions between tuples. If M is a $|T|^+$-saturated model, we define a notion of s-independence denoted $a \downarrow^s_M b$ which is symmetric, is implied by forking-independence and has bounded weight. We use it to show that two commuting types behave with respect to each other like types in a stable theory (we recover some definability and uniqueness of the non-forking extension). The guiding intuition is that of the colored order where elements have a well defined *stable part* (the image in the quotient) and in that case $a \downarrow^s_M b$ means that the stable parts are independent. We do not attempt to give any meaning to the 'stable part' of a type in general, and we do not even expect there to be a possible meaning for it. We find that the intuition "s-independence corresponds to independence of stable parts" is useful in understanding those results. Of course, it may turn out some day to be misleading.

As an application of those ideas, we prove the following 'finite–co-finite theorem' (Theorem 3.30) and give an application of it to the study of externally definable sets.