PARIS 7 LOGIQUE

CNRS

N° 91

# Séminaire de Structures Algébriques Ordonnées

2015 -- 2016

**F. Delon, M. A. Dickmann, D. Gondard, T. Servi**

Mai 2017

**Equipe de Logique Mathématique**

***Prépublications***

Institut de Mathématiques de Jussieu – Paris Rive Gauche

Universités Denis Diderot (Paris 7) et Pierre et Marie Curie (Paris 6)

Bâtiment Sophie Germain, 5 rue Thomas Mann – 75205 Paris Cedex 13
Tel. : (33 1) 57 27 91 71 – Fax : (33 1) 57 27 91 78

Les volumes des contributions au Séminaire de Structures Algébriques Ordonnées rendent compte des activités principales du séminaire de l'année indiquée sur chaque volume. Les contributions sont présentées par les auteurs, et publiées avec l'agrément des éditeurs, sans qu'il soit mis en place une procédure de comité de lecture.

Ce séminaire est publié dans la série de prépublications de l'Equipe de Logique Mathématique, Institut de Mathématiques de Jussieu–Paris Rive Gauche (CNRS -- Universités Paris 6 et 7). Il s'agit donc d'une édition informelle, et les auteurs ont toute liberté de soumettre leurs articles à la revue de leur choix.

Cette publication a pour but de diffuser rapidement des résultats ou leur synthèse, et ainsi de faciliter la communication entre chercheurs.

The proceedings of the Séminaire de Structures Algébriques Ordonnées constitute a written report of the main activities of the seminar during the year of publication. Papers are presented by each author, and published with the agreement of the editors, but are not refereed.

This seminar is published in the preprint series of the Equipe de Logique Mathématique, Institut de Mathématiques de Jussieu–Paris Rive Gauche (CNRS -- Universités Paris 6 et 7). It has the character of an informal publication aimed at speeding up the circulation of information and, hence, facilitating communication among researchers in the field. The authors are free to submit to any journal the papers preprinted in these proceedings.

# HABITS NEUFS POUR DDG…

Trente ans après sa création, les organisateurs du séminaire s'interrogeaient sur son avenir : le continuait-on, ou bien 2015 était-il le bon moment de l'arrêter ? La question fut résolue en faveur de la poursuite du séminaire, et cela pour deux raisons.

D'abord, une large majorité de collègues et amis familiers du passé du séminaire ---dont beaucoup étaient présents à Luminy--- a encouragé sa poursuite.

Ensuite, en Juillet 2015, Tamara SERVI, dont le travail scientifique est proche de certains des thèmes couverts par le séminaire, a été nommée Maître de Conférences à l'Université Paris Diderot (Paris 7). Tamara a accepté de nous rejoindre, amenant ainsi au séminaire un apport bienvenu de sang nouveau.

Depuis Octobre 2015 le surnom du séminaire est DDGS.

# NEW CLOTHS FOR DDG ...

Thirty years after its foundation, the future of the seminar after the 2015 meeting at CIRM (Luminy) was a concern for the organizers: should we continue, or was 2015 the right moment to stop? The question was eventually resolved in favor of continuing the seminar for two reasons.

First, by the encouragement of colleagues and friends familiar with the seminar's past work -- many of them present at Luminy -- who by a large majority favored continuing the seminar.

Second, by the appointment in July 2015 of Tamara SERVI as a Lecturer at the University Paris Diderot (Paris 7), whose scientific work is close to some of the seminar's themes, and who gladly joined us, providing a welcome influx of young blood.

Since October 2015 the seminar's nickname has been DDGS.

# UNIVERSITES PARIS VI et VII

Projets Logique Mathématique et Théorie des Nombres
Institut de Mathématiques de Jussieu – UMR 7586

# SEMINAIRE DE STRUCTURES ALGEBRIQUES ORDONNEES

**Responsables: F. Delon, M. Dickmann, D. Gondard, T. Servi**

**2015-2016**

---

## Liste des exposés

01/12/15 **Françoise DELON** (Equipe de Logique, IMJ-PRG, Paris 7)
Un langage qui donne les boules.

08/12/15 **Esther ELBAZ** (Equipe de logique, IMJ-PRG, Paris 7)
Anneaux de Grothendieck et fonctions de paires.

12/01/16 **Erik WALSBERG** (IMJ-PRG, Paris 6)
Geometry over first order expansions of the ordered field of real numbers that do not define the integers.

19/01/16 **Raf CLUCKERS** (Université Lille 1)
Real, p-adic, and motivic oscillatory integrals.

09/02/16 **Journée spéciale en l'honneur du 70ème anniversaire de F. Miraglia**
(dédiée principalement aux formes quadratiques).

10h00 - 11h00. **Vincent ASTIER** (Univ. College Dublin, Irlande)
Signatures of hermitian forms and positive cones on algebras with involution.

11h15 - 12h15. **Danielle GONDARD** (IMJ-PRG, Paris 6)
Towards an abstract description of the space of valuation fans.

14h15 - 15h15. **Max DICKMANN** (IMJ-PRG, Paris 7)
Faithfully quadratic rings; an overview.

15h30 – 16h30. **Eberhard BECKER** (TU Dortmund, Allemagne)
Sums of powers in function fields in one variable over R.

15/03/16 **Margaret THOMAS** (Université de Constance, Allemagne)
Smooth parameterization in o-minimal structures.

31/05/16 **José Fernando GALVAN** (Univ. Complutense, Madrid)
On the substitution theorem for rings of semialgebraic functions.

## UNIVERSITES PARIS VI et VII

Projets Logique Mathématique et Théorie des Nombres
Institut de Mathématiques de Jussieu – UMR 7586

# SEMINAIRE DE STRUCTURES ALGEBRIQUES ORDONNEES

**Responsables: F. Delon, M. Dickmann, D. Gondard, T. Servi**

**2015-2016**

---

## Liste des contributions

**Franz-Viktor KUHLMANN** (University of Silesia at Katowice, Poland)
Selected Methods for the classification of cuts and their applications
(Exposé donné le 12 octobre 2015 lors du colloque au CIRM)

**Saugata BASU** (Purdue University, Indiana, USA)
A survey of quantitative bounds on the Betti numbers of real and complex varieties, and applications.
(Exposé donné le 13 octobre 2015 lors du colloque au CIRM)

**Françoise DELON** (Equipe de Logique, IMJ-PRG, Paris 7)
Un langage qui donne les boules et
Elimination des quantificateurs dans un corps valué algébriquement clos dans le langage $L_{proj}$

**Esther ELBAZ** (Equipe de logique, IMJ-PRG, Paris 7)
Anneaux de Grothendieck et fonctions de paires.

**Erik WALSBERG** (IMJ-PRG, Paris 6) et
**Philipp HIERONYMI** (Univ. of Illinois, Urbana-Champaign, USA)
On continuous functions definable in expansions of the ordered real additive group.

## Journée spéciale en l'honneur du 70ème anniversaire de F. Miraglia.

**Vincent ASTIER** et **Thomas UNGER** (Univ. College Dublin, Irlande).
Signatures, sums of hermitian squares and positive cones on algebras with involution.

**Danielle GONDARD** (IMJ-PRG, Paris 6)
Notes on valuation fans and the real holomorphy ring.

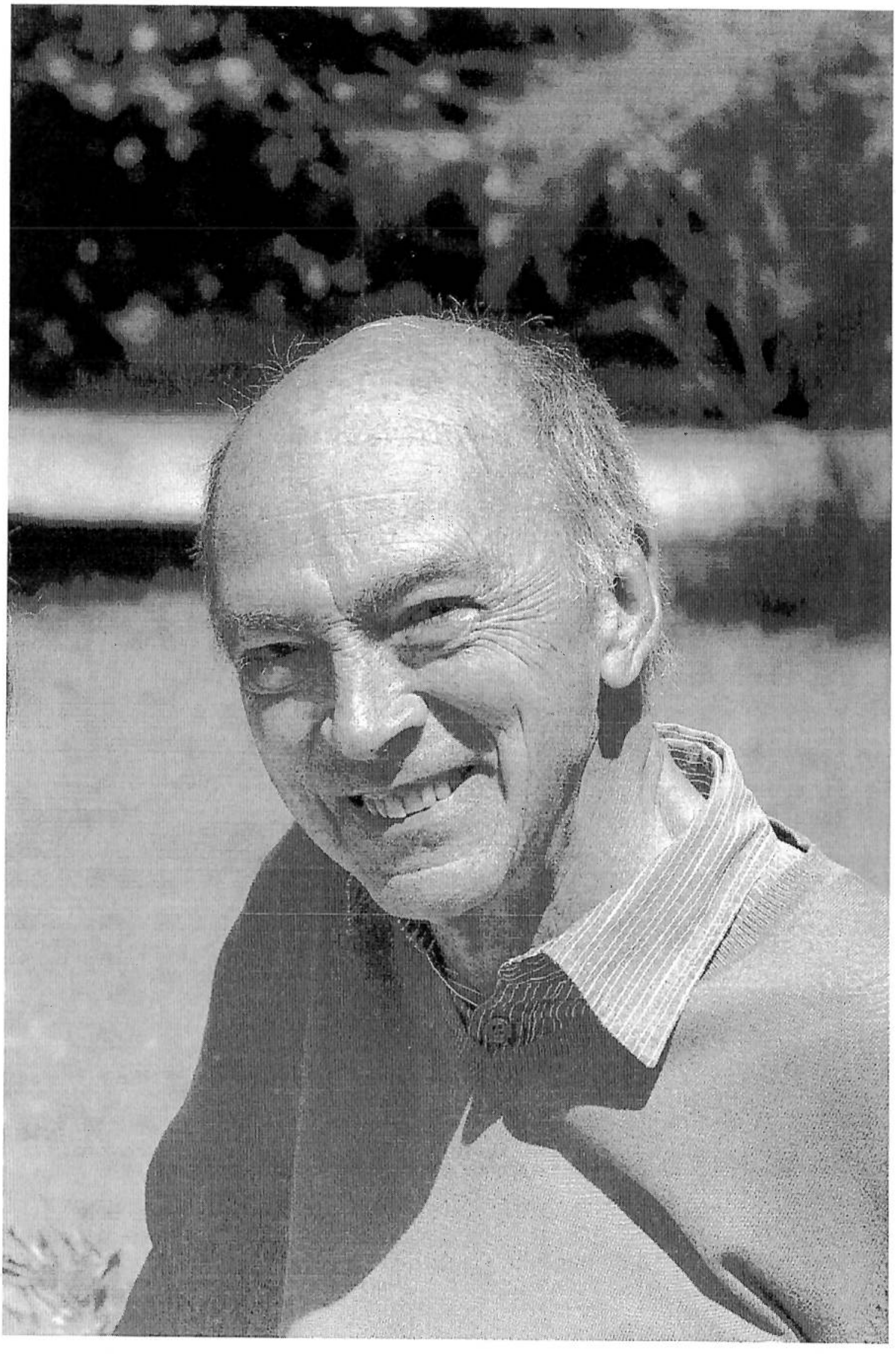
**Max DICKMANN** (IMJ-PRG, Paris 7)
Faithfully quadratic rings; an overview.

**Eberhard BECKER** (TU Dortmund, Allemagne)
A note on sums of powers in real algebraic function fields in one variable over **R**.

**Margaret THOMAS** (Fields Institute, Toronto, Canada, et Université de Constance, Allemagne)
Smooth parametrization in o-minimal structures.

**José Fernando GALVAN** et **J. M. GAMBOA** (Univ. Complutense, Madrid, Espagne)
A substitution theorem for rings of semialgebraic functions and applications.

# Hommage à François Lucas



En novembre 2009 une journée scientifique en l'honneur de François Lucas avait été organisée à Angers. Qu'il avait été doux alors d'écrire, des maths ou d'autres mots, pour François. Qu'il est triste de le faire aujourd'hui. François nous a quittés en avril 2016 et nous souhaitons lui rendre hommage dans cette édition 2015-2016 du séminaire sur les Structures Algébriques Ordonnées, séminaire dont il aura toujours été un participant assidu et un orateur régulier, et ce dès sa création en 1984. Les thèmes des recherches de François relèvent en effet tous des thématiques de ce séminaire et la chronologie de ses exposés et de ses publications scientifiques nous permet de penser que notre séminaire a rempli un de ses buts auprès de François : aider intervenants et auditeurs à mettre au point leurs idées, qu'il s'agisse de recherches personnelles, ou de la compréhension de théories ou résultats, classiques ou tout récents.

Si nous avons peut-être été utiles à François dans sa recherche, lui aura certainement joué un rôle scientifique moteur dans notre petit groupe, en organisant des journées thématiques à Angers ou des rencontres à Luminy. Il nous aura aidé par son dynamisme et sa bonne humeur, même aux heures les plus sombres de sa maladie, par sa simplicité et son intérêt pour le monde, sous tous ses aspects. Il avait toujours une exposition ou un coin des bords de Loire à nous montrer, toujours un geste ou un mot gentil.

Sa curiosité intellectuelle et sa passion pour les arts étaient immenses. Il pratiquait lui-même la gravure, au sein d'une association qu'il avait créée avec des amis, ce qui l'a amené à chercher[1] quand un local, quand une presse... Nous reproduisons à la fin de cet hommage une de ses gravures, linogravure image de sa Bretagne aimée.

Il était très actif dans le milieu associatif et syndical – il a représenté les mathématiciens au CNU autour des années 2000.

Les sites suivants comportent quelques photos :

Pour ses activités syndicales :

- http://www.angers.maville.com/actu/actudet_--l-universite-creve-et-on-veut-reformer-vite-_9-426373_actu.Htm
- http://www.snesup.fr/deces-de-francois-lucas
- http://www.lemp7.cnrs.fr/reportages/Cite-des-Sciences_France_Culture/F_Lucas_31mars2007.htm

Pour ses activités artistiques :

- http://impression-expression.com/membre/francois-lucas/
- http://apar-gravure.com/wp-content/uploads/2015/11/IMG_20150604_190408.jpg

Les actes de la rencontre d'Angers avaient été publiés par les Annales de la Faculté des Sciences de Toulouse[1]. Leur introduction contient en particulier une brève description des travaux mathématiques de François, que nous reprenons ci-dessous pour l'essentiel.

---

[1] Voir http://afst.cedram.org/afst-bin/feuilleter?id=AFST_2012_6_21_2

François soutient en 1977 sa thèse de 3ème cycle, « Équivalence élémentaire et produits », écrite sous la direction de Gabriel Sabbagh. Sa première publication, [1], en est directement issue.

Il travaille ensuite avec Françoise Delon sur les groupes abéliens ordonnées, [2], collaboration dont le travail sur les paires de groupes abéliens ordonnés divisibles reste à achever.

Puis il collabore avec Michèle Giraudet sur des structures proches : groupes cycliquement ordonnés et groupes à moitié ordonnés, [4].

François rencontre alors, à l'autre bout du monde, Daniel Gluschankof, qui va devenir son grand collaborateur et qui sera recruté maître de conférences à Angers en 1993. À Curaçao en 1988 François Lucas et Daniel Gluschankof commencent à étudier ensemble les groupes abéliens réticulés. Après avoir établi les théorèmes de représentation adéquats, ils axiomatisent des classes qui restent actuellement les plus larges connues, voir [5] , [9], [10] et aussi les articles de Daniel sur le sujet. On trouve une bibliographie de Daniel dans les actes de ce séminaire, prépublication numéro 66, publiés en Avril 1998, qui contient un hommage à Daniel.

Puis en collaboration avec Roberto Cignoli, [7], sur la question toujours ouverte de la représentabilité des spectres de $\ell$-groupes, ils réduisent sensiblement l'écart entre conditions nécessaire et suffisante.

En 1995-98, en collaboration avec Max Dickmann et Daniel Gluschankof, François étudie la version « réelle » (et purement ordinale) d'un problème classique de Kaplansky ; ils déterminent le type d'ordre des spectres réels des anneaux (commutatifs, unitaires) pour l'ordre partiel de spécialisation, [8]. Ce travail fut achevé quelques jours avant la mort brutale de Daniel en février 1998.

En 1996 François soutient son habilitation [6], « Théorie des modèles des groupes abéliens ordonnés ».

Après la disparition de Daniel Gluschankof, François entame une nouvelle collaboration et un nouveau thème de recherche. Depuis 1995, à l'initiative de Daniel Schaub, Mark Spivakovsky vient régulièrement à Angers, où il donne une série d'exposés sur la résolution des singularités dans le séminaire de Géométrie Algébrique. François, très intéressé, participe régulièrement à ce séminaire, où il est largement question de théorie des valuations. Après quelque temps et quelques exposés de part et d'autre, Mark propose un sujet qui pourrait rapprocher géomètres et théoriciens des modèles : la conjecture de Pierce-Birkhoff. Celle-ci stipule que, sur le corps des réels, toute fonction polynomiale par morceaux s'obtient en prenant la borne supérieure des bornes inférieures d'une famille finie de polynômes. James Madden en a donné une formulation purement spectrale qui permet de la considérer au-dessus d'un anneau réticulé arbitraire et non des seuls anneaux de fonctions. L'énoncé extrêmement simple de la conjecture et sa reformulation par Madden poussent François, Mark et Daniel S. à s'attaquer au sujet. L'intérêt de François pour les spectres des l-groupes, la compétence de Mark sur les valuations et les singularités, et l'intérêt de Daniel S. pour la géométrie algébrique réelle soudent cette équipe (voir [11] et [15]). Tous trois collaboreront également avec James Madden. Ils énoncent en particulier une conjecture de « connexité définissable » du spectre, [12], dont ils montrent qu'elle implique la conjecture de Pierce-Birkhoff et qui devient ainsi un critère de satisfaction de celle-ci. La technique dite des racines approchées leur permet alors de montrer que de larges classes d'anneaux satisfont Pierce-Birkhoff en dimension 2 (voir [13] et [14]).

## BIBLIOGRAPHIE SCIENTIFIQUE DE FRANÇOIS LUCAS

[15] Lucas, F., Schaub, D., Spivakovsky, M., On the Pierce-Birkhoff conjecture. J. Algebra 435 (2015), 124–158.

[14] Lucas, F., Madden, J., Schaub, D., Spivakovsky, M., Approximate roots of a valuation and the Pierce-Birkhoff conjecture. Ann. Fac. Sci. Toulouse Math. (6) 21 (2012), no. 2, 259–342.

[13] Lucas, F., Madden, J., Schaub, D., Spivakovsky, M., Introduction to the Pierce-Birkhoff conjecture and the real spectrum. (Spanish) Rev. Semin. Iberoam. Mat. 4 (2010), no. 1, 13–36.

[12] Lucas, F., Madden, J., Schaub, D., Spivakovsky, M., On connectedness of sets in the real spectra of polynomial rings. *Manuscripta Math.* 128 (2009), no. 4, 505–547.

[11] Lucas, F., Schaub, D., Spivakovsky, M., On points at infinity of real spectra of polynomial rings. Special volume in honor of Melvin Hochster. *Michigan Math. J.* 57 (2008), 587–599.

[10] Lucas, F., Lattices of antichains of a root system. General algebra and ordered sets. *Tatra Mt. Math. Publ.* 27 (2003), 177–187.

[9] Lucas, F., First-order theories of subgroups of divisible Hahn products. *Ann. Pure Appl. Logic* 121 (2003), no. 2-3, 261–279.

[8] Dickmann, M., Gluschankof, D., Lucas, F., The Order Structure of the Real Spectrum of Commutative Rings, *J. Algebra* 229 (2000), 175-204.

[7] Cignoli, R., Gluschankof, D., Lucas, F., Prime spectra of lattice-ordered abelian groups. J. Pure Appl. Algebra 136 (1999), no. 3, 217–229.

[6] Lucas, F., Théorie des modèles des groupes abéliens ordonnés ; Habilitation à diriger des recherches, Université Paris VII, juin 1996.

[5] Gluschankof, D., Lucas, F., Hyper-regular lattice-ordered groups. J. Symbolic Logic 58 (1993), no. 4, 1342–1358.

[4] Giraudet, M., Lucas, F., Groupes à moitié ordonnés. *Fund. Math.* 139 (1991), no. 2, 75–89.

[3] Lucas, F., Some applications of definable spine analysis in ordered abelian groups. *Ordered algebraic structures (Curaçao, 1988),* 123–128, Math. Appl., 55, *Kluwer Acad. Publ., Dordrecht,* 1989.

[2] Delon, F., Lucas, F., Inclusions et produits de groupes abéliens ordonnés étudiés au premier ordre. *J. Symbolic Logic* 54 (1989), no. 2, 499–511.

[1] Lucas, F., Équivalence élémentaire et produits. *C. R. Acad. Sci. Paris Sér. A-B* **287** (1978), no. 2, A41–A42.

2002 2003 DOELAN 16/30

# SELECTED METHODS FOR THE CLASSIFICATION OF CUTS, AND THEIR APPLICATIONS

F.-V. KUHLMANN

ABSTRACT. We consider four approaches to the analysis of cuts in ordered abelian groups and ordered fields, their interconnection, and various applications. The notions we discuss are: ball cuts, invariance group, invariance valuation ring, and cut cofinality.

## 1. INTRODUCTION

In these notes we deal with (Dedekind) cuts in ordered abelian groups and in ordered fields. (For the definition of the notion of a cut and other notions used in this Introduction, see Section 2.) We introduce the reader to four approaches to their classification, the links between them, and several applications. The reader should observe that a cut in an ordered field is at the same time a cut in its additive group. Hence even in the case that one is predominantly interested in cuts in ordered fields, up to a certain point their study can be fruitfully carried out in the setting of ordered abelian groups and does not need to make use of the field (or ring) multiplication. At the same time the reader should keep in mind that ordered abelian groups appear in field theory also as the value groups of valuations. In this case, cuts in the value group can for instance be generated by pseudo Cauchy sequences in the valued field. If the field is ordered and its valuation is the natural valuation induced by the ordering (see below), then it is essential to study the connection between cuts in the field and induced cuts in the value group.

The first approach to the classification of cuts is to ask whether a cut in an ordered abelian group is the upper or lower edge of a convex subgroup, or of a coset thereof. This has been used and studied more or less explicitly by many authors, and various names have been given to such cuts. We call them **ball cuts**. They appear implicitly or explicitly, sometimes with surprisingly different definitions, in [13, 14, 15, 16, 39] for the study of cuts in ordered fields, in [49] for the study of cuts in ordered abelian groups, and in several other papers cited in the references. Ball cuts will be introduced and discussed in Section 3.

Spaces of $\mathbb{R}$-places (i.e., places with residue fields embeddable in $\mathbb{R}$) of ordered fields are not well understood. It is a longstanding open problem which topological spaces appear as spaces of $\mathbb{R}$-places. Recently, ball cuts have been used to study these spaces (cf. [33, 35, 37, 26]). We describe some results in Section 3.2.

Two well known deep open problems in positive characteristic are:

1) resolution of singularities in arbitrary dimension,
2) decidability of the field $\mathbb{F}_p((t))$ of Laurent series over a finite field.

---

*Date*: 28. 2. 2017.

The author would like to thank Katarzyna Kuhlmann and Marcus Tressl for useful suggestions.

Both problems are connected with the structure theory of valued function fields of positive characteristic $p$. The main obstruction here is the phenomenon of the **defect**. Via ramification theory, the study of it can be reduced to the study of purely inseparable extensions and of Galois extensions of degree $p$. Ball cuts are essential for a classification of Galois extensions with nontrivial defect which is introduced in [30]. It will be discussed in Section 3.3.

How "broad" is a given cut? One way to answer this question is to associate to the cut the maximal set of elements that can be added to the cut sets without changing the cut. This set turns out to be a convex subgroup of the ordered group; we call it the **(additive) invariance group**. This notion was introduced by the author in his thesis ([27]) in order to handle valued fields with non-archimedean ordered value groups in connection with the model theory of valued fields. Invariance groups were also introduced by M. Tressl in his thesis ([45]), this time for the study of the model theory of ordered fields (a quick example for Tressl's use of invariance groups is given in Section 4.4). Later, A. Fornasiero and M. Mamino ([10]) used them in a detailed investigation of cuts of ordered abelian groups, which they then applied to study so-called double ordered monoids. Moreover, they have been implicitly used by several other authors (e.g. by R. Rolland in [42]), or even explicitly defined under different names (e.g. by F. Wehrung in [49], and by D. Kijima and M. Nishi in [22]). A main link to ball cuts is the fact that invariance groups can help to identify them (see Theorem 4.3 below). We will discuss invariance groups in Section 4.

In mathematics, objects that are maximal with respect to a certain property are often of particular interest. In valuation theory, this is so for **maximal valued fields**, which have the property that every proper extension will necessarily enlarge value group or residue field. Ordered abelian groups and fields carry **natural valuations** which are canonically derived from their ordering. In Section 4.2 we describe a characterization of certain ordered fields maximal with respect to their natural valuation, given by Kijima and Nishi; it makes essential use of invariance groups. Further, we discuss a generalization of this result, due to H.-J. Hüper, to the case of valuations whose valuation rings are convex under the given ordering.

In Section 4.3 we discuss an analogue for the case of ordered abelian groups. We present the Cohen-Goffman Theorem and a related result by P. Ehrlich, both of which (implicitly) use invariance groups.

The author's attention was drawn to the importance of invariance groups in the study of ordered fields by a question of J. Madden. During the Special Semester in Real Algebraic Geometry and Ordered Structures, Baton Rouge 1996, Madden showed him the definition of what we now call the **invariance valuation ring** and asked for the meaning of it. The answer to his question was first given in the manuscript [28]. Again, the invariance valuation ring was independently introduced and applied by M. Tressl (see [45, 46, 47]). The construction of invariance valuation rings appears already in Rolland's paper [42], but not in full generality.

Looking at a cut in an ordered field $K$, one may ask whether it originates in some way from a cut in the residue field of $K$ with respect to some real place. That is, one would like to know whether the cut can be translated into some "normal position" such that for some convex valuation ring $\mathcal{O}$ of $K$ with maximal ideal $\mathcal{M}$, it induces a (Dedekind) cut in the ordered residue field $\mathcal{O}/\mathcal{M}$ via the residue map.

If so, one would like to determine how this translation can be done. The invariance valuation ring is a key tool to answer these questions (see Section 5.1).

Further, in the paper [21] F. Jahnke, P. Simon and E. Walsberg introduce certain invariance valuation rings to exhibit definable valuations in ordered fields which are not dense in their real closure. The details will be discussed in Section 5.2.

A remark by M. Marshall made it clear to the author of these notes that some of his earlier results were actually a special case of a more general setting which we will now sketch. Every ordered field $K$ has a **natural valuation** $v$, whose residue field is an archimedean ordered field; its valuation ring $\mathcal{O}_v$ is the smallest of all convex valuation rings of $K$. Then every convex subgroup of the ordered additive group of $K$ is an $\mathcal{O}_v$-module.

In the present paper, we use the additive Krull notation for valuations, i.e., the ultrametric triangle law reads $v(a+b) \geq \min\{v(a), v(b)\}$ and the value group is an additively written ordered abelian group whose nonnegative elements are precisely the values of the nonzero elements of $\mathcal{O}_v$. In this notation, the map

$$M \mapsto (vK \setminus vM, vM) ,$$

where $vK$ is the value group of $(K, v)$ and $vM := \{va \mid 0 \neq a \in M\}$, is a bijection between the convex subgroups $M$ of $K$ and the cuts in the value group $vK$. This holds more generally for any (Krull) valuation $v$ of an arbitrary field $K$ and the set of all $\mathcal{O}_v$-modules $M \subseteq K$. Information about $M$ can be read off from the invariance group of the cut $(vK \setminus vM, vM)$. One can also define the invariance valuation ring of an $\mathcal{O}_v$-module. The invariance valuation ring of a cut can then be understood as the invariance valuation ring of the invariance group of the cut.

Tressl introduced the author to the definition and main properties of the **multiplicative invariance group** of a cut in an ordered field, that is, the invariance group of the cut taken in the multiplicative group of the field. For its properties, see [32], where a detailed study of ball cuts, invariance groups and invariance valuation rings is presented. Detailed studies of cuts using these concepts appear also in Tressl's papers [45, 46, 47] and T. Güldenberg's thesis [18].

After ball cuts, invariance group and invariance valuation ring, the fourth approach to the study of Dedekind cuts is to consider the pair of cardinal numbers $(\kappa, \lambda)$ where $\kappa$ is the cofinality of the lower cut set and $\lambda$ is the coinitiality of the upper cut set. Recall that the coinitiality of a linearly ordered set is the cofinality of this set under the reversed ordering. Recall further that cofinalities and coinitialities of ordered sets are regular cardinals. We call $(\kappa, \lambda)$ the **cofinality** of the cut; also the name **character** has been used in the literature.

In his groundbreaking and comprehensive work, Hausdorff constructs for any given collection of cofinalities $(\kappa, \lambda)$ which satisfy some necessary conditions, a totally ordered set where this collection is exactly the set of cofinalities of the cuts appearing in this ordering. One aim of the already cited paper [42] of Rolland is to construct ordered fields which realize a prescribed set of cut cofinalities.

A much studied property of ordered abelian groups or fields is that of being an $\eta_\alpha$-set, which is equivalent to the absence of cuts of cofinality $(\kappa, \lambda)$ with both $\kappa$ and $\lambda$ smaller than $\aleph_\alpha$. We discuss a characterization of such ordered abelian groups and fields, due to N. Alling, in Section 6.1.

In Section 6.2, we present some work of N. Yu. Galanova and G. G. Pestov which involves cut cofinalities and ball cuts.

More recently, a new aspect of cut cofinalities has been discovered. Transferring the concept of spherical completeness from ultrametric spaces to other spaces equipped with distances or topologies, the authors of [34] asked the question whether there are ordered fields, apart from the reals themselves, in which every chain of closed bounded intervals has a nonempty intersection. This happens exactly when all appearing cut cofinalities $(\kappa,\lambda)$ satisfy $\kappa\neq\lambda$. The positive answer to the question was first given by S. Shelah in the paper [43]. In joint work with Shelah and K. Kuhlmann, the author of these notes gave an alternative construction and a complete characterization of such fields in [36]. We will discuss some details in Section 6.3.

These notes are not intended to be a comprehensive survey on the general theory of cuts. However, the author hopes that they will initiate discussion and feedback so that more comprehensive information can be gathered and later be put together in a monograph on cuts.

## 2. NOTATION AND PRELIMINARIES

For general background from valuation theory, we recommend [9]. For background on ordered fields, see [38, 40].

2.1. **Cuts.** Take any ordered set $(S,<)$ (*by "ordered", we will always mean "totally ordered"*). If $S_1,S_2$ are nonempty subsets of $S$ and $a\in S$, we will write $a<S_2$ if $a<b$ for all $b\in S_2$, and we will write $S_1<S_2$ if $a<S_2$ for all $a\in S_1$.

A subset $S'$ of $S$ is called **convex in** $(S,<)$ if for every two elements $a,b\in S'$ and every $c\in S$ such that $a\leq c\leq b$, it follows that $c\in S'$. A subset $S_1$ of $S$ is an **initial segment of** $S$ if for every $a\in S_1$ and every $c\in S$ with $c\leq a$, it follows that $c\in S_1$. Symmetrically, $S_2$ is a **final segment of** $S$ if for every $a\in S_2$ and every $c\in S$ with $c\geq a$, it follows that $c\in S_2$. Note that $S_1$ is an initial segment of $S$ if and only if $S_1$ is convex and $S_1<S\setminus S_1$. Note also that $\emptyset<S$ and $S<\emptyset$ by definition; so $\emptyset$ is an initial segment as well as a final segment of $S$.

If $S_1\subseteq S$ and $S_2\subseteq S$ are such that $S_1<S_2$ and $S=S_1\cup S_2$, then we will call $(S_1,S_2)$ a **cut** in $S$. Then $S_1$ is an initial segment of $S$, $S_2$ is a final segment of $S$, and the intersection of $S_1$ and $S_2$ is empty. We write $\Lambda^L=S_1$, $\Lambda^R=S_2$, and

$$\Lambda = (\Lambda^L,\Lambda^R)\,.$$

A cut $(\Lambda^L,\Lambda^R)$ with $\Lambda^L\neq\emptyset$ and $\Lambda^R\neq\emptyset$ is called a **Dedekind cut**. If $\Lambda$ is a cut in $S$, $(T,<)$ is an extension of $(S,<)$ and $a\in T$ is such that $\Lambda^L\leq a\leq\Lambda^R$, then we will say that $a$ **realizes** $\Lambda$ (in $T$).

For any subset $M\subseteq S$, we let $M^+$ denote the cut

$$M^+ = (\{s\in S\mid \exists m\in M: s\leq m\}\,,\,\{s\in S\mid s>M\})\,.$$

That is, if $M^+=(\Lambda^L,\Lambda^R)$ then $\Lambda^L$ is the least initial segment of $S$ which contains $M$, and $\Lambda^R$ is the largest final segment having empty intersection with $M$. If $M=\emptyset$ then $\Lambda^L=\emptyset$ and $\Lambda^R=M$, and if $M=S$, then $\Lambda^L=M$ and $\Lambda^R=\emptyset$. Symmetrically, we set

$$M^- = (\{s\in S\mid s<M\}\,,\,\{s\in S\mid \exists m\in M: s\geq m\})\,.$$

That is, if $M^-=(\Lambda^L,\Lambda^R)$ then $\Lambda^L$ is the largest initial segment having empty intersection with $M$, and $\Lambda^R$ is the least final segment of $S$ which contains $M$. If $M=\emptyset$ then $\Lambda^L=M$ and $\Lambda^R=\emptyset$, and if $M=S$, then $\Lambda^L=\emptyset$ and $\Lambda^R=M$.

If $M = \{a\}$, we will write $a^+$ instead of $\{a\}^+$ and $a^-$ instead of $\{a\}^-$. These two cuts are called **principal**. Hence if $M$ has a largest element $a$, then $M^+ = a^+$ is principal, and if $M$ has a smallest element $a$, then $M^- = a^-$ is principal. The cut $(\Lambda^L, \Lambda^R)$ is principal if and only if $\Lambda^L$ has a largest element or $\Lambda^R$ has a smallest element. In the literature, a principal cut is also called **realized** or **filled**, a non-principal cut is called a **gap**, and a cut for which $\Lambda^L$ has a largest element *and* $\Lambda^R$ has a smallest element is called a **jump**. In [7], Ehrlich calls a cut $(\Lambda^L, \Lambda^R)$ **continuous** if $\Lambda^L$ is principal but not a jump.

2.2. **Valuation theory.** Take an ordered abelian group $G$. Two elements $a, b$ are **archimedean equivalent** if there is some $n \in \mathbb{N}$ such that $n|a| \geq |b|$ and $n|b| \geq |a|$. The equivalence class of $a$ is called **archimedean class of** $a$ and is denoted by $[a]$. The set $\{[a] \mid 0 \neq a \in G\}$ is totally ordered by setting $[a] < [b]$ if and only if $|a| > n|b|$ for all $n \in \mathbb{N}$. Then the class of 0 is the largest element in the set, and it only contains the element 0. The map $v :\ a \mapsto va := [a]$ is the **natural valuation** of $G$. It satisfies the triangle inequality $v(a + b) \geq \min\{va, vb\}$ and $v(-a) = va$. We call $\{va \mid 0 \neq g \in G\}$ the **value set of** $G$ **(under** $v$**)**.

If $G$ is the additive group of an ordered field $K$, then by setting $[a] + [b] := [ab]$ we obtain an addition on the set of archimedean classes that is compatible with the ordering, and the natural valuation becomes a field (Krull) valuation.

Take any extension $(L|K, v)$ of valued fields, that is, an extension $L|K$ of fields and a valuation $v$ on $L$. By $vL$ and $vK$ we denote the value groups of $v$ on $L$ and on $K$, and by $Lv$ and $Kv$ the residue fields of $v$ on $L$ and on $K$, respectively. Similarly, $vz$ and $zv$ denote the value and the residue of an element $z$ under $v$.

A valued field $(K, v)$ is called **henselian** if the extension of $v$ to every algebraic extension field $L$ of $K$ is unique, or equivalently, $(K, v)$ satisfies Hensel's Lemma. A **henselization** of $(K, v)$ is an algebraic extension of $(K, v)$ which is henselian and can be embedded over $K$ in every other henselian extension field of $(K, v)$. Henselizations exist and are unique up to valuation preserving isomorphism over $K$. Therefore, we will speak of *the* henselization of $(K, v)$ and denote it by $K^h$.

Assume that $L|K$ is finite and the extension of $v$ from $K$ to $L$ is unique. Then the Lemma of Ostrowski says that

$$(1) \qquad [L : K] \;=\; p^\nu \cdot (vL : vK) \cdot [Lv : Kv] \quad \text{with } \nu \geq 0$$

where $p$ is the **characteristic exponent** of $Kv$, that is, $p = \operatorname{char} Kv$ if this is positive, and $p = 1$ otherwise. The factor $\mathrm{d} = p^\nu$ is called the **defect** of the extension $(L|K, v)$. If $\mathrm{d} = 1$, then we call $(L|K, v)$ a **defectless extension**; otherwise, we call it a **defect extension**. Note that $(L|K, v)$ is always defectless if $\operatorname{char} Kv = 0$.

We call a henselian field $(K, v)$ a **defectless field** if every finite extension of $(K, v)$ is defectless. An arbitrary field is called a defectless field if its henselization is defectless.

The extension $(L|K, v)$ is **immediate** if for each $z \in L \setminus K$ there is $c \in K$ such that $v(z - c) > vz$; this holds if and only if the canonical embeddings of $vK$ in $vL$ and of $Kv$ in $Lv$ are onto.

For $z \in L$, we define

$$v(z - K) \;:=\; \{v(z - c) \mid c \in K\} \;\subseteq\; vL \cup \{\infty\}\,.$$

If $(L|K,v)$ is immediate, then $v(z-K)$ is a subset of $vK$ without a maximal element, and even more, it is an initial segment.

Immediate extensions of valued abelian groups can be defined as in the case of valued fields. Valued abelian groups and valued fields are called **maximal** if they do not admit proper immediate extensions.

2.3. **Pseudo Cauchy sequences.** A **pseudo Cauchy sequence** in a valued abelian group or field is a sequence $(a_\nu)_{\nu<\lambda}$ of elements, indexed by a limit ordinal $\lambda$ (which is called the **length** of the sequence), such that for all $\rho < \sigma < \tau < \lambda$,

$$v(a_\sigma - a_\rho) \; < \; v(a_\tau - a_\sigma) \,.$$

In this case, $v(a_\sigma - a_\rho) = v(a_{\rho+1} - a_\rho)$. If $(a_\nu)_{\nu<\lambda}$ is a pseudo Cauchy sequence, then the sequence of values $(v(a_{\nu+1} - a_\nu))_{\nu<\lambda}$ is strictly increasing. The set $\{b \mid \forall \nu < \lambda : v(b) > v(a_{\nu+1} - a_\nu)\}$ is called the **breadth** of the sequence $(a_\nu)_{\nu<\lambda}$. An element $a$ (in some valued extension group or field) is a **limit** of the sequence if $v(a - a_\nu) = v(a_{\nu+1} - a_\nu)$ for all $\nu < \lambda$. If $a$ is a limit of the sequence, than also $a'$ is a limit if and only if $a - a'$ is an element of the breadth.

A valued abelian group or field is called **spherically complete** if it admits a limit for every pseudo Cauchy sequence. If the valued abelian group $G'$ is an immediate extension of the valued group $G$, then every element $a \in G' \setminus G$ is the limit of a pseudo Cauchy sequence in $G$ that does not have a limit in $G$. Hence, every spherically complete valued abelian group or field is maximal. The converse is also true; in the case of valued fields this is shown by I. Kaplansky in [24], where the theory of pseudo Cauchy sequences (which he calls "pseudo-convergent sequences") is nicely laid out.

If $\alpha$ is an ordinal, then $G$ is called **$\alpha$-maximal** if every pseudo Cauchy sequence in $G$ of length less than $\aleph_\alpha$ has a limit in $G$.

2.4. **Hahn products and power series fields.** Given a linearly ordered index set $I$ and for every $\gamma \in I$ an arbitrary abelian group $C_\gamma$, we define a group called the **Hahn product** (also called **Hahn group**), denoted by $\mathbf{H}_{\gamma\in I}\, C_\gamma$. Consider the product $\prod_{\gamma\in I} C_\gamma$ and an element $c = (c_\gamma)_{\gamma\in I}$ of this group. Then the **support** of $c$ is the set $\operatorname{supp} c := \{\gamma \in I \mid c_\gamma \neq 0\}$. As a set, the Hahn product is the subset of $\prod_{\gamma\in I} C_\gamma$ containing all elements whose support is a wellordered subset of $I$, that is, every nonempty subset of the support has a minimal element. The Hahn product is a subgroup of the product group. Indeed, the support of the (componentwise) sum of two elements is contained in the union of their supports, and the union of two wellordered sets is again wellordered.

The support of every nonzero element $c$ in the Hahn product has a minimal element $\gamma_0$. This enables us to define a group valuation by setting $vc = \gamma_0$ and $v0 = \infty$; this is called the **canonical valuation** of the Hahn product $\mathbf{H}_{\gamma\in I}\, C_\gamma$.

If the $C_\gamma$ are (not necessarily archimedean) ordered abelian groups, we obtain the **ordered Hahn product**, also called **lexicographic product**, where the ordering is defined as follows. Given a nonzero element $c = (c_\gamma)_{\gamma\in I}$, let $\gamma_0$ be the minimal element of its support. Then we take $c > 0$ if and only if $c_{\gamma_0} > 0$. If all $C_\gamma$ are archimedean ordered, then the canonical valuation of the Hahn product coincides with the natural valuation of the ordered Hahn product. The **Hahn Embedding Theorem** states that every ordered abelian group $G$ can be embedded in the Hahn product with its set of archimedean classes as index sets and all $C_\gamma$ equal to the ordered group of real numbers.

Take any ordered abelian group $G$. If $H \subsetneq H'$ are convex subgroups of $G$ such that the ordering induced on $H'/H$ is archimedean (and hence $H'/H$ can be seen as an ordered subgroup of the reals), then $H'/H$ is called an **archimedean component of** $G$. If $G = \mathbf{H}_{\gamma\in I}\, C_\gamma$ and all $C_\gamma$ are archimedean ordered, then the $C_\gamma$ are precisely the archimedean components of $G$.

Take a field $k$ and an ordered abelian group $G$. Then $k((G)) := \mathbf{H}_{\gamma\in G}\, k$ is a valued abelian group. Since all supports are wellordered, a multiplication can be defined as follows: $(c_g)_{g\in G} \cdot (c'_g)_{g\in G} = (\sum_{h+h'=g} c_h \cdot c'_{h'})_{g\in G}$. Then $k((G))$ becomes a valued field, called a **power series field**. The canonical valuation of the underlying Hahn product makes it a valued field with value group $G$ and residue field $k$.

Under their canonical valuation, all Hahn products and all power series fields are spherically complete and hence maximal. All maximal fields with residue fields of characteristic 0 are power series fields, but for positive residue characteristic this is not true.

## 3. BALL CUTS

We say that a cut $\Lambda = (\Lambda^L, \Lambda^R)$ in an ordered abelian group is a **group $^+$-cut** if it is induced by the upper edge of a convex subgroup $H$ of $G$, i.e., if $\Lambda = H^+$. We will say that $\Lambda$ is a **group $^-$-cut** if it is induced by the lower edge of a convex subgroup $H$ of $G$, i.e., if $\Lambda = H^-$. In both cases, we will call $\Lambda$ a **group-cut**. Note that $0^+$ and $0^-$ are the only principal group-cuts. We call $\Lambda$ a **ball $^+$-cut** (or a **ball $^-$-cut**) if it is induced by the upper edge (or lower edge, respectively) of some coset of a convex subgroup $H$ of $G$, i.e., if it is of the form $(g+H)^+$ (or $(g+H)^-$, respectively) for some $g \in G$. Ball $^+$-cuts and ball $^-$-cuts are called **ball-cuts**, and cosets $H+g$ of convex subgroups $H$ are also called **balls**. Note that all group-cuts are ball-cuts.

Ball cuts are called **asymmetric cuts** in [13, 14, 15, 16, 39]. This name is unfortunate; it may have been chosen by the authors after they observed that there are no cuts in ordered fields that are at the same time a ball$^-$-cut and a ball$^+$-cut. But the situation is different in ordered abelian groups, as the following example shows. Consider the lexicographic ordering on $\mathbb{Z}\times\mathbb{Z}$. Then $\{(0,m) \mid m \in \mathbb{Z}\}^+ = \{(1,m) \mid m \in \mathbb{Z}\}^-$.

In Tressl's paper [47], the ball$^+$-cuts are the cuts with **signature** 1, and the ball$^-$-cuts are the cuts with signature $-1$. All non-ball cuts have signature 0. Güldenberg also uses signatures in his thesis [18], but defines them in a slightly different way.

3.1. **Monoids of cuts.** On the set of cuts in an ordered abelian group, an addition can be defined in various ways. The two immediately obvious ways to define $\Lambda_1+\Lambda_2$ are the following:

1) set $\Lambda_1 + \Lambda_2 := (\Lambda_1^L + \Lambda_2^L)^+ = \{\alpha+\beta \mid \alpha \in \Lambda_1^L, \beta \in \Lambda_2^L\}^+$,
2) set $\Lambda_1 + \Lambda_2 := (\Lambda_1^R + \Lambda_2^R)^-$.

The two additions are usually not the same, but their properties are very similar. The following fact is easy to prove:

*The idempotent elements in these monoids are precisely the group cuts.*

Monoids of cuts are studied in [49], [18], [10] and [11]. In the latter paper, the results are used for the intrinsic construction (without the use of embeddings in power series fields) of towers of complements to all (possibly fractional) ideals of the valuation ring in henselian valued fields of residue characteristic 0, and in Kaplansky fields (i.e., valued fields satisfying "Hypothesis A" in [24]) which do not admit proper immediate algebraic extensions. They are also used by N. Alling in [1] for the characterization of $\eta_\alpha$ ordered abelian groups and fields (see Section 6.1 below). Alling gives credit to A. H. Clifford ([4]) for introducing the monoid structure (but it had very probably already been observed before, when Dedekind completions of ordered abelian groups were considered).

3.2. **Applications to spaces of $\mathbb{R}$-places.** For any formally real (i.e., orderable) field $K$, the question arises which orderings induce the same natural valuations. The places associated with natural valuations are called $\mathbb{R}$-places as their residue fields are archimedean ordered and can thus be embedded in $\mathbb{R}$. We will therefore always assume that the residue field of an $\mathbb{R}$-place is a subfield of $\mathbb{R}$. The above question was answered in [35] for an interesting special case.

Take a real closed field $R$. There is a one-to-one correspondence between orderings $P$ of $R(X)$ and cuts of $R$ (see [17]). The cut $\Lambda_P = (\Lambda_P^L, \Lambda_P^R)$ corresponding to $P$ is given by $\Lambda_P^L = \{a \in R \mid a <_P X\}$ and $\Lambda_P^R = \{b \in R \mid b >_P X\}$. Conversely, if $\Lambda$ is a cut in $R$, then the set

$$P = \{f \in R(X) \mid \exists a \in \Lambda^L \; \exists b \in \Lambda^R \; \forall c \in (a,b) : \; f(c) \in \dot{R}^2\}$$

is an ordering of $R(X)$ with $\Lambda_P = \Lambda$ (here, $\dot{R} = R \setminus \{0\}$).

In [35] the following result is proved:

**Theorem 3.1.** *Two distinct orderings of $R(X)$ induce the same $\mathbb{R}$-place if and only if they correspond to the upper and lower edges of the same ball, that is, there is a convex subgroup $H$ of the additive group of $R$ and $c \in R$ such that one of the places corresponds to $(H+c)^-$ and the other to $(H+c)^+$.*

This means that the space of $\mathbb{R}$-places of $R(X)$ is obtained from the line $R$ by identifying the upper and lower edges of balls. If this is done for $R = \mathbb{R}$ then we obtain the circle (up to homeomorphism). But if $R$ is a non-archimedean ordered field, then the structure is much more complex; it may be thought of as an infinite pearl necklace in which every pearl contains a pearl necklace that is similar to the whole necklace. The rich self-similarities of this space have been exhibited in [37] by observing that the transformations $a \mapsto a+c$, $a \mapsto ca$ and $a \mapsto a^{-1}$ all transform balls into balls.

The following result is also proved in [35]:

**Theorem 3.2.** *Take an ordering on $R(X)$ which extends the ordering of $R$, and take $v$ to be the natural valuation on $R(X)$ w.r.t. this ordering. Then $X$ induces in $R$ a cut of the form $(c+H)^-$ or $(c+H)^+$ (as in Theorem 3.1) if and only if $vR \subsetneq vR(X)$, and if the former is the case, then $v(X-c)$ is rationally independent over $vR$.*

From this theorem we conclude that a cut of $R$ is a ball cut if and only if the natural value group $vR(X)$ of the corresponding ordering on $R(X)$ satisfies $[vR(X) : 2vR(X)] = 2$.

In the paper [26] P. Koprowski and K. Kuhlmann consider the more general case of an algebraic function field $F$ of transcendence degree 1 over a real closed field $R$. Choose any smooth projective model of $F$, i.e., a smooth, projective algebraic curve over $R$ with function field $F$. In [25] M. Knebusch shows that the curve consists of finitely many semialgebraic connected components, each of which can be endowed with a cyclic order. In [26] this is used to define cuts in these components; the collection of all of them is taken to be the set of cuts on the curve. The following result is proved:

**Theorem 3.3.** *The space of all cuts on the curve (endowed with the order topology) is homeomorphic to the space of all orderings on $F$ (endowed with the Harrison topology).*

Take any ordering of $F$ and let $v$ denote the natural valuation of $F$ w.r.t. this ordering. Note that the value group $vR$ of $v$ on $R$ is divisible since $R$ is real closed. Therefore, as $\operatorname{trdeg} F|K = 1$, there are only two possible cases:

a) $vF = vK$, which implies that $(vF : 2vF) = 1$,
b) $vF = vK \oplus \mathbb{Z}\alpha$ for some $\alpha \in vF \setminus vK$, whence $(vF : 2vF) = 2$.

By the Baer–Krull Theorem, in the first case there is no other ordering on $F$ that induces the same place as the given one. In the second case there is exactly one other ordering that induces the same place.

Now consider the cut that corresponds to the given ordering according to Theorem 3.3. In analogy to the case of a rational function field discussed above, the authors of [26] call this cut a **ball cut** if the second case holds. The following argument justifies this definition. Pick any element $X \in F \setminus K$. Then $F|K(X)$ is algebraic, thus $vF/vK(X)$ is a torsion group. This implies that case b) holds for $F$ with the given ordering if and only if it holds for $R(X)$ with the restriction of this ordering, as the corresponding natural valuation on $R(X)$ is just the restriction of the natural valuation on $F$. From this one obtains:

**Proposition 3.4.** *The following are equivalent:*
*1) the cut corresponding to the given ordering on $F$ is a ball cut,*
*2) for some $X \in F \setminus K$, the cut induced by $X$ in $R$ under the restriction of the ordering to $R(X)$ is a ball cut,*
*3) for each $X \in F \setminus K$, the cut induced by $X$ in $R$ under the restriction of the ordering to $R(X)$ is a ball cut.*

All results above can be obtained in an abstract setting for abstract real curves. However, once we embed the curve in an affine space we obtain a clearer picture. Note that every $n$-dimensional affine space $\mathbb{A}^n R$ over $R$ is an ultrametric space with the ultrametric generated by the natural valuation $v$ of $R$. The ultrametric distance between points $(x_1, \dots, x_n)$ and $(y_1, \dots, y_n)$ can be defined and computed as follows:

$$u((x_1, \dots, x_n), (y_1, \dots, y_n)) = \min\{v(x_i - y_i)\} = \frac{1}{2} v\left(\sum (x_i - y_i)^2\right) .$$

Therefore we can consider ultrametric balls in $\mathbb{A}^n R$. We say that an ultrametric ball $B$ in $\mathbb{A}^n R$ cuts a curve $C$ if $B \cap C \neq \emptyset$ and $(\mathbb{A}^n R \setminus B) \cap C \neq \emptyset$. In this case $B$ determines a cut (always more then one) on the curve. In [26] it is shown that such a cut is a ball cut, and the following theorem is proved:

**Theorem 3.5.** *Every ball cut on a smooth and complete real affine curve in $\mathbb{A}^n R$ is induced by some ultrametric ball. If the orderings corresponding to two ball cuts* *induce the same $\mathbb{R}$-place, then there is an ultrametric ball in $\mathbb{A}^n R$ which induces both cuts on the curve.*

The converse of the second assertion is not true, a counterexample is given in [26]. The ball mentioned in this assertion can well induce more than two cuts on the curve. It is an open question how to determine the pairs of cuts that induce the same $\mathbb{R}$-place.

3.3. **Classification of Artin-Schreier defect extensions.** An **Artin-Schreier extension** is a field extension $L|K$ of degree $p$ of fields of characteristic $p$ generated by an element $\vartheta$ that satisfies $\vartheta^p - \vartheta \in K$. Such an extension has nontrivial defect if and only if it is immediate. In this case, the cut $v(\vartheta - K)^+$ taken in the divisible hull of $vK$ enables us to distinguish two types of Artin-Schreier defect extensions. We call such an extension **dependent** if it can be derived by a transformation from a purely inseparable defect extension of degree $p$, and **independent** otherwise. In [30] the following result is proved:

**Theorem 3.6.** *An Artin-Schreier defect extension is independent if and only if the cut $v(\vartheta - K)^+$ is a group$^-$-cut.*

This classification of Artin-Schreier defect extensions is important because work by M. Temkin (see e.g. [44]) and by the author indicates that dependent defect appears to be more harmful to the above cited problems than independent defect. In the paper [6], S. D. Cutkosky and O. Piltant give an example of an extension of valued function fields consisting of a tower of two Artin-Schreier defect extensions where so-called strong monomialization fails. As the valuation on this extension is defined by use of so-called generating sequences, it is hard to determine whether the Artin-Schreier defect extensions are dependent or independent. However, Cutkosky, L. Ghezzi and S. ElHitti show that both of them are dependent (see e.g. [8]); this again lends credibility to the hypothesis that dependent defect is the more harmful one.

Moreover, the classification is an important tool in the proof of the following theorem in [30]:

**Theorem 3.7.** *A valued field is henselian and defectless if and only if each purely inseparable extension is defectless and the field does not allow any proper immediate algebraic extensions.*

This theorem in turn is used in [29] for the construction of an example showing that a certain natural axiom system for the elementary theory of $\mathbb{F}_p((t))$ ("henselian defectless valued field of characteristic $p$ with residue field $\mathbb{F}_p$ and value group a $\mathbb{Z}$-group") is not complete.

It would be desirable to have a classification of Galois defect extensions of prime degree and an analogue of Theorem 3.7 also in the case of valued fields of mixed characteristic (i.e., valued fields of characteristic 0 with residue fields of positive characteristic). The obvious problem is to find the suitable definition of "dependent", since there are no purely inseparable extensions. Some guidance can be obtained from the theory of perfectoid fields, as these allow an exchange of information between the mixed characteristic case and the case of equal positive characteristic. If we follow this guidance, then all Galois defect extensions of prime degree of perfectoid fields should be called independent. An indication that this is the right choice comes from the fact that perfectoid fields are deeply ramified fields in the sense of Section 6.6 of [12]; recent work of the author of these notes indicates that when $\vartheta$ is a suitably chosen generator of any Galois defect extension of prime degree of a deeply ramified field, then as in Theorem 3.6, $v(\vartheta - K)^+$ is a group$^-$-cut.

3.4. **Approximation of elements in henselizations.** Complete valued fields of rank 1 (i.e., with archimedean ordered value group) are henselian, but for valuations $v$ of arbitrary rank this does not hold in general. However, there is a connection between Hensel's Lemma and completions, but these completions have to be taken for residue fields of suitable coarsenings of $v$. This connection was worked out by P. Ribenboim in [41] who used **distinguished pseudo Cauchy sequences** to characterize the so called **stepwise complete** fields; it had been shown by Krull that these fields are henselian.

Take any immediate extension $(L|K, v)$ of valued fields and $z \in L \setminus K$. We call $z$ **weakly distinguished over** $K$ if $v(z - K)^+$ is a ball$^+$-cut, and we call $z$ **distinguished over** $K$ if it is a group$^+$-cut. The latter name is chosen since distinguished elements are limits of distinguished pseudo Cauchy sequences.

Now take an arbitrary valued field $(K, v)$ and extend its valuation $v$ to its algebraic closure $\tilde{K}$. Then $\tilde{K}$ contains a unique henselization $K^h$ with respect to this extension. The following result is proved in [31], answering a question from B. Teissier. It has recently been reproven by Teissier using methods from algebraic geometry.

**Theorem 3.8.** *Each element in $K^h \setminus K$ is weakly distinguished over $K$.*

Note that if $(K, v)$ is of rank 1, then its henselization lies in its completion and every element $a \in K^h \setminus K$ is distinguished over $K$ (with $v(a - K)^+ = (vK)^+$).

By "$\alpha > v(a - K)$" we mean $\alpha > v(a - c)$ for all $c \in K$. Theorem 3.8 is used in [31] to prove the following result:

**Theorem 3.9.** *Take $z \in \tilde{K} \setminus K$ such that*

$$v(a - z) \;>\; v(a - K)$$

*for some $a \in K^h$. Then $K^h$ and $K(z)$ are not linearly disjoint over $K$, that is,*

$$[K^h(z) : K^h] < [K(z) : K]$$

*and in particular, $K(z)|K$ is not purely inseparable.*

This theorem has a crucial application in [30] to the classification of Artin-Schreier defect extensions which we discussed in the previous section. The classification was originally obtained in [27] under the additional assumption that the fields in question are henselian. With the help of Theorem 3.9 this assumption can be dropped, and so the classification becomes available for valued function fields.

## 4. THE INVARIANCE GROUP

For every cut $\Lambda$ in an ordered abelian group $G$, we define

$$\mathcal{G}(\Lambda) \;:=\; \{g \in G \mid \Lambda^L + g = \Lambda^L\}$$

and call it the **invariance group** of $\Lambda$; other authors (e.g. Ehrlich in [7], following Kijima and Nishi [22]) call it the **breadth** of the cut $\Lambda$. Note that $\Lambda^L + g = \Lambda^L$ is equivalent to $\Lambda^R + g = \Lambda^R$.

The proof of the following facts is straightforward (see e.g. [32]).

**Lemma 4.1.** *Take an ordered abelian group $G$ and a Dedekind cut $\Lambda$ in $G$. Then $\mathcal{G}(\Lambda)$ is a convex subgroup of $G$, and $G$ is the disjoint union of the three convex subsets $\Lambda^L - \Lambda^R$, $\mathcal{G}(\Lambda)$ and $\Lambda^R - \Lambda^L$, with*

$$\Lambda^L - \Lambda^R < \mathcal{G}(\Lambda) < \Lambda^R - \Lambda^L .$$

**Corollary 4.2.** *The invariance group of $\Lambda$ is trivial if and only if*

$$\Lambda^R - \Lambda^L = G^{>0} .$$

The following theorem is proved in [32], but has also been stated (more or less explicitly) by other authors:

**Theorem 4.3.** *A cut $\Lambda$ in an ordered abelian group is a ball cut if and only if it is the upper or lower edge of a coset of its invariance group, i.e., if there is some $g \in G$ such that $\Lambda = (g + \mathcal{G}(\Lambda))^+$ or $\Lambda = (g + \mathcal{G}(\Lambda))^-$.*

**Remark 4.4.** R. Baer ([3]) introduced the notion **eigentlicher Schnitt**, that is, a Dedekind cut $\Lambda$ in an ordered abelian group $G$ such that for every positive $g \in G$ there are $a \in \Lambda^L$ and $b \in \Lambda^R$ such that $b - a < g$. By Lemma 4.1 this condition is equivalent to the invariance group of $G$ being trivial. Ehrlich ([7]) calls them **Veronese cuts**, and Galanova and Pestov call them **fundamental cuts**. The non-principal cuts with trivial invariance group are called **dense** in Tressl's papers [46, 47].

Several authors, e.g. Rolland in [42] and Wehrung in [49], work with the ball$^+$-cuts $\mathcal{G}(\Lambda)^+$ rather than the invariance groups themselves. The set of all of these cuts in an ordered abelian group $G$ coincides with the set of cuts $H^+$ where $H$ runs through all (proper) convex subgroups of $G$. (Note that $G$ itself is the invariance group of the two cuts $(G, \emptyset)$ and $(\emptyset, G)$, which are not Dedekind cuts.)

If $\Lambda$ is a cut in an ordered field $K$ and is positive (i.e., $0 \in \Lambda^L$), then it is also a cut in the ordered abelian multiplicative group of positive elements of $K$. Its invariance group there is called the **multiplicative invariance group of** $\Lambda$, and we denote it by $\mathcal{G}^\times(\Lambda)$.

### 4.1. Invariance group and pseudo Cauchy sequences.

From what we have said about immediate extensions and pseudo Cauchy sequences in Section 2.3, ordered abelian groups or fields that are maximal with respect to their natural valuation contain limits for all pseudo Cauchy sequences. This is why several authors employ pseudo Cauchy sequences to study and to characterize such ordered abelian groups or fields. Certain cuts can induce, or be induced by, pseudo Cauchy sequences. For example, if $(a_\nu)_{\nu<\lambda}$ is a pseudo Cauchy sequences which is also strictly increasing, then it is cofinal in the lower cut set of the cut $\Lambda = \{a_\nu \mid \nu < \lambda\}^+$, and the following holds:

**Theorem 4.5.** *The invariance group of $\{a_\nu \mid \nu < \lambda\}^+$ is equal to the breadth of the pseudo Cauchy sequence $(a_\nu)_{\nu<\lambda}$.*

If the pseudo Cauchy sequence lies in an ordered abelian group $G$, then it induces a Cauchy sequence (i.e., a pseudo Cauchy sequence with breadth $\{0\}$) in $G/\mathcal{G}(\Lambda)$.

**4.2. Ordered fields with maximal natural valuation.** Take an ordered field $K$. We will denote the ordered additive group of $K$ by $K_+$. In [22], Kijima and Nishi use the invariance group for the following result:

**Theorem 4.6.** *The following assertions are equivalent:*

*1) the natural valuation of $K$ is maximal and its residue field is $\mathbb{R}$,*

*2) for each cut $\Lambda = (\Lambda^L, \Lambda^R)$ in $K$, the induced cut $(\Lambda^L/\mathcal{G}(\Lambda), \Lambda^R/\mathcal{G}(\Lambda))$ in the ordered abelian group $K_+/\mathcal{G}(\Lambda)$ is principal.*

Here, the induced cut is $(\{a/\mathcal{G}(\Lambda) \mid a \in \Lambda^L\}, \{b/\mathcal{G}(\Lambda) \mid b \in \Lambda^R\})$; note that the two sets are disjoint by the defining property of $\mathcal{G}(\Lambda)$.

This theorem also holds for any ordered abelian group $G$ in place of the ordered field $K$ if we replace "its residue field is $\mathbb{R}$" by "all of its archimedean components are isomorphic to $\mathbb{R}$"; see the next section.

In his thesis [20] Hüper considers ordered fields with arbitrary compatible valuations (i.e., valuations whose valuation ring is convex, or equivalently, contains the valuation ring of the natural valuation). We will cite one of his main results; in its formulation he uses a notion that is derived from Baer's "eigentlicher Schnitt" (see Remark 4.4) without explicitly using invariance groups. But using them as follows puts the result in a wider context:

**Theorem 4.7.** *Take an ordered field $K$ with a compatible valuation $v$. Then the following assertions are equivalent:*

*1) the valuation $v$ is maximal,*

*2) if $H$ is a $\mathcal{O}_v$-submodule of $K$ not contained in a larger $\mathcal{O}_v$-submodule $H'$ such that there is no $\mathcal{O}_v$-submodule properly between $H'$ and $H$, and if $\Lambda$ is a cut such that $\Lambda/H := (\Lambda^L/H, \Lambda^R/H)$ is a Dedekind cut in $K_+/H$ with trivial invariance group, then $\Lambda/H$ is principal.*

Let us evaluate this theorem for the case of $v$ being the natural valuation. In this case, condition 2) can be reformulated as follows:

*2') if $H$ is a convex subgroup of $K$ which is not contained in a larger convex subgroup $H'$ such that $H'/H$ is archimedean ordered, and if $\Lambda$ is a cut such that $\Lambda/H := (\Lambda^L/H, \Lambda^R/H)$ is a Dedekind cut in $K_+/H$ with trivial invariance group, then $\Lambda/H$ is principal.*

Condition 2') can be further reformulated and thereby simplified by using the following two facts:

**Lemma 4.8.** *Take a Dedekind cut $\Lambda$ in an ordered abelian group $G$ and a proper convex subgroup $H$ of $G$. Then the following assertions hold.*

*a) $\Lambda/H$ is a Dedekind cut in $G/H$ if and only if $H \subsetneq \mathcal{G}(\Lambda)$.*

*b) $\mathcal{G}(\Lambda/H) = \{0\}$ if and only if $\mathcal{G}(\Lambda) \subseteq H$.*

In view of these facts, condition 2) is equivalent to:

*2") if $H$ is a convex subgroup of $K_+$ which is not contained in a larger convex subgroup $H'$ such that $H'/H$ is archimedean ordered, and if $\Lambda$ is a cut with invariance group $H$, then $\Lambda/H$ is principal.*

What is the role of the assumption on $H$ in conditions 2') and 2")? Well, if $H'$ is a larger convex subgroup of $K_+$ such that $H'/H$ is archimedean ordered, then $H'/H$ is an archimedean component of $K_+$ and therefore isomorphic to the additive group of $Kv$. As the theorem does not assume that the latter is equal to $\mathbb{R}$, $H'/H$ may have a non-principal Dedekind cut, which then gives rise to a non-principal Dedekind cut of $G/H$. So if we are only interested in maximality, then we have to take this case into account. However, if we assume in addition to condition 1) that $Kv$ is equal to $\mathbb{R}$, then all Dedekind cuts in archimedean components are principal, and we can drop the assumption on $H$. This shows that Theorem 4.6 is a consequence of Theorem 4.7.

**Remark 4.9.** In [23], the authors state that "the notion of maximal ordered fields was first introduced" in earlier papers of theirs, the earliest published in 1987 by Kijimi and Nishi. This statement is correct only as far as it concerns results published in journals, as the work of Hüper shows.

4.3. **Archimedean complete ordered abelian groups.** An ordered abelian group $(G,<)$ is called **archimedean complete** (a notion introduced by H. Hahn) if every proper ordered abelian group extension $(G',<)$ of $(G,<)$ introduces new archimedean classes, or in other words, the natural valuation on $(G',<)$ has a larger value set than on $G$. Hence the archimedean complete ordered abelian groups are precisely the ordered abelian groups that are maximal w.r.t. their natural valuation and whose archimedean components are as large as possible, that is, isomorphic to the additive group of real numbers. Hahn shows in [19] that archimedean complete ordered abelian groups are precisely the ones that admit an order preserving isomorphism onto a so-called Hahn product with all of its archimedean components equal to the additive group of real numbers. (Hahn products are the analogues for ordered abelian groups of the power series fields.)

Archimedean complete ordered abelian groups $G$ are characterized in the paper [5] by L. W. Cohen and C. Goffman as follows:

**Theorem 4.10.** *An ordered abelian group $(G,<)$ is archimedean complete if and only if for every proper convex subgroup $H$, the ordered factor group $G/H$ is dense and every cut in $G/H$ with trivial invariance group is principal.*

Ehrlich revisits this topic in [7]. Relying on the Theorem of Cohen and Goffman, Ehrlich proves:

**Theorem 4.11.** *An ordered abelian group $(G,<)$ is archimedean complete if and only if for every cut $\Lambda$, the induced cut in $G/\mathcal{G}(\Lambda)$ is principal, but not a jump.*

Since ordered fields admit no jumps, this theorem can be seen as an analogue of Theorem 4.6. Ehrlich shows that the induced cut has trivial invariance group; this is a special case of part b) of Lemma 4.8.

We recommend Ehrlich's paper [7] for interesting historical remarks and a detailed list of references.

4.4. **Model theory of ordered fields with cuts.** In [45, 46, 47], Tressl studies the model theory of real closed fields with a fixed cut. Given a model $M$ of an o-minimal extension $T$ of the theory of real closed fields in a language $\mathcal{L}$, he determines the model theoretic properties of $M$ in the language $\mathcal{L}(\mathcal{D})$ where $\mathcal{D}$ is a predicate for the left cut set $\Lambda^L$ of a fixed cut $\Lambda$ in $M$. If $(M_1,\Lambda_1^L)$ and $(M_2,\Lambda_2^L)$ are two structures obtained in this way, conditions are found for $(M_1,\Lambda_1^L)$ and $(M_2,\Lambda_2^L)$ to be elementarily equivalent in the language $\mathcal{L}(\mathcal{D})$ enhanced by parameters from a common elementary substructure of $M$ and $M'$. The main result is rather technical in nature, but for special classes of cuts, the situation is much easier. To illustrate this, the following theorem is taken from [47]:

**Theorem 4.12.** *Let $A \prec M_1, M_2$ be models of $T$ and let $\Lambda_1, \Lambda_2$ be non-principal cuts in $M_1, M_2$ respectively, with trivial invariance groups. Then $(M_1, \Lambda_1^L) \equiv_A (M_2, \Lambda_2^L)$ if and only if the restrictions of $\Lambda_1$ and $\Lambda_2$ to $A$ coincide.*

## 5. THE INVARIANCE VALUATION RING

The **invariance valuation ring of a cut** $\Lambda$ in an ordered field $K$ is defined as

$$\mathcal{O}(\Lambda) := \{b \in K \mid b\,\mathcal{G}(\Lambda) \subseteq \mathcal{G}(\Lambda)\}.$$

We denote its maximal ideal $\{b \in K \mid b\,\mathcal{G}(\Lambda) \subsetneqq \mathcal{G}(\Lambda)\}$ by $\mathcal{M}(\Lambda)$.

According to Lemma 4.1 $\mathcal{G}(\Lambda)$ is a convex subgroup of the ordered additive group $K_+$ of $K$, and we have already noted in the Introduction that every convex subgroup of $K_+$ is an $\mathcal{O}_v$-module, where $v$ denotes the natural valuation. In this way, the above definition becomes a special case of the following.

Take any valued field $K$ with valuation ring $\mathcal{O}_v$ and an $\mathcal{O}_v$-module $M$ in $K$. The **invariance valuation ring of an $\mathcal{O}_v$-module** $M$ in $K$ is defined as

$$\mathcal{O}(M) := \{b \in K \mid bM \subseteq M\}.$$

The relation between multiplicative invariance group and invariance valuation ring is determined in [32]. Also Tressl and Güldenberg obtain results on this topic.

### 5.1. Projecting cuts into residue fields.

Take a convex valuation ring $\mathcal{O}$ of an ordered field $K$, with maximal ideal $\mathcal{M}$. Its residue field $\mathcal{O}/\mathcal{M}$ is again an ordered field, with the ordering induced through the residue map. We will say that the cut $\Lambda$ **can be projected into the residue field** $\mathcal{O}/\mathcal{M}$ if there are elements $a, c \in K$ such that $c > 0$ and $c\Lambda + a$ induces a Dedekind cut

$$\left( ((c\Lambda^L + a) \cap \mathcal{O})/\mathcal{M}\,,\ ((c\Lambda^R + a) \cap \mathcal{O})/\mathcal{M} \right) \tag{2}$$

in $\mathcal{O}/\mathcal{M}$ via the residue map.

The following theorem shows for which convex valuation rings $\mathcal{O}$ a cut can be projected into the associated residue field. For a proof, see [32].

**Theorem 5.1.** *1) Take any convex valuation ring $\mathcal{O}$ of $(K, <)$. If $\mathcal{O}(\Lambda) \subsetneqq \mathcal{O}$, then the cut $\Lambda$ can be projected into the residue field $\mathcal{O}/\mathcal{M}$. If $\mathcal{O} \subsetneqq \mathcal{O}(\Lambda)$, then it cannot be projected into $\mathcal{O}/\mathcal{M}$.*
*2) The cut $\Lambda$ can be projected into $\mathcal{O}(\Lambda)/\mathcal{M}(\Lambda)$ if and only if $(v\mathcal{G}(\Lambda))^-$ is a ball$^+$-cut.*

### 5.2. Definable valuation rings in ordered fields.

It is obvious that if a cut (that is, its lower cut set $\Lambda^L$) is definable in some extension of the language of ordered rings, then so is $\mathcal{G}(\Lambda)$. It then follows that also the invariance valuation ring is definable.

This observation is put to work in [21], where the following is proved:

**Proposition 5.2.** *Take an ordered field $K$ with real closure $R$. If $K$ is not dense in $R$, then $K$ admits a nontrivial valuation definable in the language of ordered rings.*

The idea of the proof is as follows. If $K$ is not dense in $R$, then there is an element $r \in R \setminus K$ and a positive element $a \in R$ such that $|r - c| > a$ for all $c \in K$. Since $R|K$ is algebraic, the set $K^{>0}$ of positive elements in $K$ is coinitial in $R^{>0}$, so we can choose $a \in K$. If we set $\Lambda^L = \{c \in K \mid c < r\}$, then we obtain a cut $\Lambda$ such that $\Lambda^R - \Lambda^L \subsetneq K^{>0}$. By Corollary 4.2, its invariance group is thus nontrivial. This implies that the invariance valuation ring is not all of $K$, so the associated valuation is nontrivial. Since $r$ lies in a real closure of $K$, the set $\Lambda^L = \{c \in K \mid c < r\}$ is definable, and by what we said above, so are the invariance valuation ring and thus also the associated valuation.

The above arguments also prove the following general principle:
*If the lower cut set of some Dedekind cut with nontrivial invariance group in an ordered field is definable, then the field admits a nontrivial definable valuation ring.*

Similarly, if an $\mathcal{O}_v$-module $M$ in a valued field $(K,v)$ is definable, then so is its invariance valuation ring $\mathcal{O}(M)$. If $K \neq M \neq \{0\}$, then $\mathcal{O}(M)$ is a nontrivial valuation ring. This yields the following general principle:
*If a proper nontrivial $\mathcal{O}_v$-module in a valued field $(K,v)$ is definable, then the field admits a nontrivial definable valuation ring containing $\mathcal{O}_v$.*

Note that if a cut $\Lambda$ is definable in the value group $vK$ in a suitable language of valued fields, then the $\mathcal{O}_v$-module $\{a \in K \mid va \in \Lambda^R\}$ is also definable, and it is proper and nontrivial if and only if the cut is a Dedekind cut.

## 6. CUT COFINALITIES

Recall that by the **cofinality** of the cut $\Lambda$ we mean the pair $(\kappa, \lambda)$ where $\kappa$ is the cofinality of $\Lambda^L$, and $\lambda$ is the coinitiality of $\Lambda^R$.

### 6.1. Ordered abelian groups and fields that are $\eta_\alpha$-sets.

Take any ordinal $\alpha$. An $\eta_\alpha$-set is an ordered set $S$ such that for any two subsets $A \subseteq S$ and $B \subseteq S$ of cardinality less than $\aleph_\alpha$ with $A < B$, there is some $s \in S$ such that $A < s < B$. This is equivalent to saying that $S$ does not admit any cuts of cofinality $(\kappa, \lambda)$ where both $\kappa$ and $\lambda$ are smaller than $\aleph_\alpha$. In [1], Alling proves:

**Theorem 6.1.** *a) An ordered abelian group is an $\eta_\alpha$-set if and only if it is $\alpha$-maximal, its value set w.r.t. the natural valuation is an $\eta_\alpha$-set, and all of its archimedean components are isomorphic to $\mathbb{Z}$ or $\mathbb{R}$.*

*b) An ordered field is an $\eta_\alpha$-set if and only if it is $\alpha$-maximal, its value group w.r.t. the natural valuation is an $\eta_\alpha$-set, and its residue field is $\mathbb{R}$.*

Every $\aleph_\alpha$-saturated ordered abelian group or field is an $\eta_\alpha$-set. For the converse, the reader may note that divisible ordered abelian groups and real closed fields are o-minimal. This implies that for them, the property of being an $\eta_\alpha$-set is equivalent to that of being $\aleph_\alpha$-saturated. For results on the variety of $\eta_\alpha$ ordered abelian groups or fields, for fixed $\alpha$, see [2].

Ball cuts and invariance groups do not appear explicitly in [1] or [2]. But Rolland draws a connection in [42]. He states that an ordered abelian group is an $\eta_\alpha$-set if and only if its value set is an $\eta_\alpha$-set and for every Dedekind cut $\Lambda$ with nontrivial invariance group $\mathcal{G}(\Lambda$, the coinitiality of the upper cut set of $\mathcal{G}(\Lambda)^+$ is not less than $\aleph_\alpha$.

6.2. **Cuts in ordered power series fields.** The papers [13, 14, 15, 16] of Galanova and Pestov are devoted to the study of cuts in power series fields and in restricted power series fields (in the latter, the cardinalities of the supports of the power series are bounded by a given cardinal number). We cite three theorems from [16]. The cardinality of a set $S$ is denoted by $|S|$, and $|S|^+$ denotes its successor cardinal.

**Theorem 6.2.** *Take any ordered abelian group $G$. Then all cuts in the power series field $\mathbb{R}((G))$ are ball cuts.*

The proof of this theorem in [16] is long and technical. Let us give the sketch of a shorter and more conceptual proof. We write $K = \mathbb{R}((G))$. Every cut $\Lambda$ in $K$ is realized in some ordered field extension $L$ of $K$ (for instance, if $L$ is a $|K|^+$-saturated elementary extension of the ordered field $K$). As a power series field, $K$ is maximal w.r.t. its natural valuation. Extend $v$ to the natural valuation of $L$. Then it follows that for every $x \in L \setminus K$ there is some $a \in K$ such that $v(x-a) = \max\{v(x-c) \mid c \in K\}$ since otherwise, $x$ would be a limit of some pseudo Cauchy sequence in $K$ without a limit in $K$, contradicting the fact that $K$ is maximal. The value $v(x-a)$ can only be maximal if either $v(x-a) \notin vK$ or there is $d \in K$ such that $vd = v(x-a)$ and $d^{-1}(x-a)v \notin Kv$. But the latter cannot be the case: since $Kv = \mathbb{R}$ and $v$ on $L$ is a natural valuation, we must have that $Lv = Kv$. Hence $\gamma := v(x-a) \notin vK$. We leave it as an exercise to the reader to show that $\Lambda = (a + \{b \in K \mid vb > \gamma\})^+$ or $\Lambda = (a + \{b \in K \mid vb > \gamma\})^-$.

**Theorem 6.3.** *Take any ordered abelian group $G$ and a cardinal number $\kappa$ such that $\aleph_0 < \kappa < |G|$. Denote by $\mathbb{R}((G,\kappa))$ the subfield of $\mathbb{R}((G))$ consisting of all power series with support of cardinality less than $\kappa$. Take a non-ball cut in $\mathbb{R}((G,\kappa))$ of cofinality $(\lambda,\lambda)$. Then $\lambda$ is equal to the cofinality of $\kappa$. In particular, if $\kappa$ is regular, then $\lambda = \kappa$.*

**Theorem 6.4.** *Take a non-principal cut $\Lambda$ in some ordered field $K$ with trivial invariance group, and let $(\kappa,\kappa)$ be its cofinality. Then $\kappa$ is equal to the cofinality of $K$.*

In [42], Rolland states the existence of power series fields that admit cuts with preassigned cofinalities $(\kappa_i,\lambda_i)$, $i \in I$, where the $\kappa_i$ and $\lambda_i$ are infinite regular cardinals. The proof he gives is insufficient, but the result also follows from the work we will discuss in the next section. He uses it to show the existence and to (partially) characterize the ordered fields which admit a closed bounded interval and a continuous function which is unbounded on this interval.

6.3. **Symmetrically complete ordered abelian groups and fields.** A Dedekind cut with cofinality $(\kappa,\lambda)$ is called **symmetric** if $\kappa = \lambda$, and **asymmetric** otherwise. Note that the notion of symmetry used by Galanova and Pestov is quite different from the one defined here. However, Pestov states in [39], without proof or reference) that if a cut in an ordered field is symmetric in their sense (i.e., it is not a ball cut), then it is also symmetric in the sense of the above definition. A proof is given by Galanova in [13]. In [42], Rolland states the same in full generality for non-ball cuts in ordered abelian groups. The statement is correct, but his proof appears to have a serious gap.

A linearly ordered set $(I,<)$ is called **symmetrically complete** if every symmetric cut in $I$ has cofinality $(1,1)$, i.e., is a jump. In dense linear orderings (and hence in ordered fields) there are no jumps. Consequently, a dense linear ordering is symmetrically complete if and only if all of its cuts are asymmetric.

For example, $\mathbb{Z}$ and $\mathbb{R}$ are symmetrically complete, but $\mathbb{Q}$ is not. In $\mathbb{Z}$ and $\mathbb{R}$, every Dedekind cut is principal; in $\mathbb{Z}$ all of them have cofinality $(1,1)$, and in $\mathbb{R}$ they have cofinalities $(1,\aleph_0)$ and $(\aleph_0,1)$. In contrast, in $\mathbb{Q}$ the Dedekind cuts have cofinalities $(1,\aleph_0)$, $(\aleph_0,1)$ and $(\aleph_0,\aleph_0)$.

In [36] it is shown that a symmetrically complete ordered abelian group is spherically complete w.r.t. its natural valuation and hence a Hahn product, with all of its archimedean components isomorphic to $\mathbb{R}$. Similarly, a symmetrically complete ordered field is spherically complete w.r.t. its natural valuation and hence a power series field, with residue field $\mathbb{R}$. For Hahn products with all of its archimedean components isomorphic to $\mathbb{R}$ the set of all cut cofinalities is computed from the set of all cut cofinalities of the value set of its natural valuation, and a similar computation is done for power series fields with residue field $\mathbb{R}$. Based on this computation, a full characterization of symmetrically complete ordered abelian groups and fields is obtained. We will cite a selection of the main results.

We call an ordered set **strongly symmetrically complete** if it is symmetrically complete and does not have any cuts with cofinalities $(1,\aleph_0)$ or $(\aleph_0,1)$.

**Theorem 6.5.** *A non-archimedean ordered field is symmetrically complete if and only if it is spherically complete w.r.t. its natural valuation, has residue field* $\mathbb{R}$ *and a dense strongly symmetrically complete value group.*

*A nontrivial densely ordered abelian group is symmetrically complete if and only if it is spherically complete w.r.t. its natural valuation* $v$*, has a dense strongly symmetrically complete value set, and all archimedean components are isomorphic to* $\mathbb{R}$*. It is strongly symmetrically complete if and only if in addition, the value set has uncountable cofinality.*

In particular, it follows that symmetrically complete ordered abelian groups are divisible and symmetrically complete ordered fields are real closed.

Further, it is shown in [36] that every ordered set can be extended to a dense strongly symmetrically complete ordered set with uncountable cofinality. The reader may note that this result is not explicitly stated in Hausdorff's work. The authors of [36] also tried to give a construction that is as short as possible. It turns out that the constructed orderings are themselves lexicographic, as are the orderings on Hahn products and power series fields. Such orderings deserve to be studied in more detail.

Using the above results and the fact that every ordered set is the natural value set of some Hahn product with all components isomorphic to $\mathbb{R}$, and every ordered abelian group is the natural value group of some power series field with residue field $\mathbb{R}$, the following result of [43] is reproved:

**Theorem 6.6.** *Every ordered field can be embedded in a symmetrically complete ordered field. Every ordered abelian group can be embedded in a symmetrically complete ordered abelian group.*

## REFERENCES

[1] Alling, N. L.: *On the existence of real-closed fields that are $\eta_\alpha$-sets of power $\aleph_\alpha$*, Trans. Amer. Math. Soc. **103** (1962), 341–352

[2] Alling, N. L. – Kuhlmann, S.: *On $\eta_\alpha$-groups and fields*, Order **11** (1994), 85–92

[3] Baer, R.: *Archimedizität und Starrheit geordneter Körper*, Math. Ann. **188** (1970), 165–205

[4] Clifford, A. H.: *Completion of semi-continuous ordered commutative semigroups*, Duke Math. J. **26** (1959), 41–59

[5] Cohen, L. W. – Goffman, C.: *On completeness in the sense of Archimedes*, Amer. J. Math. **72** (1950). 747–751

[6] Cutkosky, D. – Piltant, O.: *Ramification of valuations*, Adv. Math. **183** (2004), 1–79

[7] Ehrlich, P.: *Dedekind cuts of Archimedean complete ordered abelian groups*, Algebra Universalis **37** (1997), 223–234

[8] ElHitti, S. – Ghezzi, L.: *Dependent Artin-Schreier defect extensions and strong monomialization*, J. Pure Appl. Algebra **220** (2016), 1331–1342

[9] Engler, A. J. – Prestel, A.: *Valued fields*, Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2005

[10] Fornasiero, A. – Mamino, M.: *Arithmetic of Dedekind cuts of ordered abelian groups*, Ann. Pure Appl. Logic **156** (2008), 210–244

[11] Fornasiero, A. – Kuhlmann, F.-V. – Kuhlmann, S.: *Towers of complements to valuation rings and truncation closed embeddings of valued fields*, J. Algebra **323** (2010), 574–600

[12] Gabber, O. – Ramero, L.: *Almost ring theory*, Lecture Notes in Mathematics **1800**, Springer-Verlag, Berlin, 2003

[13] Galanova, N. Yu.: *Symmetry of sections in fields of formal power fields and on the nonstandard real line*, Algebra Logika **42** (2003), 26–36, 125; translation in Algebra Logic **42** (2003), 4–20

[14] Galanova, N. Yu.: *Symmetric and asymmetric gaps in some fields of formal power series*, Serdica Math. J. **30** (2004), 495–504

[15] Galanova, N. Yu.: *An investigation of the fields of bounded formal power series by means of theory of cuts*, Acta Appl. Math. **85** (2005), 121–126

[16] Galanova, N. Yu. – Pestov, G. G.: *Symmetry of sections in fields of formal power series*, Algebra Logika **47** (2008), 174–185, 265; translation in Algebra Logic **47** (2008), 100–106

[17] Gilmer, R.: *Extension of an order to a simple transcendental extension*, Contemp. Math. **8** (1981), 113–118

[18] Güldenberg, T.: *Elementare Invarianten von Dedekindschnitten angeordneter Körper*, Diplomarbeit (MSc thesis), Universität Regensburg, 2004; available at: `http://math.usask.ca/fvk/asom3.pdf`

[19] Hahn, H.: *Über die nichtarchimedischen Grössensysteme*, Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften, Wien, Mathematisch-Naturwissenschaftliche Klasse 116 (Abteilung IIa) (1907), 601–655

[20] Hüper, H.-J.: *Über ordnungsverträglich bewertete, angeordnete Körper*, Dissertation, Univ. München, Munich, 1977

[21] Jahnke, F. – Simon, P. – Walsberg, E.: *dp-minimal valued fields*, preprint

[22] Kijima, D. – Nishi, M.: *The pseudo-convergent sets and the cuts of an ordered field*, Hiroshima Math. J. **19** (1989), 89–98

[23] Kijima, D. – Nishi, M. – Sakaibara, M.: *Maximal extensions of ordered fields*, Hiroshima Math. J. **18** (1988), 485–492

[24] Kaplansky, I.: *Maximal fields with valuations I*, Duke Math. Journ. **9** (1942), 303–321

[25] Knebusch, M: *On algebraic curves over real closed fields. I* and *II*, Math. Z. **150** (1976), 49–70 and **151** (1976), 189–205

[26] Koprowski, P. – Kuhlmann, K.: *Places, Cuts and Orderings of Function Fields*, to appear in J. Alg.

[27] Kuhlmann, F.-V.: *Henselian function fields and tame fields*, extended version of doctoral thesis, Heidelberg (1990)

[28] Kuhlmann, F.-V.: *Invariance valuation ring of cuts in ordered fields*, manuscript, Baton Rouge, (1996)

[29] Kuhlmann, F.-V.: *Elementary properties of power series fields over finite fields*, J. Symb. Logic **66** (2001), 771–791

[30] Kuhlmann, F.-V.: *A classification of Artin Schreier defect extensions and a characterization of defectless fields*, Illinois J. Math. **54** (2010), 397–448

[31] Kuhlmann, F.-V.: *Approximation of elements in henselizations*, Manuscripta math. **136** (2011), 461–474

[32] Kuhlmann, F.-V.: *Invariance group and invariance valuation ring of a cut*, in preparation; preliminary version available at `http://math.usask.ca/~fvk/CUTS.pdf`

[33] Kuhlmann, F.-V. – Kuhlmann, K.: *Embedding theorems for spaces of* $\mathbb{R}$*-places of rational function fields and their products*, Fundamenta Math. **218** (2012), 121–149

[34] Kuhlmann, F.-V. – Kuhlmann, K: *A common generalization of metric, ultrametric and topological fixed point theorems*, Forum Math. **27** (2015), 303–327; and: Correction to "A common generalization of metric, ultrametric and topological fixed point theorems", Forum Math. **27** (2015), 329–330; alternative corrected version available at: `http://math.usask.ca/fvk/GENFPTAL.pdf`

[35] Kuhlmann, F.-V. – Machura, M. – Osiak, K.: *Metrizability of spaces of* $\mathbb{R}$*-places of function fields of transcendence degree 1 over real closed fields*, Comm. in Alg. **39** (2011), 3166–3177

[36] Kuhlmann, F.-V. – Kuhlmann, K. – Shelah, S.: *Symmetrically Complete Ordered Sets, Abelian Groups, and Fields*, Israel J. Math. **208** (2015), 261–290

[37] Kuhlmann, K.: *The structure of spaces of* $\mathbb{R}$*-places of rational function fields over real closed fields*, Rocky Mountains J. Math. **46** (2016), 533–557

[38] Lam, T.Y.: Orderings, valuations and quadratic forms, CBMS Regional Conf. Ser. Math. **52**. Published for the Conf. Board of the Math. Sciences, Washington (1983)

[39] Pestov, G. G.: *On the theory of sections in ordered fields*, Sibirsk. Mat. Zh. **42** (2001), 1350–1360; translation in Siberian Math. J. **42** (2001), 1123–1131

[40] Prestel, A.: *Lectures on formally real fields*, Lecture Notes in Mathematics **1093**, Springer-Verlag, Berlin, 1984

[41] Ribenboim, P.: *Théorie des valuations*, Les Presses de l'Université de Montréal (1964)

[42] Rolland, R.: *Étude des coupures dans les groupes et corps ordonnés*, Real algebraic geometry and quadratic forms (Rennes, 1981), 386–405, Lecture Notes in Math. **959**, Springer, Berlin – New York, 1982

[43] Shelah, S.: *Quite Complete Real Closed Fields*, Israel J. Math. **142** (2004), 261–272

[44] Temkin, M.: *Inseparable local uniformization*, J. Algebra **373** (2013), 65–119

[45] Tressl, M.: *Dedekind cuts in polynomially bounded, o-minimal expansions of real closed fields*, doctoral thesis, Regensburg (1996)

[46] Tressl, M.: *Model completeness of o-minimal structures expanded by Dedekind cuts*, J. Symbolic Logic **70** (2005), 29–60

[47] Tressl, M.: *The elementary theory of Dedekind cuts in polynomially bounded structures*, Ann. Pure Appl. Logic **135** (2005), 113–134

[48] Tressl, M.: *Pseudo completions and completions in stages of o-minimal structures*, Arch. Math. Logic **45** (2006), 983–1009

[49] Wehrung, F.: *Monoids of intervals of ordered abelian groups*, J. Algebra **182** (1996), 287–328

INSTITUTE OF MATHEMATICS, UNIVERSITY OF SILESIA AT KATOWICE, BANKOWA 14, 40-007 KATOWICE, POLAND

*E-mail address:* `fvk@math.us.edu.pl`

# A SURVEY OF QUANTITATIVE BOUNDS ON THE BETTI NUMBERS OF REAL AND COMPLEX VARIETIES AND APPLICATIONS

SAUGATA BASU

Abstract. We survey old results as well as recent developments in the quantitative study of topology of real algebraic varieties and semi-algebraic sets. We also indicate the newer applications of these results, especially in the fields of combinatorics and discrete geometry.

## 1. Introduction

Throughout the paper we denote by R a real closed field, and C the algebraic closure of R. We can even assume that $\mathrm{R} = \mathbb{R}$, and $\mathrm{C} = \mathbb{C}$ without any loss of generality, since all the results of the paper follow for arbitrary real closed fields (resp. algebraically closed fields of characteristic 0), from the corresponding result for $\mathbb{R}$ (resp. $\mathbb{C}$), by an appropriate use of the Tarski-Seidenberg transfer principle (resp. Lefschetz principle) (see for example [12, Chapter 7]). We also fix a field $\mathbb{F}$ to which the coefficients of the various cohomology groups considered in the paper will belong.

For any field K, and finite subsets $\mathcal{P} \subset \mathrm{K}[X_1, \ldots, X_n]$, let $\mathrm{Z}(\mathcal{P}, \mathrm{K}^n)$ be the set of common zeros of $\mathcal{P}$ in $\mathrm{K}^n$. Similarly, for finite subsets of homogeneous polynomials $\mathcal{P} \subset \mathrm{K}[X_0, \ldots, X_n]$, we denote by $\mathrm{Z}(\mathcal{P}, \mathbb{P}^n_{\mathrm{K}})$ be the set of common zeros of $\mathcal{P}$ in $\mathbb{P}^n_{\mathrm{K}}$. Given a real or complex variety $X$, we denote by $\mathrm{H}^i(X, \mathbb{F})$ (resp. $\mathrm{H}_i(X, \mathbb{F})$, $\mathrm{H}^i_c(X, \mathbb{F})$) the $i$-th cohomology group (resp. $i$-th homology, $i$-th cohomology group with compact support) with coefficients in the field $\mathbb{F}$. (We refer the reader to [12, Chapter 6] for the definition of homology/cohomology groups of semi-algebraic sets defined over arbitrary real closed fields, noting that they are isomorphic to the singular homology/cohomology groups in the special case of $\mathrm{R} = \mathbb{R}$.)

We denote by $b^i(X, \mathbb{F})$ (resp. $b^i_c(X, \mathbb{F})$) the dimension of $\mathrm{H}^i(X, \mathbb{F})$ (resp. $\mathrm{H}^i_c(X, \mathbb{F})$), and by $b(X, \mathbb{F}) = \sum_{i \geq 0} b^i(X, \mathbb{F})$ (resp. $b_c(X, \mathbb{F}) = \sum_{i \geq 0} b^i_c(X, \mathbb{F})$).

## 2. Method of critical points

**2.1. Bounding the sum of the Betti numbers of varieties.** The problem of bounding explicitly the Betti numbers of real and complex varieties have been considered for a long time and there have been many applications of these bounds in combinatorics and discrete geometry (see for example, [11] for a survey).

---

*Date*: April 10, 2017.

1991 *Mathematics Subject Classification.* Primary 14F25; Secondary 68W30.

*Key words and phrases.* Betti numbers, semi-algebraic sets, o-minimal.

The author was partially supported by NSF grants CCF-1618981 and DMS-1620271.

The first results are due to Oleĭnik and Petrovskiĭ [49], Thom [55] and Milnor [47] who proved the following bounds.

**Theorem 1.** [49, 55, 47] *Let $\mathcal{P} \subset \mathrm{R}[X_1, \ldots, X_k]$ be a finite set polynomials of degrees at most $d$, and let $V = \mathrm{Z}(\mathcal{P}, \mathrm{R}^k)$. Then,*

$$b(V, \mathbb{F}) \leq \mathrm{Aff}_{\mathrm{R}}(k, d) := d(2d-1)^{k-1} = (O(d))^k. \tag{2.1}$$

The bound in Theorem 1 also holds in the projective case.

**Theorem 2.** [49, 55, 47] *Let $\mathcal{P} \subset \mathrm{R}[X_0, \ldots, X_k]$ be a finite set homogeneous polynomials of degrees at most $d$, and let $V = \mathrm{Z}(\mathcal{P}, \mathbb{P}^k_{\mathrm{R}})$. Then,*

$$b(V, \mathbb{F}) \leq \mathrm{Proj}_{\mathrm{R}}(k, d) := d(2d-1)^{k-1} = (O(d))^k. \tag{2.2}$$

The main technique used in proving these bounds are infinitesimal perturbations and Morse theory. An initial perturbation is made to the real variety, so that the given (possibly highly singular) variety has the same Betti numbers as a semi-algebraic set defined by a single polynomial inequality, which moreover has a smooth boundary. The problem of bounding the Betti numbers of the original variety then reduces (via Alexander duality) to the problem of bounding the Betti numbers of this hypersurface. The sum of the Betti numbers of the hypersurface is then bounded by the number of critical points of a generic linear function and Bezout's theorem.

There are several important points to note.

(A) Firstly, because of the obvious fact, that any finite set of polynomial eqations over R can be replaced by one polynomial defining the same real variety by taking a sum of squares (at the cost of doubling the degree), the bound on the Betti numbers of real varieties is independent of the number of defining polynomials.

(B) On the other hand, taking a sum of squares leads to losing important information about the degrees as well as the Newton polytopes of the defining polynomials (the degree of the sums of square is twice the maximum degree of the given polynomials). Thus, this method is unsuitable if one wants to obtain bounds with a more refined dependence on the degree squence, as is needed in recent applications in incidence questions.

(C) The method does not yield any better bounds on the individual Betti numbers.

(D) The bounds are (for every fixed $d$) singly exponential in the number of variables $k$. Moreover, this exponential dependence on $k$ is unavoidable even if the variety $V$ is a non-singular hypersurface defined by one polynomial, and we consider just a single Betti number of $V$ (for example $b_0(V, \mathbb{F})$ or $b_{k-1}(V, \mathbb{F})$) instead of their sum, as the following examples show.

**Example 1.** Let $P = \sum_{i=1}^{k} X_i^2(X_i - 1)^2 - \varepsilon$, with $0 < \varepsilon \ll 1$, and $P^h$ denote the homogenization of $P$. Let $V = \mathrm{Z}(P, \mathrm{R}^k)$, and $V^h = \mathrm{Z}(P^h, \mathbb{P}^k_{\mathrm{R}})$. Now notice that $\deg(P) = 4$, and

$$b^0(V, \mathbb{F}) = b^{k-1}(V, \mathbb{F}) = b^{k-1}(V^h, \mathbb{F}) = 2^k,$$
$$b^0(V^h, \mathbb{F}) = 2^{k+1}.$$

2.2. **Bounding Betti numbers of complex varieties.** While Theorems 1 and 2 deal only with real varieties, they can be used to bound the Betti numbers of complex varieties, since every complex affine variety in $\mathrm{C}^k$ defined by $r$ polynomials of degrees bounded by $d$, can be considered after separating the real and imaginary parts of the defining polynomials as a real affine variety in $\mathrm{R}^{2k}$ defined by $2r$ polynomials of degree at most $d$. It then follows directly from Theorem 1 that:

**Theorem 3.** *Let $\mathcal{P} \subset \mathrm{C}[Z_1, \ldots, Z_k]$ be a finite set polynomials of degrees at most $d$, and let $V = \mathrm{Z}(\mathcal{P}, \mathrm{C}^k)$. Then,*

$$b(V, \mathbb{F}) \leq \mathrm{Aff}_{\mathrm{C}}(k, d) := d(2d-1)^{2k-1} = (O(d))^{2k}. \tag{2.3}$$

Using an argument involving the Hopf fibration and the Gysin exact sequence one also derives a similar bound in the projective case.

**Theorem 4.** *Let $\mathcal{P} \subset \mathrm{C}[Z_0, \ldots, Z_k]$ be a set of homogeneous polynomials of degrees at most $d \geq 2$, and let $V = \mathrm{Z}(\mathcal{P}, \mathbb{P}^k_{\mathrm{C}})$. Then,*

$$b(V, \mathbb{F}) \leq \mathrm{Proj}_{\mathrm{C}}(k, d) := kd(2d-1)^{2k+1} = (O(d))^{2k+2}. \tag{2.4}$$

*Proof.* Let $\mathbf{S}^{2k+1} \subset \mathrm{C}^{k+1} = \mathrm{R}^{2k+2}$ denote the unite sphere defined by $|Z_0|^2 + \cdots + |Z_k|^2 = 1$. Consider the Hopf fibration $\phi : \mathbf{S}^{2k+1} \to \mathbb{P}^k_{\mathrm{C}}$, defined by $(z_0, \ldots, z_k) \mapsto (z_0 : \cdots : z_k)$. We denote by $\widetilde{V} = \phi^{-1}(V)$. We have the following commutative diagram:

$$\begin{tikzcd}
\widetilde{V} \arrow[r, "i"] \arrow[d, "\phi|_{\widetilde{V}}"'] & \mathbf{S}^{2k+1} \arrow[d, "\phi"] \\
V \arrow[r, "i"] & \mathbb{P}^k_{\mathrm{C}}
\end{tikzcd}$$

Note that $\widetilde{V}$ is a $\mathbf{S}^1$-bundle over $V$. It follows from the Gysin exact sequence [53, page 260] of this bundle that for each $n \geq 0$,

$$\begin{aligned}
b^n(V, \mathbb{F}) &\leq b^{n-2}(V, \mathbb{F}) + b^n(\widetilde{V}, \mathbb{F}) \\
&\leq b^{n-4}(V, \mathbb{F}) + b^{n-2}(\widetilde{V}, \mathbb{F}) + b^n(\widetilde{V}, \mathbb{F}) \\
&\vdots \\
&\leq \sum_{i \geq 0} b^{n-2i}(\widetilde{V}, \mathbb{F}).
\end{aligned}$$

It follows that

$$\begin{aligned}
b(V, \mathbb{F}) &\leq \sum_{i=0}^{2k} \lceil (2k-i)/2 \rceil b^i(\widetilde{V}, \mathbb{F}) \\
&\leq kb(\widetilde{V}, \mathbb{F}).
\end{aligned}$$

The theorem now follows from Theorem 1. □

*Remark* 1. With a little more care (for example, using [16, Theorem 32] instead of Theorem 1 as in the proof above), it is possible to prove a bound of $(O(d))^{2k}$ on $b(\widetilde{V}, \mathbb{F})$. This would also improve the bound on $b(V, \mathbb{F})$ in Theorem 4 to $(O(d))^{2k}$ .

## 3. MAYER-VIETORIS INEQUALITIES AND BOUNDING THE BETTI NUMBERS OF SEMI-ALGEBRAIC SETS

In certain applications, such as in applications in discrete geometry and for proving lower bounds, one needs bounds not just on the Betti numbers of real varieties but also on those of more general *semi-algebraic* sets – i.e. subsets of $\mathrm{R}^k$ defined by quantifier-free Boolean formulas whose atoms are of the form $P \sim 0$, $P \in \mathrm{R}[X_1, \ldots, X_k]$ and $\sim$ is one of $=, >, <$.

We first introduce a few notation.

**Notation 1.** For any finite family of polynomials $\mathcal{P} \subset \mathrm{R}[X_1, \ldots, X_k]$, we call an element $\sigma \in \{0, 1, -1\}^{\mathcal{P}}$, a *sign condition* on $\mathcal{P}$. For any semi-algebraic set $Z \subset \mathrm{R}^k$, and a sign condition $\sigma \in \{0, 1, -1\}^{\mathcal{P}}$, we denote by $\mathcal{R}(\sigma, Z)$ the semi-algebraic set defined by

$$\{\mathbf{x} \in Z \mid \mathbf{sign}(P(\mathbf{x})) = \sigma(P), P \in \mathcal{P}\},$$

and call it the *realization* of $\sigma$ on $Z$. More generally, we call any Boolean formula $\Phi$ with atoms, $P \sim 0, P \in \mathcal{P}$ where $\sim$ is one of $=, >$, or $<$, to be a $\mathcal{P}$*-formula*. We call the realization of $\Phi$, namely the semi-algebraic set

$$\mathcal{R}(\Phi, \mathrm{R}^k) = \{\mathbf{x} \in \mathrm{R}^k \mid \Phi(\mathbf{x})\}$$

a $\mathcal{P}$*-semi-algebraic set*. Finally, we call a Boolean formula without negations, and with atoms $P \sim 0, P \in \mathcal{P}$ where $\sim$ is one of $\leq, \geq$, to be a $\mathcal{P}$*-closed formula*, and we call the realization, $\mathcal{R}(\Phi, \mathrm{R}^k)$, a $\mathcal{P}$*-closed semi-algebraic set*.

Finally, if $\mathcal{Q} \subset \mathrm{R}[X_1, \ldots, X_k]$ is another finite set of polynomials, and $\Phi$ a $\mathcal{P}$ (resp. $\mathcal{P}$-closed) formula, we call the formula

$$\Psi := \bigwedge_{Q \in \mathcal{Q}} (Q = 0) \wedge \Phi$$

is called a $(\mathcal{Q}, \mathcal{P})$ (resp. $(\mathcal{Q}, \mathcal{P})$-closed) formula, and we call the realization, $\mathcal{R}(\Psi, \mathrm{R}^k)$, a $(\mathcal{Q}, \mathcal{P})$ *(resp.* $(\mathcal{Q}, \mathcal{P})$*-closed) semi-algebraic set.*

It turns out semi-algebraic sets which are evidently closed by virtue of the shape of their defining formulas, i.e. the $\mathcal{P}$-closed semi-algebraic sets. are much easier to handle, since the problem of bounding their Betti numbers can be reduced relatively easily to the same problem for algebraic sets discussed in §2.1 above. The main technical tools are certain infinitesimal perturbations coupled with the following inequalities which are derived from the Mayer-Vietoris spectral sequence of a covering of a closed semi-algebraic sets, by a finite number of closed semi-algebraic subsets. We first recall these inequaities.

### 3.1. Mayer-Vietoris inequalities.

Let $S_1, \ldots, S_N \subset \mathrm{R}^k$, $N \geq 1$, be closed semi-algebraic subsets of $\mathrm{R}^k$. For $J \subset [1, n]$, we denote

$$S_J = \bigcap_{j \in J} S_j,$$
$$S^J = \bigcup_{j \in J} S_j.$$

**Proposition 1.** *A. For* $i \geq 0$,

$$b_i(S^{[1,s]},\mathbb{F}) \leq \sum_{j=1}^{i+1} \sum_{\substack{J\subset\{1,\ldots,s\}\\ \mathrm{card}(J)=j}} b_{i-j+1}(S_J,\mathbb{F}). \tag{3.1}$$

*B.*

$$b_i(S_{[1,s]},\mathbb{F}) \leq \sum_{j=1}^{k-i} \sum_{\substack{J\subset\{1,\ldots,s\}\\ \mathrm{card}(J)=j}} b_{i+j-1}(S^J,\mathbb{F}) + \binom{s}{k-i} b_k(S^\emptyset,\mathbb{F}). \tag{3.2}$$

*Proof.* See [12, Proposition 7.33]. □

3.2. **Bounding the Betti numbers of $(\mathcal{Q},\mathcal{P})$-closed semi-algebraic sets.** Using certain certain infinitesimal perturabtion arguments and Proposition 1 one obtains the following bound.

**Theorem 5.** *Let $\mathcal{Q}$ and $\mathcal{P} \neq \emptyset$ be finite subsets of* $\mathrm{R}[X_1,\ldots,X_k]$. *Let $Z = \mathrm{Z}(\mathcal{Q},\mathrm{R}^k)$, and let $k' > 0$ be the dimension of $Z = \mathrm{Z}(\mathcal{Q},\mathrm{R}^k)$. Suppose that* $\mathrm{card}(\mathcal{P}) = s$, *and the degrees of the polynomials in $\mathcal{P}\cup\mathcal{Q}$ are bounded by $d$. Then for any $(\mathcal{Q},\mathcal{P})$-closed semi-algebaric set $S$,*

$$b(S,\mathbb{F}) \leq \sum_{i=0}^{k'}\sum_{j=1}^{k'-i}\binom{s+1}{j}6^j d(2d-1)^{k-1}.$$

A second result proved using similar techniques is sometimes useful in practice (see for example, [5] for a recent application).

**Theorem 6.** *Let $\mathcal{Q}$ and $\mathcal{P} \neq \emptyset$ be finite subsets of* $\mathrm{R}[X_1,\ldots,X_k]$. *Let $Z = \mathrm{Z}(\mathcal{Q},\mathrm{R}^k)$, and let $k' > 0$ be the dimension of $Z = \mathrm{Z}(\mathcal{Q},\mathrm{R}^k)$. Suppose that* $\mathrm{card}(\mathcal{P}) = s$, *and the degrees of the polynomials in $\mathcal{P}\cup\mathcal{Q}$ are bounded by $d$. Then,*

$$\sum_{\sigma\in\{0,1,-1\}^{\mathcal{P}}} b_i(\mathcal{R}(\sigma,Z),\mathbb{F}) \leq \sum_{1\leq j\leq k'-i}\binom{s}{j}4^j d(2d-1)^{k-1}.$$

3.3. **Bounding the Betti numbers of arbitrary semi-algebraic sets.** Surprisingly, the problem of proving singly exponential bounds for arbitrary semi-algebraic sets remained open for a long time. The technique of using Mayer-Vietoris inequalities to reduce to the algebraic case, while controlling the number of algebraic sets to consider, could not be applied directy since inequalities in Proposition 1 require the sets involved to be closed. This problem was finally remedied by Gabrielov and Vorobjov, who showed that an arbitrary $(\mathcal{Q},\mathcal{P})$-semi-algebraic set is semi-algebraically homotopy equivalent to a certain $(\mathcal{Q},\mathcal{P}')$-closed semi-algebraic set, where $\mathcal{P}'$ is a set of polynomials of cardinality only sightly bigger than that of $\mathcal{P}$, and having the same maximum degree as $\mathcal{P}$. The problem of bounding the Betti numbers of arbitrary $(\mathcal{Q},\mathcal{P})$-semi-algebraic sets, can then be reduced to bounding the Betti numbers of this $(\mathcal{Q},\mathcal{P}')$-closed semi-algebraic set, for which one can use Theorem 5.

3.3.1. *Approximation of an arbitrary semi-algebraic set by a closed one.* In fact Gabrielov and Vorobjov gave two different constructions of the $(\mathcal{Q}, \mathcal{P}')$-closed semi-algebraic set mentioned above, with the cardinality of the new set of polynomials been slightly different in each case. We describe below the second of their two constructions.

Let $\mathcal{Q}, \mathcal{P} \subset \mathrm{R}[X_1, \ldots, X_k]$ be finite sets and $\Phi$, a $(\mathcal{Q}, \mathcal{P})$-formula. Then, $\Phi$ is a disjunction of fomulas of the kind,

$$\bigwedge_{Q \in \mathcal{Q}} (Q = 0) \wedge \bigwedge_{P \in \mathcal{P}_=} (P = 0) \wedge \bigwedge_{P \in \mathcal{P}_>} (P > 0) \wedge \bigwedge_{P \in \mathcal{P}_<} (P < 0), \tag{3.3}$$

where $\mathcal{P}$ is a disjoint union of $\mathcal{P}_=, \mathcal{P}_<, \mathcal{P}_>$.

For $\delta, \varepsilon \in \mathrm{R}$, we denote by $\Phi_{\delta,\varepsilon}$ the formula obtained by replacing in each disjunct (3.3), $P = 0$ by $-\varepsilon \leq P \leq \varepsilon$, $P > 0$ by $P \geq \delta$ and $P < 0$ by $P \leq -\delta$.

The following theorem is easily derived from the main theorem in [35].

**Theorem 7.** *For each $(\mathcal{Q}, \mathcal{P})$-formula $\Phi$, $r > 0$, and $0 < \varepsilon_0 \ll \delta_0 \ll \varepsilon_1 \ll \delta_1 \ll \cdots \ll \varepsilon_m \ll \delta_m \ll 1$, denoting by $\Phi_m = \bigvee_{i=0}^{m} \Phi_{\delta_i,\varepsilon_i}$, for $0 \leq i < m$, $\mathrm{H}_i(\mathcal{R}(\Phi \wedge (||X||^2 \leq r), \mathrm{R}^k), \mathbb{F}) \cong \mathrm{H}_i(\mathcal{R}(\Phi_m \wedge (||X||^2 \leq r, \mathrm{R}^k), \mathbb{F})$.*

*Remark* 2. Notice that $\Phi_m$ is a $(\mathcal{Q}, \mathcal{P}_m)$-closed formula, where $\mathcal{P}_m \cup_{i=0}^{m} \{P \pm \delta_i, \varepsilon_i \mid P \in \mathcal{P}\}$. Notice that $\mathrm{card}(\mathcal{P}_m) = 4(m+1)\mathrm{card}(\mathcal{P})$, and $\max_{P \in \mathcal{P}} \deg(P) = \max_{P \in \mathcal{P}_m} \deg(P)$.

Theorem 7 yields in conjunction with Theorem 5 the following theorem.

**Theorem 8.** *Let $\mathcal{Q}$ and $\mathcal{P} \neq \emptyset$ be finite subsets of $\mathrm{R}[X_1, \ldots, X_k]$. Let $Z = \mathrm{Z}(\mathcal{Q}, \mathrm{R}^k)$, and let $k' > 0$ be the dimension of $Z = \mathrm{Z}(\mathcal{Q}, \mathrm{R}^k)$. Suppose that* $\mathrm{card}(\mathcal{P}) = s$, *and the degrees of the polynomials in $\mathcal{P} \cup \mathcal{Q}$ are bounded by $d$. Then for any $(\mathcal{Q}, \mathcal{P})$-semi-algebraic set $S$,*

$$b(S, \mathbb{F}) \leq \sum_{i=0}^{k'} \sum_{j=1}^{k'-i} \binom{4(k+1)s+1}{j} 6^j d(2d-1)^{k-1}.$$

Theorem 7 has other applications as well. For example, it is a crucial ingredient of the extension of previously known lower bounds on membership-testing in locally closed semi-algebraic sets in terms of the Betti numbers of such sets ([59, 48]) to arbitrary semi-algebraic sets [36].

## 4. METHOD OF PERTURBATIONS AND SMITH INEQUALITY

Another approach for bounding the Betti numbers of real varieties has been via complex algebraic geometry and Smith inequalities. In this approach, one makes use of the fact that the Betti numbers of non-singular complex projective complete intersection varieties are determined only by the degree sequence of the defining polynomials. The main idea is to prove an upper bound on the Betti numbers of a given real variety (possibly highly singular), in terms of the Betti numbers of some non-singular complete intersection varieties. The $\mathbb{Z}_2$-Betti numbers of the complexification of the latter varieties, which are determined only by their degree sequence can then be used in conjunction with Smith inequalities to bound the Betti numbers of the original real variety. This approach has proved to be particularly useful in contexts where one needs bounds depending in a more refined way on the sequence of degrees of the polynomials defining a real variety, rather than a bound depending only on the maximum degree as in Theorem 1 above. One drawback thought is that this approach produces bounds only on the $\mathbb{Z}_2$-Betti numbers.

One example, of the above method is the following theorem.

**Theorem 9.** [6]
*Suppose that*

$$2 \leq d_1 \leq d_2 \leq \frac{1}{k+1} d_3 \leq \frac{1}{(k+1)^2} d_4 \leq \cdots \leq \frac{1}{(k+1)^{\ell-2}} d_\ell.$$

*Then,*

$$b_0(V_\ell) \quad \leq \quad O(1)^k \sum_{\tau=(\tau_0,\tau_1,\ldots,\tau_{\ell-1})} F(k,\tau) \left( d_\ell^{\tau_{\ell-1}} \prod_{1 \leq i < \ell} ((k - \tau_{i-1} + 1) d_i)^{\tau_{i-1} - \tau_i} \right)$$

*where the sum on the right hand side is taken over all* $\tau \in \mathbb{N}^\ell$, *with* $k = \tau_0 \geq \tau_1 \geq \cdots \geq \tau_{\ell-1} \geq 0$, *and* $\tau_i \leq k_i$, *for each* $i, 1 \leq i < \ell$, *and*

$$F(k,\tau) = (k - \tau_{\ell-1} + 1) \binom{k - \tau_{\ell-1}}{\tau_0 - \tau_1, \tau_1 - \tau_2, \ldots, \tau_{\ell-2} - \tau_{\ell-1}}.$$

*This implies that*

$$b_0(V_\ell) \quad \leq \quad O(1)^\ell O(k)^{2k} \left( \prod_{1 \leq j < \ell} d_j^{k_{j-1} - k_j} \right) d_\ell^{k_{\ell-1}},$$

*and in particular if* $\ell \leq k$,

$$b_0(V_\ell) \leq O(k)^{2k} \left( \prod_{1 \leq j < \ell} d_j^{k_{j-1} - k_j} \right) d_\ell^{k_{\ell-1}}.$$

Aside from its applications in discrete geometry (particularly, in incidence problems tackled using the polynomial partitioning method [60, 19]), Theorem 9 also remedies a well-known anomaly – which is that the naive statement of Bezout inequality, that the number of isolated solutions in $\mathrm{C}^n$ of a system of $n$ polynomial equations in $n$ variables is bounded by the product of their degrees is no longer true for isolated solutions in $\mathrm{R}^n$. This is illustrated by the following well known example [31].

**Example 2.** Let $k = 3$, and let

$$\begin{aligned} Q_1 &= X_3, \\ Q_2 &= X_3, \\ Q_3 &= \sum_{i=1}^{2} \left( \prod_{j=1}^{d} (X_i - j)^2 \right). \end{aligned}$$

The real variety defined by $\{Q_1, Q_2, Q_3\}$ is 0-dimensional, and has $d^2$ isolated (in $\mathrm{R}^3$) points, whereas the degree sequence is $(d_1, d_2, d_3) = (1, 1, 2d)$, and thus the bound predicted by the naive Bezout inequality is equal to $2d$.

The technique behind proving Theorem 9 involve successive approximations of possibly singular varieties with ones which are non-singular complete intersections. The number of connected components of these are then bounded using classical formulas (coming from complex algebraic geometry) expressing the Euler-Poincaré characteristic of non-singular complex complete intersections in terms of degree sequences, and Smith inequalities. The main difficulty though lies in showing that the number of connected components of the given variety is bounded by the number of connected components of (the real part) of these complete intersections. The approximation scheme used in Theorem 9 is rather complicated and the following example and pictures are for illustrative purposes only. The following figures show these approximation varieties in a special case.

**Example 3.** Let $k = 3$, $\ell = 3$, and

$$\begin{aligned} Q_1 &= (X_1^2 + X_2^2 + X_3^2 - 1)(X_3^2 + (X_1^2 + \frac{1}{2}X_2^2 - 1)^2), \\ Q_2 &= (X_3^2 + (X_1^2 + \frac{1}{2}X_2^2 - 1)^2), \\ Q_3 &= X_3^2 + X_2^2 + (X_1 - 1)^2. \end{aligned}$$

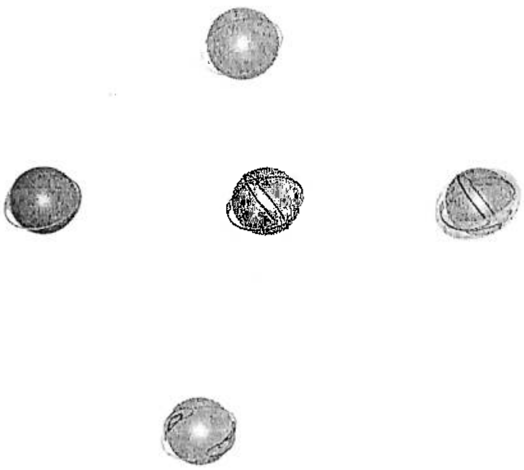


FIGURE 1. The successive approximating varieties.

The variety $V_1$ (shown in the top line of Figure 1) is bounded, and equal to the union of the unit sphere $S \subset \mathrm{R}^3$ (shown in orange), and an ellipse, $C$ (shown in green), contained in the plane $\mathrm{span}(e_1, e_2)$, with $S \cap C = \{(\pm 1, 0, 0)\}$ (shown in red). The variety $V_2 = C$, and $V_3 = \{(1, 0, 0)\}$. The successive approximation varieties in this case is shown in the second line (see [6] for details). The final approximating variety to $V_3$ (red points) consists of a curve with 4 connected components (in blue and purple).

It is easy to see from Figure 1 that the approximating varieties, while approximating the given variety in the sense of Hausdorff limit (which is sufficient for proving a bound on the number of connected components), do not approximate the *topology* of the given variety. Thus, using this technique one is not able to prove a similar upper bound as in Theorem 9 on the sum of all the Betti numbers – as in the case of the original bounds of Oleĭnik and Petrovskiĭ [49], Thom [55] and Milnor [47].

*Remark* 3. Note that since the real dimension of each variety $V_i$ is at most the complex dimension of $V_i$, Theorem 9 remains true if we replace real dimension by complex dimension in the statement. This observation is important in the application of Theorem 9 to incidence problems (see [19]).

Theorem 9 helped to make progress on several incidence questions. For example, it plays a crucial role in Zahl's new proof of Szemeredi-Trotter theorem over the complex numbers [60] (i.e. Szemeredi-Trotter theorem for points and complex lines in $\mathrm{C}^2$), as well as in recent work on bounding point-hypersurface incidences in $\mathrm{R}^4$ [19].

The bound in Theorem 9 has also been extended to the semi-algebraic case. With the same assumptions as in Theorem 9, suppose additionally that $\mathcal{P} \subset \mathrm{R}[X_1, \ldots, X_k]$ is a finite family of polynomials with $\deg(P) \leq d$ for all $P \in \mathcal{P}$, and $\mathrm{card}(\mathcal{P}) = s$, and suppose that $d_\ell \leq \frac{1}{k+1} d$.

**Theorem 10.** [6]

$$\sum_{\sigma \in \{0,1,-1\}^{\mathcal{P}}} b_0(\mathcal{R}(\sigma, V_\ell)) \quad \leq \quad \sum_{j=0}^{k_\ell} 4^j \binom{s}{j} O(1)^k \Delta \tag{4.1}$$

*where $\Delta$ is defined by*

$$\Delta = \sum_{\tau = (\tau_0, \tau_1, \ldots, \tau_\ell)} F(k, \tau) d^{\tau_\ell} \left( \prod_{1 \leq i \leq \ell} ((k - \tau_{i-1} + 1) d_i)^{\tau_{i-1} - \tau_i} \right),$$

*where the sum is taken over all $\tau \in \mathbb{N}^{\ell+1}$, with $k = \tau_0 \geq \tau_1 \geq \cdots \geq \tau_\ell \geq 0$, and $\tau_i \leq k_i$, for each $i, 1 \leq i \leq \ell$, and*

$$F(k, \tau) = (k - \tau_\ell + 1) \binom{k - \tau_\ell}{\tau_0 - \tau_1, \tau_1 - \tau_2, \ldots, \tau_{\ell-1} - \tau_\ell}.$$

*This implies that*

$$\sum_{\sigma \in \{0,1,-1\}^{\mathcal{P}}} b_0(\mathcal{R}(\sigma, V_\ell)) \quad \leq \quad O(1)^\ell O(k)^{2k} (sd)^{k_\ell} \left( \prod_{1 \leq j \leq \ell} d_j^{k_{j-1} - k_j} \right).$$

*In particular, if $\ell \leq k$,*

$$\sum_{\sigma \in \{0,1,-1\}^{\mathcal{P}}} b_0(\mathcal{R}(\sigma, V_\ell)) \quad \leq \quad O(k)^{2k} (sd)^{k_\ell} \left( \prod_{1 \leq j \leq \ell} d_j^{k_{j-1} - k_j} \right).$$

With the same assumptions as in Theorem 10, let for $P \in \mathcal{P}$, $d_P = \deg(P)$, and for any subset $\mathcal{I} \subset \mathcal{P}$ let

$$d_{\mathcal{I}} \quad = \quad (k+1)^{\binom{\mathrm{card}\mathcal{I}}{2} + (k_\ell - \mathrm{card}\mathcal{I})(\mathrm{card}\mathcal{I} - 1)} \left( \prod_{P \in \mathcal{I}} d_P \right) (\max_{P \in \mathcal{I}} d_P)^{k_\ell - \mathrm{card}\mathcal{I}}. \tag{4.2}$$

We have the following variant of Theorem 10 (the extra precision with respect to the degrees of polynomials in $\mathcal{P}$ might be useful in applications in incidence geometry).

Using notation introduced in Theorem 10 and (4.2) above:

**Theorem 11.** [6]

$$\sum_{\sigma \in \{0,1,-1\}^{\mathcal{P}}} b_0(\mathcal{R}(\sigma, V_\ell)) \leq \sum_{\substack{\mathcal{I} \subset \mathcal{P} \\ j = \mathrm{card}\mathcal{I} \leq k_\ell}} 4^j O(1)^\ell O(k)^{2k} d_{\mathcal{I}} \left( \prod_{1 \leq j \leq \ell} d_j^{k_{j-1}-k_j} \right).$$

## 5. Multidegree bounds

In certain applications one is interested in tight bounds on the Betti numbers of real varieties and semi-algebraic sets in terms of more finer invariants than the total degrees of the polynomials – for example, one might want to have a bound in terms of the degrees of the polynomials in each variable. Such a bound is useful when the degrees of the polynomials in different variables are very different. The following results were obtained in [16].

**Notation 2.** Given, $\mathbf{k} = (k_1, \ldots, k_p), \mathbf{d} = (d_1, \ldots, d_p) \in \mathbb{N}^p$, and $j > 0$, we denote by $k = \sum_{i=1}^p k_i$ and

$$G_{\mathrm{gen}}(\mathbf{d}, \mathbf{k}, j) = 1 + (-1)^{k-j+1} + (k-j+2)^2 \binom{k}{j-1} \binom{k}{\mathbf{k}}^{-1} \frac{(1+p)^{3k-j+1}}{p(p+2)} d_1^{k_1} \cdots d_p^{k_p}.$$

**Theorem 12.** [16] *Let* $\mathcal{Q} = \{Q_1, \ldots, Q_\ell\} \subset \mathrm{R}[\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(p)}]$ *be a finite set of polynomials with* $\ell > 0$, *where for* $1 \leq i \leq p$, $\mathbf{X}^{(i)} = (X_1^{(i)}, \ldots, X_{k_i}^{(i)})$, *and* $\deg_{\mathbf{X}^{(i)}}(Q) \leq d_i$, $d_i \geq 2$, *for all* $Q \in \mathcal{Q}$. *Let also* $V = \mathrm{Z}(\mathcal{Q}, \mathrm{R}^k)$, *where* $k = \sum_{i=1}^p k_i$. *Denote by* $\mathbf{d} = (d_1, \ldots, d_p)$ *and* $\mathbf{k} = (k_1, \ldots, k_p)$. *Then,*

$$\begin{aligned} b(V, \mathbb{Z}_2) &\leq G_{\min}(\mathbf{d}, \mathbf{k}, \ell) \\ &\leq O(1)^k p^{3k} d_1^{k_1} \cdots d_p^{k_p}, \end{aligned}$$

*where* $G_{\min}(\mathbf{d}, \mathbf{k}, \ell)$ *equals*

$$\min\left(3 + \sum_{j=1}^{k} \binom{\ell}{j} 2^j (G_{\mathrm{gen}}(\mathbf{d}', \mathbf{k}, j) + G_{\mathrm{gen}}(\mathbf{d}', \mathbf{k}, j+1)),\ \frac{1}{2} G_{\mathrm{gen}}(2\mathbf{d}, \mathbf{k}, 1)\right),$$

$\mathbf{d}' = (d_1', \ldots, d_p')$, *and for* $1 \leq i \leq p$, $d_i'$ *is the least even integer* $\geq d_i$.

**Theorem 13.** [16] *Let* $\mathcal{P} = \{P_1, \ldots, P_s\} \subset \mathrm{R}[\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(p)}]$ *be a finite set of polynomials with* $s > 0$, *where for* $1 \leq i \leq p$, $\mathbf{X}^{(i)} = (X_1^{(i)}, \ldots, X_{k_i}^{(i)})$, *and* $\deg_{\mathbf{X}^{(i)}}(P) \leq d_i$, $d_i \geq 2$, *for all* $P \in \mathcal{P}$. *Denote by* $\mathbf{d} = (d_1, \ldots, d_p)$ *and* $\mathbf{k} = (k_1, \ldots, k_p)$. *Then, for each* $i, 0 \leq i \leq k-1$,

$$\begin{aligned} \sum_{\sigma \in \{0,1,-1\}^{\mathcal{P}}} b_i(\mathcal{R}(\sigma, \mathrm{R}^k), \mathbb{Z}_2) &\leq \sum_{j=1}^{k-i} \binom{s}{j} 4^j G_{\min}(\mathbf{d}, \mathbf{k}, j) \\ &\leq O(1)^k s^{k-i} p^{3k} d_1^{k_1} \cdots d_p^{k_p}. \end{aligned}$$

*Furthermore, if* $S$ *is any* $\mathcal{P}$*-closed semi-algebraic set, then*

$$\begin{aligned} b(S, \mathbb{Z}_2) &\leq \sum_{i=0}^{k} \sum_{j=1}^{k-i} \binom{s+1}{j} 6^j G_{\min}(\mathbf{d}, \mathbf{k}, j) \\ &\leq O(1)^k s^k p^{3k} d_1^{k_1} \cdots d_p^{k_p}. \end{aligned}$$

*Remark* 4. Theorems 12 and 13 are proved using a framework which can be summarized as follows. Using infinitesimal perturbation, and inequalities derived from the Mayer-Vietoris exact sequence, the problem is reduced to bounding the sum of the $\mathbb{Z}_2$-Betti numbers of a particular class of semi-algebraic sets, to bounding the same for a set of real algebraic varieties, which are non-singular complete intersections in affine space. The perturbations need to be chosen carefully so that the degree dependencies of the various blocks of variables in the original set of polynomials are preserved. One then uses Smith inequalities and a result of Khovanskiĭ [40] to bound the $\mathbb{Z}_2$-Betti numbers of these varieties. Similar techniques for the simpler problem of bounding the number of connected components of real algebraic varieties was used earlier by Benedetti, Loeser and Risler [21].

## 6. QUADRATIC CASE

Semi-algebraic sets defined by few quadratic inequalities are topologically simpler. This was first noticed by Agrachev [3, 2, 1] who proved a bound which is polynomial in the number of variables and exponential in the number of inequalities for *generic* quadratic inequalities. The technique introduced by Agrachev was very important in later developments as well. Independently, using a different technique (closer to the spirit of Morse theoretic arguments) Barvinok [7] proved the following theorem (no genericity assumption is required).

**Theorem 14.** [7] *Let* $S \subset \mathrm{R}^k$ *be defined by* $P_1 \geq 0, \ldots, P_s \geq 0$, $\deg(P_i) \leq 2$, $1 \leq i \leq s$. *Then,*

$$b(S, \mathbb{Z}_2) \leq k^{O(s)}.$$

This bound was later sharpened in [9], where the authors using the technique discussed in §4 prove the following bound.

**Theorem 15.** [9] *Let* $\mathcal{P} = \{P_1, \ldots, P_s\} \subset \mathrm{R}[X_1, \ldots, X_k]$, $s \leq k$. *Let* $S \subset \mathrm{R}^k$ *be defined by*

$$P_1 \geq 0, \ldots, P_s \geq 0$$

*with* $\deg(P_i) \leq 2$. *Then, for* $0 \leq i \leq k-1$,

$$b_i(S, \mathbb{Z}_2) \leq \frac{1}{2}\left(\sum_{j=0}^{min\{s,k-i\}} \binom{s}{j}\binom{k+1}{j} 2^j\right).$$

*In particular, for* $2 \leq s \leq k/2$, *we have*

$$b_i(S, \mathbb{Z}_2) \leq \frac{1}{2} 3^s \binom{k+1}{s} \leq \frac{1}{2}\left(\frac{3e(k+1)}{s}\right)^2.$$

This was further sharpened in the case of algebraic sets by Lerario in [45, Theorem 15], where the following nearly optimal result was proved.

**Theorem 16.** [45] *Let* $\mathcal{Q} \subset \mathrm{R}[X_0, \ldots, X_k]$ *be a set of* $\ell$ *quadratic forms, and* $V = \mathrm{Z}(\mathcal{Q}, \mathbb{P}^k_{\mathrm{R}})$ *be the* projective *variety defined by* $\mathcal{Q}$. *Then,*

$$b(V, \mathbb{Z}_2) \leq (O(k))^{\ell-1}.$$

A striking aspect of the above bounds that the dependence on the dimension $k$ is polynomial for fixed number of equations. This is very special to the quadratic case. In higher degrees, the real variety defined by one polynomial can have exponentially large (in $k$) Betti numbers.

One natural question is whether there exists common generalizations of the singly exponential bounds as in Theorem 1, and the polynomial bounds in the quadratic case as in the above theorems.

Finally, in [18] the authors also prove a result that generalizes the bounds on Betti numbers of general semi-algebraic sets (defined by $s$ polynomials having degrees bounded by $d$, cf. Theorem 5), as well as the bounds in the quadratic case (cf. Theorems 14, 15, 16). More precisely they prove:

**Theorem 17.** [18] *Let* $\mathcal{P}_1 \subset \mathrm{R}[X_1,\ldots,X_{k_1}]$, *a finite set of polynomials with*

$$\deg_X(P) \leq d, P \in \mathcal{P}_1, \mathrm{card}(\mathcal{P}_1) = s,$$

*and let* $\mathcal{P}_2 \subset \mathrm{R}[X_1,\ldots,X_{k_1},Y_1,\ldots,Y_{k_2}]$, *a finite set of polynomials with*

$$\deg_X(P) \leq d, \deg_Y(P) \leq 2, P \in \mathcal{P}_2, \mathrm{card}(\mathcal{P}_2) = m,$$

*Let* $S \subset \mathrm{R}^{k_1+k_2}$ *be a* $(\mathcal{P}_1 \cup \mathcal{P}_2)$*-closed semi-algebraic set. Then*

$$b(S,\mathbb{Z}_2) \quad \leq \quad k_2^2(O(k_2+s+m)k_2d)^{k_1+2m}. \tag{6.1}$$

*In particular, for* $m \leq k_2$, $b(S,\mathbb{Z}_2) \leq k_2^2(O(s+k_2)k_2d)^{k_1+2m}$.

*Remark* 5. In particular, if in Theorem 17, $\mathcal{P}_1 = \emptyset$ (and hence $s = 0$), and $m, k_1 < k_2$, we get

$$b(S,\mathbb{Z}_2) \quad \leq \quad k_2^2(O(m+k_2)k_2d)^{k_1+2m}. \tag{6.2}$$

Using different techniques, akin to those described in §4, the following theorems are proved in [16].

**Theorem 18.** [16] *With the same notation as in Theorem 17, for each* $i$, $0 \leq i \leq k-1$ *and assuming* $m \leq k_2$,

$$b(S,\mathbb{Z}_2) \quad \leq \quad (O(k_2))^{k_1+m+3}(O(sd))^{k_1}, \text{ for } m, k_1 < k_2.$$

**Theorem 19.** [16] *Let* $\mathcal{Q} = \{Q_1,\ldots,Q_\ell\} \subset \mathrm{R}[X_1,\ldots,X_{k_1},Y_1,\ldots,Y_{k_2}]$ *be a finite set of polynomials with* $\ell > 0$, $\deg_{\mathbf{X}}(Q) \leq d$, $d \geq 2$, *and* $\deg_{\mathbf{Y}}(Q) \leq 2$ *for all* $Q \in \mathcal{Q}$. *Let* $V$ *denote* $\mathrm{Z}(\mathcal{Q},\mathrm{R}^k)$. *Then, with* $\ell, k_1 \leq k_2$,

$$b(V,\mathbb{Z}_2) \quad \leq \quad (O(k_2))^{\ell+k_1}d^{k_1}. \tag{6.3}$$

For projective varieties in $\mathbb{P}^k_{\mathrm{R}}$ defined by a fixed number of homogeneous quadratic polynomials we have the following bound that is asymptotically a slight improvement over the tightest bound known previously [15, Theorem 15] (namely, the bound $(O(k))^{\ell-1}$).

**Theorem 20.** [16] *For each fixed* $\ell > 0$, *and for each set* $\mathcal{P} \subset \mathrm{R}[X_0,\ldots,X_k]$ *of homogeneous polynomials of degree* 2 *of* $\mathrm{card}(\mathcal{P}) \leq \ell$,

$$b(\mathrm{Z}(\mathcal{P},\mathbb{P}^k_{\mathrm{R}}) \quad \leq \quad \left(O\left(\frac{k}{\ell}\right)\right)^{\ell-1}.$$

## 7. TOPOLOGICAL COMPLEXITY OF IMAGES OF SEMI-ALGEBRAIC SETS UNDER POLYNOMIAL MAPS

As a consequence of the Tarski-Seidenberg principles, semi-algebraic sets are closed under taking images under polynomial maps. It is often of interest to have good bounds on the Betti numbers of images of semi-algebraic sets under polynomial maps (often just projections on a subset of the coordinates). Now, since there are known effective bound on the complexity of real quantifier elimination, on possible approach to having such bounds is to first obtain a a semi-algebraic description of the image of the given set, and then use the bounds described previously to bound its Betti numbers. This approach often leads to non-optimal upper bounds. Moreover, in certain case, when the complexity of a polynomial is measured by some quatity other than the degree (for example, the number of monomials), no effective bounds might be known on the description of the image in terms of such a measure (see for example [34]).

Gabrielov, Vorobjov and Zell [33] introduced a different technique relying on a certain spectral sequence argument to prove directly a bound on th eBetti numbers of th eimage without going through effective quantifier elimination. The advantage of this method is that it applies even in situations when effective quantifier elimination is not available (for example, in the case of sets defined by fewnomials or more generally by Pfaffian functions).

The main topological tool used is the following theorem.

**Theorem 21.** [33] *Let $f : X \to Y$ be a proper semi-algebraic map. Then, for each $n \geq 0$,*

$$b_n(f(X), \mathbb{F}) \quad \leq \quad \sum_{p+q=n} b^q(W_f^p(X), \mathbb{F}),$$

*where*

$$W_f^p(X) = \underbrace{X \times_f \cdots \times_f X}_{(p+1)-times},$$

*is the $(p+1)$-fold fiber product of $X$ over $f$.*

*Remark* 6. The assumption of properness can be relaxed slightly but cannot entirely be done away with.

With Theorem 21 at hand it is possible to prove the following theorem.

**Theorem 22.** [33] *Let $\mathcal{P} \subset \mathrm{R}[Y_1, \ldots, Y_m, X_1, \ldots, X_k]$ be a family of polynomials and with $\deg(P) \leq d, P \in \mathcal{P}$, $\mathrm{card}(\mathcal{P}) = s$. Let $\pi : \mathrm{R}^{m+k} \to \mathrm{R}^m$ be the projection map to the first $m$ co-ordinates, and let $S$ be a bounded $\mathcal{P}$-closed semi-algebraic set. Them,*

$$b(\pi(S), \mathbb{F}) \quad = \quad (O(sd))^{(k+1)m}.$$

The polynomial map $\pi$ under which the image taken in Theorem 22 is a projection and is thus a linear map. Sometimes, it is useful to make a distinction between the degree of the map, and the degrees of the polynomials defining the semi-algebraic set $S$. The following theorem which makes this distinction is proved using the same topological tool (namely, Theorem 21), but uses a multi-degree argument to bound the Betti numbers of the fibered products.

**Theorem 23.** [10]. *Let* $\mathcal{F} = \{F_1, \ldots, F_m\}, \mathcal{G} \subset \mathrm{R}[X_1, \ldots, X_k]$, *with* $\deg(F) \leq d, F \in \mathcal{F}$, *and* $\deg(G) \leq D, G \in \mathcal{G}$, *and let* $\mathrm{card}(\mathcal{G}) = s$. *Let* $\mathbf{F} : \mathrm{R}^k \rightarrow \mathrm{R}^m$ *denote the polynomial map* $x \mapsto (F_1(x), \ldots, F_m(x))$, *and let* $T \subset \mathrm{R}^k$ *be a bounded* $\mathcal{G}$*-closed semi-algebraic set. Suppose also that* $d \geq D$.

*Then, for* $0 \leq i \leq m$,

$$b_i(\mathbf{F}(T), \mathbb{Z}_2) \quad \leq \quad O(i)^{\alpha_i}(m+s)^{\alpha_i} d^{(i+1)k} D^m$$

*where* $\alpha_i = (i+1)k + m$.

## 8. SYMMETRIC SEMI-ALGEBRAIC SETS

Another topic that has attracted attention recently in the area of quantitative semi-algebraic geometry is the study of symmetric real varieties and semi-algebraic sets defined by symmetric polynomials of fixed degrees. It is easily seen that such varieties can have Betti numbers which are exponentially large in the dimension of the ambient space. Nevertheless, it is possible to prove a polynomial upper bound on the *equivariant* Betti numbers of such varieties. In this simple situation the equivariant Betti numbers are equal to the ordinary Betti numbers of the quotient $V/\mathfrak{S}_k$ (see [15] for more details).

The following theorem is proved in [15].

**Theorem 24.** *Let* $P \in \mathrm{R}[X_1, \ldots, X_k]$ *be a symmetric polynomial of degree* $d$ *and* $\mathbb{F}$ *a field of characteristic* 0. *Then, the sum of the* $\mathfrak{S}_k$*-equivariant Betti numbers of* $\mathrm{Z}(P, \mathrm{R}^k)$ *is bounded by* $k^{O(d)}$. *Moreover, the equivariant Betti numbers vanish in dimension greater than* $2d$.

The proof Theorem 24 relies on prior work of Riener [50] and Timofte [56, 57] on the so called "degree principle" (which states that if a real symmetric polynomial in $\mathrm{R}^k$ of degree bounded by $d$ has isolated real zeros, then the number of distinct coordinates in each such zero is bounded by $d$), in addition to some some non-trivial perturbation arguments and equivariant Morse-theory.

More recently, Theorem 24 has been strengthened and generalized in several directions. Essentially tight bounds on the vanishing and on the equivariant Betti numbers, which improve the bound in Theorem 24 has been proved in [14] using results of Arnold [4], Giventhal [37] and Kostov [44] on the topology of Vandermonde varieties and hyperbolic polynomials, as well as certain perturbation and Mayer-Vietoris type arguments. In another direction, the polynomial bound on the equivariant Betti number in Theorem 24 has been extended to a bound on the multiplicities of all irreducible representations of $\mathfrak{S}_k$ (i.e. the Specht-modules) that appear in the isotypic decomposition of $\mathrm{H}^*(V, \mathbb{F})$ as an $\mathfrak{S}_k$-module [13].

There are appplications of these results in a non-equivariant setting as well – for example, improving the best known bounds on the Betti numbers of projections of semii-algebraic sets when the degree is fixed (cf. Theorem 22). The main idea behind this application is that in Theorem 21, the fibered products that appear can be replaced by their quotients under the action of appropriate symmetric groups, and the results on bounding the Betti numbers of these quotients can then be used instead of the standard bounds.

## 9. FEWNOMIALS, ADDITIVE COMPLEXITY

### 9.1. Fewnomial Bounds.

Real algebraic geometry is distinguished from complex algebraic geometry in that it is possible to bound the number of real zeros of a polynomial $P \in \mathrm{R}[X]$ in terms of the number of monomials appearing with non-zero coefficients. By Descarte's rule of signs (see for example [12, Chapter 2]), if the number of monomials in $P$ with non-zero coefficients is $m$, then the number of real roots of $P$ cannot exceed $2m - 1$.

A generalization of such a result to the multi-variate case was was proved by Khovanskiĭ [42, 41] as a special case of a much more general result on Pfaffian functions.

**Theorem 25.** [42] *A system of $k$ polynomials in $\mathbb{R}[X_1, \dots, X_k]$ having $m + k + 1$ distinct monomials has at most $2^{\binom{m+k}{2}}(k+1)^{m+k}$ non-degenerate positive solutions.*

The main technical tool is a generalized Rolle's theorem [42].

It has been a long-standing open problem in real algebraic geometry to improve Khovanskiĭ's bound, and find an analog of Descartes' rule in the multi-variate case. But this turned out to be a difficult problem and only partial progress has been made so far.

Bihan considered real polynomial systems in $\mathbb{R}[X_1, \dots, X_k]$ supported on *circuits*. This means that the exponent vectors appearing in the monomials with non-zero coefficients in the polynomial system is of cardinality $k + 2$ and span $\mathbb{R}^k$. A circuit $\mathcal{C} \subset \mathbb{R}^k$ has a unique minimal affinely dependent subset, and denote by $m(\mathcal{C})$ the dimension of the affine span of this subset.

The following theorem appears in [22].

**Theorem 26.** [22] *The number of positive solutions to a generic real polynomial system supported on a circuit $\mathcal{C} \subset Z^k$ is at most $m(\mathcal{C}) + 1$. Therefore, the number of real solutions to a generic real polynomial system supported on a circuit $\mathcal{C} \subset Z^k$ is at most $2^k(m(\mathcal{C}) + 1)$. These bounds are tight.*

Using a tool from convex geometry, namely Gale-duality, Bihan and Sottile [24] improved the bound in Theorem 25 (with certain added assumptions) to $O(1)2^{\binom{m}{2}}k^m$. They also extended their bound to sums of Betti numbers using stratified Morse theory [25].

Very recently Bihan and Dickenstein [23] obtained an analog of Descartes' rule in the multi-variate case, where the exponent vectors $\mathcal{A} \subset \mathbb{Z}^k$ form a circuit. The matrix of coeffcients of a system of $k$ polynomials in $\mathbb{R}[X_1, \dots, X_k]$ with support $\mathcal{A}$ induces an ordering on $\mathcal{A}$. They show that under certain additional constraints on the set of coefficients, the number of real zeros of a system of $k$ polynomials in $\mathrm{R}[X_1, \dots, X_k]$ with support in a circuit, is bounded by the number of sign variations of the corresponding ordered sequence of coefficients in any affine relation of the given exponents $\mathcal{A}$. The main ingredients of their proof are a generalization of Descartes rule of signs for vectors spaces of analytic real-valued functions, and the classical Gale duality.

Special systems of polynomials where all but one of the polynomials are trinomials were considered by Li, Rojas and Wang [46]. They proved that:

**Theorem 27.** [46] *The number of non-degenerate positive real solutions of the system*

$$F_1 = F_2 = \cdots = F_k = 0,$$

*where each $F_i \in \mathbb{R}[X_1, \dots, X_k]$, and $F_1, \dots, F_{k-1}$ are trinomials, and $F_k$ has at most $m$ monomials in its support, is bounded by $k + k^2 + \cdots + k^{m-1}$.*

Motivated a connection between the problem of proving ng lower bounds in complexity theory and upper bounds on the number of solutions of almost sparse polynomial systems, Koiran, Portier and Tavenas [43] obtained a generalization in another direction.

**Theorem 28.** [43] *Let* $P, Q \in \mathrm{R}[X, Y]$, *where* $0 < \deg(P) \leq d$ *and the number of monomials in* $Q$ *bounded by* $m$. *Then,*

$$b_0(\mathrm{Z}(\{P, Q\}, \mathbb{Z}_2) = O(d^3 m + d^2 m^3).$$

The main technical tool used is bounding the number of real zeros of a sum of a finite number of real analytic functions (in one variable) in terms of the number of real zeros of their Wronskians. No genericity is assumed, but note the restriction that $\deg(P) > 0$.

9.2. **Additive complexity.** A notion of "complexity" that generalizes that of measuring the complexity of a real polynomial by the number of monomials in its support, is that of "additive complexity". Roughly speaking the additive complexity of a polynomial (see [20]) is bounded from above by the number of additions in any straight line program (allowing divisions) that computes the values of the polynomial at generic points of $\mathbb{R}^k$.

Additive complexity of real univariate polynomials was first considered in the context of computational complexity theory by Borodin and Cook [26], who proved an effective bound on the number of real zeros of an univariate polynomial in terms of its additive complexity. This result was further improved upon by Grigoriev [38] and Risler [51] who applied Khovanskiĭ's results on fewnomials [42] to obtain singly exponential (in the additive complexity and $k$) bounds on the number of real solutions of systems of polynomials in $\mathbb{R}^k$.

A surprising fact conjectured in [20], and proved by Coste [27] and van den Dries [58], is that the number of topological types of real algebraic varieties defined by polynomials of bounded additive complexity is finite.

## 10. BOUNDS ON TOPOLOGICAL TYPES

Most of the quantitative bounds described in the previous section were on the topology of semi-algebraic sets. A related but different problem is to prove quantitative bounds on the number of distinct topological types occuring amongst the fibers of some semi-algebraic map.

Suppose that $S \subset \mathbb{R}^k$ is a semi-algebraic set and $\pi_S : S \to \mathbb{R}^n$ be a semi-algebraic map. We can assume without any loss of generality that $\pi_S$ is the restriction to $S$ of the projection map $\pi : \mathbb{R}^{m+n} \to \mathbb{R}^n$, where $k = m + n$.

The following statement is a version of Hardt's triviality theorem.

**Theorem 29** ([39, 29]). *There exists a finite partition of* $\mathbb{R}^n$ *into semi-algebraic sets* $\{T_i\}_{i \in I}$ *such that* $S$ *is semi-algebraically trivial over each* $T_i$.

Theorem 29 implies that for each $i \in I$ and any point $\mathbf{y} \in T_i$, the pre-image $\pi_S^{-1}(T_i)$ is semi-algebraically homeomorphic to $\pi_S^{-1}(\mathbf{y}) \times T_i$ by a fiber preserving homeomorphism. In particular, for each $i \in I$, all fibers $\pi_S^{-1}(\mathbf{y}), \mathbf{y} \in T_i$ are semi-algebraically homeomorphic.

Hardt's theorem is a corollary of the existence of *cylindrical algebraic decompositions* of semi-algebraic sets. Since the decompositions can be effectively computed (see [10]), this implies a double exponential (in $mn$) upper bound on the cardinality of $I$ and hence on the number of semi-algebraic homeomorphism types of the fibres of the map $\pi_S$. Apparently, no better bounds than the double exponential bound are known, even though it seems reasonable to conjecture a single exponential upper bound on the number of semi-algebraic homeomorphism types of the fibres of the map $\pi_S$.

In [17], the authors consider the weaker problem of bounding the number of distinct *semi-algebraic homotopy types*, occurring amongst the set of all fibres of $\pi_S$. The main results of this paper are *single exponential* upper bounds on the *semi-algebraic homotopy types*, occurring amongst the set of all fibres of $\pi_S$.

**Theorem 30.** [17] *Let $\mathcal{P} \subset \mathrm{R}[X_1,\dots,X_m,Y_1,\dots,Y_n]$, with $\deg(P) \le d$ for each $P \in \mathcal{P}$ and cardinality $\#\mathcal{P} = s$. Then, there exists a finite set $A \subset \mathrm{R}^n$, with*

$$\#A \le (s^m nd)^{O(nm)},$$

*such that for every $\mathbf{y} \in \mathrm{R}^n$ there exists $\mathbf{z} \in A$ such that for every $\mathcal{P}$-semi-algebraic set $S \subset \mathrm{R}^{m+n}$, the set $\pi_S^{-1}(\mathbf{y})$ is semi-algebraically homotopy equivalent to $\pi_S^{-1}(\mathbf{z})$. In particular, for any fixed $\mathcal{P}$-semi-algebraic set $S$, the number of different homotopy types of fibres $\pi_S^{-1}(\mathbf{y})$ for various $\mathbf{y} \in \pi(S)$ is also bounded by*

$$(s^m nd)^{O(nm)}.$$

Notice that the bound in Theorem 30 is single exponential in $mn$. The following example shows that the single exponential dependence on $m$ is unavoidable.

**Example 4.** Let $P \in \mathrm{R}[X_1,\dots,X_m] \hookrightarrow \mathrm{R}[X_1,\dots,X_m,Y]$ be the polynomial defined by

$$P := \sum_{i=1}^{m} \prod_{j=1}^{d} (X_i - j)^2.$$

The algebraic set defined by $P = 0$ in $\mathrm{R}^{m+1}$ with coordinates $X_1,\dots,X_m,Y$, consists of $d^m$ lines all parallel to the $Y$ axis. Consider now the semi-algebraic set $S \subset \mathrm{R}^{m+1}$ defined by

$$(P = 0) \ \wedge \ (0 \le Y \le X_1 + dX_2 + d^2X_3 + \cdots + d^{m-1}X_m).$$

It is easy to verify that, if $\pi : \mathrm{R}^{m+1} \to \mathrm{R}$ is the projection map on the $Y$ coordinate, then the fibres $\pi_S^{-1}(y)$, for $y \in \{0,1,2,\dots,d^m - 1\} \subset \mathrm{R}$ are 0-dimensional and of different cardinality, and hence have different homotopy types.

## 11. O-MINIMAL BOUNDS

Various uniform finiteness results in semi-algebraic geometry, such as the finiteness of the number of connected components or more generally Betti numbers etc. of semi-algebraic sets, actually hold in the more general context of o-minimal structures [58]. It is thus a natural question whether there exists extensions of quantitative bounds in semi-algebraic geometry to the more general context of o-minimal geometry.

**11.1. Combinatorial vs algebraic.** Notice that in the bounds on the Betti numbers of semi-algebraic sets discussed in §3, the bounds can be split into a product of two parts. One part that is a function of the number of polynomials used in the definition of the semi-algeraic set, and the other part that depends on the degrees. We refer to the part that depends on the number of polynomials as the the *combinatorial part* of the bounds, while the part that depends on the degrees as the *algebraic part.*

In the context of an arbitrary o-minimal structures (unlike for example, in the case of semi-algebraic or semi-Pfaffian sets), there is no way to associate a "complexity" measure (such as the degree or sparsity of a polynomial) to arbitrary definable functions. Thus, the notion of "algebraic complexity" does not generalize in an useful way to arbitrary o-minimal structures. However, it is possible to recover the "combinatorial part" (i.e. the part of depending only on the number of polynomials) of the various bounds on the Betti numbers of semi-algebraic sets in the general context of o-minimal structures.

We consider definable sets which are defined using unions and intersections of at most $n$ fibers of some fixed definable map (in an *o-minimal structure*) $\pi : T \to \mathrm{R}^\ell$, where $T \subset \mathrm{R}^{k+\ell}$ is a definable set.

One can can obtain tight bounds on the Betti numbers terms of $n$ and $k$. The role of the algebraic complexity is played by a constant that depends only on the particular definable family. In this way, we are able to generalize the notion of combinatorial complexity to definable sets over an arbitrary o-minimal structure. For example, the following result is proved by the PI in [8].

**Definition 1.** Let $\mathcal{S}(\mathrm{R})$ be an o-minimal structure on a real closed field R and let $T \subset \mathrm{R}^{k+\ell}$ be a definable set. Let $\pi_1 : \mathrm{R}^{k+\ell} \to \mathrm{R}^k$ (resp. $\pi_2 : \mathrm{R}^{k+\ell} \to \mathrm{R}^\ell$), be the projections onto the first $k$ (resp. last $\ell$) co-ordinates.

$$\begin{tikzcd} & T \subset \mathrm{R}^{k+\ell} \arrow[dl, "\pi_1"'] \arrow[dr, "\pi_2"] & \\ \mathrm{R}^k & & \mathrm{R}^\ell \end{tikzcd}$$

We will call a subset $S$ of $\mathrm{R}^k$ to be a $(T, \pi_1, \pi_2)$-set if

$$S = \pi_1(\pi_2^{-1}(\mathrm{y}) \cap T)$$

for some $\mathrm{y} \in \mathrm{R}^\ell$, and when the context is clear we will denote $T_{\mathrm{y}} = \pi_1(\pi_2^{-1}(\mathrm{y}) \cap T)$.

**Definition 2.** Let $\mathcal{A} = \{S_1, \ldots, S_n\}$, such that each $S_i \subset \mathrm{R}^k$ is a $(T, \pi_1, \pi_2)$-set. For $I \subset \{1, \ldots, n\}$, we let $\mathcal{A}(I)$ denote the set

$$\bigcap_{i \in I \subset [1\ldots n]} S_i \;\cap \bigcap_{j \in [1\ldots n]\setminus I} \mathrm{R}^k \setminus S_j, \tag{11.1}$$

and we will call such a set to be a basic $\mathcal{A}$-set. We will denote by $\mathcal{C}(\mathcal{A})$ the set of non-empty connected components of all basic $\mathcal{A}$-sets.

We will call definable subsets $S \subset \mathrm{R}^k$ defined by a Boolean formula whose atoms are of the form, $x \in S_i, 1 \le i \le n$, a $\mathcal{A}$-set. A $\mathcal{A}$-set is thus a union of basic $\mathcal{A}$-sets. If $T$ is closed, and the Boolean formula defining $S$ has no negations, then $S$ is closed by definition (since each $S_i$ being homeomorphic to the intersection of

$T$ with a closed set $\pi^{-1}(\mathbf{y})$ for some $\mathbf{y} \in \mathrm{R}^\ell$ is closed) and we call such a set an $\mathcal{A}$-closed set.

Moreover, if $V$ is any closed definable subset of $\mathrm{R}^k$, and $S$ is an $\mathcal{A}$-set (resp. $\mathcal{A}$-closed set), then we will call $S \cap V$ an $(\mathcal{A}, V)$-set (resp. $(\mathcal{A}, V)$-closed set).

**Theorem 31.** [8] *Let $\mathcal{S}(\mathrm{R})$ be an o-minimal structure over a real closed field $\mathrm{R}$ and let $T \subset \mathrm{R}^{k+\ell}$ be a closed definable set. Then, there exists a constant $C = C(T) > 0$ depending only on $T$, such that for any $(T, \pi_1, \pi_2)$-family $\mathcal{A} = \{S_1, \ldots, S_n\}$ of subsets of $\mathrm{R}^k$ the following holds.*

*For every $i, 0 \leq i \leq k$,*

$$\sum_{D \in \mathcal{C}(\mathcal{A})} b_i(D) \leq C \cdot n^{k-i}.$$

Let $V$ be a closed definable subset of $\mathrm{R}^k$ of dimension $k' \leq k$. For any $(T, \pi_1, \pi_2)$-family, $\mathcal{A} = \{S_1, \ldots, S_n\}$, of subsets of $\mathrm{R}^k$, and $I \subset \{1, \ldots, n\}$, we let $\mathcal{A}(I, V)$ denote the set

$$V \;\cap \bigcap_{i \in I \subset [1 \ldots n]} S_i \;\cap \bigcap_{j \in [1 \ldots n] \setminus I} \mathrm{R}^k \setminus S_j, \tag{11.2}$$

and we call a connected component of $\mathcal{A}(I, V)$ a cell of the arrangement restricted to $V$.

Let $\mathcal{C}(\mathcal{A}, V)$ denote the set of all non-empty cells of the arrangement $\mathcal{A}$ restricted to $V$, and we call the cardinality of $\mathcal{C}(\mathcal{A}, V)$ the combinatorial complexity of the arrangement $\mathcal{A}$ restricted to $V$. Similarly, we define the topological complexity of an arrangement $\mathcal{A}$ restricted to $V$ to be the number

$$\sum_{D \in \mathcal{C}(\mathcal{A}, V)} \sum_{i=0}^{k'} b_i(D).$$

We have the following generalization of Theorem 31.

**Theorem 32.** [8] *Let $\mathcal{S}(\mathrm{R})$ be an o-minimal structure over a real closed field $\mathrm{R}$ and let $T \subset \mathrm{R}^{k+\ell}$, $V \subset \mathrm{R}^k$ be closed definable sets with $\dim(V) = k'$. Then, there exists a constant $C = C(T, V) > 0$ depending only on $T$ and $V$, such that for any $(T, \pi_1, \pi_2)$-family, $\mathcal{A} = \{S_1, \ldots, S_n\}$, of subsets of $\mathrm{R}^k$, and for every $i, 0 \leq i \leq k'$,*

$$\sum_{D \in \mathcal{C}(\mathcal{A}, V)} b_i(D) \leq C \cdot n^{k'-i}.$$

Now, let as before $\mathcal{S}(\mathrm{R})$ be an o-minimal structure over a real closed field $\mathrm{R}$, and let $T \subset \mathrm{R}^{k+\ell}$, $V \subset \mathrm{R}^k$ be closed definable sets with $\dim(V) = k'$.

**Theorem 33.** [8] *Let $\mathcal{S}(\mathrm{R})$ be an o-minimal structure over a real closed field $\mathrm{R}$, and let $T \subset \mathrm{R}^{k+\ell}$, $V \subset \mathrm{R}^k$ be closed definable sets with $\dim(V) = k'$. Then, there exists a constant $C = C(T, V) > 0$ such that for any $(T, \pi_1, \pi_2)$-family, $\mathcal{A}$ with $|\mathbb{A}| = n$, and an $\mathcal{A}$-closed set $S_1 \subset \mathrm{R}^k$, and an $\mathcal{A}$-set $S_2 \subset \mathrm{R}^k$,*

$$\sum_{i=0}^{k'} b_i(S_1 \cap V) \leq C \cdot n^{k'} \quad \textit{and,}$$

$$\sum_{i=0}^{k'} b_i(S_2 \cap V) \leq C \cdot n^{2k'}.$$

It is also possible to study cylindrical decomposition quantitatively in the o-minimal setting.

The fact that given any finite family $\mathbb{A}$ of definable subsets of $\mathrm{R}^k$, there exists a cylindrical definable cell decomposition (cdcd) of $\mathrm{R}^k$ adapted to $\mathbb{A}$ is classical (see [28, 58]). The following theorem is a quantitative version of this result.

Since we need to consider several different projections, we adopt the following convention. Given $m$ and $p$, $p \leq m$, we will denote by denote by $\pi_m^{\leq p} : \mathrm{R}^m \to \mathrm{R}^p$ (resp. $\pi_m^{> p} : \mathrm{R}^m \to \mathrm{R}^{m-p}$) the projection onto the first $p$ (resp. the last $m-p$) coordinates.

The following theorem appears in [8].

**Theorem 34.** [8] *Let $\mathcal{S}(\mathrm{R})$ be an o-minimal structure over a real closed field* $\mathrm{R}$*, and let $T \subset \mathrm{R}^{k+\ell}$ be a closed definable set. Then, there exist constants $C_1, C_2 > 0$ depending only on $T$, and definable sets,*

$$\{T_i\}_{i \in I},\ T_i \subset \mathrm{R}^k \times \mathrm{R}^{2(2^k-1)\cdot\ell},$$

*depending only on $T$, with $|I| \leq C_1$,*

*such that for any $(T, \pi_1, \pi_2)$-family, $\mathcal{A} = \{S_1, \ldots, S_n\}$ with $S_i = T_{\mathbf{y}_i}, \mathbf{y}_i \in \mathrm{R}^\ell, 1 \leq i \leq n$, some sub-collection of the sets*

$$\pi_{k+2(2^k-1)\cdot\ell}^{\leq k}\left({\pi_{k+2(2^k-1)\cdot\ell}^{>k}}^{-1}(\mathbf{y}_{i_1}, \ldots, \mathbf{y}_{i_{2(2^k-1)}}) \cap T_i\right),$$
$$i \in I,\ 1 \leq i_1, \ldots, i_{2(2^k-1)} \leq n,$$

*form a cdcd of $\mathrm{R}^k$ compatible with $\mathbb{A}$. Moreover, the cdcd has at most $C_2 \cdot n^{2(2^k-1)}$ cells.*

The combinatorial complexity bound in Theorem 34 compares favorably with the combinatorial parts of similar quantitative results on cylindrical decomposition of semi-algebraic sets (see for instance Section 11.1 in [10]), as well as sub-Pfaffian sets (see main result in [32]).

Moreover, since a doubly exponential dependence on $k$ is unavoidable (see [30]), the complexity bound in Theorem 34 is very close to the best possible.

*Remark* 7. There are several applications of the o-minimal bounds described above. The bound in Theorem 31 implies bounds on the Vapnik-Chervonenkis dimension of definable families of sets in an o-minimal structure, which is related to not having the independence property (NIP) of o-minimal theories. Theorem 34 implies amongst other things that o-minimal theories are *distal*, and the precise quantitative bound leads to combinatorial theorems about graphs with edge relations definable in an o-minimal structure. We refer the reader to the survey by Scanlon [52], and the Bourbaki-seminar article by Starchenko [54] for more details about these connections with combinatorics.

## REFERENCES

1. A. A. Agrachev and R. V. Gamkrelidze, *Computation of the euler characteristic of intersections of real quadrics*, Sov. Math., Dokl. **37** (1988), no. 2, 297–300 (Russian, English). 11
2. A.A. Agrachev, *Homology of intersections of real quadrics*, Sov. Math., Dokl. **37** (1988), no. 2, 493–496 (Russian, English). 11
3. ———, *Topology of quadratic maps and Hessians of smooth maps*, Algebra, Topology, Geometry, Itogi Nauki i Tekhniki, Akad. Nauk SSSR, Vsesoyuz. Inst. Nauchn. Tekhn. Inform.,

vol. 26, VINITI, Moscow, 1988, Translated in J. Soviet Mathematics. 49 (1990), no. 3, 990-1013., pp. 85–124 (Russian, English). 11

4. V. I. Arnold, *Hyperbolic polynomials and Vandermonde mappings*, Funktsional. Anal. i Prilozhen. **20** (1986), no. 2, 52–53. MR 847139 14
5. Boris Aronov, Edward Y Miller, and Micha Sharir, *Eliminating depth cycles among triangles in three dimensions*, Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, 2017, pp. 2476–2494. 5
6. Sal Barone and Saugata Basu, *On a real analog of Bezout inequality and the number of connected components of sign conditions*, Proc. Lond. Math. Soc. (3) **112** (2016), no. 1, 115–145. MR 3458147 7, 8, 9, 10
7. A. I. Barvinok, *On the Betti numbers of semialgebraic sets defined by few quadratic inequalities*, Math. Z. **225** (1997), no. 2, 231–244. MR 98f:14044 11
8. S. Basu, *Combinatorial complexity in o-minimal geometry*, Proc. London Math. Soc. (3) **100** (2010), 405–428, (an extended abstract appears in the Proceedings of the ACM Symposium on the Theory of Computing, 2007). 18, 19, 20
9. S. Basu and M. Kettner, *A sharper estimate on the Betti numbers of sets defined by quadratic inequalities*, Discrete Comput. Geom. **39** (2008), no. 4, 734–746. 11
10. S. Basu, R. Pollack, and M.-F. Roy, *Algorithms in real algebraic geometry*, Algorithms and Computation in Mathematics, vol. 10, Springer-Verlag, Berlin, 2003, Revised version of the first edition online at urlhttp://perso.univ-rennes1.fr/marie-francoise.roy/. MR 1998147 (2004g:14064) 17, 20
11. ———, *Betti number bounds, applications and algorithms*, Current Trends in Combinatorial and Computational Geometry: Papers from the Special Program at MSRI, MSRI Publications, vol. 52, Cambridge University Press, 2005, pp. 87–97. 1
12. ———, *Algorithms in real algebraic geometry*, Algorithms and Computation in Mathematics, vol. 10, Springer-Verlag, Berlin, 2006 (second edition). MR 1998147 (2004g:14064) 1, 5, 15
13. S. Basu and C. Riener, *On the isotypic decomposition of cohomology modules of symmetric semi-algebraic sets: polynomial bounds on multiplicities*, ArXiv e-prints (2015). 14
14. ———, *On the equivariant Betti numbers of symmetric semi-algebraic sets: vanishing, bounds and algorithms*, ArXiv e-prints (2016). 14
15. ———, *Bounding the equivariant Betti numbers of symmetric semi-algebraic sets*, Advances in Mathematics **305** (2017), 803–855. 14
16. S. Basu and A. Rizzie, *Multi-degree bounds on the Betti numbers of real varieties and semi-algebraic sets and applications*, ArXiv e-prints (2015). 3, 10, 12, 14
17. S. Basu and N. Vorobjov, *On the number of homotopy types of fibres of a definable map*, J. Lond. Math. Soc. (2) **76** (2007), no. 3, 757–776. MR 2377123 17
18. Saugata Basu, Dmitrii V. Pasechnik, and Marie-Françoise Roy, *Bounding the Betti numbers and computing the Euler-Poincaré characteristic of semi-algebraic sets defined by partly quadratic systems of polynomials*, J. Eur. Math. Soc. (JEMS) **12** (2010), no. 2, 529–553. MR 2608951 (2011d:14102) 12
19. Saugata Basu and Martí n Sombra, *Polynomial partitioning on varieties of codimension two and point-hypersurface incidences in four dimensions*, Discrete Comput. Geom. **55** (2016), no. 1, 158–184. MR 3439263 7, 9
20. R. Benedetti and J.-J. Risler, *Real algebraic and semi-algebraic sets*, Actualités Mathématiques, Hermann, Paris, 1990. 16
21. Riccardo Benedetti, François Loeser, and Jean-Jacques Risler, *Bounding the number of connected components of a real algebraic set*, Discrete Comput. Geom. **6** (1991), no. 3, 191–209. MR 1090179 (92e:14053) 11
22. Frederic Bihan, *Polynomial systems supported on circuits and dessins d'enfants*, J. Lond. Math. Soc. (2) **75** (2007), no. 1, 116–132. MR 2302733 15
23. Frédéric Bihan and Alicia Dickenstein, *Descartes rule of signs for polynomial systems supported on circuits*, International Mathematics Research Notices (2016), rnw199. 15
24. Frédéric Bihan and Frank Sottile, *New fewnomial upper bounds from Gale dual polynomial systems*, Mosc. Math. J. **7** (2007), no. 3, 387–407, 573. MR 2343138 15
25. ———, *Betti number bounds for fewnomial hypersurfaces via stratified Morse theory*, Proc. Amer. Math. Soc. **137** (2009), no. 9, 2825–2833. MR 2506438 15
26. Allan Borodin and Stephen Cook, *On the number of additions to compute specific polynomials*, SIAM J. Comput. **5** (1976), no. 1, 146–157. MR 0395313 (52 #16110) 16

27. M. Coste, *Topological types of fewnomials*, Singularities Symposium—Łojasiewicz 70 (Kraków, 1996; Warsaw, 1996), Banach Center Publ., vol. 44, Polish Acad. Sci., Warsaw, 1998, pp. 81–92. MR 1677331 (2000b:14075) 16
28. Michel Coste, *An introduction to o-minimal geometry*, Istituti Editoriali e Poligrafici Internazionali, Pisa, 2000, Dip. Mat. Univ. Pisa, Dottorato di Ricerca in Matematica. 20
29. ———, *An introduction to semi-algebraic geometry*, Istituti Editoriali e Poligrafici Internazionali, Pisa, 2000, Dip. Mat. Univ. Pisa, Dottorato di Ricerca in Matematica. 16
30. J. H. Davenport and J. Heintz, *Real quantifier elimination is doubly exponential*, Journal of Symbolic Computation **5** (1988), no. 1/2, 29–35. 20
31. W. Fulton, *Intersection theory*, second ed., Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics], vol. 2, Springer-Verlag, Berlin, 1998. MR 1644323 (99d:14003) 7
32. A. Gabrielov and N. Vorobjov, *Complexity of cylindrical decompositions of sub-pfaffian sets*, J. Pure and Applied Algebra **164** (2001), 179–197. 20
33. A. Gabrielov, N. Vorobjov, and T. Zell, *Betti numbers of semialgebraic and sub-Pfaffian sets*, J. London Math. Soc. (2) **69** (2004), no. 1, 27–43. MR 2025325 (2004k:14105) 13
34. Andrei Gabrielov, *Counterexamples to quantifier elimination for fewnomial and exponential expressions*, Mosc. Math. J. **7** (2007), no. 3, 453–460, 574. MR 2343142 13
35. Andrei Gabrielov and Nicolai Vorobjov, *Approximation of definable sets by compact families, and upper bounds on homotopy and homology*, J. Lond. Math. Soc. (2) **80** (2009), no. 1, 35–54. MR 2520376 6
36. ———, *On topological lower bounds for algebraic computation trees*, Found. Comput. Math. **17** (2017), no. 1, 61–72. MR 3600849 6
37. A. Givental, *Moments of random variables and the equivariant morse lemma*, Russian Mathematical Surveys **42** (1987), no. 2, 275–276. 14
38. D. Yu. Grigor′ev, *Lower bounds in the algebraic complexity of computations*, Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **118** (1982), 25–82, 214, The theory of the complexity of computations, I. MR 659083 (84j:68021) 16
39. R. Hardt, *Semi-algebraic local-triviality in semi-algebraic mappings*, Amer. J. Math. **102** (1980), no. 2, 291–302. MR 564475 (81d:32012) 16
40. A. G. Hovanskiĭ, *Newton polyhedra, and the genus of complete intersections*, Funktsional. Anal. i Prilozhen. **12** (1978), no. 1, 51–61. MR 487230 (80b:14022) 11
41. A. G. Khovanskiĭ, *Fewnomials and Pfaff manifolds*, Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Warsaw, 1983) (Warsaw), PWN, 1984, pp. 549–564. MR 87d:58014 15
42. ———, *Fewnomials*, Translations of Mathematical Monographs, vol. 88, American Mathematical Society, Providence, RI, 1991, Translated from the Russian by Smilka Zdravkovska. MR 92h:14039 15, 16
43. Pascal Koiran, Natacha Portier, and Sébastien Tavenas, *On the intersection of a sparse curve and a low-degree curve: a polynomial version of the lost theorem*, Discrete Comput. Geom. **53** (2015), no. 1, 48–63. MR 3293488 16
44. V.P. Kostov, *On the geometric properties of vandermonde's mapping and on the problem of moments*, Proceedings of the Royal Society of Edinburgh: Section A Mathematics **112** (1989), no. 3-4, 203–211. 14
45. A. Lerario, *Complexity of intersections of real quadrics and topology of symmetric determinantal varieties*, ArXiv e-prints (to appear in J. Eur. Math. Soc.) (2012). 11, 12
46. Tien-Yien Li, J. Maurice Rojas, and Xiaoshen Wang, *Counting real connected components of trinomial curve intersections and m-nomial hypersurfaces*, Discrete Comput. Geom. **30** (2003), no. 3, 379–414. MR 2002964 15
47. J. Milnor, *On the Betti numbers of real varieties*, Proc. Amer. Math. Soc. **15** (1964), 275–280. MR 0161339 (28 #4547) 2, 9
48. J. L. Montaña, J. E. Morais, and Luis M. Pardo, *Lower bounds for arithmetic networks. II. Sum of Betti numbers*, Appl. Algebra Engrg. Comm. Comput. **7** (1996), no. 1, 41–51. MR 1464537 (98j:68060) 6
49. I. G. Petrovskiĭ and O. A. Oleĭnik, *On the topology of real algebraic surfaces*, Izvestiya Akad. Nauk SSSR. Ser. Mat. **13** (1949), 389–402. MR 0034600 (11,613h) 2, 9

50. Cordian Riener, *On the degree and half-degree principle for symmetric polynomials*, J. Pure Appl. Algebra **216** (2012), no. 4, 850–856. MR 2864859 14
51. J.-J. Risler, *Additive complexity and zeros of real polynomials*, SIAM J. Comput. 14 (1985), no. 1, 178–183. MR 774937 (86g:68081) 16
52. T Scanlon, *O-minimality*, Gazette Mathématique **149** (2016), 33–39. 20
53. Edwin H. Spanier, *Algebraic topology*, McGraw-Hill Book Co., New York, 1966. MR 0210112 (35 #1007) 3
54. S. Starchenko, *Nip, keisler measures and combinatorics*, Seminaire BOURBAKI (2015-2016), no. 1114. 20
55. R. Thom, *Sur l'homologie des variétés algébriques réelles*, Differential and Combinatorial Topology (A Symposium in Honor of Marston Morse), Princeton Univ. Press, Princeton, N.J., 1965, pp. 255–265. MR 0200942 (34 #828) 2, 9
56. Vlad Timofte, *On the positivity of symmetric polynomial functions. II. Lattice general results and positivity criteria for degrees 4 and 5*, J. Math. Anal. Appl. **304** (2005), no. 2, 652–667. MR 2126558 (2006a:05168) 14
57. ______, *On the positivity of symmetric polynomial functions. III. Extremal polynomials of degree 4*, J. Math. Anal. Appl. **307** (2005), no. 2, 565–578. MR 2142445 (2006d:05183) 14
58. L. van den Dries, *Tame topology and o-minimal structures*, London Mathematical Society Lecture Note Series, vol. 248, Cambridge University Press, Cambridge, 1998. MR 1633348 (99j:03001) 16, 17, 20
59. Andrew Chi-Chih Yao, *Decision tree complexity and Betti numbers*, J. Comput. System Sci. **55** (1997), no. 1, part 1, 36–43, 26th Annual ACM Symposium on the Theory of Computing (STOC '94) (Montreal, PQ, 1994). MR 1473048 (99b:68106) 6
60. Joshua Zahl, *A Szemerédi-Trotter type theorem in $\mathbb{R}^4$*, Discrete Comput. Geom. 54 (2015), no. 3, 513–572. MR 3392965 7, 9

DEPARTMENT OF MATHEMATICS, PURDUE UNIVERSITY, WEST LAFAYETTE, IN 47906, U.S.A.
*E-mail address*: `sbasu@math.purdue.edu`

# Un langage qui donne les boules

Françoise Delon

Un corps algébriquement clos ou un corps réels clos élimine les imaginaires : les quotients par des relations d'équivalence définissables peuvent être définissablement assimilés à des éléments du corps. Il en va différemment pour les corps valués algébriquement clos. Ainsi une valuation ou un élément du corps de restes n'admet pas de représentant canonique dans le corps : il n'existe pas d'injection définissable $vK \hookrightarrow K^n$ ou $K/v \hookrightarrow K^m$. Haskell, Hrushovski et Macpherson ont montré qu'il est suffisant d'ajouter une infinité de sortes, dites géométriques, qui sont des images linéaires de certains produits de boules. Des langages naturels où considérer les corps valués sont donc les suivants, rangés par richesse croissante :

$\mathcal{L}_{div} = \mathcal{L}_{ring} \cup \{div\}$, à une seule sorte, pour les éléments du corps valué,

$\mathcal{L}_{\Gamma}$, à deux sortes, pour les éléments du corps valué et ceux du groupe de valuation,

$\mathcal{L}_{\Gamma R}$, à trois sortes, pour les éléments du corps valué, du groupe de valuation et du corps résiduel,

$\mathcal{L}_{Boules}$, à trois sortes, pour les éléments du corps valué, les boules ouvertes et les boules fermées,

$\mathcal{L}_{G}$ avec une infinité de sortes, les sortes géométriques.

Le langage $\mathcal{L}_{\Gamma R}$ est inclus dans $\mathcal{L}_{Boules}$ puisque les valuation peuvent être identifiées aux boules fermées contenant 0 et les restes aux boules ouvertes de rayon 1 contenues dans la boule fermée de rayon 1 contenant l'origine. Le langage que nous allons décrire, $\mathcal{L}_{proj}$, est intermédiaire entre $\mathcal{L}_{\Gamma}$ et $\mathcal{L}_{Boules}$. Il a trois sortes, pour les éléments du corps, ceux du groupe de valuation et pour les boules fermées de l'anneau de valuation. La présence de l'inverse du corps rend nécessaire de se limiter à des boules de rayon borné si l'on veut séparer boules ouvertes et boules fermées. En effet, si $B := \{x \in K; v(x) \geq \gamma\}$, alors $(B \setminus \{0\})^{-1} = \{x \in K; v(x) \leq -\gamma\}$, dont le complémentaire est une boule ouverte, définissable sans quantificateurs à partir de $B$. L'intérêt de $\mathcal{L}_{proj}$ est une présentation uniforme des anneaux quotients $O/\gamma O$, où $O$ est l'anneau des entiers de la valuation et $\gamma O$ l'idéal des éléments de valuation $\geq \gamma$.

Tous ces langages ont le même pouvoir d'expression, au sens où ils ont les mêmes imaginaires.

Adaptée aux corps $p$-adiquement clos, cette approche fournit une axiomatisation uniforme des anneaux quotients $\mathbb{Z}/p^n\mathbb{Z}$ (à $p$ fixé, pour commencer). C'est l'objet d'un travail avec Franziska Jahnke.

# Élimination des quantificateurs dans un corps valué algébriquement clos dans le langage $\mathcal{L}_{proj}$

Françoise Delon

Résumé. Nous considérons les corps valués dans un langage à trois sortes, pour les éléments du corps, du groupe de valuation, et pour les boules fermées de l'anneau de valuation $O$. Nous présentons une expansion par définition dans ces sortes qui permet d'éliminer les quantificateurs dans les corps valués algébriquement clos de caractéristique résiduelle nulle. Nous obtenons donc la description de la structure induite par le corps valué sur l'ensemble des boules de $O$, qui n'est autre que l'arbre canonique de la $C$-relation sur $O$.

## 1 Définition du langage $\mathcal{L}_{proj}$

Dans un corps valué, une boule $B$ contenant 0 et de rayon valuationnel positif est un idéal de l'anneau de valuation $O$. L'ensemble des boules de $O$ de même rayon porte donc une structure d'anneau, isomorphe à $O/B$, et l'ensemble de toutes les boules de $O$ apparaît comme le système projectif constitué par tous ces anneaux, lorsque le rayon varie. Cela explique l'introduction des objets définissables suivants.

On commence avec le langage $\mathcal{L}_{val}$ à deux sortes classique des corps valués, qui comprend :
- une sorte $C$ pour les éléments du corps, munie du langage $\{0, 1, +, -, \cdot, ^{-1}\}$
- une sorte $\Gamma$ pour les valuations, munie du langage $\{0, +, -, \leq, \infty\}$
- un symbole de fonction pour la valuation $v : C \to \Gamma$.

Il est connu que les CACV non trivialement éliminent les quantificateurs dans $\mathcal{L}_{val}$. Le langage $\mathcal{L}_{proj}$ que nous introduisons maintenant est une extension par définition de ce langage dans laquelle ils les éliminent également, du moins en caractéristique résiduelle nulle.

Introduisons les objets définissables suivants : $O := \{x \in C; v(x) \geq 0\}$, $\Gamma^+ := \Gamma^{>0}$, et pour chaque $\gamma \in \Gamma^+$, $\gamma O := \{x \in C; v(x) \geq \gamma\}$ et $O_{<\gamma} := O/\gamma O$. Munis des projections $\pi_{\gamma,\delta} : O_{<\gamma} \to O_{<\delta}$, $x + \gamma O \mapsto x + \delta O$, pour $\gamma > \delta$ dans $\Gamma^+$, les $O_{<\gamma}$ constituent un système projectif d'anneaux. Par abus de notation, nous écrirons également, pour $\gamma > \delta > 0$ dans $\Gamma$ et $x \in \gamma O$, $x + \delta O := \pi_{\gamma,\delta}(x)$. Chaque $O_{<\gamma}$ porte la trace de la valuation : c'est l'application $v_\gamma : O_{<\gamma} \to \Gamma_{<\gamma} := \{x \in \Gamma; 0 \leq x < \gamma\} \cup \{\infty\}$, $x + \gamma O \mapsto v(x)$ si $x \notin \gamma O$, et qui envoie 0 sur l'infini. Elle satisfait formellement les axiomes habituels d'une valuation, mais prend ses valeurs non plus dans un groupe (sauf pour $\gamma = \infty$) mais dans un segment initial du groupe ordonné $vK$, union l'infini. Sauf pour $\gamma = \infty$ l'anneau $O_{<\gamma}$ n'est plus intègre, mais il hérite de l'anneau de valuation $O$ une propriété de divisibilité :

$$O_{<\gamma} \models \forall x, y\ \exists z\ [\ \infty > v_\gamma(x) \geq v_\gamma(y) \to x = zy\ ].$$

Si $z$ et $z'$ dans $O_{<\gamma}$ vérifient tous deux $x = zy = z'y$ alors $v_\gamma(z - z') \geq \gamma - v_\gamma(y)$. Donc $z + (\gamma - v_\gamma(y))O = z' + (\gamma - v_\gamma(y))O$, nous noterons $xy^{-1}$ cet élément de $(\gamma - v_\gamma(y))O$.

Le langage $\mathcal{L}_{proj}$ est composé de $\mathcal{L}_{val}$ auquel on ajoute une nouvelle sorte pour l'ensemble des boules fermées de $O$, à savoir $\mathcal{B} := \{x + \gamma O; x \in O, \gamma \in \Gamma^+\}$. Cette sorte représente donc l'union disjointe des anneaux $O_{<\gamma}$. Le langage $\mathcal{L}_{proj}$ donne ainsi une définition des $O_{<\gamma}$ qui soit sans quantificateurs et uniforme en $\gamma$. On ajoute en outre des symboles pour
- la somme (uniformément présentée et partielle) $\mathcal{B} \times \mathcal{B} \to \mathcal{B}$ : $(x + \gamma O) + (y + \delta O)$ est défini exactement lorsque $\gamma = \delta$ et vaut alors $(x + y) + \gamma O$ ;
- le symétrique $\mathcal{B} \to \mathcal{B}$, $x + \gamma O \mapsto (-x) + \gamma O$,
- la valuation $\mathcal{B} \to \Gamma^+$, $x + \gamma O \mapsto v_\gamma(x) \in \Gamma_{<\gamma}$,
- la multiplication (totale) : $(x + \gamma O, y + \delta O) \mapsto xy + \varepsilon O$, où $\varepsilon := min\{\gamma + v_\delta(y + \delta O), \delta + v_\gamma(x + \gamma O)\}$ ;
- l'inverse binaire (partiel) qui à $(x + \gamma O, y + \delta O)$ avec $v_\gamma(x) \leq v_\delta(y) \neq \infty$, associe $yx^{-1} \in O_{<\delta - v_\gamma(x)}$ ;
- la fonction $O \times \Gamma^+ \to \mathcal{B}$, $(x, \gamma) \mapsto x + \gamma O$, qui représente de façon uniforme la projection $\pi_\gamma : O \to O_{<\gamma}$,
- et enfin pour la fonction $\mathcal{B} \times \Gamma \to \mathcal{B}$, $(x + \gamma O, \delta) \mapsto x + \delta O$, définie pour $\gamma > \delta$ dans $\Gamma^+$ et qui représente de façon uniforme les projections $\pi_{\gamma,\delta} : O_{<\gamma} \to O_{<\delta}$

pour tous $\gamma \geq \delta$ dans $\Gamma^+$.

Tout corps valué, disons $K$, porte sa $\mathcal{L}_{proj}$-structure $K_{proj} := ((O_K)_{<\gamma})_{\gamma \in \overline{vK}^+}$, notation dans laquelle nous n'explicitons pas les fonctions $\pi_{\gamma,\delta}$. Plus généralement, pour un sous-anneau $A$ de $O$, $A_{proj}$ désigne le système projectif des $A_{<\gamma} := A/A \cap \gamma O$, indexé par $\gamma \in \overline{vA}$. Dans de tels systèmes, les fonctions $\pi_{\gamma,\delta}$ sont surjectives. Ce n'est évidemment pas le cas dans une $\mathcal{L}_{proj}$-sous-structure quelconque de $K_{proj}$.

**Théorème 1.** *Les cacv non trivialement et de caractéristique résiduelle nulle éliminent les quantificateurs dans le langage* $\mathcal{L}_{proj}$.

[Commentaire perso.
Pour $x \in A_\gamma, x \neq 0$, $y \in A_\delta, y \neq 0$ et $z \in A_{\gamma+v(y)}$ avec $\gamma + v(y) \leq \delta + v(x)$, la relation $xy = z$ doit être définissable sans quantificateurs. Que le langage contienne nécessairement la fonction $(x, y) \mapsto z$ ou $(x, z) \mapsto y$ n'est pas clair pour moi. N'est pas non plus clair ce que j'utilise de ces deux fonctions dans ma preuve, certainement l'inverse binaire faible $A_\gamma \times A_\gamma \to A_{<\gamma}$.]

Les anneaux de la forme $O_{<\gamma}$ portent une application, que nous appellerons valuation tronquée, qui satisfait les axiomes habituels d'une valuation et prend ses valeurs dans $\bar{G} := G \dot{\cup} \{\infty\}$ où $G$ est une structure plus générale que le semi-groupe des éléments positifs d'un gao.

# 2 Pré-semi-groupes abéliens ordonnés tronqués

**Définition 2.** *On appelle* semi-GAO *le semi-groupe des éléments positifs d'un gao,* pré-semi-gao *une sous-structure d'un semi-GAO dans le langage* $\{0, +, <\}$.
*Un* presse-GAOT *(pour : pré-(semi-groupe abélien ordonné) tronqué) est ou bien un pré-semi-gao, ou bien proprement tronqué; l'addition n'est alors que partielle et on ajoint à* $G$ *un nouvel élément, noté « $\infty$ », sur lequel on envoie toutes les sommes indéfinies, pour donner une structure* $\bar{G} := G\dot{\cup}\{\infty\}$ *du langage* $\{0, \infty, +, <\}$*, dont on requiert qu'elle soit modèle des axiomes (universels) suivants :*
*1.* $+$ *est une loi binaire, associative et commutative, dont* $0$ *est élément neutre,*
*2.* $<$ *est un ordre linéaire, dont* $0$ *est le premier élément, et* $\infty$ *le dernier,*
*3.* $x + \infty = \infty$*,*
*4. régularité initiale :*
$x < y \rightarrow (x + z < y + z \vee x + z = y + z = \infty)$.

**Exemple.** Un segment $[0, a[$ ou $[0, a[$ dans un groupe abélien ordonné, équippé de l'« addition tronquée ».

**Quelques constatations.**
0. $x \leq y \rightarrow x + z \leq y + z\,;\ x \leq x + z$.
1. $G$ est un monoïde régulier ordonné ssi il est stable par $+$.
2. Si $I$ est un segment initial de $G$, $I \cup \{\infty\}$ porte une structure canonique de presse-GAOT : $a + b$ est défini comme dans $G$ si c'est un élément de $I$ et est sinon défini comme étant $\infty$.
3. Si $I$ est un segment final de $\bar{G}$, $\{0\} \cup I$ est un presse-gaotte.

**Fait 3.** *Tout segment initial* $I$ *non vide et stable par addition de* $G$ *définit sur* $G$ *une relation d'équivalence compatible avec la loi et l'ordre (c'est à dire à classes convexes) :* $a$ *et* $b < a$ *sont équivalents lorsque* $I$ *contient un élément* $i$ *pour lequel* $a \leq b + i$*. La structure quotient est un presse-GAOT noté* $G/I$*.*

**Notation.** Par régularité initiale, l'ensemble $\Delta_G := \{x \in G; \bigwedge_{n \in \mathbb{N}} nx \neq \infty\}$ est convexe et stable par addition. C'est le sous-monoïde régulier maximal de $G$.

**Définition 4.** *Nous dirons que* $G$ *est* archimédien *lorsque pour tous* $x, y$ *dans* $G$ *différents de* $0$*, il existe un entier* $n$ *pour lequel* $nx > y$*, et* brutalement archimédien *lorsque pour tout* $x > 0$ *dans* $G$*, il existe un entier* $n$ *pour lequel* $nx = \infty$.

**Fait 5.** *La structure quotient* $G/\Delta_G$ *est brutalement archimédienne. En particulier, si* $G$ *est stable par addition, alors* $\Delta_G = G$.

Toujours d'après l'axiome 4, pour $a, b \in G$ avec $a \geq b$, il existe au plus un $x \in G$ vérifiant $a = b + x$ ; s'il existe cet élément sera noté $a - b$. En particulier $a - a$ existe et est égal à 0.

**Définition 6.** *Nous dirons que* $G$ *est* clos, *ou bien qu'il est un GAOT, lorsque* $a - b$ *existe pour tous* $a \geq b \in G$ *; on a alors*
$a + (b - c) = (a - c) + b = (a + b) - c$ *pour tous* $a, b \geq c \in G$.
*Pour* $x \in G$*, si* $G$ *n'est pas un monoïde et si* $\inf\{y \in G; x + y = \infty\}$ *existe, on le note* $\infty - x$*. On dit que* $G$ a une fonction $\infty - x$ *si* $\infty - x$ *existe pour tout* $x \in G$ *et vérifie* $\infty - (\infty - x) = x$.

**Remarque.** Si $mz \neq 0$, on a $(m + 1)z - mz \leq z$. Mais l'égalité peut être stricte lorsque $(m + 1)z = \infty$. (Lorsque $(m + 1)z = \infty$, $(m + 1)z - mz = z$ ssi $z = \min\{x \in G; (m + 1)x = \infty\}$.)

**Fait 7.** *Un presse-GAOT* $G$ *clos a les propriétés suivantes.*

1. $\Delta_G$ *est la partie positive d'un gao (ie c'est un semi-gao).*
2. *Si* $G$ *est de plus archimédien, alors ou bien il est isomorphe à un segment initial de* $(\mathbb{N}, +)$ *muni de l'addition tronquée, ou bien il est dense et de plus, pour tout* $\gamma \in G$ *et tout entier* $n$ *non nul,* $\gamma \pm nG := (\gamma + nG) \cup \{g \in G; \gamma \in g + nG\}$ *est dense dans* $G$.

Démonstration. 1 est clair. Supposons maintenant $G$ archimédien.
- Si 0 a un successeur, « 1 », dans $G$, pour tout $x \in G$, par clôture de $G$, il n'y a rien entre $x$ et $x + 1$. Ainsi, ou bien $G$ est le monoïde $(\mathbb{N}, +)$, ou bien il existe un premier entier $> 0$ tel que $n1$ soit nul dans $G$ et alors $G$ est $(\mathbb{N}, +)$ tronqué à $n$.
- Si 0 est limite, disons d'une suite strictement décroissante $(x_\alpha)_{\alpha < \alpha_G}$ de $G$, définissons la suite $(y_n)_{n \in \omega}$ comme suit : $y_0 = x_0$ et $y_{n+1}$ est le plus grand $x_\alpha$ tel que $(n + 1)y_{n+1} < y_n$. De cette façon, pour tout entier $n_0$, la suite $(n_0 y_n)_n$ tend vers 0. Pour $x, y \in G$, $0 < x < y < \infty$, il existe donc $z \in G$, $0 < n_0 z < \min\{x, y - x\}$ ; puisque $G$ est archimédien, il existe un entier $N$ tel que $y < N n_0 z$. L'existence d'un entier $m$ vérifiant $m n_0 z \leq x$

et $y \leq (m+1)n_0 z$ impliquerait $y - x \leq n_0 z$, ce qui est une contradiction ; il existe donc un entier $m$ pour lequel $x \leq m n_0 z \leq y$ ; ainsi $n_0 G$ est dense. Par clôture de $G$, l'application $x \mapsto x + \gamma$ définit une bijection entre $\{x \in G; x + \gamma < \infty\}$ et $(\gamma + G) \setminus \{\infty\}$. En conséquence, pour tout $\gamma \in G$, $\gamma + n_0 G$ est dense $\{x \in G; x \geq \gamma\}$. La preuve est donc achevée si, étant donné $x \in G^+$, $x < \gamma$, nous trouvons $g \in G$, vérifiant $n_0 g < \gamma$ et $\gamma - n_0 g \leq x$, ou encore $\gamma - x \leq n_0 g < \infty$. Or $\gamma - x < \gamma$, donc par densité de $n_0 G$ il existe $g \in G$, $\gamma - x \leq n_0 g \leq \gamma$. □

Reprenons tout cela du point de vue des anneaux.

# 3 Anneaux quasi valués

## 3.1 Valuations tronquées

**Définition 8.** *Soit un anneau $A$, un presse-GAOT $G$ et une application $v : A \to G \cup \{\infty\}$. On appelle $v$* valuation tronquée *lorsqu'elle vérifie*

$v(x) = \infty$ *ssi* $x = 0$
$v(x+y) \geq \min\{v(x), v(y)\}$
$v(xy) = v(x) + v(y)$.

*On dira que $A$ est* quasi-valué.

Quelques remarques.

1. Une valuation tronquée $v$ sur un anneau $A$ est une valuation classique ssi $A$ est intègre ssi $vA \setminus \{\infty\}$ est stable par addition.

2. Comme classiquement (mais ici $I$ n'a pas besoin d'être stable par addition), si $I$ est un segment initial de $vA \setminus \{\infty\}$, $P_I := \{x \in A; v(x) > I\}$ est un idéal de $A$, $v$ induit sur $A/P_I$ une valuation tronquée à valeurs dans $\bar{I} := I \cup \{\infty\}$, et $P_I$ est premier ssi $I$ est stable par addition. En particulier pour $I = \{0\}$, $A/P_I$ est un anneau intègre, dont on appelle la caractéristique *caractéristique résiduelle* de $(A, v)$. On note $A/v := A/P_{\{0\}}$ et, pour $x \in A$, $x/v := x + P_{\{0\}}$.

3. Tout segment initial $I$ de $vA \setminus \{\infty\}$ stable par addition définit sur $A$ une valuation tronquée, notée $v/I$, à valeurs dans $vA/I$ ; on dit que $v/I =: w$ est *plus grossière* que $v$, et on note $w \leq v$.

## 3.2 Anneaux henseliens

Soit $(A, v)$ un anneau quasi-valué.

**Définition 9.** *Un polynôme unitaire* $P \in A[X]$ *et* $x \in A$ *sont dits* en position de Hensel *lorsque* $x/v$ *est racine résiduelle simple de* $P/v$.
$(A, v)$ *est dit* henselien *lorsque, pour tout* $P \in A[X]$ *et* $x \in A$ *en situation de Hensel,* $P$ *admet une (unique) racine* $y \in A$ *de reste* $y/v = x/v$.

Soit $P$ unitaire $\in A[X]$, $x, y \in A$ deux racines de $P$ ; on a ainsi dans $A$, $0 = (x-y)(P'(x) + (x-y)\dots)$. Si $v(P'(x)) = 0$ et $v(x-y) > 0$, alors $\infty = v(x-y) + v(P'(x)) = v(x-y)$, donc $x = y$, cad qu'une racine résiduelle simple de $P$ se relève en au plus une racine de $P$. Cette unicité est une chose importante puisque les anneaux considérés ne sont pas intègres.

**Proposition 10.** *Soit un anneau quasi-valué* $(A, v)$ *dans lequel tous les éléments de valuation nulle sont inversibles.*
*1. Si* $v$ *est brutalement archimédienne, alors elle est henselienne.*
*2. Soit* $I$ *un segment initial de* $vA$ *et* $\bar{v}$ *la valuation à valeurs dans* $\bar{I}$ *induite par* $v$ *sur* $A/P_I$*. Alors* $\bar{v}$ *est henselienne si* $v$ *l'est.*
*3. Si* $I$ *est de plus stable par addition et* $w := v/I$*, alors* $v$ *est henselienne ssi* $\bar{v}$ *(qui est classique) et* $w$ *le sont.*

Démonstration. 1. Si $P$ et $x$ sont en situation de Hensel et qu'on prend $y := x - P'(x)^{-1}P(x) \in A$, alors $P(y) = P(x)P'(x)^{-1}(y-x)z$ avec $z \in A$, et $v(P(y)) \geq v(x-y)$. En posant $x_0 = x$ et $x_{n+1} = x_n - P'(x_n)^{-1}P(x_n)$, on définit comme classiquement une suite $(x_n)_{n \in \mathbb{N}}$ de $A$ qui vérifie $x_n/v = x/v$ et $v(x_{n+1} - x_n) = v(P(x_n)) \geq nv(P(x))$, et donc stationne sur un zéro de $P$ relevant $x/v$. Cela montre le premier point de l'énoncé.
2. Si $\bar{P} \in (A/P_I)[X]$ et $\bar{x} \in A/P_I$ sont en situation de Hensel pour $\bar{v}$ et qu'on relève $\bar{P}$ en $P \in A[X]$ unitaire et $\bar{x}$ en $x \in A$, alors $P$ et $x$ sont en situation de Hensel pour $v$.
3. Considérons d'abord le cas où $I = \Delta_G$. Parce que $w$ est brutalement archimédienne et grâce au point 1, l'équivalence à prouver devient : $v$ henselienne ssi $\bar{v}$ l'est. Le point 2 a établi une des directions, reste l'autre. Si $P$ et $x$ sont en situation de Hensel pour $v$, $P/w$ et $x/w$ le sont pour $\bar{v}$, donc il existe une racine $y \in A/w$ de $P/w$ qui vérifie $y/\bar{v} = (x/w)/\bar{v} = x/v$, et qui est nécessairement simple ; ainsi $P'(t)$ est inversible dans $A$ pour tout relèvement $t \in A$ de $y$, ce qui permet de construire une racine $z \in A$ de $P$ de reste $z/w = y$, donc $z/v = x/v$. Dans le cas général, puisque $I$ est stable par addition, il est

contenu dans $\Delta_G$. Soit $\bar{v}_{\Delta_G}$ la valuation à valeurs dans $\Delta_G$ induite par $v$ sur $A/P_{\Delta_G}$ et $\bar{w} \leq \bar{v}_{\Delta_G}$ à valeurs dans $\Delta_G/I$. Par le point 2, $v$ est henselienne ssi $\bar{v}_{\Delta_G}$ l'est, ssi $(A/P_I, \bar{v})$ et $(A/P_{\Delta_G}, \bar{w})$ le sont (en effet toutes les valuations en présence sont ici classiques, et l'équivalence 3 est connue dans le cas classique). Par le point 2, $(A/P_{\Delta_G}, \bar{w})$ est henselienne ssi $(A, w)$ l'est. □

**Proposition 11.** *Un anneau quasi-valué $(A, v)$ vérifiant $v(x) \leq v(y) \Rightarrow x|y$ admet une clôture henselienne, unique à isomorphisme près au-dessus de $A$. À l'intérieur d'un anneau valué, cette clôture de $A$ est unique (pas seulement à isomorphisme près).*

Démonstration. Si $A$ est intègre, alors c'est un anneau de valuation et la théorie classique montre l'existence et l'unicité de la clôture henselienne de $FrA$. Soit dans le cas général $w := v/\Delta_{vA}$, qui est brutalement archimédienne. Si $B$ est la clôture henselienne de $\bar{A} := A/w$ (intègre), $B$ est obtenu par adjonctions successives d'élément de la forme $\bar{x}$ racine de $\bar{P} \in \bar{A}[X]$, $\bar{x}$ et $\bar{P}$ en situation de Hensel pour $\bar{v}$. On peut de plus supposer $\bar{P}$ de degré $d$ minimal sans racine relevant $\bar{x}/\bar{v}$. Il y a alors une unique façon de valuer $A[x]$ où $x$ annule $P \in A[X]$ unitaire relevant $\bar{P}$ et $x/v = \bar{x}/\bar{v}$. En effet, si $Q \in A[X]$ est de degré $< d(P)$ et si $q \in A$ est un coefficient de valuation ($v$ !) minimale de $Q$, par l'hypothèse sur $A$, $Q$ s'écrit $qR$ avec $R \in A[X]$ et $(R/w)(x/w) \neq 0$. L'extension de $A$ obtenue en ajoutant de tels $x$ de façon itérée est la clôture henselienne (pour $v$ !) de $A$. □

**Proposition 12.** *Soit $(A, v)$ un anneau quasi-valué de caractéristique résiduelle nulle, et $w := v/\Delta_{vA}$.*

1. *Supposons que pour tous $x, y \in A$ avec $w(x) = w(y) = 0$ et $v(x) \leq v(y)$, $x$ divise $y$ dans $A$. Soit $A_0$ un sous-anneau de $A$ trivialement valué par $w$ et maximal pour cette propriété. Alors $(A_0, v)$ est un relèvement de $(A/w, v/w)$.*
2. *Supposons que pour tous $x, y \in A$ avec $w(x) = w(y)$ et $v(x) \leq v(y)$, $x$ divise $y$ dans $A$. Soit $A_0$ un relèvement de $A/w$ dans $A$ comme ci-dessus. Soit un anneau $C$ contenant $A$ et portant une valuation tronquée prolongeant $w$, et un anneau intègre $B$, $A_0 \subseteq B \subseteq C$ et sur lequel $w$ est triviale. Alors le sous-anneau de $C$ engendré par $B$ au-dessus de $A$ est isomorphe à $A \otimes_{A_0} B$, il y a une unique façon de le valuer en étendant les valuations sur $A$ et $B$, et son anneau résiduel modulo $w$ est canoniquement isomorphe à $B$.*

Démonstration.

**Fait 13.** *Un anneau de valuation de caractéristique* 0 *(ça n'est pas important ici) s'obtient à partir d'un sous-anneau* $A' := \mathbb{Q}[t_i, b_j; i \in I, j \in J]$, *où les* $t_i$ *sont algébriquement indépendants et les* $b_j$ *entiers sur les* $t_i$ *(et les* $b_j$ *antérieurs), en lui ajoutant des éléments de la forme* $xy^{-1}$ *où* $x, y \in A'$ *et* $v(x) \geq v(y)$. $\dashv$

L'énoncé précédent, plus le fait que $(A, w)$ est brutalement archimédien, permettent de reproduire la caractérisation classique des relèvements du corps résiduel dans un corps henselien de caractéristique résiduelle nulle.

Pour la seconde partie également, la preuve classique s'adapte. Le passage au reste $A \to A/w, x \mapsto x/w$, induit un isomorphisme $A_0 \simeq A/w$, dont on nomme $\sigma$ l'inverse. Montrons par induction sur $n$ qu'il y a une seule façon de valuer une somme $x := \Sigma^n a_i b_i$ où $a_i \in A$ et $b_i \in B$ ; on peut supposer qu'on a $w(a_1) = \cdots = w(a_m) < w(a_{m+1}), \cdots, w(a_n)$ pour un entier $m \leq n$, et $v(a_1) \leq v(a_i)$ pour tout $i \leq m$ ; on a alors

$$\Sigma^n a_i b_i = a_1(b_1 + \Sigma_2^m (a_i a_1^{-1}) b_i) + \Sigma_{m+1}^n a_i b_i =$$

$$a_1([b_1 + \Sigma_2^n b_i \sigma(a_i a_1^{-1}/w)] + \Sigma_2^n b_i (a_i a_1^{-1} - \sigma(a_i a_1^{-1}/w))) + \Sigma_{m+1}^n a_i b_i;$$

si le terme entre crochets $y$ (notez que $y \in B$) n'est pas nul, $v(x) = v(a_1) + v(y)$ et sinon nous sommes ramenées à valuer une somme de $n-1$ termes. Cela montre également $A[B_0] \simeq A \otimes_{A_0} B$ et la dernière assertion. ☐

## 3.3 Les questions de divisibilité

Soit $(A, v)$ un anneau quasi-valué. Comme on l'a remarqué lors de la définition de $\mathcal{L}_{proj}$, pour $x, y, z \in A$, si $x = yz$, la donnée de $x$ et $y$ détermine $z \mod \delta$ pour tout $\delta \in vA$ tel que $v(y) + \delta < \infty$. Dans cette section, nous utilisons systématiquement cette unicité.

Un presse-GAOT $G$ étendant $vA$ permet de définir le système projectif des anneaux quasi-valués $(A_{<\gamma})_{\gamma \in G}$, où $A_{\geq\gamma} := \{x \in A; v(x) \geq \gamma\}$ et $A_{<\gamma} := A/A_{\geq\gamma}$, équipé de la multiplication et de la division mixtes (partielles).

**Définition 14.** *On dit que* $A$ *est* inversible *lorsqu'il vérifie :* $v(x) \leq v(y) \Rightarrow x$ *divise* $y$.
*Il est dit* 0-inversible *lorsqu'il vérifie :* $v(x) = v(y) \Rightarrow x$ *divise* $y$.

*Supposons que dans $A$ tous les éléments de valuation nulle sont inversibles. On appelle* système d'inverses binaires *de $A$ un système projectif d'anneaux quasi-valués $\mathcal{B} = (B_\delta, \pi_{\delta,\varepsilon})_{\varepsilon \leq \delta \in \Gamma}$, où :*
*- $\Gamma$ est un semi-GAOT étendant $vA$, avec fonction $\infty - x$,*
*- la multiplication et les inverses binaires sont définies partout dans le système projectif des $B_\delta$, au sens où, pour tous $x \in B_\gamma$ et $y \in B_\delta$, on a des éléments $xy \in B_\varepsilon$ où $\varepsilon := min\{\gamma + v(y), \delta + v(x)\}$, et $yx^{-1} \in B_{\delta - v(x)}$ si $v(x) \leq v(y)$, avec l'associativité et la distributivité de la multiplication partielle sur la somme, et les relations $(xy^{-1})y = x$, $(xy^{-1})z^{-1} = x(yz)^{-1}$, etc, chaque fois que tous les termes de l'égalité sont définis,*
*- $\mathcal{B}$ étend le système projectif associé à $A$ et $\Gamma$,*
*- $B_\infty = A$.*

1. Pour une valuation $v$ classique, $A$ est inversible ssi $A$ est l'anneau de l'unique valuation sur le corps des quotients de $A$ qui prolonge $v$.
2. $A$ inversible ssi ($A$ 0-inversible et $vA$ clos).
3. Si $A$ est 0-inversible alors, pour tous $a \in A$ et $\delta \in vA$ vérifiant $\delta + v(a) < \infty$, la multiplication par $a$ définit un isomorphisme $[a] : (A_{=\delta}, +) \simeq (A_{=\delta+v(x)}, +)$, où $A_{=\delta} := A_{\geq\delta}/A_{>\delta}$. On exprime cela en disant que $A$ « a toutes ses fibres égales ». Réciproquement, si dans $A$ tous les éléments de valuation nulle sont inversibles, alors $A$ est 0-inversible ssi toute ses fibres sont égales.
4. Soit $I$ un segment initial de $vA \setminus \{\infty\}$. Si $(A, v)$ est inversible, alors $(A/P_I, \bar{v})$ l'est. Idem avec 0-inversible.
5. Si $A$ est inversible, $(A_{<\gamma})_{\gamma \in vA}$ est un système d'inverses binaires de $A$.
6. Que $A$ soit inversible n'implique pas que $vA$ ait une fonction $x \mapsto \infty - x$. L'hypohèse que $\Gamma$ ait une telle fonction n'est faite ici que pour simplifier les énoncés, et bien sûr parce qu'elle sera réalisée dans le contexte qui nous intéresse.
7. Un système d'inverses binaires suppose par définition que dans $A$ tout élément de valuation nulle soit inversible. En particulier $A/v$ est un corps.

Le but est maintenant, étant donné un anneau quasi-valué équipé d'un système d'inverses binaires, de le clore par inverses binaires de façon canonique. Nous procédons par étapes.

**Lemme 15.** *Soient donnés un anneau quasi-valué $A$ équipé d'un système d'inverses binaires $\mathcal{B} = (B_\gamma)_{\gamma \in \Gamma}$, où $\Gamma$ est brutalement archimédien. Définissons $B_0 :=$ inj lim $B_\gamma$. Il existe alors des extensions $A'$ de $A$ et $\mathcal{B}'$, également indexé par $\Gamma$, de $\mathcal{B}$, telles que $\mathcal{B}'$ soit un système d'inverses binaires de $A'$,*

*$A'/v = B_0 = B'_\gamma/v$ pour tout $\gamma \in \Gamma$, et $vA' = vA$, et dont le type d'isomorphisme au-dessus de $\mathcal{B}$ est uniquement déterminé. En particulier $A'$ est* 0-*inversible.*

Démonstration. Soit à ajouter à $A$ un élément $z$ vérifiant $v(z) = 0$ et $z/v \notin A/v$. Si $z/v$ est algébrique sur $A/v$, on considère le polynôme minimal unitaire $p \in A/v[X]$ de $z/v$ sur $A/v$, qu'on relève en $P \in A[X]$, unitaire de même degré ; puisque $\Gamma$ est brutalement archimédien, l'anneau à construire est henselien, donc dès qu'on ajoute $z/v$ à $A/v$, on ajoute à $A$ la racine de $P$ de reste $z/v$ ; on suppose donc désormais $P(z) = 0$.

Assertion : le type d'isomorphisme au dessus de $A$ de l'anneau quasi-valué $A[z]$ est uniquement déterminé par $P$, $z/v$ et le système des inverses binaires.

Preuve. On doit savoir valuer $Q(z)$ pour $Q \in A[X]$, $Q$ de degré < celui de $P$ lorsque $z/v$ est algébrique sur $A/v$. Si $q$ est le coefficient de plus basse valuation de $Q$, $q^{-1}Q$ est défini dans $B_{\infty - v(q)}[X]$, est de degré au plus égal à celui de $Q$, et a un coefficient égal à 1 ; donc $(q^{-1}Q)/v \not\equiv 0$, $(q^{-1}Q(z))/v) \neq 0$, ie $v(q^{-1}Q(z)) = 0$, et $v(Q(z)) = v(q)$. On augmente les $B_\gamma$ de façon correspondante : si $\pi_\gamma P$ a un zéro $z_\gamma \in B_\gamma$ de reste $z/v$, alors $B'_\gamma = B_\gamma$ et $\pi_\gamma(z) = z_\gamma$ ; sinon cela signifie que $z/v \notin B_\gamma/v$, on ajoute alors $\pi_\gamma(z)$ à $B_\gamma$ comme on a ajouté $z$ à $A$, et il y a une unique façon de le faire. Il y a également une et une seule façon d'étendre les multiplication et division mixtes. ☐

**Lemme 16.** *Soit un anneau quasi-valué $A$ dont le corps de restes est de caractéristique nulle. Supposons $A$ équipé d'un système $\mathcal{B}$ d'inverses binaires où $\Gamma$ est brutalement archimédien et vérifiant $A/v = B_\gamma/v$ pour tout $\gamma \in \Gamma$. Alors $A$ admet une unique clôture inversible dont le système projectif se plonge dans $\mathcal{B}$.*

Démonstration. Puisque $\Gamma$ est brutalement archimédien, la clôture cherchée $B$ sera henselienne. Par le lemme précédent on peut supposer $A$ et les $B_\gamma$ 0-inversibles, avec toutes leurs fibres non nulles égales. Nous devons maintenant ajouter à $A$ (et à certains $B_\gamma$) de nouvelles valuations, le corps résiduel commun restera le même. Il en résulte que si un élément $z \in B$ a une valuation $\zeta$ telle que $n\zeta \in vA$ pour un entier $n$ non nul, alors il y a $a \in A$ tel que $v(a - z^n) > v(a) = v(z^n)$, et donc $a \in B^n$.

Nous cherchons à ajouter à $A$ un élément $z$ vérifiant $y = zx$ avec $x, y \in A$, $\gamma := v(x) \leq v(y) =: \delta$ et $v(z) \notin vA$. Cet élément est connu avec d'autant plus de précision que $\gamma$ est petit. Un premier élément de complication vient de ce qu'il n'y a pas nécessairement de $\delta$ et $\gamma$ minimaux réalisant $v(z) = \delta - \gamma$. Le deuxième est illustré par l'exemple suivant : soit les anneaux $R := \mathbb{Q}[X^q, X^3 Z, XZ^2; q \in \mathbb{Q}^{>0}]$ et $A := R/R_{\geq\gamma}$, $A$ quasi-valué de façon à ce que $v\mathbb{Q} = 0$, $vX = 1$ et $v(Z)$ est irrationnel, $0 < v(Z) < 1$, (donc $X$ et $Z$ sont algébriquement indépendants sur $\mathbb{Q}$) ; pour tout rationnel $\gamma > \gamma$, l'équation $XZ^2 = Xz^2$ (en $z$) détermine $\pm z$ au-dessus de $A_\gamma$ avec plus de précision que ne le fait l'équation $X^3 Z = X^3 z$.

Au total, considérons l'information sur $z$ contenue dans les relations suivantes :

. $v(z) \notin vA$ ;

. si $nv(z) \in vA$ et $iv(z) \notin vA$ pour $i \in \{1, \ldots, n-1\}$, alors $z^n = y^{(n)} \in A$ ;

. $\bigwedge_{\alpha<\alpha_i} y_\alpha^{(i)} = z^i x_\alpha^{(i)}$ pour tout entier $i$ ($i < n$ si $z^n \in A$ et $z^i \notin A$ pour $i \in \{1, \ldots, n-1\}$) et des $y_\alpha^{(i)}, x_\alpha^{(i)} \in A \setminus \{0\}$ tels que $\infty > v(y_\alpha^{(i)}) \geq v(x_\alpha^{(i)}) =: \gamma_\alpha^{(i)}$, avec $(\gamma_\alpha^{(i)})_{\alpha<\alpha_i}$ strictement décroissante ; appelons $I_i$ le segment final de $vA \setminus \{\infty\}$ engendré par les $\gamma_\alpha^{(i)}$ ;

. $y \neq z^i x$ pour tous $x, y \in A$ avec $v(x) \notin I_i$.

Parce que $A$ est 0-inversible avec $A/v = B_\gamma/v$ pour tout $\gamma \in \Gamma$, la dernière condition ci-dessus équivaut à ce que, pour tout $x \in A$ avec $v(x) \notin I_i$, alors $v(x) + iv(z) \notin vA$. Si 0 est adhérent à un des $I_i$ (id est $0 = \inf I_i$) dans $\Gamma$, $z^i$ est connu mod $\infty - \varepsilon$ pour tout $\varepsilon \in \Gamma^{>0}$, cad exactement, et on réalise d'abord $z^i$ avant de réaliser $z$. Nous supposons donc désormais :

$(*)$ 0 n'est adhérent à aucun $I_i$.

Assertion : le type d'isomorphisme au dessus de $A$ de l'anneau quasi-valué $A[z]$ est uniquement déterminé.

Preuve. La valuation de $Q(z)$ doit être imposée, pour $Q = \Sigma a_i X^i \in A[X]$, de degré $< n$ si $z^n \in A$ comme ci-dessus. Si $v(a_i) \geq v(x_\alpha^{(j)})$ pour un $j \leq i$ et un $\alpha < \alpha_j$ alors, dans $B_{\infty - v(x_\alpha^{(j)})}$,

$$a_i z^i = (a_i (x_\alpha^{(j)})^{(-1)})(x_\alpha^{(j)} z^j) z^{i-j} = (a_i (x_\alpha^{(j)})^{(-1)}) y_\alpha^{(j)} z^{i-j}.$$

On applique cette réduction pour $i$ le terme de degré maximal de $Q$ auquel elle s'applique. On itère par degré décroissant. On se ramène ainsi à un polynôme $Q$ qui, pour un $\delta \in \bigcup_{i \leq d(Q)} I_i$, vérifie : pour tout $i \leq d(Q)$, ou bien $v(a_i) \geq \delta$ ou bien $v(a_i) < v(x_\alpha^{(j)})$ pour tous $j < i$ et $\alpha < \alpha_j$. Ainsi les valuations des monômes $a_i z^i$ avec $v(a_i) < \delta$ sont toutes différentes : si $v(a_i z^i) = v(a_j z^j)$ avec $j > i$, alors $v(a_i) = v(a_j) + (j - i)v(z)$ donc $v(a_j) \in I_{j-i}$, contradiction. Il ne reste une indétermination sur $v(Q(z))$ que lorsque tous les coefficients de $Q$ sont de valuation $\geq \delta$. On considère alors comme précédemment un coefficient $q$ de valuation minimale, et le polynôme $q^{-1}Q \mod \infty - v(q)$, auquel on applique le raisonnement précédent. L'indétermination ne peut se reproduire indéfiniment à cause de l'hypothèse $(*)$ et du caractère brutalement archimédien de $\Gamma$. $\square$

**Proposition 17.** *Soit un anneau quasi-valué $A$ de caractéristique résiduelle nulle et équipé d'un système $\mathcal{B} = (B_\gamma)_{\gamma \in \Gamma}$ d'inverses binaires. Soit $B_0 :=$ inj lim $B_\gamma$. Alors il existe une extension inversible $A'$ de $A$ et une extension $\mathcal{C}$ de $\mathcal{B}$, également indexée par $\Gamma$, vérifiant $A'/v = B_0 = C_\gamma/v$ pour tout $\gamma \in \Gamma$, et telles que $A'_{proj}$ se plonge dans $\mathcal{C}$ au-dessus de $A_{proj}$. Il existe de telles extensions minimales et elles sont toutes isomorphes au-dessus de $\mathcal{B}$.*

Démonstration. Soit $\Delta$ le sous-monoïde régulier maximal de $\Gamma$ et $w \leq v$ la valuation correspondante sur $A$ et les $B_\gamma$. Afin de nous ramener au cas archimédien, nous construisons canoniquement à partir de $\mathcal{B}$ un système d'inverses binaires de $(A, w)$. Pour $\bar{\gamma} = \gamma + \Delta \in \Gamma/\Delta$, définissons $B'_{\bar{\gamma}} := \operatorname{inj\,lim}_{\delta \in \Delta} B_{\gamma+\delta}$. Les $\pi_{\gamma,\delta}$ induisent des $\pi_{\bar{\gamma},\bar{\delta}}$.
Première extension de $\mathcal{B}$.
Ainsi $B'_{\bar{\gamma}}$ satisfait la condition de la proposition 12 (1) : si $v(x) \leq v(y) \in \Delta$ avec $x, y \in B_\gamma$, alors $yx^{-1} \in B_{\gamma - v(x)}$, or $\overline{\gamma - v(x)} = \bar{\gamma}$. On prend un sous-anneau $b_\gamma$ de $B_\gamma$ trivialement valué par $w$, et on montre comme dans la proposition 12 (2), mais en utilisant maintenant le système $\mathcal{B}$, que l'anneau quasi-valué $(B'_{\bar{\gamma}} Fr(b_\gamma), v)$ est uniquement déterminé. Précisons. Le passage au reste $B'_{\bar{\gamma}} \to B'_{\bar{\gamma}}/w : x \mapsto x/w$, induit un isomorphisme $b_\gamma \simeq B'_{\bar{\gamma}}/w$, dont on nomme l'inverse $\sigma$. On montre par induction sur $n$ qu'il y a une seule façon de valuer une somme $x := \Sigma^n a_i c_i$ où $a_i \in Fr(b_\gamma)$ et $c_i \in B'_{\bar{\gamma}}$ ; on peut supposer tous les $w(b_i)$ égaux non infini, soit $\delta$ cette valeur commune ; on a ainsi dans $B'_{\overline{\gamma-\delta}}$ ,

$$\pi_{\bar{\gamma},\overline{\gamma-\delta}}(\Sigma^n a_i c_i) = \pi_{\bar{\gamma},\overline{\gamma-\delta}}(c_1(a_1 + \Sigma_2^n a_i c_i c_1^{-1})) =$$

$$\pi_{\bar{\gamma},\overline{\gamma-\delta}}(c_1([a_1 + \Sigma_2^n a_i \sigma(c_i c_1^{-1}/w)] + \Sigma_2^n a_i(c_i c_1^{-1} - \sigma(c_i c_1^{-1}/w))));$$

si le terme entre crochets n'est pas nul, $v(x) = \delta$ et sinon nous sommes ramenées à valuer une somme de $n - 1$ termes. Nous sommes donc fondées à poser $B''_{\bar{\gamma}} := B'_{\bar{\gamma}} \otimes_{b_\gamma} Fr(b_\gamma)$. Parce que $Fr(b_\gamma)$ est unique au-dessus de $b_\gamma$, $B''_{\bar{\gamma}}$ l'est au-dessus de $B'_{\bar{\gamma}}$. En particulier les $B''_{\bar{\gamma}}$, $\bar{\gamma} \in \Gamma/\Delta$, forment un système projectif. Les multiplication et division mixtes s'étendent uniquement à ce système. Par le lemme 16 $(A, w)$ admet une unique clôture inversible $A'$ compatible avec le système $\mathcal{B}''$. La clôture cherchée est $\{x \in A'; v(x) \geq 0\}$. $\square$

# 4 Preuve de l'eq dans les cvac

**Théorème 6.** *Les cacv non trivialement et de caractéristique résiduelle nulle éliminent les quantificateurs dans le langage $\mathcal{L}_{proj}$.*

Ce théorème découle immédiatement de la proposition suivante qui ramène l'eq dans $\mathcal{L}_{proj}$ à l'eq bien connue dans $\mathcal{L}_{val}$.

**Proposition 18.** *Soit $M$ un cacv non trivialement et de caractéristique résiduelle nulle, et une $\mathcal{L}_{proj}$-sous-structure dénombrable $\mathcal{A} \subseteq M_{proj}$. Il existe alors un sous-corps $K$ de $M$ ayant les propriétés suivantes :*
*- $K$ est dénombrable,*
*- $\mathcal{A} \subseteq K_{proj}$,*
*- $K$ se plonge au-dessus de $\mathcal{A}$ dans tout cacv $\aleph_1$-saturé $M^*$ tel que $\mathcal{A} \subseteq_{proj} M^*_{proj}$,*
*- le type d'isomorphisme de $K$ au-dessus de $\mathcal{A}$ ne dépend que de $\mathcal{A}$.*

Le reste de cette section est consacrée à la preuve de cette proposition.

Le corps valué $M$ est considéré dans le langage $\mathcal{L}_{proj}$. Son anneau de valuation $O(M)$ s'identifie à $O_{<\infty}$, chaque $O_{<\gamma}(M)$ a une structure d'anneau, de zéro $\pi_\gamma(0)$ (l'image par $\pi_\gamma$ du 0 de $O$) et d'unité $\pi_\gamma(1)$, et porte la valuation $v_\gamma$. Une $\mathcal{L}_{proj}$-sous-structure de $\mathbb{M}$ est de la forme $\mathcal{K} = \langle K, \Delta, (A_\gamma)_{\gamma \in \bar{\Delta}^+} \rangle$, où $K$ est un sous-corps de $M$, $\Delta$ un sous-groupe de $vM$ contenant $vK$, chaque $A_\gamma$ un sous-anneau de $O_{<\gamma}(M)$, et le système des $A_\gamma$ est stable par les fonctions de $\mathcal{L}_{proj}$. Les $(A_\gamma)_{\gamma \in \bar{\Delta}^+}$ forment un système projectif d'anneaux. La présence des inverses binaires dans le langage fait que ce système contient, pour chaque $\gamma$, un système d'inverses binaires de $A_\gamma$ au sens de la section 2.4. Chaque $A_\gamma / v_\gamma$ est en particulier un corps.

Grâce à la section précédente, on peut supposer le corps $K$ et les anneaux $A_\gamma$ henseliens, et les $A_\gamma$ inversibles. L'unicité des clôtures (propositions 17 et 11), la proposition 10, (2), et la remarque suivant la définition 14, (3), imposent la cohérence avec le système projectif (projections, multiplications et divisions : à vérifier !).

(*) Pour $\delta \in \Delta$ vérifiant $\gamma/2$ (il existe dans $vM$) $< \delta < \delta + 2\varepsilon < \gamma$ et $x \in A_\gamma$ avec $v(x) = \delta$, la multiplication par $x$ définit un isomorphisme

$$\{y \in A_\gamma; \delta \leq v(y) < \delta + \varepsilon\} \to \{y \in A_{\gamma+\delta}; 2\delta \leq v(y) < 2\delta + \varepsilon\}$$

donc la translation $+\gamma$ définit un isomorphisme du segment $[\delta, \delta + \varepsilon[$ de $vA_\gamma$ vers le segment $[2\delta, 2\delta + \varepsilon[$ de $vA_{\gamma+\delta}$. Par inversibilité de $A_\gamma$, les segments $[\delta, \delta + \varepsilon[$ et $[0, \varepsilon[$ de $vA_\gamma$ sont isomorphes, de même dans $vA_{\gamma+\delta}$. En conséquence le segment $[0, \varepsilon[$ est le même dans $vA_\gamma$ et $vA_{\gamma+\delta}$, et $vA_\gamma = vA_{\gamma+\delta} \upharpoonright \gamma$. La multiplication par $x$ définit également un isomorphisme entre la fibre en $\delta$ de $A_\gamma$ et la fibre en $2\delta$ de $A_{\gamma+\delta}$. Cela va suffire pour conclure dans le cas archimédien.

## 4.1 Le cas archimédien

**Proposition 19.** *Soit $M$ un cacv non trivialement et de caractéristique résiduelle nulle et $\mathcal{A} = (A_\gamma)_{\gamma \in \bar{\Delta}^+} \subseteq M_{proj}$, où $\Delta$ est un sous-groupe archimédien de $vM$. Il existe alors des sous-corps $K$ de $M$ henseliens minimaux vérifiant $\mathcal{A} \subseteq K_{proj}$. Ils sont tous isomorphes au-dessus de $\mathcal{A}$ et leur type d'isomorphisme est indépendant de $M$. Ils sont dénombrables si les $A_\gamma$ le sont.*

Démonstration. On a vu que les $A_\gamma$ peuvent être supposés inversibles et henseliens.
Si un $vA_\gamma$ est dense, par application de $\pi_{\gamma,\gamma'}$ tout $vA_{\gamma'}$ avec $\gamma' < \gamma$ l'est aussi. Mais c'est également vrai pour $\gamma' > \gamma$ avec de plus $vA_\gamma = vA_{\gamma'} \upharpoonright \gamma$ : parce que $\Delta$ est archimédien, $\gamma' < \gamma + n\delta$ pour un entier $n$, et (*) permet de conclure par induction sur $n$. Si $A$ est la limite projective des $A_\gamma$, alors $vA \upharpoonright \gamma = vA_\gamma$ pour tout $\gamma$. Par l'argument analogue sur les fibres, tous les $A_\gamma$ et $A$ ont même corps résiduel, $k$. Un relèvement de $k$ dans $A$ se projette en un relèvement de $k$ dans chaque $A_\gamma$. On peut donc étendre les $A_\gamma$ de façon à ce que leur corps de restes commun soit algébriquement clos (par la proposition 12 ?). Le semi-groupe de valuation $\bigcup vA_\gamma$ de $A$ se relève en un sous-semi-groupe multiplicatif $G$ de $A$. Le corps des fractions de $A$ est le complété d'un corps $K$ obtenu par induction à partir de $k(G)^h$ en réalisant des suites PC non C de type transcendant, parfaitement déterminées parce que les $A_\gamma$ le sont, et en quantité dénombrable si les $A_\gamma$ sont dénombrables.

Ce corps $K$ est le corps cherché.
Si un $v_\gamma A_\gamma$, donc tout $v_\gamma A_\gamma$, est discret, le système $\mathcal{A}$ peut être dégénéré : on appellera ainsi un système engendré par multiplication mixte par un anneau $A_2 = k + kX$ où $k$ est un corps et $X$ un élément de valuation 1 ; on a dans ce cas $A_{n+1} = k + kX^n$ pour chaque entier $n$ ; ce système est clos par division binaire, ses limites projective et injective sont $k$. D'une manière générale, ou bien $A_3 = k + kX^2$, alors $A_4 = k + kX^3$ sinon par divisibilité et multiplication (internes) $A_4 = k + kX + kX^2 + kX^3$, contradiction, et par induction $A_{n+1} = k + kX^n$ et nous sommes face à l'exemple précédent, ou bien $A_3 = k + kX + kX^2$ et par induction $A_{n+1} = k + kX + \cdots + kX^n$. Dans les deux cas, le corps $K$ cherché est $k(X)^h$.
$\Delta$ est ici n'importe quel sous-semi-groupe de $\mathbb{R}^{\geq 0}$, avec alors $A_r = A_{[r]}$ où $[r]$ est la partie entière du réel $r$. □

## 4.2 Le cas général

Pour $\mathcal{A} = (A_\gamma)_{\gamma \in \Delta}$ et un sous-groupe convexe $C$ de $\Delta$, la restriction de $\mathcal{A}$ à $C$ est simplement le système $\mathcal{A} \restriction C = (A_\gamma)_{\gamma \in C}$. Si $C$ est un sous-groupe convexe principal et si $C'$ est son prédécesseur, on définit $(\mathcal{A} \restriction C)/C'$ de la façon suivante : c'est le système indéxé par $\Delta/C$ où pour $\bar{\gamma} = \gamma + C$ avec $\gamma \in \Delta$ et $w = v/C$, $A_{\bar{\gamma}} := A_\gamma \otimes_{A_\gamma/w} Fr(A_\gamma/w)$, au sens du lemme machin. Ce système est inversible, et il est henselien par ? ?. On appelle $(\mathcal{A} \restriction C)/C'$ la composante archimédienne $C/C'$ de $\mathcal{A}$.

1. On supprime les composantes archimédiennes dégénérées de $\mathcal{A}$ de la façon suivante : si $C$ est un sous-groupe convexe principal de $\Delta$ et $C'$ son prédécesseur, définissons $A_C := O(K_{min}(C/C'))$, où dans la parenthèse extérieure se trouve le corps minimal associé à la composante archimédienne $C/C'$ de $\mathcal{A}$, et si $C/C'$ est dégénéré et $\gamma \in C \setminus C'$ remplaçons $A_\gamma$ par $O_{<\gamma}(Fr(A_{1+C'}(X)^h)$. Le système reste projectif, et stable par multiplication et division mixtes parce que, en-dessous de $C'$ et au-dessus de $C$ rien n'est changé, et qu'entre les deux tout se passe dans un corps.

2. Désormais, pour tout sous-groupe convexe principal $C$ de $\Delta$, $C'$ son prédécesseur, et tout $\gamma \in C \setminus C'$, la projection $\pi_{C,\gamma} : A_C \to A_\gamma$ est surjective. Le système $\mathcal{A}$ se déduit donc du système $(A_C)_{C \in X}$ (où $X$ est l'ensemble des sous-groupes convexes principaux de $\Delta$) et la proposition 19 se déduit immédiatement de son analogue ci-dessous pour les anneaux de valuation.

**Proposition 20.** *Soit $M$ un cacv non trivialement et de caractéristique résiduelle nulle. Soit $W$ une chaîne de valuations sur $M$, ayant un élément maximal $v$ et contenant la valuation triviale. Soit $\mathbb{M}$ le système projectif $((O(M/w, v/w))_{w\in W}, (\pi_{w>w'})_{w,w'\in \bar{W}})$ et $\mathcal{A} := (A_w)_{w\in W} \subseteq \mathbb{M}$. Supposons $(M, v)$ $\aleph_1$-saturé, $W$ et les $A_w$ dénombrables. Alors il existe un sous-corps $K$ de $M$ ayant les propriétés suivantes :*

- *$K$ est dénombrable,*
- *$\mathcal{A} \subseteq \mathbb{K}$,*
- *$K$ se plonge au-dessus de $\mathcal{A}$ dans tout cacv $\aleph_1$-saturé $M^*$ tel que $\mathcal{A} \subseteq \mathbb{M}^*$,*
- *le type d'isomorphisme de $K$ au-dessus de $\mathcal{A}$ ne dépend que de $\mathcal{A}$.*

Attention : $A_w$ ne désigne en rien l'anneau de la valuation $w$, qui, selon nos notations, serait noté $O(M, w)$.
Dans cet énoncé, le corps $M$ est le corps de fractions de $O(M/v_0)$ pour $v_0$ la valuation triviale.
Exunt multiplication et divisions mixtes puisque les anneaux sont désormais intègres.

Nous commençons par étendre $A_{v_0}$ de façon à lui adjoindre tous les restes réalisés par un anneau $A_\gamma$.

**Lemme 21.** *Soit un corps $M$ équipé d'une chaîne $W$ de valuations, ayant un élément maximal $v$ et contenant la valuation triviale. Soit $\mathbb{M}$ le système projectif $(O(M/w)_{w\in W}, (\pi_{w,w'})_{w,w'\in W})$, $\mathcal{A} := (A_w)_{w\in W} \subseteq \mathbb{M}$ et $k := \bigcup A_w/v$. Supposons $(M, v)$ henselien et $\aleph_1$-saturé, $W$ et les $A_w$ dénombrables. Alors il existe un système projectif $\mathcal{B} = (B_w)_{w\in W} \supseteq \mathcal{A}$ ayant les propriétés suivantes :*

- *les $B_w$ sont dénombrables,*
- *$B_w/v = k$ pour tout $w \in W$,*
- *$\mathcal{B}$ se plonge au-dessus de $\mathcal{A}$ dans $\mathbb{M}^*$ pour tout corps valué henselien $\aleph_1$-saturé $M^*$ tel que $\mathcal{A} \subseteq \mathbb{M}^*$,*
- *le type d'isomorphisme de $\mathcal{B}$ au-dessus de $\mathcal{A}$ ne dépend que de $\mathcal{A}$.*

Démonstration. Si $r \in k$ est algébrique sur $A_v/v$ de polynôme minimal unitaire $p$ sur $A_v/v$, si $P \in A_v[X]$ est un relèvement unitaire de $p$, et si $M^*$ est henselien, il contient un unique zéro $x$ de $P$ de reste $r$, et $A_w/v$ un unique zéro de $P/w$ de reste $r$. Ainsi le type d'isomorphisme de $x/w$ sur $A_w$ est déterminé pour chaque $w \in W$ et donc aussi le type de $x$ sur $\mathcal{A}$. Cela permet de clore algébriquement $A_v/v$ dans $k$. Les anneaux $A_w[x/w]$ restent henseliens.
Soit maintenant à ajouter à $A_v$ un élément $x$ de reste $r \in k$ transcendant sur $A_v/v$. Si $r$ reste transcendant sur $k$, le type d'isomorphisme de $x$ sur $\mathcal{A}$ est

déterminé.
Soit sinon la suite $1 = d_1, d_2, \dots$ strictement croissante, indexée par un segment initial (qu'on ne présume ni propre ni fini) de $\mathbb{N}$, des degrés de $r$ sur les $A_w/v$. Il lui correspond la suite d'intervalles $W_i$ de $W$ tels que :

$W_i < W_{i+1}$
$\bigcup W_i$ est un segment initial de $W$
si $w \in W_i$, $[(A_w/v)[r], (A_w/r)] = d_i$
si $w \in W \setminus \bigcup W_i$, $r$ est transcendant sur $A_w/v$.

Soit $v_i := \inf W_{i+1}$ et $w_0 := \inf v_i$. Par le cas algébrique et par induction sur $i$, il apparaît une suite de polynômes irréductibles $p_i \in A_{v_i}[X]$, de degré $d_i$ et vérifiant $p_i | (p_{i+1}/v_i)$, tels que chaque $\mathcal{A}/v_i$ admette (à isomorphisme près) une unique extension minimale ajoutant un zéro à $p_i$ (en fait : un unique zéro $r_i$ de reste $r_i/v_{i-1} = r_{i-1}$, et $r_1/v = r$). Soit $\mathcal{B}_i$ cette extension et $\mathcal{B}'$ la réunion des $\mathcal{B}_i$. Si $M^*$ est $\aleph_1$-saturé il contient un élément $x$ de reste $x/v_i = r_i$ pour chaque $i$. Pour $w > w_0$, $x/w$ est transcendant sur $A_w$. Ainsi le type de $x$ au-dessus de $\mathcal{A}$ est déterminé. On ferme par clôture henselienne. On itère. □

Démonstration de la proposition 18.
Soit $W = \{w_i ; i \in \omega\}$. Réénumérons $W$ avec répétition de façon à ce que chaque valuation apparaisse cofinalement : $W = \{w_\alpha ; \alpha \in \alpha_0\}$ où $\alpha_0 := \omega \times \omega$ et $w_{i,j} := w_i$. Construisons par induction une suite croissante $(\mathcal{A}_\alpha)_{\alpha < \alpha_0}$ de systèmes projectifs : $\mathcal{A}_0 := \mathcal{A}$ ; si $\mathcal{A}_\alpha$ est connu, considérons le système $\mathcal{B}_\alpha := (B_{\alpha,w})_{w \in W, w > w_\alpha}$ où $B_{\alpha,w} := O(Fr A_{\alpha,w}, w_\alpha)$, son extension $\mathcal{B}'_\alpha$ donné par le lemme 32 (appliqué avec la valuation $v/w_\alpha$ en place de $v$), et enfin $\mathcal{A}_{\alpha+1}$ donné par

$A_{\alpha+1,w} := O(B_{\alpha,w})$ si $w > w_\alpha$
$A_{\alpha+1,w} := A_{\alpha,w}$ si $w < w_\alpha$.

On prend l'union aux étapes limites. Le corps des fractions de $A_{\alpha_0,v}$ est le corps cherché : $\forall w \in W, \forall x \in A_{\alpha,w}, \exists \alpha_1 < \alpha_0$ pour lequel $x \in A_{\alpha_1,w}$, mais alors il existe $\alpha_2$ tel que $\alpha_1 < \alpha_2 < \alpha_0$ et $w = w_{\alpha_2}$, et donc $x \in A_{\alpha_2+1,v}/w$. □

# Références

[1] Deirdre Haskell, Ehud Hrushovski and Dugald Macpherson, *Definable sets in algebraically closed valued fields : elimination of imaginaries*, Journal für die Reine und Angewandte Mathematik, 597 (2006) 175–236.

# Anneaux de Grothendieck et fonctions de paires sans cycles

Esther Elbaz

Paris, 15 mars 2017

# Table des matières

Introduit en théorie des modèles en 2000 par Jan Krajicek et Thomas Scanlon [1], le concept d'anneaux de Grothendieck d'une structure élémentaire est une généralisation de la notion, connue en géométrie algébrique, d'anneau de Grothendieck d'une variété.

En théorie des modèles, ces anneaux permettent une meilleure compréhension de la façon dont les ensembles définissables interagissent entre eux. Ils offrent aussi une analyse des propriétés combinatoires de la structure : il existe un parallèle entre les propriétés combinatoires d'une structure et les propriétés algébriques de son anneau de Grothendieck.

Le but de cet exposé est de calculer l'anneau de Grothendieck des structures $M$ dotées d'une fonction de paire sans cycles.

Une bijection $p$ de $M^2$ dans $M$ est dite sans cycles, ou encore localement libre, si pour tout terme $t(x_1, \dots, x_j)$ formée à partir de $p$, et qui n'est pas une simple variable, on a, pour tous $a_1, \dots, a_n$ de $M$, $t(a_1, \dots, a_n) \neq a_1$. Toutes ces "fonctions-paires" sont élémentairement équivalentes (dés que l'ensemble $M$ est non vide).

Ces fonctions-paires constituent l'exemple le plus simple d'une théorie stable qui n'est pas limite de théories superstables : c'est en effet un seul énoncé, l'existence d'une bijection de $M^2$ dans $M$, et non une infinité, qui empêche la superstabilité.

Plusieurs articles ont été consacrés à leur théorie. Dans [13] Belegradek a étudié leurs propriétés de stabilité. J. F. Pabion [14] a utilisé une fonction-paire pour $\omega_1$-saturer les modèles de l'Arithmétique. Dans [15], Bruno Poizat a fait remarqué combien il était utile pour cela de disposer d'une théorie non superstable sans propriété de recouvrement fini. Cette théorie de fonctions-paires a également été utilisée dans [16]. Dans [17] Elisabeth Bouscaren et Bruno Poizat ont, entre autre chose, étudié cette théorie pour trouver des contre-exemples aux propriétés des théories superstables.

Après avoir rappelé la construction de l'anneau de Grothendieck d'une structure, nous énoncerons un critère qui lorsqu'il est satisfait assure qu'une structure ait un anneau de Grothendieck non trivial et de caractéristique nulle. Nous montrerons ensuite comment les structure dotée d'une fonction de paire sans cycles entrent dans le cadre d'application du critère. Celui-ci nous permettra d'établir que leur anneau de Grothendieck est $\mathbb{Z}[X]/(X - X^2)$.

# 1 Anneaux de Grothendieck

## 1.1 Construction de l'anneau

Soit $M$ une structure. Son anneau de Grothendieck est construit à partir de ses sous-ensembles définissables en identifiant ceux qui sont en bijection définissable. Cet anneau dépend donc du langage utilisé. Ce qui suit détaille la construction de cet anneau.

### 1.1.1 Le premier quotient

Soit $M$ une structure et soit $Def(M)$ l'ensemble des sous-ensembles définissables de $M^n$ pour $n$ variant dans $\mathbb{N}$.

Soit $\sim$ la relation d'équivalence définie sur $Def(M)$ par :

$A \sim B$ si et seulement si $A$ est en bijection définissable avec $B$.

On note $\tilde{Def}(M)$ l'ensemble $Def(M)$ quotienté par $\sim$.

Si $A \in \tilde{Def}(M)$, on note $[A]$ la classe d'équivalence de $A$.

### 1.1.2 Les lois du futur anneaux

On peut alors munir $\tilde{D}ef(M)$ des opérations $+$ et $\times$ suivantes.

La loi $+$ correspond à l'union disjointe. Elle est définit par :
$[A] + [B] = [A' \cup B']$ où $[A] = [A'], [B] = [B']$ et $A' \cap B' = \emptyset$.

L'élément neutre de $+$ correspond à la classe d'équivalence de l'ensemble vide.

La loi $\times$ correspond au produit cartésien. Elle est définit par : $[A] \times [B] = [A \times B]$.

L'élément neutre de $\times$ correspond à la classe d'équivalence d'un singleton.

### 1.1.3 On quotiente une deuxième fois

$(\tilde{D}ef(M), +, \times)$ n'est pas un anneau. En effet, la loi $+$ n'est pas inversible. Elle n'est même pas simplifiable. En effet, $[A] + [B] = [A] + [C]$, n'implique pas $[A] = [C]$.

Par exemple, soit $\mathbb{N}$ l'ensemble des entiers naturels, considéré comme une structure naturelle du langage $\{+, =\}$.

On peut trouver une bijection définissable entre $\mathbb{N} \cup \{(0,0)\}$ et $\mathbb{N}$ (par exemple la fonction $f$ telle que $f((0,0)) = 0$ et $f(n) = n + 1$ pour tout $n \in \mathbb{N}$.)

Ainsi $[\mathbb{N}] + [\{(0,0)\}] = [\mathbb{N}]$. Mais il n'y a évidemment pas de bijection entre $\{(0,0)\}$ et le vide et $[\{(0,0)\}]$ n'est donc pas égal à 0.

Afin d'obtenir un monoïde simplifiable, on quotiente $(\tilde{D}ef(M), +)$ par la relation d'équivalence $\simeq$ définie par $a \simeq b$ si et seulement si il existe $c \in \tilde{D}ef(M)$ tel que $a + c = b + c$.

Comme, tout monoïde simplifiable, $(\tilde{D}ef(M), +)$ quotienté par $\simeq$, se plonge dans un anneau unique à isomorphisme près et minimal pour la relation d'inclusion.

C'est cet anneau qu'on appelle anneau de Grothendieck de $M$. Il est noté $K_0(M)$.

## 1.2 Le principe des tiroirs surjectif

Reprenons l'exemple de $\mathbb{N}$ considéré comme une structure du langage $\{+, =\}$. Nous avons vu que $[\mathbb{N}] + [\{(0,0)\}] = [\mathbb{N}]$ donc, dans $K_0(M)$, $[\{(0,0)\}] = 0$. Autrement dit $1 = 0$ (puisque la classe d'équivalence d'un singleton correspond à l'élément neutre de la multiplication). Et $K_0(M)$ est trivial.

Ceci est un exemple particulier du lemme plus général :

**Lemme 1.1.** *Soit $M$ une structure. Il y a équivalence entre :*

- *Il existe $A$ un ensemble définissable de $\cup_{n \in \mathbb{N}} M^n$ en bijection définissable avec lui même privé d'un point.*
- *L'anneau de Grothendieck de $M$ est trivial.*

*Démonstration.* L'implication directe se démontre exactement comme nous l'avons fait dans le cas de $\mathbb{N}$ : puisque $A$ est en bijection définissable avec $A - \{a\}$, $[A - \{a\}] = [A]$, $[\{a\}] = 0$ c'est-à-dire $1 = 0$. $K_0(M)$ est donc trivial.

La réciproque se démontre tout aussi facilement. Si $K_0(M)$ est trivial, alors $1=0$. Ce qui signifie en remontant à $Def(M)$ qu'il existe $A$ un ensemble définissable tel que $[A] + 1 = [A]$. Autrement dit en prenant $a$ un élément de $\cup_n M^n$ n'appartenant pas à $A$, $A \cup \{a\}$ est en bijection avec $A$, c'est-à-dire lui-même privé d'un point. ☐

La propriété de ne pas admettre d'ensembles définissables en bijection définissable avec eux-même privés d'un point, est appelé principe des tiroirs surjectif, ou en anglais onto-pigeonhole principle, abrégé onto-PHP.

C'est une propriété du premier ordre.

Ainsi ce principe, qui correspond à une propriété combinatoire des structures, détermine si l'anneau de Grothendieck de ces structures est ou non trivial.

### 1.3 Exemples d'anneaux de Grothendieck connus

Toute structure finie a un anneau de Grothendieck isomorphe à $\mathbb{Z}$.

Krajicek et Scanlon ont démontré dans [1] que l'anneau de Grothendieck du corps réel clos est également isomorphe à $\mathbb{Z}$ : l'image d'un ensemble définissable de $\mathbb{R}$ dans $K_0(\mathbb{R})$ est déterminée par sa caractéristique d'Euler.

Grâce au principe des tiroirs surjectif, on a pu démontrer que l'anneau de Grothendieck de plusieurs corps valués dans le langage $L_{anneaux}$ est trivial : Lou van den Dries, dans [4], a ainsi établi la trivialité de l'anneau de Grothendieck de $\mathbb{Q}_p$ comme $L_{anneaux}$-structure. Dans [5], Raf Cluckers a obtenu le même résultat pour l'anneau de Grothendieck de certaines séries de Laurent dans le langage $L_{anneaux}$.

Deirdre Haskell et Raf Cluckers l'ont eux démontré dans [7] pour $\mathbb{Q}_p$ et, plus généralement, pour les corps $\mathbb{Z}$-valués qui vérifient certaines conditions dont l'hensélienníté du corps et l'existence d'une composante angulaire.

Denef et Loeser ([6]) ont montré que l'anneau de Grothendieck de $\mathbb{C}$ considéré comme $L_{anneau}$-structure admet l'anneau $\mathbb{Z}[X;Y]$ comme quotient. Krajicek et Scanlon ont démontré par la suite que $\mathbb{Z}[X_i : i \in \mathbb{R}]$ s'injecte dans $K_0(\mathbb{C})$.

Des résultats concernant les modules sont également connus.

Perera a montré dans ([10]) que pour tout $R$-module infini $M$ où $R$ est un anneau de division infini, $K_0(M)$ est isomorphe à $\mathbb{Z}[X]$.

Prest a conjecturé [11, Ch. 8, Conjecture A] que tout module a un anneau de Grothendieck non trivial. Kuber a démontré dans [12] cette conjecture pour les modules $M$ tels que $M$ est élémentairement équivalent à $M \oplus M$. Il a, pour ces modules, pu calculer explicitement cet anneau : c'est le quotient d'un anneau de polynôme à une infinité de variables.

# 2 Un critère pour assurer la non trivialité de l'anneau de Grothendieck d'une structure

## 2.1 Rappels sur les topologies noethériennes

**Définition 2.1.** Un espace topologique est dit noethérien si toute suite décroissante de fermés est stationnaire.

**Définition 2.2.** Un sous-ensemble d'un espace topologique est dit irréductible s'il est non-vide et ne peut pas être écrit comme l'union de deux sous-ensembles fermés propres.

**Lemme 2.3.** *Dans un espace noethérien, tout fermé non vide* $X$ *peut s'écrire comme une union finie* $X = \bigcup_i X_i$ *de fermés irréductibles* $X_i$*. Cette décomposition est de plus unique si l'on impose* $X_i \not\subseteq X_j$ *pour tout* $i \neq j$*.*

**Définition 2.4.** La décomposition du lemme précédent est appelée décomposition irréductible de $X$, les $X_i$ en sont ses composantes irréductibles.

**Définition 2.5.** Soit $X$ un fermé irréductible non vide.

On définit inductivement la dimension de $X$ :

- $\dim(X) = 0$ si $C$ est un singleton
- $\dim(X) = \sup\{\dim(F) + 1 | F \subsetneq X, \text{ fermé irréductible non vide}\}$

Si $X$ est un fermé non nécessairement irréductible, alors on définit $\dim(X) := \dim(X_1)$ où $X_1$ est une composante irréductible de $X$ de dimension maximale.

Pour énoncer notre critère nous devons introduire la définition suivante :

**Définition 2.6.** Un ensemble $A$ est dit **décomposable** s'il peut s'écrire sous la forme $A = \bigcup_{j=1}^{h}(B_j \setminus C_j)$ où pour tout $j$, $B_j$ et $C_j$ sont des fermés. On dit que $A = \bigcup_{j=1}^{h}(B_j \setminus C_j)$ est une **décomposition** de $A$.

**Lemme 2.7.** *Soit* $A$ *un ensemble décomposable. Il est possible de raffiner la décomposition précédente pour qu'elle vérifie les conditions suivantes :*

1. *pour tout* $i$, $B_i$ *est irréductible*
2. *pour tout* $i$, $C_i \subsetneq B_i$
3. *pour tous* $i \neq j$, $C_i \subsetneq B_j$ *implique* $C_i \subsetneq C_j$
4. *pour tous* $i \neq j$ *tels que* $B_i \cap C_j \neq \emptyset$, *si* $B_i \cap C_j = \cup_k B'_k$ *est sa décomposition en irréductibles, alors il existe* $i_1, \ldots, i_k$ *tels que* $B_{i_1} = B'_1, \ldots, B_{i_k} = B'_k$

**Définition 2.8.** Une décomposition qui vérifie toutes ces conditions est dite **réduite**.

## 2.2 Enoncé du critère

**Proposition 2.9.** *Soit* $M$ *une structure. Supposons que pour toute injection définissable* $g$ *d'un ensemble définissable* $A \subseteq M^n$ *dans lui-même privé d'un point, il existe* $T$ *et* $T'$ *deux topologies noethériennes sur* $M^n$*telles que :*

- *$A$ est décomposable pour la topologie $T$.*

— *Il existe, pour la topologie T, une décomposition réduite de A,* $A = \bigcup_{j=1}^{h}(B_j \setminus C_j)$ *telle que pour tout j, g se prolonge en une injection définissable* $\tilde{g}$ *sur* $B_j$*.*

— *g est telle que pour tout fermé irréductible* $B \subset B_j$, $\tilde{g}(B))$ *est un fermé irréductible de* $T'$*.*

*Alors M vérifie l'onto-PHP et son anneau de Grothendieck est de caractéristique nulle.*

Sur tout module $M$ tel que $M$ est élémentairement équivalent à $M \oplus M$, il est possible de définir une topologie $T = T'$ qui vérifie les conditions de cette proposition. Ce critère permet donc de retrouver le résultat obtenu par Amit Kuber : l'anneau de Grothendieck de ces modules est non trivial et de caractéristique nulle.

Le critère est en fait un cas particulier du lemme suivant :

**Lemme 2.10.** *Soit M une structure qui satisfait les conditions énoncées dans 2.9. Soit g une injection définissable définie sur* $A \subseteq M$ *et dont l'image contient A privé d'un nombre fini de points. Alors* $g(A) = A$*.*

En effet, la non trivialité de l'anneau de Grothendieck est équivalente à ce lemme d'après le principe surjectif des tiroirs. Par ailleurs, que l'anneau de Grothendieck soit de caractéristique nulle est équivalent à ce que toute injection définissable $g$ d'un ensemble $A$ dans lui-même et dont l'image contient $A$ privé d'un ensemble fini de points, est en fait une bijection.

Nous ne donnerons pas la démonstration de ce lemme mais nous essayons d'en expliquer le principe sur un exemple.

Soit $M$ une structure qui satisfait aux conditions énoncées dans 2.9. Soit $g$ une injection définissable définie sur $A \subseteq M^n$ et dont l'image contient $A$ privé d'un nombre fini de points.

Supposons que $A$ soit l'union de deux fermés irréductibles $X$ et $Y$ et que $(X \setminus B) \cup (Y \setminus B') \cup B \cup B'$ soit une décomposition réduite de $A$ qui satisfait les conditions du critère.

Pour simplifier encore supposons que $B$ et $B'$ sont irréductibles et non réduits à un point.

D'après les hypothèses du critère, $g(B)$ et $g(B')$ sont des fermés irréductibles, et $g(X \setminus B) = X' \setminus C$ (respectivement $g(Y \setminus B') = Y' \setminus C'$), où $X', Y', C, C'$ sont des fermés irréductibles et $C \subsetneq X'$, $C' \subsetneq Y'$.

Puisque $g(A) \subseteq A$, $X' \subseteq A$ (respectivement $Y' \subseteq A$).

$A$ est donc égale à $g(B) \cup g(B') \cup X' \cup Y'$. Par unicité de la décomposition en fermés irréductibles l'un des ensembles $g(B), g(B'), X', Y'$ est égale à $X$ (respectivement à $Y$).

Puisque $g$ est injective, $\{X', Y'\} = \{X, Y\}$.

$g(B)$ et $g(B')$ sont alors inclus dans $C$ ou dans $C'$. Ils ne peuvent pas appartenir au même ensemble puisque sinon le complémentaire de $g(A)$ contiendrait un fermé irréductible non réduit à un point. Supposons que $g(B) \subseteq C$ et $g(B) \subseteq C'$.

$C \setminus g(B)$ (respectivement $C' \setminus g(B')$) est soit vide soit infini. En effet, puisque $C$ (respectivement $C'$) est un fermé irréductible, il ne peut exister un ensemble fini $F$ tel que $C = g(B) \cup F$ (respectivement $C' = g(B') \cup F$).

Puisqu'on a supposé que le complémentaire de $g(A)$ dans $A$ est fini, cela implique que $g(A) = A$.

Dans le cas général, ce sont les mêmes principes qui s'appliquent. Soit $g$ une injection définissable définie sur un ensemble $A$ et telle que $g(A)$ soit égale à $A$ privé d'un nombre fini de points. Soit $A = \bigcup_{j=1}^{h}(B_j \setminus (\cup_i X_{j_i})$ où les $X_{i,j}$ sont des fermés irréductibles, est une décomposition réduite de $A$ qui vérifie les conditions du critères. Les hypothèses du critères impliquent que $g(A)$ s'écrive comme combinaison booléenne des fermés irréductibles $g(B_j)$ et $g(X_{j_i})$.

Les propriétés des fermés irréductibles et les propriétés de $g$ impliquent qu'il n'est pas possible d'écrire une combinaison booléenne de ces fermés qui soit égale à $A$ privé d'un nombre fini de points.

# 3 Théorie des fonctions de paires sans cycles

## 3.1 Rappels

Soit $M$ un ensemble et soit $\Theta$ une bijection de $M$ dans $M^2$. On dit alors que $\Theta$ est une "fonction de paire", qui permet de définir deux autres fonctions :

- la fonction $g$ appelée "membre gauche" qui à tout $x \in M$ associe $y$ tel qu'il existe $z \in M$ avec $\Theta(x) = (y, z)$
- la fonction $d$ appelée "membre droit" qui à tout $x \in M$ associe $z$ tel qu'il existe $y \in M$ avec $\Theta(x) = (y, z)$.

Autrement dit, $\Theta(x) = (g(x), d(x))$.
Si $x \in M$, nous noterons parfois $x_g := g(x)$ et $x_d := d(x)$.

On dit que $\Theta$ est sans cycles, si pour tout terme $t(x_1, \ldots, x_n)$ formé à partir de $\Theta$ et qui n'est pas une variable, on a pour tout $a_1, \ldots, a_n \in M$, $t(a_1, \ldots, a_n) \neq a_1$.

Dans le langage constitué d'un symbole de fonction pour le membre gauche et pour le membre droit (et du signe égal), la théorie des fonctions de paires sans cycles admet l'élimination des quantificateurs et est complète ([17]). C'est dans ce langage, $L = \{g, d, =\}$ que nous nous plaçons, désormais.

## 3.2 Arbres binaires

Soit $T$ un arbre binaire infini.

**Définition 3.1.** On définit la profondeur d'un nœud récursivement :

- par convention, sa racine est de profondeur $0$;
- si un nœud est de profondeur $p$, alors ses fils sont de profondeur $p+1$.

Il y a ainsi, un nœuds à la profondeur 0, deux à la profondeur $1, \ldots,$ et $2^p$ à la profondeur $p$. On compte donc $2^{p+1} - 1$ nœuds à profondeur inférieure ou égale à $p$.

**Définition 3.2.** Numérotation des nœuds
On énumère les nœuds de $T$ en leur associant une suite finie d'éléments de $\{g, d\}$ de la façon suivante :

- à la racine on associe la suite vide;
- si un nœud est numéroté par une suite $\{u_1, \ldots, u_n\}$, alors son fils gauche est numéroté par $\{u_1, \ldots, u_n, g\}$ et son fils droit par $\{u_1, \ldots, u_n, d\}$.

Les nœuds de profondeur $p$ sont ainsi numérotés par des suites à $p$ éléments.

**Notation 1.** *Nous noterons $T_g$ (respectivement $T_d$) le sous-arbre gauche (respectivement droit) de $T$, c'est-à-dire l'arbre de racine $g(x)$ (respectivement l'arbre de racine $d(x)$).*

Dans la suite, tous les arbres considérés seront des arbres binaires infinis. Pour alléger la rédaction, nous parlerons simplement d'arbres binaires sans préciser à chaque fois qu'ils sont infinis.

## 3.3 Définitions

Soit $M$ un modèle de la théorie des fonctions de paires sans cycles. Par formule, on entendra formule de $L(M)$.

**Définition 3.3. Arbre associé à un élément**
Soit $x$ un élément de $M$.
On associe à $x$ un arbre binaire étiqueté par des éléments de $M$ de la façon suivante :
sa racine est étiquetée par $x$ ;
si un nœud est étiqueté par $y$, alors son fils gauche est étiqueté par $y_g$ et son fils droit par $y_d$.

On note $T_x$ cet arbre.

On énumère les nœuds de $T_x$ par des suites finies d'éléments de $\{g, d\}$ comme expliqué à la section précédente.
Si $i$ est une suite finie de $\{g, d\}$, on note $x_i$ le terme qui étiquette le $i$-ième nœud de $T_x$.

Cette notation est bien cohérente avec celle proposée plus haut : $g(x)$ qui est noté $x_g$ correspond bien au fils gauche de l'arbre de racine $x$ ; et $d(x)$ noté $x_d$ correspond bien à son fils droit.

Les nœuds de $T_x$ décrivent l'ensemble des termes faisant intervenir comme unique variable $x$.

**Définition 3.4. Profondeur d'une formule, d'un ensemble, d'une fonction définissable**
Soit $\phi$ une formule. On appelle **profondeur en** $x$ de $\phi$ le plus grand entier $n$ tel que $\phi$ fait intervenir un terme de $x$ situé à profondeur $n$ dans l'arbre $T_x$.
Soit $A$ un ensemble définissable. On appelle **profondeur en** $x$ de $A$ le plus petit entier $n$ tel qu'il existe une formule définissant $A$ qui soit de profondeur $n$ en $x$.
Soit $h$ une application définissable. On appelle **profondeur en** $x$ de $h$ le plus petit entier $n$ tel qu'il existe une formule définissant le graphe de $h$ qui soit de profondeur $n$ en $x$.

**Définition 3.5. Formule primitive**
On considère les formules $\phi$ à deux variables libres $x$ et $y$, qui s'écrivent $\wedge_j \phi_j$ où les $\phi_j$ ont une des formes suivantes :

1. $t(x) = c$
2. $t(x) = t'(x)$
3. $t(x) = t'(y)$
4. $t(x) \neq t'(y)$
5. $t(x) \neq c$
6. $t(x) \neq t'(y)$

où $t(x), t'(x)$ (respectivement $t'(y)$) sont des termes en $x$ (respectivement en $y$), et $c$ est une constante.
Une telle formule est dite **primitive**. Si les formules $\phi_j$ sont toutes de la forme 1, 2 ou 3 alors on dit que $\phi$ est **primitive positive**.

**Définition 3.6. Arbre associé à une formule primitive**
Soit $\phi = \wedge_j \phi_j$ une formule primitive à une seule variable libre. On lui associe un arbre binaire dont le $i$-ème nœud est étiqueté par la liste des formules $\phi_j$ qui font intervenir le terme correspondant à $x_i$.
Par abus, nous pourrons également dire que cet arbre est associé à $\phi(M)$ ou qu'il définit $\phi(M)$.

Réciproquement à un tel arbre peut être associée une formule :

**Définition 3.7. Formule primitive associée à un arbre**
Soit $T$ un arbre binaire étiqueté par des listes de formules de la façon suivante :
a. Le $i$-ème nœud est étiqueté par une liste finie de formules $\phi_j$ de la forme :
$t(x) = c$ ou
$t(x) = t'(x)$ ou
$t(x) \neq c$ ou
$t(x) \neq t'(x)$
où $t(x)$ est le terme correspondant à $x_i$, $t'(x)$ est un terme ayant pour unique variable $x$, et $c$ est une constante.
b. Seul un nombre fini de nœuds est étiqueté par une liste non vide.

On peut associer à $T$ la formule $\wedge_j \phi_j$ obtenue en prenant la conjonction de toutes les formules qui apparaissent dans les étiquettes de $T$. On note cette formule $\phi(T)$.

*Exemple* 3.8. 1. Soit $\phi$ la formule $d(x) \neq g(x)$. On associe à $\phi$ l'arbre dont les deux fils de la racine portent chacun la condition $d(x) \neq g(x)$. Les autres nœuds de l'arbre sont étiquetés par la liste vide. $\phi$ est réalisée par tous les éléments dont l'arbre est tel que les deux fils de la racine portent des éléments distincts.

2. Soit $c \in M$. Considérons $\phi$ la formule $(d(x) = c) \wedge (g(x) = d(x))$. L'arbre associé à $\phi$ est l'arbre dont le premier fils gauche porte l'étiquette $\{g(x) = d(x)\}$, et le premier fils droit, l'étiquette $\{(d(x) = c), (g(x) = d(x))\}$. Les autres nœuds sont étiquetés par la liste vide.
(La formule $\phi$ définit le même ensemble que $(d(x) = c) \wedge (g(x) = d(x)) \wedge (g(x) = c)$. Mais elles ont des arbres différents.)

**Remarque 1.** *Soit $T$ un arbre étiqueté comme en 4.7. Supposons que la racine de $T$ soit étiquetée par la liste vide, alors $\phi(T)(x)$ est équivalente à $\phi(T_g)(g(x)) \wedge \phi(T_d)(d(x))$.*
*Puisque $\Theta$ est bijective, pour tout $y \in \phi(T_g)(M)$ et tout $z \in \phi(T_d(M))$, l'arbre $T'$ tel que $T'_g = T_y$, $T'_d = T_z$, et dont la racine est étiquetée par la liste vide, est l'arbre d'un élément de $T(M)$.*

subsectionForme des injections définissables

**Lemme 3.9.** ***Forme des injections définissables*** *Soit $A$ un sous-ensemble définissable de $M$ et $h$ une application définissable injective de $A$ dans $A$. Le graphe de $h$ peut être défini par une formule de la forme $\vee_j(\phi_j(x) \wedge \phi'_j(x,y))$ où :*
*a. chaque $\phi_j(x)$ est une formule primitive*
*b. les ensembles $\phi_j(M)$ sont tous disjoints,*
*c. chaque $\phi'_j(x,y)$ est une disjonction de conjonctions de formules primitives positives.*

*Démonstration.* Par élimination des quantificateurs, on peut supposer qu'on a partitionné $A$ en une union disjointe d'ensembles $A_j$ définis par des formules $\phi_j$ telles que les points a et c soient satisfaits.
Soit $\phi'_j(x,y)$ des formules définissant le graphe de $h$ sur $A_j$. Puisque $h$ est une fonction, il est évident que les formules $\phi'(x,y)$ doivent vérifier le point b. - Vérifions qu'on peut également satisfaire le point d.
Toujours par élimination des quantificateurs, et quitte à raffiner, on peut supposer que les $\phi'_j(x,y)$ sont des formules primitives. Il s'agit de montrer que ces formules sont en fait équivalentes à des formules primitives positives.

Autrement dit si une formule définit un singleton et s'écrit comme une conjonction de formules de la forme

1. $t(y) = t'(x)$
2. $t(y) = t'(y)$
3. $t(y) \neq t'(x)$
4. $t(y) \neq t'(y)$
5. $t(y) = c$ où $c \in M$
6. $t(y) \neq c$

où $t(x), t'(x)$ (respectivement $t(y), t'(y)$) sont des termes faisant intervenir uniquement la variable $x$ (respectivement $y$), il s'agit de montrer, qu'on peut en fait se restreindre aux formules de la forme 1 ou 5.

Soit $\phi'_j(x, y)$ une conjonction d'égalités ou de négations d'égalités de la forme $t(y) = t'(x)$, $t(y) = t'(y)$, ou $t(y) = c$ où $t(y), t'(y), t'(x)$ sont des termes et $c$ une constante.
Á $x$ fixé, $\phi'_j(x, y)$ est équivalente à une conjonction de formules de la forme $t(y) \neq c$ ou $t(y) = c$ où $t(y)$ est un terme et $c$ une constante.

Le résultat découlera donc du lemme suivant :

**Lemme 3.10.** *Soit*

$$\psi(y) = (\wedge_{j=1}^{l} t_j(y) = c_j) \wedge (\wedge_{j=1}^{n} t'_j(y) \neq c'_j) \wedge (\wedge_{j=1}^{m} t''_j(y) = t'''_j(y)) \wedge (\wedge_{j=1}^{k} t''''_j(y) \neq t'''''_j(y))$$

*où $t_j(y), t'_j(y), t''_j(y), t'''_j(y), t''''_j(y), t'''''_j(y)$ sont des termes en une seule variable et $c_j, c'_j$ des constantes.*
*Si $\psi(y)$ définit un ensemble fini, alors elle est en faite équivalente à $\wedge_{j=1}^{l} t_j(y) = c_j$.*

La suite de cette sous-section est essentiellement consacrée à la démonstration de ce lemme.

Commençons par remarquer qu'un ensemble fini défini par une conjonction d'égalités de la forme $t(y) = c$ où $t(y)$ est un terme faisant intervenir uniquement la variable $y$ et $c$ est une constante, est soit un singleton, soit vide :

**Lemme 3.11.** *Soit $\psi(y) = (\wedge_{j=1}^{n} t_j(y) = c_j)$ où les $c_j$ sont des constantes et les $t_j(y)$ des termes non constants faisant intervenir uniquement la variable $y$. Supposons que $A := \psi(M)$ est non vide. Pour tout $i$, soit $A_i := \{c | \exists y \in A, y_i = c\}$. Pour tout $i \in \{g, d\}^n$ où $n \in \mathbb{N}$, $A_i$ est soit infini soit réduit à un singleton.*
*De plus le fait que $A_i$ soit réduit à un singleton ne dépend pas du choix des $c_j$ mais uniquement de l'ensemble $\{i | \exists j, t_j(y)$ soit égale à $y_i\}$.*

*Démonstration.* Cela est une conséquence immédiate du fait que $\Theta$ est une bijection et peut se démontrer par récurrence sur $q$ la profondeur de $\psi$.
Pour $q = 1$, puisque $\Theta$ est une bijection, la propriété est évidente.

Supposons la propriété vérifiée pour $q \in \mathbb{N}^*$. Montrons qu'elle est vraie pour $q + 1$.

Soit $T$ l'arbre associé à $\psi$.

Si la racine de $T$ n'est pas étiquetée par la liste vide, alors $T(M)$ est un singleton et le résultat est évident. Nous pouvons donc supposer que la racine de $T$ est étiquetée par la liste vide, et que $\phi(T)(x)$ est équivalente à $\phi(T_g)(g(x)) \wedge \phi(T_d)(d(x))$.

Par hypothèse de récurrence, $T_g(M)$ et $T_d(M)$ vérifient la propriété du lemme. Soit $i_1, \ldots, i_n$ les éléments de la suite finie $i$. Soit $\hat{i}$ la suite $i_2, \ldots, i_n$.
Puisque $\Theta$ est une bijection, $\phi(M)_i = \phi(T_{i_1})(M)_{\hat{i}}$. L'arbre $T$ tout entier la vérifie donc aussi.
Cela achève la démonstration de la propriété au rang $n$. □

**Remarque 2.** *1. En particulier si $A$ est un ensemble fini défini par une formule de la forme $\psi(y) = (\wedge_{j=1}^{n} t_j(y) = c_j)$ où les $c_j$ sont des constantes et les $t_j(y)$ des termes faisant intervenir uniquement la variable $y$. Alors $A$ est un singleton (car $A = A_\emptyset$).*
*2. Supposons que $g$ est une fonction définie sur $B - C$ où $B$ et $C$ sont des ensembles définissables, et que le graphe de $g$ est défini par $(\wedge_j t_j(y) = t'_j(x)) \wedge (\wedge_i t'_i(y) = c_i)$, où les $c_i \in M$, et $t_j(x), t_j(y), t'_i(y)$ sont des termes faisant intervenir une seule variable libre. Alors $g$ peut se prolonger à $B$ tout entier. En effet, pour tout $x_0 \in B - C$ fixé, la formule $(\wedge_j t_j(y) = t'_j(x_0)) \wedge (\wedge_i t'_i(y) = c_i)$ définit un singleton. Puisque $x_0$ est fixé, cette formule est en fait de la forme des formules considérées dans le lemme 2. Le fait qu'une telle formule définit un singleton est en fait uniquement déterminé par l'ensemble l'ensemble $\{k | \exists j, t_j(y) = y_k$ ou $t'_i(y) = y_k\}$. La formule $(\wedge_j t_j(y) = t'_j(x)) \wedge (\wedge_i t'_i(y) = c_i)$ définit donc un singleton pour tout $x \in B$ : on peut donc prolonger $g$ à $B$ tout entier, en posant $g(x) := y$ tel que $\wedge_j (t_j(y) = t'_j(x)) \wedge (\wedge_i t'_i(y) = c_i)$.*

*Démonstration.* Preuve du lemme 3.10
Montrons le résultat pas récurrence sur $(m, k, n) \in \mathbb{N}^3$ muni de l'ordre lexicographique.

**Etape 1 :** Commençons par montrer par récurrence sur $n$ que pour tout $n \in \mathbb{N}^*$, le résultat est vrai pour $(m, k, n) = (0, 0, n)$.

Pour $n = 1, k = 0, m = 0$, soit $\psi(y) = (\wedge_{j=1}^{l} t_j(y) = c_j) \wedge (t'(y) \neq c')$. Soit $i$ tel que $t'(y) = y_i$. Soit $A$ l'ensemble défini par $\wedge_{j=1}^{l} t_j(y) = c_j$.

Par le lemme précédent $A$ et $A_i$ sont chacun soit infini, soit réduit à un singleton.
Si $A$ est un singleton, alors $(\wedge_{j=1}^{l} t_j(y) = c_j) \wedge (t'(y) \neq c')$ qui est non vide par hypothèse, est équivalente à $(\wedge_{j=1}^{l} t_j(y) = c_j)$.

Supposons $A$ infini.
On remarque que $c' \notin A_i$ : en effet, sinon puisque $(\wedge_{j=1}^{l} t_j(y) = c_j) \wedge (t'(y) \neq c')$ définit un ensemble non vide, $A_i$ contient au moins deux éléments et donc par le lemme précédent est infini. Mais cela est impossible puisqu'alors $(\wedge_{j=1}^{l} t_j(y) = c_j) \wedge (t'(y) \neq c')$ définit un ensemble infini.
Puisque $c' \notin A_i$, il est clair que $(\wedge_{j=1}^{l} t_j(y) = c_j) \wedge (t'(y) \neq c')$ est équivalente à $\wedge_{j=1}^{l} t_j(y) = c_j$ : la propriété est bien démontrée pour $n = 1$.

Soit $n \in \mathbb{N}^*$. Supposons la propriété démontrée pour $(0, 0, m)$ quelque soit $m < n$. Montrons la pour $(0, 0, n)$.

Soit $\psi(y) = (\wedge_{j=1}^{l} t_j(y) = c_j) \wedge (\wedge_{j=1}^{n} t'_j(y) \neq c'_j)$ une formule définissant un ensemble fini.
Soit $i$ tel que $t'_n(y)$ correspond à $y_i$. Soient $i_1, \ldots, i_k$ les éléments de $[|1; n-1|]$ tels que $t'_{i_j}(y)$ correspond à $y_i$.
Soit $A$ l'ensemble que définit $\wedge_{j=1}^{l} t_j(y) = c_j$.

Supposons que $(\wedge_{j=1}^{l} t_j(y) = c_j) \wedge (\wedge_{j=1, j \notin \{n, i_1, \ldots, i_k\}}^{n} t'_j(y) \neq c'_j))$ définit un ensemble fini.
Par hypothèse de récurrence $(\wedge_{j=1}^{l} t_j(y) = c_j) \wedge (\wedge_{j=1, j \notin \{n, i_1, \ldots, i_k\}}^{n} t'_j(y) \neq c'_j))$ est équivalente à $\wedge_{j=1}^{l} t_j(y) = c_j$.

Soit $A$ l'ensemble que définit cette formule. On procède comme dans le cas $n = 1$. D'après le lemme précédent, $A$ et $A_i$ sont soient infinis, soit réduits à un singleton.
Si $A$ est un singleton, alors $(\wedge_{j=1}^{l} t_j(y) = c_j) \wedge (t'_n(y) = c'')$ qui est non vide par hypothèse, est équivalente à $(\wedge_{j=1}^{l} t_j(y) = c_j)$ et la propriété est démontrée.
Supposons $A$ infini.
On remarque que pour tout $j \in \{n, i_1, \ldots, i_k\}$, $c_j \notin A_i$. En effet, sinon puisque $(\wedge_{j=1}^{l} t_j(y) = c_j) \wedge (\wedge_{j=1}^{n} t'_j(y) \neq c'_j)$ définit un ensemble non vide, $A_i$ contient au moins $k+1$ éléments et donc par le lemme précédent est infini. Mais cela est impossible puisqu'alors $(\wedge_{j=1}^{l} t_j(y) = c_j) \wedge (\wedge_{j=1}^{n} t'_j(y) \neq c'_j)$ définit un ensemble infini.
Ainsi, puisque pour tout $j \in \{n, i_1, \ldots, i_k\}$, $c_j \notin A_i$, il est clair que $(\wedge_{j=1}^{l} t_j(y) = c_j) \wedge (\wedge_{j=1, j \notin \{i_1, \ldots, i_k\}}^{n} t'_j(y) \neq c'_j)$ est équivalente à $\wedge_{j=1}^{l} t_j(y) = c_j$ : la propriété est bien démontrée.

Supposons que $(\wedge_{j=1}^{l} t_j(y) = c_j) \wedge (\wedge_{j=1, j \notin \{i_1, \ldots, i_k\}}^{n} t'_j(y) \neq c'_j)$ définit un ensemble infini.
Soit $A$ l'ensemble que définit $(\wedge_{j=1}^{l} t_j(y) = c_j)$ et $B$ celui défini par $(\wedge_{j=1}^{l} t_j(y) = c_j) \wedge (\wedge_{j=1, j \notin \{i_1, \ldots, i_k\}}^{n} t'_j(y) \neq c'_j)$.
Puisque $\Theta$ est une bijection, et que pour tout $j \notin \{n, i_1, \ldots, i_k\}$, $t'_j(y) \neq y_i$, on a $B_i = A_i$.
D'après le lemme précédent, $A_i$ et donc $B_i$ est soit réduit à un singleton soit infini.
Supposons que $B_i$ soit réduit à un singleton. Alors par le même argument que précédemment, $c_j \notin B_i$ pour tout $j \in \{n, i_1, \ldots, i_k\}$ et donc $(\wedge_{j=1}^{l} t_j(y) = c_j) \wedge (\wedge_{j=1, j \notin \{n, i_1, \ldots, i_k\}}^{n} t'_j(y) \neq c'_j) \wedge (\wedge_{j \in \{n, i_1, \ldots, i_k\}} t'_j(y) \neq c'_j)$ est équivalente à $(\wedge_{j=1}^{l} t_j(y) = c_j) \wedge (\wedge_{j=1, j \notin \{n, i_1, \ldots, i_k\}}^{n} t'_j(y) \neq c'_j)$. Par hypothèse de récurrence cette dernière formule est équivalente à $(\wedge_{j=1}^{l} t_j(y) = c_j)$.

Supposons que $B_i$ soit infini. Alors puisque $\{n, i_1, \ldots, i_k\}$ est fini, $\psi(M)_i$ est aussi infini. Mais cela contredit que $\psi$ définit un ensemble fini. Cette contradiction achève la récurrence : pour tout $n \in \mathbb{N}^*$, le résultat est vrai pour $(m, k, n) = (0, 0, n)$.

**Etape 2 :** Montrons à présent par récurrence sur $k \in \mathbb{N}^*$, que pour tout $k \in \mathbb{N}^*$, pour tout $n \in \mathbb{N}^*$, le résultat du lemme est vrai pour $(m, k, n) = (0, k, n)$.

Soit $n \in \mathbb{N}^*$.
Démontrons que le résultat est vrai pour $(0, 1, n)$.

Soit
$$\psi(y) = (\wedge_{j=1}^{l} t_j(y) = c_j) \wedge (\wedge_{j=1}^{n} t'_j(y) \neq c'_j) \wedge (t''_j(y) = t'''_j(y))$$
où $t_j(y), t'_j(y), t''_j(y), t'''_j(y)$ sont des termes en une seule variable et $c_j, c'_j$ des constantes.

Si $(\wedge_{j=1}^{l} t_j(y) = c_j) \wedge (\wedge_{j=1}^{n} t'_j(y) \neq c'_j)$ défini un ensemble fini, alors par ce qui précède elle est équivalente à $(\wedge_{j=1}^{l} t_j(y) = c_j)$ et définit en fait un singleton. Il est alors est évident que $\psi(y)$ est équivalente à $(\wedge_{j=1}^{l} t_j(y) = c_j)$.

Supposons que $(\wedge_{j=1}^{l} t_j(y) = c_j) \wedge (\wedge_{j=1}^{n} t'_j(y) \neq c'_j)$ est un ensemble infini. Si $t''(y)$ ou $t'''(y)$ correspondent à l'un des termes $t_j(y)$ ou $t'_j(y)$, alors la formule est en fait équivalente à une formule du type précédemment considéré et le résultat s'en suit.

Supposons que $t''(y)$ ou $t'''(y)$ ne correspondent à aucun des termes $t_j(y)$ ou $t'_j(y)$. Soit $y_t$ le terme correspondant à $t''(y)$ et $y_s$ celui correspondant à $t'''(y)$.

Si $A_s$ ou $A_t$ est un singleton $\psi$ est en fait équivalente à une formule du type précédemment considéré (c'est-à-dire une formule ne faisant apparaître aucune relation entre deux termes non constants en $y$).

Supposons que $A_s$ et $A_t$ sont infinis. On peut supposer qu'il n'y a aucun de lien de filiations entre $y_s$ et $y_t$ (grâce à la propriété que $\Theta$ est sans cycles). On peut également supposer que $y_t$ et $y_s$ ne sont fils d'aucun terme $y_i$ sur lequel porte une condition du type $y_i = c_i$ (puisqu'alors cette condition détermine également la valeur de ses descendants). Deux cas de figures se présentent donc : soit $y_s$ ou $y_t$ est père d'un des termes $y_i$ sur lequel porte une condition du type $y_i = c_i$ ; soit $y_t$ et $y_s$ sont tous les deux indépendants de tous ces termes.

Dans le premier cas, puisque $y_t = y_s$, l'arbre de racine $y_s$ coïncide avec celui de racine $y_t$ et les conditions portant sur leurs descendants doivent donc être les mêmes et alors $A_s = A_t$. Il est alors clair que $\psi$ ne peut pas définir un ensemble fini. Cette contradiction achève la récurrence.

Dans le deuxième cas, puisque $\Theta$ est une bijection $A_s = A_t = M$ et il est clair que $\psi$ ne peut pas un ensemble fini. Et cela achève la récurrence de la même façon.

Supposons le résultat démontré pour $k \in \mathbb{N}^*$. Démontrons-le pour $k+1$. Soit

$$\psi(y) = (\wedge_{j=1}^{l} t_j(y) = c_j) \wedge (\wedge_{j=1}^{n} t'_j(y) \neq c'_j) \wedge (\wedge_{j=1}^{k+1} t''_j(y) = t'''_j(y))$$

où $t_j(y), t'_j(y), t''_j(y), t'''_j(y)$ sont des termes en une seule variable et $c_j, c'_j$ des constantes.

On démontre le résultat de la même façon que pour $k = 1$. En effet, puisque $\Theta$ n'admet aucun cycle, on peut supposer qu'il n'existe aucun lien de filiation entre $t''_{k+1}(y)$ (respectivement $t'''_{k+1}(y)$), et les autres termes $t''_i(y), t'''_i(y)$.

Le résultat est donc démontré pour $m = 0$ et $k, n$ quelconques.

**Etape 3 :** Il reste à démontrer que le résultat est vrai pour $m$ quelconque. Cela se démontre par récurrence de façon tout à fait similaire à ce qui a été fait à l'étape 2.

$\square$

$\square$

**Remarque 3.** *1. Soient $x_i$ et $x_j$ deux termes en $x$. Si $x_i$ correspond à un descendant de $x_j$ (dans l'arbre associé à $x$), alors la formule $(x_i = t_i) \wedge (x_j = t_j)$ où $t_j, t_i$ sont des termes quelconques est soit insatisfaisable, soit équivalente à $(x_j = t_j)$.*

*2. Plus généralement, soit $\phi(x)$ une conjonction de formules de la forme $x_j = t_j$ où $x_j$ est un terme en une seule variable libre $x$, et $t_j$ est un terme quelconque.*
*Si $\phi$ est satisfaisable, alors elle est équivalente à la conjonction $\psi$ des formules $x_j = t_j$ qui sont telles que :*

- *$x_j = t_j$ apparaît dans la conjonction définissant $\phi$*
- *il n'existe pas de formule $x_i = t_i$ qui apparaisse dans la conjonction définissant $\phi$ et qui soit telle que $x_j$ correspond à un descendant de $x_i$.*

*La formule $\psi$ sera appelée la **forme normale** de $\phi$.*

**Définition 3.12.** Soit $g$ une fonction dont le graphe est défini par $(\wedge_j (t_j(y) = t_j(x))) \wedge (\wedge_j t'_j(y) = c_j)$.
On dit que la forme normale de cette formule est une **forme normale** de $g$ .

**Lemme 3.13.** *Soit $g$ une fonction injective définie par une forme normale sur un ensemble $A - B$. $g$ se prolonge sur $A$ en une fonction injective définie par la même formule normale.*

*Démonstration.* C'est une conséquence immédiate de la remarque 2.

$\square$

## 3.4 Ensembles simples et application du critère

Nous allons à présent donner des définitions qui nous permettront de définir une topologie qui satisfasse les conditions du critère 2.9.
Pour cela nous allons considérer les ensembles que nous appellerons "simples" et qui, à profondeur fixée, seront un analogue des ensembles irréductibles : ces ensembles simples seront tels qu'ils ne peuvent s'écrire comme union non triviale d'autres ensembles simples et ils permettront de "reconstituer" tout ensemble définissable de profondeur $p$ par des combinaisons booléennes.

Pour cela nous allons considérer les ensembles de profondeur $p$ et définir les ensembles simples en fixant toutes les relations possibles qu'il peut exister entre les $x_i$ où $i \in \{g,d\}^p$ et éventuellement en fixant certains de ces éléments $x_i$ comme étant égale à une constante. Fixer toutes les relations possibles qu'il peut exister entre les $x_i$, revient à se donner une partition $\cup_{j=1}^{n} I_j$ de $\{g,d\}^p$ telle que pour tous $i,j \in \{g,d\}^p$, $x_i = x_j$ si et seulement si il existe $k$ avec $i,j \in I_k$.
Choisir un élément défini par une telle formule (en supposant que celle-ci soit satisfaisable) revient à choisir $n$ éléments de $M$.
Étant donné que nous considérons les ensembles de profondeur $p$, une fois fixées toutes les relations possibles entre les éléments $x_i$, les seules conditions que l'ont peut ajouter sont celles qui fixent certains éléments comme étant égaux à une constante donnée.
En faisant cela, on diminue le degré de liberté que nous avions lors du choix d'un élément satisfaisant une formule. Intuitivement on diminue la dimension. Deux ensembles simples inclus l'un dans l'autre seront nécessairement dans ce cas de figure.
Deux ensembles simples ne peuvent ainsi être inclus l'un dans l'autre que si le degré de liberté dans le choix d'éléments diminue.

Ce qui suit formalise tout cela.

**Définition 3.14.** Soit $\cup_{j=1}^{n} I_j$ une partition de $\{g,d\}^p$ et soit $C$ un ensemble de couples $(i,c)$ où $i \in \{g,d\}^p$ et $c \in M$.
On associe à $\cup_{j=1}^{n} I_j$ et $C$, l'ensemble $A$ tel que :

— Pour tout $1 \leq j \leq n$, pour tout $x \in A$, et tout $j_1, j_2 \in I_j$, $x_{j_1} = x_{j_2}$

— Pour tout $(i,c) \in C$, pour tout $x \in A$, $x_i = c$.

Supposons de plus que pour tout $i \in \{g,d\}^p$ tel qu'il n'existe pas $c \in M$ avec $(i,c) \in C$, on peut trouver $x$ et $x'$ dans $A$ tels que $x_i \neq x'_i$. (L'ensemble $C$ est "maximal".)
L'ensemble $A$ est appelé ensemble simple de profondeur $p$ (la profondeur n'est pas précisée quand il n'y a pas d'ambigüité).
On appelle $\cup_{j=1}^{n} I_j$ la décomposition de $\{g,d\}^p$ associée à $A$ et $C$ son ensemble de constantes.

**Remarque 4.** *1. Si $B$ est un ensemble simple et $C$ un ensemble définissable strictement inclus dans $B$ (et de même profondeur), alors $C$ est également un ensemble simple dont la décomposition est la même que celle associée à $B$, et dont l'ensemble de constantes est strictement inclus dans celui de $B$.*

*2. Il est évident que si $A, B, C$ sont des ensembles simples, alors $A = B \cup C$ implique que $B = \emptyset$ ou $C = \emptyset$.*

**Proposition 3.15.** *Les ensembles simples de profondeur p forment la base de fermés d'une topologie noethérienne sur $M$ dont ils sont les fermés irréductibles.*

*Démonstration.* Soit $F$ l'ensemble dont les éléments sont les unions finis d'ensembles simples.
Il est immédiat de vérifier que les axiomes définissant des fermés sont satisfaits pas les éléments de

$F$.
$F$ est clairement stable par union finie.
L'ensemble $M$ est la réunion de tous les ensembles simples dont l'ensemble des constantes est vide (il y a $2^p$ tels ensembles qui correspondent à toutes les partitions possibles de $[|1; 2^p|]$). Il est donc un élément de $F$.
Il est évident que $F$ est stable par intersection.
Il est également évident qu'un ensemble simple ne peut s'écrire comme union non triviale de deux autres éléments de $F$.
La Noethériennité est immédiate à vérifier. Deux ensembles simples sont inclus l'un dans l'autre si et seulement si ils ont la même décomposition et que l'ensemble de constantes du premier est inclus dans celui du deuxième. Comme un ensemble de constantes a au plus $2^p$ éléments, il ne peut pas exister de chaînes infinie d'ensembles simples. ☐

**Remarque 5.** *1. Toute décomposition de $\{g, d\}^p$ ne permet pas de définir un ensemble simple. Par exemple, en profondeur 2, la décomposition associée à l'arbre (c'est-à-dire telle que $x_g = x_{\{g,d\}} = x_{\{d,d\}}$ et $x_d = x_{\{d,g\}} = x_{gg}$ avec de plus $x_g \neq x_d$.) n'est pas compatible avec l'absence de cycles.*

**Remarque 6.** *Soit $B$ un ensemble simple de profondeur $p$.*
*Chaque élément $x$ de $B$ correspond à un $2^p$ uplets : la liste des termes de $x$ situés à profondeur $p$ sur l'arbre de $x$.*

*La dimension de $B$ en tant que fermé d'une topologie noethérienne correspond au plus grand entier $d$, tel qu'il existe $x \in B$ dont le $2^p$-uplet de ses termes de profondeur $p$ contient $d$ éléments distincts.*

Informellement, la dimension correspond au nombre de paramètres qu'on peut choisir pour les éléments de $B$.

Il est évident que ces ensembles simples permettent par combinaison booléenne d'obtenir tous les ensembles définissable de profondeur inférieure ou égale à $p$

**Proposition 3.16.** *Soit $A$ un ensemble définissable de profondeur inférieure à $p$. Il est possible d'écrire $A$ comme une combinaison booléenne d'ensembles simples de profondeur $p$. De plus, si $A = \cup_j(B_j - (\cup_{j_i} B_{j_i}))$ et que cette écriture est sans répétition (i.e. il n'existe pas de $j \neq j'$ tels que $B_j = B_{j'}$ et il n'existe pas de $i, j, j'$ tels que $B_{j_i} = B_j$ ou $B_{j_i} = B_{j'}$), alors cette écriture est unique.*

On peut alors définir une notion de dimension pour tout ensemble définissable de profondeur inférieure à $p$ :

**Définition 3.17.** Soit $A$ un ensemble définissable de profondeur $p$ dont la décomposition en ensembles simples est $A = \cup_j(B_j - (\cup_{j_i} B_{j_i}))$
On définit la dimension de $A$ par $\dim(A) := max_j(\dim(B_j))$.

**Définition 3.18.** Soit $B$ un ensemble simple. On appelle **rang dans** $B$ d'une famille de variables $x_{i_1}, \ldots, x_{i_r}$ le plus grand nombre de variables qui peuvent être choisies distinctes.
Soit $g$ une fonction définissable définie sur $B$. Soit $x_{i_1}, \ldots, x_{i_r}$ les variables apparaissant dans $g$. On appelle **rang de** $g$ le rang dans $B$ de $x_{i_1}, \ldots, x_{i_r}$.

**Lemme 3.19.** *Une fonction $g$ définie sur un ensemble $A$ injective a un rang égale à la dimension de $A$.*

*Démonstration.* Pour que $g$ soit injective, il faut que pour tout $1 \leq i \leq n$, il existe $j_i \in I_i$ tel que $x_{j_i}$ intervient dans la définition de $g$.
Sinon, $g(y)$ est déterminé par la conjonction d'une formule $\psi(y)$ qui ne fait intervenir que les $x_i$ avec $i \notin I_i$ et d'une formule $\psi(y)$ qui ne fait intervenir que $x_{j_i}$. Il existe une infinité d'éléments $x$ de $A$ qui ont la même image par $g$.

□

Cela nous permet d'élargir le domaine de définition d'une fonction injective définie sur un ensemble $B - C$ où $B$ est un ensemble simple.

**Lemme 3.20.** *Soit $A$ un ensemble définissable dont l'adhérence est un ensemble simple . Soit $g$ une injection définissable dont le graphe est défini sur $A$ par la formule normale $\phi = (\wedge_j(t_j(y) = t'_j(x))) \wedge (\wedge_i t'_i(y) = c_j)$. Alors $g$ peut être prolongée à une injection définissable sur $\overline{A}$.*

*Démonstration.* $A$ est de la forme $B - \cup_j B_j$ où $B$ et les $B_j$ sont des ensemble simples. Nous avons déjà remarqué (3.13 ) que $g$ se prolonge à $B$ tout entier. Il reste à montrer l'injectivité de son prolongement. Puisque chaque $B_j$ est un ensemble simple inclus dans $B$, si une famille de variables $x_{i_1}, \ldots, x_{i_r}$ a un rang dans $B$ égal à la dimension de $B$, alors il est clair que cette même famille a un rang dans $B_j$ égal à la dimension de $B_j$ : la formule normale définissant $g$ se prolonge donc en une injection sur $B_j$. □

**Lemme 3.21.** *Soit $A$ un ensemble simple de profondeur $p$. Soit $g$ une injection dont le graphe est défini par une formule normale. Alors, il existe un entier $q$ tel que $g(A)$ est un ensemble simple de profondeur $q$.*

*Démonstration.* Soit $(\wedge_j t_j(y) = t_j(x)) \wedge (\wedge_j t'_j(y) = c_j)$ une formule normale définissant $g$ sur $A$, et soit $q$ la profondeur maximale des termes $t_j(y)$ et $t'_j(y)$. Il est évident qu'il est possible d'écrire $g(A)$ comme un ensemble simple de profondeur $q$. □

**Lemme 3.22.** *Les modèles de la théorie des fonctions de paires satisfont le critère 2 : leur anneau de Grothendieck est non trivial et de caractéristique nulle.*

*Démonstration.* Soit $g$ une injection définissable de domaine de définition $A$. Prenons une partition de $A$ en ensembles définissables $A_j$ telle que pour tout $j$, $g$ est définissable par une formule normale sur $A_j$. Soit $q$ le plus grand entier tel que $g$ est définie par une formule de profondeur $q$ sur $A_j$ Il résulte de ce qui précède que

- pour tout $j$, $g$ se prolonge en une injection définissable sur $\overline{A_j}$ l'adhérence de $A_j$.
- pour tout fermé irréductible $B \subset A_j$, $g(B)$ peut être vu comme un ensemble simple de profondeur $q$.

Le critère s'applique donc. □

## 3.5 Calcul de l'anneau de Grothendieck

**Corollaire 3.23.** *La théorie des fonctions de paires sans cycles a comme anneau de Grothendieck $\mathbb{Z}[X]/(X - X^2)$.*

*Démonstration.* Soit $X$ la classe de $M$ dans l'anneau de Grothendieck.
Il est évident que $X = X^2$. De plus, puisque, d'après la proposition ??, l'anneau de Grothendieck est de caractéristique non nulle, $\mathbb{Z}[X]/(X^2 - X)$ s'injecte dans l'anneau de Grothendieck de $M$.
Puisque $M$ est isomorphe à $M^2$, tout ensemble définissable de $M^n$ avec $n$ entier, est isomorphe à un ensemble définissable de $M$.
Soit $A$ un ensemble définissable de $M$ définit par $\phi$ une conjonction d'égalités sur les termes. Montrons que si $A$ n'est pas fini, alors il est isomorphe à $M$.
Soit $T$ l'arbre associé à $\phi$ et $p$ sa profondeur. Soit $i$ le nombre de nœud de $T$ qui ne correspondent pas à l'un des termes apparaissant dans $\phi$. $A$ est isomorphe à $M^i$ : une façon d'avoir un élément de $A$ est de placer sur chacun des nœud ne correspondant pas à l'un des termes apparaissant dans $\phi$, un élément arbitraire de $M$.
$A$ est donc isomorphe à $M$.

Tout ensemble définissable est combinaison booléenne d'ensemble défini par une conjonction d'égalités sur les termes.
Il s'en suit que si $A$ est un ensemble définissable de $M$, alors sa classe dans l'anneau de Grothendieck de $M$ est un élément de $\mathbb{Z}[X]/(X - X^2)$.

☐

## 3.6 Exemple de modèles de la théorie des fonctions de paires sans cycles

Il existe de nombreux exemples de fonctions de paires sur $\mathbb{N}$. Par exemple, lafonction de paire de Cantor : $(m,n) \mapsto \frac{1}{2}(m+n)(m+n+1)+m$. Mais cette fonction n'est pas sans cycle : $0 \mapsto 0$.
Les fonctions de paire auxquelles on pense le plus naturellement sont souvent définies de façon algébrique et vérifient une certaine propriété de décroissance : si $a \in N$, alors son image $(b,c)$ est telle que $b < a$ ou $c < a$... Elles ont donc un cycle en 0.
Nous allons ici donner l'exemple d'une fonction de paire sans cycles.

Considérons un tableau dont les lignes et les colonnes sont indexées par $\mathbb{N}$.
On inscrit dans chaque case du tableau un entier en procédant ainsi : on considère le tableau comme une union de carrés de côté $[|0,n|]^2$, on commence par la case $(0,0)$ puis si on suppose $n \geq 0$ et que les cases des $n$ premiers carrés ont déjà été énumérées, on considère le carré de côté $n+1$ dont on commence par énumérer les cases de la dernière ligne $n+1$ avant de remonter le long de la dernière colonne.

Voici ce que cela donne pour les trois premiers carrés :

| | | |
|---|---|---|
| 0 | 3 | 8 |
| 1 | 2 | 7 |
| 4 | 5 | 6 |

Cela définit une bijection de $\mathbb{N}$ dans $\mathbb{N}^2$. Cette bijection correspond en fait à la fonction définie ainsi :
pour $n = k^2 + t$ où $t \leq k$, on pose $\Theta(n) = (t,k)$
pour $n = k^2 + t$ où $k \leq t < 2k+1$, on pose $\Theta(n) = (k, k-t)$.

$\Theta$ n'est pas sans cycles : en effet, 0 est envoyé sur $(0,0)$.
Nous allons un peu modifié $\Theta$ afin qu'elle soit sans cycles. Le problème avec $\Theta$ est en fait qu'elle

vérifie une propriété de "'décroissance"' : pour tout $n \in \mathbb{N}$ son image $(a, b)$ est telle que $a \leq n$ et $b \leq n$.
Pour remédier à cela, nous allons faire du point 0, un point qui fasse "'diverger vers l'infini"'.
On procède ainsi :
pour définir les images des entiers, nous allons à nouveau faire appel au tableau dont les lignes et les colonnes sont indexées par $\mathbb{N}$.
Sur toutes les cases de coordonnées de la forme $(2^k, 2^k)$ où $k > 1$, on inscrit le chiffre $2^{k-1}$. Sur $(2, 2)$, on inscrit 0.
Puis on inscrit les chiffres restant (qui ne sont pas des puissances de 2) l'un après l'autre (il est entendu qu'on les énumère dans l'ordre croissant) en parcourant le tableau de la même façon qu'on l'avait parcouru lorsqu'on a définit $\Theta$.

Voici ce que cela donne pour les trois premiers carrés :

| | | |
|---|---|---|
| 1 | 6 | 11 |
| 3 | 5 | 10 |
| 7 | 9 | 0 |

Il est évident que $p'$ est sans cycles.
En effet, $p'$ est strictement "'décroissante"' sur les entiers strictement positifs qui ne sont pas une puissance de 2 (dans le sens où $p'(n) = (a, b)$ avec $a < n$ et $b < n$).
Et sur $\{2^k | k \in \mathbb{N}\} \cup \{0\}$, $p'$ est strictement croissante.

## 4 Bibiographie

1. J. Krajicek and T. Scanlon, Combinatorics with definable sets : Euler characteristics and Grothendieck rings, this Bulletin, vol. 6 (2000), pp. 311-330.
2. L. van den Dries, Tame topology and o-minimal structures, London Math. Soc. Lecture Note Series, Vol. 248, (1998), Cambridge University Press.
3. J. Krajicek, Uniform familes of polynomila equations over a finite field ans structures admitting an Euler characteristic of definable sets, Proc. London Mathematical Society.
4. Lou van den Dries, Tame topology and o-minimal structures, Lecture note series, vol. 248, Cambridge University Press, 1998.
5. R. Cluckers. Grothendieck rings of laurent series fields. Journal of Algebra, to appear. available at arXiv :math.LO/0210350.
6. J. Denef and F. Loeser, Definable sets, motives and p-adic integrals, Journal of the American Mathematical Society.
7. Raf Cluckers and Deirdre Haskell, Grothendieck Rings of $\mathbb{Z}$-Valued Fields,The Bulletin of Symbolic Logic, Vol. 7, No. 2 (Jun., 2001), pp. 262-269.
8. A. Kuber, Grothendieck rings of theories of modules, Preprint, arXiv :1302.4229, 2013.
9. J. Denef and F. Loeser, Germs of arcs on singular algebraic varieties and motivic integration, Invent. Math., 123 (1999), 201-232.
10. S. Perera, Grothendieck Rings of Theories of Modules, Doctoral Thesis, University of Manchester, 2011.

11. M. Prest, Model Theory and Modules, London Math. Soc., Lecture Notes Ser., Vol. 130, Cambridge University Press, 1988.
12. M. Prest, Purity, Spectra and Localisation, Encyclopedia of Mathematics and its Applications, Vol. 121, Cambridge University Press, 2009.
13. V. Belegradek, Theory of models of locally free algebras, Trudy Instituta Matematiki Sibirskogo Otdelenlja Akademii Nauk SSSR
14. J.-F. Pabion, Saturated models of Peano arithmetic, The Journal of Symbolic Logic, Vol. 53, No. 2 (Jun., 1988), vol. 47, pp. 625-637.
15. B. Poizat, Deux remarquesa propos de la propriétée recouvrement fini, The Journal of Symbolic Logic, Vol. 53, No. 2 (Jun., 1988),vol. 49, pp. 803-807
16. B. Poizat, Cours de théorie des modèles, Nur al-Mantiq wal-Ma'rifah, Villeurbanne
17. E. Bouscaren, B. Poizat, Des belles paires aux beaux uples, The Journal of Symbolic Logic, Vol. 53, No. 2 (Jun., 1988), pp. 434-442

# ON CONTINUOUS FUNCTIONS DEFINABLE IN EXPANSIONS OF THE ORDERED REAL ADDITIVE GROUP

PHILIPP HIERONYMI AND ERIK WALSBERG

ABSTRACT. For every expansion of the ordered real additive group one of the following holds: every continuous definable function $[0,1] \to \mathbb{R}$ is $C^2$ on an open dense subset of $[0,1]$, or every definable $C^2$ function $[0,1] \to \mathbb{R}$ is affine, or every continuous function $[0,1] \to \mathbb{R}$ is definable. The first case holds for any $NTP_2$ expansion of $(\mathbb{R},<,+)$, more generally for any expansion that does not interpret the monadic second order theory of one successor.

## 1. INTRODUCTION

Throughout $\mathcal{R} = (\mathbb{R},<,+,\dots)$ is an expansion of the ordered additive group of real numbers, and "definable" means "$\mathcal{R}$-definable, possibly with parameters". This paper continues our earlier work in [6, 10] on definable sets in such expansions. This endeavor is part of the general program of investigating geometric and topological properties of definable sets in expansions of the real line, as outlined by Miller [12].

An $\omega$-**orderable set** is a definable set that admits a definable ordering with order type $\omega$. We say that such a set is **dense** if it is dense in some open subinterval of $\mathbb{R}$. We divide expansions of $(\mathbb{R},<,+)$ into three distinct categories:

(A) $\mathcal{R}$ does not define a dense $\omega$-orderable set,
(B) $\mathcal{R}$ defines a dense $\omega$-orderable set, but does not define some compact set,
(C) $\mathcal{R}$ defines every compact subset of every $\mathbb{R}^k$.

It is easy to see that a type C expansion defines a dense $\omega$-orderable set (see [6, Theorem 3.9(i)]), so the three cases are indeed exclusive. In this paper we study consequences of this trichotomy on definable continuous functions, and show that such functions behave very differently in each case.

We first give some motivation for this trichotomy. A type C structure is not model-theoretically tame in any sense as it interprets second order arithmetic. Every projective subset of $[0,1]^k$ is definable in such a structure and even simple questions about basic properties of projective sets, such as Lebesgue measureability, are independent of ZFC. Although not stated in precisely this way, the argument in [7] shows that every expansion of the field of real numbers is type A or type C. Therefore any model-theoretically tame expansion of the real field is type A. Indeed a stronger result is true: any expansion of $(\mathbb{R},<,+)$ that defines multiplication by

---

*Date*: March 17, 2017.

This is an early preprint version. Later versions might contain significant changes. Comments are welcome! The first author was partially supported by NSF grant DMS-1300402.

$\lambda$ for uncountably many $\lambda \in \mathbb{R}$ is either type A or type C (see [6, Theorem C]).

We regard type A as the ultimate generalization of o-minimality in the setting of expansions of $(\mathbb{R}, <, +)$. All topological assumptions on unary definable sets that have been proposed as generalizations of o-minimality, imply type A. For example, if every definable subset of $\mathbb{R}$ has interior or is nowhere dense, then $\mathcal{R}$ is type A as a dense $\omega$-orderable set is dense and codense in an interval. Model-theoretic tameness conditions also imply type A. By [10, Theorem A] an expansion which admits a dense $\omega$-orderable set interprets the monadic second order theory of one successor and for that reason violates all known Shelah-style combinatorial tameness properties such as NIP or $\mathrm{NTP}_2$ (see e.g. Simon [15] for definitions). So all $\mathrm{NTP}_2$ expansions of $(\mathbb{R}, <, +)$ are type A.

Type B structures lie in between the two extremes of type A and type C. While such an expansion fails any combinatorial tameness, there are type B expansions with decidable theories. Natural examples of such type B expansions are $(\mathbb{R}, <, +, x \mapsto \sqrt{2}x, \mathbb{Z})$ (see [8]) and $(\mathbb{R}, <, +, C)$, where $C$ is the middle-thirds Cantor set (see Balderrama and Hieronymi [1]). Interestingly, type B expansions have received little attention within tame geometry and model theory, but have appeared in theoretical computer science (see for example Boigelot, Rassart, and Wolper [3]) and fractal geometry (see Charlier, Leroy, and Rigo [5]). One reason might be that all known examples of type B expansions are bi-interpretable with the monadic second order theory of one successor. This theory was shown to be decidable by Büchi [4] using automata-theoretic rather then model-theoretic methods.

Let us return to continuous functions. First observe that in type C expansions there is no restriction at all on the behavior of continuous definable functions as every continuous function from $[0,1]$ to $\mathbb{R}$ is definable in such a structure. This is in stark contrast to the situation for type A expansions. Indeed, weak analogues of known results for o-minimal structures such as the monotonicity theorem hold for all type A structures (see [10, Theorem E]). In this paper we prove the following regularity result.

**Theorem A.** *Suppose $\mathcal{R}$ is type A. Let $f : I \to \mathbb{R}$ be a definable continuous function on an open interval $I$. Then for every $k \in \mathbb{N}$ there is a definable open dense $U \subseteq I$ on which $f$ is $C^k$.*

In the o-minimal setting this result is due to Laskowski and Steinhorn [11], and our proof of Theorem A uses ideas from their work. In particular, it also relies crucially on a classical theorem of Boas and Widder [2]. Note that the assumption of continuity is necessary, as $(\mathbb{R}, <, +, \mathbb{Q})$ is of type A and the characteristic function of $\mathbb{Q}$ is nowhere $C^1$. Of course, even when $f$ is continuous, there need not be an open dense definable set on which $f$ is smooth, as such a result fails already in the o-minimal setting by Rolin, Speissegger and Wilkie [14]. As a corollary to Theorem A, we obtain the following generalization of a theorem of Peterzil [13].

**Corollary A.** *Suppose that $\mathcal{R}$ is type A. Then one of the following holds:*

1. *For every continuous definable $f : [0,1] \to \mathbb{R}$ there is an open dense $U \subseteq [0,1]$ such that $f$ is affine on each connected component of $U$.*

(2) *There are definable functions* $\oplus, \otimes : I^2 \to I$ *on an open interval* $I$ *such that* $(I, <, \oplus, \otimes)$ *is isomorphic to* $(\mathbb{R}, <, +, \cdot)$.

We also give strong restrictions on continuous functions definable in type B structures. As noted above a type B expansions only defines multiplication by $\lambda$ for countably many $\lambda \in \mathbb{R}$ by [6, Theorem C]. Our first result for type B expansions here is the following counterpart to Theorem A.

**Theorem B.** *Suppose* $\mathcal{R}$ *is type B. Then every definable* $C^2$ *function* $f : [0,1] \to \mathbb{R}$ *is affine.*

Theorem B indicates that type B structures enjoy linearity properties. The following result gives further evidence in this direction.

**Theorem C.** *Suppose* $\mathcal{R}$ *is type B. Then*

(1) *a definable family of linear functions* $[0,1] \to \mathbb{R}$ *contains only finitely many distinct elements,*
(2) *every definable continuous function on a bounded interval is bounded,*
(3) *every definable continuous function* $\mathbb{R} \to \mathbb{R}$ *is eventually bounded above by a linear function,*
(4) *there is no interval* $I \subsetneq \mathbb{R}$ *and definable functions* $\oplus, \otimes : I^2 \to I$ *such that* $(I, <, \oplus, \otimes)$ *is an ordered field isomorphic to* $(\mathbb{R}, <, +, \cdot)$.

The proofs of Theorem B and C depend crucially on a slight generalization of Hieronymi and Tychonievich [9, Theorem A] first observed in [6, Proposition 3.8]. We establish Theorem B by showing that continuous functions in type B expansions are weakly periodic in a certain sense (see Theorem 3.4). This weak periodicity immediately rules out strictly convex definable functions, and therefore Theorem B follows. For Theorem C, we show that a type B expansion cannot define a weak pole, that is a family of continuous functions that surject arbitrarily small intervals onto a fixed nonempty open interval (see Definition 3.6). We then note that any expansion of $(\mathbb{R}, <, +)$ which satisfies the negation of any one of the statements in Theorem C defines a weak pole.

It is natural to ask if Theorem B can be strengthened to assert that every $C^1$ function definable in a Type B structure is affine. We do not know the answer to this question. We show the following:

**Theorem D.** *Suppose* $\mathcal{R}$ *defines a* $C^1$ *function* $[0,1] \to \mathbb{R}$ *with nonconstant derivative. Then one of the following holds:*

(1) $\mathcal{R}$ *interprets* $(\mathbb{R}, <, +, \cdot, \mathbb{Z})$.
(2) *there is an open interval* $I$ *and definable functions* $\oplus, \otimes : I^2 \to I$ *such that* $(I, <, \oplus, \otimes)$ *is isomorphic to* $(\mathbb{R}, <, +, \cdot)$.

Theorem C and D together yield the following corollary.

**Corollary B.** *If a type B expansion defines a* $C^1$ *function* $[0,1] \to \mathbb{R}$ *with nonconstant derivative then it interprets* $(\mathbb{R}, <, +, \cdot, \mathbb{Z})$.

It follows that every $C^1$ function definable in a type B structure with decidable theory is affine. Note that this covers the examples of type B structures above.

"Most" expansions are type C. The following Theorem gives a precise instance of this idea.

**Theorem E.** *Let $C([0,1])$ be the set of continuous functions $[0,1] \to \mathbb{R}$ equipped with the topology of uniform convergence. Then the set of $f \in C([0,1])$ such that $(\mathbb{R}, <, +, f)$ is type C is comeager in $C([0,1])$.*

Theorem E states that a generic bounded continuous function defines *all* bounded continuous functions over $(\mathbb{R}, <, +)$. The proof of Theorem E may be applied to many other Polish spaces of continuous functions $[0,1] \to \mathbb{R}$.

**Acknowledgements.** We thank Samantha Xu for useful conversations on the topic of this paper.

**Notations.** Let $X \subseteq \mathbb{R}^n$. We denote by $\mathrm{cl}(X)$ the closure of $X$, by $\mathrm{Int}(A)$ the interior of $X$, and by $\mathrm{Bd}(X)$ the boundary $\mathrm{cl}(X) \setminus \mathrm{Int}(X)$ of $X$. Whenever $X \subseteq \mathbb{R}^{m+n}$ and $x \in \mathbb{R}^m$, then $X_x$ denotes the set $\{y \in \mathbb{R}^n : (x,y) \in X\}$. We always use $i, j, k, l, m, n$ for natural numbers and $r, s, t, \lambda, \epsilon, \delta$ for real numbers. Given $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ we let $\|x\| := \max\{|x_1|, \dots, |x_n|\}$ be the $l_\infty$ norm of $x$.

## 2. DIFFERENTIABILITY IN TYPE A STRUCTURES

Throughout this section $\mathcal{R}$ is type A. In this section we prove Theorem A. The reader will find it helpful to have copies of [6, 10] handy, as we repeatedly make us of results from these papers.

### 2.1. Prerequisites.

Before diving into the proof of Theorem A, we establish a few basic facts about type A structures for later use. Recall that $X \subseteq \mathbb{R}^n$ is ***$D_\Sigma$*** if there is a definable family $\{Y_{r,s} : r, s > 0\}$ of compact subsets of $\mathbb{R}^n$ such that $X = \bigcup_{r,s} Y_{r,s}$, and $Y_{r,s} \subseteq Y_{r',s'}$ if $r \leq r'$ and $s \geq s'$, for all $r, r', s, s' > 0$. We say that the family $\{Y_{r,s} : r, s > 0\}$ witnesses that $X$ is $D_\Sigma$.

**Lemma 2.1.** *Let $I \subseteq \mathbb{R}$ be an open interval and $X \subseteq I \times \mathbb{R}_{>0}$ be $D_\Sigma$ such that $X_x$ is finite for every $x \in I$. Then there is an open subinterval $J \subseteq I$ and an $\epsilon > 0$ such that $J \times [0, \epsilon]$ is disjoint from $X$.*

*Proof.* Let $\pi : I \times \mathbb{R}_{>0} \to \mathbb{R}$ be the projection onto the first coordinate. Observe that $\pi(X)$ is $D_\Sigma$, and hence by [6, Theorem D] either has interior or is nowhere dense. Therefore we can reduce to the case that $\pi(X) = I$. Let $\{B_{s,t} : s, t \in \mathbb{R}_{>0}\}$ be a definable family of compact sets that witnesses that $X$ is $D_\Sigma$. Let

$$C_{s,t} = \pi(X \setminus B_{s,t}) \quad \text{and} \quad D_{s,t} = I \setminus C_{s,t} \quad \text{for all } s, t > 0.$$

Note that each $C_{s,t}$ is $D_\Sigma$. As $\pi(X) = I$, we have $x \in D_{s,t}$ if and only if $X_x \subseteq (B_{s,t})_x$. Since each $X_x$ is finite, every $x \in I$ is contained in some $D_{s,t}$. Thus $\bigcup_{s,t} D_{s,t} = I$. By the Baire Category theorem there are $s, t \in \mathbb{R}_{>0}$ such that $D_{s,t}$ is somewhere dense. Since $D_{s,t}$ is closed, $D_{s,t}$ has interior. Let $J$ be an open subinterval whose closure is contained in the interior of $D_{s,t}$. Then

$$X \cap (\mathrm{cl}(J) \times \mathbb{R}_{>0}) = B_{s,t} \cap (\mathrm{cl}(J) \times \mathbb{R}_{>0}).$$

As $\mathrm{cl}(J) \times \{0\}$ and $B_{s,t}$ are disjoint compact subsets of $\mathbb{R}^2$, there is an $\epsilon > 0$ such that no point in $B_{s,t}$ lies within distance $\epsilon$ of any point in $\mathrm{cl}(J) \times \{0\}$. It follows that $J \times [0, \epsilon]$ is disjoint from $B_{s,t}$ and thus disjoint from $X$. □

**Definition 2.2.** We say that $D \subseteq \mathbb{R}_{>0}$ is a **sequence set** if it is discrete and bounded with closure $D \cup \{0\}$.

It is easy to see that $(D,>)$ has order type $\omega$. By [10, Lemma 3.2] $\mathcal{R}$ either defines a sequence set or every bounded nowhere dense definable subset of $\mathbb{R}$ is finite.

**Lemma 2.3.** *Let $D$ be a definable sequence set and $X \subseteq D \times \mathbb{R}$ be definable such that $X_d$ is nowhere dense for each $d \in D$. Then $\bigcup_{d\in D} X_d$ is nowhere dense.*

*Proof.* Set

$$Y := \{(d,x) \in D \times \mathbb{R} \ : \ \exists e \in D \ e \geq d \wedge (e,x) \in X\}.$$

As $(D,>)$ has order type $\omega$, the set $\{e \in D \ : \ d \leq e\}$ is finite for every $d \in D$. Therefore $Y_d$ is nowhere dense for every $d \in D$. As $Y_d \subseteq Y_e$ when $d \geq e$, the family $\{Y_d \ : \ d \in D\}$ is increasing. By [10, Lemma 3.3] $\bigcup_{d\in D} Y_d$ is nowhere dense. It follows directly that $\bigcup_{d\in D} X_d$ is nowhere dense. □

2.2. **Proof of Theorem A.** Let $f : I \to \mathbb{R}$, where $I = (a,b)$ is an open interval, and $h = (h_1,\dots,h_k) \in \mathbb{R}^k$. We define the **generalized $k$-th difference of $f$** as follows:

$$\Delta^0 f(x) := f(x).$$

and for $k \geq 1$

$$\Delta^k_h f(x) := \Delta^{k-1}_{(h_1,\dots,h_{k-1})} f(x+h_k) - \Delta^{k-1}_{(h_1,\dots,h_{k-1})} f(x).$$

Note that for given $h$, the function $\Delta^k_h f$ is defined on the interval $(a, b-k\|h\|)$.

Let $J$ be a subinterval of $\mathbb{R}$ and $k \in \mathbb{N}$. A tuple $(u,x) \in \mathbb{R}^k_{\geq 0} \times J$ is **$(J,k)$-suitable** if $x + k\|u\| \in J$. We denote the set of such pairs by $S_{J,k}$. Note that $S_{J,k}$ is open and $\Delta^k_h f(x)$ is defined for each $(h,x) \in S_{J,k}$.

We apply a theorem of Boas and Widder. We first fix some notation.

**Definition 2.4.** Let $f : I \to \mathbb{R}$ be continuous. We say $H^f_k$ **holds on** $I$ if either

- $\Delta^k_{(h,\dots,h)} f(x) \geq 0$ for all $(x,h) \in I \times \mathbb{R}_{>0}$ with $((h,\dots,h),x) \in S_{I,k}$ or
- $\Delta^k_{(h,\dots,h)} f(x) \leq 0$ for all $(x,h) \in I \times \mathbb{R}_{>0}$ with $((h,\dots,h),x) \in S_{I,k}$.

If $I = (a,b)$ and $(x,h) \in I \times \mathbb{R}_{>0}$, then $((h,\dots,h),x) \in S_{I,k}$ if and only if $a < x < x + kh < b$.

**Fact 2.5** ([2, Theorem]). *Let $f : I \to \mathbb{R}$ be continuous and $k \geq 2$. If $H^f_k$ holds on $I$, then $f^{(k-2)}$ exists and is continuous on $I$.*

We now prove a lemma:

**Lemma 2.6.** *Let $f : J \to \mathbb{R}$ be a continuous definable function and $D$ be a definable sequence set. If $\Delta^k_{(h,d)} f(x) \geq 0$ for all $((h,d),x) \in S_{J,k} \cap (\mathbb{R}^{k-1} \times D) \times J$, then $\Delta^k_u f(x) \geq 0$ for $(u,x) \in S_{J,k}$.*

*Proof.* By continuity of $f$, it is enough to show that $\{(u,x) \in S_{J,k} \ : \ \Delta^k_u f(x) \geq 0\}$ is dense in $S_{J,k}$. Let $U \subseteq S_{J,k}$ be open. Let $(u_1,u_2,x) \in U$. As $D$ is a sequence set, there are $d_1,\dots,d_n \in D$ such that $(u_1,\sum_{i=1}^n d_i, x) \in U$. It is left to show the following claim: For every $j \in \{1,\dots,n\}$, $(u_1,\sum_{i=1}^j d_i, x) \in S_{J,k}$ and $\Delta^k_{(u_1,\sum_{i=1}^j d_i)} f(x) \geq 0$.

First observe that since $\sum_{i=1}^n d_i < u_2$ and $(u_1,u_2,d) \in S_{J,k}$, we have $(u_1,\sum_{i=1}^j d_i, x) \in S_{J,k}$. We now show the second statement of the claim by applying induction to $j$.

For $j = 1$, $\Delta^k_{(u_1,d_1)} f(x) \geq 0$ by our assumptions on $D$. So now let $j > 1$ and suppose $\Delta^k_{(u_1,\sum_{i=1}^{j-1} d_i)} f(x) \geq 0$. Since $(u_1, \sum_{i=1}^{j} d_i, x) \in S_{J,k}$, it follows immediately that $(u_1, d_j, x + \sum_{i=1}^{j-1} d_i) \in S_{J,k}$. Thus $\Delta^k_{(u_1,d_j)} f(x + \sum_{i=1}^{j-1} d_i) \geq 0$. Using the definition of the $\Delta^k$, we obtain

$$\begin{aligned}
\Delta^k_{(u_1,\sum_{i=1}^{j} d_i)} f(x) &= \Delta^{k-1}_{u_1} \Delta^1_{\sum_{i=1}^{j} d_i} f(x) \\
&= \Delta^{k-1}_{u_1} \left( f\left( x + \sum_{i=1}^{j} d_i \right) - f(x) \right) \\
&= \Delta^{k-1}_{u_1} \left( f\left( x + \sum_{i=1}^{j} d_1 \right) - f\left( x + \sum_{i=1}^{j-1} d_i \right) + f\left( x + \sum_{i=1}^{j-1} d_i \right) - f(x) \right) \\
&= \Delta^k_{(u_1,d_j)} f\left( x + \sum_{i=1}^{j-1} d_i \right) + \Delta^k_{(u_1,\sum_{i=1}^{j-1} d_i)} f(x) \geq 0.
\end{aligned}$$

□

Theorem A follows from Proposition 2.7 and Fact 2.5.

**Proposition 2.7.** *Let $f : I \to \mathbb{R}$ be a continuous definable function. Then there is a definable open dense subset $U$ of $I$ such that for each connected component $J$ of $U$ either*

- *$\Delta^k_h f(x) \geq 0$ for all $(h, x) \in S_{J,k}$, or*
- *$\Delta^k_h f(x) \leq 0$ for all $(h, x) \in S_{J,k}$.*

*Proof.* Let $a, b \in \mathbb{R}$ be such that $I = (a, b)$. We first treat the case when $\mathcal{R}$ defines a sequence set $D$. We proceed by induction on $k$. The case $k = 0$ follows immediately from the weak monotonicity theorem for type A structures (see [6, Fact 3.3]).

Let $k > 0$. Observe that for $d \in D$, $\Delta^k_{h,d} f = \Delta^{k-1}_h \Delta^1_d f$ and that $\Delta^1_d f$ is defined on the interval $(a, b - d)$. By the induction hypothesis we find for every $d \in D$ a dense definable open set $U_d \subseteq (a, b - d)$ such that for each connected component $J$ of $U_d$, either $\Delta^{k-1}_h \Delta^1_d f(x) \geq 0$ for all $(h, x) \in S_{J,k-1}$, or $\Delta^{k-1}_h \Delta^1_d f(x) \leq 0$ for all $(h, x) \in S_{J,k-1}$. For $d \in D$ set

$$X_d = \Big( (a, b - d) \setminus U_d \Big) \cup \{b - d\}.$$

By Lemma 2.3 $\bigcup_{d \in D} X_d$ is nowhere dense. Set $U := I \setminus \mathrm{cl}(\bigcup_{d \in D} X_d)$ and observe that $U$ is definable, open and dense in $I$. Let $J$ be a connected component of $U$. Then for each $d \in D$, either

(i) $(a, b - d) \cap J = \emptyset$ or
(ii) $J \subseteq (a, b - d)$ and one of the following is true:
  (a) $\Delta^{k-1}_h \Delta^1_d f(x) \geq 0$ for all $(h, x) \in S_{J,k-1}$, or
  (b) $\Delta^{k-1}_h \Delta^1_d f(x) \leq 0$ for all $(h, x) \in S_{J,k-1}$.

Since $D$ is a sequence set, there are infinitely many $d \in D$ for which (ii) holds. Denote the set all such $d \in D$ by $D'$. Let $D'' := \{d \in D' : \Delta^{k-1}_h \Delta^1_d f(x) \geq 0 \text{ for all } (h, x) \in S_{J,k-1}\}$. Then either $D' \setminus D''$ is infinite or $D'$ is infinite. Suppose that $D''$ is infinite. We now want to show that $\Delta^k_u f(x) \geq 0$ for all $(u, x) \in S_{J,k}$. By Lemma 2.6 it is enough to show that $\Delta^k_{h,d} f(x) \geq 0$ for all $((h, d), x) \in S_{J,k} \cap$

$(\mathbb{R}^{k-1} \times D'') \times J$. Let $(h, d, x) \in S_{J,k} \cap (\mathbb{R}^{k-1} \times D'') \times J$. By definition of $S_{J,k}$, we get that $x + k\|(h,d)\| \in J$. Thus $x + (k-1)\|h\| \in J$ and hence $(h, x) \in S_{J,k-1}$. Since $d \in D''$, we get

$$\Delta^k_{h,d} f(x) = \Delta^{k-1}_h \Delta^1_d f(x) \geq 0.$$

The case when $D' \setminus D''$ is infinite may be handled similarly.

We now suppose that $\mathcal{R}$ does not define a sequence set. Then by [10, Lemma 3.2] every bounded nowhere dense definable subset of $\mathbb{R}$ is finite. Set

$$V_1 := \{(x,h) \in I \times \mathbb{R}_{>0} \ : \ \Delta^k_{(h,\ldots,h)} f(x) \geq 0\}$$
$$V_2 := \{(x,h) \in I \times \mathbb{R}_{>0} \ : \ \Delta^k_{(h,\ldots,h)} f(x) \leq 0\}.$$

Both $V_1$ and $V_2$ are closed. Let $W := (I \times \mathbb{R}_{>0}) \setminus (\operatorname{Int} V_1 \cup \operatorname{Int} V_2)$. Then $W$ is $D_\Sigma$. Since $W \subseteq (V_1 \setminus \operatorname{Int} V_1) \cup (V_2 \setminus \operatorname{Int} V_2)$, $W$ is nowhere dense and therefore has no interior. Let $\pi : I \times \mathbb{R}_{>0} \to I$ be the coordinate projection onto $I$. Consider

$$Y := \{x \in I \ : \ \dim W_x \geq 1\}.$$

By [6, Fact 2.14(2)] $Y$ is $D_\Sigma$. By [6, Theorem 3] $\dim Y = 0$, so $Y$ is nowhere dense. Now let $U$ be the complement of $\operatorname{cl}(Y)$ in $I$. We that $\dim W_x = 0$ for all $x \in U$. In particular, each $W_x$ is nowhere dense and hence finite. Consider

$$Z := \{x \in U \ : \ \forall \delta, \epsilon > 0 \ (x - \delta, x + \delta) \times (0, \epsilon) \cap W \neq \emptyset\}.$$

We will show that $Z$ is nowhere dense. Suppose $J$ is an open subinterval of $I$ in which $Z$ is dense. Observe that $(J \times \mathbb{R}_{>0}) \cap W$ is $D_\Sigma$. Applying Lemma 2.1 to this set we get a subinterval $J'$ and an $\epsilon > 0$ such that $J' \times (0, \epsilon)$ is disjoint from $W$. This contradicts the density of $Z$ in $J$. Thus $Z$ is nowhere dense. Let $U'$ be the complement of $\operatorname{cl}(Z)$. Let $J''$ be a connected component of $U'$. Let $x \in J''$. As $x \notin Z$, there are $\delta, \epsilon > 0$ such that $(x - \delta, x + \delta) \times (0, \epsilon) \cap W = \emptyset$. It follows from connectedness that $(x - \delta, x + \delta) \times (0, \epsilon)$ is contained in $V_1$ or $V_2$. Thus $H^f_k$ holds on $(x - \delta, x + \delta)$. □

## 3. CONTINUOUS FUNCTIONS IN TYPE B STRUCTURES

The goal of this section is to prove Theorem B and C. The key ingredient is the following generalization of [9, Theorem A].

**Fact 3.1** ([6, Proposition 3.8]). *If $\mathcal{R}$ defines an order $(D, \prec)$, an open interval $U \subseteq \mathbb{R}$, and a function $g \colon \mathbb{R}^3 \times D \to D$ such that*

(i) *$(D, \prec)$ has order type $\omega$ and $D$ is dense in $U$,*
(ii) *for every $a, b \in U$ and $e, d \in D$ with $a < b$ and $e \preceq d$,*

$$\{c \in \mathbb{R} : g(c, a, b, d) = e\} \cap (a, b) \text{ has nonempty interior,}$$

*then $\mathcal{R}$ is type C.*

The following fact is a corollary of Fact 3.1 which is often easier to apply. Loosely speaking, it says that every expansion of $(\mathbb{R}, <, +)$ that defines two sufficiently independent $\omega$-orderable sets, is type C.

**Fact 3.2** ([1, Lemma 3.7]). *If there exists two dense $\omega$-orderable subset $C$ and $D$ of $(0, 1)$ such that $(C - C) \cap (D - D) = \{0\}$, then $\mathcal{R}$ is type C.*

3.1. **Proof of Theorem B.** Throughout this section we suppose that $\mathcal{R}$ is type B. We apply Fact 3.2 to continuous functions definable in type B expansions.

**Lemma 3.3.** *Let $I$ be an open interval in $\mathbb{R}$, $D$ be a dense $\omega$-orderable subset of $I$, and $f : I \to \mathbb{R}$ be definable, continuous and nonconstant. Then for every open interval $J \subseteq I$ there are $d_1, d_2, d_3, d_4 \in J \cap D$ such that*

- $d_1 \neq d_2$, $d_3 \neq d_4$,
- $d_1 - d_2 = f(d_3) - f(d_4)$.

*Proof.* Let $J \subseteq I$ be an open subinterval of $I$. Let $a, b \in \mathbb{R}$ be such that $J = (a, b)$. By the intermediate value theorem there is an open interval $J' \subseteq f(J)$. Since $D$ is dense in $I$, we have that $f(D \cap J) \cap J'$ is dense in $J'$. Let $a', b' \in \mathbb{R}$ be such that $(a', b') = J'$. By lowering $b'$, we can assume that $b' - a' < b - a$. Then both $(-a + D) \cap (0, b' - a')$ and $-a' + f(D \cap J) \cap J'$ are dense $\omega$-orderable subsets of $(0, b' - a')$. By Fact 3.2, there are $d_1, d_2, d_3, d_4 \in J$ such that $d_1 \neq d_2$, $d_3 \neq d_4$ and

$$d_1 - d_2 = (-a + d_1) - (-a + d_2) = -a' + f(d_3) - \big( - a' + f(d_4)\big) = f(d_3) - f(d_4).$$

□

**Theorem 3.4** (Weak periodicity). *Let $I$ be a bounded open interval and $f : I \to \mathbb{R}$ be definable and continuous. Then for every open interval $J \subseteq I$ there are $x, y \in J$ and $\delta > 0$ such that $x \neq y$ and for all $\epsilon_1, \epsilon_2 \in \mathbb{R}$ with $|\epsilon_1|, |\epsilon_2| < \delta$*

$$f(x + \epsilon_1) - f(y + \epsilon_1) = f(x + \epsilon_2) - f(y + \epsilon_2).$$

*Proof.* Since $\mathcal{R}$ is type B, there is a dense $\omega$-orderable subset $D$ of $I$. Let $J \subseteq I$ be an open interval. We can directly reduce to the case that $f$ is nonconstant on any open subinterval of $J$. Let $C$ be the set

$$\{(d_1, d_2, d_3, d_4) \in D^4 \ : \ d_1 \neq d_2, d_3 \neq d_4\}.$$

For $d = (d_1, d_2, d_3, d_4) \in C$, let $A_d \subseteq \mathbb{R}$ be the set of $t$ such that

- $t + d_3, t + d_4 \in J$, and
- $d_1 - d_2 = f(t + d_3) - f(t + d_4)$.

For each $d \in C$, the set $A_d$ is closed in $J$ by continuity of $f$. Let $s > 0$ and $J'$ be an open subinterval of $J$ such that $t + J' \subseteq J$ for all $t \in (0, s)$. Let $t \in (0, s)$. Consider the function $g_t : J' \to \mathbb{R}$ that maps $c \in J'$ to $f(t + c)$. Applying Lemma 3.3 to $g_t$ we obtain a $d = (d_1, d_2, d_3, d_4) \in C$ such that:

$$d_1 - d_2 = g_t(d_3) - g_t(d_4) = f(t + d_3) - f(t + d_4).$$

Thus $t \in A_d$. Therefore $(0, s) \subseteq \bigcup_{d \in C} A_d$. By the Baire Category Theorem $A_d$ has interior for some $d \in C$. Fix such a $d = (d_1, d_2, d_3, d_4) \in C$ and let $J''$ be an open interval in the interior of $A_d$. The the function $J'' \to \mathbb{R}$ given by $t \mapsto f(t + d_3) - f(t + d_4)$ is constant. The statement of the Theorem follows. □

Recall that if $f : [0, 1] \to \mathbb{R}$ is strictly convex if and only if

$$\frac{f(y) - f(x)}{y - x} < \frac{f(y') - f(x')}{y' - x'} \quad \text{for all } 0 < x < y < x' < y' < 1.$$

**Corollary 3.5.** *There is no strictly convex definable function $f : [0, 1] \to \mathbb{R}$.*

*Proof.* Suppose that $f : [0,1] \to \mathbb{R}$ is continuous and definable. By Theorem 3.4 there are $x, y \in [0,1]$ and $\delta > 0$ such that $x < y$ and for all $\epsilon_1, \epsilon_2 \in \mathbb{R}$ with $|\epsilon_1|, |\epsilon_2| < \delta$

$$f(x+\epsilon_1) - f(y+\epsilon_1) = f(x+\epsilon_2) - f(y+\epsilon_2).$$

Taking $\epsilon_2 = 0$ and $\epsilon_1 < y - x, 1 - y$ we have

$$f(x+\epsilon_1) - f(x) = f(y+\epsilon_1) - f(y) \quad \text{and } 0 < x < x+\epsilon_1 < y < y+\epsilon_1 < 1.$$

So $f$ is not strictly convex. ☐

*Proof of Theorem B.* Set

$$\begin{aligned} B_+ &:= \{x \in [0,1] \ : \ f''(x) > 0\}, \\ B_0 &:= \{x \in [0,1] \ : \ f''(x) = 0\}, \\ B_- &:= \{x \in [0,1] \ : \ f''(x) < 0\}. \end{aligned}$$

By continuity of $f'$, $B_+, B_-$ are open and $B_0$ is closed. Suppose $B_+$ is non-empty. Let $I$ be an open interval in $B_+$. Then $f'$ is strictly increasing on $I$. Since the integral of a strictly increasing function is strictly convex, the restriction of $f$ to $I$ is strictly convex. This contradicts Corollary 3.5. Thus $B_+ = \emptyset$. Similarly we can show that $B_- = \emptyset$. Thus $B_0 = [0,1]$. ☐

3.2. **Proof of Theorem C.** Recall that a **pole** is a definable homeomorphism between a bounded and an unbounded interval. In the following we will consider structures that might not define a pole, but still admit a definable family of maps that continuously surject arbitrarily small intervals onto a fixed open interval.

**Definition 3.6.** A **weak pole** is a definable family $\{h_d : d \in E\}$ of continuous maps $h_d : [0,d] \to \mathbb{R}$ such that for some $\delta > 0$:

(i) $E \subseteq \mathbb{R}_{>0}$ contains a sequence set,
(ii) $[0,\delta] \subseteq h_d([0,d])$ for all $d \in E$.

It is easy to see that if $\mathcal{R}$ admits a weak pole whenever it defines a pole.

**Theorem 3.7.** *Suppose that $\mathcal{R}$ defines a weak pole and a dense $\omega$-orderable set. Then $\mathcal{R}$ is type C.*

*Proof.* Let $(D, \prec)$ be a dense $\omega$-orderable set and $\{h_d : d \in E\}$ be a weak pole. Since $\mathcal{R}$ expands $(\mathbb{R}, <, +)$, we can assume that $E$ is closed, $D$ is dense in $[0,1]$ and $[0,1] \subseteq h_d([0,d])$ for all $d \in E$. Let $Z = \{(a,b) \in [0,1]^2 \ : \ a < b\}$ and let $\lambda : \mathbb{R}_{>0} \to E$ map $x$ to $\max(-\infty, x] \cap E$. We now define $g : [0,1] \times Z \times D \to D$ to be the function that maps $(c,a,b,d)$ to

$$\begin{cases} d, & \text{if } c - a > \lambda(b-a); \\ \prec\text{-minimal } e \in D_{\preceq d} \text{ s.t. } h_{\lambda(b-a)}(c-a) - e \text{ is minimal}, & \text{otherwise.} \end{cases}$$

We will now show that $g$ satisfies the assumptions of Fact 3.1. For this, let $a, b \in Z$ and $d, e \in D$ with $e \preceq d$. Since $[0,1] \subseteq h_{\lambda(b-a)}([0,\lambda(b-a)])$, there is $z \in [0, \lambda(b-a)]$ such that $h_{\lambda(b-a)}(z) = e$. Since $D_{\preceq d}$ is finite and $h_{\lambda(b-a)}$ is continuous, there is an open interval $I$ around $z$ such that for each $y \in I$, $e$ is the only element in $D_{\preceq d}$ such that $h_{\lambda(b-a)}(y) - e$ is minimal. Let $c \in (a,b)$ such that $c - a = z$. From the above argument it follows immedieately that $g(x,a,b,d) = e$ for all $x \in c + I \cap (a,b)$. Thus (ii) of Fact 3.1 holds for our choice of $g$. ☐

The first statement of Theorem C follows immediately from Theorem 3.7 and the following Proposition.

**Proposition 3.8.** *Let* $\{f_x : x \in \mathbb{R}^l\}$ *be a definable family of linear functions* $[0,1] \to \mathbb{R}$ *that has infinitely many distinct elements. Then* $\mathcal{R}$ *admits a weak pole.*

*Proof.* After replacing each $f_x$ with $|f_x|$ if necessary we suppose that each $f_x$ takes nonnegative values. Set $B = \{f_x(1) : x \in \mathbb{R}^l\}$. Let $g : B \times [0,1] \to \mathbb{R}$ maps $(\lambda, t)$ to $f_y(t)$ for any $y \in \mathbb{R}^l$ with $f_y(1) = \lambda$. Note that $g$ is well-defined and $g(\lambda, t) = \lambda t$ for all $t \in [0,1]$. We now extend $g$ to the closure of $B$. Let $\lambda \in \operatorname{cl}(B)$. We declare

$$\tilde{g}(\lambda, t) = \lim_{\lambda' \in B, \lambda' \to \lambda} g(\lambda', t) \quad \text{for all } t \in [0,1].$$

After replacing $g$ by $\tilde{g}$ and $B$ by $\operatorname{cl}(B)$ we suppose $B$ is closed. Then $B$ is a closed and infinite subset of $\mathbb{R}_{>0}$. Therefore one of the following holds:

(1) $B$ is unbounded.
(2) $B$ has an accumulation point.

We first suppose that $B$ is unbounded. Let $\{h_d : d \in \mathbb{R}_{>0}\}$ be the definable family of functions $h_d : [0,d] \to \mathbb{R}$ given by declaring $h_d(t) = g(\lambda, t)$ where $\lambda$ is the minimal element of $B$ such that $g(\lambda, d) \geq 1$. Then $h_d(t) \geq d^{-1}t$ for all $t \in [0,d]$. It directly follows that $\{h_d : d \in \mathbb{R}_{>0}\}$ is a weak pole.

Now suppose that (2) holds. Let $\lambda$ be an accumulation point of $B$. We declare

$$\psi(\lambda', t) := |g(\lambda, t) - g(\lambda', t)| \quad \text{for all } \lambda' \in B, t \in [0,1].$$

Note that $\psi$ is definable and $\psi(\lambda', t) = |\lambda - \lambda'| t$ for $\lambda' \in B, t \in [0,1]$. Set $C = \{|\lambda - \lambda'| : \lambda' \in B\}$. Let $\{h_d : d \in C\}$ be the definable family of functions $h_d : [0,d] \to \mathbb{R}$ where $h_d$ is the compositional inverse of $t \mapsto \psi(\lambda', t)$ where $\lambda' \in B$ and $d = |\lambda - \lambda'|$. Note that $C$ contains arbitrarily small positive elements as $\lambda$ is an accumulation point of $B$. Then $h_d$ satisfies $h_d(t) = d^{-1}t$. It follows that $\{h_d : d \in C\}$ is a weak pole. ☐

We now prove the second and third statement of Theorem C. We first recall a basic real analysis fact.

**Lemma 3.9.** *Let* $I$ *be an open interval,* $f : I \to \mathbb{R}$ *a function, and* $\epsilon, \delta > 0$ *be such that* $|f(t) - f(t')| < \delta$ *whenever* $t, t' \in I$ *satisfy* $|t - t'| < \epsilon$. *Then*

$$|f(t) - f(t')| \leq \frac{\delta}{\epsilon}|t - t'| + 2\delta \quad \textit{for all } t, t' \in I.$$

*It follows that if* $f$ *is bounded when* $I$ *is bounded and* $f$ *is eventually bounded from above by a linear function when* $I = \mathbb{R}$.

*Proof.* Let $t, t' \in \mathbb{R}$. If $|t - t'| < \epsilon$ then $|f(t) - f(t')| < \delta$ and the lemma holds. We therefore suppose that $|t - t'| \geq \epsilon$ and further suppose without loss of generality that $t < t'$. Let $t = x_0 < \ldots < x_n = t'$ be such that $|x_i - x_{i+1}| < \epsilon$ for all $1 \leq i \leq n-1$ where $n = \left\lceil \frac{|t-t'|}{\epsilon} \right\rceil + 1$. We compute

$$\begin{aligned} |f(t) - f(t')| &\leq \sum_{i=0}^{n-1} |f(x_i) - f(x_{i+1})| < n\delta \\ &= \delta \left\lceil \frac{|t - t'|}{\epsilon} \right\rceil + \delta < \frac{\delta}{\epsilon}|t - t'| + 2\delta. \end{aligned}$$

Thus the lemma holds in the case $|t - t'| \geq \epsilon$. □

The second and third statements of Theorem C follow from the Proposition below, Lemma 3.9, and Theorem 3.7.

**Proposition 3.10.** *Suppose $\mathcal{R}$ does not define a weak pole. Let $I$ be an open interval and let $f : I \to \mathbb{R}$ be continuous and definable. Then $f$ is uniformly continuous.*

*Proof.* We suppose that $f$ is not uniformly continuous and show that $\mathcal{R}$ defines a weak pole. Let $\delta > 0$ be such that for all $\epsilon > 0$ there are $t, t' \in I$ such that $|f(t) - f(t')| \geq \delta$ and $|t - t'| \leq \epsilon$. For every $\epsilon > 0$ let

$$A_\epsilon := \{t \in I : |f(t) - f(t')| \geq \delta \text{ for some } t \leq t' \leq t + \epsilon\}.$$

Note that each $A_\epsilon$ is closed in $I$ and nonempty. Let $p$ be a fixed element of $I$. Let $g_0(\epsilon)$ be the maximal element of $A_\epsilon \cap (\infty, p]$ if $A_\epsilon \cap (\infty, p] \neq \emptyset$ and the minimal element of $A_\epsilon \cap [p, \infty)$ otherwise. Note that $g_0 : \mathbb{R}_{>0} \to I$ is definable. Let $g_1(\epsilon)$ be the least $t' \in [g_0(\epsilon), g_0(\epsilon) + \epsilon]$ such that $|f(g_0(\epsilon)) - f(t')| \geq \delta$. Then $g_1 : \mathbb{R}_{>0} \to I$ is definable and for all $\epsilon > 0$:

$$0 < g_1(\epsilon) - g_0(\epsilon) \leq \epsilon \quad \text{and} \quad |f(g_1(\epsilon)) - f(g_0(\epsilon))| \geq \delta.$$

We consider the definable family of functions $h_\epsilon : [0, g_1(\epsilon) - g_0(\epsilon)] \to \mathbb{R}$ given by declaring

$$h_\epsilon(t) = |f(g_0(\epsilon) + t) - f(g_0(\epsilon))|.$$

Each $h_\epsilon$ is continuous, it follows from the intermediate value theorem that $[0, \delta]$ is contained in the image of every $h_\epsilon$. Thus $\{h_\epsilon : \epsilon \in \mathbb{R}_{>0}\}$ is a weak pole. □

It is easy to see that $\mathcal{R}$ defines a weak pole when the last statement in Theorem C holds. This completes our proof of Theorem C.

## 4. WHAT ABOUT $C^1$-FUNCTIONS?

In this section we prove Theorem D.

**Definition 4.1.** We say that $\mathcal{R}$ is of **field-type** if there is an interval $I \subseteq \mathbb{R}$ together with definable functions $\oplus, \otimes : I^2 \to I$ such that $(I, <, \oplus, \otimes)$ is an ordered field isomorphic to $(\mathbb{R}, <, +, \cdot)$.

**Theorem 4.2.** *Let $I$ be a closed interval and let $f : I \to \mathbb{R}$ be a definable $C^1$ function with non-constant derivative. If there is an open subinterval of $I$ on which $f'$ is strictly increasing or strictly decreasing, then $\mathcal{R}$ is of field-type. If $f'$ is not strictly increasing or strictly decreasing on any open subinterval, then $\mathcal{R}$ interprets $(\mathbb{R}, <, +, \cdot, \mathbb{Z})$. In particular, if $f$ is $C^2$ then $\mathcal{R}$ is of field-type.*

*Proof.* Let $I = [a, b]$. After shrinking $I$ if necessary we suppose that one of the following two cases holds:

(a) $f'$ is strictly increasing or strictly decreasing on $I$.
(b) there is no open subinterval of $I$ on which $f'$ is strictly increasing or strictly decreasing

After shrinking $I$ if necessary we suppose $f'(a) \neq f'(b)$. After replacing $f$ with $-f$ if necessary we suppose $f'(a) < f'(b)$. If case (a) this implies that $f'$ is strictly increasing. Let $f'(a) < q < f'(b)$ be rational. It follows from continuity of $f'$ and the intermediate value theorem that there is an $a < x < b$ such that $f'(x) = q$.

Continuity of $f'$ implies that the set of such $x$ is closed, let $a'$ be the maximal element of $[a,b]$ such that $f'(a') = q$. Note that $a' < b$. After replacing $a$ with $a'$ if necessary we suppose that $f'(a) = q$ and $f'(x) > q$ for all $a < x \leq b$. Let $h : [a,b] \to \mathbb{R}$ be given by $h(x) = f(x) - (x-a)q$. Note that $h'(x) = f'(x) - q$. Thus $h'(a) = 0$ and $h'(x) > 0$ for all $a < x \leq b$. In particular $h'$ is strictly increasing when $f'$ is, so $h'$ is strictly increasing in case $(a)$. Note that $h$ is definable as $q$ is rational. After replacing $f$ with $h$ if necessary we suppose $f'(a) = 0$. Let $N \in \mathbb{N}$ be such that $f'(b) \geq \frac{1}{N}$. After replacing $f$ with $Nf$ if necessary we suppose $f'(b) \geq 1$. Let $b'$ be the minimal element of $[a,b]$ such that $f'(b') = 1$. After replacing $b$ with $b'$ if necessary we suppose that $f'(b) = 1$ and that $0 < f'(x) < 1$ for all $a < x < b$.

**Claim 4.3.** *Let* $\{g_x : x \in X\}$ *be a definable family of functions defined on closed intervals with right endpoint zero such that each* $g'_x(0)$ *exists for all* $x \in X$. *Then the relations* $g'_x(0) < g'_x(0)$, $g'_x(0) \leq g'_y(0)$, *and* $g'_x(0) = g'_y(0)$ *are definable on* $X$.

We only prove the first claim, the latter two follow. We show that $g'_x(0) < g'_y(0)$ if and only if there is a $a < z < b$ such that

$$\star \quad g_x(\epsilon) + [f(z+\epsilon) - f(z)] < g_y(\epsilon) \quad \text{for sufficiently small } \epsilon > 0.$$

Suppose that $\star$ holds. Let $a < z < b$ be such that

$$g_x(\epsilon) + [f(z+\epsilon) - f(z)] < g_y(\epsilon) \quad \text{for sufficiently small } \epsilon > 0.$$

Dividing by $\epsilon$ and taking the limit $\epsilon \to 0$ we have $g'_x(0) + f'(z) \leq g'_y(0)$. As $z > a$ we have $f'(z) > 0$ so $g'_x(0) < g'_y(0)$.

Now suppose that $g'_x(0) < g'_y(0)$. Let $\delta > 0$ be such that $g'_x(0) + \delta < g'_y(0)$. As $f'$ is continuous and $f'(a) = 0$ there is a $a < z < b$ such that $f'(z) < \delta$. Then $g'_x(0) + f'(z) < g'_y(0)$, that is

$$\lim_{\epsilon \to 0} \frac{g_x(\epsilon)}{\epsilon} + \lim_{\epsilon \to 0} \frac{f(z+\epsilon) - f(z)}{\epsilon} < \lim_{\epsilon \to 0} \frac{g_y(\epsilon)}{\epsilon}.$$

Hence $\star$ holds. Claim 4.3 follows.

Applying Claim 4.3 to the definable family $g_x(t) = f(x+t) - f(x)$ we see that the relations $f'(x) < f'(y)$, $f'(x) \leq f'(y)$ and $f'(x) = f'(y)$ are definable on $I$.

We let $E \subseteq [a,b]$ be the set of $x$ such that $f'(y) < f'(x)$ for all $a \leq y < x$. It follows from Claim 4.3 that $E$ is definable. For every $0 \leq t \leq 1$ the set $\{a \leq z \leq b : f'(z) \geq t\}$ is closed and nonempty and thus has a minimal element $x$ which must satisfy $f'(z) = t$. This minimal elements is in $E$. Thus for every $0 \leq t \leq 1$ there is an $x \in E$ such that $f'(x) = t$. Note also that if $x, y \in E$ and $x < y$ then $f'(x) < f'(y)$.

In case $(a)$ we trivially have $E = I$. If $E$ contains an open interval then $f'$ must be strictly increasing on that interval. Thus $E$ cannot have interior in case $(b)$. A somewhere dense subset of $\mathbb{R}$ closed under limits of increasing sequences has interior. Claim 4.4 below thuss implies that $E$ is nowhere dense in case $(b)$.

**Claim 4.4.** *Every limit of an increasing sequence of elements of* $E$ *is in* $E$.

Let $\{x_i\}_{i \in \mathbb{N}}$ be a strictly increasing sequence of elements of $E$ with limit $x$. We have $f'(x_i) < f'(x_j)$ whenever $i < j$. As $f'(x) = \lim_{i \to \infty} f'(x_i)$ we have $f'(x_i) < f'(x)$ for all $i$. Suppose $0 \leq y < x$. Then $y < x_j < x$ for some $j$. As $x_j \in E$ we have $f'(y) < f'(x_j)$, hence $f'(y) < f'(x)$. Thus $x \in E$. This proves Claim 4.4.

Given $x \in E \setminus b$ and $0 \leq t \leq b - x$ we let $f_x(t) = f(x+t) - f(x)$. We declare $f_b(t) = 1$ for all $t > 0$. Then $f'_x(0) = f'(x)$ for all $x \in E$. As $f$ is strictly increasing, each $f_x$ is strictly increasing. After translating $I$ if necessary we suppose $a = 0$, so $E$ is a subset of $[0, b]$. We declare

$$E_1 = -(E \setminus \{0, b\}) + b, \quad E_2 = -(E \setminus \{0\}), \quad E_3 = -E_1.$$

Then $E, E_1, E_2, E_3$ are pairwise disjoint as they are subsets of $[0, b], (b, 2b), [-b, 0)$, and $(-2b, -b)$ respectively. We let $F = E \cup E_1 \cup E_2 \cup E_3$. Note that in case $(a)$ we have $F = (-2b, 2b)$, so $F$ is an interval. We construct a definable family of functions $\{g_x : x \in F\}$ with the following two properties:

(1) For all $t \in \mathbb{R}$ there is a unique $x \in F$ such that $g'_x(0) = t$.
(2) If $x, y \in F$ and $x < y$ then $g'_x(0) < g'_y(0)$.

If $x \in E$ then we let $g_x = f_x$. If $x \in E_1$ then we let $g_x$ be the compositional inverse of $f_{2b-x}$. As each $f_x$ is strictly increasing each $f_x$ has compositional inverse. We have $g'_x(0) = f'_{2b-x}(0)^{-1}$ for all $x \in E_1$. It follows that for every $t > 1$ there is a unique $x \in E_1$ such that $f'_x(0) = t$. If $x \in E_2$ then we let $g_x = -f_{-x}$. We then have $g'_x(0) = -f'_{-x}(0)$ for all $x \in E_2$. It follows that for every $-1 \leq t < 0$ there is a unique $x \in E_2$ such that $g'_x(0) = t$. If $x \in E_3$ then we let $g_x = -g_{-x}$. It follows that for every $t < -1$ there is a unique $x \in E_3$ such that $g'_x(0) = t$. Conditions (1) and (2) above follow.

We now define functions $\oplus, \otimes : F^2 \to F$. Given $x, y \in F$ we let $x \oplus y$ be the unique element of $F$ such that

$$g'_{x \oplus y}(0) = (g_x + g_y)'(0)$$

and $x \otimes y$ be the unique element of $F$ such that

$$g'_{x \otimes y}(0) = (g_x \circ g_y)'(0).$$

It follows from Claim 4.3 that $\oplus$ and $\otimes$ are definable. For all $x, y \in F$:

$$g'_{x \oplus y}(0) = g'_x(0) + g'_y(0) \text{ and } g'_{x \otimes y}(0) = g'_x(0) g'_y(0).$$

So $x \mapsto g'_x(0)$ gives an isomorphism $(F, <, \oplus, \otimes) \to (\mathbb{R}, <, +, \cdot)$. As observed above, $F$ is an interval in case $(a)$ and is nowhere dense in case $(b)$. Thus $\mathcal{R}$ is of field type in case $(a)$. The next claim shows that $\mathcal{R}$ interprets second order arithmetic in case $(b)$.

**Claim 4.5.** *If $F$ is nowhere dense then $\mathcal{R}$ interprets second order arithmetic.*

Suppose $F$ is nowhere dense. Let $U$ be the complement of the closure of $F$. Claim 4.4 implies that the left endpoint of every bounded connected component of $U$ is in $E$. We let $D \subseteq E$ be the definable set of left endpoints of bounded connected components of $U$. We define an $\omega$-order $\prec$ on $D$. Let $\delta : D \to \mathbb{R}$ be the definable function such that $\delta(d)$ is the length of connected component of $U$ with left endpoint $d$. We declare $d \prec d'$ if $\delta(d') < \delta(d)$ or if $\delta(d') = \delta(d)$ and $d < d'$. It is easy to see that $\prec$ is an $\omega$-order on $D$, this is shown in Section 2 of [10]. Furthermore $D$ is dense in $F$ as $F$ is nowhere dense. Consider the structure $(F, <, \oplus, \otimes, D, \prec)$. This structure is isomorphic to an expansion of $\mathbb{R}$ which admits a dense $\omega$-orderable set. Let

$$Z = \{x \in F : g'_x(0) \in \mathbb{Z}\}.$$

Then $(F, <, \oplus, \otimes, Z)$ is isomorphic to $(\mathbb{R}, <, +, \cdot, \mathbb{Z})$. It follows from [7] that $(F, <, \oplus, \otimes, D, \prec)$ defines $Z$. This completes the proof of Theorem D. ☐

Theorem D and Corollary A follow from the previous theorem.

**Corollary 4.6.** *Let* $f : [0,1] \to \mathbb{R}$ *be nowhere* $C^k$ *for some* $k$.

1. *If* $f$ *is continuous then* $(\mathbb{R}, <, +, f)$ *interprets the monadic second order theory of one successor.*
2. *If* $f$ *is* $C^1$ *then* $(\mathbb{R}, <, +, f)$ *interprets* $(\mathbb{R}, <, +, \cdot, \mathbb{Z})$.
3. *If* $f$ *is* $C^2$ *then* $(\mathbb{R}, <, +, f)$ *defines all compact subsets of all* $\mathbb{R}^k$.

*Proof.* Item (1) follows from Theorem A and [10, Theorem A]. Item (3) follows from Theorem A and Theorem B. Suppose $f$ is $C^1$. As $f$ is nowhere $C^k$ it follows from Theorem A that $(\mathbb{R}, <, +, f)$ admits a dense $\omega$-orderable set. If $f$ or $-f$ is strictly convex on some open subinterval then $(\mathbb{R}, <, +, f)$ defines all compact sets by Corollary 3.5 and hence interprets $(\mathbb{R}, <, +, \cdot, \mathbb{Z})$. If $f$ and $-f$ are not strictly convex on any open subinterval then $f'$ is not strictly increasing or strictly decreasing on any open subinterval. It follows from the proof of Theorem D that $(\mathbb{R}, <, +, f)$ interprets $(\mathbb{R}, <, +, \cdot, \mathbb{Z})$. ☐

## 5. GENERIC FUNCTIONS ARE TYPE C

In this section we describe the proof of Theorem E without giving full details. It is classically known that the set of somewhere differentiable continuous functions $[0,1] \to \mathbb{R}$ is meager in $C([0,1])$. It therefore suffices to show that the collection of all continuous function $[0,1] \to \mathbb{R}$ definable in type B expansions is meager. By Theorem 3.4 it suffices to show that the set of $f \in C([0,1])$ such that for some $0 < x, y < 1$ and $0 < \delta \leq y - x$ we have

$$f(x+\epsilon) - f(x) = f(y+\epsilon) - f(y) \quad \text{for all } 0 < \epsilon \leq \delta$$

is meager. It is enough to show that for each $n \geq 1$ the set $A_n$ of $f \in C([0,1])$ such that for some $0 < x, y < 1$ we have $\frac{1}{n} \leq y - x$ and

$$f(x+\epsilon) - f(x) = f(y+\epsilon) - f(y) \quad \text{for all } 0 < \epsilon \leq \frac{1}{n}$$

is nowhere dense. Each $A_n$ is a closed subset of $C([0,1])$, so it suffices to show that each $A_n$ has empty interior in $C([0,1])$. For any $n, f \in C([0,1])$, and $\epsilon > 0$ it is easy to construct a piecewise linear $g \in C([0,1])$ such that $\|f - g\| < \epsilon$ and $g \notin A_n$. So $A_n$ has empty interior.

## REFERENCES

[1] W. Balderrama and P. Hieronymi. Definability and decidability in expansions by generalized Cantor sets. *Preprint arXiv:1701.08426*, 2017.

[2] R. P. Boas, Jr. and D. V. Widder. Functions with positive differences. *Duke Math. J.*, 7:496–503, 1940.

[3] B. Boigelot, S. Rassart, and P. Wolper. On the expressiveness of real and integer arithmetic automata (extended abstract). In *Proceedings of the 25th International Colloquium on Automata, Languages and Programming*, ICALP '98, pages 152–163, London, UK, UK, 1998. Springer-Verlag.

[4] J. R. Büchi. On a decision method in restricted second order arithmetic. In *Logic, Methodology and Philosophy of Science (Proc. 1960 Internat. Congr .)*, pages 1–11. Stanford Univ. Press, Stanford, Calif., 1962.

[5] É. Charlier, J. Leroy, and M. Rigo. An analogue of Cobham's theorem for graph directed iterated function systems. *Adv. Math.*, 280:86–120, 2015.

[6] A. Fornasiero, P. Hieronymi, and E. Walsberg. How to avoid a compact set. arXiv:1612.00785, 2016.

[7] P. Hieronymi. Defining the set of integers in expansions of the real field by a closed discrete set. *Proc. Amer. Math. Soc.*, 138(6):2163–2168, 2010.
[8] P. Hieronymi. Expansions of the ordered additive group of real numbers by two discrete subgroups. *J. Symbolic Logic*, 81(3):1007–1027, 2016.
[9] P. Hieronymi and M. Tychonievich. Interpreting the projective hierarchy in expansions of the real line. *Proc. Amer. Math. Soc.*, 142(9):3259–3267, 2014.
[10] P. Hieronymi and E. Walsberg. Interpreting the monadic second order theory of one successor in expansions of the real line. *Israel J. Math.*, to appear, arXiv:1601.04555, 2016.
[11] M. C. Laskowski and C. Steinhorn. On o-minimal expansions of Archimedean ordered groups. *J. Symbolic Logic*, 60(3):817–831, 1995.
[12] C. Miller. Tameness in expansions of the real field. In *Logic Colloquium '01*, volume 20 of *Lect. Notes Log.*, pages 281–316. Assoc. Symbol. Logic, Urbana, IL, 2005.
[13] Y. Peterzil. Reducts of some structures over the reals. *J. Symbolic Logic*, 58(3):955–966, 1993.
[14] J.-P. Rolin, P. Speissegger, and A. J. Wilkie. Quasianalytic Denjoy-Carleman classes and o-minimality. *J. Amer. Math. Soc.*, 16(4):751–777, 2003.
[15] P. Simon. *A guide to NIP theories*, volume 44 of *Lecture Notes in Logic*. Cambridge University Press, 2015.

Department of Mathematics, University of Illinois at Urbana-Champaign, 1409 West Green Street, Urbana, IL 61801
*E-mail address*: `phierony@illinois.edu`
*URL*: `http://www.math.uiuc.edu/~phierony`

*E-mail address*: `erikw@illinois.edu`

# SIGNATURES, SUMS OF HERMITIAN SQUARES AND POSITIVE CONES ON ALGEBRAS WITH INVOLUTION

VINCENT ASTIER AND THOMAS UNGER

ABSTRACT. This expository article is a synopsis of a number of talks given by us in Paris (Astier) and Konstanz (Unger) in 2015 and 2016 and aims to provide a coherent picture of our recent efforts in extending the real theory of quadratic forms over fields in the noncommutative direction towards hermitian forms over algebras with involution.

## 1. INTRODUCTION

Hilbert, in his famous Paris talk in 1900 posed the following as his 17th problem [?]: *ob nicht jede definite Form als Quotient von Summen von Formenquadraten dargestellt werden kann.* Translated: is every nonnegative $n$-ary polynomial over a field $F$ a sum of squares of rational functions over $F$? Partial answers were known to Hilbert, as well as the fact that a positive semidefinite polynomial need not be a sum of squares of polynomials.

The full, affirmative, answer was obtained by Artin in 1927, and built on joint work with Schreier, published in the same year. To be precise: let $F$ be a field of characteristic different from 2 with space of orderings $X_F$ (these are the orderings on $F$ that are compatible with the arithmetic of $F$). The Artin-Schreier theorem [?] says that $F$ admits an ordering if and only if $-1$ is not a sum of squares in $F$. Artin's theorem [?] says that an element is positive at all $P \in X_F$ if and only if it is a sum of squares in $F$. This result is crucial in his solution of Hilbert's 17th problem.

Quadratic forms over $F$ come into the picture via the Witt ring $W(F)$ (the ring of Witt equivalence classes of non-degenerate quadratic forms over $F$). Orderings $P \in X_F$ correspond to signature homomorphisms $\mathrm{sign}_P : W(F) \to \mathbb{Z}$. This goes back to Sylvester's *Law of Inertia* [?]. In addition to the fundamental ideal of $W(F)$, the prime ideals of $W(F)$ are given by $\mathrm{sign}_P^{-1}(0)$ and $\mathrm{sign}_P^{-1}(p\mathbb{Z})$ (where $p$ is an odd prime) with $P \in X_F$ (Lorenz-Leicht [?]). Given $q \in W(F)$, the total signature map $\mathrm{sign}\, q : X_P \to \mathbb{Z}, P \mapsto \mathrm{sign}_P\, q$ is continuous with respect to the Harrison topology on $X_F$ and the discrete topology on $\mathbb{Z}$. Furthermore, the torsion part of $W(F)$ consists of those $q$ with $\mathrm{sign}\, q = 0$ (Pfister's local-global principle [?]).

The above results (well-documented in [?] and [?]), and some of their extensions to commutative rings, are among the foundations of real algebra, see for example [?] or Lam's expository paper [?]. In a series of recent papers [?], [?],

---

2010 *Mathematics Subject Classification.* 16K20, 11E39, 13J30, 16W10.

[?], [?], [?] (and also [?]) we extended these results in the noncommutative direction, more precisely to hermitian forms over central simple $F$-algebras with involution.

The theory of central simple algebras with involution was developed by Albert in the 1930s [?] and is still a topic of current research as testified by *The Book of Involutions* [?]; see also [?] and the copious references therein for a list of open problems in this area. A large part of present day research in algebras with involution is driven by the deep connections with linear algebraic groups, first observed by Weil [?]; see also Tignol's *2 ECM* exposition [?]. Some work has been done on algebras with involution over formally real fields, for example [?], [?], but this part of the theory is relatively underdeveloped. This observation, together with the fact that algebras with involution are a natural generalization of quadratic forms, are motivating factors for our research.

## 2. SIGNATURES

Let $(A, \sigma)$ be an *$F$-algebra with involution*, by which we mean that $A$ is a finite dimensional simple $F$-algebra with centre a field $K \supseteq F$ and $\sigma$ is an $F$-linear anti-automorphism of $A$ of order 2 (which implies that $[K : F] \leqslant 2$). Let $W(A, \sigma)$ denote the Witt group of $(A, \sigma)$, i.e. the $W(F)$-module of Witt equivalence classes of non-degenerate hermitian forms $h : M \times M \to A$, where $M$ is a finitely generated right $A$-module (cf. [?, Chap. I] or [?, Chap. 7]). We identify hermitian forms with their Witt class in $W(A, \sigma)$, unless indicated otherwise. Given an ordering $P \in X_F$ we wish to define a signature at $P$, i.e. a morphism of groups

$$W(A, \sigma) \to \mathbb{Z}.$$

Following the approach of [?] we do this by extending scalars to a real closure $F_P$ of $F$ at $P$ and realizing that, by Morita equivalence, the Witt group of any $F_P$-algebra with involution is isomorphic to either $\mathbb{Z}$, 0 or $\mathbb{Z}/2\mathbb{Z}$. In the last two cases, the only sensible definition is to take the signature at $P$ to be identically zero. In this case we call $P$ a *nil-ordering* and we write $\mathrm{Nil}[A, \sigma]$ for the set of all nil-orderings, noting that it only depends on the Brauer class of $A$ and the type of $\sigma$. Furthermore, $\mathrm{Nil}[A, \sigma]$ is clopen in $X_F$, cf. [?, Corollary 6.5].

In the first case, the Witt group $W(A \otimes_F F_P, \sigma \otimes \mathrm{id})$ is isomorphic to one of $W(F_P)$, $W((-1,-1)_{F_P}, -)$ or $W(F_P(\sqrt{-1}), -)$, where $-$ denotes (quaternion) conjugation, each one in turn being isomorphic to $\mathbb{Z}$ via the usual Sylvester signature of quadratic or hermitian forms. The composite map $s_P$, given by

$$W(A, \sigma) \longrightarrow W(A \otimes_F F_P, \sigma \otimes \mathrm{id}) \longrightarrow \mathbb{Z}\ ,$$

enables us to define a signature. The map $s_P$ is independent of the choice of $F_P$ [?, Prop. 3.3], but a different choice of Morita equivalence may result in a sign change [?, Prop. 3.4] and, conversely, such a sign change can always be obtained by taking a well-chosen different Morita equivalence.

At first sight, one way to fix a sign would be to demand that $s_P(\langle 1 \rangle_\sigma)$ is positive, as is the case for quadratic forms. This is the approach taken in [?], but it may not always work, since it may happen that $s_P(\langle 1 \rangle_\sigma)$ is in fact 0, as illustrated in [?, Rem. 3.11 and Ex. 3.12]. Our solution to this dilemma is to show that there exists a hermitian form $\eta$ over $(A,\sigma)$, called a *reference form*, such that $s_P(\eta)$ is always nonzero whenever $P \in \widetilde{X}_F := X_F \setminus \mathrm{Nil}[A,\sigma]$, cf. [?, Prop. 3.2]. Using this, given $P \in \widetilde{X}_F$, we define the *signature at $P$ with respect to the reference form $\eta$*,

$$\mathrm{sign}_P^\eta : W(A,\sigma) \to \mathbb{Z},$$

to be the map $s_P$, multiplied by $-1$ in case $s_P(\eta) < 0$, so that $\mathrm{sign}_P^\eta(\eta) > 0$.

The map $\mathrm{sign}_P^\eta$ does not depend on the Morita equivalence used in its computation and so we may use the explicit Morita equivalence presented in [?] in all practical situations.

**Remark 2.1.** In case $(A,\sigma) = (F, \mathrm{id}_F)$, we may take $\eta = \langle 1 \rangle$ and $\mathrm{sign}_P^\eta$ is then the usual Sylvester signature $\mathrm{sign}_P$ of quadratic forms.

**Remark 2.2.** The signature map is defined for all hermitian forms over $(A,\sigma)$, not just the non-degenerate ones as the notation above (which makes use of $W(A,\sigma)$) might suggest. It suffices to replace a form by its non-degenerate part, or alternatively, to replace $W(A,\sigma)$ by $\mathfrak{Herm}(A,\sigma)$, the category of hermitian forms over $(A,\sigma)$.

**Remark 2.3.** A reference *tuple* of hermitian forms of dimension one can be used instead of the reference form, cf. [?, Thm. 6.4] and [?, §3]. In fact, this is the approach used in [?].

We collect some immediate properties of the signature map:

**Proposition 2.4** (Properties of the signature map [?, Thm 2.6]).

1. *Let $h$ be a hyperbolic form over $(A,\sigma)$, then* $\mathrm{sign}_P^\eta h = 0$.
2. *Let $h_1, h_2 \in W(A,\sigma)$, then* $\mathrm{sign}_P^\eta(h_1 \perp h_2) = \mathrm{sign}_P^\eta h_1 + \mathrm{sign}_P^\eta h_2$.
3. *Let $h \in W(A,\sigma)$ and $q \in W(F)$, then* $\mathrm{sign}_P^\eta(q \cdot h) = \mathrm{sign}_P q \cdot \mathrm{sign}_P^\eta h$.
4. *(Going-up) Let $h \in W(A,\sigma)$ and let $L/F$ be an algebraic extension of ordered fields. Then*
$$\mathrm{sign}_Q^{\eta\otimes L}(h \otimes L) = \mathrm{sign}_{Q\cap F}^\eta h$$
*for all $Q \in X_L$.*

Property (4) is complemented by the following *going-down* result:

**Theorem 2.5** (Knebusch trace formula [?, Thm 8.1]). *Let $L/F$ be a finite extension of ordered fields and assume $P \in X_F$ extends to $L$. Let $h \in W(A\otimes_F L, \sigma \otimes \mathrm{id})$. Then*

$$\mathrm{sign}_P^\eta(\mathrm{Tr}^*_{A\otimes_F L} h) = \sum_{P\subseteq Q\in X_L} \mathrm{sign}_Q^{\eta\otimes L} h,$$

*where $\mathrm{Tr}^*_{A\otimes_F L} h$ denotes the Scharlau transfer induced by the $A$-linear homomorphism $\mathrm{id}_A \otimes \mathrm{Tr}_{L/F} : A \otimes_F L \to A$.*

**Theorem 2.6** (Preservation under Morita equivalence [?, Thm 4.2]). *Let $(B,\tau)$ be an $F$-algebra with involution, Morita equivalent to $(A,\sigma)$, and assume that $\sigma$ and $\tau$ are of the same type. Let $\zeta : W(A,\sigma) \xrightarrow{\sim} W(B,\tau)$ be the induced isomorphism of Witt groups. Then*

$$\mathrm{sign}_P^\eta h = \mathrm{sign}_P^{\zeta(\eta)} \zeta(h)$$

*for all* $h \in W(A,\sigma)$ *and all* $P \in X_F$.

**Theorem 2.7** (Pfister's local-global principle [?, Thm 4.1]). *Let* $h \in W(A,\sigma)$. *Then* $h$ *is a torsion form if and only if* $\operatorname{sign}^\eta_P h = 0$ *for all* $P \in X_F$.

**Theorem 2.8** (Continuity of the total signature [?, Thm 7.2]). *Let* $h \in W(A,\sigma)$. *The total signature of* $h$, $\operatorname{sign}^\eta h : X_F \to \mathbb{Z}, P \mapsto \operatorname{sign}^\eta_P h$, *is continuous.*

These two theorems motivate the following results, familiar from the quadratic forms case. Let $C(X_F, \mathbb{Z})_{[A,\sigma]}$ denote the ring of continuous functions from $X_F$ to $\mathbb{Z}$, that are zero on $\operatorname{Nil}[A,\sigma]$.

**Theorem 2.9** ([?, Prop. 4.3]). *For every* $f \in C(X_F, \mathbb{Z})_{[A,\sigma]}$ *there exists* $n \in \mathbb{N}$ *such that* $2^n f \in \operatorname{Im}\operatorname{sign}^\eta$. *In other words, the cokernel of* $\operatorname{sign}^\eta$ *is a* 2*-primary torsion group.*

The *stability index* of $(A,\sigma)$ is the smallest $k \in \mathbb{N}$ such that $2^k C(X_F,\mathbb{Z})_{[A,\sigma]} \subseteq \operatorname{Im}\operatorname{sign}^\eta$ if such a $k$ exists and $\infty$ otherwise. It is independent of the choice of $\eta$. The group $\operatorname{coker}\operatorname{sign}^\eta$ is up to isomorphism independent of the choice of $\eta$. We denote it by $S_\eta(A,\sigma)$ and call it the *stability group* of $(A,\sigma)$.

**Theorem 2.10** ([?, Thm. 4.10]). *Let* $W_t(A,\sigma)$ *denote the torsion subgroup of* $W(A,\sigma)$. *The sequence*

$$\begin{tikzcd} 0 \arrow[r] & W_t(A,\sigma) \arrow[r] & W(A,\sigma) \arrow[r, "\operatorname{sign}^\eta"] & C(X_F,\mathbb{Z})_{[A,\sigma]} \arrow[r] & S_\eta(A,\sigma) \arrow[r] & 0 \end{tikzcd}$$

*is exact. The groups* $W_t(A,\sigma)$ *and* $S_\eta(A,\sigma)$ *are* 2*-primary torsion groups.*

## 3. IDEALS AND MORPHISMS

Let $R$ be a commutative ring and let $M$ be an $R$-module. We introduce ideals of $R$-modules as follows: An *ideal* of $M$ is a pair $(I,N)$ where $I$ is an ideal of $R$ and $N$ is a submodule of $M$ such that $I \cdot M \subseteq N$. An ideal $(I,N)$ of $M$ is *prime* if $I$ is a prime ideal of $R$ (we assume that all prime ideals are proper), $N$ is a proper submodule of $M$, and for every $r \in R$ and $m \in M$, $r \cdot m \in N$ implies that $r \in I$ or $m \in N$.

These definitions are in part motivated by the following natural example: The pair $(\ker\operatorname{sign}_P, \ker\operatorname{sign}^\eta_P)$ is a prime ideal of the $W(F)$-module $W(A,\sigma)$ whenever $P \in \widetilde{X}_F$ .

We obtain a classification à la Harrison and Lorenz-Leicht [?]:

**Theorem 3.1** ([?, Props. 6.5, 6.7]). *Let* $(I,N)$ *be a prime ideal of the* $W(F)$*-module* $W(A,\sigma)$.

(1) *If* $2 \notin I$, *then one of the following holds:*
   (*i*) *There exists* $P \in \widetilde{X}_F$ *such that* $(I,N) = (\ker\operatorname{sign}_P, \ker\operatorname{sign}^\eta_P)$.
   (*ii*) *There exist* $P \in \widetilde{X}_F$ *and a prime* $p > 2$ *such that* $(I,N) = \big(\ker(\pi_p \circ \operatorname{sign}_P), \ker(\pi \circ \operatorname{sign}^\eta_P)\big)$, *where* $\pi_p : \mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}$ *and* $\pi : \operatorname{Im}\operatorname{sign}^\eta_P \to \operatorname{Im}\operatorname{sign}^\eta_P / (p \cdot \operatorname{Im}\operatorname{sign}^\eta_P)$ *are the canonical projections.*

(2) *If* $2 \in I$, *then* $I = I(F)$, *the fundamental ideal of* $W(F)$. *Furthermore, a pair* $(I(F),N)$ *is a prime ideal of* $W(A,\sigma)$ *if and only if* $N$ *is a proper submodule of* $W(A,\sigma)$ *with* $I(F) \cdot W(A,\sigma) \subseteq N$.

**Remark 3.2.** When $2 \notin I$, $N$ is completely determined by $I$. This is however not the case when $2 \in I$, cf. [?, Ex. 6.8].

There is a notion of morphism linked in the usual way to the above notion of ideal, cf. [?, Lemmas 5.6, 5.7 and 5.8]: Let $R$ and $S$ be commutative rings, let $M$ be an $R$-module and $N$ an $S$-module. We say that a pair $(\varphi, \psi)$ is an *$(R,S)$-morphism (of modules) from $M$ to $N$* if

(1) $\varphi : R \to S$ is a morphism of rings (and in particular $\varphi(1) = 1$);
(2) $\psi : M \to N$ is a morphism of additive groups;
(3) for every $r \in R$ and $m \in M$, $\psi(r \cdot m) = \varphi(r) \cdot \psi(m)$.

We call an $(R,S)$-morphism $(\varphi, \psi)$ *trivial* if $\psi = 0$. We denote the set of all $(R,S)$-morphisms from $M$ to $N$ by $\mathrm{Hom}_{(R,S)}(M,N)$ and its subset of non-trivial $(R,S)$-morphisms by $\mathrm{Hom}^*_{(R,S)}(M,N)$.

Let $(\varphi, \psi_1)$ and $(\varphi, \psi_2)$ be $(R,S)$-morphisms of modules from $M$ to $N_1$ and $N_2$ respectively. We say that $(\varphi, \psi_1)$ and $(\varphi, \psi_2)$ are *equivalent* if there is an isomorphism of $\mathrm{Im}\,\varphi$-modules $\vartheta : \mathrm{Im}\,\psi_1 \to \mathrm{Im}\,\psi_2$ such that $\psi_2 = \vartheta \circ \psi_1$. We write $\sim$ for the relation "being equivalent".

The pair $(\mathrm{sign}_P, \mathrm{sign}^\eta_P)$ is again a natural example of a $(W(F), \mathbb{Z})$-morphism from $W(A,\sigma)$ to $\mathbb{Z}$ and is trivial if and only if $P \in \mathrm{Nil}[A,\sigma]$.

The classification of prime ideals of $W(A,\sigma)$ yields the following description of signatures as morphisms:

**Theorem 3.3** ([?, Prop. 7.4]). *The map that sends $P \in \widetilde{X}_F$ to the pair* $(\mathrm{sign}_P, \mathrm{sign}^\eta_P)$ *induces a bijection between $\widetilde{X}_F$ and the equivalence classes with respect to $\sim$ of* $\mathrm{Hom}^*_{(W(F),\mathbb{Z})}(W(A,\sigma), \mathbb{Z})$.

Theorems ?? and ?? give us

**Corollary 3.4.** *There is a bijective correspondence between the prime ideals of $W(A,\sigma)$, the non-zero signatures of hermitian forms over $(A,\sigma)$ and the equivalence classes of* $\mathrm{Hom}^*_{(W(F),\mathbb{Z})}(W(A,\sigma), \mathbb{Z})$ *with respect to $\sim$.*

## 4. SUMS OF HERMITIAN SQUARES

In the field case, Pfister's local-global principle can be used to give a short proof of the fact that sums of squares are exactly the elements that are non-negative at every ordering. In [?] we showed that the same approach yields a similar result for $F$-division algebras with involution and, with some extra effort, for all $F$-algebras with involution.

Let $A^\times$ denote the set of invertible elements of $A$, $\mathrm{Sym}(A,\sigma)$ the set of $\sigma$-symmetric elements of $A$ and $\mathrm{Sym}(A,\sigma)^\times := \mathrm{Sym}(A,\sigma) \cap A^\times$. We say that an element $a \in \mathrm{Sym}(A,\sigma)$ is *$\eta$-maximal at an ordering $P \in X_F$* if $\mathrm{sign}^\eta_P \langle a \rangle_\sigma$ is maximal among all $\mathrm{sign}^\eta_P \langle b \rangle_\sigma$ for $b \in \mathrm{Sym}(A,\sigma)$. In the field case, this means $\mathrm{sign}_P \langle a \rangle = 1$, in other words $a \in P \setminus \{0\}$. For elements $b_1, \ldots, b_t \in F^\times$ we denote the *Harrison set* $\{P \in X_F \mid b_1, \ldots, b_t \in P\}$ by $H(b_1, \ldots, b_t)$.

**Theorem 4.1** ([?, Thm. 3.6]). *Let $b_1, \ldots, b_t \in F^\times$ and $Y = H(b_1, \ldots, b_t)$. Assume that $a \in \mathrm{Sym}(A,\sigma)^\times$ is $\eta$-maximal at all $P \in Y$. Let $u \in \mathrm{Sym}(A,\sigma)$. The following statements are equivalent:*

(*i*) *$u$ is $\eta$-maximal at all $P \in Y$.*
(*ii*) *$u \in D_{(A,\sigma)}(2^k \times \langle\langle b_1, \dots, b_t\rangle\rangle \langle a\rangle_\sigma)$ for some $k \in \mathbb{N}$.*

The presence of the element $a$ as well as the hypothesis on $\eta$-maximality correspond in the field case to the fact that 1 belongs to every ordering. Here 1 does not play a particular role since it may not have maximal signature at some orderings. We replace it by the element $a$ and only consider a set of orderings $Y$ on which $a$ has maximal signature.

This result is reminiscent of Procesi and Schacher's theorem [?, Thm. 5.4] and can also be used to address the question they raised in [?, p. 404], of whether or not the following property is always true:

**(PS):** for every $u \in \mathrm{Sym}(A, \sigma)$, the form $T_{(A,\sigma,u)}$ is positive semidefinite at all $P \in X_\sigma$ if and only if $u \in D_{(A,\sigma)}(2^s \times \langle 1\rangle_\sigma)$ for some $s \in \mathbb{N}$,

where $T_{(A,\sigma,u)}(x, y) := \mathrm{Trd}_A(\sigma(x)uy)$ for $x, y \in A$. The general answer to this question is negative as shown in [?], but we can now describe cases where the answer is positive, and also propose a natural reformulation (inspired by signatures of hermitian forms) of the question that has a positive answer. Following [?, Def. 1.1] we let

$$\begin{aligned} X_\sigma &:= \{P \in X_F \mid \sigma \text{ is positive at } P\} \\ &= \{P \in X_F \mid T_{(A,\sigma,1)} \text{ is positive semidefinite at } P\}. \end{aligned}$$

**Proposition 4.2** ([?, Cor. 4.17]). *If $X_\sigma = \widetilde{X}_F$, then property (PS) holds.*

And, if we introduce the property

**(PS'):** for every $u \in \mathrm{Sym}(A, \sigma)$, the form $T_{(A,\sigma,u)}$ is positive semidefinite at all $P \in \widetilde{X}_F$ if and only if $u \in D_{(A,\sigma)}(2^s \times \langle 1\rangle_\sigma)$ for some $s \in \mathbb{N}$,

we obtain

**Theorem 4.3** ([?, Thm. 4.18]). *Property (PS') holds if and only if $\widetilde{X}_F = X_\sigma$.*

## 5. POSITIVE CONES

The results presented thus far suggest that there could be a notion of "ordering" on central simple algebras with involution, whose behaviour would be similar to that of orderings on fields. The purpose of this final section is to present such a notion.

**Definition 5.1** ([?, Def. 3.1]). A *positive cone* $\mathscr{P}$ on $(A, \sigma)$ is a subset $\mathscr{P}$ of $\mathrm{Sym}(A, \sigma)$ such that

(P1) $\mathscr{P} \neq \varnothing$;
(P2) $\mathscr{P} + \mathscr{P} \subseteq \mathscr{P}$;
(P3) $\sigma(a) \cdot \mathscr{P} \cdot a \subseteq \mathscr{P}$ for every $a \in A$;
(P4) $\mathscr{P}_F := \{u \in F \mid u\mathscr{P} \subseteq \mathscr{P}\}$ is an ordering on $F$.
(P5) $\mathscr{P} \cap -\mathscr{P} = \{0\}$ (we say that $\mathscr{P}$ is *proper*).

We say that a positive cone $\mathscr{P}$ is *over* $P \in X_F$ if $\mathscr{P}_F = P$.

**Remark 5.2.** Axiom (P4) is necessary if we want our positive cones to consist of positive semidefinite (PSD) matrices with respect to $P$ only, or negative semidefinite (NSD) matrices with respect to $P$ only, in the case of $(M_n(F), t)$, see [?, Rem. 3.13].

If $\mathscr{P}$ is a positive cone, then $-\mathscr{P}$ is also a positive cone. This is due to the fact that positive cones are meant to contain elements of maximal signature, and the sign of the signature can vary with a change of the reference form.

It can also be shown that there is a positive cone over $P \in X_F$ on $(A, \sigma)$ if and only if $P \in \widetilde{X}_F$.

**Examples 5.3.**

(1) Let $P \in \widetilde{X}_F$. We define

$$\mathscr{M}_P^\eta(A,\sigma) := \{a \in \mathrm{Sym}(A,\sigma)^\times \mid a \text{ is } \eta\text{-maximal at } P\} \cup \{0\}.$$

If $A$ is a division algebra, $\mathscr{M}_P^\eta(A,\sigma)$ is a positive cone over $P$ on $(A,\sigma)$.

(2) The set of PSD matrices, and of NSD matrices with respect to some $P \in X_F$ are both positive cones over $P$ on $(M_n(F), t)$.

We obtain the desired results linking positive cones and $W(A,\sigma)$:

**Proposition 5.4** ([?, Prop. 7.11]). *The following statements are equivalent:*

*(i) $(A,\sigma)$ is formally real (i.e. there is at least one positive cone on $(A,\sigma)$;*
*(ii) $W(A,\sigma)$ is not torsion;*
*(iii) $\widetilde{X}_F \neq \varnothing$.*

Positive cones are well-behaved under Morita equivalence: If $(A,\sigma)$ and $(B,\tau)$ are Morita equivalent, then there is an inclusion-preserving bijection between their sets of positive cones. This bijection can be made explicit in the case of scaling, or when $(A,\sigma) = (D,\vartheta)$ is an $F$-division algebra with involution and $(B,\tau) = (M_n(D), \vartheta^t)$, using descriptions of positive cones reminiscent of the characterizations of positive semidefinite matrices, cf. [?, §4.1, §4.2].

The notion of positive cone can be seen as somewhat equivalent to that of preordering or pre-semiordering, so it is natural to consider in more detail the maximal positive cones. They can be completely described and match the examples provided above. To see this we define, for $P \in X_F$ and $S \subseteq \mathrm{Sym}(A,\sigma)$,

$$\mathscr{C}_P(S) := \Big\{\sum_{i=1}^{k} u_i \sigma(x_i) s_i x_i \;\Big|\; k \in \mathbb{N},\ u_i \in P,\ x_i \in A,\ s_i \in S\Big\}$$

(the smallest, possibly non-proper, positive cone over $P$ containing $S$), and we denote by $X_{(A,\sigma)}$ the set of all maximal positive cones on $(A,\sigma)$.

**Theorem 5.5** ([?, Thm. 7.5]).

$$X_{(A,\sigma)} = \{-\mathscr{C}_P(\mathscr{M}_P^\eta(A,\sigma)), \mathscr{C}_P(\mathscr{M}_P^\eta(A,\sigma)) \mid P \in \widetilde{X}_F\}.$$

*Moreover, for each $\mathscr{P} \in X_{(A,\sigma)}$, there exists $\varepsilon \in \{-1,1\}$ such that $\mathscr{P} \cap A^\times = \varepsilon \mathscr{M}_P^\eta(A,\sigma) \setminus \{0\}$.*

In particular, the only maximal positive cones over $P$ on $(D,\vartheta)$ are $\mathscr{M}_P^\eta(D,\vartheta)$ and $-\mathscr{M}_P^\eta(D,\vartheta)$ and therefore the examples above are essentially the only maximal positive cones on $(A,\sigma)$, cf. [?, Props. 4.3 and 4.6]. It follows that the PSD matrices over $P$ and the NSD matrices over $P$ are the only maximal positive cones over $P$ on $(M_n(F), t)$.

Using this description, it is possible to make the link with the results presented in Section ??, and to obtain results similar to the Artin-Schreier and Artin theorems.

**Theorem 5.6** ([?, Thm. 7.9]). *The following statements are equivalent:*

*(i)* *$(A, \sigma)$ is formally real;*
*(ii)* *There is $a \in \mathrm{Sym}(A, \sigma)^\times$ and $P \in X_F$ such that $\mathscr{C}_P(a) \cap -\mathscr{C}_P(a) = \{0\}$;*
*(iii)* *There is $b \in \mathrm{Sym}(A, \sigma)^\times$ such that $\langle b \rangle_\sigma$ is strongly anisotropic.*

The second statement is a trivial consequence of the first one, but it is still included here to point out that while the element $a$ in it obviously belongs to a positive cone (namely $\mathscr{C}_P(a)$), the element $b$ in the third statement may not belong to any positive cone on $(A, \sigma)$, contrary to what could be expected from the field case (see [?, Rem. 7.10]).

**Theorem 5.7** ([?, Thm. 7.14]). *Let $b_1, \ldots, b_t \in F^\times$, let $Y = H(b_1, \ldots, b_t)$ and let $a \in \mathrm{Sym}(A, \sigma)^\times$ be such that, for every $\mathscr{P} \in X_{(A,\sigma)}$ with $\mathscr{P}_F \in Y$, $a \in \mathscr{P} \cup -\mathscr{P}$. Then*

$$\bigcap \{\mathscr{P} \in X_{(A,\sigma)} \mid \mathscr{P}_F \in Y \text{ and } a \in \mathscr{P}\} = \bigcup_{s \in \mathbb{N}} D_{(A,\sigma)}(2^s \times \langle\langle b_1, \ldots, b_t \rangle\rangle \langle a \rangle_\sigma).$$

The element $a$ in this theorem plays the same role as the element $a$ in Theorem ??, and chooses a positive cone from $\{\mathscr{P}, -\mathscr{P}\}$ in a uniform way. This is not necessary in the field case, because 1 belongs to every ordering. In the special case where $a = 1$ can be used for this purpose, we obtain a result more similar to the usual one:

**Corollary 5.8** ([?, Cor. 7.15]). *Assume that for every $\mathscr{P} \in X_{(A,\sigma)}$, $1 \in \mathscr{P} \cup -\mathscr{P}$. Then*

$$\bigcap \{\mathscr{P} \in X_{(A,\sigma)} \mid 1 \in \mathscr{P}\} = \Big\{ \sum_{i=1}^{s} \sigma(x_i) x_i \,\Big|\, s \in \mathbb{N},\ x_i \in A \Big\}.$$

The hypothesis of Corollary ?? is exactly $X_\sigma = \widetilde{X}_F$ in the terminology of Section ??. More precisely, as seen therein, this property characterizes the algebras with involution for which there is a positive answer to (PS'), cf. [?, Section 4.2].

We finish this survey by a presentation of our main result concerning the topology of $X_{(A,\sigma)}$. We define, for $a_1, \ldots, a_k \in \mathrm{Sym}(A, \sigma)$,

$$H_\sigma(a_1, \ldots, a_k) := \{\mathscr{P} \in X_{(A,\sigma)} \mid a_1, \ldots, a_k \in \mathscr{P}\}.$$

We denote by $\mathcal{T}_\sigma$ the topology on $X_{(A,\sigma)}$ generated by the sets $H_\sigma(a_1, \ldots, a_k)$, for $a_1, \ldots, a_k \in \mathrm{Sym}(A, \sigma)$, and by $\mathcal{T}_\sigma^\times$ the topology on $X_{(A,\sigma)}$ generated by the sets $H_\sigma(a_1, \ldots, a_k)$, for $a_1, \ldots, a_k \in \mathrm{Sym}(A, \sigma)^\times$.

**Proposition 5.9** ([?, Prop. 8.2]). *The topologies $\mathcal{T}_\sigma$ and $\mathcal{T}_\sigma^\times$ are equal.*

**Theorem 5.10** ([?, Prop. 8.17]). *$\mathcal{T}_\sigma$ is a spectral topology on $X_{(A,\sigma)}$.*

Recall that spectral topologies are precisely the topologies of the spectra of commutative rings and that they can be characterized in topological terms, cf. [?].

The topology $\mathcal{T}_\sigma$ is also well-behaved under Morita equivalence:

**Proposition 5.11** ([?, Prop. 8.18]). *Let* $(A, \sigma)$ *and* $(B, \tau)$ *be two Morita equivalent* $F$*-algebras with involution. The spaces* $(X_{(A,\sigma)}, \mathcal{T}_\sigma)$ *and* $(X_{(B,\tau)}, \mathcal{T}_\tau)$ *are homeomorphic.*

## REFERENCES

[1] A. A. Albert, *Involutorial simple algebras and real Riemann matrices*, Ann. of Math. (2) **36** (1935), no. 4, 886–964.

[2] E. Artin, *Über die Zerlegung definiter Funktionen in Quadrate*, Abh. Math. Sem. Univ. Hamburg **5** (1927), no. 1, 100–115.

[3] E. Artin and O. Schreier, *Algebraische Konstruktion reeller Körper*, Abh. Math. Sem. Univ. Hamburg **5** (1927), no. 1, 85–99.

[4] V. Astier and T. Unger, *Signatures of hermitian forms and the Knebusch trace formula*, Math. Ann. **358** (2014), no. 3-4, 925–947.

[5] V. Astier and T. Unger, *Signatures of hermitian forms and "prime ideals" of Witt groups*, Adv. Math. **285** (2015), 497–514.

[6] V. Astier and T. Unger, *Positive cones on algebras with involution*, `http://arxiv.org/abs/1609.06601` (2016).

[7] V. Astier and T. Unger, *Signatures of hermitian forms, positivity, and an answer to a question of Procesi and Schacher*, `http://arxiv.org/abs/1511.06330` (2016).

[8] V. Astier and T. Unger, *Stability index of algebras with involution*, Contemporary Mathematics, American Mathematical Society, 2017 (to appear).

[9] A. Auel, E. Brussel, S. Garibaldi, and U. Vishne, *Open problems on central simple algebras*, Transform. Groups **16** (2011), no. 1, 219–264.

[10] E. Bayer-Fluckiger and R. Parimala, *Classical groups and the Hasse principle*, Ann. of Math. (2) **147** (1998), no. 3, 651–693.

[11] J. Bochnak, M. Coste, and M.-F. Roy, *Real algebraic geometry*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3), vol. 36, Springer-Verlag, Berlin, 1998.

[12] D. Hilbert, *Mathematische Probleme*, Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse **3** (1900), 253–297.

[13] M. Hochster, *Prime ideal structure in commutative rings*, Trans. Amer. Math. Soc. **142** (1969), 43–60.

[14] I. Klep and T. Unger, *The Procesi-Schacher conjecture and Hilbert's 17th problem for algebras with involution*, J. Algebra **324** (2010), no. 2, 256–268.

[15] M.-A. Knus, *Quadratic and Hermitian forms over rings*, Grundlehren der Mathematischen Wissenschaften, vol. 294, Springer-Verlag, Berlin, 1991.

[16] M.-A. Knus, A. Merkurjev, M. Rost, and J.-P. Tignol, *The book of involutions*, American Mathematical Society Colloquium Publications, vol. 44, American Mathematical Society, Providence, RI, 1998.

[17] T. Y. Lam, *An introduction to real algebra*, Rocky Mountain J. Math. **14** (1984), no. 4, 767–814, Ordered fields and real algebraic geometry (Boulder, Colo., 1983).

[18] T. Y. Lam, *Introduction to quadratic forms over fields*, Graduate Studies in Mathematics, vol. 67, American Mathematical Society, Providence, RI, 2005.

[19] D. W. Lewis and J.-P. Tignol, *On the signature of an involution*, Arch. Math. (Basel) **60** (1993), no. 2, 128–135.

[20] D. W. Lewis and T. Unger, *A local-global principle for algebras with involution and Hermitian forms*, Math. Z. **244** (2003), no. 3, 469–477.

[21] D. W. Lewis and T. Unger, *Hermitian Morita theory: a matrix approach*, Irish Math. Soc. Bull. (2008), no. 62, 37–41.

[22] F. Lorenz and J. Leicht, *Die Primideale des Wittschen Ringes*, Invent. Math. **10** (1970), 82–88.

[23] A. Pfister, *Quadratische Formen in beliebigen Körpern*, Invent. Math. **1** (1966), 116–132.

[24] C. Procesi and M. Schacher, *A non-commutative real Nullstellensatz and Hilbert's 17th problem*, Ann. of Math. (2) **104** (1976), no. 3, 395–406.

[25] A. Quéguiner, *Signature des involutions de deuxième espèce*, Arch. Math. (Basel) **65** (1995), no. 5, 408–412.

[26] W. Scharlau, *Quadratic and Hermitian forms*, Grundlehren der Mathematischen Wissenschaften, vol. 270, Springer-Verlag, Berlin, 1985.

[27] J.J. Sylvester, *A demonstration of the theorem that every homogeneous quadratic polynomial is reducible by real orthogonal substitutions to the form of a sum of positive and negative squares*, Philosophical Magazine, IV (1852), 138–142.

[28] J.-P. Tignol, *Algebras with involution and classical groups*, European Congress of Mathematics, Vol. II (Budapest, 1996), Progr. Math., vol. 169, Birkhäuser, Basel, 1998, pp. 244–258.

[29] A. Weil, *Algebras with involutions and the classical groups*, J. Indian Math. Soc. (N.S.) **24** (1960), 589–623 (1961).

SCHOOL OF MATHEMATICS AND STATISTICS, UNIVERSITY COLLEGE DUBLIN, BELFIELD, DUBLIN 4, IRELAND

*E-mail address*: `vincent.astier@ucd.ie, thomas.unger@ucd.ie`

# On Valuation Fans and the Real Holomorphy Ring
## *Dedicated to Professor Francisco Miraglia*

Danielle Gondard-Cozette

**Abstract.** In this survey we recall basic facts in Real Algebra, and study valuation fans, a refined notion of fans. We explain how valuation fans are involved in the theory of residually real-closed henselian fields. Then we deal with other mathematical notions, such as $\mathbb{R}-places$ and the real holomorphy ring which are also linked to valuation fans. Finally we consider abstract spaces of orderings and Marshall's problem, recall some results, and present the complete real spectrum of a ring.

## 1. Background in Real Algebra.

### 1.1. Preorderings, orderings.

**Definition 1.1.** A preordering $T$ of $K$ is a subset $T \subseteq K$, satisfying:

$$\begin{aligned} T+T &\subseteq T,\ T\cdot T\subseteq T,\ 0,1\in T,\ \ -1\notin T \\ \text{and } T^* &= T\backslash\{0\} \text{ is a subgroup of } K^*=K\backslash\{0\}. \end{aligned}$$

**Definition 1.2.** A preordering $T$ is called a quadratic preordering if $K^2 \subseteq T$. If $K^{2n} \subseteq T$, $T$ is said to be of level n. Preorderings with no level do exist.

Zorn's lemma shows the existence of maximal quadratic preorderings; these are just the usual orderings, and are characterized by:

**Definition 1.3.** A subset $P$ of $K$ is an ordering if:

$$P+P\subseteq P,\ P\cdot P\subseteq P,\ P\cup -P=K,\ \ -1\notin P.$$

From these properties one can deduce that $0,1 \in P$, $P\cap -P=\{0\}$ and $\sum K^2 \subseteq P$. Here, and throughout the paper, $\sum K^{2n}$ denotes the set of all finite sums of $2n$-$th$ powers.

We can also call $P$ a positive cone: to any such ordering $P$ one can associate a binary relation $\leq_P$. This is a total order relation compatible with the field structure, defined as follows:

$$b-a\in P \Leftrightarrow a\leq_P b.$$

Then $P$ is the set of elements positive for the order relation $\leq_P$ .

The set of orderings of a field $K$ will be denoted by $\chi(K)$; it is also denoted in the literature by $SperK$ (so as to coincide with the usual notation for rings).

The field $\mathbb{R}$ admits only one ordering, and its set of positive elements is $\mathbb{R}^2$.

The field $\mathbb{Q}(\sqrt[2]{2}) := \{a + b\sqrt[2]{2} \mid a, b \in \mathbb{Q}\}$ admits two orderings, one making $\sqrt[2]{2}$ positive and the other making $\sqrt[2]{2}$ negative.

$\mathbb{R}((X))$, the power series field, admits also two orderings making X infinitesimal positive or negative.

$\mathbb{R}(X)$ admits infinitely many orderings. For any $a \in \mathbb{R}$ one can define orderings $P_{a,+}$ and $P_{a,-}$ making $X - a$ infinitesimal positive or negative respectively. $\mathbb{R}(X)$ admits also the orderings $P_+$ and $P_-$ making $\frac{1}{X}$ infinitesimal positive or negative respectively.

### 1.2. Real Valuations.

**Definition 1.4.** A Krull valuation $v$ on a field $K$ is a surjective map

$$v : K^* \twoheadrightarrow \Gamma$$

where $\Gamma$ is a totally ordered abelian group (called the value group), such that
(1) $v(xy) = v(x) + v(y)$ for any $x, y$ in $K^*$;
(2) $v(x + y) \geq \min\{v(x), v(y)\}$, for any $x, y$ in $K^*$, with $x + y$ in $K^*$.

The *valuation ring* of $v$ is

$$A_v := \{x \in K \mid x = 0 \text{ or } v(x) \geq 0\}$$

and its *maximal ideal* is

$$I_v := \{x \in K \mid x = 0 \text{ or } v(x) > 0\}.$$

$k_v := A_v / I_v$ is called the *residue field* of the valuation.
$U_v := A_v \setminus I_v$ denotes the *group of units.*

**Definition 1.5.** A valuation $v$ on a field $K$ is said to be real if and only if the residue field $k_v$ is real (meaning $-1 \notin \sum k_v^2$).

A field admits real valuations if and only if it is real. Of course a real field admits real valuations, at least the trivial one.

The converse implication follows from the Baer-Krull theorem which ensures that if $k_v$ admits an ordering, then $K$ admits also at least one ordering.

Given an ordering $P$ in a field $K$, the convex hull of $\mathbb{Q}$ in $K$ is:

$$A(P) := \{x \in K \mid \exists r \in \mathbb{Q} \quad r \pm x \in P\}.$$

$A(P)$ is a valuation ring in $K$ with unique maximal ideal:

$$I(P) := \{x \in K \mid \forall r \in \mathbb{Q}^{+*} \quad r \pm x \in P\}.$$

where $\mathbb{Q}^{+*} = \{r \in \mathbb{Q} \mid r > 0\}$.

$A(P)$ is clearly a subring of $K$; it is a valuation ring because $b \notin A(P)$ implies $b^{-1} \in A(P)$: let $b \notin A(P)$, assume $b > 0$, since $b \notin A(P)$ we have in particular $1 < b$, therefore $0 < b^{-1} < 1$ which implies that $b^{-1} \in A(P)$ because $A(P)$ is convex in $K$ with respect to $P$.

We will see below that the valuation associated to $A(P)$ is compatible with the ordering $P$ and pushes down on the residue field an (archimedean) ordering, hence this valuation is real.

### 1.3. Compatibility of an ordering with a valuation.

**Definition 1.6.** An ordering $P$ is said to be compatible with a valuation $v$ if and only if $1 + I_v \subset P$.

Then $\overline{P}$, induced by $P$ on the residue field $k_v$, is an ordering of $k_v$. Clearly $\overline{P}$ is closed under addition and multiplication and $\overline{P} \cup -\overline{P} = k_v$. If $-1$ was in $\overline{P}$ we would have $-1 = \overline{a}$ for some $a \in P \cap A(P)$. Then $1 + a \in I(P)$, hence $-a \in 1 + I(P) \subset P$, so we would get $a = 0$ which is impossible.

The trivial valuation, sending every non-zero element of $K$ to 0, is compatible with any ordering of $K$.

Note that the valuation $v$ associated to an ordering $P$ of $K$ with valuation ring

$$A(P) := \{x \in K \mid \exists r \in \mathbb{Q} \;\; r \pm x \in P\}$$

is compatible with $P$. In fact $I(P) := \{x \in K \mid \forall r \in \mathbb{Q}^{+*} \;\; r \pm x \in P\}$ being the maximal ideal of $A(P)$ we have $1 + I(P) \subset P$. Hence the valuation is compatible with $P$. Then $\overline{P}$ induced by $P$ on the residue field $k_v$ is an archimedean ordering; we already know that $\overline{P}$ is an ordering, this ordering $\overline{P}$ is archimedean: for any $x \in A(P)$ there exists some $r \in \mathbb{Z}$ such that $-r <_P x <_P r$, hence in the residue field we have $-r <_{\overline{P}} \overline{x} <_{\overline{P}} r$, and therefore $\overline{P}$ is an archimedean ordering of $k_v$.

**Theorem 1.7.** *Let $P$ be an ordering of $K$, and $v$ be a valuation on $K$; the following are equivalent:*

*(1) $0 <_P a \leq_P b \Rightarrow v(a) \geq v(b)$ in $\Gamma$ (the value group of $v$).*

*(2) The valuation ring $A_v$ is convex in $K$ with respect to $P$.*

*(3) The maximal ideal $I_v$ of $A_v$ is convex in $K$ with respect to $P$ .*

*(4) $v$ is compatible with $P$ (i.e. $1 + I_v \subset P$).*

**Proof.** (1) $\Rightarrow$ (2) $A_v$ convex in $K$ means that if $x <_P y <_P z$, with $x, z \in A_v$ then $y \in A_v$, or equivalently $0 <_P a <_P b$ with $b \in A_v$ implies $a \in A_v$.

From (1) we deduce that $v(a) \geq v(b) \geq 0$ in $\Gamma$ hence $a \in A_v$.

(2) $\Rightarrow$ (3) Assume $0 <_P a <_P b$ with $b \in I_v$ then $0 <_P b^{-1} <_P a^{-1}$. Since $b^{-1} \notin A_v$ using (2) we deduce $a^{-1} \notin A_v$, hence $a \in I_v$, $I_v$ being the ideal of non invertible elements of $A_v$.

(3) $\Rightarrow$ (4) Let $m \in I_v$, if $1 + m \notin P$ then $1 + m \in -P$, so $1 + m <_P 0$ hence $0 <_P 1 <_P -m$. Using the convexity of $I_v$ in $K$ for $P$, since $-m \in I_v$ too, this yields $1 \in I_v$ which is impossible.

(4) $\Rightarrow$ (1) Assume $0 <_P a \leq_P b$ and $v(a) < v(b)$ in $\Gamma$ ; then we deduce $0 < v(b) - v(a) = v(\frac{b}{a})$, hence $\frac{b}{a} \in I_v$, and also $-\frac{b}{a} \in I_v$ and $a \neq b$. From (4) we get $1 + (-\frac{b}{a}) \in P$, so $\frac{a-b}{a} >_P 0$, hence $a >_P b$ which is impossible.

**Theorem 1.8.** *Let $\mathcal{F}$ be the family of all valuation rings of $K$ compatible with a given ordering $P$, then:*

*(1) the valuation rings in $\mathcal{F}$ form a chain under inclusion;*

*(2) the smallest element of $\mathcal{F}$ is $A(P)$.*

**Proof.** (1) Suppose $A, B \in \mathcal{F}$ and $A \nsubseteq B$, let $a \in A \backslash B$ and $a > 0$. We prove that $B \subset A$. Consider $0 < b \in B$, by the convexity of $B$ in $K$ we cannot have $0 < a \leq b$, so we must have $0 < b \leq a$. From the convexity of $A$ in $K$, we deduce $b \in A$.

(2) Let $A \in \mathcal{F}$, $A$ is convex in $K$ and contains $\mathbb{Z}$, hence $A$ contains $A(P)$ the convex hull of $\mathbb{Q}$ in $K$.

Note that any subring of $K$ containing a valuation ring must itself be a valuation ring, hence $\mathcal{F}$ consists of all subrings of $K$ containing $A(P)$. Remark also that $A \subset A'$ implies $I' \subset I$.

**Theorem 1.9.** *(Baer-Krull) Let $A$ be a real valuation ring of $K$, and let $v$ be the associated valuation. Let $\overline{P}$ be an ordering in the residue field $k_v$. Denote $\chi_{v,\overline{P}}$ the set of all orderings $P_i$ in $K$ inducing the given $\overline{P}$ in $k_v$. Then there is a bijection between $\chi_{v,\overline{P}}$ and* $\mathrm{Hom}(\Gamma, \mathbb{Z}/2)$ *where $\Gamma$ denotes the value group of $v$.*

For the proof we refer the reader to [5].

## 2. Fans (level 1 case).

In this section we mainly follow the notations and proofs of [48].

**2.1. Quadratic preorderings.** The compatibility of a quadratic preordering with a valuation can be of two types. Given $T$ a quadratic preordering in a real field $K$, $v$ a valuation on $K$ is *compatible* with $T$ if it is compatible with *some* ordering $P$ containing $T$.

$v$ is called *fully compatible* with $T$ if it is compatible with *every* ordering $P$ containing $T$. In this case $T$ induces on the residue field $k_v$ a quadratic preordering $\overline{T}$. This pushdown preordering $\overline{T}$ is defined to be the image of $T \cap A_v$ under the natural map from the valuation ring $A_v$ to the residue field $k_v$.

Below we give alternative characterizations.

**Definition 2.1.** Given $T$ a quadratic preordering in a real field $K$, and $v$ a valuation on $K$ with unique maximal ideal $I_v$ in the associated valuation ring $A_v$:

(1) $v$ is fully compatible with $T$ if and only if $1 + I_v \subset T$.

(2) $v$ is compatible with $T$ if and only if $(1 + I_v) \cap -T = \emptyset$.

(3) $v$ is compatible with $T$ if and only if $\overline{T}$ is a preordering in the residue field $k_v$.

We set $\chi_{/T} := \{P \textit{ ordering} \mid P \supset T\}$.

A way of building fully compatible preorderings is to use the "wedge product" introduced in 1978 by Becker in [6], and by Becker and Brocker in [10].

**Definition 2.2.** Let $K$ be a real field, let $A$ be a valuation ring in $K$, and $\pi : A \longrightarrow k_v$ be the projection map. Let $T$ be a preordering of $K$ and let $S$ be a preordering of $k_v$ such that $S \supset \overline{T}$. The wedge product is defined by $T \wedge S := T \cdot \pi^{-1}(S\backslash\{0\})$.

We refer the reader to Lam's book ([48], p.21) to verify that $T \wedge S$ is a preordering in $K$, fully compatible with $v$, and such that residually $\overline{T \wedge S} = S$.

Again refering to [48] (3.3 p.22), remark that the wedge product $T \wedge S$ can also be defined for $S$ a preordering of $k_v$ and $T = T^* \cup \{0\}$ where $T^*$is a subgroup of $K^*$. Then $T \wedge S$ is a preordering in $K$, and if $K^2 \subseteq T$ then $T \wedge S$ is a quadratic preordering.

There is also an alternative definition for the wedge product:

$T \wedge S = \cap \left\{\text{orderings } P \;\middle|\; P \supset T \text{ and } \overline{P} \in \chi_{/S}\right\}$

**2.2. Fans of level 1.** In the context of preorderings fans were first presented by Becker and Köpping in [15].

**Definition 2.3.** Let $K$ be a real field and let $T$ be a quadratic preordering in $K$. $T$ is a fan if and only if for any $S \supset T$, such that $-1 \notin S$ and such that $S^* = S\backslash\{0\}$ is a subgroup of $K^*$ satisfying $[K^* : S^*] = 2$, $S$ is an ordering in $K$.

Note that if $T$ is a fan any preordering containing $T$ is again a fan. There is an alternative useful characterization of a fan given in [48] (p.40), with proof of equivalence:

**Proposition 2.4.** *A preordering $T$ is a fan if and only if for any $a \in K^*\backslash -T$ we have $T + aT \subset T \cup aT$. Such an element $a$ is said to be $T$-rigid.*

First examples of fans are the trivial fans : these are orderings $P$ and intersection of two orderings $P_1 \cap P_2$.

Another example is the pullback $\widehat{S}$ of a trivial fan $S$ in $k_v$ . Namely $\widehat{S} = K^2 \wedge S = K^2 \cdot \pi^{-1}(S\backslash\{0\})$ is a fan in $K$. In fact Bröcker's trivialization theorem given later in 2.6 says that all fans arise in this way.

Fans are well behaved for compatibility with real valuations.

**Theorem 2.5.** *Let $K$ be a real field, $v$ a valuation on $K$, and $T$ a preordering in $K$ . Then the followings hold:*

*(a) If $v$ is compatible with $T$, $T$ is a fan implies that $\overline{T}$ is a fan in $k_v$;*

*(b) If $v$ is fully compatible with $T$, $T$ is a fan if and only if $\overline{T}$ is a fan in $k_v$.*

**Proof.**

(a) We use proposition 2.4 characterizing a fan. Let $b \in A\backslash I$ such that $\overline{b} \notin -\overline{T}$ we shall show that $\overline{b}$ is $\overline{T}$*-rigid*. $T$ being a fan let $t_1 + t_2 b \in T + bT \subset T \cup bT$ hence there exist $t_3$ or $t_4$ such that $t_1 + t_2 b = t_3$ or $t_1 + t_2 b = t_4 b$. Going down to $k_v$ we get $\overline{t_1} + \overline{t_2}\overline{b} = \overline{t_3}$ or $\overline{t_1} + \overline{t_2}\overline{b} = \overline{t_4}\overline{b}$ hence $\overline{t_1} + \overline{t_2}\overline{b} \in \overline{T} \cup \overline{bT}$, and $\overline{T}$ is a fan.

(b) We use the definition of a fan. Assume $v$ is fully compatible with $T$ and $\overline{T}$ is a fan we have to prove that $T$ is a fan. Let $W \supset T$ be such that $-1 \notin W$, $W^* = W\backslash\{0\}$ is a subgroup of $K^*$ and $[K^* : W^*] = 2$, we have to prove that $W$ is an ordering. We first show that $\overline{W}$ is an ordering. If $-1 = \overline{w}$ for some $w \in W \cap A$, then $-1 = w + m$ for some $m \in I$, so $-w = 1 + m \in 1 + I \subset T \subset W$ hence $-1 \in W$ which is impossible. Since $\overline{T}$ is a fan and $\overline{W^*}$ a subgroup of $k_v^*$ such that $[k_v^* : \overline{W^*}] = 2$, $\overline{W}$ is an ordering. Form the wedge product $W \wedge \overline{W} = W \cdot \pi^{-1}(\overline{W}\backslash\{0\}) = W \cdot (1 + I) \subset W \cdot T \subset W$, since from [48] (p.22) $W \cdot \pi^{-1}(\overline{W}\backslash\{0\}) = W \cdot (1 + I)$ ); then $W \wedge \overline{W} \subset W$ holds, hence $W = W \wedge \overline{W}$ is an ordering.

**2.3. Trivialization of fans.** A remarkable result is Bröcker's theorem on trivialization of fans ([20]).

**Theorem 2.6.** *Let $K$ be a real field and $T \subset K$ be a fan. Then there exists a valuation $v$, fully compatible with $T$, such that the pushdown $\overline{T}$ in the residue field $k_v$ is a trivial fan.*

The theorem follows from propositions 2.7 and 2.8 below. We use the proof given by Lam ([48], p. 94).

**Proposition 2.7.** *Let $T$ be a non-trivial fan in the field $K$. Then there exists a non-trivial valuation $v$ on $K$, fully compatible with $T$.*

The proof of proposition 2.7 requires three lemmas.

**Lemma 1.** *Let $G$ be an ordered group (written additively), and $H$ be a subgroup of $G$. If $H$ does not contain a non-trivial convex subgroup of $G$, then for any positive element $h \in H$ there exists $g \in G\backslash H$ such that $0 < g < h$.*

**Proof of lemma 1.** Let $C := \{g \in G \mid \exists n \in \mathbb{N} \quad -nh \leq g \leq nh\}$. $C$ is the convex hull of the subgroup of $G$ generated by $h$, hence a convex subgroup. Assume there does not exist an element $g$ as in the statement, then for any $g \in G$, $0 \leq g \leq h$ implies $g \in H$. By easy induction on $n$ it follows that for any $n \in \mathbb{N}$, $-nh \leq g \leq nh$ implies $g \in H$. Hence $\{0\} \neq C \subseteq H$, contradicting the assumption that $H$ does not contain a non-trivial convex subgroup of $G$.

**Lemma 2.** *Let $T$ be a fan in the field $K$. Let $v_1$ be a valuation on $K$ with value group $\Gamma_1$; if $v_1(T^*)$ does not contain a non-trivial convex subgroup of $\Gamma_1$, then $v_1$ is fully compatible with $T$.*

**Proof of lemma 2.** We claim that the condition: "for every $m$ in the unique maximal ideal $\mathcal{M}_1$, and for every $t \in U_1 \cap T$, a unit belonging to $T$, $t + m \in T$ implies that $1 + \mathcal{M}_1 \subset T$" entails that $v_1$ is fully compatible with $T$.

We distinguish two cases:

*Case 1.* Assume $v_1(m) \notin v_1(T^*)$.

In this case $(T \cdot m) \cap U_1 = \emptyset$; so in particular $m \notin -T$, since $v_1(m) > 0$. Since $T$ is a fan, $t + m \in T + T \cdot m = T \cup T \cdot m$. We have to show that $t + m \in T$.

Clearly $t+m \in U_1$ because $v_1(t+m)=0$ since $v_1(t)=0$ and $v_1(m)>0$. Since $(T\cdot m)\cap U_1=\emptyset$ we get $t+m\notin T\cdot m$ hence $t+m\in T$.

*Case 2.* Assume $v_1(m)\in v_1(T^*)$.

Apply lemma 1 to $H:=v_1(T^*)$. Since $v_1(m)$ is a positive element of $H$ there exists $x$ such that $v_1(x)\notin H$ and $0<v_1(x)<v_1(m)$. Now let $t+m=t'+m'$ where $t':=t+x$ and $m'=m-x$. From $x\in\mathcal{M}_1$ we get $t'\in U_1$, and since $v_1(m')\notin v_1(T^*)$, case 1 gives $t'\in T$. Finally from $v_1(x)<v_1(m)$ we get $v_1(m')=v_1(m-x)=\min\{v_1(m),v_1(x)\}=v_1(x)\notin v_1(T^*)$. Thus using again case 1, we get $t'+m'\in T$, and hence $t+m\in T$.

**Lemma 3.** *Let $T\subset K$ be a non trivial fan and $P\in\chi_{/T}$. Let $v_P: K^*\longrightarrow\Gamma$ be the canonical valuation associated with $P$; then $v_P(T^*)\neq\Gamma$. In particular $v_P$ is not the trivial valuation so every ordering in $\chi_{/T}$ is non archimedean.*

For the proof of this last lemma we refer to Lam [48], corollary 12-11 of lemma 12-10 p. 95.

**Proof of proposition 2.7.** Given a non trivial fan $T\subset K$, fix $v_0: K^*\twoheadrightarrow\Gamma_0$ such that $v_0(T^*)\neq\Gamma_0$ (for instance, take $P\in\chi_{/T}$ and let $v_0$ be the valuation $v_P$ associated with $A(P)$). Now consider the convex subgroups of $\Gamma_0$ contained in $v_0(T^*)$; they form a chain under inclusion. The union of them $\Delta$ is the largest convex subgroup contained in $v_0(T^*)$. By quotienting we can coarsen the valuation $v_0$ into a valuation $v_1: K^*\twoheadrightarrow\Gamma_1:=\Gamma_0/\Delta$. Then $v_1(T^*)$ cannot contain a non-trivial convex subgroup of $\Gamma_1$. Hence, by lemma 2, $v_1$ is fully compatible with $T$. Since $[\Gamma_1:v_1(T^*)]=[\Gamma_0:v_0(T^*)]>1$, $v_1$ is a non trivial valuation.

**Proposition 2.8.** *For any preordering $T$ in a field $K$, the followings are equivalent:*

*(1) $T$ is a fan in $K$.*

*(2) There exists a valuation $v_1$ on $K$, fully compatible with $T$, such that, with respect to $v_1$, $T$ pushes down to a trivial fan, hence* $\left[\overline{K}^*:\overline{T}^*\right]\leq 4$.

**Proof of proposition 2.8.**

(2)⇒(1) Trivially if $v_1$ exists, is fully compatible with $T$, and pushes down to a trivial fan $\overline{T}$, then $T$ is a fan.

(1)⇒ (2) From the previous proposition we know that there exists a valuation $v$ fully compatible with $T$, hence $\overline{T}$ is a fan in the residue field $k_v$.

If $\left[k_v^*:\overline{T}^*\right]\geq 8$, then $\overline{T}$ would be a non-trivial fan, and applying lemma 3 to $\overline{T}$ in $k_v$ we would get a non-trivial valuation on $k_v$ fully compatible with $\overline{T}$. But from proposition 12-3 in [48], $k_v$ has no non-trivial valuation fully compatible with $\overline{T}$. Then just take $v_1=v$.

For the geometric point of view on fans we refer to [2] and [1].

## 3. Valuation fans and examples.

From now on preorderings are no more supposed to be quadratic.

Let us recall the definition of a general preordering. A preordering $T$ in a field $K$ is a subset $T \subseteq K$, satifying:

$T + T \subseteq T$, $T \cdot T \subseteq T$, $0, 1 \in T$, $-1 \notin T$, $T^* = T \backslash \{0\}$ is a subgroup of $K^*$.

### 3.1. Valuation fans (of any level).

**Definition 3.1.** (Jacob, [41]). Let $K$ be a field; a valuation fan in $K$ is a preordering $T$ such that there exists $v$ a real valuation on $K$, $v$ fully compatible with $T$ (meaning $1 + I_v \subset T$), and $T$ induces an archimedean ordering on the residue field $k_v$.

More precisely, a preordering $T$ in $K$ is a valuation fan if and only if $A(T) = \{x \in K \mid \exists r \in \mathbb{Q}\ r \pm x \in T\}$ is a valuation ring with associated valuation $v$ fully compatible with $T$, and $\overline{T}$ in $k_v$ is an (archimedean) ordering.

There is an alternative characterization for valuation fans given in [42], which is sometimes useful in model theory:

**Proposition 3.2.** *A preordering $T$ in a field $K$ is a valuation fan if and only if for any $x \notin \pm T$ we have either $1 \pm x \in T$ or $1 \pm x^{-1} \in T$.*

Usual orderings $P$ are valuation fans (of level 1, i.e. $\sum K^2 \subset P$).

It is I think important for real algebraic geometry to understand minimal valuation fans of level 1. They are defined as valuation fans not properly containing any valuation fan which is a quadratic preordering. Of course such a minimal valuation fan $T_0$ pushes down an archimedean ordering in the residue field of $K$ for the valuation associated to the valuation ring given by: $A(T_0) = \{x \in K \mid \exists r \in \mathbb{Q}\ r \pm x \in T_0\}$.

**3.2. Orderings of higher level.** Further examples of valuation fans are provided by Becker's orderings of higher level.

**Definition 3.3.** (Becker, [6]). Let $K$ be a commutative real field, $P \subset K$ is an ordering of level $n$ if: $\sum K^{2n} \subset P$, $P + P \subset P$, $P.P \subset P$, $-1 \notin P$, $P^*$ is a subgroup of $K^*$ and $K^*/P^*$is cyclic.

When $K^*/P^* \simeq \mathbb{Z}/2n\mathbb{Z}$, then the ordering is said to be of exact level $n$.

A very interesting paper on sums of d-th powers in rings with some relation to orderings of higher level is [43].

The orderings of level 1 are the usual total orderings.

If $K = \mathbb{R}((X))$, there exist two usual orderings:

$$P_+ := K^2 \cup XK^2, P_- := K^2 \cup -XK^2$$

And for every integer $n \geq 1$ there exist two orderings of exact level $n$:

$$P_{n,+} := K^{2n} \cup X^n K^{2n}, P_{n,-} := K^{2n} \cup -X^n K^{2n}.$$

These higher level orderings have important links with sums of powers; we refer the reader to [9] and just mention the following important theorems from [6] :

**Theorem 3.4.** *(Becker, [6]) Let $K$ be a real field, then:*
$\sum K^{2n} = \cap \{P_i \mid P_i$ *ordering of level dividing* $n\}$.

**Theorem 3.5.** *(Becker, [6]). Let $K$ be a real field, and let $p$ be a prime. The followings are equivalent:*
*(1)* $\sum K^2 \neq \sum K^{2p}$.
*(2) $K$ admits an ordering of exact level $p$.*

**3.3. Another approach with signatures.** Usual orderings can also be studied in terms of signatures. A signature is a group morphism, $\sigma : K^* \longrightarrow \{\pm 1\}$, with additivily closed kernel; then $P = ker\sigma \cup \{0\}$ is an ordering of $K$.

This notion of a signature has a higher level analog:

**Definition 3.6.** (Becker, [8]). A signature of level $n$ on a field $K$ is a morphism of abelian groups:

$$\sigma : K^* \to \mu_{2n}$$

such that the kernel is additively closed, where $\mu_{2n}$ denotes the group of $2n$-th roots of 1.

Clearly if $\sigma$ is a signature of level $n$, then $P = ker\sigma \cup \{0\}$ is an ordering of higher level with exact level dividing $n$.

But there exists also a much more general notion of signature involving valuation fans:

**Definition 3.7.** ( Schwartz, [55]). A generalized signature in a field $K$ is a morphism of abelian groups, $\sigma : K^* \to G$, such that the kernel is a valuation fan.

## 4. Algebraic closure of a field equipped with a valuation fan.

Several notions of a closure, under algebraic extensions, of a field equipped with either higher level orderings or higher level signatures, either valuation fans or generalized signatures, have been introduced and studied in the literature.

All these notions of closure can be unified in one theory, the theory of Henselian Residually Real-Closed fields (HRRC fields).

In this section we present, without any proof, the main features of this theory, from an algebraic point of view.

**Definition 4.1.** (Becker, Berr, G. , [11]). A field $K$ is henselian residually real-closed (HRRC) if and only if it admits an henselian valuation $v$ with real-closed residue field $k_v$.

Recall that a valuation $v$ on a field $K$, with valuation ring $A_v$, is *henselian* if it satisfies Hensel's lemma : *"For any monic polynomial $f \in A_v[X]$, if $\overline{f}$ has a simple root $\beta \in k_v$, then $f$ has a root $b \in A_v$ such that $\overline{b} = \beta$ "*. The henselian residually real-closed fields have been variously named in the literature: they are called real henselian fields in Brown [16], [17], fields real-closed with respect to a signature in Schwartz [55] and almost real-closed fields in Delon-Farre [27].

**4.1. Examples of HRRC fields.** The basic examples of henselian residually real-closed fields arise in a classical way as follows (see [33]): given $R$ a real-closed field, and $\Gamma$ a totally ordered abelian group, let $R((\Gamma)) = \{\sum_\gamma a_\gamma t^\gamma \mid \gamma \in \Gamma, a_\gamma \in R\}$ be the set of generalized power series with support well ordered, where $support \sum_\gamma a_\gamma t^\gamma = \{\gamma \in \Gamma \mid a_\gamma \neq 0\}$. In $K = R((\Gamma))$ one can define:

- Product by: $t^\gamma t^\delta = t^{\gamma+\delta}$;
- Addition by: $\sum_\gamma a_\gamma t^\gamma + \sum_\delta b_\delta t^\delta = \sum_\alpha (a_\alpha + b_\alpha) t^\alpha$;
- Order by: $\sum_\gamma a_\gamma t^\gamma >_K 0 \Leftrightarrow a_m >_R 0$, where $m = \min(support \sum_\gamma a_\gamma t^\gamma)$;
- Valuation by: $v : R((\Gamma)) \to \Gamma$ and $v(\sum_\gamma a_\gamma t^\gamma) = m = \min(support \sum_\gamma a_\gamma t^\gamma)$.

It is well-known that $R((\Gamma))$ is a field, admitting $v$ as a henselian valuation with real-closed residue field $R$ and value group $\Gamma$; hence $R((\Gamma))$ is an HRRC field.

**4.2. Subtheories of the theory of HRRC fields.** Let $v$ be a real valuation on a field $K$, $k_v$ its residue field, $\Gamma_v$ its value group, and let $S$ be a set of primes. Relations between various subtheories of the theory of HRRC fields are described by the following diagram where arrows indicate subtheories.

*Henselian Residually Real-Closed Fields (HRRC)*
$v$ henselian valuation , $k_v$ real-closed field
closed for **generalized signature**, or for **valuation fan**
↓
*HRRC fields of type $S$* ($p \notin S \Rightarrow \Gamma_v$ $p$-divisible)
↙ ↘

*$S$-generalized real-closed fields* ($S$ finite)
if $p \notin S$ then $\Gamma_v$ is $p$-divisible
if $p \in S$ and $\Gamma_v$ is not $p$-divisible then
$\Gamma_v / p\Gamma_v \simeq \mathbb{Z}/p\mathbb{Z}$ ; closed for
**higher level ordering** or **chain signature**
↓ ↘

*Rolle fields*
$\Gamma_v$ odd divisible
HRRC field of type $\{2\}$
↙ ↓

*Real-closed fields*
$\emptyset$-generalized real-closed
$\Gamma_v$ divisible
closed for a **usual order**

*Chain-closed fields*
$\{2\}$-generalized real-closed
$\Gamma_v / 2\Gamma_v \simeq \mathbb{Z}/2\mathbb{Z}$
closed for an **ordering of level $2^k$**

In the diagram above, most of the theories correspond to some notion of closure, under algebraic extensions, of a field equipped with some object. With an *ordering* (real-closed field), with an *ordering of exact level a power of* 2 (chain-closed field), with an *ordering of exact level a power of p where p is prime* ($\{p\}$-real-closed fields), with an *ordering of exact level n* ($S$-generalized real-closed fields of exact type $S$ ($p \in S \Leftrightarrow \Gamma_v$ not $p$-divisible, and for all $p \in S$, $p \mid n$ ), or with a *valuation fan* (henselian residually real-closed field).

**4.3. On the question of the uniqueness of closure.** For a field equipped with a usual ordering it is well known that the real closure is unique up to $K$-isomorphism.

Even for chain-closed fields this is not true anymore. In order to recover the uniqueness of the closure, up to $K$-isomorphism, one needs to consider a closure for a whole chain of orderings with levels powers of 2 in the sense of Harman:

**Definition 4.2.** (Harman, [40]). A 2-primary chain of orderings in a field $K$ is:

$$(P_n)_{n\in\mathbb{N}} = (P_0,P_1,...,P_n,...)$$

$P_0$ being a usual ordering and $P_n$ an ordering of level $2^{n-1}$, such that

$$P_n \cup -P_n = (P_0 \cap P_{n-1}) \cup -(P_0 \cap P_{n-1}).$$

**Theorem 4.3.** *A field $K$ equipped with a 2-primary chain of orderings admits a closure under algebraic extensions unique up to $K$ -isomorphism. The closure is called a chain-closed field and it is equal to the intersection of two real-closures of $K$ for $P_0$ and $P_1$.*

For generalized real-closed fields, in order to recover the uniqueness up to $K$-isomorphism, Niels Schwartz has introduced the notion of chain signature.

**Definition 4.4.** (Schwartz, [54]). A chain signature on a field $K$ is a homomorphism:

$$\varphi : K^* \to \{1,-1\} \times \widehat{\mathbb{Z}}$$

such that ker$\varphi$ is a valuation fan, where $\widehat{\mathbb{Z}} = \prod \widehat{\mathbb{Z}}_p$ and $\widehat{\mathbb{Z}}_p$ denotes the additive group of p-adic integers.

One can recover orderings of higher level by taking:
$P_n(\varphi) = \varphi^{-1}(1 \times n\widehat{\mathbb{Z}}) \cup \{0\}$.

**Theorem 4.5.** *A field $K$ equipped with a chain signature $\varphi$ admits a closure under algebraic extensions unique up to $K$ -isomorphism. This closure is a HHRC field.*

In the more general situation of a field equipped with a valuation fan we can also ensure the uniqueness of the closure by considering a field equipped, not only with a single valuation fan, but with a whole chain of valuation fans.

From Brown's work we can derive the following:

**Theorem 4.6.** *Let $R$ and $R'$ be two HRRC fields, algebraic extensions of a field $K$, then the followings are equivalent:*
*(1) $R$ and $R'$ are $K$-isomophic.*
*(2) $R^{2n} \cap K = R'^{2n} \cap K$ for all $n \in \mathbb{N}$ .*

In fact these $T_n = R^{2n} \cap K$ are valuation fans, which form a chain of valuation fans $(T_n)_{n \in \mathbb{N}}$ as defined below; this chain is said to be induced on $K$ by $R$.

**Definition 4.7.** (Becker, Berr, G. [11]). A chain of valuation fans in a field $K$ is defined as $(T_n)_{n \in \mathbb{N}}$ such that:
(1) $K^{2n} \subset T_n$;
(2) $T_{n.m} \subset T_n$;
(3) $(T_n)^m \subset T_{n.m}$;
(4) $T_n^*/T_{n.m}^* \subset T_1^*/T_{n.m}^*$ is the subgroup of elements of exponent $m$.

With this notion we have been able in [11] to obtain the following theorem:

**Theorem 4.8.** *Any field $K$, equipped with a chain of valuation fans $(T_n)_{n \in \mathbb{N}}$, admits a closure under algebraic extensions $R$, unique up to $K$-isomorphism. Then $R$ is a HRRC field, and $R$ induces on $K$ a chain of valuation fans $(T_n)_{n \in \mathbb{N}}$ (i.e. $T_n = R^{2n} \cap K$ for all $n$).*

**4.4. Properties of HRRC fields.** Henselian residually real-closed fields have a lot of nice properties ; we list, again without any proof, some of them below. Main reference is [11].

Let $K$ be an HRRC field then:

(1) $K$ is a real field;

(2) Every algebraic extension of $K$ is a radical extension;

(3) $K$ a HRRC field of type $S$ has no real extension of degree $p \in \mathbb{P} \backslash S$.
Note that whenever $2 \in S$, one can replace (3) by (3') "$K$ has no extension of degree $p \in \mathbb{P} \backslash S$ ";

(4) $\forall n \in \mathbb{N}$ , $K$ is *n-pythagorean* : $K^{2n} + K^{2n} = K^{2n}$;

(5) $K$ is *hereditarily pythagorean*, i.e., every algebraic extension is again a pythagorean field;

(6) $\forall n \in \mathbb{N}$ , $K^{2n}$ is a *fan* (refer to definion 2.3, or to characteriztion 2.4 for such preorderings);

(7) $\forall n \in \mathbb{N}$ , $K^{2n}$ is a *valuation fan*, i.e. it is a preordering such that:
$\forall x \notin \pm K^{2n}$ either $1 \pm x \in K^{2n}$ or $1 \pm x^{-1} \in K^{2n}$;

(8) All real valuations on $K$ are henselian;

(9) The set of real valuation rings in $K$ is totally ordered by inclusion;

(10) The *smallest* real valuation ring in $K$ is:

$$A(K^2) = A(K^{2n}) = H(K)$$

where $A(T) = \{x \in K \mid \exists n \in \mathbb{N} \ \ n \pm x \in T\}$, $T$ being a valuation fan, and where $H(K)$ is the real holomorphy ring (i.e. the intersection of all real valuation rings);

(11) $K$ admits a unique $\mathbb{R}$-place which can be defined using the valuation ring $A(K^2)$ and the associated valuation;

(12) Jacob's ring $J(\bigcap_{n \in \mathbb{N}} K^{2n})$ is the *biggest* valuation ring with real-closed residue field. This ring is defined as follows. If $T$ is a valuation fan, the ring $J(T)$ is equal to $J_1(T) \cup J_2(T)$ where:

$$\begin{array}{rcl} J_1(T) & = & \{x \in K \mid x \notin \pm T \text{ and } 1 + x \in T\} \\ & & \text{and} \\ J_2(T) & = & \{x \in K \mid x \in \pm T \text{ and } xJ_1(T) \subset J_1(T)\}). \end{array}$$

**4.5. On the model theory of HRRC fields.** These fields have been studied from a model theoretic point of view; the previous theories are all elementary theories, with nice first order axiomatizations (see [11], [24], [25], [28], [34] and [36]).

A Rolle field is an ordered field where Rolle theorem holds for polynomials. These fields have been introduced by Brown, Craven and Pelling [21]. Below is an axiomatization for the theory of Rolle fields; these axioms are first order in the language of fields, hence the theory is elementary.

**Theorem 4.9.** *(G. [34]) :*
*(1) axioms for commutative fields ;*
*(2) "K formally real " :*
*for each $n \geqslant 1$*

$$\forall x_1 ... \forall x_n \ \rceil(-1 = x_1^2 + ... + x_n^2)$$

*(3) "K does not have any algebraic extension of odd degree" :*
*for each $p \geqslant 0$*

$$\begin{array}{rcl} & & \forall x_0 ... \forall x_{2p+1} \exists y \\ (x_{2p+1} & = & 0 \vee x_0 + x_1 y + ... + x_{2p+1} y^{2p+1} = 0) \end{array}$$

*(4) "$K^2$ is a fan " :*

$$\forall x \forall y \forall z \exists t (x = -t^2 \vee y^2 + xz^2 = t^2 \vee y^2 + xz^2 = xt^2)$$

*(5)"K is pythagorean at level 2 " :*

$$\forall x \forall y \exists z (x^4 + y^4 = z^4)$$

Remark that the three first sets of axioms are the same as in the theory of real-closed fields; to get a real-closed field axiomatization, just replace (4) and (5) by

$$\forall x \exists y (x = y^2 \vee x = -y^2)$$

In [G1] it is also shown that :

**Theorem 4.10.** *For any Rolle field $K$ having a finite number of orders $2^n$, there exists $n+1$ orders $P_i$, such that $K$ is the intersection of $n+1$ real closures $R_i$ of $K$ ordered by $P_i$.*

The theory of HRRC fields is also elementary and the next theorem gives a first order axiomatization.

**Theorem 4.11.** *(Becker, Berr, G. [11]) : The class of HRRC fields admits the following axiomatization :*

*(1) $R$ is a real commutative field ;*
*(2) $R$ is a hereditarily pythagorean field ;*
*(3) for all $n \in \mathbb{N}$, $R^{2n}$ is a valuation fan.*

**Corollary 4.12.** *The class of HRRC fields is an elementary class.*

**Remark 4.13.** The class of HRRC fields of type S is also an elementary class, just add to the axiomatization in theorem 4.11 :
(4) for all $p \in \mathbb{P} \backslash S$ , $K^2 = K^{2p}$.

Corollary 4.11 follows from B. Jacob ([41]), who first proved that the class of hereditarily pythagorean fields is elementary.

An alternative proof from [11] for "the class of hereditarily-pythagorean fields is elementary" is given below. It uses the characterization by Becker ([6], thm. 4, p. 94) of hereditarily pythagorean fields :

$$\sum K(X)^2 = K(X)^2 + K(X)^2$$

which is equivalent to :

$$\sum K[X]^2 \subset K(X)^2 + K(X)^2$$

By Cassel's theorem this is also equivalent to :

$$\sum K[X]^2 = K[X]^2 + K[X]^2 \quad (*)$$

Remark that if $f, g, h \in K[X]$ satisfy $f^2 = g^2 + h^2$, the degrees of $g$ and $h$ are less or equal to the degree of $f$ because $K$ is formally real.

Hence $(*)$ is expressible by an infinite sequence of first order sentences in the language of fields.

The significance of Jacob's ring for the model theory of these fields appears in [42], and also later with the transfer theorem obtained by Delon and Farre [27] and given below.

We first recall that in the following theorem, $\equiv$ denotes elementary equivalence and $\prec$ elementary inclusion. The second symbol $\prec$ means that every closed first order formula with parameters in the smaller model holds in one model if and only if it holds in the other.

**Theorem 4.14.** *(Delon, Farré, [27]) : Let $K$ and $L$ be HRRC fields, then :*

*(i) $K \equiv L \Leftrightarrow \Gamma_{J(K)} \equiv \Gamma_{J(L)}$ ;*

*(ii) if $K \subset L$ then $K \prec L \Leftrightarrow \Gamma_{J(L)}$ extends $\Gamma_{J(K)}$, and $\Gamma_{J(K)} \prec \Gamma_{J(L)}$ , where the $\Gamma'$s are the value groups of the Jacob rings of $K$ and $L$.*

In [27] the authors established a bijection between theories of HHRC fields and certain theories of ordered abelian groups. This bijection preserves completeness and sometimes decidability. Finally they proved that the only model-complete theory among these is the theory of real-closed fields.

They also characterized definable real valuation rings in such fields and have shown that these valuation rings were in bijection with the definable convex subgroups of the value group of the Becker ring.

In case there is only one real (henselian) valuation ring with real-closed residue field, i.e. the Becker ring equals the Jacob ring, then the model theory works well, and we are able to get real algebraic results such as a Nullstellensatz or Hilbert's 17th problem at level n; we refer the reader to [12], [13], [26].

## 5. $\mathbb{R}$-places, and the real holomorphy ring.

**5.1. $\mathbb{R}$-place associated to an ordering.** For a complete presentation of these notions one can refer to [48], or in a more geometrical setting to [57], [58] and [59].

Let $K$ be a real field and $P$ be an ordering on $K$. Let $v$ denote the valuation associated to the valuation ring $A(P)$. From previous results we know that $(k_v,\overline{P})$ can be uniquely embedded in $(\mathbb{R},\mathbb{R}^2)$ since $\overline{P}$ is archimedean. Denote this embedding by $i$ and let $\pi$ be the canonical mapping from $K$ into $k_v \cup \{\infty\}$ (where if $a \notin A(P)$, then $\pi(a)=\infty$).

**Definition 5.1.** The $\mathbb{R}$-place associated to $P$ is $\lambda_P : K \to \mathbb{R}\cup\{\infty\}$ defined by the following commutative diagram:

$$\begin{tikzcd}
K \arrow[rr, "\lambda_P"] \arrow[dr, "\pi"'] & & \mathbb{R}\cup\{\infty\} \\
& k_v \cup \{\infty\} \arrow[ur, "i"'] &
\end{tikzcd}$$

Explicitly $\lambda_P(a)=\infty$ when $a \notin A(P)$, and $\lambda_P(a)=\inf\{r \in \mathbb{Q} \mid a \leq_P r\} = \sup\{r \in \mathbb{Q} \mid r \leq_P a\}$ if $a \in A(P)$. In fact it is known that any $\mathbb{R}$-place arises in this way from some ordering $P$ (see [48], 9.1).

**5.2. The space of $\mathbb{R}$-places.** The space of $\mathbb{R}$-places of a field $K$ is the set $M(K) = \{\lambda_P \mid P \in \chi(K)\}$, where $\chi(K)$ denotes the space of orderings of $K$. $M(K)$ is equipped with the coarsest topology making continuous the evaluation mappings defined for every $a \in K$ by:

$$\begin{aligned} e_a \ : M(K) &\longrightarrow \mathbb{R} \cup \{\infty\} \\ \lambda_P &\mapsto \lambda_P(a) \end{aligned}$$

Recall that the usual topology on $\chi(K)$ is the Harrison topology generated by the open-closed Harrison sets:

$$\mathcal{H}(a) = \{P \in \chi(K) \mid a \in P\}.$$

With this topology $\chi(K)$ is a compact totally disconnected space. Craven has shown in [23] that every compact totally disconnected space is homeomorphic to the space of orderings $\chi(K)$ of some field $K$.

Now consider the mapping $\Lambda$ defined by:

$$\begin{aligned} \Lambda : \chi(K) &\longrightarrow M(K) \\ P &\mapsto \lambda_P \end{aligned}$$

With the previous topologies on $\chi(K)$ and $M(K)$ the mapping $\Lambda$ is a continuous, surjective and closed mapping.

$M(K)$ equipped with the above topology is a compact Hausdorff space. Remark that this topology on $M(K)$ is also the quotient topology inherited from the above topology on $\chi(K)$.

**5.3. The Real Holomorphy Ring.** We now provide some facts on the real holomorphy ring which has heavy links with orderings and $\mathbb{R}$-places.

**Definition 5.2.** The real holomorphy ring, denoted $H(K)$, is the intersection of all real valuation rings of $K$.

From the results in part 1 we obtain $H(K) = \bigcap_{P \in \chi(K)} A(P)$.

We also have:

$$H(K) = A(\sum K^2) = \{a \in K \mid \exists n \in \mathbb{N},\ n \geq 1,\ n \pm a \in \sum K^2\}.$$

$H(K)$ is a Prüfer ring with quotient field $K$ (see [48], p.85). Recall that a Prüfer ring is a ring $R \subset K$ such that, for any prime ideal $\mathfrak{p}$ in $R$, the localization $R_{\mathfrak{p}}$ is a valuation ring in $K$ .

In the sequel we denote the real spectrum of the real holomorphy ring of $K$ by:

$$Sper(H(K)) = \{\alpha = (\mathfrak{p},\overline{\alpha}),\ \mathfrak{p} \in Spec(H(K)),\ \overline{\alpha}\ ordering\ of\ quot(H(K)/\mathfrak{p})\}.$$

Relations between $\chi(K)$, $M(K)$ and $H(K)$ are given in [14] by the next theorem.

**Theorem 5.3.** *(Becker, G. [14]) The following diagram is commutative:*

$$\begin{tikzcd}
\chi(K) \arrow[rr, "sper\ i"] \arrow[d, "\Lambda"] & & MinSperH(K) \arrow[d, "sp"] \\
M(K) \arrow[r, "res"] & \mathrm{Hom}(H(K),\mathbb{R}) \arrow[r, "j"] & MaxSperH(K)
\end{tikzcd}$$

*where the horizontal mappings are homeomorphisms, and the vertical ones continuous surjective mappings (see definitions below).*

Hence $\chi(K)$ the space of orderings of $K$ is homeomorphic to $MinSperH(K)$, and the space $M(K)$ of $\mathbb{R}$-places on $K$ is homeomorphic to $MaxSperH(K)$.

The mappings in the above diagram are defined as follows:

$\Lambda : \chi(K) \longrightarrow M(K)$ is given by $P \mapsto \lambda_P$.

$sper\ i : \chi(K) \longrightarrow MinSperH(K)$ is given by $P \mapsto P \cap H(K)$.

$sp : MinSperH(K) \longrightarrow MaxSperH(K)$ is given by $\alpha \longmapsto \alpha^{\max}$ ,
where $\alpha^{\max}$ is the unique maximal specialization of $\alpha$.

$res : M(K) \longrightarrow \mathrm{Hom}(H(K),\mathbb{R})$ is given by $\lambda \mapsto \lambda_{|H(K)}$.

$j : \mathrm{Hom}(H(K),\mathbb{R}) \longrightarrow MaxSperH(K)$ is given by $\varphi \mapsto \alpha_\varphi$,
where $\alpha_\varphi = \varphi^{-1}(\ \mathbb{R}^2)$ or, using the notation for the real spectrum,
$\alpha_\varphi = (\ker\varphi, \overline{\alpha})$ with $\overline{\alpha} = \mathbb{R}^2 \cap quot(\varphi(H(K))$.

All the spaces in the diagram are compact and the topologies of $M(K)$ and $MaxSperH(K)$ are the quotient topologies inherited through $\Lambda$ and *sp*.

**Remark 5.4.** $\chi(K)$ may be seen as a space of maximal valuation fans of level 1, and $M(K)$ might be associated to the space of minimal valuation fans of level 1 using the preorderings $T_\lambda = \cap\{P_i \mid P_i \in \Lambda^{-1}(\lambda)\}$, where $\Lambda^{-1}(\lambda) = \{P_i \mid \lambda_{P_i} = \lambda\}$.

## 6. On the Abstract Side.

The space of orderings of a field, studied in relation with quadratic forms and real valuations, have been the origin of the theory of abstract spaces of orderings (1979-80) and of Marshall's problem:

*"Is every abstract space of orderings the space of orderings of some field ?"*

**6.1. Abstract spaces of orderings (level 1 case).** Abstract space of orderings have been introduced using signatures by Marshall in [50]:

**Definition 6.1.** An abstract space of orderings is $(X, G)$, where $G$ is a group of exponent 2 (hence abelian), -1 a distinguished element of $G$, and $X$ a subset of $Hom(G, \{1, -1\})$ such that:

(1) $X$ is a closed subset of $Hom(G, \{1, -1\})$;

(2) $\forall \sigma \in X \quad \sigma(-1) = -1$;

(3) $\bigcap\limits_{\sigma \in X} \ker \sigma = 1$ (where $\ker \sigma = \{a \in G \mid \sigma(a) = 1\}$);

(4) For any $f$ and $g$ quadratic forms over $G$:

$$D_X(f \oplus g) = \cup\{D_X\,\langle x, y\rangle \mid x \in D_X(f),\ y \in D_X(g)\}.$$

In the above definition $D_X(f)$ denotes the set $\{a \in G$ represented by $f\}$, i.e. there exists $g$ such that $f \equiv_X \langle a\rangle \oplus g$ where $f \equiv_X h$ if and only if $f$ and $h$ have same dimension, and have for any $\sigma \in X$ same signature.

On the side of fans, seen as sets of signatures on a field, a four elements fan of level 1 is characterized by: $\sigma_0\sigma_1\sigma_2\sigma_3 = 1$ and it corresponds to the fan seen as a preordering: $T = \bigcap\limits_{i=0}^{3} \ker \sigma_i \cup \{0\}$.

In the abstract situation, abstract fans have been defined by Marshall.

**Definition 6.2.** An abstract fan is an abstract space of orderings $(X, G)$ such that $X = \{\sigma \in Hom(G, \{1, -1\}) \mid \sigma(-1) = -1\}$.

It is also characterized by: if $\sigma_0, \sigma_1, \sigma_2 \in X$ then the product $\sigma_0\sigma_1\sigma_2 \in X$.

What was expected to correspond to the space of $\mathbb{R}$-places of the field case in the context of abstract spaces of orderings is called a $P$-structure and has been defined as follows by Marshall in [51].

**Definition 6.3.** A $P$-structure is an equivalence relation on a space of orderings $(X, G)$ such that the canonical mapping $\Lambda : X \to M$, where $M$ is the set of equivalence classes, satisfies:

(1) Each fiber is a fan;

(2) If $\sigma_0\sigma_1\sigma_2\sigma_3 = 1$ then $\{\sigma_0, \sigma_1, \sigma_2, \sigma_3\}$ has a non empty intersection with at most two fibers.

Marshall has proved in [51] that every abstract space of orderings has a $P$-structure, generally not unique. But unlike the case of the space of $\mathbb{R}$-places in a field, this $P$-structure $M$, equipped with the quotient topology, is not always Hausdorff. Hence we have to improve this notion to fit with the space of $\mathbb{R}$-places in the field case.

**6.2. Abstract spaces of signatures (higher level).** In the higher level case, one can also define abstract spaces of signatures (similar to 3.3 in the field case).

**Definition 6.4.** An abstract space of signatures of level $2^n$ is $(X, G)$, $G$ abelian group of exponent $2^n$, $X \subset \mathrm{Hom}(G, \mu_{2^n})$ such that:

(0) $\forall \sigma \in X$, $\forall k \in \mathbb{N}$ with $k$ odd, $\sigma^k \in X$;
(1) $X$ is a closed subset of $\mathrm{Hom}(G, \mu_{2^n})$ ;
(2) $\forall \sigma \in X \ \ \sigma(-1) = -1$ ($-1$ distinguished element of $\mu_{2^n}$);
(3) $\bigcap_{\sigma \in X} \ker \sigma = 1$ (where $\ker \sigma = \{a \in G \mid \sigma(a) = 1\}$);
(4) For any $f$ and $g$ forms over $G$

$$D_X(f \oplus g) = \cup \{D_X \langle x, y\rangle \mid x \in D_X(f),\ y \in D_X(g)\}.$$

In fields, the space of $\mathbb{R}$-places is known as soon as one knows the usual orderings and the orderings of level 2. Using this idea in the abstract situation we have been able to obtain in [38] a theorem which can be seen as the first case of a $P$-structure which looks like an abstract space of $\mathbb{R}$-places.

**Theorem 6.5.** *Let $(X, G)$ be a subspace of a space of signatures $(X', G')$ with 2-power exponent.*

*For $\sigma_0, \sigma_1 \in X$, define $\sigma_0 \sim \sigma_1$ if $\sigma_0\sigma_1 = \tau^2 \in X'^2$.*

*Then the followings are equivalent:*

*(1) If $\sigma_0\sigma_1\sigma_2\sigma_3 = 1$, then either $\sigma_0$ is in relation by $\sim$ with exactly one of the $\sigma_1, \sigma_2, \sigma_3$, or $\sigma_0$ is in relation by $\sim$ with everyone of the $\sigma_1, \sigma_2, \sigma_3$.*

*(2) $\sim$ defines a $P$-structure on $X$.*

Moreover in this case the induced $P$-structure defined on $X$ by $\sim$ has a Hausdorff topology.

The key idea for proving the theorem is that, in the field case studied by Harman in [40], for any $P_2$, ordering of level 2, holds for some orderings $P_0, P_1$:
$a^2 \in P_2 \iff a \in P_2 \cup -P_2 = (P_0 \cap P_1) \cup (-(P_0 \cap P_1))$.
Hence on the side of abstract signatures we get $\tau(a^2) = \tau(a)^2 = \sigma_0(a)\sigma_1(a)$.

## 7. Some open problems and the complete real spectrum

**7.1. The space of valuation fans.** Study in the field case the space of level 1 valuation fans $\mathcal{VF}(K)$, and its relation with $SperH(K)$. The motivation comes from the fact that $\chi(K)$, isomorphic to $MinSperH(K)$, consists of valuation fans $P_i$, and that to a $\mathbb{R}$-place $\lambda$ in $M(K)$, which is isomorphic to $MaxSperH(K)$, can be associated a valuation fan of level 1: $T_\lambda = \cap\{P_i \mid P_i \in \Lambda^{-1}(\lambda)\}$ where $\Lambda^{-1}(\lambda) = \{P_i \mid \lambda_{P_i} = \lambda\}$.

Then we could define a notion of abstract space of valuation fans of level 1, or of minimal valuation fans of level 1. Then use abstract space of valuation fans of level 1 to solve Marshall's problem of realizability of abstract spaces of orderings, or state Marshall's problem in other terms.

**7.2. The space of $\mathbb{R}$-places.** Construct a finer theory for abstract spaces of orderings taking into account the $\mathbb{R}$-places. For example, $\mathbb{Q}(2^{\frac{1}{2}})$ and $\mathbb{R}((X))$ have isomorphic spaces of orderings, but the first one has two $\mathbb{R}$-places and no ordering of level 2, and the second one has only one $\mathbb{R}$-place but has a 2-primary chain of higher level orderings. A preliminary step is to characterize the topological spaces which are realizable as spaces of $\mathbb{R}$-places. Partial results in that direction have been recently obtained in [30], [45], [46] and [49].

It will be useful to study for a field $K$ the space of connected components of the space of $\mathbb{R}$-places of $K$, $\pi_0(M(K))$. This might be some kind of space of orderings. Another question in this area is: in which cases are the connected components of $M(K)$ homeomorphic?

**7.3. The complete real spectrum.** With Murray Marshall, in [39], we considered rings instead of fields.

Let $A$ be any commutative ring with 1. We define a big object $\mathrm{Sper}^c A$ which we call the *complete real spectrum* of $A$. There are various connections between this and the valuation spectra considered by R. Huber and M. Knebush in Contemporary Math. 155 (1994). Roughly speaking, the complete real spectrum is related to the valuation spectrum in the same way that the real spectrum is related to the prime spectrum. We define a topology on $\mathrm{Sper}^c A$ and prove that $\mathrm{Sper}^c A$, with this topology, is a spectral space.

In [51, Sect. 8.6] another sort of attempt is made to overcome shortcomings of the real spectrum of $A$ by introducing the space of real places of $A$, which we denote here by $M_A$. By definition, $M_A$ consists of pairs $(\mathfrak{p}, \lambda)$ where $\mathfrak{p}$ is a real prime of $A$ and $\lambda$ is a place from the residue field $k(\mathfrak{p})$ into the field of real numbers. This takes care of the real places in a satisfactory way but does not keep track of all real valuations on the $k(\mathfrak{p})$ and all the orderings on the corresponding residue fields of $k(\mathfrak{p})$. Still, the $M_A$ construction in [51] is closely related to the complete real spectrum construction described below.

**Definition 7.1.** (Marshall, G. [39]) The elements of $\mathrm{Sper}^c A$ are triples $(\mathfrak{p}, v, P)$ where $\mathfrak{p}$ is a real prime of $A$, $v$ is a real valuation (more precisely, an equivalence class of real valuations) on the residue field $k(\mathfrak{p})$, and $P$ is an ordering on the residue field $B_v/M_v$ of $v$. Here, $B_v \subseteq k(\mathfrak{p})$ denotes the valuation ring of $v$ and $M_v$ its maximal ideal.

It is possible to give another definition of the complete real spectrum.

**Definition 7.2.** The elements of $\mathrm{Sper}^c A$ are pairs $(\mathfrak{p}, Q)$ where $\mathfrak{p}$ is a real prime of $A$ and $Q$ is an element of $\mathrm{Sper}\, H_{k(\mathfrak{p})}$.

There are natural maps

$$(\mathfrak{p}, v, P) \mapsto \mathfrak{p}, \quad (\mathfrak{p}, v, P) \mapsto (\mathfrak{p}, v)$$

from $\mathrm{Sper}^c A$ into $\mathrm{Spec}\, A$ (the prime spectrum of $A$) and from $\mathrm{Sper}^c A$ into $\mathrm{Spv}\, A$ (the valuation spectrum of $A$), and a natural map

$$(\mathfrak{p}, P) \mapsto (\mathfrak{p}, 0, P)$$

(where 0 denotes the trivial valuation on $k(\mathfrak{p})$) from $\mathrm{Sper}\, A$ into $\mathrm{Sper}^c A$.

There is also the specialization map

$$(\mathfrak{p}, Q) \mapsto (\mathfrak{p}, Q')$$

from $\mathrm{Sper}^c A$ onto the space of real places $M = M_A$ defined in [51]. Here, $Q'$ denotes the unique maximal specialization of $Q$ in $\mathrm{Sper}\, H_{k(\mathfrak{p})}$; also see [14]. The composite map $\mathrm{Sper}\, A \to M_A$ is just the P-structure map $\Lambda$ considered in [51].

$\mathrm{Sper}^c A$ is given a topology and turned to be a spectral space. Subbasic open sets in $\mathrm{Sper}^c A$ are defined using pairs of elements of $A$. For $(a, b) \in A \times A$, we define:

$$U(a,b) = \{(\mathfrak{p}, v, P) \in \ \mathrm{Sper}^c A: \ v(a) = v(b) \neq \infty \text{ and } \frac{a+\mathfrak{p}}{b+\mathfrak{p}} + M_v > 0 \text{ at } P\}$$

Here, $v(a)$ is standard shorthand notation for $v(a + \mathfrak{p})$.

**Remark 7.3.** It is interesting to note that the complete real spectrum of a formally real field $K$ is naturally identified with the real spectrum of its real holomorphy ring $H_K$ involved in the diagram 5.3.

## 8. References

[1] C. Andradas, L. Bröcker, J. Ruiz, *Constructible Sets in Real Geometry.* Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 33, Springer 1996.

[2] C. Andradas and J. Ruiz, *Algebraic and Analytic Geometry of Fans.* Memoirs of AMS 115 (1995), no. 553.

[3] E. Artin and O. Schreier, Algebraischer Konstruktion reeller Körper. *Hamb. Abh.* **5** (1926), 85–99.

[4] R. Baer, Über nicht-archimedisch geordnete Körper. *Sitz. ber. Heidelberger Akad. Abh.* 1927, 3–13.

[5] J. Bochnak, M. Coste and M.-F. Roy, *Géométrie algébrique réelle.* Springer 1987.

[6] E. Becker, *Hereditarily Pythagorean Fields and Orderings of Higher Level.* IMPA Lecture Notes 29, Rio de Janeiro, 1978.

[7] E. Becker, Valuations and real places in the theory of formally real fields. In *Géométrie Réelle et Formes Quadratiques, Proc. Rennes (1981)* (ed. by M. Coste, L. Mahé, M.-F. Roy), Lecture Notes in Mathematics 959, Springer 1982, 1–40.

[8] E. Becker, Extended Artin-Schreier theory of fields. *Rocky Mountain J. of Math.* **14** (1984), 881–897.

[9] E. Becker, The real holomorphy ring and sums of 2n-th powers. In *Géométrie Réelle et Formes Quadratiques, Proc. Rennes (1981)* (ed. by M. Coste, L. Mahé, M.-F. Roy), Lecture Notes in Mathematics 959, Springer 1982, 139–181.

[10] E. Becker and L. Bröcker, On the description of the reduced Witt ring. *J. of Algebra* **52** (1978), 328–346.

[11] E. Becker, R. Berr, and D. Gondard, Valuation fans and residually real-closed henselian fields. *J. of Algebra* **215** (1999), 574–602.

[12] E. Becker, R. Berr, F. Delon, and D. Gondard, Hilbert's 17th problem for sums of 2nth powers. *J. Reine Angew. Math.* **450** (1994), 139–157.

[13] E. Becker and D. Gondard, On rings admitting orderings and 2-primary chains of orderings of higher level. *Manuscripta Math.* **65** (1989), 63–65.

[14] E. Becker and D. Gondard, Notes on the space of real places of a formally real field. In *Real Analytic and Algebraic Geometry, Proc. Trento 1992* (ed. by F. Broglia, M. Galbiati and A. Tognoli), W. de Gruyter 1995, 21–46.

[15] [BK] E. Becker and E. Kopping, Reduzierte quadratische Formen und Semiordnung reeller Körper. *Abh. Math. Sem. Univ. Hamburg* **46** (1977), 143–177.

[16] R. Brown, Automorphism and isomorphism of henselian fields. *Trans. A.M.S.* **307** (1988), 675–703.

[17] R. Brown, Superpythagorean fans. *J. of Algebra* **42** (1976), 483–494.

[18] R. Brown, Real places and ordered fields. *Rocky Mount. J. Math* **1**(1971), 633–636.

[19] R. Brown, Real-valued places on the function field of an algebraic curve. *Houston J. Math.* **6** (1980), 227-243.

[20] L. Bröcker, Characterization of fans and hereditarily pythagorean fields. *Math. Z.* **151** (1976), 149–163.

[21] R. Brown, T. Craven and M. Pelling, Ordered fields satisfying Rolle's theorem. *Ill. J. Math.* **30** (1980), 66–78.

[22] J.-L. Colliot Thélène, Eine Bemerkung zu einem Satz von E. Becker und D. Gondard. *Math. Zeitschrift* **249** (2005), 541–543.

[23] T. C. Craven, The topological space of orderings of rational function field. *Duke Math. J.* **41** (1974), 339–347.

[24] F. Delon, Corps et anneaux de Rolle. *Proc. Amer. Math. Soc.* **97** (1986), 315–319.

[25] F. Delon, Compléments sur les corps chaîne-clos. In *séminaire Structures Algébriques Ordonnées*, Publications de l'Université Paris VII, vol. 33, 1990, 13–20.

[26] F. Delon and D. Gondard, 17ème problème de Hilbert sur les corps chaîne-clos. *J. Symbolic Logic* **56**, (1991), 853–861.

[27] F. Delon and R. Farre, Some model theory for almost real-closed fields. *J. Symbolic Logic* **61** (1996), 1121–1151.

[28] M. Dickmann, The model theory of chain-closed fields. *J. Symbolic Logic* **53** (1988), 921–930.

[29] O. Endler, *Valuation Theory.* Springer-Verlag, Berlin-New-York 1972.

[30] I. Efrat and K. Osiak, Topological spaces as spaces of $\mathbb{R}$-places. *J. Pure Applied Alg.* **215** (2011), 830–846.

[31] A. Endler and A. Prestel, *Valued fields.* Springer monographs in Mathematics, Springer-Verlag, Berlin (2005).

[32] R. Farré, A Positivstellensatz for chain-closed fields, $\mathbb{R}((t))$ and some related fields. *Arch. Math.* **57** (1991), 446–455.

[33] L. Fuchs, *Partially ordered algebraic systems.* Pergamon Press, Oxford-London-New York-Paris, Addison-Wesley publishing Co. 1963.

[34] D. Gondard-Cozette, Axiomatisations simples des théories des corps de Rolle. *Manuscripta Mathematica* **69** (1990), 267–274.

[35] D. Gondard-Cozette, Chainable fields and real algebraic geometry. In *Real Analytic and Algebraic Geometry, Proc. Trento (1988)* (ed. by M. Galbiati and A. Tognoli). Lecture Notes in Mathematics 1420, Springer-Verlag 1990, 128–148.

[36] D. Gondard, Théorie du premier ordre des corps chaînables et des corps chaîne-clos. *C. R. Acad. Sc. Paris* **304** (1987).

[37] D. Gondard-Cozette, $\mathbb{R}$-places et géométrie algébrique réelle. In *Séminaire Structures Algébriques Ordonnées*, prépublication 75, Equipe de Logique, Université Paris VII, 2003.

[38] D. Gondard and M. Marshall, Towards an abstract description of the space of $\mathbb{R}$-places. *Contemporary Mathematics* **253**, A.M.S. 2000, 77–113.

[39] D. Gondard and M. Marshall, Real holomorphy rings and the complete real spectrum. *Ann. Fac. Sci. Toulouse Math.* **6** (2010), 5774.

[40] J. Harman, Chains of higher level orderings. *Contemporary Mathematics* **8**, A.M.S. 1982, 141–174.

[41] B. Jacob, Fans, real valuations, and hereditarily-pythagorean fields. *Pacific J. Math.* **93** (1981), 95–105.

[42] B. Jacob, The model theory of generalized real-closed fields. *J. reine angew. Math.* **323** (1981), 213–220.

[43] J.-R. Joly, Sommes de puissances d-èmes dans un anneau commutatif. *Acta Arithmetica* **17** (1970), 31–114.

[44] W. Krull, Allgemeine bewertungstheorie. *J. reine angew. Math.* **167** (1931), 160–196.

[45] F.-V. Kuhlmann and K. Kuhlmann, Embedding theorems for spaces of $\mathbb{R}$-places of rational function fields and their products. *Fundamenta Math.* **218** (2012), 121–149.

[46] F.-V. Kuhlmann, M. Machura and K. Osiak, Spaces of $\mathbb{R}$-*places* of function fields over real closed fields. *Comm. in Alg.* **39** (2011), 3166–3177.

[47] F.-V. Kuhlmann and A. Prestel, On places of algebraic function fields. *J. reine angew. Math.* **353** (1984), 181–195.

[48] T. Y. Lam, *Orderings, Valuations and quadratic forms.* Regional Conference Series in Mathematics 52, AMS 1983.

[49] M. Machura, M. Marshall and K. Osiak, Metrizability of the space of $\mathbb{R}$-places of a real function field. *Math. Z.* **266** (2010), 237–242.

[50] M. Marshall, A simple system of axioms for spaces of signatures. *Pure and Applied Mathematics* **57** (1989), 159–164.

[51] M. Marshall, *Spaces of Orderings and Abstract Real Spectra.* Lecture Notes in Mathematics 1636, Springer 1996.

[52] P. Ribenboim, *Arithmétique des corps.* Hermann 1972.

[53] P. Ribenboim, *Théorie des valuations.* Les Presses de l'Université de Montréal, 1964.

[54] C. Scheiderer, *A short remark on a theorem by Becker and Gondard.* `http://www.math.uni-konstanz.de/~scheider/pub.html`.

[55] N. Schwartz, Chain signatures and real closure. *J. reine angew. Math.* **347** (1984), 1–20.

[56] N. Schwartz, Signatures and real closures of fields. In *Séminaire Structures Algébriques Ordonnées,* Publications de l'Université Paris VII **33** (1990), 65–78.

[57] H.-W. Schülting, On real places of a field and their holomorphy ring. *Communications in Algebra* **10** (1982), 1239–1284.

[58] H.-W. Schülting, The strong topology on real algebraic varieties. *Contemporary Mathematics* **8** (1982), 141–174.

[59] H.-W. Schülting, Real holomorphy rings in real algebraic geometry. In *Géométrie Réelle et Formes Quadratiques, Proc. Rennes (1981)* (ed. by M. Coste, L. Mahé and M.-F. Roy), Lecture Notes in Mathematics 959, Springer 1982, 433–442.

[60] J.-P. Serre, Extensions de corps ordonnés. *C. R. Acad. Sci. Paris* **229** (1949), 576–577.

Danielle Gondard-Cozette, Institut de Mathematiques de Jussieu - Paris rive gauche,
Université Pierre et Marie Curie, 4 place Jussieu, 75252 Paris cedex 05, France
E-mail: danielle.gondard@imj-prg.fr

# Faithfully quadratic rings; an overview

M. Dickmann
Institut de Mathématiques de Jussieu – Paris Rive Gauche
CNRS – Universités Paris 6 et 7

In this talk we gave an overview, without proofs, of the main results of the joint monograph [DM2] [1]

The theory of **reduced special groups** (abbreviated RSG) is an abstract theory of quadratic forms developed in the 1990s ([DM1]) modeled on quadratic form theory over fields of characteristic $\neq 2$.

In [DM2] we apply the theory of RSGs to study (diagonal) quadratic forms with invertible coefficients over certain classes of (commutative, unitary) rings, $A$, called **$T$-faithfully quadratic**, where $T$ is $A^2$ or a preorder of $A$. The axioms defining $T$-faithfully quadratic rings guarantee that:

(1) The structure $G_T(A)$, consisting of the group $A/T$ endowed with a suitable notion of representation of elements by quadratic forms, is a special group.

(2) The intrinsic theory of quadratic forms based on a suitable notion of **$T$-isometry** is identical to the formal theory of quadratic forms carried by the special group $G_T(A)$.

This later identity guarantees that a number of fundamental structural results known to be valid in the classical theory of quadratic forms over fields of characteristic 0 or $\neq 2$, are also valid over $T$-faithfully quadratic rings [2].

The following classes of rings are proved to be $T$-faithfully quadratic in [DM2]:

(i) Rings with many units are (with a few exceptions) *completely* faithfully quadratic, i.e., $T$-faithfully quadratic for $T = A^2$ and all preorders $T$.

(ii) Reduced $f$-rings, $A$, are $T$-faithfully quadratic for all preorders $T$ containing the natural partial order $T_{\sharp}^A$ of $A$.

(iii) In the important particular case of (ii), where $A = \mathcal{C}(X) =$ the ring of continuous real-valued functions on a topological space $X$, it follows that $\mathcal{C}(X)$ is *completely* faithfully quadratic.

(iv) *Strictly representable* preordered rings $\langle A, T \rangle$ [3].

In all these cases we determine the structure of the associated RSG $G_T(A)$.

---

[1] A summary of these results also appears in [DM3].

[2] These include, among others, the Arason-Pfister Hauptsatz, Milnor's Witt ring conjecture mod 2, Marshall's signature conjecture, uniform upper bounds for the Pfister index of quadratic forms.

[3] I.e., preordered rings having a representation dense in an algebra $\mathcal{C}(X)$ for some compact Hausdorff space $X$, where the elements of $T$ are represented by non-negative functions and those of $T^{\times}$ by strictly positive functions.

## References

[DM1] M. Dickmann, F. Miraglia, **Special Groups. Boolean-Theoretic Methods in the Theory of Quadratic Forms**, Memoirs Amer. Math. Soc. **689** (2000), 247 pp.

[DM2] M. Dickmann, F. Miraglia, **Faithfully Quadratic Rings**, Memoirs Amer. Math. Soc. **1128** (2015), xi + 141 pp.

[DM3] M. Dickmann, F. Miraglia, **Faithfully Quadratic Rings. A Summary of Results**, Banach Center Publ. **108**, Polish Academy of Sciences, Warsaw (2016), 37–48.

# A NOTE ON SUMS OF POWERS IN REAL ALGEBRAIC FUNCTION FIELDS IN ONE VARIABLE OVER ℝ

E.BECKER

*Dedicated to Professor and Friend F. Miraglia on the Occasion of his Seventieth Birthday.*

## CONTENTS

## 1. INTRODUCTION

This note contributes to the study of sums of powers, mainly of even exponents, in real algebraic function fields in one variable over the field of real numbers ℝ. Shortly after E. Artin's general solution of Hilbert's 17th problem in 1927, cf. [1], it was E. Witt who studied the representation of functions as sums of squares for algebraic function fields in one variable over ℝ [18]. His seminal paper gave an astounding complete answer in this special case. The present author was able to extend Artin's method and general results and to develop methods for the study of sums of $2n$-th powers in formally real fields, cf. [4] and the references quoted there. In this paper we return to the case studied by E. Witt with the objective to present some specific features of sums of powers of higher exponents.

Quite generally, let $K$ denote any field. Set

$$\sum_1^r K^d = \{\sum_{i=1}^r {x_i}^d | x_i \in K\}, \sum K^d = \bigcup_{r=1}^{\infty} \sum_1^r K^d.$$

In the first two sections we deal with the quantitative aspects of how many sums of $d$-th powers are needed to represent the elements in $\sum K^d$. As it is quite common we call

$$p_d(K) = \min\{r \mid \sum K^d = \sum_1^r K^d\}$$

the $d-$th Pythagoras number of $K$. If the minimum does not exist we set $p_d(K) = \infty$.

We are dealing with the task of obtaining reasonable bounds for these invariants, mainly for even exponents $d = 2n$. To keep things simple we only work with extension fields of ℝ. The main results are concerned with formally real fields exclusively. Notably, the study of higher Pythagoras numbers for fields of positive characteristic is a completely different business and requires other techniques, cf. [11] e.g.

From Witt's work we know $p_2(F) = 2$ if $F$ is an algebraic function field function in one variable over $\mathbb{R}$. Allowing a more general setting in the first section we review older proofs to display general bounds on $p_{2n}(K)$, in particular under the assumption of $p_2(K) \leq 2$. A few remarks are added for Pythagorean fields, i.e. formally real fields with $p_2(K) = 1$.

Starting with section 2 we focus on formally real algebraic function fields in one variable over $\mathbb{R}$ , always denoted by $F$. The general bounds of section 1 can be improved in this case. To get the sharper bounds we study a certain topological representation

$$\Phi : F \to C(\gamma, \mathbb{P}^1)$$

of the elements of $F$ as continuous functions on a compact smooth real curve $\gamma$ into the projective line $\mathbb{P}^1$ over $\mathbb{R}$. The image of $\phi$ is shown to be a dense subset, a fact that accounts for better bounds.

The last section is devoted to a Unique Factorization Property of non-zero sums of squares in $F$ with respect to the group of the totally positive units of the real holomorphy ring $H$ of $F$. This latter ring turns out to be the pre-image of $C(\gamma, \mathbb{R})$ under the representation map $\Phi : F \to C(\gamma, \mathbb{P}^1)$. In addition, $H$ is a Dedekind domain a fact which yields the Unique Factorization Property.

## 2. General Extension Fields of $\mathbb{R}$

We start by listing basic notions and facts for formally real fields which will be used to derive general bounds for the higher Pythagoras numbers, cf. [3, 4, 8, 12, 15] for more details. Most of them were first introduced and studied by Dubois, [8].

A field $K$ is called formally real if $-1 \notin \sum K^2$, equivalently if $K$ admits an ordering $P$. i.e. a subset $P \subseteq K$ satisfying the conditions

$$P + P \subseteq P, PP \subseteq P, -1 \notin P, P \cup -P = K.$$

From Artin we know: $\sum K^2 = \bigcap P$ where $P$ ranges over all orderings $P$ of $K$.

Associated with orderings are **real places** $\lambda : K \to \mathbb{R} \cup \infty$. Quite generally, a mapping $\lambda : K \to L \cup \infty$ , $L$ a field augmented by an element $\infty$, subject to the following conditions

(1) $V_\lambda := \{x \in K \mid \lambda(x) \neq \infty\}$ is a valuation ring,
(2) $\lambda : V_\lambda \to L$ is a ring homomorphism,
(3) $\lambda(x) = \infty \Rightarrow \lambda(x^{-1}) = 0$.

is called a $L$-valued place of $K$.

A field is formally real if and only if it admits an orderings, if and only if it admits a real place. Formally real fields contain $\mathbb{Q}$. All real places of a formally real field are induced by its orderings. The way this works is as follows: given an ordering $P$ consider the ring

$$A(P) = \{x \in K \mid n \pm x \in P \text{ for some } n \in \mathbb{N}\}.$$

with its maximal ideal

$$I(P) = \{x \in K \mid \frac{1}{n} \pm x \in P \text{ for all } n \in \mathbb{N}\}.$$

This ring is a valuation ring with its residue field $A(P)/I(P)$ contained in $\mathbb{R}$. We therefore get a real place $\lambda_P : K \to \mathbb{R} \cup \infty$, and all real places on $K$ arise in this manner.

A valuation ring of a field $K$ is called a **real valuation ring** if its residue field is formally real. A field $K$ admits a real valuation ring if and only it is formally real. The valuation rings $A(P)$ are obviously real valuation rings, and each real valuation ring contains a ring $A(P)$ for some ordering $P$ of $K$.

The **real holomorphy ring** $H(K)$ of a formally real field $K$ is a fundamental invariant which incorporate in a subtle manner most of the relevant "reality" features of the field. It is defined as follows:

$$H(K) = \bigcap V = \bigcap A(P)$$

where $V$ ranges over all real valuation rings and $P$ over all orderings of $K$. We set

$$\mathbb{E}(K) = H(K)^*, \mathbb{E}_+(K) = H(K)^* \cap \sum K^2$$

for its group of units, resp. for the group of totally positive units.

**Proposition 1.** (1) *$H(K)$ is a Prüfer ring, i.e. every finitely generated fractional ideal is invertible,*

(2) *more precisely: $(a_1, \ldots, a_r)^{2n} = (\sum_{i=1}^{r} a_i^{2n})$ for any finitely generated fractional ideal of $H(K)$,*

(3) *given $q_1, \ldots, q_r \in \sum K^2$ then $(q_1, \ldots, q_r) = (\sum_{i=1}^{r} q_i)$.*

The next invariant we need is the **space of real places** of $K$, defined as a set as follows:

$$M(K) = \{\lambda : K \to \mathbb{R} \cup \infty\}.$$

It is known that $K$ is the field of quotients of $H(K)$. Each $a \in H(K)$ gives rise to a function $\hat{a} : M(K) \to \mathbb{R}$ by evaluation:

$$\hat{a} : M(K) \to \mathbb{R}, \lambda \mapsto \lambda(a).$$

We endow $M(K)$ with the strong topology with respect to the family of functions $\hat{a}, a \in H(K)$, i.e. the coarsest topology such that all maps $\hat{a}$ are continuous. With this topology, $M(K)$ is a compact space, referred to as the **space of real places** of $K$. Consequently, we arrive at a representation, i.e. a ring homomorphism:

$$\Phi : H(K) \to C(M(K), \mathbb{R}), a \mapsto \hat{a}.$$

It is this representation which plays a fundamental role. The Banach-algebra $C(M(K), \mathbb{R})$ is considered under its natural topology of uniform convergence.

**Proposition 2.** (1) *The image of $\phi$ is dense,*

(2) *for each $a \in H : a \in \mathbb{E}_+(K) \Leftrightarrow \hat{a} > 0$ everywhere on $M(K)$.*

We now turn to first results on bounds for higher Pythagoras numbers $p_d(K), d \geq 2$. The invariant $p_2(K)$ is simply called the Pythagoras number of the field $K$. As said earlier we confine ourselves to the case

$$\mathbb{R} \subseteq K.$$

The following two cases are more easily disposed of: the **exponent $d$ is odd** or **$K$ is not formally real.**

If $d$ is odd then $-1 = (-1)^d$, if $K$ is not formally real it is proven, cf. [4, 11], that $-1 \in \sum K^d$, say, $-1 = \sum_1^{s_d} x_i^d$. Then using the identity

$$d!X = \sum_{h=0}^{d-1} (-1)^{d-1-h} \binom{d-1}{h} [(X+h)^d + (-h^d)]$$

one concludes that $p_d \leq s_d(d+1)$ not saying that this (too) easily obtained bound is a very sharp one. In fact, if $d = 3$ then $p_3 \leq 3$ (instead of $p_3 \leq 4$) follows from the socalled Richmond's identity, quoted from [9, p.10]:

$$r = \left[\frac{s(1+t^3)}{3(1-t+t^2)}\right]^3 + \left[\frac{s(3t-1-t^3)}{3(1-t+t^2)}\right]^3 + \left[\frac{s(3t-3t^2)}{3(1-t+t^2)}\right]^3, r, s \in K, s \neq 0, t = \frac{3r}{s^3}.$$

Only little is known about the the "stufe of higher level $s_d$". So, this topic is completely left aside in this note.

**For the rest of this section we additionally assume $K$ to be a formally real field.**

We will make extensive use of the socalled **Hilbert identity**:

$$(*) \quad \left(\sum_{i=0}^{k} X_i{}^2\right)^n + 2nX_0{}^2\left(\sum_{i=0}^{k} X_i{}^2\right)^{n-1} = \sum_{i=1}^{m} \beta_i L_i(X_0, \cdots, X_k)^{2n}$$

where $k, n \in \mathbb{N}, \beta_i \in \mathbb{Q}, \beta_i > 0$, linear forms $L_i \in \mathbb{Q}[X_0, \cdots, X_k]$ and $m \leq \binom{2n+k}{k}$. This identity was used by Hilbert in his solution of the famous Waring problem. That we can bound the number of terms by the binomial coefficient is due to Caratheodory's theorem, [14, p.27, prop. 2.3]. This fact is important to us, it was not made use of in [4] and had led to larger upper bounds. The proof of the following result is taken from the PhD thesis of Natea Hunde [13], it optimizes the arguments from [4].

**Proposition 3.** *Let $p := p_2(K) < \infty$ then $p_{2n} \leq p^2\binom{2n+p}{p}$.*

*Proof.* In Hilbert's identity the coefficients $\beta_i$ can be merged with the linear forms since they are $2n$-th powers, in other words we can assume that all $\beta_i = 1$.
**1st step:** We show that $\sum_i a_i^{2n} = \epsilon \sum_1^p b_j^{2n}$ where $\epsilon$ is a totally positive unit $\epsilon$, in particular $\epsilon = \sum_1^p x_i^2$ with all $x_i \in H$ necessarily. To see the first claim we pass to a sequence of fractional ideals: $(\sum_i a_i^{2n}) = (\ldots, a_i, \ldots)^{2n} = [(\ldots, a_i, \ldots)^2]^n = (\sum_i a_i^2)^n = (\sum_{j=1}^p b_j^2)^n = \cdots = (\sum_j b_j^{2n})$.
**2nd step:** We now consider the totally positive unit $\epsilon$ and assume $\hat{\epsilon} < 1$. Then for all $i$ we get $\widehat{x_i^2} < 1$, hence $1 - x_i^2 = \sum_{j=1}^p y_{ij}^2$, by prop. 2. We get $1 = x_i^2 + \sum_j y_{ij}^2$. Plugging in $x_i, y_{i1}, \ldots, y_{ip}$ in Hilbert's identity we find for each index $i$

$$1 + 2nx_i^2 = \textstyle\sum_1^N x_{ij}^{2n}, N = \binom{2n+p}{p}.$$

Summing over all $i$ and dividing by $p$ this means

$$1 + (2n/p)\epsilon = \textstyle\sum_1^{Np} y_j^{2n}.$$

**3rd step:** We finally drop the condition imposed on the positive unit $\epsilon$ but want to get back to the previous situation. First of all, by the density of the image of

$H$ in $C(M,\mathbb{R})$ and the second statement in prop. 2, we can find a totally positive unit $\gamma$ such that

$$1 < \hat{\epsilon} \cdot \widehat{\gamma^{2n}} < 1 + 2n/p \quad \text{everywhere on } M.$$

Next define another totally positive unit $\eta$ by $1 + (2n/p) \cdot \eta = \epsilon\gamma^{2n}$. This unit satisfies the condition in the second step. Consequently, $\epsilon\gamma^{2n}$ turns out to be a sum of $Np$ $2n$-th powers. Then the same claim is true for $\epsilon$. After multiplication with the sum $\sum_1^p b_j^{2n}$ we finally get the claim on $p_{2n}$. □

It seems worth emphasizing some of the results so far:

**Corollary 1.** (1) *$p_3 \leq 3, p_d \leq d+1$ in general if $d$ is odd,*

(2) *if $p_2 = 1$, i.e. $K$ is a pythagorean field, then $p_3 \leq 3$ and $p_d \leq d+1$ for all $d$,*

(3) *if $p_2 \leq 2$ then $p_{2n} \leq 2(2n+1)(2n+2)$.*

These bounds for $p_{2n}$ might be correct assymptotically but are definitely too large for small exponents and small quadratic Pythagoras number $p_2$. Progress in those cases came from new ideas published much later in the paper of Choi, Lam, Reznick and Prestel [7]. J. Schmid [16, 17] took up their ideas and presented substantial improvements as will be explained below. Interesting ideas, somewhat similar, are presented in Ellison's paper on Waring's problem for fields [9].

To begin with we note that the proof of the last proposition has shown, under the hypothesis of a finite Pythagoras number, that every totally positive unit of the real holomorphy ring is a sum of $2n$th powers in $K$ for every n. In [4, 1.6] this result was proven without this assumption: $\mathbb{E}_+ \subseteq \bigcap_n \sum K^{2n}$. It was greatly improved by J. Schmid [17, 3.3]. He could prove

**Theorem 1.** *Every totally positive unit of $H$ is a sum of $n$th powers of totally positive units, for every $n \in \mathbb{N}$:*

$$\mathbb{E}_+ = \sum \mathbb{E}_+^n.$$

The original Waring problem had asked whether every natural number can be represented as a bounded number of $n$-th powers of natural numbers. Schmid's theorem can be understood as a qualitative variant of the Waring problem, now posed in the context of general formally real fields: the set of natural numbers is replaced by the set of totally positive units of the real holomorphy rings. Motivated by this interpretation we are led to introduce a **Waring bound** for $\mathbb{E}_+$:

$$w_n := \min\{k \mid \mathbb{E}_+ = \sum_1^k \mathbb{E}_+^n\} \text{ or } w_n = \infty \text{ if no such k exists .}$$

The relationship between the invariants $p_{2n}$ and $w_n$ is not well understood to date. Little is known but it seems very worthwhile comparing these two invariants since one is led to study hidden properties of the fields in question. The discussion in the next section is just one example.

A sample of general results is presented in the following proposition. As indicated we get smaller bounds than predicted by the previous proposition.

**Proposition 4.** (1) *$w_2 \leq p_2 + 1, w_3 \leq 3$,*

(2) *if $p_2 = 1$ then $p_{2n} \leq w_n, w_2 \leq 2$, hence $p_4 \leq 2, p_6 \leq 3$,*

(3) *if $p_2 \leq 2$ then $p_{2n} \leq (n+1)w_n$, in particular $p_4 \leq 9, p_6 \leq 12$.*

*Proof.* (1) The first inequality is proven in [17, 2.4]. To prove the second inequality we apply Richmond's identity from above. Let $r \in \mathbb{E}_+$. In analogy to the argument in the 3rd step in the proof of prop. 3 we find $r \in \mathbb{E}_+$ such that $2/3 < \widehat{3r}/\hat{s}^3 < 1$. Then $t := 3r/s^3 \in \mathbb{E}_+$ is such that all entries in the brackets are totally positive units. Hence $w_3 \leq 3$.
For all other claims we consider the relation

$$s := \sum_i a_i^{2n} = [\sum_i a_i^{2n}/q^n] \cdot q^n \text{ where } q = \sum_i a_i^2.$$

The element in [] is a totally positive unit $\epsilon$ as it is a unit in all real valuation rings and it is a sum of squares. Hence, $\epsilon = \sum_1^{w_n} \epsilon_i^n$ where all $\epsilon_i \in \mathbb{E}_+$.
(2) From $p_2 = 1$ we get that each $\epsilon_i = \eta_i^2$ and that $q = b^2$ for some field elements, hence $s = \sum_1^{w_n} (\eta_i b)^{2n}$. This means $p_{2n} \leq w_n$, and the other statements follow.
(3) In this case we have $\epsilon_i = \eta_1^2 + \eta_2^2, q = a^2 + b^2$. This leads to

$$s = \sum_i [(\eta_1^2 + \eta_2^2)(a^2 + b^2)]^n = \sum_1^{w_n} (a_i^2 + b_i^2)^n$$

since the product of two sums of two squares is a sum of two squares. In the book of Reznick "Sums of powers of real linear forms" we find identities

$$(u^2 + v^2)^n = \sum_1^{n+1} w_i^{2n},$$

cf. [14, 9.5, p.124]. This completes the proof. □

The precise relationship between $w_2$ and $p_2$ is still not clarified. We add a few preliminary remarks.
First of all, it can be proven, using prop. 2, (2), that $w_2 = 1$ implies that the space of real places $M$ is connected.
If $K$ is a pythagorean field, i.e. $p_2 = 1$ then both situation $w_2 = 1$ and $w_2 = 2$ occur: $w_2(\mathbb{R}) = 1$ , whereas the pythagorean closure of $\mathbb{R}(X)$ admits a space of real places highly disconnected, implying $w_2 = 2$ for this field.
The case of $p_2 = 2$ is already very interesting. So far, no field with $w_2 = 3$ has been found. To prove that $w_2 = 2$ is not all easy in the cases where it could be achieved. In [7] the result $w_2 = 2$ was obtained for the rational function field $\mathbb{R}(X)$, in an equivalent form. The method used there was expanded by Schmid in his papers [16, 17] and it was claimed that $w_2 = 2$ for all formally real algebraic function field in one variable over $\mathbb{R}$ [16]. Unfortunately, the claim was based on an insufficient argument, an immediate remedy wasn't visible. Finally the claim turned out to be correct. The only proof known so far uses some topological machinery displayed in the next section. Hence we get $p_4 \leq 6$ for all such fields.

The problem of comparing $w_n$ and $p_{2n}$ stems from the fact that we deal with elements of the real holomorphy ring H in the first instance whereas we have to consider arbitrary field elements in the second case. The first situation allows to apply the representation $\Phi : H \to C(M, \mathbb{R})$ with its well understood properties, see proposition 2. Luckily, this representation can be extended to all of $K$. It is well known that for each $a \in K$ the function

$$\hat{a} : M \to \mathbb{P}^1, \lambda \mapsto \lambda(a) \text{ if } a \in V_\lambda, \lambda \mapsto \infty \text{ otherwise}$$

is continuous. This fact leads to a map

$$\Phi : K \to C(M, \mathbb{P}^1), a \mapsto \hat{a}$$

where we get $\Phi^{-1}(C(M, \mathbb{R})) = H$. In general, the properties of this general map are not really understood. However, in the case of formally real function fields in one variable over $\mathbb{R}$ the image is shown to be dense and this fact is used to derive the equality $w_2 \leq 2$ whence $p_4 \leq 6$ for such fields.

## 3. The Function Field Case

In this section we deal with algebraic function fields $F|\mathbb{R}$ in one variable, in other words with finite extensions of the rational function field $\mathbb{R}(T)$. They are always denoted by $F$.

Real valuation rings in $F$ must contain the base field $\mathbb{R}$. Hence they are discrete valuation rings $V$, contain $\mathbb{R}$ and have residue field $\mathbb{R}$: such valuation rings of $F$ are referred to as real prime divisors of F. The other discrete valuation rings of $F$ which contain $\mathbb{R}$ must have residue field $\mathbb{C}$, they are called complex prime divisors. Both types are referred to as prime divisors. From the general theory of our fields we know that the prime divisors determine the structure of the field. In this sense, the real prime divisors determine the "reality" features we are interested in.

In fact, each real place $\lambda \in M$ gives rise to a real prime divisor $V_\lambda := \{a \in F \mid \lambda(a) \neq \infty\}$, and this assignment constitutes a bijection between $M$ and the set of real prime divisors. Later on, we will identify the real prime divisors with the points on a $C^\infty$ compact curve with finitely many components each one diffeomorphic to $S^1$. This well known topological interpretation allows to apply tools from differential topology which are not available in a purely algebraic setting.

For the remainder of this section we assume $F$ to be formally real, referred to as a ***real function field*** for short. As indicated above we want to study the representation $\Phi : F \to C(M, \mathbb{P}^1), a \mapsto \hat{a}$ where we consider $C(M, \mathbb{P}^1)$ in its compact-open topology. We will pass to a more accessible, still equivalent representation by using stereographic projections.

Let $K$ denote any formally real field and set $S^1(K) = \{(a,b) \in K^2 \mid a^2+b^2 = 1\}$. From the basic properties of the real holomorphy ring we see that the condition $a^2 + b^2 = 1$ implies that $a, b \in H$. Hence, $S^1(K) = S^1(H)$.
In addition, set $S^1 = S^1(\mathbb{R}) \subseteq \mathbb{C}$ and as always $\mathbb{P}^1 = \mathbb{P}^1(\mathbb{R})$.
The stereographic projection

$$st : S^1(K) \to \mathbb{P}^1(K), (a,b) \mapsto a/(1-b) \text{ if } b \neq 1, (0,1) \mapsto \infty$$

is a bijection with inverse map

$$\iota : \mathbb{P}^1(K) \to S^1(K), t \mapsto (\tfrac{2t}{1+t^2}, \tfrac{1-t^2}{1+t^2}) \text{ if } t \neq \infty, \infty \mapsto (0,1).$$

In the case of $K = \mathbb{R}$ both maps $st, \iota$ are homeomorphism. In particular, $\iota$ induces a homeomorphism

$$\iota_* : C(M, \mathbb{P}^1) \to C(M, S^1), f \mapsto \iota \circ f.$$

where the compact-open topology on $C(M, S^1)$ is nothing but the topology of uniform convergence defined by the maximum norm. Using the stereographic projection for $F$ we obtain the following commutative diagram

$$\begin{tikzcd} K \arrow[r, "\Phi"] \arrow[d, "\iota"] \arrow[dr, phantom, "///"] & C(M,\mathbb{P}^1(\mathbb{R})) \arrow[d, "\iota_*"] \\ S^1(H) \arrow[r, "\Psi"] & C(M,S^1) \end{tikzcd}$$

The map $\Psi : S^1(H) \to C(M,S^1)$ is given by the assignment $(a,b) \mapsto \hat{a} + i\hat{b}$ where $i = \sqrt{-1}$.

In this diagram the map $\iota$ is injective with image $S^1(H) \setminus \{(0,1)\}$ and $\iota_*$ is a homeomorphism as mentioned already. In particular, we see that the image $\Phi(K)$, respectively its closure, is naturally identified with the image of $S^1(H) \setminus \{(0,1)\}$, respectively its closure, under $\Psi$. Passing to the closure of $im\, \Phi$ allows us to "catch" the missing function $\Psi(0,1)$ which is the constant function sending every point to $(0,1)$. One easily verifies that the constant functions $f_n = (2n/(n^2+1), (n^2-1)/(n^2+1))$ converge to $\Psi(0,1)$. Hence, the closure of $\Phi(K)$ is identified via $\iota_*$ with the closure of $\Psi(S^1(H))$.

Now follows the crucial fact for our real function field $F$:

**Proposition 5.** *If the image of $S^1(H) \xrightarrow{\Psi} C(M,S^1)$ is dense then $w_2(F) \leq 2$.*

*Proof.* Consider any $\epsilon \in \mathbb{E}_+$. Since $p_2(F) = 2$ by Witt we find $a, b \in F$ such that $a^2 + b^2 = \epsilon$. Necessarily $a, b \in H$. In view of $\hat{\epsilon} > 0$ everywhere on $M$ we can pass to the continuous function

$$f = \frac{1}{\sqrt{\hat{\epsilon}}}(\hat{a} + i\hat{b}) : M \to S^1.$$

By assumption $f$ can be approximated as closely as needed by a function $\Psi(u,v)$, $u^2 + v^2 = 1$. To simplify notation we are using the symbol $\approx$ to state that the approximation is good enough for the argument to follow.
In our situation, $f \cdot (\hat{u} - i\hat{v}) \approx 1$. After multiplying we obtain

$$(\widehat{au+vb}) + i(\widehat{bu-av}) \approx \sqrt{\hat{\epsilon}}.$$

We set $c = au + vb, d = bu - av$ then $c^2 + d^2 = a^2 + b^2 = \epsilon, \hat{c} \approx \sqrt{\hat{\epsilon}}, \hat{d} \approx 0$. The element $c$ turns out to a totally positive unit already by prop. 2, the same cannot be said in the case of $d$. However $\hat{d}$ can be chosen as small as needed to guarantee that both elements $c+d, c-d$ are totally positive units, again by prop. 2. Hence,

$$\epsilon = (\frac{c+d}{\sqrt{2}})^2 + (\frac{c-d}{\sqrt{2}})^2$$

and we get that $\epsilon \in \mathbb{E}_+^2 + \mathbb{E}_+^2$. □

The classical Stone-Weierstraß approximation theorem is concerned with dense subsets of $C(X,\mathbb{R})$, $X$ any compact space. A similar result for $C(X,S^1)$, as general as the Stone-Weierstraß approximation theorem, is not known to the author. Yet, there are favorable conditions under which the closure of certain subsets in $C(X,S^1)$ can be described.

**Proposition 6.** *Let $X$ be a compact space and $G$ a subgroup of $C(X,S^1)$. If its closure $\bar{G}$ (which is a subgroup again) contains all non-surjective functions $X \to S^1$ then $\bar{G}$ is the union of all homotopy classes of elements in $G$.*

As a matter of fact, the image $G := \Psi(S^1(H))$ satisfies the conditions above. So it remains to show that every continuous function $M \to S^1$ is homotopic to functions of the type $\Psi(u,v) = \hat{u} + i\hat{v}$ where $u, v \in H, u^2 + v^2 = 1$. This amounts to produce sufficiently many functions originating from $F$ for which it is possible to determine the homotopy type. As to the first requirement we can apply some of the fundamental results of Witt in [18]. Secondly the homotopy type is determined by the (topological) degree, e.g. cf. [10, chapter 5]. At the end we succeed in proving the density of the image of $\Phi$. Details of this approach will be published somewhere.

Instead, we will be following a different route by appealing to the results in [2, Chapter 13] on polynomial or regular mappings with values in spheres. That approach is based on the study of algebraic versus topological vector bundles on real algebraic varieties. We focus on this method.

Either approach requires to substitutes the space of real places by a smooth compact real algebraic curve $\gamma$. This curve is a compact $C^\infty$ curve, hence by a general theorem it has finitely many connected components each of which is diffeomorphic to $S^1$. We sketch the transition from the space $M$ to this curve $\gamma$.

The function field $F$ is the quotient field of the real holomorphy ring $H$. Hence, we can find an affine algebra $A \subseteq H$ (over $\mathbb{R}$) the quotient field of which is $F$ as well. Consider the integral closure $B$ of $A$ in $F$. As a Prüfer ring the ring $H$ is integrally closed in $F$, hence $B \subseteq H$. From commutative algebra we know that $B$ is an affine algebra, integrally closed of dimension 1. In particular, $B$ turns out to be Dedekind domain what implies that $H$, as an overring of $B$, is a Dedekind domain as well.

It is a general fact that the valuation rings of a formally real field containing its real holomorphy ring are exactly its real valuation rings. In our situation, the valuation overrings of $H$ are exactly the real prime divisors of $F$. Let $V$ be such a real prime divisor, $\mathfrak{m}$ its maximal ideal. The intersection ideals $\mathfrak{p} := \mathfrak{m} \cap H$ and $\mathfrak{q} := \mathfrak{m} \cap B$ are maximal ideals of either rings. The localizations $B_\mathfrak{q}$ is contained in the localization $H_\mathfrak{p}$. Since both rings are discrete valuation rings they have to be equal. Conversely, let $\mathfrak{q} \subseteq B$ be any maximal ideal with residue field $\mathbb{R}$ then the local ring $B_\mathfrak{q}$ has to be a real prime divisor, hence it contains $H$.

Now, set $\Gamma = SpecB, \gamma := \Gamma(\mathbb{R})$ the set of closed real points. It is a smooth affine curve with function field $F$ and the following properties:

(1) The center map $c : M \to \gamma, \lambda \mapsto ker(\lambda_{|B})$ is bijective,

(2) $H = \bigcap_{x\in\gamma} \mathcal{O}_x(\gamma) = B_{1+\sum B^2}$. Cf. [2, 4.1] to prove the last equality. As a result we obtain that each $a \in H$ can be evaluated at any point $x \in \gamma$, hence gives rise to a function $\tilde{a} : \gamma \to \mathbb{R}$.

(3) For each $a \in H, \lambda \in M$ we have $\tilde{a}(c(\lambda)) = \hat{a}(\lambda)$.
Imposing the strong topology on $\gamma$ with respect to the functions $\tilde{a} \in H$ we find that the center map continuous bijection. This finally leads to :

(4) $\gamma$ is a compact space, the center map $c$ a homeomorphism.
The final step consists of realizing $\gamma$ as a curve in some $\mathbb{R}^N$. To this end choose a presentation $B \simeq \mathbb{R}[T_1, \ldots, T_N]/I$ and consider the set of zeros $X := \mathcal{Z}(I)$ in $\mathbb{R}^N$.

(5) $X$ is a compact $C^\infty$ curve and naturally homeomorphic to $\gamma$ and $M$, the latter via the center map.

(6) $F$ is interpreted as the field $\mathcal{F}$ of rational functions on $X$ which induces an isomorphism between $H$ and $\mathcal{R}(X)$, the ring of rational functions which are defined on $X$ [2, 3.2].

Consequently, the representation $\Phi : F \to C(M, \mathbb{P}^1)$ can be substituted by the evaluation map $ev : \mathcal{F} \to C(X, \mathbb{P}^1)$. $S^1(H)$ is mapped onto $\{(f,g) \mid f, g \in \mathcal{R}(X), f^2 + g^2 = 1\}$ which is nothing but the set of regular functions $\mathcal{R}(X, S^1)$. So, at the end we are concerned with the question whether

$$\mathcal{R}(X, S^1) \text{ is a dense subset of } C(X, S^1)$$

which is equivalent to our original question whether the image $\Psi(S^1(H))$ is dense in $C(M, S^1)$.

In [2, 13.3.3] density is shown with respect to $C^\infty(M, S^1)$ which, in turn, is dense in $C(M, S^1)$ by [10, 2.2, p.44]. Therefore we have proven the main theorem in this section:

**Theorem 2.** *The image of* $S^1(H) \xrightarrow{\Psi} C(M, S^1)$ *is dense.*

As said before this implies $w_2 \leq 2$. It is not hard to see that in real function fields always "$w_2 > 1$" the proof of which we ommit. Anyway, we want to state

**Theorem 3.** *Let* $F$ *be a real function field in one variable over* $\mathbb{R}$ *then* $w_2 = 2$ *and* $p_2 = 2, p_3 \leq 3, p_4 \leq 6$.

So far it is known that $p_3(\mathbb{R}(X)) = 3$ and $p_4(\mathbb{R}(X)) \geq 3$. To conjecture that $p_3 = 3, p_4 \geq 3$ holds for all real function fields seems reasonable. As to the value of $p_4$, however, the author is not in the position of offering a substantiated conjecture.

It may seem that methods and results from differential topology are too big a shot for getting the "innocent-looking" result $w_2 \leq 2$. However one can prove the following result for formally real fields with $H$ a Dedekind domain:
the two statements

(1) $w_2 \leq 2$,
(2) $p_2 \leq 2$ and the map $S^1(H) \to C(M, S^1)$ has dense image,

are equivalent.
At this moment no simpler proof of the density result is known.

## 4. UNIQUE FACTORIZATION OF SUMS OF SQUARES

Apart from general remarks we still assume that we are dealing with a formally algebraic function in one variable over $\mathbb{R}$, denoted by $F$.

It is known that Dedekind domains rings are not unique factorization domains unless they are principal ideal domains, [6, VII,3.1]. So, real holomorphy rings are rarely unique factorization domains.
In the present situation one can even show that no maximal ideal of $H$ is principal. We sketch the proof and start by transferring everything to the geometric model $\gamma$ and interpreting the elements of $F$ as rational functions $\gamma \to \mathbb{P}^1$. The next ingredient is the decomposition of $\gamma$ into finitely many circles $S^1$. An element $a \in H, a \neq 0$ then gives rise to continuous functions $S^1 \to \mathbb{R}$. Such a function can be understood, using a properly chosen stereographic projection, as a function $\mathbb{R} \to \mathbb{R}$ with existing limits $\lim_{x \to +\infty}$ and $\lim_{x \to -\infty}$ which are equal and not zero. If $a$ were a generator of a maximal ideal $\mathfrak{m}$ then we would end up with a continuous function as above which changes sign but admits just one zero. That contradicts the intermediate value theorem.
On the other hand it is known that any ideal of a Dedekind domain domain can be generated by two elements. In our situation of real holomorphy rings this can be

seen as follows. Pick an ideal $\mathfrak{a}$. Then $\mathfrak{a}^2 = (a_1^2 + \cdots + a_r^2), a_1, \ldots, a_r \in \mathfrak{a}$. Using $p_2 \leq 2$ we get $\mathfrak{a}^2 = (a^2 + b^2) = (a, b)^2$ implying $\mathfrak{a} = (a, b)$.

So, unique factorization is not valid in $H$. However, if we restrict attention to the non-zero sums of squares in $H$ the situation is quite different: we do have unique factorization and a complete list of prime elements is given by the sum of squares-generators of the squares of the maximal ideals of H as we will see in the sequel.

We set

$$\Sigma := \sum H^2 \setminus \{0\}. \text{ Note that } H \cap \sum K^2 = \Sigma \cup \{0\}.$$

Given $f, g \in \Sigma$ we write, as usual $f|g$ iff there is $h \in \Sigma$ with $fh = g$. Note that $f|g$ is equivalent to saying that $f$ divides $g$ in $H$. This follows from the fact $\sum H^2 = H \cap \sum K^2$. Hence, in terms of $H$-ideals, $f|g$ is equivalent to
$(g) = gH \subseteq (f)$.
$f$ and $g$ are called associate, denoted by $f \sim g$, if $f|g$ and $g|f$, equivalently, if $g = \epsilon f$ for some unit $\epsilon \in \mathbb{E}_+$.
An element $f$ is called **irreducible** if it is not a unit and any decomposition $f = gh$ enforces that one of the factors is a totally positive unit. $f$ is called a **prime element**, or prime for short, if $f$ is not a unit and whenever $f|gh$ we get $f|g$ or $f|h$. Clearly, prime elements are irreducible. To prove the converse we start with the following lemma.

**Lemma 1.** *Let $f, g \in \Sigma$, $f$ irreducible. Then the following statements are equivalent:*

(1) $f \nmid g$,
(2) $(f, g) = H$,
(3) $f + g \in \mathbb{E}_+$.

*Proof.* We start with the equation $f = \frac{f}{f+g}(f+g)$ where all entries are in $\sum H^2$. If $f/(f+g)$ were a unit, a forteriori, a totally positive unit, then $f|g$ would follow. So, the implication (1) $\Rightarrow$ (3) is proven. The equation also implies $(f, g) = (f+g)$ for any two sums of squares in $H$, hence (3) $\Rightarrow$ (2). The remaining implication is obvious. ☐

**Lemma 2.** *Let $f \in \Sigma$ and $\mathfrak{m}$ a maximal ideal of $H$. If $f \in \mathfrak{m}$ then even $f \in \mathfrak{m}^2$.*

*Proof.* We know that $\mathfrak{m}$ is a real prime ideal. Hence $f = \sum a_i^2 \in \mathfrak{m}$ implies that all $a_i \in \mathfrak{m}$. This gives the claim. ☐

**Proposition 7.** *Every irreducible element in $\Sigma$ is a prime element.*
*Given $p \in \Sigma$ the following statements are equivalent:*

(1) *$p$ is a prime element,*
(2) *$(p) = \mathfrak{m}^2$ for some maximal ideal $\mathfrak{m}$ of $H$.*

*Proof.* Assume $p$ to be an irreducible element and that $p|fg$ . If $p \nmid f$ then by the first lemma: $p + f \in \mathbb{E}_+$, say $p + f = \epsilon$. Multiplying with $g$ yields that $p$ divides $g$. Hence, $p$ is a prime element.
Let's continue with a prime element $p$. Then the $H$-ideal $(p)$ is contained in a maximal ideal $\mathfrak{m}$ of $H$. We claim that $(p) = \mathfrak{m}^2$. First of all, by the second lemma, $(p) \subseteq \mathfrak{m}^2$ . To prove equality, pick any $f \in \mathfrak{m}$ and assume $f^2 \notin (p)$. Then, by the

lemma above, $p + f^2$ is a unit: a contradiction. Therefore, for any $f \in \mathfrak{m}$ we get $f^2 \in (p)$. Choosing any other $g \in \mathfrak{m}$ we have that $f^2, g^2, (f+g)^2 \in (p)$ enforcing $fg \in (p)$. Thus, $\mathfrak{m}^2 \subseteq (p)$ and equality results.
Finally, consider an element $p \in \Sigma$ subject to $(p) = \mathfrak{m}^2$ as stated. Assume $p|fg$ . Then $fg \in \mathfrak{m}$ implies, say, $f \in \mathfrak{m}$. The second lemma states that $f \in \mathfrak{m}^2$ as $f$ is a sum of squares. Hence, $f \in (p)$ meaning $p|f$ in $\Sigma$. ☐

We can now fix a complete list $\mathcal{P}$ of pairwise non-associate prime elements:

$$\mathcal{P} = \{p_\mathfrak{m} \mid p_\mathfrak{m} \in \Sigma, (p_\mathfrak{m}) = \mathfrak{m}^2, \mathfrak{m} \text{ a maximal ideal of } H\}.$$

In the following we will make use of the discrete valuations attached to the real prime divisors. The real prime divisors are just the localizations $H_\mathfrak{m}, \mathfrak{m}$ running through the maximal ideals of $H$. The discrete valuation with value group $\mathbb{Z}$ attached to $\mathfrak{m}$ is denoted by $v_\mathfrak{m}$. They are referred to as the real valuations of $F$. We transfer from the general theory of Dedekind domains that every non-zero principal fractional ideal $(a) = aH$ admits a presentation

$$(a) = \prod_\mathfrak{m} \mathfrak{m}^{v_\mathfrak{m}(a)}.$$

Using this fact the prime elements of $\Sigma$ are characterized by the following properties:

$v(p) = 2$ for a unique real valuation $v$ and $w(p) = 0$ for all real valuations $w \neq v$.

In the sequel we fix a complete system of prime elements $\mathcal{P} := \{p\}$ and set $v_p = v_\mathfrak{m}$ if $(p) = \mathfrak{m}^2$.

**Theorem 4.** *Every $a \in \Sigma$ admits a unique presentation $a = \epsilon \prod_{p \in \mathcal{P}} p^{e_p}$ where $\epsilon \in \mathbb{E}_+, e_p \in \mathbb{N} \cup \{0\}$, all but finitely many $e_p = 0, 2e_p = v_p(a)$.*

*Proof.* Any such valuations is real valuations, hence the exponents are even natural numbers: $v_\mathfrak{m} = 2e_\mathfrak{m}$ This leads to $f = \epsilon \prod_\mathfrak{m} p_\mathfrak{m}^{e_\mathfrak{m}} \epsilon \in \mathbb{E}_+$. Such a decomposition is unique since it can be interpreted as a decomposition of the principal ideal $(f)$ as a product of powers of maximal ideal which is unique by the basic property of a Dedekind domain. ☐

Now consider any non-zero element $a \in \sum K^2$. It can be written as $a = f/g, f, g \in \sum H^2$. So we arrive at a presentation

$a = \epsilon \prod_{p \in \mathcal{P}} p^{e_p}$ where $\epsilon \in \mathbb{E}_+, e_p \in \mathbb{Z}$, all but finitely many $e_p = 0, 2e_p = v_p(a)$.

In order to arrive at an explicit factorization one has to come up with a list of explicit generators of the ideals $\mathfrak{m}^2$. There are cases where this has been achieved. In the PhD thesis of Natea Hunde [13] the cases of rational and elliptic function fields have been dealt with explicitly. In particular, for $\mathbb{R}(X)$ the maximal ideals of $H$ are indexed by the points $\alpha \in \mathbb{P}^1$:

$$\mathfrak{m}_\alpha = \left(\frac{X-\alpha}{1+X^2}, \frac{(X-\alpha)^2}{1+X^2}\right) \text{ if } \alpha \in \mathbb{R}, \mathfrak{m}_\infty = \left(\frac{1}{1+X^2}, \frac{X}{1+X^2}\right).$$

The square of a maximal ideal is generated by the sum of the squares of its generators which can be multiplied by a totally positive unit if convenient. In this way, we get for $\mathbb{R}(X)$ the following complete systems of prime elements:

$$\frac{(X-\alpha)^2}{1+X^2} \text{ where } \alpha \in \mathbb{R} \text{ and } \frac{1}{1+X^2}.$$

Surprisingly, the multiplicative presentation above can be turned into an additive presentation of $a$ as a sum of mixed powers. A general theory of sums of mixed powers is due to R. Berr [5].

**Theorem 5.** *Let* $a \neq 0, a \in \sum K^2, a = \epsilon \prod_{p\in\mathcal{P}} p^{e_p}$. *Then*

$$a = \eta \sum_{p\in\mathcal{P}, e_p\neq 0} b_p^{e_p}$$

*for some* $\eta \in \mathbb{E}_+, b_p \neq 0, b_p \in \sum K^2, v_p(b_p) = 2.$

*Proof.* We consider only those prime elements with $e_p \neq 0$. So, $a = \epsilon \prod_{e_p\neq 0} p^{e_p}$. For each such $p$ choose $l_p \in \mathbb{Z}$ with $l_p e_p \geq e_q$ for all other prime elements $q$ occurring. Setting $b_p := p \prod_{q\neq p} q^{l_p}$ we calculate

$$a^{-1} * b_p^{e_p} = \epsilon^{-1} \prod_{q\neq p} q^{l_p e_p - e_q} =: c_p.$$

For each $p$ we get: $b_p \in \sum K^2, v_p(b_p) = 2, c_p \in \sum H^2$. Moreover, $(\{c_p \mid e_p \neq 0\}) = H$. Using the proposition 1,(3) above this leads to $\sum_p c_p = \omega \in \mathbb{E}_+$ whence the claim follows. ☐

**Corollary 2.** *Let* $a \neq 0, a \in \sum K^2$ *then* $a \in \sum_{p:v_p(a)\neq 0}(\sum K^{|v_p(a)|})$.

*Proof.* We are using the additive presentation of the element $a$ from above. Using Hilbert's identity one gets that $b_p^{e_p} \in \sum K^{2|e_p|}$. Note that the multiplicative inverse of a sum of squares is again a sum of squares and that $2e_p = v_p(a)$. Furthermore, the unit $\eta$ is a sum of $2n$-th powers for every $n$ since $\mathbb{E}_+ \subseteq \cap_n \sum K^{2n}$. This completes the proof. ☐

## REFERENCES

[1] E.Artin, *Über die Zerlegung definiter Funktionen in Quadrate*, Hamb. Abh. 5 (1927), 225-231

[2] J.Bochnak, M.Coste, M.-F.Roy, *Real Algebraic Geometry*,Springer-Verlag 1998

[3] E. Becker, *Valuations and real places in the theory of formally real fields*,in: Geometrie algebrique reelle et formes quadratiques, Lecture Notes in Math. No. 959 (1982), 1-40.

[4] E.Becker, *The real holomorphy ring and sums of* $2n$*-th powers*, Lecture Notes in Mathematics 959 (1982), 139-181

[5] R. Berr, *Sums of mixed powers in fields and orderings of prescribed level*, Math. Z. 210 (1992), 513 - 528

[6] N. Bourbaki, *Commutative Algebra*,Hermannn Publisher 1972

[7] M. D. Choi, T. Y. Lam, A. Prestel, B. Reznick, *Sums of 2m-th powers of rational functions in one variable over real closed fields*, Math. Z. 221 (1996), 93 - 112

[8] D.Dubois, *Real commutative algebra I. Places*, Revista Matematica Hispana-Americana 39 (1979), 57-65

[9] W.J. Ellison *Waring's problem for fields*, arxiv.org/pdf/1303.4818 (2013)

[10] M.W. Hirsch *Differential topology*, Springer-Verlag 1976

[11] J.R. Joly *Sommes des puissances d-iemes dans un anneau commutatif*, Acta Arithm. 17 (1970), 37-114

[12] T.Y. Lam, *Orderings, Valuations, and Quadratic Forms*, CBMS Regional Conf. Series in Math.,Vol. 52, Amer. Math. Soc., Providence, RI, 1983

[13] Natea Hunde, *The real holomorphy ring and sums of powers in rational and elliptic function fields over the field of real numbers*, Addis Ababa University, Ethiopia, 2017

[14] B. Reznick, *On Sums of Even Powers of Real Linear Forms*, Mem. AMS 463, 1992.

[15] H. W. Schülting, *On real places of a field and their holomorphy ring*, Comm. Algebra 10 (1982), 1239 - 1284

[16] J. Schmid, *Sums of fourth powers of real algebraic functions*, Manuscripta Math. 83 (1994), 361-364

[17] J. Schmid, *On totally positive units of real holomorphy rings*, Israel J. of Math. 85 (1994),339-350

[18] E. Witt, *Zerlegung algebraischer Funktionen in Quadrate, Schiefkörper über reellen Funktionenkörpern*, J. reine angew, Mathematik 171 (1939), 4-11

# Smooth parameterization in o-minimal structures

Margaret E. M. Thomas

The Fields Institute for Research in Mathematical Sciences*
Zukunftskolleg and Department of Mathematics and Statistics, University of Konstanz†
margaret.thomas@wolfson.oxon.org

## 1 Introduction

The counting theorem of Pila and Wilkie [21] opened up one of the most important developments in model theory in recent years. The theorem provides a bound on the density of rational points for sets definable in o-minimal expansions of the real field, a result which has had several stunning number-theoretic applications due to its incorporation into the so-called "Pila–Zannier method" for solving special points problems. This method was established initially to provide a new proof of the Manin–Mumford Conjecture [22] and since developed and exploited by Pila, Tsimerman and others to give, amongst other results, the first unconditional proofs of cases of the André-Oort Conjecture (e.g. [19, 20, 27]).

Central to the proof of the Pila–Wilkie Theorem is an o-minimal version of Yomdin-Gromov parameterization [29, 28, 10], one type of 'smooth parameterization'. This technique in general involves decomposing sets into the images of finitely many functions which have controlled higher-order derivatives. Originally introduced to study topological entropy and volume growth in smooth dynamics, it also has other important geometric and arithmetical consequences. We will discuss several types of smooth parameterization and their connection to both o-minimality and questions of diophantine geometry.

## 2 Yomdin–Gromov $r$-parameterization

Let $n$ be a non-negative integer and let $X \subseteq [0,1]^n$. In the original result of Yomdin and Gromov, $X \subseteq [0,1]^n \subseteq \mathbb{R}^n$ is assumed to be a semi-algebraic set (i.e. a set definable in the real field $\overline{\mathbb{R}} := \langle \mathbb{R}, +, \cdot, 0, 1, < \rangle$ [25]); in the result of Pila and Wilkie, $X \subseteq [0,1]^n \subseteq R^n$ is more generally assumed to be a set definable in any o-minimal expansion of a real closed field $\langle R, +, \cdot, 0, 1, < \rangle$.

---

*The Fields Institute for Research in Mathematical Sciences, 222 College Street, Second Floor, Toronto, Ontario, M5T 3J1, Canada

†Zukunftskolleg and Department of Mathematics and Statistics, University of Konstanz, Fach 216, Universitätsstraße 10, D-78457 Konstanz, Germany

Let $r$ be a non-negative integer. The results of Yomdin–Gromov/Pila–Wilkie tell us that, under the appropriate assumption on $X$, there exists a finite set $\Phi_r$ of $C^r$ maps $\phi\colon (0,1)^{\dim X} \to [0,1]^n$ such that

(i) $X = \bigcup_{\phi\in\Phi_r} \operatorname{Im}(\phi)$;

(ii) $\left|\left|\phi^{(\alpha)}\right|\right|_\infty \leq 1$, for all $|\alpha| \leq r$.

We call a set $\Phi_r$ satisfying (i) and (ii) an *$r$-parameterization* of $X$.

The central application of this result to diophantine geometry is the following theorem of Pila and Wilkie, which captures a sense in which the number of rational points lying on a set conveys information about the algebraic nature of the set.

For a rational tuple $\overline{q} = (\frac{a_1}{b_1}, \dots, \frac{a_n}{b_n}) \in \mathbb{Q}^n$, with $\gcd(a_i, b_i) = 1$, for all $1 \leq i \leq n$, the height of $\overline{q}$ is defined to be $ht(\overline{q}) = \max_{1\leq i\leq n}\{|a_i|, |b_i|\}$. For any set $Y \subseteq \mathbb{R}^n$, define $Y(\mathbb{Q}, H) := \{\overline{q} \in \mathbb{Q}^n \mid ht(\overline{q}) \leq H\}$. Define $Y^{\text{alg}}$ to be the union of all infinite, semi-algebraic subsets of $Y$, and define $Y^{\text{trans}} := Y \setminus Y^{\text{alg}}$.

**Pila–Wilkie Theorem.** *Let $X \subseteq [0,1]^n$ be definable in an o-minimal expansion of the real field $\overline{\mathbb{R}} := \langle \mathbb{R}, +, \cdot, 0, 1, < \rangle$. Let $\epsilon > 0$. There exists a positive constant $c = c(X, \epsilon)$ such that, for all $H \geq 1$, $\#X^{trans}(\mathbb{Q}, H) \leq cH^\epsilon$.*

Thus, if such a definable set $X$ contains more than $H^\epsilon$ many rational points of height at most $H$, for some $\epsilon > 0$, then $X$ contains an infinite, connected, semi-algebraic set, and in fact contains a semi-algebraic curve in that case.

There are three central ingredients to the proof of this theorem and other similar results of this kind, which we sketch roughly. The procedure is made more involved by the fact that, in order to obtain an inductive argument (on the dimension of $X$), it is necessary to work rather in certain (definable) families of sets, but for simplicity we present the ideas in terms of a single set $X$.

1. Obtain a suitable smooth parameterization for $X$ with sufficient control over the number of parameterizing functions required and over any parameters involved (in this case, an $r$-parameterization for a suitable $r = r(\epsilon)$). This allows us to reduce to the case of sets which are the images of parameterizing functions.

2. Realise the points $X^{\text{trans}}(\mathbb{Q}, H)$ being counted as zeros of finitely many polynomials of a suitable degree $d$ (in this case $d$ is given in terms of $\epsilon$), where we have sufficient control over the number of polynomials needed in terms of $d$ and $H$ (in this case, this results in $O(H^\epsilon)$-many).

3. Obtain a so-called zero estimate, a bound on the size of sets of the form $X^{\text{trans}} \cap V(p)$, where $V(p)$ is the zero set of a polynomial $p$ of degree $d$ (in this case, a bound of size $O(H^{\epsilon})$).

The proof doesn't provide a method for computing the constant $c$ in the statement of the Pila–Wilkie Theorem in terms of $\epsilon$ and some definition of $X$. Indeed, at the level of generality of sets definable in o-minimal expansions of the real field, such an effective constant cannot be obtained; this is not even possible for the graphs of all one-variable, transcendental, restricted analytic functions. However, in certain cases, say when $X$ can be defined using functions satisfying some reasonable algebraic differential equations, the question makes sense, and indeed might be interesting in view of the many applications of the Pila-Wilkie Theorem to diophantine geometry.

We present here a first result in that direction, which is for surfaces implicitly defined from restricted Pfaffian functions. This is an effective case of Pila's earlier result giving this subpolynomial bound for subanalytic surfaces [16, 17], a precursor to the Pila–Wilkie Theorem. An analytic function $f$ on an open subset of $\mathbb{R}^n$ is Pfaffian if it satisfies a triangular system of polynomial differential equations. Pfaffian functions have a natural measure of complexity given by the dimension of the open set, the number of equations in the system and the degrees of the polynomials involved. (For precise definitions the reader is referred to [14, 12].) This allows us to make the following statement.

**Theorem 1** ([12]). *Let $a$, $b$ be real numbers such that $0 \leq a < b \leq 1$ and let $B$ and $\epsilon$ be positive real numbers. Suppose that $f, g\colon (a,b) \to (0,1)$ are functions implicitly defined from Pfaffian functions of complexity at most $B$ and that $f < g$. Suppose further that $F\colon (f,g)_{(a,b)} \to \mathbb{R}$ is also implicitly defined from Pfaffian functions of complexity at most $B$, where $(f,g)_{(a,b)}$ denotes the cell $\{(x,y) \in (a,b) \times (0,1) \mid f(x) < y < g(x)\}$. There exists a positive real number $c$, bounded effectively in $B$ and $\epsilon$, such that, for all $H \geq 1$, $\#graph(F)^{trans}(\mathbb{Q}, H) \leq cH^{\epsilon}$.*

Our proof follows the same structure as the proof of the Pila-Wilkie Theorem. Consider the three steps outlined above. Step 2 can be seen to be effective from the original proof. For effectiveness in Step 3, we appeal to an earlier result of [13] in which we obtain in our setting an effective zero estimate of size $O((\log H)^{\gamma})$, for some effective positive constant $\gamma = \gamma(B)$, which is clearly more than sufficient here.

Thus, in order to obtain an effective constant $c$, it remains to consider Step 1, $r$-parameterization. Here we cannot appeal directly to the o-minimal version of $r$-parameterization proved by Pila and Wilkie, for it involves the use of the (ineffective) compactness theorem (of first-order logic). Our main contribution is therefore an effective version of this $r$-parameterization result in the setting of our theorem. To prove this, we first prove a uniform parameterization result for families of curves, where the base of the family is an interval in the real line. The setting of functions implicitly defined from Pfaffian functions is sufficiently general that we can remain within it when carrying out the induction required. For more details, see [12].

An alternative proof of $r$-parameterization for subanalytic families was recently given using complex analytic methods by Binyamini and Novikov [2]. During the writing of this note, the arXiv preprint [1] appeared, which makes use of this alternative approach to obtain effective Pila–Wilkie bounds for Noetherian varieties. While the setting of [1] encompasses that of the Theorem 1, the constant obtained in our case is more uniform, in that it depends only on the complexity of the Pfaffian functions involved in defining $f$ (and not on any other input data, such as the coefficients of the polynomials appearing in the system of differential equations defining them).

Beyond identifying cases in which one can obtain $r$-parameterization in an effective way (which includes an effective bound on the number of parameterizing functions required), another refinement which can be sought with the aim of improving diophantine applications is an explicit uniform bound on the number of functions required for $r$-parameterizations across a family of sets. This is a feature of the original result of Yomdin and Gromov, namely, that the number of parameterizing functions required for an $r$-parameterization of a given semi-algebraic set $X \subseteq [0,1]^n$ is bounded in terms of $r$, $n$, the dimension of $X$ and the degrees of the polynomials involved in describing $X$ (see [4]). This idea has been extended in [8] to certain reducts of $\mathbb{R}_{\text{an}}$ and of the expansion of $\mathbb{R}_{\text{an}}$ by power functions; in fact these methods improve the bound known the semi-algebraic case to one which is polynomial in $r$, settling an open conjecture about entropy in smooth dynamics [30, 5].

## 3 Mild parameterization

A set $X \subseteq [0,1]^n$ has *mild parameterization* if there exists a finite set $\Psi$ of $C^\infty$ maps $\psi\colon (0,1)^{\dim X} \to [0,1]^n$ such that

(i) $X = \bigcup_{\psi \in \Psi} \operatorname{Im}(\psi)$;

(ii) there exist constants $B > 0$, $C \geq 0$, such that $\left\|\psi^{(\alpha)}\right\|_\infty \leq \alpha!(B\,|\alpha|)^{C|\alpha|}$, for all $|\alpha| \leq r$.

Functions which are $C^\infty$ and satisfy condition (ii) are said to be *mild*. Pila introduced mild parameterization as part of ongoing work towards improving the bound of the Pila–Wilkie Theorem from subpolynomial to polylogarithmic. Such an improvement is not possible in the full generality of o-minimal expansions of the real field, as demonstrated by counter-examples due to Pila [15]. However, Wilkie has conjectured that such an improvement should hold for sets definable in the real exponential field $\mathbb{R}_{\mathrm{exp}} := \langle \overline{\mathbb{R}}, \exp\rangle$.

**Wilkie's Conjecture.** *Let $X \subseteq [0,1]^n$ be definable in $\mathbb{R}_{\mathrm{exp}}$. There exist positive constants $c = c(X)$ and $\gamma = \gamma(X)$ such that, for all $H \geq 1$, $\# X^{trans}(\mathbb{Q}, H) \leq c(\log H)^\gamma$.*

To date, this conjecture has been established for curves and for surfaces with mild parameterization [13], for which certain explicit examples are known [18, 6]. In addition, several results establishing the polylogarithmic bound for other examples of definable sets are now known, for example for $\mathbb{R}_{\mathrm{an}}$-definable Pfaffian surfaces [13, 8] or sets definable in the expansion of the real field by restricted elementary functions [3].

The connection to mild parameterization can be seen by considering the proof strategy outlined above. In [18], Pila proved a result which could provide an analogue to Step 2 in this context, namely that if $X \subseteq [0,1]^n$ has a mild parameterization, then $X(\mathbb{Q}, H)$ lies in the union of $O((\log H)^{\gamma_1})$ zero sets of polynomials of degree $O((\log H)^{\gamma_2})$, for constants $\gamma_1 = \gamma_1(X)$, $\gamma_2 = \gamma_2(X)$. In order to make use of this fact, however, it remains to establish mild parameterization in a sufficiently uniform way (in order to carry out the induction) in settings in which we may also obtain Step 3.

Let us summarise what is known about mild parameterization for sets definable in o-minimal structures. We say that a structure *has (definable) mild parameterization* if all of its definable sets have a mild parameterization (given by parameterizing functions definable in the same structure).

On account of uniformization, the structure $\mathbb{R}_{\mathrm{an}}$ has mild parameterization, but it moreover can be seen to have definable mild parameterization [11]; this fact also holds true for any reduct of $\mathbb{R}_{\mathrm{an}}$ by the same argument. Indeed, if one considers the expansions of the real field by the quasi-analytic classes $\mathcal{C}$ of restricted $C^\infty$ functions studied by Rolin, Speissegger and Wilkie [24], expansions which they proved to be o-minimal and polynomially bounded, then we showed that all definable sets in such an expansion $\mathbb{R}_{\mathcal{C}}$ have a "$\mathcal{C}$-parameterization"; this has the natural definition in our context as a finite covering for which the parameterizing functions lie in the class $\mathcal{C}$. One instance of such a quasi-analytic class $\mathcal{C}$ is the class of restricted analytic functions; another is the class of all restricted functions definable in a fixed o-minimal, polynomially bounded structure. Since restricted analytic functions are mild, the aforementioned results of [11] follow.

More recently, Rolin and Servi considered expansions of the real field by generalised quasi-analytic classes $\mathcal{A}$, likewise proving that they are o-minimal and polynomially bounded [23]. In addition, they established a "parametrisation theorem" for sets definable in such structures using functions from the class $\mathcal{A}$, which can be adapted to form what may be called an "$\mathcal{A}$-parameterization" in our setting. Consequently, classes $\mathcal{A}$ consisting of mild functions also give rise to o-minimal structures with definable mild parameterization. One new example given in this way, settling a question of [11], is $\mathbb{R}_G$, the expansion of the real field by a certain class of Gevrey functions, a structure which was shown to be o-minimal (and polynomially bounded) by van den Dries and Speissegger [9].

Unfortunately, there is no hope for definable mild parameterization to be obtained for all o-minimal expansions of the real field, even ones which are polynomially bounded and have strong analytic properties. There exists an o-minimal, polynomially bounded expansion of the real field, $\mathbb{R}_h$, by a one-variable function $h$ such that the structure $\mathbb{R}_h$ has analytic cell decomposition but the graph of $h$ does not have a mild parameterization definable in $\mathbb{R}_h$; in fact, it does not have a mild parameterization which is definable in any o-minimal, polynomially bounded expansion of the real field [26]. By different methods, the graph of a single irrational power function $x^\alpha$ restricted to $(0, 1)$, for $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ - which is not mild but does have a mild parameterization given by the three functions $(\exp(-\frac{1}{\alpha t}), \exp(-\frac{1}{t}))$, $((1-t)e^{-1/\alpha} + t, ((1-t)e^{-1/\alpha} + t)^\alpha)$ and $(e^{-1/\alpha}, e^{-1})$ defined on $(0, 1)$ - is also known not to have a mild parameterization by functions definable in any o-minimal, polynomially bounded expansion of the real field [7]. This means in particular that no o-minimal, polynomially bounded expansion of the real field with field of exponents strictly larger than $\mathbb{Q}$ can have definable mild parameterization.

One feature unfortunately lacking from the mild parameterizations obtained thus far is the uniformity required to exploit fully Pila's analogue to Step 2 in the above proof strategy. We therefore close by noting the recent development of a modified form of parameterization known as analytic quasi-parameterization [8], which involves many-variable complex analytic maps given in terms of both Weierstrass polynomials and mild maps. It has been shown that in suitable reducts of $\mathbb{R}_{\text{an}}$ such parameterizations may be obtained with a certain uniformity, although at the expense of increasing the size of the set being parameterized to a larger one of the same dimension. This improves the result of [13] concerning $\mathbb{R}_{\text{an}}$-definable Pfaffian surfaces, mentioned above, to one which is uniform in families. However, this approach still does not provide sufficient uniformity to complete the proof sketch, and so it does not provide any new higher dimensional instances of sets attaining the polylogarithmic bound, even where Step 3 is known. Exploiting this strategy of parameterization to approach Wilkie's Conjecture therefore remains a highly active area of investigation.

## References

[1] Binyamini, G. Density of algebraic points on Noetherian varieties. Preprint, arxiv.org/abs/1704.00442.

[2] Binyamini, G., and Novikov, D. The Pila-Wilkie theorem for subanalytic families: a complex analytic approach. Preprint, arxiv.org/abs/1605.04537.

[3] Binyamini, G., and Novikov, D. Wilkie's conjecture for restricted elementary functions. Preprint, arxiv.org/abs/1605.04671.

[4] Burguet, D. A proof of Yomdin-Gromov's algebraic lemma. *Israel J. Math. 168* (2008), 291–316.

[5] Burguet, D., Liao, G., and Yang, J. Asymptotic $h$-expansiveness rate of $C^\infty$ maps. *Proc. Lond. Math. Soc. (3) 111*, 2 (2015), 381–419.

[6] Butler, L. A. Some cases of Wilkie's conjecture. *Bull. Lond. Math. Soc. 44*, 4 (2012), 642–660.

[7] Çiray, D., and Thomas, M. E. M. Mild parameterization for certain o-minimal expansions of the real field. In Preparation.

[8] Cluckers, R., Pila, J., and Wilkie, A. J. Uniform parameterization of subanalytic sets and diophantine applications. Preprint, arxiv.org/abs/1605.05916.

[9] van den Dries, L., and Speissegger, P. The real field with convergent generalized power series. *Trans. Amer. Math. Soc. 350*, 11 (1998), 4377–4421.

[10] Gromov, M. Entropy, homology and semialgebraic geometry. *Astérisque*, 145-146 (1987), 5, 225–240. Séminaire Bourbaki, Vol. 1985/86.

[11] Jones, G. O., Miller, D. J., and Thomas, M. E. M. Mildness and the density of rational points on certain transcendental curves. *Notre Dame J. Form. Log. 52*, 1 (2011), 67–74.

[12] Jones, G. O., and Thomas, M. E. M. Effective Pila - Wilkie bounds for surfaces implicitly defined from Pfaffian functions. In Preparation.

[13] Jones, G. O., and Thomas, M. E. M. The density of algebraic points on certain Pfaffian surfaces. *Q. J. Math. 63*, 3 (2012), 637–651.

[14] Khovanskii, A. G. *Fewnomials*, vol. 88 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1991. Translated from the Russian by Smilka Zdravkovska.

[15] Pila, J. Geometric postulation of a smooth function and the number of rational points. *Duke Math. J. 63*, 2 (1991), 449–463.

[16] PILA, J. Integer points on the dilation of a subanalytic surface. *Q. J. Math. 55*, 2 (2004), 207–223.

[17] PILA, J. Rational points on a subanalytic surface. *Ann. Inst. Fourier (Grenoble) 55*, 5 (2005), 1501–1516.

[18] PILA, J. Counting rational points on a certain exponential-algebraic surface. *Ann. Inst. Fourier (Grenoble) 60*, 2 (2010), 489–514.

[19] PILA, J. O-minimality and the André-Oort conjecture for $\mathbb{C}^n$. *Ann. of Math. (2) 173*, 3 (2011), 1779–1840.

[20] PILA, J., AND TSIMERMAN, J. The André-Oort conjecture for the moduli space of abelian surfaces. *Compos. Math. 149*, 2 (2013), 204–216.

[21] PILA, J., AND WILKIE, A. J. The rational points of a definable set. *Duke Math. J. 133*, 3 (2006), 591–616.

[22] PILA, J., AND ZANNIER, U. Rational points in periodic analytic sets and the Manin-Mumford conjecture. *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl. 19*, 2 (2008), 149–162.

[23] ROLIN, J.-P., AND SERVI, T. Quantifier elimination and rectilinearization theorem for generalized quasianalytic algebras. *Proc. Lond. Math. Soc. (3) 110*, 5 (2015), 1207–1247.

[24] ROLIN, J.-P., SPEISSEGGER, P., AND WILKIE, A. J. Quasianalytic Denjoy-Carleman classes and o-minimality. *J. Amer. Math. Soc. 16*, 4 (2003), 751–777.

[25] TARSKI, A. *A decision method for elementary algebra and geometry.* University of California Press, Berkeley and Los Angeles, Calif., 1951. 2nd ed.

[26] THOMAS, M. E. M. An o-minimal structure without mild parameterization. *Ann. Pure Appl. Logic 162*, 6 (2011), 409–418.

[27] TSIMERMAN, J. A proof of the Andre-Oort conjecture for $\mathcal{A}_g$. Preprint, arxiv.org/abs/1506.01466.

[28] YOMDIN, Y. $C^k$-resolution of semialgebraic mappings. Addendum to: "Volume growth and entropy". *Israel J. Math. 57*, 3 (1987), 301–317.

[29] YOMDIN, Y. Volume growth and entropy. *Israel J. Math. 57*, 3 (1987), 285–300.

[30] YOMDIN, Y. Local complexity growth for iterations of real analytic mappings and semicontinuity moduli of the entropy. *Ergodic Theory Dynam. Systems 11*, 3 (1991), 583–602.

# A SUBSTITUTION THEOREM FOR RINGS OF SEMIALGEBRAIC FUNCTIONS AND APPLICATIONS

JOSÉ F. FERNANDO, J.M. GAMBOA

ABSTRACT. The present paper is a survey collecting the main results exposed by the first author in a talk presented in the seminar "Structures Algébriques Ordonnées" on May 31st, 2016. Let $R \subset F$ be an extension of real closed fields and $\mathcal{S}(M,R)$ the ring of (continuous) semialgebraic functions on a semialgebraic set $M \subset R^m$. We show that every $R$-homomorphism $\varphi : \mathcal{S}(M,R) \to F$ is essentially the evaluation homomorphism at a certain point $p \in F^m$ *adjacent* to the extended semialgebraic set $M_F$. This type of result is commonly known in Real Algebra as Substitution Theorem. In case $M$ is locally closed, the results are neat, whereas the non locally closed case requires a more subtle approach and some auxiliary results of geometric nature (*weak continuous extension theorem, appropriate immersion* of semialgebraic sets) that have interest on their own. We also afford the same problem for the ring $\mathcal{S}^*(M,R)$ of bounded (continuous) semialgebraic functions getting results of a rather different nature. Finally, we use substitution theorems to understand better some prime and maximal ideals of the rings $\mathcal{S}(M,R)$ and $\mathcal{S}^*(M,R)$.

## 1. INTRODUCTION

A very general question in Mathematics is to find, given a nonempty set $X$, a distinguished subset $S \subset X$ such that the elements of suitable class of maps $\varphi : X \to Y$ are determined by the restriction $\varphi|_S : S \to Y$. With this generality the unique option is to choose $S := X$, but if we restrict ourselves to the case in which $X$ is a vector space and $\varphi$ is linear, everybody knows that we can choose as $S$ an arbitrary basis of $X$. This kind of results is ubiquitous in mathematics; as a very basic one recall that if $A$ is a commutative ring the images $\varphi(u_1), \ldots, \varphi(u_m)$ determine every $A$-algebras homomorphism $\varphi : A[u_1, \ldots, u_m] \to B$. In the same vein, let $\mathbb{K} := \mathbb{R}$ or $\mathbb{C}$ and denote $\mathcal{A}_m$, indistinctly, the rings $\mathbb{K}\{\mathrm{x}\}$ or $\mathbb{K}[[\mathrm{x}]]$ of convergent and formal power series in $m$ variables $\mathrm{x} := (\mathrm{x}_1, \ldots, \mathrm{x}_m)$. Let $A$ and $B$ be two *analytic or formal algebras,* that is, two quotients $A := \mathcal{A}_m/\mathfrak{a}$ and $B := \mathcal{A}_n/\mathfrak{b}$ where $\mathfrak{a}$ and $\mathfrak{b}$ are, respectively, ideals of $\mathcal{A}_m$ and $\mathcal{A}_n$. Then it follows from Krull's intersection theorem that every homomorphism $\varphi : A \to B$ is completely determined by the images $\varphi(\mathrm{x}_i + \mathfrak{a})$ for $i = 1, \ldots, m$.

In a different context the maximal ideals of the ring $\mathcal{C}(K)$ of $\mathbb{R}$-valued continuous functions on a compact topological space $K$ have the form $\mathfrak{m}_p := \{f \in \mathcal{C}(K) : f(p) = 0\}$ for some $p \in K$ as a direct consequence of Uryshon's theorem. Consequently, all ring homomorphisms $\mathcal{C}(I) \to \mathbb{R}$ are evaluations at a suitable point.

---

*Date*: February 28th, 2017.

2010 *Mathematics Subject Classification.* Primary 14P10, 54C30; Secondary 12D15, 13E99.

*Key words and phrases.* Semialgebraic set, ring of semialgebraic functions, extension of coefficients, evaluation homomorphisms, Substitution Theorem, appropriately embedded semialgebraic set, weak continuous extension property, ring of bounded semialgebraic function, semialgebraic bounded-closure, $z$-ideal, depth, ideals associated to semialgebraic and formal paths.

Authors supported by Spanish MTM2014-55565 and GAAR Grupos UCM 910444.

In Real Algebra this kind of results are usually called 'Substitution Theorems' in honour to G. Efroymson's theorem [E] that we recall here. Let $\mathcal{N}(M)$ denote the ring of Nash functions on a Nash manifold $M \subset \mathbb{R}^m$ and let $F$ be a real closed field containing $\mathbb{R}$. Consider the $i$th projection

$$\pi_i : M \to \mathbb{R},\ x := (x_1, \ldots, x_m) \mapsto x_i$$

for $i = 1, \ldots, m$ (which is trivially a Nash function). Efroymson Theorem shows that the images $\varphi(\pi_1), \ldots, \varphi(\pi_m)$ determine every $\mathbb{R}$-homomorphism $\varphi : \mathcal{N}(M) \to F$. This result was extended in [FG1, §6.1] to the case in which $R$ is a real closed field and $M \subset R^m$ is a semialgebraic set, that is, a boolean combination of sets defined by polynomial equations and inequalities. In particular it applies if $M$ is a Nash set, that is, the zero set of a Nash function $f : U \to R$ where $U$ is an open semialgebraic subset of $R^m$ that contains $M$.

Along this work a function $f : M \to R$ is a *semialgebraic function* if its graph is a semialgebraic subset of $R^{m+1}$ and in addition it is continuous, whereas $f : M \to \mathbb{R}$ is a *Nash function* if $M$ is a Nash manifold and $f$ is a semialgebraic function of class $\mathcal{C}^\infty$. We denote the ring of (continuous) semialgebraic functions on $M$ by $\mathcal{S}(M, R)$ and its subring constituted by those that are bounded by $\mathcal{S}^*(M, R)$. We use the notation $\mathcal{S}^\diamond(M, R)$ when referring to both of them indistinctly.

To explain the name *Substitution* (that can be replaced by *Evaluation*) we recall an elementary but useful construction in semialgebraic geometry.

1.A. **Extension of coefficients.** Let $M \subset R^m$ be a semialgebraic set and let $F$ be a real closed field containing $R$. There exists a unique semialgebraic subset $M_F \subset F^m$ called the *extension of $M$ to $F$* such that $M = M_F \cap R^m$. The set $M_F$ is the subset of $F^m$ satisfying the same equalities and inequalities that occur in a definition of $M$ and by Tarski-Seidenberg Theorem (or equivalently by Artin-Lang Theorem) it does not depend on the chosen representation of $M$. All what is needed concerning extension of coefficients appears in [DK1] and [BCR, §5].

Given another semialgebraic set $N \subset R^n$ and a semialgebraic map $f : M \to N$, there exists a unique semialgebraic map $f_F : M_F \to N_F$ called *extension of $f$ to $F$* that fulfills $f_F|_M = f$. This follows applying the extension of semialgebraic sets to the graph of $f$.

By [BCR, 7.3.1], the extension of semialgebraic functions induces a well-defined injective $R$-homomorphism

$$\mathtt{i}_{M,F} : \mathcal{S}(M, R) \hookrightarrow \mathcal{S}(M_F, F),\ f \mapsto f_F.$$

For each point $\mathrm{p} \in M_F$ let us denote the *evaluation* at $\mathrm{p}$ as the $R$-homomorphism

$$\mathrm{ev}_{M_F,\mathrm{p}} : \mathcal{S}(M_F, F) \to F,\ g \mapsto g(\mathrm{p}).$$

This way one gets the $R$-homomorphism $\psi_{\mathrm{p}} := \mathrm{ev}_{M_F,\mathrm{p}} \circ \mathtt{i}_{M,F} : \mathcal{S}(M, R) \to F$. With this terminology and taking into account that $\mathcal{N}(M) \subset \mathcal{S}(M, \mathbb{R})$ Efroymson's Theorem quoted above says that if $M \subset \mathbb{R}^m$ is a Nash manifold and $\varphi : \mathcal{N}(M) \to F$ is an $\mathbb{R}$-homomorphism, then the point $\mathrm{p} := (\varphi(\pi_1), \ldots, \varphi(\pi_m)) \in M_F$ and for each function $f \in \mathcal{N}(M)$ the equality $\varphi(f) = f_F(\mathrm{p})$ holds, that is, $\varphi = \psi_{\mathrm{p}}$. Thus every $\mathbb{R}$-homomorphism $\varphi : \mathcal{N}(M) \to F$ is the evaluation at a suitable point in $M_F$.

It seems worthwhile mentioning the existence of other classes of $\mathbb{R}$-algebras homomorphisms that are evaluations at some point. For example, all $\mathbb{R}$-homomorphisms of the ring of smooth functions $\mathcal{C}^\infty(X, \mathbb{R})$ on a differentiable manifold $X$ (or an open subset of a Banach space) into $\mathbb{R}$ (see [BL1, KMS]) or the already mentioned above between analytic or formal $\mathbb{K}$-algebras. In [BL] the authors proved that many subalgebras $A$ of the $\mathbb{R}$-algebra $\mathcal{C}(X)$ of $R$-valued continuous functions on a completely regular topological space $X$ enjoys the following *evaluative property:*

for each $\mathbb{R}$-homomorphism $\varphi : A \to \mathbb{R}$ and every sequence $\{f_n\} \subset A$ there exists a point $p \in X$ such that $\varphi(f_n) = f_n(p)$. Under mild conditions, the $\mathbb{R}$-homomorphisms of the ring $\mathcal{C}^k(U)$ of $\mathcal{C}^k$ functions on an open subset $U$ of a Banach space into $\mathbb{R}$ are evaluations at some point [GGJ, GL, J].

1.B. **Core of a homomorphism.** The situation in the semialgebraic context is harder. Fix an $R$-homomorphism $\varphi : \mathcal{S}(M,R) \to F$. Its *core* is defined as $\mathrm{p}_\varphi := (\varphi(\pi_1), \ldots, \varphi(\pi_m)) \in F^m$ and, in general, it does not belong to $M_F$. This is a first obstruction to get a Substitution theorem, that is, the equality $\varphi = \psi_{\mathrm{p}_\varphi}$. In addition, in dealing with the ring $\mathcal{S}^*(M,R)$, the set $M$ is bounded if and only if all projections $\pi_i : M \to R$ belong to $\mathcal{S}^*(M,R)$. Consequently, the equality $\varphi(f) = f_F(\mathrm{p}_\varphi)$ for each $R$-homomorphism $\varphi : \mathcal{S}^*(M,R) \to F$ and all $f \in \mathcal{S}^*(M,R)$ makes sense only if $M$ is bounded. This can always be assumed because the map

$$\Phi : \mathcal{B}_m(0,1) \to R^m,\ x \mapsto \frac{x}{\sqrt{1-\|x\|^2}},$$

where $\mathcal{B}_m(0,1)$ denotes the open ball of $R^m$ of radius 1 centered at the origin, is a semialgebraic homeomorphism and $\Phi^{-1}(M)$ is a bounded semialgebraic set. Thus, we can replace $M$ by the bounded semialgebraic set $\Phi^{-1}(M)$ if needed.

1.C. **Adjacency.** Let $M \subset R^m$. A point $\mathrm{p} \in F^m$ is *adjacent to* $M$ if $\mathrm{p} \in N_F$ for each locally closed semialgebraic set $N \subset R^m$ that contains $M$. The *adjancency of* $M$ *in* $F^m$ is the pro-constructible set $\widehat{M}_F$ of points of $F^m$ that are adjacent to $M$. Notice that $M_F \subset \widehat{M}_F \subset \mathrm{Cl}(M)_F$. If $M$ is locally closed or if $R = F$, then $M_F = \widehat{M}_F$. The set $\widehat{M}_F$ is carefully studied in [S3, I.3.20] and it holds that for each $R$-homomorphism $\varphi : \mathcal{S}(M,R) \to F$ its core $\mathrm{p}_\varphi \in \widehat{M}_F$. In particular, if $M$ is locally closed in $R^m$, we have $\mathrm{p}_\varphi \in M_F$ .

It must be pointed out that if $M$ is bounded and $\varphi : \mathcal{S}^*(M,R) \to F$ is an $R$-homomorphism we can only say that $\mathrm{p}_\varphi \in \mathrm{Cl}(M)_F$. This is the price we pay because the ring $\mathcal{S}^*(M,R)$ contains functions with empty zero set that are not units.

*Proof.* Write $\mathrm{Cl}(M) := \bigcup_{i=1}^r \{g_{i1} \geq 0, \ldots, g_{is} \geq 0\}$ with $g_{ij} \in R[\mathrm{x}]$ and assume by contradiction that $\mathrm{p}_\varphi \notin \mathrm{Cl}(M)_F$. As $M$ is bounded, each $h_{ij} := g_{ij}|_M \in \mathcal{S}^*(M,R)$, so we may assume $h_{i1,F}(\mathrm{p}_\varphi) < 0$ for each index $i = 1, \ldots, r$. The product $h := \prod_{i=1}^r (h_{i1} - |h_{i1}|)^2$ is the zero function on $M$, so $\varphi(h) = 0$. As $\varphi(|h_{i1}|) = |h_{i1,F}(\mathrm{p}_\varphi)|$, we conclude

$$0 = \varphi(h) = \prod_{i=1}^r (\varphi(h_{i1}) - \varphi(|h_{i1}|))^2 = \prod_{i=1}^r (h_{i1,F}(\mathrm{p}_\varphi) - |h_{i1,F}(\mathrm{p}_\varphi)|)^2 > 0,$$

which is a contradiction. □

In addition, some kind of converse is true. Namely, given a point $\mathrm{p} \in \mathrm{Cl}(M)_F$ there exists by the Curve Selection Lemma [BCR, 2.5.5] a semialgebraic path $\alpha : [0,1]_F \to F^m$ such that $\alpha(0) = \mathrm{p}$ and $\alpha((0,1]_F) \subset M_F$. Then

$$\varphi : \mathcal{S}^*(M,R) \to F,\ f \mapsto \lim_{t \to 0^+} (f_F \circ \alpha)(t)$$

is a well-defined homomorphism whose core is p.

## 2. MAIN RESULTS

We introduce right now the necessary ingredients to state the Substitution Theorem in the semialgebraic setting [Fe2]. This result implies that if $M$ is a locally closed semialgebraic set, then $\varphi(f) = f_F(\mathrm{p}_\varphi)$ for each $R$-homomorphism $\varphi : \mathcal{S}(M,R) \to F$ and each $f \in \mathcal{S}(M,R)$ (see Corollary 2.8). Theorem 2.10 affords the Substitution Theorem for $R$-homomorphisms $\mathcal{S}^*(M,R) \to F$. It is harder to prove Theorem 2.11 that gives the best possible result about the behaviour of the $R$-homomorphisms $\mathcal{S}(M,R) \to F$ when $M$ is an arbitrary semialgebraic set. The full proofs of these results appear in [Fe2].

As the coordinate projections are involved in its definition, the core of a homomorphism $\varphi : \mathcal{S}^\diamond(M,R) \to F$ depends on how $M$ is embedded in $R^m$. Consequently, 'a good immersion' increases the possibility to have a Substitution Theorem. As a general real closed field $R$ is not complete (that is, Cauchy sequences in $R$ need not to be convergent in $R$) as it happens with the field $\mathbb{R}$ of real numbers, there are few compact semialgebraic sets in $R^n$. For instance, if $R := \mathbb{R}(\{\mathtt{t}^*\})$ is the field of meromorphic Puiseux series with coefficients in $\mathbb{R}$, the circle $\mathbb{S}^1 := \{(x,y) \in R^2 : x^2 + y^2 = 1\}$ is not compact, although it is bounded and closed in $R^2$. The notion that substitutes compactness is *bounded-closedness.*

We say that a subset $S \subset \mathbb{R}^m$ is *bounded-closed* if it is bounded and closed in $R^m$ and it is *locally bounded-closed* if each point in $S$ has a bounded-closed semialgebraic neighborhood in $S$. In addition, $S$ is *locally closed in* $R^m$ if there exist an open subset $U$ of $R^m$ and a closed subset $C$ of $R^m$ such that $S = U \cap C$. If $S$ is locally closed in $R^m$, then $\rho_0(S) := \mathrm{Cl}(S) \setminus S$ is a closed subset of $R^m$ and $S = (S \setminus \rho_0(S)) \cap \mathrm{Cl}(S)$. Thus, if $M$ is a locally closed semialgebraic subset of $R^m$, then it is the intersection of an open and a closed semialgebraic subsets of $R^m$. In addition, one can check that local closedness and local bounded-closedness are equivalent notions.

The ring $\mathcal{S}(M,R)$ has a nice behaviour if $M$ is locally closed. One of the main reasons is that in this case Łojasiewicz's inequality holds [BCR, 2.6.7]. To treat the general situation a useful approach is to compare $\mathcal{S}(M,R)$ and $\mathcal{S}(M_{\mathrm{lc}},R)$ where

$$M_{\mathrm{lc}} := \mathrm{Cl}(M) \setminus \mathrm{Cl}(\rho_0(M))$$

is the largest locally closed and dense subset of $M$. It holds that $M_{\mathrm{lc}}$ is an open semialgebraic subset of $M$ and it can be characterized as the set of points of $M$ that have a bounded-closed neighborhood in $M$.

**Definitions 2.1.** Let $M \subset R^m$ be a semialgebraic set and let $\eta(M) \subset R^m$ be the semialgebraic subset consisting of those points $q \in \rho_0(M)$ such that either $0 \leq \dim(\rho_0(M)_q) < \dim(M_q) - 1$ or $M_q$ is not semialgebraically connected. A semialgebraic set $M \subset R^m$ is *appropriately embedded* if $\eta(M)$ is empty.

Next theorem guarantees that an appropriate embedding with certain additional properties always exists.

**Theorem 2.2** (Appropriate embedding). *Assume that $M \subset R^m$ is a bounded semialgebraic set. Then there exists a semialgebraic neighborhood $U$ of $\eta(M)$ in $R^m$ such that the difference $N := M \setminus U$ is appropriately embedded and a surjective semialgebraic map $h : \mathrm{Cl}(N) \to \mathrm{Cl}(M)$ such that $h|_N : N \to M$ is a semialgebraic homeomorphism.*

To state the Substitution Lemma in the semialgebraic set we need some additional definitions.

**Definition 2.3.** Let $\mathrm{p} \in F^m$ and denote $\mathcal{C}_{M,\mathrm{p}}$ the family of all closed semialgebraic subsets $N$ of $M$ with $\mathrm{p} \in N_F$. The *semialgebraic length* of $\mathrm{p}$ is the nonnegative integer

$$\ell_M(\mathrm{p}) := \min\{\dim(N) : \ N \in \mathcal{C}_{M,\mathrm{p}}\},$$

*Remarks* 2.4. The previous invariant has different algebraic and geometrical meanings.

(i) The *semialgebraic depth* of a prime ideal $\mathfrak{p}$ of $\mathcal{S}(M,R)$ was introduced in [Fe1] as

$$\mathrm{d}_M(\mathfrak{p}) := \min\{\dim(Z_M(f)) : f \in \mathfrak{p}\},$$

where $Z_M(f) := \{x \in M : f(x) = 0\}$, and it plays in semialgebraic geometry a similar role to that of dimension of zero sets of prime ideals in Algebraic Geometry. As each closed semialgebraic subset $C$ of $M$ is the zero set of the semialgebraic function $M \to R$, $x \mapsto \operatorname{dist}(x, C)$, we have

$$\begin{aligned}\mathrm{d}_M(\ker(\psi_{\mathrm{p}})) &= \min\{\dim(Z_M(f)) : f \in \ker(\psi_{\mathrm{p}})\} = \min\{\dim(Z_M(f)) : f_F(\mathrm{p}) = 0\} \\ &= \min\{\dim(Z_M(f)) : \mathrm{p} \in (Z_M(f))_F\} = \min\{\dim(N) : N \in \mathcal{C}_{M,\mathrm{p}}\} = \ell_M(\mathrm{p}).\end{aligned}$$

(ii) In [S5] it is proved that the quotient field $F := \operatorname{qf}(\mathcal{S}(M,R)/\mathfrak{p})$ is real closed for each prime ideal $\mathfrak{p}$ in $\mathcal{S}(M,R)$. Let $\mathrm{p}_\varphi \in F^m$ be the core of the canonical homomorphism

$$\varphi : \mathcal{S}(M,R) \to \mathcal{S}(M,R)/\mathfrak{p} \hookrightarrow F, \ f \mapsto f + \mathfrak{p}.$$

Assume that $\mathrm{p}_\varphi \in M_F$. Then $\mathfrak{p} = \{f \in \mathcal{S}(M,R) : \ f_F(\mathrm{p}_\varphi) = 0\}$ and $\ell_M(\mathrm{p}_\varphi) = \mathrm{d}_M(\mathfrak{p})$.

(iii) Suppose that $M$ is bounded-closed and let $R(\mathrm{p})$ be the smallest subfield of $F$ that contains $R$ and the coordinates of the tuple $\mathrm{p} \in F^n$. Denote $\kappa(\mathrm{p})$ the quotient field of $\mathcal{S}(M,R)/\ker(\psi_{\mathrm{p}})$. Then,

$$\ell_M(\mathrm{p}) = \mathrm{d}_M(\ker(\psi_{\mathrm{p}})) = \operatorname{tr\,deg}_R(\kappa(\mathrm{p})) = \operatorname{tr\,deg}_R(\operatorname{qf}(R[\mathrm{x}]/(\ker(\psi_{\mathrm{p}}) \cap R[\mathrm{x}]))) = \operatorname{tr\,deg}_R(R(\mathrm{p})).$$

Next example illustrates the relationship between both invariants we have introduced.

**Example 2.5.** Let $F := \mathbb{R}(\{\mathtt{t}^*\})$ be the field of meromorphic Puiseux series with coefficients in $\mathbb{R}$. Consider the point $\mathrm{p} := (\mathtt{t}, e^{\mathtt{t}}) \in F^2$. Then $\mathbb{R}(\mathrm{p}) = \mathbb{R}(\mathtt{t}, e^{\mathtt{t}})$, so $\operatorname{tr\,deg}_R(R(\mathrm{p})) = 2$.

*Proof.* Let us check that $\ell_{\mathbb{R}^2}(\mathrm{p}) = 2$. We have to prove that no one-dimensional closed semialgebraic subset $N$ of $\mathbb{R}^2$ satisfies $\mathrm{p} \in N_F$. Suppose by contradiction that there exists a one-dimensional closed semialgebraic subset $N$ such that $\mathrm{p} \in N_F$. By [BCR, 2.9.10] we may assume $N$ is the union of two points and a Nash manifold that is Nash diffeomorphic to the open interval $(0,1)$. Thus, there exists a non-zero polynomial $P \in \mathbb{R}[\mathtt{t}, \mathtt{x}]$ such that $N \subset Z_{\mathbb{R}^2}(P)$. Consequently, $N_F \subset Z_{F^2}(P_F)$, so $P_F(\mathtt{t}, e^{\mathtt{t}}) = 0$ and this implies $P(t, e^t) = 0$ for each $t \in \mathbb{R}$ (because $f : \mathbb{R} \to \mathbb{R}$, $t \mapsto e^t$ is an analytic function). Hence, $f : \mathbb{R} \to \mathbb{R}$, $t \mapsto e^t$ is a semialgebraic function and by [BCR, 2.6.1] its growth is upperly bounded by certain monomial $ct^k$ for $t$ large, which is false. □

**Lemma 2.6** (Substitution Lemma). *Let $\mathrm{p} \in F^m$ be adjacent to $M$ and denote $X := \operatorname{Cl}(M)$.*

(i) *Suppose that $\mathrm{p} \in M_F$. Then $\psi_{\mathrm{p}} : \mathcal{S}(M,R) \to F$ is the unique homomorphism $\mathcal{S}(M,R) \to F$ whose core is $\mathrm{p}$.*

(ii) *If $\ell_X(\mathrm{p}) = \dim(M_{F,\mathrm{p}})$ then $\mathrm{p} \in M_{\mathrm{lc},F}$.*

(iii) *Suppose that $M$ is appropriately embedded, $\ell_X(\mathrm{p}) = \dim(M_{F,\mathrm{p}}) - 1$ and $\mathrm{p} \notin M_F$. Then $\psi_{\mathrm{p}} : \mathcal{S}(M,R) \to F$ is the unique homomorphism whose core is $\mathrm{p}$. In addition, given $f \in \mathcal{S}(M,R)$ there exists a semialgebraic set $M_f \subset X$ containing $M$ such that $\mathrm{p} \in M_{f,F}$ and $f$ admits an extension $\widehat{f} \in \mathcal{S}(M_f, R)$ satisfying the equality $\psi_{\mathrm{p}}(f) = \widehat{f}_F(\mathrm{p})$.*

(iv) *Suppose that* $\mathrm{p} \notin M_F$ *and* $\ell_X(\mathrm{p}) \leq \dim(M_{F,\mathrm{p}}) - 2$. *Then there exist infinitely many homomorphisms* $\mathcal{S}(M,R) \to F$ *whose core is* $\mathrm{p}$.

The proof of the Substitution Lemma is based on a fruitful use of universal properties of the real closure of the ring of polynomials with coefficients in $R$ (see [S2, S3, S4, S5]) and the following weak continuous extension property that has its own interest.

**Theorem 2.7** (Weak continuous extension property). *Suppose that the germ* $M_q$ *is semialgebraically connected for each point* $q \in \operatorname{Cl}(M)$. *Then for each* $f \in \mathcal{S}(M,R)$ *there exist an open semialgebraic neighborhood* $U$ *of* $M$ *in* $\operatorname{Cl}(M)$ *and a semialgebraic set* $Y \subset \operatorname{Cl}(M) \setminus M$ *such that* $\dim(Y_q) \leq \dim(M_q) - 2$ *for all* $q \in Y$ *and* $f$ *can be extended continuously to* $U \setminus Y$.

Notice that Lemma 2.6 provides a very satisfactory Substitution Theorem for $R$-homomorphisms $\mathcal{S}(M,R) \to F$ in case $M$ is locally closed.

**Corollary 2.8.** *If* $M$ *is locally closed, then every homomorphism* $\varphi : \mathcal{S}(M,R) \to F$ *is an evaluation homomorphism.*

*Proof.* As $M$ is locally closed, the core $\mathrm{p}_\varphi \in M_F$. Thus, $\varphi = \operatorname{ev}_{M_F,\mathrm{p}_\varphi} \circ \mathtt{i}_{M,F}$ because both are homomorphisms whose core is $\mathrm{p}_\varphi$. Consequently, for each $f \in \mathcal{S}(M,R)$,

$$\varphi(f) = (\operatorname{ev}_{M_F,\mathrm{p}_\varphi} \circ \mathtt{i}_{M,F})(f) = f_F(\mathrm{p}_\varphi) = f_F(\varphi(\pi_1), \ldots, \varphi(\pi_m)),$$

as required. □

In the proof of Lemma 2.6 one supplies the lack of local closedness of a semialgebraic set $M$ by substituting it by the largest locally closed and dense subset $M_{\mathrm{lc}}$ of $M$. Another approach, that is mainly fruitful in dealing with the ring $\mathcal{S}^*(M,R)$ of bounded semialgebraic functions, is to use *semialgebraic bounded-closures*.

**Definition 2.9.** A *semialgebraic bounded-closure* of $M$ is a pair $(X, \mathtt{j})$ where $X$ is a bounded-closed semialgebraic set contained in some $R^n$ and $\mathtt{j} : M \to R^n$ is a semialgebraic homeomorphism onto $\mathtt{j}(M) \subset X \subset \operatorname{Cl}(\mathtt{j}(M))$.

A distinguished class of semialgebraic bounded-closures was introduced in [FG3]. A *brimming semialgebraic bounded-closure* of $M$ for a prime ideal $\mathfrak{p}$ of $\mathcal{S}^\diamond(M,R)$ is a semialgebraic bounded-closure $(X, \mathtt{j})$ of $M$ such that the real closed quotient fields $\kappa(\mathfrak{p}) := \operatorname{qf}(\mathcal{S}^\diamond(M,R)/\mathfrak{p})$ and $\kappa(\mathfrak{p}_X) = \operatorname{qf}(\mathcal{S}(X)/\mathfrak{p}_X)$, where $\mathfrak{p}_X := \mathfrak{p} \cap \mathcal{S}(X)$, are isomorphic. It was proved in [FG3] that for each prime ideal $\mathfrak{p}$ there exists a brimming semialgebraic bounded-closure of $M$ with respect to $\mathfrak{p}$.

We are ready to understand how the homomorphisms $\varphi : \mathcal{S}^*(M,R) \to F$ look like. The statement of the following result holds also for the ring $\mathcal{S}(M,R)$.

**Theorem 2.10** (Substitution Theorem I). *Let* $\varphi : \mathcal{S}^\diamond(M,R) \to F$ *be an* $R$*-homomorphism and let* $(X, \mathtt{j})$ *be a brimming bounded-closure of* $M$ *for* $\mathfrak{p} := \ker(\varphi)$. *Let* $\mathrm{p}$ *be the core of* $\psi := \varphi \circ \mathtt{j}^*$ *and denote* $A := \mathcal{S}(X,R)/\mathfrak{p}_X$. *Then* $\psi$ *induces the homomorphism*

$$\widehat{\psi} : \operatorname{qf}(A) = \operatorname{qf}(\mathcal{S}^\diamond(M,R)/\mathfrak{p}) \hookrightarrow F, \ \tfrac{[a_1]}{[a_2]} \mapsto \tfrac{\psi(a_1)}{\psi(a_2)} = \tfrac{a_{1,F}(\mathrm{p})}{a_{2,F}(\mathrm{p})}$$

*and the following diagram is commutative*

$$\begin{tikzcd}
\operatorname{qf}(A) \arrow[r, equal] & \operatorname{qf}(\mathcal{S}^\diamond(M,R)/\mathfrak{p}) \arrow[r, hook, "\widehat{\psi}"] & F \\
A \arrow[u, hook] \arrow[r, hook, "\overline{\mathtt{j}}^*"] & \mathcal{S}^\diamond(M,R)/\mathfrak{p} \arrow[u, hook] \arrow[ur, hook, "\overline{\varphi}"] & \\
\mathcal{S}(X,R) \arrow[u] \arrow[r, hook, "\mathtt{j}^*"'] \arrow[uurr, bend right, "\psi"] & \mathcal{S}^\diamond(M,R) \arrow[u] \arrow[uur, bend right, "\varphi"'] &
\end{tikzcd}$$

*Thus $\varphi$ is completely determined by $\mathfrak{p}$: for each $f \in \mathcal{S}^\diamond(M,R)$, there exist $a_1, a_2 \in \mathcal{S}(X,R)$ such that $a_{2,F}(\mathfrak{p}) \neq 0$, $a_2 f - a_1 \in \mathfrak{p}$ and $\varphi(f) = \frac{a_{1,F}(\mathfrak{p})}{a_{2,F}(\mathfrak{p})}$.*

*Proof.* Recall that $\operatorname{qf}(A)$ is a real closed field. By Lemma 2.6 we have $\psi = \operatorname{ev}_{X_F,\mathfrak{p}} \circ \mathtt{i}_{X,F}$, so $\widehat{\psi} : \operatorname{qf}(A) \to F$ is the unique homomorphism between the real closed fields $\operatorname{qf}(A)$ and $F$. On the other hand, the injective homomorphism $\overline{\varphi} : \mathcal{S}^\diamond(M,R)/\mathfrak{p} \hookrightarrow F$ induced by $\varphi$ extends to the quotient field in a unique way:

$$\widehat{\overline{\varphi}} : \operatorname{qf}(\mathcal{S}^\diamond(M,R)/\mathfrak{p}) \hookrightarrow F, \frac{[f]}{[g]} \mapsto \frac{\varphi(f)}{\varphi(g)}.$$

As claimed in the diagram above, $\operatorname{qf}(A) = \operatorname{qf}(\mathcal{S}^\diamond(M,R)/\mathfrak{p})$ because $(X, \mathtt{j})$ is a brimming bounded-closure of $M$, so $\widehat{\overline{\varphi}} = \widehat{\psi}$. In addition, for each $f \in \mathcal{S}^\diamond(M,R)$ there exist $a_1, a_2 \in \mathcal{S}(X,R)$ such that $a_2 \notin \mathfrak{p}_X$ and $a_2 f - a_1 \in \mathfrak{p}$. In particular, $\varphi(a_i) = \psi(a_i) = a_{i,F}(\mathfrak{p})$ and $a_{2,F}(\mathfrak{p}) = \varphi(a_2) \neq 0$. Thus,

$$a_{2,F}(\mathfrak{p})\varphi(f) - a_{1,F}(\mathfrak{p}) = \varphi(a_2)\varphi(f) - \varphi(a_1) = \varphi(a_2 f - a_1) = 0,$$

that is, $\varphi(f) = \frac{a_{1,F}(\mathfrak{p})}{a_{2,F}(\mathfrak{p})}$, as required. □

As one can expect, the most involved and interesting situation concerning Substitution theorems for the ring $\mathcal{S}(M,R)$ appears when $M$ is not necessarily locally closed. The precise statement in this general context is the following.

**Theorem 2.11** (Substitution Theorem II). *Let $\varphi : \mathcal{S}(M,R) \to F$ be an $R$-homomorphism. Then there exist an appropriately embedded semialgebraic subset $N \subset R^n$, a semialgebraic embedding $h : N \hookrightarrow M$ such that $N_0 := h(N)$ is closed in $M$ and a point $\mathfrak{p} \in \widehat{N}_F$ with either $\mathfrak{p} \in N_F$ or $\ell_{\operatorname{Cl}(N)}(\mathfrak{p}) = \dim(N_{F,\mathfrak{p}}) - 1$, such that the following diagram is commutative*

$$\begin{tikzcd}
\mathcal{S}(M,R) \arrow[r, "\varphi"] \arrow[d, "\mathtt{j}^*"'] & F \\
\mathcal{S}(N_0,R) \arrow[r, "h^*"] & \mathcal{S}(N,R) \arrow[u, "\psi_\mathfrak{p}"']
\end{tikzcd}$$

*where $\mathtt{j} : N_0 \hookrightarrow M$ is the inclusion, $\psi_\mathfrak{p} : \mathcal{S}(N,R) \to F$ is the unique homomorphism whose core is $\mathfrak{p}$, $h^* : \mathcal{S}(N_0,R) \to \mathcal{S}(N,R)$, $f \mapsto f \circ h$ and $\mathtt{j}^* : \mathcal{S}(M,R) \to \mathcal{S}(N_0,R)$, $f \mapsto f \circ \mathtt{j}$.*

*Sketch of proof.* The proof of this result is rather involved (see [Fe2] for full details) and we only present a brief sketch.

**Step 1.** *Construction of the semialgebraic set $N_0$ in the statement.* Let $\mathfrak{p}_0 := \ker(\varphi)$ and let $(X, \mathtt{j})$ be a brimming semialgebraic bounded-closure of $M$ for $\mathfrak{p}_0$. Suppose that $X \subset R^m$ and let $Z$ be the smallest algebraic subset of $R^m$ containing $X$. Let $\mathcal{P}(Z)$ be the ring of polynomial functions on $Z$ and denote $\operatorname{Spec}_r(\mathcal{S}(X,R))$ the real spectrum of $\mathcal{S}(X,R)$ (the reader can find in [BCR, VII] the background concerning real spectra of commutative rings with unity). The inclusion $\mathcal{P}(Z) \hookrightarrow \mathcal{S}(X,R)$ induces a continuous map

$$\phi : \operatorname{Spec}_r(\mathcal{S}(X,R)) \to \operatorname{Spec}_r(\mathcal{P}(Z))$$

between the real spectra of both rings, whose image is the constructible subset $\widetilde{X}$ of $\operatorname{Spec}_r(\mathcal{P}(Z))$ associated to $X$. A basic but essential property of rings of semialgebraic functions is that the support map

$$\operatorname{supp} : \operatorname{Spec}_r(\mathcal{S}(X,R)) \to \operatorname{Spec}(\mathcal{S}(X,R)),\ \alpha \mapsto \operatorname{supp}_\alpha := \alpha \cap (-\alpha)$$

is a homeomorphism when $\operatorname{Spec}_r(\mathcal{S}(X,R))$ and $\operatorname{Spec}(\mathcal{S}(X,R))$ are respectively endowed with the usual and the Zariski topologies. By [CC, §3]

$$\rho := \phi \circ \operatorname{supp}^{-1} : \operatorname{Spec}(\mathcal{S}(X,R)) \to \widetilde{X}$$

is a homeomorphism. Let $\alpha_0 := \rho(\mathfrak{p}_0)$. As $M_{\mathrm{lc}}$ is dense in $X$, so is by Artin-Lang's Theorem $\widetilde{M}_{\mathrm{lc}}$ in $\widetilde{X}$. Thus $\alpha_0$ is adherent to the constructible set $\widetilde{M}_{\mathrm{lc}}$. By [Rz, Thm. I] it holds

$$d := \dim_{\alpha_0}(\widetilde{M}_{\mathrm{lc}}) - \dim(\alpha_0) \geq 0, \tag{2.1}$$

so $\widetilde{M}_{\mathrm{lc}}$ contains $d$ points $\alpha_1, \ldots, \alpha_d$ such that $\alpha_d \subsetneq \cdots \subsetneq \alpha_1 \subsetneq \alpha_0$. The prime ideals

$$\mathfrak{p}_i := \rho^{-1}(\alpha_i)\mathcal{S}(M_{\mathrm{lc}},R) \cap \mathcal{S}(M,R), \quad \text{for} \quad i = 1, \ldots, d$$

satisfy $\mathfrak{p}_d \subsetneq \cdots \subsetneq \mathfrak{p}_1 \subsetneq \mathfrak{p}_0$. Pick $f \in \mathfrak{p}_1$ such that $\dim(Z_M(f)) = \mathrm{d}_M(\mathfrak{p}_1)$. Then the semialgebraic set $N_0 := Z_M(f)$ satisfies the required properties.

**Step 2.** *Construction of the semialgebraic set $N$ in the statement.* The ideal $\mathfrak{p}_1$ contains the kernel of the surjective homomorphism $\psi : \mathcal{S}(M,R) \to \mathcal{S}(N_0,R)$, $f \mapsto f|_{N_0}$. Thus $\ker(\psi) \subset \mathfrak{p}_0$ and the quotient $\mathfrak{q}_0 := \mathfrak{p}_0/\ker(\psi)$ is identified as a prime ideal of $\mathcal{S}(N_0,R)$. By Theorem 2.2 there exists an appropriately embedded semialgebraic set $N \subset R^m$ and a surjective semialgebraic map $H : Y := \operatorname{Cl}(N) \to Y_0 := \operatorname{Cl}(N_0)$ whose restriction $h := H|_N : N \to N_0$ is a semialgebraic homeomorphism. Let $h^* : \mathcal{S}(N_0,R) \to \mathcal{S}(N,R)$, $g \mapsto g \circ h$ be the homomorphism induced by the map $h$. It holds that $Y$ is a brimming semialgebraic bounded-closure of $N$ for $h^*(\mathfrak{q}_0)$, so $N$ satisfies the conditions in the statement.

**Step 3.** *Construction of the point $\mathrm{p} \in F^n$ in the statement.* Consider the $R$-homomorphism $\overline{\varphi} : \mathcal{S}(N,R) \to F$ such that $\varphi = \overline{\varphi} \circ h^* \circ \mathtt{i}^*$ where $\mathtt{i}^* : \mathcal{S}(M,R) \to \mathcal{S}(N_0,R)$, $g \mapsto g|_{N_0}$ is the homomorphism induced by the inclusion $\mathtt{i} : N_0 \hookrightarrow M$. Let $\mathrm{p} := (\overline{\varphi}(\pi_1), \ldots, \overline{\varphi}(\pi_m))$, where $\pi_i : R^m \to R$ is the projection onto the $i$-th coordinate. By Proposition 2.4 we have

$$\begin{aligned}\dim(N_{F,\mathrm{p}}) = \dim(Y_{F,\mathrm{p}}) \geq \mathrm{d}_Y(\mathrm{p}) &= \operatorname{tr\,deg}_R \operatorname{qf}(\mathcal{S}(Y,R)/(h^*(\mathfrak{q}_0) \cap \mathcal{S}(Y,R))) \\ &= \dim(N_0) - 1 = \dim(N) - 1 \geq \dim(N_{F,\mathrm{p}}) - 1.\end{aligned}$$

If $\mathrm{p} \in N_F$, which includes the case $\mathrm{d}_Y(\mathrm{p}) = \dim(N_{F,\mathrm{p}})$, or if $\mathrm{d}_Y(\mathrm{p}) = \dim(N_{F,\mathrm{p}}) - 1$ and $\mathrm{p} \notin N_F$, it follows from Lemma 2.6 that $\overline{\varphi} = \psi_{\mathrm{p}} : \mathcal{S}(N,R) \to F$ is the unique $R$-homomorphism from $\mathcal{S}(N,R)$ to $F$ whose core is $\mathrm{p}$, as required. □

## 3. APPLICATIONS

The use of infinitesimals in Real Algebraic Geometry comes back to pioneer works of A. Robinson, G. Kreisel and G. Efroymson among others. The trick is to introduce in the arguments a positive infinitesimal element $\varepsilon > 0$ with respect to $\mathbb{R}$, that is, to work over the field $\mathbb{R}(\varepsilon)$ where

$\varepsilon$ is smaller than each positive real numer and afterwards to consider $\lim_{\varepsilon\downarrow 0}$. In more recent articles, see for instance [B, R], the method is employed in a more refined way to study the complexity of algorithms in semialgebraic geometry.

Let $R[[\mathtt{t}]]$ stands for the ring of formal power series in one variable with coefficients in $R$ and $R((\mathtt{t}))$ for its quotient field. A formal power series is *algebraic* if it is algebraic over the field of rational functions $R(\mathtt{t}) := \operatorname{qf}(R[\mathtt{t}])$. The subring (resp. subfield) of $R[[\mathtt{t}]]$ (resp. $R((\mathtt{t}))$) of all algebraic series is denoted with $R[[\mathtt{t}]]_{\mathrm{alg}}$ (resp. $R((\mathtt{t}))_{\mathrm{alg}}$). Given a formal power series $\xi \in R((\mathtt{t}))$, we denote its order with $\omega(\xi)$. We endow the previous rings with their respective unique orderings $\leq$ in which $\mathtt{t}$ is positive and infinitesimal with respect to $R$ (here appears the connection with the classical procedures). We denote the real closed field of Puiseux series with $F_1 := R((\mathtt{t}^*))$ and the real closed field of algebraic Puiseux series with $F_0 := R((\mathtt{t}^*))_{\mathrm{alg}}$, see [BCR, 1.3.6]. Basu and Roy employ in [B, R] the field $F_1$ to get information about semialgebraic subsets of $R^m$. We will use in this section substitution theorems applied to homomorphisms $\mathcal{S}^\diamond(M,R) \to F_i$ for $i = 0,1$, to produce prime and maximal ideals of $\mathcal{S}^\diamond(M,R)$ with nice properties.

**Definition 3.1.** A *formal path* is a tuple $\alpha := (\alpha_1,\dots,\alpha_m) \in R[[\mathtt{t}]]^m$. If $\alpha \in R[[\mathtt{t}]]^m_{\mathrm{alg}}$, there exists $\varepsilon > 0$ such that the map $[0,\varepsilon] \to R^m$, $t \mapsto \alpha(t)$ is semialgebraic. Conversely, each semialgebraic map $[0,1] \to R^m$ defines an element $\alpha \in R[[\mathtt{t}]]^m_{\mathrm{alg}}$. The elements of $R[[\mathtt{t}]]^m_{\mathrm{alg}}$ are called *semialgebraic paths.* If $\alpha \in F_i^m$ for $i = 0,1$ we denote

$$\psi_\alpha := \operatorname{ev}_{M_{F_i},\alpha} \circ \mathtt{i}_{M,F_i} : \mathcal{S}(M,R) \to F_i \quad \text{and}$$
$$\varphi_\alpha := \operatorname{ev}_0 \circ \psi_\alpha|_{\mathcal{S}^*(M,R)} : \mathcal{S}^*(M,R) \to R,\ f \mapsto \psi_\alpha(f)(0).$$

We recall some concepts, notations and elementary results that will appear in the sequel (see [FG3] for further details). An ideal $\mathfrak{a}$ of $\mathcal{S}^\diamond(M,R)$ is *free* if $\bigcap_{f\in\mathfrak{a}} Z_M(f) = \varnothing$. Otherwise it is called a *fixed* ideal. Given a point $p \in M$ the *fixed maximal ideal associated to* $p$ is the kernel $\mathfrak{m}_p^\diamond$ of the surjective homomorphism $\mathcal{S}^\diamond(M,R) \to R$, $f \mapsto f(p)$. We denote

$$\mathfrak{m}_p := \mathfrak{m}_p^\diamond \quad \text{if} \quad \mathcal{S}^\diamond(M,R) = \mathcal{S}(M,R) \quad \text{and} \quad \mathfrak{m}_p^* := \mathfrak{m}_p^\diamond \quad \text{if} \quad \mathcal{S}^\diamond(M,R) = \mathcal{S}^*(M,R).$$

An ideal $\mathfrak{a}$ of $\mathcal{S}(M,R)$ is a *z-ideal* if given $f,g \in \mathcal{S}(M,R)$ such that $Z_M(f) \subset Z_M(g)$ and $g \in \mathfrak{a}$, then $f \in \mathfrak{a}$. It was proved in [GJ] that the ring of continuous functions on a completely regular topological space is a *Gelfand ring,* that is, each prime ideal is contained in a unique maximal ideal. Essentially the same proof shows that $\mathcal{S}^\diamond(M,R)$ is a Gelfand ring. Let $\mathfrak{m}$ be a maximal ideal of $\mathcal{S}(M,R)$. Then $\mathfrak{m} \cap \mathcal{S}^*(M,R)$ is a prime ideal of $\mathcal{S}^*(M,R)$ and we denote $\mathfrak{m}^*$ the unique maximal ideal of $\mathcal{S}^*(M,R)$ containing $\mathfrak{m} \cap \mathcal{S}^*(M,R)$. The set of prime ideals of $\mathcal{S}^\diamond(M,R)$ containing a prime ideal $\mathfrak{p}$ in $\mathcal{S}^\diamond(M,R)$ is a finite chain.

*Remarks* 3.2. Let $M \subset R^m$ be a semialgebraic set, let $\alpha \in M_{F_1}$ be a formal path and denote $p := \alpha(0) \in \operatorname{Cl}(M)$.

(i) There exists a unique homomorphism $\psi_\alpha : \mathcal{S}(M,R) \to F_1$ such that $\psi_\alpha(\pi_i) = \alpha_i$, that is, whose core is $\alpha$. Let us check that $\psi_\alpha(\mathcal{S}^*(M,R)) \subset R[[\mathtt{t}^*]]$. Given $f \in \mathcal{S}^*(M,R)$ and $L \in R$ such that $|f(x)| < L$ for each $x \in M$ we have $L - f(x) > 0$ and $f(x) + L > 0$ for each point $x \in M$. Thus there exist $h_1, h_2 \in \mathcal{S}^*(M,R)$ such that $h_1^2 = L - f$ and $h_2^2 = f + L$. Consequently,

$$L - \psi_\alpha(f) = \psi_\alpha(h_1^2) = \psi_\alpha(h_1)^2 \geq 0 \quad \text{and} \quad \psi_\alpha(f) + L = \psi_\alpha(h_2^2) = \psi_\alpha(h_2)^2 \geq 0,$$

so $|\psi_\alpha(f)| \leq L$. Hence, $\psi_\alpha(f) \in R[[\mathtt{t}^*]]$.

(ii) With the notations in 3.1, $\mathfrak{m}^*_\alpha := \ker(\varphi_\alpha)$ is a maximal ideal of $\mathcal{S}^*(M,R)$. Note that if $p \in M$, then $\mathfrak{m}^*_\alpha = \mathfrak{m}^*_p$, whereas $\mathfrak{m}^*_\alpha$ is a *free* maximal ideal of $\mathcal{S}^*(M,R)$ if $p \in \operatorname{Cl}(M) \setminus M$. In the latter case we call $\mathfrak{m}^*_\alpha$ *the free maximal ideal of* $\mathcal{S}^*(M,R)$ *associated with* $\alpha$.

(iii) Let $\alpha \in R[[\mathtt{t}]]_{\mathrm{alg}} \cap M_{F_0}$ be a non-constant semialgebraic path. We have $\operatorname{tr\,deg}_R(R(\alpha)) = 1$. Then $\mathfrak{p}_\alpha := \ker(\psi_\alpha)$ is a prime $z$-ideal and by [FG3, Thm. 1.3] $\operatorname{qf}(\mathcal{S}^*(M,R)/\mathfrak{p}_\alpha)$ is isomorphic to $R((\mathtt{t}^*))_{\mathrm{alg}}$, because $\operatorname{qf}(\mathcal{S}^*(M,R)/\mathfrak{p}_\alpha)$ is a real closed field of transcendence degree 1 over $R$.

**Proposition 3.3.** *Let* $p \in \operatorname{Cl}(M)\setminus M$ *and let* $\alpha \in F_1$ *be such that* $p = \alpha(0)$. *Then* $\mathfrak{m}_\alpha := \ker(\psi_\alpha)$ *is the free maximal ideal of* $\mathcal{S}(M,R)$ *satisfying* $\mathfrak{m}_\alpha \cap \mathcal{S}^*(M,R) \subset \mathfrak{m}^*_\alpha$. *In addition,*

(i) *The real closed field* $\mathcal{S}(M,R)/\mathfrak{m}_\alpha$ *has transcendence degree* $\mathrm{d}_M(\mathfrak{m}_\alpha) = \operatorname{tr\,deg}_R(R(\alpha))$.
(ii) *There is no prime ideal between* $\mathfrak{m}_\alpha \cap \mathcal{S}^*(M,R)$ *and* $\mathfrak{m}^*_\alpha$

*Proof.* Suppose $\mathfrak{m}_\alpha$ is not a maximal ideal. Then there exists a prime ideal $\mathfrak{q}$ of $\mathcal{S}(M,R)$ such that $\mathfrak{m}_\alpha \subsetneq \mathfrak{q}$ and we choose $f \in \mathfrak{q}\setminus\mathfrak{m}_\alpha$. Substituting $f$ by $f/(1+|f|)$ if necessary we may assume that $f$ is bounded. Write

$$\|\alpha(\mathtt{t}) - p\| := c\mathtt{t}^d + \cdots \quad \text{where} \quad c \neq 0 \quad \text{and} \quad d := \omega(\|\alpha(\mathtt{t}) - p\|) \in \mathbb{Q}^+.$$

As $\psi_\alpha(f) \in R[[\mathtt{t}^*]]\setminus\{0\}$ we can write $\psi_\alpha(f)(\mathtt{t}) := a\mathtt{t}^b + \cdots$ where $a \neq 0$ and $b := \omega(\psi_\alpha(f)) \in \mathbb{Q}^+$. The closed semialgebraic subset

$$Z := \{x \in M : |a| \cdot \|x - p\|^{b/d} \leq 2c^{b/d} \cdot |f(x)|\}$$

of $M$ satisfies $Z = Z_M(g)$, where $g := \operatorname{dist}(\cdot, Z) \in \mathcal{S}(M,R)$. Observe that

$$|a| \cdot \|\alpha(\mathtt{t}) - p\|^{b/d} = c^{b/d}|a|\mathtt{t}^b + \cdots \quad \text{and} \quad \psi_\alpha(|f|)(\mathtt{t}) = |a|\mathtt{t}^b + \cdots,$$

so $\alpha \in Z_{F_1}$. Consequently, $\psi_\alpha(g) = g_{F_1}(\alpha) = 0$ or, equivalently, $g \in \mathfrak{m}_\alpha \subset \mathfrak{q}$.

Then $h := f^2 + g^2 \in \mathfrak{q}$. As $p \notin M$, we have $Z_M(h) = \varnothing$, so $h \in \mathfrak{q}$ is a unit in $\mathcal{S}(M,R)$, which is a contradiction. Thus, $\mathfrak{m}_\alpha$ is the unique maximal ideal of $\mathcal{S}(M,R)$ satisfying $\mathfrak{m}_\alpha \cap \mathcal{S}^*(M,R) \subset \mathfrak{m}^*_\alpha$.

Next we prove statements (i) and (ii).

(i) It follows from Proposition 2.4 (iv) that the real closed field $\mathcal{S}(M,R)/\mathfrak{m}_\alpha$ has transcendence degree $\mathrm{d}_M(\mathfrak{m}_\alpha) = \operatorname{tr\,deg}_R(R(\alpha))$.

(ii) Let $\mathfrak{p}_0 := \mathfrak{m}_\alpha \cap \mathcal{S}^*(M,R) \subsetneq \cdots \subsetneq \mathfrak{p}_r = \mathfrak{m}^*_\alpha$ be the collection of all prime ideals of $\mathcal{S}^*(M,R)$ containing $\mathfrak{p}_0$. Let $(X, \mathtt{j})$ be a semialgebraic bounded-closure of $M$ such that there exists a chain of prime ideals $\mathfrak{q}_0 \subsetneq \cdots \subsetneq \mathfrak{q}_r$ of $\mathcal{S}(X,R)$ with $\mathfrak{q}_i = \mathfrak{p}_i \cap \mathcal{S}(X,R)$. We may assume $X \subset R^m$ and notice that $\mathfrak{q}_0 = \ker(\psi_{\mathtt{j}\circ\alpha})$, where $\mathtt{j} \circ \alpha \in R[[\mathtt{t}^*]]^m$. After reparameterizing $\alpha$ if necessary, we may assume $\mathtt{j} \circ \alpha \in R[[\mathtt{t}]]^m$. Proceeding as in (i) one finds a semialgebraic function $h \in \mathfrak{q}_1$ such that $Z_X(h) = \{q\}$, where $q := \mathtt{j}(p)$, so $\mathfrak{q}_1 = \mathfrak{m}_q$, that is, $r = 1$ as required. □

**Proposition 3.4.** *Suppose that* $M$ *is non bounded-closed and pure dimensional and let* $\alpha \in M_{F_1}$ *be a formal path such that* $p := \alpha(0) \in M$.

(i) *Then* $\mathfrak{p}_\alpha := \ker(\psi_\alpha)$ *is a prime* $z$*-ideal of* $\mathcal{S}(M,R)$ *of coheight* 1 *contained in* $\mathfrak{m}_p$.
(ii) *For each* $\ell = 1, \ldots, d := \dim(M)$ *there exists a free maximal ideal* $\mathfrak{m}_\alpha$ *of* $\mathcal{S}(M,R)$ *associated to a formal path* $\alpha \in R[[\mathtt{t}]]^m$ *such that the transcendence degree over* $R$ *of the real closed field* $\mathcal{S}(M,R)/\mathfrak{m}_\alpha$ *equals* $\ell$.

*Proof.* (i) Using a semialgebraic triangulation of $M$ we construct a bounded-closed semialgebraic subset $K$ of $M$ such that $p \in K$ and $\alpha \in K_{F_1}$. As $K$ is closed in $M$ the homomorphism

$\phi : \mathcal{S}(M,R) \to \mathcal{S}(K,R)$ induced by the inclusion of $K$ in $M$ is surjective by Tietze's semialgebraic theorem [DK2]. Thus $\mathfrak{q} = \phi^{-1}(\phi(\mathfrak{q}))$ for each prime ideal $\mathfrak{q}$ of $\mathcal{S}(M,R)$ that contains $\mathfrak{p}_\alpha$. Proceeding similarly to part (i) in Proposition 3.3 one finds $h \in \mathfrak{q}$ such that $Z_M(h) = \{p\}$. Consequently, $\mathfrak{q} = \mathfrak{m}_p$, so $\mathfrak{p}_\alpha$ is a prime $z$-ideal of coheight one.

(ii) Using a semialgebraic triangulation of $M$, we can choose a formal path $\alpha \in \mathbb{R}[[\mathtt{t}]]^m \cap M_{F_1}$ such that $\alpha(0) \in \operatorname{Cl}(M) \setminus M$ and $\operatorname{tr}\deg_R(R(\alpha)) = \ell$. □

It is possible to describe all prime $z$-ideals $\mathfrak{p}$ of $\mathcal{S}(M,R)$ with $\mathrm{d}_M(\mathfrak{p}) = 1$.

**Proposition 3.5.** *Suppose that* $M \subset R^m$ *is bounded and let* $\mathfrak{p}$ *be a prime* $z$*-ideal of* $\mathcal{S}(M,R)$ *such that* $\mathrm{d}_M(\mathfrak{p}) = 1$. *Then* $\mathfrak{p} = \ker(\psi_\alpha)$ *for a suitable semialgebraic path* $\alpha : [0,1] \to R^m$ *with* $\alpha((0,1]) \subset M$. *In addition,* $R((\mathtt{t}))_{\mathrm{alg}} = \mathrm{qf}(\mathcal{S}(M,R)/\mathfrak{p})$. *In particular, if* $\mathfrak{p}$ *is free then it is a maximal ideal of* $\mathcal{S}(M,R)$.

To finish we present two easy consequences of the substitution theorems in case $R := \mathbb{R}$ is the field of real numbers.

**Example 3.6.** Let $X$ be a closed disc centered at the origin $p := (0,0)$ of $\mathbb{R}^2$, which is a semialgebraic bounded-closure of the semialgebraic set $M := X \setminus \{p\}$ and, for a suitable positive real number $\varepsilon$, let $\alpha : (0,\varepsilon) \to M$, $t \mapsto (t, e^t - 1)$. As $\alpha \in F_1$, the ideal $\mathfrak{m}^*_\alpha = \ker(\varphi_\alpha)$ is a maximal ideal of $\mathcal{S}^*(M,R)$ and $\mathfrak{m}_\alpha = \ker(\psi_\alpha)$ is the unique maximal ideal of $\mathcal{S}(M,\mathbb{R})$ such that $\mathfrak{m}_\alpha \cap \mathcal{S}^*(M,\mathbb{R}) \subset \mathfrak{m}^*_\alpha$. In addition, $\mathfrak{m}_\alpha$ is a minimal prime ideal of $\mathcal{S}(M,\mathbb{R})$ because $\mathrm{d}_M(\mathfrak{m}_\alpha) = 2$. As the chain $\mathfrak{m}_\alpha \cap \mathcal{S}^*(M,\mathbb{R}) \subsetneq \mathfrak{m}^*_\alpha$ does not admit refinaments, $\mathrm{ht}(\mathfrak{m}^*_\alpha) = 1$. However $\mathrm{ht}(\mathfrak{m}^*_\alpha \cap \mathcal{S}(X,\mathbb{R})) = 2$, because $\mathfrak{m}^*_\alpha \cap \mathcal{S}(X,\mathbb{R}) = \mathfrak{m}_p$.

Let us present a Nullstellensatz for the ring $\mathcal{S}^*(M,\mathbb{R})$ whose statement is reminiscent of classical Hilbert's Nullstellensatz. It just involves the set $\widetilde{\partial}M$ of maximal ideals of $\mathcal{S}^*(M,\mathbb{R})$ associated to semialgebraic paths $\alpha \in M_{F_0}$. We omit the proof, which is based on the density of $\widetilde{\partial}M$ in the space $\beta^*_sM$ of maximal ideals of $\mathcal{S}^*(M,\mathbb{R})$ endowed with its Zariski topology (see [FG2]). For each $f \in \mathcal{S}^*(M,\mathbb{R})$ let $\mathcal{Z}^*(f) := \{\mathfrak{m}^*_\alpha \in \widetilde{\partial}M : f \in \mathfrak{m}^*_\alpha\}$ and for each ideal $\mathfrak{a}$ of $\mathcal{S}^*(M,\mathbb{R})$ let us denote

$$\mathcal{Z}^*[\mathfrak{a}] := \{\mathcal{Z}^*(f) : f \in \mathfrak{a}\} \quad \text{and} \quad \mathcal{J}(\mathcal{Z}^*[\mathfrak{a}]) := \{f \in \mathcal{S}^*(M,\mathbb{R}) : \mathcal{Z}^*(f) \in \mathcal{Z}^*[\mathfrak{a}]\}.$$

We have the following result.

**Theorem 3.7** (Nullstellensatz). *For each ideal* $\mathfrak{a}$ *in* $\mathcal{S}^*(M,\mathbb{R})$ *we have* $\mathcal{J}(\mathcal{Z}^*[\mathfrak{a}]) = \sqrt{\mathfrak{a}}$.

## REFERENCES

[B] S. Basu: Algorithms in Real Algebraic Geometry: a survey, inside [S. Basu, J. Huisman, K. Krzysztof, V. Powers, J.-P. Rolin: Real Algebraic Geometry. *Real Algberaic Geometry*, Rennes, **2011**, Jun 2011, Rennes, France. hal-00609687]

[BL] P. Biström, M. Lindström: Function algebras on which homomorphisms are point evaluation on sequences. *Manuscripta Math.* **73**, no. 1, (1991) 179–185.

[BL1] P. Biström, M. Lindström: Homomorphisms on $C^\infty(E)$ and $C^\infty$-bounding sets. *Monatsh. Math.* **115**, no. 3, (1993) 257–266.

[BCR] J. Bochnak, M. Coste, M.-F. Roy: Real algebraic geometry. *Ergeb. Math.* **36**, Springer-Verlag, Berlin (1998).

[CC] M. Carral, M. Coste: Normal spectral spaces and their dimensions. *J. Pure Appl. Algebra* **30** (1983), 227–235.

[DK1] H. Delfs, M. Knebusch: On the homology of algebraic varieties over real closed fields. *J. Reine Angew. Math.* **335** (1982), 122–163.

[DK2] H. Delfs, M. Knebusch: Separation, Retractions and homotopy extension in semialgebraic spaces. *Pacific J. Math.* **114** (1984), no. 1, 47–71.

[E] G.A. Efroymson: Substitution in Nash functions. *Pacific J. Math.* **63** (1976), no. 1, 137–145.

[Fe1] J.F. Fernando: On chains of prime ideals in rings of semialgebraic functions. *Q.J. Math.* **65**, no. 3, (2013), 893–930.

[Fe2] J.F. Fernando: On the substitution theorem for rings of semialgebraic functions. *J. Inst. Math. Jussieu* **14** (2015), no. 4, 857–894.

[FG1] J.F. Fernando, J.M. Gamboa: On the irreducible components of a semialgebraic set. *Internat. J. Math.* **23** (2012), no. 4, 1250031, 40 pp.

[FG2] J.F. Fernando, J.M. Gamboa: On Łojasiewicz's inequality and the Nullstellensatz for rings of semialgebraic functions. *J. Algebra* **399** (2014), 475–488.

[FG3] J.F. Fernando, J.M. Gamboa: On the Krull dimension of rings of continuous semialgebraic functions. *Rev. Mat. Iberoam.* **31**, no. 3, (2015), 753–766.

[GGJ] M.I. Garrido, J. Gómez Gil, J.A. Jaramillo: Homomorphisms on function algebras. *Canad. J. Math.* **46** (1994) no. 4, 734–745.

[GJ] L. Gillman, M. Jerison: Rings of continuous functions. *The Univ. Series in Higher Nathematics* **1**, D. Van Nostrand Company, Inc. (1960).

[GL] J. Gómez, J.G. Llavona: Multiplicative functionals on function algebras. *Rev. Mat. Univ. Complut. Madrid* **1** (1988), no. 1-3, 19–22.

[J] J.A. Jaramillo: Topologies and homomorphisms on algebras of differentiable functions. *Math. Japon.* **35** (1990), no. 2, 343–349.

[KMS] A. Kriegl, P. Michor, W. Schachermayer: Characters on algebras of smooth functions. *Ann. Global Anal. Geom.* **7** (1989), no. 2, 85–92.

[R] M.F. Roy: Using infinitesimals for algorithms in real algebraic geometry. IRMAR (UMR CNRS 6625), Univ. Rennes, April **12** (1989).

[Rz] J.M. Ruiz: A dimension theorem for real spectra. *J. Algebra* **124** (1989), no. 2, 271–277.

[S2] N. Schwartz: Real closed rings. Algebra and order (Luminy-Marseille, 1984), 175–194, *Res. Exp. Math.*, 14, Heldermann, Berlin (1986).

[S3] N. Schwartz: The basic theory of real closed spaces. *Mem. Amer. Math. Soc.* **77** (1989), no. 397.

[S4] N. Schwartz: Rings of continuous functions as real closed rings. Ordered algebraic structures (Curaçao, 1995), 277–313, Kluwer Acad. Publ., Dordrecht (1997).

[S5] N. Schwartz: Real closed spaces. Ordered fields and real algebraic geometry (Boulder, Colo., 1983). *Rocky Mountain J. Math.* **14** (1984), no. 4, 971–972.

[Z] W.R. Zame: Homomorphisms of rings of germs of analytic functions. *Proc. Amer. Math. Soc.* **33** (1972), 410–414.

DEPARTAMENTO DE ÁLGEBRA, FACULTAD DE CIENCIAS MATEMÁTICAS, UNIVERSIDAD COMPLUTENSE DE MADRID, 28040 MADRID (SPAIN)

*E-mail address*: `josefer@mat.ucm.es`, `jmgamboa@mat.ucm.es`