Notes written by Alun C. Stokes

Course Notes on O-minimality and the Pila-Wilkie Theorem

Presented by Gareth Jones as part of the Fields Institute Graduate Course on O-minimality and Applications

May 2022

Springer Nature

As the instructor for this module, these lecture notes are dedicated to Dr Gareth Jones, whose presentation of the Pila-Wilkie Theorem and its proof was both illuminating and intellectually challenging for this young graduate student.

Preface

This collection of notes results from a series of 8 lectures given by Dr Gareth Jones at the Fields Institute over the course of 19 Jan to 11 Feb 2022. The vast majority of the material presented here was given as part of this first module in the three-module course on o-minimality and Applications. Primarily, attempt was made to uniformize notation and formatting across the various lectures, and some extra details added where otherwise missing or left as an exercise to the viewer – or even more likely, of some confusion to this author.

This current iteration of the notes is to be considered preliminary, and the style and particularities of their presentation are by no means final. As the owner of the sole pair of eyes to have read this document, no assurances are made that it is free of mistakes or oversights (in fact, one might be more willing to make promises to the contrary) – but to the best of our ability, we have tried our best to minimize the sure-to-come lists of errata and corrigenda upon review.

This document intends to act as an easily (to an early graduate student in mathematics) accessible introduction to o-minimality in the context of expansions of the real field and their application in proving the Pila-Wilkie theorem. In particular, o-minimality is not discussed in full generality – rather, the role that o-minimality of ordered fields plays in results on semi-algebraic sets, cell decompositions and parameterization by these decompositions, and of course, how this all comes together to prove the eponymous theorem of Pila and Wilkie are discussed. The ordering of content presented is not strictly adherent to the delineation given in the lectures, but it does broadly follow.

These notes are written, as most things are, from particular authorial perspective – which is to say, not by those with more than little to any background in model theory. There may be at points a belabouring of ideas that another would find trivial, unnecessary, or otherwise not necessarily worth the space they take up on the page. The purpose of compiling these notes is not just to archive the lecture series given but also to make it somewhat more accessible by clarifying what we ourselves had to further investigate and understand as we attended the course. This does not mean any significant amount of material is added above and beyond the course content itself – rather, more so that some ‘one-liners’ in the original presentation are afforded just a few more in these notes, and some contextualization added in the first chapter to bridge the gap for a student for whom o-minimality is completely novel.

Where not otherwise cited, facts should be supposed to have been taken from the lectures — which themselves will periodically include references or suggestions for additional reading material. All citations are included for exteriorly sourced information, with the full list of references listed at the end of the document. This is contrary to the Springer Nature style of chapter-end citations, but the struggle with incorporating both this feature and BibTeX at the hands of this authorial team was nothing short of nightmarish. And, as with so many things in life, BibTeX should always come first.

Although at times dry, the content here, especially once one delves into it and splashes about in the waters of its intrigue, is quite the scene to behold. However, we would be remiss to assume this view is shared by one new to the subject area, or only acquainted with logics and not number theory, vice versa – or (may the good lord help you) neither. Attempt at levity is made throughout in order to keep things interesting, whilst of course still rigorous, complete, and unobtrusive. Appropriateness of this attempt is yet to be determined, but it is the philosophy of this author that texts are to be enjoyed – perhaps even the otherwise driest more so than any other; how else can we be expected to pay sufficient and adequate attention unless we are sufficiently engaged, amused, and interested all at once. Perhaps this is just one author's opinion unshared by the mathematical community at large – but either way, you, my reader dearest, have no choice in the matter, and so shall have to decide for yourself the merit and suitability of the approach for the matters at hand. With all that in mind, please enjoy.

Stokes, Alun C. *McMaster University*
May 2022 *MATH 712*

Acknowledgements

I would like to thank Dr Gareth Jones for his wonderful series on the content presented, along with Drs Ya'acov Peterzil and Jacob Tsimerman, for jointly organizing the course – and each of whom instructed fascinating courses of their own, building broadly on the material that this document attempts to capture. I would further like to thank the Fields Institute for hosting this graduate course and allowing me a seat to attend (of the figurative type, of course), and finally, I am eminently grateful to Dr Patrick Speissegger as the instructor of record for MATH 712 at McMaster University, and his accommodation and understanding with the preparation of this manuscript. More so for the second part there, if I'm being honest.

Contents

Acronyms

The following are used varyingly judiciously throughout.

CD Cell Decomposition
PW Pila-Wilkie
DLO Dense Linear Order
FOL First-Order Language
MT Monotonicity Theorem
FT Finiteness Theorem
UF Uniform Finiteness
QE Quantifier Elimination
DC Definable Choice
CS Curve Selection

Part I
O-minimality and Necessary Concepts

This introductory part will focus primarily on setting up the setting in which we find ourselves working for most of the course. Highlights include "*what even is o-minimality?*" and "*but how does that imply Pila-Wilkie?*" – and perhaps a favourite of mine: "*what is the Pila-Wilkie theorem?*". While the latter two are answered in much more detail, later on, this part will take us quickly through some of the significant results we'll find ourselves needing in a bit. In particular, we prove the monotonicity theorem (MT), define cell decompositions (CD) and much later on *smooth* cell decompositions – then prove the Cell Decomposition theorem itself, Later on, we discuss dimensionality and the agreement between the definable (or geometric) and algebraic (or model-theoretic) notions of the quantity – something that will come to be quite useful to us later on. We finally end off with that promised discussion of smooth cell decompositions, although in comparison to other sections, it really is more a series of results and sketches (if that) of how one might prove these things, as they broadly build on either proof structures we will have seen before, or shouldn't be too hard in conjunction with a dusting off of evertone's favourite: Calculus: Early Transcendentals – of which I'm sure we all have a copy somewhere.

The prepared reader should find themselves acquainted with preliminary ideas in mathematical logic (nothing further than one would find in an undergraduate class on the subject) and the basics of field theory. Again, nothing further than an undergraduate would necessarily be expected to have encountered.

A reader should leave this section feeling themselves reasonably well-acquainted with some of the tools they might see themselves using in general sorts of proofs about o-minimal structures and how they may go about proving or disproving something to be o-minimal. Some preliminary notions of how this all fits into counting points of bounded height on curves may be coming to surface by the end of this part, but the novice reader would be well-forgiven were that not the case. They should, however, feel comfortable identifying what is 'clearly' a definable set and be able to do some rudimentary reasoning on how we can use the finiteness and definability of one of a pair of complementary sets to say something about the definability of the other. The idea of cells and cell decompositions should be reasonably understood (at least *intuitively* if not in full technical detail). The reader should have a map of sorts in their mind that sequentializes and connects the discussed matter in a reasonable and meaningful way – as these preliminary ideas form the basis for the larger lemmata and theorems to come. By the end of this part, the well-established reader should find the proof presented rather intuitive, and even (hopefully) find that they belabour the points they make *too* much for how clearly obvious they are.

Chapter 1
Mapping the Landscape and Telegraphing our Journey

Abstract How one starts with the Tarskian ideas introduced in their first course in mathematical logic and ends up studying o-minimality to the end of proving the Pila-Wilkie theorem is perhaps and unsurprisingly eminently unclear to any who do not already know the methodology. So we don't become too disenfranchised and uninspired as we work our way through the logical results to the end of a number theoretic result, we layout a brief map of what is to come and how each part logically follows from its antecedent. This is not perhaps the most interesting course for one interested solely in number theory or simply mathematical logics, but where these two unlikely friends collide creates something wondrous and beautiful. We then go on with just a few words in preparation for what is to come; some definitions, expectations of the experience (or lack thereof) on the part of the reader, and a general outline are given. The overly excited reader may feel free to skip right onto Chapter 2, but this section serves as just a bit of an *amuse-bouche* for those not so ready to jump right in.

1.1 And Pila-Wilkie is?

Perhaps the best and most prudent question to be asking one's self currently, if for no other reason than determine the worth of their time in reading this whole affair, is what the statement of the Pila-Wilkie theorem *actually* is? And what, supposing the reader knows the context in which we define o-minimality, could that have to do with a number theoretic result like the Pila-Wilkie theorem? Well, dear reader, we hope in this brief first section to enlighten you to the big ideas upon which we will ruminate for the remainder of this course, and provide a coarse outline of how these come to build on one other in order to start from the relatively basic, to the phenomenal. To live up to this section's name, however, we now state the Pila-Wilkie theorem– first informally, and then as we will come to prove it. In all cases, however, note that we are speaking to o-minimal expansions of the *real* field, and we will not be covering the theorem and all that leads up to it in full generality.

Theorem 1.1 (Pila-Wilkie theorem (Informal)) *Let $\tilde{\mathbb{R}}$ an o-minimal expansion of $(\mathbb{R}, <)$. Then **transcendental** definable sets have very **few** rational points.*

Easily understandable, seems reasonable, and (maybe?) doable without too much fuss, doesn't it seem? Here now is the formal statement, which requires the following crash course in notation; we denote by H the usual rational multiplicative height function – but it also doubles, when not used as a function, as an upper-bound on rational numbers we are interested in (I didn't decide the notation). As well, when taking vectorial input, the height is given by the element-wise maximum. Suppose $X \subseteq \mathbb{R}^n$. Denote $X(\mathbb{Q}) := X \cap \mathbb{Q}$ and further $X(\mathbb{Q}, H) \subseteq X(\mathbb{Q})$ with element height as given by the function, H, and bounded by the constant, H. Finally, the superscripts *tr* and *alg* refer to the transcendental and algebraic parts of the set they sit atop (note of course that necessarily we have $X \setminus X^{\mathrm{tr}} = X^{\mathrm{alg}}$). What we then want, and what Pila-Wilkie gives us, are good bounds on $|X^{\mathrm{tr}}(\mathbb{Q}, H)|$.

Theorem 1.2 (Pila-Wilkie theorem (Formal)) *Let $X \subseteq \mathbb{R}^n$ definable in $\tilde{\mathbb{R}}$ an o-minimal expansion of $(\mathbb{R}, <)$. Then for any ε there exists some c such that for all H,*

$$|X^{tr}(\mathbb{Q}, H)| \leq c \cdot H^{\varepsilon},$$

which is a very fine thing indeed – that last bit being an editorial note, and not (necessarily) part of the honest theorem.

1.2 There but Not Back Again

How we shall come to prove this is not direct by most meanings or usages of the term, and in fact, is not going to follow the proof originally given by the eponyms for the theorem. Rather, we will more so be following a later, arguably nicer proof given just this year (if you are reading this in 2022) by Bhardwaj and van den Dries [**?**], that makes clever use of semialgebraic cell decomposition for a large part of the original proof, and elsewhere use a proof of the Yomdin-Gromov theorem given by Binyamini and Novikov [**?**] also very recently.

The broad strokes are as follows, with some bits left out that are in majority just definitional. Broadly, this encompasses most of Part 1 of this report – as this is, after all, a graduate course on this topic. We start from the very basics with definably, o-minimality, the Finiteness Theorem in 2 variables (which goes on to be used more broadly), and then cell decompositions. With basics in place, we define definable choice and curve selection, and then go on to show that definable and algebraic dimension are in agreement. Closer to the majority of the matter, we introduce semialgebraicity, and go on to prove the above mentioned theorems that, when all put together, allows us to produce a proof of Pila-Wilkie. Some parts are taken for granted (for example, assuming uniform finiteness without proof), and in general, we treat only expansions of the real field rather than a more generic structure – and this allows us to use a few tricks such as the Baire Category theorem at one point. In whole, however, this set of notes constitutes most all but some of the trickier details in a fully self-contained proof of the Pila-Wilkie theorem, and the reader is directed to where any gaps may be addressed when present.

1.3 Before the Lectures Proper

We feel somewhat compelled to address an aspect of this topic that we felt was slightly neglected in the curriculum. Had one seen the list of attendees to these lectures, the reason for skipping over such ‘trivialities’ as we are about to point out briefly is clear — with several attendees being former students of Pila or Wilkie themselves. Still, for the *not-even-amateur* logician, the following certainly bears some explicit mention.

In the absence of the results we come to find, o-minimality may appear a relatively unmotivated idea to study. Of course, as the pure mathematicians we are (or hope one day to be), why *shouldn't* the mere concept of further understanding be enough to compel our interest? Still, the progression from introductory mathematical logic and the importance and usefulness of quantifier elimination (QE) to o-minimality is one that was unapparent to this author until being noted elsewhere. We don't claim this to be a failure of the course so much as a failure in personal preparation, but should the reader find themselves similarly underprepared, then they will find themselves thankful for this little pretext.

To keep things brief, we will say just this: in a predicate logical system, we are interested in quantifier elimination. The result of the elimination of quantifiers is essentially the answer to the question a quantified statement asks. Perhaps the most famous example of this is the existence of real roots of quadratic equations. We ask the quantified ‘question’:

$$\exists x \in \mathbb{R}.(a \cdot x^2 + b \cdot x + c = 0 \wedge a \neq 0) \tag{1.1}$$

— that is, does there exist such an x? The quantifier eliminated equivalent form is

$$b^2 - 4 \cdot a \cdot c \geq 0 \wedge a \neq 0, \tag{1.2}$$

the first half of which should be recognized as the quadratic discriminant (and the second just to ensure non-degeneracy). Here, quantifier elimination gives the exact and deterministic characterization of the answer to the quantified statement – and it is this property that motivates its study. We trust at this point that the motivation has been sufficiently belaboured.

It is known that first-order theories with QE (that is, decidability for the theory can be reduced to the question of satisfaction of quantifier-free sentence in the theory) are model complete. In the interest of not straying too far, we leave it to the reader to believe or convince themselves that this is a desirable property. While this is not the focus of how we define o-minimality in this course, a structure is indeed o-minimal exactly if every formula given by no more than one free variable and some subset of M-parameters is equivalent to a quantifier-free formula defined only by these parameters, and the ordering on the structure [**?**]. Thus, for anyone finding themselves perhaps unconvinced upfront of the merit of some of the ideas explored here (outside the Pila-Wilkie Theorem), we hope this motivates the sequence we are about to take on. And for everyone else, we hope that this section did not bore you too thoroughly.

1.4 Preliminary Definitions

Throughout, we will be working with models $\mathcal{M} = (M, <)$ of the theory of dense linear orders (DLO) *without endpoints*. For now, M will be fixed, but we will look at some specific instances later on. Perhaps then, one of the most important definitions, to begin with, is that of *definability*.

Definition 1.1 (Definability of sets (without parameters)) For $n \in \mathbb{N}$, we say a set $A \subseteq M^n$ is *definable without parameters* if there exists some formula in our model, φ, satisfied exactly by the elements of A.

Definition 1.2 (Definability of sets) For $X \subseteq M$ and $n \in \mathbb{N}$, then we say a set $A \subseteq M^n$ is *definable with parameters from* X if there exists some formula in our model, φ, and elements $b_1, \ldots, b_m$, such that φ is satisfied exactly by the elements of A along with the parameters in X.

Notice then that definability without a parameter is simply the case of definability with parameters coming from the empty set. These definitions immediately and naturally induce the idea of definable functions and definable points. In particular, a function is definable in parameters if its graph is definable by those same parameters in $\mathcal{M} = (M, <)$. Similarly, an element a is definable in $\mathcal{M}$ (with parameters) if the singleton $\{a\}$ is definable in $\mathcal{M}$ by those same parameters. This isn't something we will need to consider too extensively.

As ever, when introduced to a novel space, we are interested in what its open intervals look like. We have the following characterization:

Definition 1.3 (Open Interval) A set, $A \subset M$ is an open interval in $(M, <)$ if A is of one of the following forms:

- (a, b) with $a < b \in M$
- $(-\infty, a)$ with $a \in M$
- (a, ∞) with $a \in M$

We say further that intervals of the first type — that is, those having finite bounds — are *bounded*. Easy to miss but important to note is that the endpoints must sit inside our domain. So, for example, in $(\mathbb{Q}, <)$, the set $(-5, \sqrt{7})$ is *not* an open interval.

We imbue M with the order topology and M^n with the product topology. We then define what it means to be an o-minimal expansion.

Definition 1.4 (O-minimal expansion) Taking $\mathcal{M} = (M, <, \dots)$ an expansion of $(M, <)$, we say $\mathcal{M}$ is o-minimal if every definable (with parameters) subset of M is given by a finite union of open intervals and points.

If we weaken the above and ask only for *convex* sets (which are a superset of our open intervals) in place of open intervals, then the above would define *weak o-minimality* — but that won't be a topic of discussion here.

For the etymologically inclined, it is noted that the 'o' in o-minimal comes from the shortening of 'order-minimality'. For more information on the history and development of the idea of o-minimality, one may reference **Tame Topology & O-minimal Structures** [**?**] or **Definable Sets in Ordered Structures I** [**?**] and **II** [**?**].

Some (arguably) simple examples of o-minimal structures are given by expansions of the real field. Consider, for example, $\overline{\mathbb{R}} = (\mathbb{R}, <, +, -, \cdot, 0, 1)$ and the further expansion $\mathbb{R}_{\exp} = (\overline{\mathbb{R}}, \exp)$, both of which are o-minimal. Mind not to mistake the use of 'simplicity' as an indication that these are trivial or did not require particular and considerable consideration — rather, just that they have a relatively simple-seeming form. For now and going forward, we fix $\mathcal{M}$ an o-minimal structure and move on to our first theorem.

Chapter 2
Setting it all up

Abstract We now begin properly with a from-the-basics definition of the objects at play: field expansions, monotonicity, cells and decompositions into them, semi-algebraicity and similarly fundamental ideas are each defined and contextualized. Note that we will not be discussing topological definitions in general. That is to say, the reader is assumed to be familiar with the basic point-set topology and the ordinary sorts of topologies we see cropping up (e.g. order, product) – not that topological ideas won't be discussed. As well, basic knowledge of mathematical logic is assumed; first-order languages (FOL), $\mathcal{L}-$Structures, relations, and satisfiability are all presumed familiarities. With definability now a part of our tool-set, we start by proving a few theorems fundamental to results later in this course.

2.1 On Monotonicity

What constitutes a 'nice' property of a function is generally non-contentious; injectivity and surjectivity are often useful – together even more so – and it would be the odd mathematician to turn their nose up at a function being bounded, supposing they weren't chasing a nasty counterexample or engaging in some other such endeavour. At present, we will focus on the property of *monotonicity* and when we can determine a definable function to be monotonic in the context of open intervals. The following was proved in [**?**] by Pillay and Steinhorn:

Theorem 2.1 (The Monotonicity Theorem) *Suppose $f\colon I \to M$ is a definable function for $I \subset M$ an open interval. Then there exist $a_1, \ldots, a_k \in I$ such that on each adjacent interval, (a_j, a_{j+1}) (where $I = (a_0, a_{k+1})$) f is either constant, or strictly monotonic and continuous. Further, if f is definable over some $A \subseteq M$, then so too are $a_1, \ldots, a_k$ definable over A.*

Hence, we will refer to this simply as the Monotonicity theorem, abbreviated by MT. It is perhaps not immediately apparent why this should be true, or even that we should be interested that it is. The answer to the second point is that this piece-wise continuity and monotonicity of definable functions is a relatively rigid condition, and this (not just here but for structures in general) allows us to say a good bit about them. Observe that if we have some $X \subseteq M$ definable and infinite, then X must contain some open interval. This should be relatively intuitive, even if a proof doesn't come to you immediately, given what we've covered thus far. As for why the Monotonicity theorem holds, we show this by piecing together three lemmata that should make the picture a bit more clear. Throughout, take $J \subset I$ as an open interval. To not get bogged down in the minutiae of their proofs as we go through — not that they are particularly challenging — but in any case, we will state all three and then prove them sequentially.

Lemma 2.1 *There is an open interval, $J' \subseteq J$, on which f is constant or injective.*

Lemma 2.2 *If f is injective on J, then there is an open interval, $J' \subseteq J$ on which f is strictly monotonic.*

and finally,

Lemma 2.3 *If f is a strictly monotonic function on J, then there exists some open interval $J' \subseteq J$ on which J is continuous.*

Taking these lemmata for granted, it is not terribly difficult to see how the Monotonicity theorem falls out. The fun then is in proving these three facts — which is nice, as they are not terribly complicated.

We start where any sensible person would.

Proof (Lemma 2.1) Suppose there is some $y \in M$ such that its preimage under f intersected with J is infinite. This necessarily implies the existence of $J' \subseteq J$ an open interval on which f takes constant value, and so we can assume for any $y \in M$ that we have $f^{-1}(y) \cap J$ is finite. Then, we must have f(J) infinite, and so contains interior with subset (a, b), for $a < b$. Taking

$$q \colon (a, b) \to J$$
$$q \colon y \mapsto \min\{x \in J : f(x) = y\},$$

we get q injective — and so this is an open interval $J' \subseteq q((a, b))$ on which f is injective. □

Proof (Lemma 2.2) Write me □

Proof (Lemma 2.3) This we can get quite quickly. Suppose such a strictly monotone function exists on J. Clearly, f cannot be constant (else monotonicity would be non-strict), and so by o-minimality of f, we get that the image of J under f contains some open interval, $J' \subseteq \text{image}(f)$, on which we have preimage a sub-interval of J. We get monotonicity on this interval by Lemma 2.1 and non-constancy (and thus monotonicity) of f; this must be a bijection (either order-preserving or reversing, but bijective either way), and so we are finished. □

Proof (Theorem 2.1) We now combine these three lemmata to get our result. Take A the set of all $x \in I$ (coming from our original theorem statement) such that f is both continuous and strictly monotone at x. We know that taking the restriction of f to some open sub-interval on which f is defined maintains both continuity and monotonicity by Lemmata 2 and 3 — and so taking the set difference of A from I, the original open interval, we cannot have *any* open intervals. There are then thus only finitely many points, and the theorem follows. □

Take note that the proof provided here is *not* precisely the one that was given in the lecture, but rather a bit more condensed, less roundabout method of achieving the result. The strategy is the same, however, differing only in presentation.

Two Exercises Lec1 pg 4

The following result is a special case in 2 dimensions of what is referred to as the *Finiteness theorem*, abbreviated FT. We first prove this special case and then take a brief detour to talk about cell decompositions before we can address the more general theorem.

2.2 The (Planar) Finiteness Theorem

Theorem 2.2 (Finiteness Theorem in M^2) *Suppose $A \subseteq M^2$ and that for each $x \in M$, the fibre A_x above x — that is, the set of y with $(x, y) \in A$ — is finite. Then, there exists some $N \in \mathbb{N}$ such that $|A_x| \leq N$ for all $x \in M$*

Proof (of Theorem 2.2)) We define a point $(a, b) \in M^2$ to be *normal* if it sits in an open box, $I \times J$ satisfying

- $(I \times J) \cap A = \emptyset$
- $(a, b) \in A$
- There exists a continuous $f : I \to M$ such that $(I \times J) \cap A = \operatorname{graph}(f)$.

Similarly, for points with only one finite endpoint, we say some (a, ∞) (resp. $(a, -\infty)$) is *normal* if there exists open interval I such that $a \in I$ and some $b \in M$ such that

$$(I \times (b, \infty)) \cap A = \emptyset$$

and again, respectively taking $(b, -\infty)$ for the other case.

Supposing we take the set $\{ (a, b) \in M^2 : (a, b) \text{ is normal} \}$, it easily follows that this set is definable, and similarly so for the $\{\pm\infty\}$ cases. We now define functions $f_1, f_2, \dots, f_n$ by the property that

$$\operatorname{dom}(f_k) = \{ x \in M : |A_x| \geq k \}.$$

We have the property that $f_k(x)$ is the k-th element of A_x — and so we get the definability of each f_k by the finiteness of each fibre.

Fixing some $a \in M$ and taking $n \geq 0$ maximal such that all of $f_1, \ldots, f_n$ are defined and *continuous* on an open interval around a. We then say that a is

- **good** if $a \notin \mathrm{cl}\,(\mathrm{dom}\,(f_{n+1}))$ and otherwise
- **bad** if a *is* in this closure.

We partition into $G = \{\, a \in M \;:\; a \text{ is good} \,\}$ and $B = \{\, a \in M \;:\; a \text{ is bad} \,\}$. What we will now show is that G is definable — which we do by showing that for any $a \in B$, there is a minimal $b \in M \cup \{\pm\infty\}$ such that (a, b) is *not* normal.

Let $a \in B$. We use the following notation for convenience:

$$\lambda(a, -) = \begin{cases} \lim\limits_{x \to a^-} f_{n+1}(a) & : \ f_{n+1} \text{ defined on } (t, a) \text{ for some } t < a. \\ \infty & : \ \text{else} \end{cases}$$

$$\lambda(a, 0) = \begin{cases} f_{n+1}(a) & x \in \mathrm{dom}\,(f_{n+1}) \\ \infty & : \ \text{else} \end{cases}$$

$$\lambda(a, +) = \begin{cases} \lim\limits_{x \to a^+} f_{n+1}(a) & : \ f_{n+1} \text{ defined on } (a, t) \text{ for some } a < t. \\ \infty & : \ \text{else} \end{cases}$$

Take $\beta(a) = \min\{\lambda(a, -),\ \lambda(a, 0),\ \lambda(a, +)\}$. It is not difficult to see then that $\beta(a)$ is simply the least $b \in M \cup \{\pm\infty\}$ such that (a, b) is not normal. Were we instead to take some $a \in G$, then (a, b) must *always* be normal for any $b \in M \cup \{\pm\infty\}$. So, B can be given as

$$B = \{\, a \in M \;:\; \exists b \in M \cup \{\pm\infty\} \text{ s.t. } (a, b) \text{ is not normal} \,\},$$

and as such, is definable.

If we take some $a \in G$, then $|A_x|$ is constant on an open interval about a by definition of G. By showing that B is finite, we get our desired result. Supposing B to be *infinite*, we can partition B into

$$\begin{aligned} B_+ &= \{\, a \in B \;:\; \exists y \text{ s.t. } y > \beta(a),\ (a, y) \in A \,\} \\ B_- &= \{\, a \in B \;:\; \exists y \text{ s.t. } y < \beta(a),\ (a, y) \in A \,\}, \end{aligned}$$

both evidently definable sets. By the infinitude of B, so too must at least one of B_-, B_+ be infinite — and further, so must one of

- $B_+ \cap B_-$
- $B_+ \setminus B_-$
- $B_- \setminus B_+$

- $B \setminus (B_+ \cup B_-)$.

We can then apply the Monotonicity theorem (Theorem 2.1) to each case to reach a contradiction by showing that assuming non-finiteness, we *should* be able to find a normal point with first coordinate a – contradicting the 'badness' of any point in B. Thus, B is *finite*, and so there must be some finite upper bound on the cardinality of all fibres, A_x, and our proof is complete. □

2.3 Cell Decompositions

We start with a few definitions that should hopefully feel motivated in anticipation of the higher-dimensional analogues of what we have seen already.

Definition 2.1 (Cells in M^n) For a sequence $(i_1, \ldots, i_n)$ for each $i_j \in \{0, 1\}$, we define $(i_1, \ldots, i_n)$-*cells* of M^n inductively as follows:

1. A 0-cell is a point in M, and a 1-cell an open interval (both in M^1).
2. Supposing $(i_1, \ldots, i_n)$-cells are defined for M^n,
   a. we define an $(i_1, \ldots, i_n, 0)$-cell to be a definable set given by graph (f) for f a continuous, definable function on an $(i_1, \ldots, i_n)$-cell.
   b. Perhaps predictably then, we define an $(i_1, \ldots, i_n, 1)$-cell to be a definable set of the form $(f, g)_C = \{ (x, y) \in C \times M \ : \ f(x) < y < g(x) \}$ for f, g continuous, definable functions on an $(i_1, \ldots, i_n)$-cell, with $C \subset M^n$. Note that we may also allow $f \equiv -\infty$ or $g \equiv \infty$.

As usual, we denote a projection map by π, and for any $(i_1, \ldots, i_n)$-cell we can define the projection

$$\pi \colon M^n \to M^k$$

for k the sum of $i_1, \ldots, i_n$, such that the restriction of π to our $(i_1, \ldots, i_n)$-cell is a homeomorphism onto its image.

It is not hard to see that what we are doing here is just projecting away from the coordinate 0 parts of the cell. This can be thought of as a canonical coordinate projection that any cell comes naturally equipped with, which is quite a fine thing to have at hand.

In what should hopefully be predictable at this point, we wish now to define what it means to *decompose* our space into cells. At some point, we will cease prefacing these definitions with 'as usual, we do so by induction' – but that point is yet to come. So, as usual, we proceed by defining cell decompositions by induction.

Definition 2.2 (Cell Decomposition *of* M) A *cell decomposition* of M is a finite set defined by some strictly increasing finite sequence $a_1, \ldots, a_k$ that form the set

$$\{(-\infty, a_1), (a_1, a_2), \ldots, (a_k, \infty), \{a_1\}, \{a_2\}, \ldots \{a_k\}\}.$$

That is, all the sequential open intervals (including those with infinite endpoints) plus the singleton sets. Just for the sake of belabouring the point, this is a definitionally *definable* set.

As we have time and time before, we now up the dimension by induction to define general cell decompositions:

Definition 2.3 (Cell Decomposition *of* M^{n+1}) A cell decomposition of M^{n+1} is a finite partition, $\mathfrak{D}$, of M^{n+1} into cells, such that

$$\{\, \pi(C) \;:\; C \in \mathfrak{D} \,\}$$

is itself a decomposition of M^n, with each π the respective projection as discussed above.

We trust the conjunction of these two definitions into an understanding of cell decompositions in arbitrary dimensions is clear by induction. An idea that may seem a bit apropos until a bit later on is that of *compatibility* – but rest assured, dear reader, that this will all come together shortly.

Definition 2.4 (Compatibility) We call a cell decomposition $\mathfrak{D}$ of M^n *compatible* with a subset, $X \subseteq M^n$ if for each cell, $C \in \mathfrak{D}$, either $C \cap X$ is empty, or C is a subset of X.

2.4 The Cell Decomposition Theorem

The importance of this last point will be made clear in the theorem we have built up to the Cell Decomposition Theorem – which essentially says that these compatible decompositions exist and are compatible with any finite collection of definable sets and, most importantly, that definable functions are continuous on each cell in such a decomposition (defined in the domain of the function, of course). Properly, and as proved by Knight, Pillay, and Steinhorn [**?**]:

Theorem 2.3 (Cell Decomposition) *Take $n \in \mathbb{N}$.*

> *Note the use (about to be made) of subscripts on statements $(I)_n$ and $(II)_n$ to denote the dimension of M^n to which each statement refers. This is going to be notationally useful in the proof, but may seem a bit queer at present, and without introduction.*

Then

$(I)_n$ *Suppose $X_1, \ldots, X_k \subseteq M^n$ are definable sets. Then there is a cell decomposition of M^n compatible with each X_j.*

$(II)_n$ *If $f\colon X \to M$ is definable, then there is a cell decomposition, $\mathfrak{D}$ of M^n compatible with X s.t. the restriction $f|_C$ is continuous for each $C \in \mathfrak{D}$.*

Further, and in analogy to the Monotonicity theorem, if our $X_1, \ldots, X_k$ or f (depending on case (I) or (II)) are definable over $A \subset M$, then we can take the cells in $\mathfrak{D}$ to be similarly definable over A; that is, with the same *parameters.*

This last point is perhaps a bit unfair to mention, as we will not be providing a proof for it – though for the sake of interest, it would feel incomplete to not at least analogize with Theorem 2.1. In truth, what follows is not a full proof of Cell Decomposition, but a special case where we take M to be $\mathbb{R}$ and use the yet unproven (or even stated) result of uniform finiteness. *We take this result entirely for granted in the lectures due to the oddity that a complete (unassuming) proof somewhat 'bootstraps' uniform finiteness into the induction we do on $(I)_j$ and $(II)_j$, proving it as we go along. This is because uniform finiteness is actually itself an immediate consequence of the Cell Decomposition Theorem (which makes the proof a fun little oddity). For our purposes, we take it as assumedly true – in part due to the length of this proof even* with *that assumption – and trust that our dear intelligent reader sees plainly how we could fix this in the absence of the assumption.*

Uniform finiteness is a generalization of the finiteness theorem we proved earlier (Theorem 2.2), but with potentially many parameters and higher dimensions. As with the argument for the Cell Decomposition Theorem, we will similarly restrict our attention to the case where $M = \mathbb{R}$. This special case is as follows:

Proposition 2.1 (Uniform Finiteness (for $\mathbb{R}$)) *Suppose $X \subset \mathbb{R}^{n+1}$ is definable with each fibre X_x finite for $x \in \mathbb{R}^n$. Then there is some $N \in \mathbb{N}$ such that $|X_x| \le N$ for all $x \in \mathbb{R}^n$.*

The following proof is due to van den Dries [**?**] which, for no reason other than interest's sake, we mention went on to inspire the later work of Pillay and Steinhorn in [**?**].

Proof (Cell Decomposition (Theorem 2.3)) We proceed by induction on parameter n. The base cases are both already done for us; $(\mathrm{I})_1$ is immediate from the definition of o-minimality, and $(\mathrm{II})_1$ is given by the Monotonicity theorem. What we go on to show is two inductive facts that 'bounce off' one another in a sense, to allow us to prove both $(\mathrm{I})_n$ and $(\mathrm{II})_n$ for all n. These are

(a) Given $(\mathrm{I})_1, \ldots, (\mathrm{I})_n$ and $(\mathrm{II})_1, \ldots, (\mathrm{II})_{n-1}$, we can conclude $(\mathrm{II})_n$; and
(b) Given $(\mathrm{I})_1, \ldots, (\mathrm{I})_n$ and $(\mathrm{II})_1, \ldots, (\mathrm{II})_n$, we can conclude $(\mathrm{I})_{n+1}$. □

That these two facts together give us the desired result should be clear. Getting there requires a bit more effort, and so we simply begin with (a). Thus we wish to prove $(\mathrm{II})_n$: that for a definable $f\colon X \to M$, there is a cell decomposition $\mathfrak{D}$ of M^n compatible with X and having continuity of $f|_C$ for each cell, $C \in \mathfrak{D}$.

Suppose $f: X \to M$ is such a definable function. We assume (because we already have) $(\mathrm{I})_1$. By this, we may assume X is a cell. If X is not already an open cell, then recall that we can take its image under the canonical projection away from zero coordinates. Since we do not refer here to the dimension of X, we assume that it is open or has been made so as described and then use our inductive hypothesis to conclude. So, we suppose $X \in \mathfrak{D}$ is an open cell on which f is continuous. Take

$$X' = \{ x \in X : f \text{ is continuous and definable at } x \}.$$

Clearly, X' is definable, and we are supposing that we know X' to be open in X. Using inductive assumption $(\mathrm{I})_n$, we get a cell decomposition, $\mathfrak{D}$ with $\mathbb{R}^n$ compatible with $X \setminus X'$ and with X'. If some $C \in \mathfrak{D}$ is an open cell contained in X, we get continuity of f on C by density; that is, $C \cap X' \neq \emptyset$, and so $C \subseteq X'$ and it follows that $f|_C$ is continuous. Supposing however that C was *not* an open cell, we apply the aforementioned projection construction, and the argument just presented holds (up to a change in dimension).

This would be all well and good to end off (a) with, were it not predicated on the yet unjustified density of X' in X – and so we now prove this. Suppose $B \subseteq X$ is an open box. We will show that there must exist a point in B at which f is continuous. In analogy to our proof of monotonicity, we know that if B' is an open box contained inside of B, then f takes on infinitely-many values on B' (following from $(\mathrm{I})_n$). This is the obvious case. Supposing otherwise, we proceed as follows:

Construct a sequence of open boxes, $(B_j)_{1\leq j \leq n}$ in B, and sequence $(I_j)_{1\leq j\leq n}$, of open intervals, each I_j having length less than $\frac{1}{j}$, with the closure, $\mathrm{cl}\,(B_{n+1}) \subseteq B$, and $f(B_n) \subseteq I_n$. Then, by compactness. we get that the intersection of all B_n is non-empty, and at some point in this intersection, f is continuous. This is of course just our claim – we now go on to *prove* this by construction.

To get I_1, simply consider $f(B) \subseteq \mathbb{R}$ – meaning

$$f(B) = \bigcup_{p\in\mathbb{N}} J_p \cup F$$

for F a finite set, and J_p a countable set of open intervals of length less than 1. Then, B is given by

$$B = \left(\bigcup_{p\in\mathbb{N}} f^{-1}(J_p) \cap B \right) \cup \left(\bigcup_{r\in F} f^{-1}(r) \cap B \right).$$

To each half of the middle cup, we can apply $(\mathrm{I})_n$ to determine the contents of each of the respective *big* cups to be a finite union of cells – and so B must be an *countable* union of cells, each of which is contained in one of these sets. Perhaps coming a bit out of left field, we apply the Baire Category Theorem to conclude that by the openness of B, so too must be one of these cells be open.

Notice that this is one reason we restrict ourselves to working over $\mathbb{R}$ – the Baire Category Theorem simply does not hold in any DLO model. So this argument could not be broadened beyond the reals (or compact spaces) as we are currently undertaking it.

This *cannot* be one of $f^{-1}(r) \cap B$, as it would then contain a box on which we took the value of r, an so this open cell must be in one of $f^{-1}(J_p) \cap B$ for some p. Taking J_1 to be that J_p, and B_1 to be an open box contained in $f^{-1}(J_1) \cap B$, with $\mathrm{cl}\,(B_1) \subseteq B$. As desired, we then have $f(B_1) \subset I_1$. Clearly the first step in an induction, we then (incompletely) note that, having $I_1, \ldots, I_n, B_1, \ldots, B_n$ constructed, we repeat exactly as above to finish the induction.

And with that, we can give ourselves a *light* patting on the back – for as much as we've done so far, this is just the end of the proof of (a). To get the 'bounced-back' half of the induction, we now go on to prove (b); that is, given $(I)_1, \ldots, (I)_n$, $(II)_1, \ldots, (II)_n$, we may derive $(\mathrm{I})_{n+1}$.

For reasons of breaking up this lengthy proof into its two constituent sections, please enjoy the following horizontal line:

Try not to have too much fun with that, now. We move on to proving (b); recall our assumptions that $(I)_1, \ldots, (I)_n$ and $(II)_1, \ldots, (II)_n$ hold. We want now to prove $(\mathrm{I})_{n+1}$. First, we start with a small proposition.

Proposition 2.2 *Suppose $\mathfrak{D}_1$, $\mathfrak{D}_2$ are cell decompositions of $\mathbb{R}^{n+1}$ with a common refinement – that is, another cell decomposition, $\mathfrak{D}$ of R^{n+1} compatible with all cells in each of $\mathfrak{D}_1$ and $\mathfrak{D}_2$. Terminology-wise, we say that $\mathfrak{D}$* refines *$\mathfrak{D}_1$ and $\mathfrak{D}_2$, or that is is an* refinement *of $\mathfrak{D}_1$ and $\mathfrak{D}_2$.*

Proof (Cell Refinement (Proposition 2.2)) For the purpose of transparency, we note that this proof was left out of the lecture and as an exercise for the interested (or obligated) viewer. The following takes inspiration from van den Dries [**?**]. We note that this could be made a bit cuter if we had the machinery of *dimension* that we will soon define, but in either case, this proof is relatively trivial. We have our $\mathfrak{D}_1$ and $\mathfrak{D}_2$ two decompositions of the trivially definable subset of, $\mathbb{R}^{n+1}$: $\mathbb{R}^{n+1}$ itself. We can then simply take a decomposition of the ambient space (which here is the *whole* space) containing our definable subset. We have seen previously that we can take this decomposition to partition each cell of $\mathfrak{D}_1 \cup \mathfrak{D}_2$ – and the 'restriction' of this decomposition to our definable set (again, just to ensure this is sufficiently belaboured, this is not actually a restriction since our definable set is $\mathbb{R}^{n+1}$), we are left with our everywhere (on cells in $\mathfrak{D}_1 \cup \mathfrak{D}_2$) compatible decomposition. □

Now, if some $A \subseteq \mathbb{R}$ is definable, we define its type, $\tau(A)$ as follows:

Let $a_1, \ldots, a_L$ strictly increasing be the points in the boundary of A. We let $\tau(A)$ then act as an indicator function on sequential intervals, (a_j, a_{j+1}), defined as the positive unit (1) when that interval sits inside A, and otherwise the negative unit (-1). We set $a_0 = -\infty$ and $a_{L+1} = \infty$ (which is starting to become sort of an out-of-bounds normalcy), and define

$$\tau_{2j+1} = 1$$

if $(a_j, a_{j+1}) \subseteq A$ (and of course -1 otherwise). Note, of course, that this would then mean that the given interval is contained in the complement of A. For even numbers, we have

$$\tau_{2j} = 1$$

if $a_j \in A$ and naturally, -1 if $a_j \notin A$. Then, we have $\tau(A) = (\tau_1, \ldots, \tau_{2 \cdot L+1}) \in \{0, 1\}^{2 \cdot L+1}$ a sequence of length $2 \cdot L + 1$ consisting of ± 1s.

To ground ourselves for a moment, consider

$$\tau((1, 2] \cup \{3\})) = (-1, -1, +1, +1, -1, +1, -1),$$

which can immediately convince ourselves is non-unique, as this is the same sequence induced by $\tau((9, 10] \cup \{5, 7\})$.

At this point, you may be a bit confused (if the recollection is even still there) as to why we made such a fuss early on about *uniform finiteness* – when we've yet to see it used. Well, for the anxious amongst you, satisfaction will come soon, as we now make use of that perhaps unjustified assumption.

By UF, we get the following:

Proposition 2.3 *If we have a definable $X \subseteq \mathbb{R}^{n+1}$, then the set of the types of fibres – that is*

$$\{ \tau(X_x) : x \in \mathbb{R}^n \}$$

– is finite. Further, for each given choice of type, the set of fibres giving rise to that type is definable.

Perhaps starting this sentence with 'of course' would be unfair, but it shouldn't be hard to see, intuit, or at least *guess* that the set of $x \in \mathbb{R}^n$ giving rise to any *particular* type is usually empty. As before, this proof was left as an exercise to the responsible party, and so please forgive any clear indications of amateurism – were they absent, there may be something a little suspect going on.

Proof (Proof of Proposition 2.3) This we will not belabour even slightly. We have assumed UF and so simply appeal to UF in the case of M^n, by which we get that each fibre is finite – and so must have finite types belonging to its elements. In the case of each type, it is given by some point or open interval in a fibre and so is defined by points and open intervals. Supposing there were infinitely-many *different* such types, we would have to be working in a non-finite dimension. Thus, definability falls out, almost as if by accident. $\square$

Now, with these two propositions, we are just about ready to put together our proof of (b). That is, we now prove $(\mathrm{I})_{n+1}$.

By Proposition 2.2 (on cell-refinement), we can assume $k = 1$ – which you'll recall is the number of sets we have from our original statement of the theorem. Then, by Proposition 2.3 and $(\mathrm{I})_n$, we get a cell decomposition, $\mathfrak{D}$ of $\mathbb{R}^n$ such that for each $C \in \mathfrak{D}$ there is an L and a $\tau \in \{\pm 1\}^{2 \cdot L+1}$ such that

$$\tau(X_x) = \tau$$

for all $x \in C$. Fixing such C, τ, and L, we get definable functions $f_1, \ldots, f_L$ with each $f_1 < \ldots < f_L$ and such that either

1. $(f_i, f_{i+1})_C \subseteq X$; or
2. $(f_i, f_{i+1})_C \cap X = \emptyset$,

with the normal condition of $f_0 = -\infty$, $f_{L+1} = \infty$. Notice also that the same holds for the graphs of f_j – that they are either contained in or disjoint from X, excluding of course the cases at infinities (which we allowed above). Now, with our small army of definable functions, all defined on this cell $C \in \mathbb{R}^n$, we can use $(\mathrm{II})_n$ (which we proved in our induction in part (a)) to conclude that we may partition C into finitely-many cells, such that f is continuous on each cell. With the end very nearly in sight, we apply $(\mathrm{I})_n$ to get a cell decomposition of $\mathbb{R}^n$ (and in fact of all cells from the above) compatible with all the resulting cells. Finally, taking graphs over those cells will give us the desired cell decomposition of $\mathbb{R}^{n+1}$. And with that, we have 'bounced back' such that we may repeat this induction *ad infinitum*, and the Cell Decomposition Theorem is proved (in our special case) for all n. Please note the now properly-placed well-deserved QED-symbol. Revel a little, if you must. □

It is now at *this* point that the reader not only *may* but is encouraged to give themselves their well-deserved, no-bars-held, hearty pat on the back for the fortitude it took to get through that. Perhaps also giving the above a quick re-read wouldn't be such a bad idea, as there are some bits and subtleties that this author needed a few passes to feel entirely comfortable with. For a proof that doesn't make the assumptions we did here, the reader is directed to [**?**], but by no means necessarily encouraged towards it —- just made aware of its existence. With that done, we are through one of the more laborious parts of this first part of the course. For the masochists in the audience, we note that there is more length and labour to come of this variety. Still, for the normal amongst us, it is with a relaxation that we should move on to discuss *definable connectedness*, followed by *dimensionality* – and the problem that mathematicians have for coming up with distinct words for distinct concepts. But first – which is a phrase we perhaps begin sentences with all too often – we have a bit of miscellany to address.

Chapter 3
In Absence of Aproposia: A Brief Foray From the Curriculum

Abstract Here, we mention a point or two that would otherwise go overlooked. Necessity is entirely eschewed, and this chapter can be safely skipped without any loss in understanding of the course as it was intended to be presented. One is perhaps encouraged to leave, in fact. This short chapter exists only for the interested, dedicated, obliged, or otherwise neurotic reader. In short, this encapsulates that which has no other place thus far, nor is deserving of its own section (should we not consider a chapter a section in and of itself).

3.1 Let's Just Get it Out of the Way

One of the first thoughts the critical or unassuming reader – which is likely all of you – had when reading this chapter title was whether *that* word was really even a word *at all*. If you are at all a curious person, searching the internet or your favourite etymological website or reference text for the word 'aproposia,' you would have failed in your task – unless the task you set out was to assure yourself that it is, in fact, *not* a common or even extant word. Were 'apropos' of Latin root, then perhaps this linguistic abomination may make more sense, but considering the French origin of 'apropos,' no such logic applies. Nonetheless, there is little the reader can do about this choice of wording (save for one particular professor). And since we felt its hopefully clear meaning and appealing sound appropriate to the nature of the topic, you should then be glad at all, dear reader, that Latin is no longer the language of science you would be expected to learn to have what would be considered 'valid' opinions on its workings. If anything, feel free to use it as word to befuddle and confuse your dear friends and colleagues. While we will let you get away with calling everything and anything 'normal', we insist you take 'aproposia' as valid and unilaterally call that a fair trade.

3.2 Onto the Miscellany

In which we again abuse language, in the sense that there is only one fact of miscellaneous variety. Sometimes, certain benign ridiculousness must be allowed to amuse your readers or even just oneself. What we do, after all, is exciting not in its protracted execution, but in the few moments where it all comes together. It's how we prevent the onset of early insanity when getting into the thick of these sorts of ideas.

Remark 3.1 Recall that an o-minimal structure requires all definable sets *with parameters* to be given by finite unions of points and open intervals (as defined by the model). If we only assume this for sets definable *without* parameters, the resulting theory is *legitimately* and provably weaker than what we get with o-minimality. This was in passing mentioned to be potentially true earlier on, but in one of the question and answer sessions held for this course, it was pointed out that it is *in fact* true by a gentleman with the given name Chris. In a moment, we will be referring to him by his family name, Miller. This awkward wording will be clear in just a moment.

An easy (in the sense of being counter-exemplary) way to show this is due to Dolich, Miller, and our old friend Steinhorn [**?**]. This can be expressed (though perhaps not proven, as the length of their paper implies) quite compactly by constructing the model

$$\mathcal{M} = (\mathbb{R}, <, V)$$

for V the Vitali set (defined by the Vitali *relation*):

$$V = \{\, (x, y) \in \mathbb{R}^2 \ :\ x - y \in \mathbb{Q} \,\}.$$

The only $\emptyset$-definable subsets of this are $\emptyset$ itself and $\mathbb{R}$ – and so this fits the definition of being $\emptyset$ o-minimal, but given any defined parameter, we end up with the rationals definable; clearly, this is *not* o-minimal. If you recall the short mention made earlier, it may interest the reader to note that this is *weakly* o-minimal.

Exercise 3.1 As an exercise of *this* author to the reader, attempt to prove that the above expansion admits QE. *Hint: You would be well-served to first read Chapter 5 on dimensionality and definable closures.*

Putting an end to this brief foray into intrigue with a splash of ridiculousness, we now move back on to a consequential idea once one defines cell decompositions: connectedness.

Chapter 4
Connectedness, and What it Entails

Abstract It is likely that one unacquainted too thoroughly with this topic of study has been imagining in their minds open intervals as they might imagine them in $\mathbb{R}$. Unfortunately for such a reader, they are about to be disabused of that rather idyllic notion – and we will speak to when we legitimately *may* presume that things are (what we will come to call) *definably connected*. At this point, we should have an acronym to express that the importance of this will not be immediately apparent but will soon come to be and be worked into our standard model of understanding these objects we find ourselves working with. The creation of such an acronym is left as an exercise to the reader.

4.1 But *Why* Aren't Things (Always) Connected?

The easy answer is: sometimes we just don't work over connected domains. Take, for example, $(\mathbb{Q}, <)$ – where we may write $\mathbb{Q} = (-\infty, \sqrt{2}) \cup (\sqrt{2}, \infty)$.

> Recall, of course, that neither interval is open in $\mathbb{Q}$ by non-membership of the irrational endpoints in the domain.

In particular, we find that a convenient definition of a definably connected set, $X \in M^n$ is

Definition 4.1 (Definable Connectedness) Some $X \in M^n$ is definably connected if X is *not* the union of two disjoint non-empty definable open subsets of X.

Example 4.1 Open intervals, we claim, are definably connected. So too are cells – although this one we reason a bit about. By Cell Decomposition, we know that definable sets have finitely-many definably connected components which are maximal definably connected subsets. So by uniformity, it seems that we can always take a cell to not be the union of two disjoint non-empty definable open subsets of itself.

We formalize this idea with the following proposition.

Proposition 4.1 *Support some* $X \in M^{m+n}$ *is definable. Then, there exists some* $N \in \mathbb{N}$ *such that if* $x \in M^m$*. then* X_x *has* at most N *definably connected components.*

Corollary 4.1 *Given* $\mathcal{N}$*, a structure elementarily equivalent to* $\mathcal{M}$ *(satisfies exactly the same first order sentences in our language) for* M *o-minimal, then* N *is also o-minimal. In short,*

$$\mathcal{M} \equiv \mathcal{N} \ \wedge \mathcal{M} \text{ o} - \text{minimal} \implies \mathcal{N} \text{ o} - \text{minimal}.$$

For the interested and more informed reader than expected necessarily, it is noted that the property of *minimality* (not o-minimality) is *not* preserved under elementary equivalence as o-minimality is. This is one of the motivations for the idea of *strong*-minimality, but for all intents and purposes here, we pretend there is no notion of a *strong*-o-minimality.

This corollary will come to be quite important later, so if nothing else from here, keep that fact in the back of the mind as we go forward.

4.2 Definable Choice & Curve Selection

In the interest of time, space, audience, and simply relevance, in most cases, we will not be providing examples as we have before (as in the case of expansion by the Vitali relation) and instead just assume we take $\mathcal{M}$ an o-minimal expansion of an *ordered* field, $(M, <, +, \cdot, 0, 1)$. Not addressed here, but as a good exercise to the interested reader, attempt to show that this must necessarily be a real closed field. We will think in abstractness only, and the reader who even tries to think of an example should be rather a bit ashamed of what they've done.

Without any faff, we get right into the point of this section.

Proposition 4.2 (Definable Choice)

1. *Given a definable family,* $X \subseteq M^{n+m}$ *with* π *the projection map onto the first* n *coordinates, then there is a definable map,* $f : \pi X \to M^n$ *with* graph $(f) \subseteq X$.
2. *Given* E *a definable equivalence relation on a definable set,* $X \subseteq M^n$*, then* E *has a set of representatives.*

Proof (Proofs of the above (Proposition 4.2) First, we go on to show that if some $X \subset M^n$ is definable and non-empty, then we may definably pick some element, $e(X) \in X$. As ever, we induct on n.

Suppose $n = 1$, then either

1. X has a least element, and so we let $e(X)$ be that; and otherwise

2. X has a left-most interval, (with respect to our order), and splitting by cases, we can take
   - $e(X) = 0$ if $(a, b) = (-\infty, \infty)$;
   - $e(X) = b - 1$ if $a = -\infty$, $b \in M$;
   - $e(X) = a + 1$ if $a \in M$, $b = \infty$; and
   - $e(X) = \dfrac{a+b}{2}$ if both are finite.

Note, of course, that our arithmetic is well-defined here due to the field expansion we are working with. We now induct a bit differently to prove each of cases 1 and 2;

- In the first case (definable without parameters), we put $f(x) = e(X_n)$ for $x \in \pi X$ (since our fibre is non-empty; and then for
- we take $\{\, e(A) \;:\; A \text{ is an equivalence class of } E \,\}$ a definable set or representation, as desired. □

We may go on to refer to *definable choice* as DC, stated in the acronym section upfront. From this, we can then go on to prove a neat and useful little result called *curve selection*.

Proposition 4.3 (Curve Selection) *Suppose $X \subset M^n$ is definable and $a \in \mathrm{fr}\,(X)$ (the frontier of X, defined by to the closure of X less X). Then, there is a continuous definable injective $\gamma\colon (0, \varepsilon) \to X$ for some $\varepsilon > 0$ with*

$$\lim_{t \to 0^+} \gamma(t) = a.$$

Predictably, this is going to use the result we just proved on definable choice, so we just jump right in.

Proof (of Proposition 4.3) Let $|x| = \max\{|x_1|, |x_2|, \ldots, |x_n|\}$. Since $a \in \mathrm{fr}\,(X)$, the set

$$\{\, |a - x| \;:\; x \in X \,\}$$

is a definable set with arbitrarily small positive elements, and contains some interval $(0, \varepsilon)$.

If t is in this $(0, \varepsilon)$, then the set

$$\{\, x \in X \;:\; |a - x| = t \,\}$$

is non-empty. Thus, by DC, we get a definable $\gamma\colon (0, \varepsilon) \to X$ such that $|a - \gamma(t)| = t$ for some t in our interval. Clearly then, t is injective with limit a as t approaches 0 (on the right). As a throwback, we can apply the Monotonicity Theorem and reduce ε to reach the assumption that γ is continuous. □

We've perhaps teased at it now for a bit, and the particularly knowledgable or prescient reader will have seen this coming, but it is at this point we move on to a spot of dimension theory. The discussion here is interesting in that we approach it both from the angle of definability (as expected) and algebraicity and see what comes to pass.

Chapter 5
Definable Dimension: The First Go-Round

Abstract As so often one finds in mathematics, dimension is one of those ideas that it almost seems each individual mathematician has a notion of how it should be defined – and more truthfully, it has hugely varying meanings across the vast spectra of mathematical disciplines. And perhaps even more, unfortunately, it so often seems the case that the average mathematician was raised on a dictionary of no more than 30 or so words – which they go on to use and reuse and smash together into all-new words, for the most part equally inequivalently to how their office-mate might be doing the same thing in their preferred area of abstractness. To say the same in so many fewer words: we often find that the same words mean different things to different mathematicians – Here, we are going to *dimension* from the logician's point of view (at least in the context of o-minimality as we have been thus far). Little will be terribly new here, although we will encounter some interesting techniques, some longer proofs, and a fair deal of dallying about with unions, intersections, and differences of sets; all this to say, a set theorist's dream. Those of you who feel we are being a bit sly in our phrasing are correct, and you will come to see why, if you aren't already keenly aware of the big surprise we are half-heartedly keeping from you, in Chapter 6 – when a very exciting realization will come to pass. However, for now we stick to the exciting realizations that come with continuing on along the logical path we have set for ourselves thus far.

We continue in our assumption from the previous section – that is, that we take $\mathcal{M}$ an o-minimal expansion of an *ordered* field, $(M, <, +, \cdot, 0, 1)$, and prove things about this construction in generality. Strictly speaking, this is not necessary here, even speaking less strictly than we have about this assumption earlier on. It is simply a matter of convenience and lack of loss of generality.

5.1 So How Does One Define Dimension?

Tempting though it is to answer "depends on who you ask" and move on with our day, the question does bear significant thought. We start with the following definition.

Definition 5.1 (Dimension) Suppose $X \subset M^n$ is definable and non-empty. Then we set

$$\dim(X) = \max\left\{ \sum_{j \in j_1, \ldots, j_n} j : X \text{ contains a } (j_1, \ldots, j_n)\text{-cell} \right\}$$

to be the dimension of our subset of M^n, with the dimension of $\emptyset$ being $-\infty$. Intuitively, this hopefully makes sense – dimension being given by the largest object (of our current interest) sitting inside our space.

At first blush, this seems to be a reasonable definition of dimension given the manner in which we've been defining the rest of our toolkit – and with further blushing, we will come to find it even more reasonable than it may even have initially appeared. For those finding this *unreasonable*, we would encourage a re-reading of some of the earlier definitions, specifically on cells and their decompositions, or, barring that, just setting fire to this manuscript and going on about your day. Either way, we say a little something about what subsets with non-empty interior may tell us.

Lemma 5.1 *If $X \subset M^n$ has interior,* $\mathrm{int}(X)$*, non-empty, then there is is a definable injective map, $f: X \to M^n$ with image of X under f containing an open cell.*

This lemma will go one to be quite useful in a moment when we wish to say some useful things about, for example (and in particular), the *dimension invariant of definable bijections*.

Proof (of Lemma 5.1) We promise you now, barring life-threatening or otherwise dire circumstances, that this will be the last preface to a proof of this sort – and trust you, dear reader, to recognize when we are setting up a proof by inductions on n. But for now, we proceed by induction on n.

When $n = 1$, then $X \subseteq M^n$ is a non-empty subset of M, and so is infinite. Since f is injective, $f(X)$ so too is it infinite (relying as well on definability to conclude this), and so $f(X)$ contains an open interval – which we know to be an open-cell.

Suppose now that $n > 1$, and our inductive hypothesis holds for all k between 1 and $n - 1$. By CD, we will simply assume that X is itself already an open cell and that f is continuous. So, by CD we can take some cell decomposition $\mathfrak{D}$ of $f(X)$. Should one of the $C \in \mathfrak{D}$ be open, we are done – and so we assume none are (to address all possible cases). Then, since X is the union of the preimages of the cells in $\mathfrak{D}$, some $f^{-1}(C_j)$ contains an open cell, $C \subseteq M^n$. Say then that C is contained in the preimage of, without loss of generality, $f^{-1}(C_1)$. Then, restricting f to C, we have

$$f|_C : C \to C_1$$

continuous, definable, and injective. Recall the homeomorphic nature of the projection away from 0-coordinates. Similarly, in extensions of fiends (this is not about to be proven), open cells are homeomorphic to the ambient space, and so since we expand a field, any $A \in M^n$ are definably homeomorphic to M^m. So, putting this together, we have C homeomorphism to M^n and C_1 homeomorphism to M^ℓ for some $\ell < n$ – and so C is definable homeomorphic to M^n, C_1 for M^ℓ, and thus we have a continuous definable injective

$$g : M^n \to M^\ell$$
$$h : M^\ell \to M^\ell$$

where we define h by

$$h : y \mapsto g(0, y)$$

with, to be clear, 0 coming from $M^{n-\ell}$ (in case things weren't seeming all above-board here). By our inductive hypothesis, we can finally conclude that the image $h(N^\ell)$ *has* interior. Letting some $b \in \operatorname{int}\left(h(M^\ell)\right)$ and $a \in M^\ell$ with $h(a) = b$, we can say by continuity of g that for $x \in M^{n-\ell} \setminus \{0\}$ sufficiently small, that $g(x, a)$ will sit in the image of h, and so be achievable by some argument to h. We will call this $a' \in M^n$ satisfying $f(a') = g(x, a) = g(0, a')$ – clearly contradicting injectivity.

As such, we have that $f(X)$ contains an open cell, and we can now pronounce our proposition proved. □

5.2 Dimension Under Functions

Now, we will go on to show something very sensible indeed – and something that, should it *not* hold, cause great concern. That is, invariance of domain under definable bijection.

Proposition 5.1 (A Bit on Definable Bijections)

1. *If $X \subseteq Y \subseteq M^n$ all definable, then the dim*

$$\dim(X) \leq \dim(Y) \leq n.$$

2. *If $X \subseteq M^m$, $Y \subseteq M^n$ are both definable and $f : X \to Y$ is a definable bijection between them*

$$\dim(X) = \dim(Y).$$

3. *If $X, Y \subseteq M^n$ are both definable, then*

$$\dim(X \cup Y) = \max\{\dim(X), \dim(Y)\}$$

Hopefully these statements don't come across too contentious, nor necessarily difficult to prove, but nonetheless, we state our own:

Proof

1. This is almost as free as it gets; by containment of X in Y, we already have that no cell in X can have dimension larger than one in Y. Similarly, we have bounding above of the dimension of any cells by the ambient space, which we know to have dimension n. That felt like perhaps a lot to say for what is quite clear.
2. Assume we have proven 3, which we do first. To be clear,

```
GOTO 3
COMPREHEND_3()
GOTO 2
```

And you're back! So now, let's prove 2. Suppose $d = \dim(X)$ and $e = \dim(Y)$. By symmetry, it clearly suffices to show that $d \leq e$ (or vice versa) as one could then simply switch X for Y.
Since X has dimension d, it must contain an $(i_1, \ldots, i_m)$-cell with indices $i_1, \ldots, i_m$ summing to d. Let a cell $C \subseteq X$ an $(i_1, \ldots, i_m)$-cell of the same dimension, d. We have that f sends X to Y, and so composing with a definable bijective $C \to M^d$, we get a definable injection

$$g\colon M^d \to Y.$$

Cell decomposing Y into $C_1, \ldots, C_k$ (each a cell, together having union Y), we have then that

$$M^d = g^{-1}(C_1) \cup \cdots \cup g^{-1}(C_k).$$

By openness of M^d, as before we must have that some $g^{-1}(C_j)$ (without loss of generality, suppose C_1) contains some open cell, $D \subseteq M^d$, such that $D \subseteq g^{-1}(C_1)$. Since C_1 sits in Y, we generically say it is an $(j_1, \ldots, j_n)$-cell. For sake of contradiction, suppose these coordinates have sum less than d, and denote this quantity e'. Then, we have

$$h'\colon D \xrightarrow{g} C_1 \xrightarrow{\sim} M^{e'}$$

a definable injective function. We can imagine $M^{e'}$ as 'sitting inside' D, since we know by our assumption that its dimension is necessarily lesser. So now, considering 0 in $M^{d-e'}$, we have the map

$$h''\colon D \to M^{e'} \times \{0\} \subseteq M^d$$

also both definable and injective– but then since we have an open cell, D, we contradict Lemma 5.1. Thus, we have that $d \leq e$, and so must also have $d = e$ by symmetry of the argument, and we are done.

3. Let $d = \dim(X \cup Y)$. and let C a cell in the union that witnesses the dimension, d (which to be clear, and going forward, will refer to a $(c_1, \ldots, c_n)$-cell with $\sum_{j=1}^{n} c_j = \mathrm{d}$). Now, let $\varphi \colon C \to M^n$ a definable bijection. Then, we have (from before, and not non-obviously)
$$M^d = \varphi(C \cap X) \cup \varphi(C \cap Y)$$
open (by M^d open), and so one of the constituent parts of the union must be also. Hence, one of the two contains an open cell, D (without loss of generality, suppose $\varphi(C \cap X)$) such that
$$\varphi^{-1}(D) \subseteq X$$
and so
$$\dim\left(\varphi^{-1}(D)\right) = \dim(D) = d.$$
So, then $\dim(X) \geq d \geq \dim(X)$, and the rest follows. □

Now with a bit under our belt about this notion of dimension that we would reasonably expect, we go on to prove a few more interesting facts that will come to use later on. In particular, we first show something about the additivity of the dimension of fibres.

5.3 Concerning Fibres

Proposition 5.2 *Suppose $X \subseteq M^m \times M^n$ is definable, and some positive natural number $d \leq n$. Denote*
$$X(d) = \{\, x \in M^n : \ \dim(X_x) = d \,\},$$
recalling, of course, that X_x denotes the fibre of X over x (for any who have forgotten). Then, $X(d)$ is definable and
$$\dim\left(\bigcup_{x \in X(d)} (\{x\} \times X_x)\right) = \dim(X(d)) + d$$

Proof (of Proposition 5.2) Suppose $\mathfrak{D}$ is a cell decomposition of of $M^m \times M^n$ compatible with X. If some $C \in \mathfrak{D}$ is an $(i_1, \ldots, i_{m+n})$-cell, then its projection onto its first m components, πC, is an $(i_1, \ldots, i_m)$-cell and each fibre, C_x is an $(i_{m+1}, \ldots, i_{m+n})$-cell for $x \in \pi C$. Clearly, we have that
$$\dim(C) = \dim(\pi C) + \dim(C_x) \tag{5.1}$$

for each $x \in \pi C$.

Now, setting $C' \in \pi\mathfrak{D}$ (the cells given by this projection on all cells in $\mathfrak{D}$) and letting $C^{(1)}, \ldots, C^{(k)} \subset \mathfrak{D}$ the cells contained in X whose projection is C'. Then for $x \in C'$, the fibre

$$X_x = C_x^{(1)} \cup \ldots \cup C_x^{(k)}$$

and so

$$\dim(X_x) = \max\left\{ \dim(C_x)^{(1)}, \ldots, \dim(C_x)^{(k)} \right\}.$$

Combining this with 5.1, we have that

$$\dim(X_x) = \max\left\{ \dim\left(C^{(1)}\right), \ldots, \dim\left(C^{(k)}\right) \right\} - \dim(C'),$$

which is completely independent of X – that is, *where* we fibre makes no difference in the dimension. Let d be this dimension, given by the maximum. Then $\dim(C_x) = d$ for all $x \in C'$, so $C' \subseteq X(d)$, and thus $X(d)$ is a finite union of cells and therefore, as desired, definable. So we have the definability part of the proposition – we now need to prove the given equality.

We have that our d is defined as above by

$$d = \max\left\{ \dim\left(C^{(1)}\right), \ldots, \dim\left(C^{(k)}\right) \right\} - \dim(C')$$
$$d = \dim\left(\bigcup_{i=1}^{k} C^{(i)} \right) - \dim(C')$$

Recall, this union is the part of X laying above C', and so we can further improve to get

$$d = \dim\left(\bigcup_{x=C'} \{x\} \times C_x^{(i)} \right) - \dim(C')$$

and so

$$\dim\left(\bigcup_{x=C'} \{x\} \times C_x^{(i)} \right) = d + \dim(C')$$

which, taking the union over all $C' \in \pi\mathfrak{D}$ with $C' \subseteq X(d)$, gives us the result, knowing that the maximal dimension over all C' is the dimension of $X(d)$. With that, we have gotten exactly what we want. $\square$

As a corollary, we get something of a more general form (in several parts) of the previous proposition, which goes as follows:

Corollary 5.1

1. *If $X \in M^{m+n}$ is definable, then X has dimension given*

$$\dim(X) = \max_{0\le d\le n} \{\dim(X(d)) + d\}$$
$$\ge \dim(\pi X)$$

   for π as in the proposition.
2. *Suppose $X \subseteq M^n$, $f : X \longrightarrow M^m$ definable. Denote*

$$X_f(d) = \left\{ x \in M^m : \dim\left(f^{-1}(x)\right) = d \right\}.$$

   Then $X_f(d)$ is definable and further,

$$\dim\left(f^{-1}(X_f(d))\right) = \dim\left(X_f(d)\right) + d$$

   and $\dim(X) \ge \dim(f(X))$.
3. *Consider now a special case of taking products, where $X \subseteq M^m$, $Y \subseteq M^n$ are definable. Then, perhaps predictably,*

$$\dim(X \times Y) = \dim(X) + \dim(Y).$$

Easy as they are, these proofs were left to the viewer, so please enjoy this unassessed attempt at just that. Note that as we know part 3 to simply be a special case following from part 2 we only prove the first two parts of the corollary.

Proof (of Corollary 5.1)

1. That the dimension of X is not lesser than that of its projection onto a subset of its coordinates is obvious, and really doesn't bear much consideration. Further, immediately applying Proposition 5.2, simply noting that taking the union over fibres of subsets of X of each possible dimension, $0 \le d \le n$, gives us this result exactly, we are done.
2. In essence, here what we are doing is again applying Proposition 5.2, but now to the graph of our X, $f(X)$. Again, by a similar argument as in the previous part of this proof, and the direct application of the formula as given in the proposition, the result we desire falls out.
3. Follows from the previous. □

Exercise 5.1 Show that: Supposing $X, Y \subseteq M^{1+n}$ with Y non-empty, and that for each $x \in M$, either $X_x = \emptyset$ or $\dim(X_x) < \dim(Y_x)$, then $\dim(X) < \dim(Y)$.

We now use this result to show the following, perhaps quite intuitive lemma.

Lemma 5.2 *Suppose $X \subseteq M^{n+1}$ is definable. Then*

$$\{x \in M : \mathrm{cl}(X_x) \ne \mathrm{cl}(X)_x\}$$

*is finite. That is, there are only finitely-many points at which the closure of the fibre is **not** the fibre of the closure.*

Proof (of Lemma 5.2) Note that we always have that the closure of the fibre is always contained in the fibre of the closure – that is, $\mathrm{cl}\,(X_x) \subseteq \mathrm{cl}\,(X)_x$. Suppose, for purposes of contradiction, that the set defined above *is* infinite. As such (with definability), it contains an open interval which we shall call I. By definition of our set, for each $x \in I$, there is an open box, $B \subseteq M^n$, witnessing the difference – that is, that

$$B \cap X_x = \emptyset$$

and simultaneously,

$$B \cap \mathrm{cl}\,(X)_x \neq \emptyset$$

We know the family of open boxes in M^n to be definable, and so by definable choice (do recall that we said it would become important), we can get each such box as a *definable function* (whose notation we overload by referring to it as B) of the prescribed $x \in I$. By monotonoicity, we may assume that B is continuous, and so taking

$$U = \{\ (x, y) \in I \times M^n : \ y \in B(x)\ \},$$

we have U open in $I \times M^n$, and $U \cap X = \emptyset$. However, we must *also* have then that $U \cap \mathrm{cl}\,(X) \neq \emptyset$ – which is a clear contradiction. Thus, our original assumption of the infinitude of our set was incorrect, and the result is proven. □

Now, using this lemma and a bit from before, we can prove the following theorem:

Theorem 5.1 *Suppose $X \subseteq M^n$ is non-empty, and definable. Then,* $\dim(\mathrm{fr}\,(X)) < \dim(X)$.

This one is quite intuitive, and its proof not terribly involved – but the result itself is quite useful, and sees much use in inductive arguments about these sorts of ideas. We will not be using this result quite often, but later modules in this lecture series will make much more judicious usage.

Proof (of Theorem 5.1) We start with the base case of $X \subseteq M^1$. We trust that this follows without comment.

Now consider $X \subseteq M^{n+1}$, and suppose that the result holds for M^k for any $1 \leq k \leq n$ (in fact we only need weak induction here, but we assume this regardless for reasons of not mattering). For each coordinate, $i_1, \ \ldots, \ i_{n+1}$, consider the set

$$\mathrm{cl}\,(X)_i = \left\{ x \in M^{n+1} : x \in \mathrm{cl}\left(X \cap \pi_i^{-1}(\pi_i(x))\right) \right\}$$

for π_i predictably the projection map onto the i-th coordinate. Note, of course, that $\pi_i(X) \subseteq \pi(X)$.

Without any loss of generality, we make our arguments with respect to x_1 here – although there is no reason they should not similarly hold for any other coordinate. Suppose we have some $x \in \mathrm{cl}(X) \setminus \mathrm{cl}(X)_1$ (a set defined similarly to the frontier of *all* of X). Then we can write this x as $x = (x_1, x')$ where $x' \in \mathrm{cl}(X)_{x_1}$ and further $x' \notin \mathrm{cl}(X_x)$. By the above lemma (Lemma 5.2) there are only finitely-many possible such x_1 – and so the set difference above lies in a finite union of hyperplanes, each of which projecting onto a single point under π_1. So, there exist points $a_{1,1}, \ldots, a_{1,k_1} \in M$ such that

$$\mathrm{cl}(X) \setminus \mathrm{cl}(X)_1 \subset \bigcup_{j=1}^{k_1} \pi_1^{-1}(a_{1,j}). \tag{5.2}$$

As before, it the choice of coordinate 1 was arbitrary, and this same argument holds for *all* coordinates. Thus, permuting coordinates, we get that for each remaining coordinate, we get some $\left(a_{i,1}, \ldots, a_{i,k_i}\right)$, for i varying over the remaining coordinates, such that 5.2 holds for the respective i. Orthogonality of these planes is clear by the differing i's, and so the closure minus the union of these closures sits within a set that can be characterized as

$$\mathrm{cl}(X) \setminus \bigcup_{i=1}^{n+1} \mathrm{cl}(X)_i \subseteq \left\{ (a_1, j_1, \ldots, a_{n+1}, j_{n+1}) \ : \ \begin{array}{l} j_1 = 1, \ldots, k_1, \\ j_{n+1} = 1, \ldots, k_{n+1} \end{array} \right\}.$$

The thing to take notice of here is that this set on the right is *finite*, and so our difference in the closure of X and union of closures of the projections onto each coordinate is contained in a finite set.

Recalling what we proved about the dimension of the union of a set of sets (Proposition 5.2), we get that

$$\begin{aligned} \dim(\mathrm{fr}(X)) &= \dim(\mathrm{cl}(X) \setminus X) \\ &\leq \max\{\dim(\mathrm{cl}(X)_i \setminus X), \ 0\}. \end{aligned}$$

The problem is now reduced to showing simply that $\dim(\mathrm{cl}(X)_i \setminus X)$ has dimension less than X for each i. This doesn't turn out to be too difficult, and we jump right in with little fuss.

Without loss of generality (a phrase we are beginning to believe may have a nice initialism), take $i = 1$, and some point $a \in M$. Then

$$\dim\left(\mathrm{cl}\left(X \cap \pi_1^{-1}(a)\right) \setminus \left(X \cap \pi_1^{-1}(a)\right)\right) < \dim\left(\pi_1^{-1}(a)\right)$$

so long as this is non-empty, by the inductive hypothesis. Otherwise, if it *is empty*, then we consider the whole faff above, by which we of course mean

$$\mathrm{cl}\left(X \cap \pi_1^{-1}(a)\right) \setminus \left(X \cap \pi_1^{-1}(a)\right) = (\mathrm{cl}(X)_1 \setminus X) \cap \pi_1^{-1}(a),$$

and so (and we don't claim obviousness necessarily) for each $a \in M$, with $X \cap \pi_1^{-1}(a) \neq \emptyset$, we have that

$$\dim\left(\mathrm{cl}\left(X\right)_1 \setminus X\right) \cap \pi_1^{-1}(a) < \dim\left(X \cap \pi_1^{-1}\right).$$

Notice that the left half of this inequality can be identified by the fibre over a, and the right as the fibre over x, so that

$$\dim\left(\mathrm{cl}\left(X\right)_1 \setminus X\right)_a < \dim\left(X_a\right).$$

So, by Exercise 5.1, we get that

$$\dim\left(\mathrm{cl}\left(X\right)_1 \setminus X\right) < \dim\left(X\right).$$

Pulling the usual and, to be honest at this point, tired and tiring rabbit out of our (frankly garish) hat of permuting coordinates, we see (in a twist we are certain no one saw coming) that this argument works for *all* coordinates, not just $i = 1$ – and so the result holds for i from 2 up to $n + 1$. We now have that one of the following holds:

1. $\dim\left(\mathrm{cl}\left(X\right) \setminus X\right) < \dim\left(X\right)$
2. $\dim\left(X\right) = 0$

The former is exactly what we wanted, and should the dimension of X be 0, then X is finite, and so closed – and we are finished. □

And now, we are done with dimensionality! Sort of. In a way. But in another way, not at all – and for those who gave the introduction more than a cursory skim or have seen the Table of Contents, you will know that this has only been the first half in our investigation into the idea of dimension. Without belabouring beyond reason, as we will be sure to do so in the next section as well, prepare now to investigate a *seemingly* less motivated (at least from the previous material) definition and discussion on dimension – but one that will become vital to our proofs to come.

Chapter 6
Model-Theoretic Dimension: Once More with Feeling

Abstract What we just worked through was broadly a very natural extension of the ideas we'd been toying with so far, and the idea of defining dimensionality by cells is a very sensible one indeed. However, the antsy and number theoretically-inclined amongst you may be getting a little weary of just how little of that sort has cropped up so far. Fear not then, our brethren and sistren in the fine and noble study of numbers and the little things they do – as we now examine the idea of dimensionality from an *algebraic* perspective, and begin to work in the ideas that will come to unite the logical world of o-minimality with that of our own (not to expose our biases too blatantly). We will come to the wonderful conclusion that, in fact, these two notions are one and the same, and we can start to drop the veil of pretense we have been vaguely shrouding this whole affair in – although you would be in the know had you read the introduction in its entirety, or even just the Table of Contents. Importantly, this idea will lead us to one of the important tools that allows the Pila-Wilkie theorem to work at all; that being this connection between algebra, number theory and o-minimality that together can put together something bigger and more wonderful than each constituent on its own. So, dear reader, please enjoy a rehashing of much of what you saw just a chapter ago, but now through an entirely different lens – one that will perhaps expose you to some ideas and inner workings you'd not thought of or seen previously.

As in Chapters 4 and now 5, we will continue in our assumption that we have $\mathcal{M}$ an o-minimal expansion of an *ordered* field, $(M, <, +, \cdot, 0, 1)$, and prove things about this construction in generality.

Although at no point will we be diving deeply into the depths of model theory and all the fun/horror that this may entail (depending on who you ask), this is going to be the part of the course that relies most heavily on *some* knowledge of model theory, and there's little to do to get around that besides having little heuristics in one's head that make things *kinda make sense* enough to not worry about it too much. All this said, the reader who's not even heard the term 'model theory' in their life should not be intimidated, as we will be doing our best in holding you hand (or not as you see fit) through the dicier parts of this section, and so much as we hate to keep using this phraseology, belabouring that which may be very obvious to one of a different background. While this may not have been mentioned before, as a series of papers (or more likely document on a computer of some variety) it is not only easy, but vastly personally rewarding to jump over sections you feel eminently comfortable about with a reckless abandon. Well, this is true for all but I hope the one professor whom I hope does get to at least the vast majority of words written here. But for anyone else – skip judiciously, and skip as if you've much better things to do (which you likely do).

6.1 So How Does One Define Dimension?

Tempting though it is to answer "depends on who you ask" and move on with our day, the question does bear significant thought. We start with the following definition (and promise we are not trying to induce déjà vu):

Definition 6.1 (Algebraic Dimension) Unlike in the previous chapter, we will *not* be starting with a definition of algebraic dimension, and working from there – you simply are not yet prepared, given the content discussed so far. We now back up a little bit, and will return to this topic when more sufficiently prepared. Suffice it to say for now that our definition of dimension will be reliant on the idea of *algebraic* (and then, equivalently, *definable*) closures (in the model-theoretic sense).

6.2 Some Recollections or Introductions

Here we recall for some, and for others introduce for the first time, some of the rudimentary ideas of model theory, and in particular of how one goes about defining a model-theoretic or algebraic closure of a set – and then go on in a similar way to before to use this to define dimensionality.

6.2.1 On Closures of Various Types

Definition 6.2 (Algebraic Closure) We say some $A \subseteq M$, then the **model-theoretic algebraic closure of** A, which we will denote by $\operatorname{acl}(A)$, is the union of all finite A-definable subsets (in $\mathcal{M}$) of M.

Definition 6.3 (Definable Closure) Again, given $A \subseteq M$, then the **definable closure of** A, denoted $\operatorname{dcl}(A)$, is given by the union of all A-definable singletons.

Note perhaps a more careful attention to parameters being paid here than in previous sections – precision here with respect to the topic is, not to say of greater importance than elsewhere, but can much more easily lead to confusion if not properly, clearly, and rigorously attended to.

In our setting, in particular in reference to the order-structure our models possess, these seemingly disparate definitions of closure actually coincide – that is,

$$\operatorname{acl}(A) = \operatorname{dcl}(A) .$$

For convenience then, we will in general simply work with the definable closure.

6.2.2 Some Easy Properties

Proposition 6.1 (Basically Free) *Let $A \subseteq M$. Then*

1. $A \subseteq \operatorname{dcl}(A)$
2. $\operatorname{dcl}(\operatorname{dcl}(A)) = \operatorname{dcl}(A)$
3. $\operatorname{dcl}(A) = \bigcup \left\{ \operatorname{dcl}(\{a_1, \ldots, a_n\}) : n \in \mathbb{N},\ a_j \in A \right\}$

The third of these properties is referred to as having *finite character*. These each feel reasonably obvious enough that we can go without proof, and the interested reader should have little to no trouble coming up with one themselves. The *non*-trivial property, however, that we *will* be proving is referred to as the *Exchange lemma*, which we will curiously notate as a theorem.

6.2.3 A Less Easy Property

Theorem 6.1 (Exchange (Pillay & Steinhorn [?])) *Suppose $A \subseteq M$, and $a, b \in M$. Then if $b \in \operatorname{dcl}(A \cup \{a\})$, but $b \notin \operatorname{dcl}(A)$, then we may 'swap' a and b, concluding that $a \in \operatorname{dcl}(A \cup \{b\})$.*

Whilst this may seem apropos of nothing and not terribly well-motivated, with a bit of squinting, head-scratching, and perhaps dusting off a copy of *Linear Algebra and Applications 6e* or some other such undergraduate textbook, you may realize that this is in fact a generalization of a fact you know quite well – the Steinitz Exchange Lemma; that is, that any two bases of isomorphic (finite) vector spaces contain the same number of elements. Much like many model-theoretic results, and as is sort of the point of model theory, we see this sort of thing often – where something abstract-feeling and seemingly far-removed has immediate consequences should specification to any particular model be made. Now let's prove this, relying heavily on the Monotonicity theorem to do so.

Proof (of Exchange (Theorem 6.1)) Adding constants for elements of A, we may suppose that $A = \emptyset$. So, then every $b \in \operatorname{dcl}(\{a\})$ and *not* in $\operatorname{dcl}(\emptyset)$. Since $b \in$

dcl $(\{a\})$, then there is an $\emptyset$-definable function, f, with $a \in \operatorname{dom}(f) \subseteq M$ and such that $f(a) = b$ (by definition).

Since $\operatorname{dom}(f)$ is $\emptyset$-definable, it is a finite union of $\emptyset$-definable points and $\emptyset$-definable intervals. Supposing a was one such point, then a would be $\emptyset$-definable, and by b in the definable closure of a, we would also have b $\emptyset$-definable – but we know it not to be by assumption, and so this is a contradiction. Thus, a is *not* a point in $\operatorname{dom}(f)$, and we can thus assume $\operatorname{dom}(f)$ to be an $\emptyset$-definable open interval, I, with $a \in I$. By Monotonicity theorem, we can assume that f is strictly monotonic (or constant) on I. If constant, then

$$b = f(a) = f(m)$$

for m the midpoint of I (which must be $\emptyset$-definable by I $\emptyset$-definable), and further by f $\emptyset$-definable, so too is b $\emptyset$-definable. Again, this contradiction implies f is non-constant, and so must be strictly monotonic.

Then, $f^{-1}(b) = a$ is well defined, and so a is definable from b – that is, $a \in \operatorname{dcl}(\{b\})$ as desired. □

Having those easy properties above (in particular finite character) along with exchange means that the concept of definable closure is what is referred to as a *pregeometry*. That is, dcl is a pregeometry in any model of the theory of $\mathcal{M}$. For those of us who have not previously encountered the idea of pregeometries, this is significant because a pregeometry comes equipped with a notion of *dimension* – which is exactly what we're after right now. Even if you are unfamiliar with the term, I can assure you, dear reader, that you are intimately familiar with several examples of these objects. Vector spaces, projective and affine spaces, and algebraically closed fields are all examples of pregeometries– although it should be noted that these are of quite different classifications (within the realm of this area of study) from one another [**?**].

Suppose $A \subseteq M$, $a \in M^n$. We define

$$\dim({}^{a}/_{A}) = \min\left\{ |a'| : \ a' \text{ a subtuple of } a \text{ and } a \in \operatorname{dcl}(A \cup a')^n \right\}.$$

An alternative characterization is as follows: we call some $X \subseteq M$ *independent* over A if for every $x \in X$, we have $x \notin \operatorname{dcl}(A \cup (X \setminus \{x\}))$. Then, the dimension as first defined is given by the *maximum* cardinality of a subtuple which is independent over A. That is,

$$\dim({}^{a}/_{A}) = \max\{|X| : \ \forall x \in X \,.\, x \notin \operatorname{dcl}(A \cup (X \setminus \{x\}))\}$$

We assume that $\mathcal{M}$ is sufficiently saturated and that all parameter sets are small relative to this saturation claim. We will not be discussing at any great length what this means, but we will note that this is actually a stronger condition than we necessarily need to take – it is just a convenience that will make things a bit easier down the line. Consider now the following definition:

Definition 6.4 Suppose $X \subseteq M^n$ is definable over A. Then we define the dimension of X to be, predictably,

$$\dim(X) = \max\{\ \dim(a/A)\ :\ a \in X\ \}.$$

Remark 6.1 Without the assurance of sufficient saturation, the points $\{a/A\}$ may not even exist, and so instead we could quantify over all elementary extensions of our model – but for our purposes, we acknowledge, and go on to ignore this small potential snag.

Although we have been making reference to some set of parameters, A throughout, it turns out that as long as you take a "small" set (again, relative to the saturation), then the choice of A does not truly affect the dimension of the definable X. To be clear, we will not be clear about quantifying appropriate "smallness" relative to the saturation in this course, and simply sweep the issue under the rug for perhaps a secondary, more rigorous course on the topic.

Informally, we can envision that for two sufficiently small sets of parameters, $A \subseteq B \subseteq M$, we can clearly see that for definable $X \subseteq M^n$, we get $\dim(X)_B \leq \dim(X)_A$. Supposing the dimension with respect to A to be k with point $a \in X$ witnessing the dimension, without loss of generality we can take the first $(a_1, \ldots, a_k)$ to be independent over A – but then by saturation, we can similarly find k independent such $(b_1, \ldots, b_k)$ over B with type matching that for our a. Then, our $(b_1, \ldots, b_k)$ extends to $b \in X$, and so $\dim(X)_B \geq k$ as well. Thus, the choice of parameters is irrelevant (again, given this caveat).

6.3 Reconciling Dimensionality

Ultimately, our goal here is to show that the dimension we are currently discussing is, in fact, equivalent to that discussed in the previous chapter. To this end, consider the following lemma, which topologically characterizes the independence of points.

Lemma 6.1 *Suppose that $a \in M^n$, and $A \subseteq M$ (small). Then, the coordinates, $(a_1, \ldots, a_n)$ of a are independent over A if and only if every A-definable set $X \subseteq M^n$ with $a \in X$ has non-empty interior.*

Proof (of Lemma 6.1) We start with the easy direction, and use a proof by contradiction. Suppose $a_1, \ldots, a_n$ are *dependent* over A. Without loss of generality, take $a_n \in \operatorname{dcl}(A \cup \{a_1, \ldots, a_{n-1}\})$. Then, our language contains some formula, φ, with parameters from A, such that a_n is the unique x satisfying $\varphi(a_1, \ldots, a_{n-1}, x)$. Now, let

$$X = \{\ x \in M^n\ :\ \varphi(x_1, \ldots, x_n) \text{ holds and } x_n \text{ uniquely satisfies } \varphi(x_1, \ldots, x_n)\ \}.$$

By uniqueness of x_n, we must have that X has empty interior and $a \in X$. Now, for the other direction we use cell decomposition by induction on n.

If $n = 1$, and if a_1 is independent over A, then a_1 is not in any finite A-definable set. So, if $a_1 \in X$ for X A-definable, then X is infinite, and so contains an open interval. Thus, it has non-empty interior.

Supposing this holds for M^n, we take $a \in M^{n+1}$ with coordinates $(a_1, \ldots, a_{n+1})$ independent over A, and suppose $X \subseteq M^{n+1}$ to be A-definable and having $a \in X$.

It bears mentioning here, as we are about to use this fact but have neither mentioned, proved, nor intend to prove it, but the following is true: If we have a set of cells decomposed over some set of parameters, A, then we can take each of those cells to be defined over that same set of parameters.

Recall that by cell decomposition, we get to suppose that $X = C$ is a cell defined over A. As such, C may either be

- $C = \operatorname{graph}(f)$ for $f: C' \to M$ A-definable – in which case $a_{n+1} = f(a_1, \ldots, a_n)$, leading to a contradiction of independence over A; or
- $C = (f, g)_{C'}$ for $f, g: C' \to M$ continuous and A-definable with $f < g$ (or one of the functions $\pm\infty$). This latter case must be true by exhaustion.

So, we get that $(a_1, \ldots, a_n) \in C'$. We have that $a_1, \ldots, a_n$ are independent over A and that C' is an A-definable set containing them – and so must have interior. By the inductive hypothesis, C' is an open cell, and then so too must be C, meaning X has non-empty interior.

With both directions proved, we are now finished. □

Immediately putting that lemma to work, we can show the following:

Proposition 6.2 *Let $X \subseteq M^n$ an A-definable set, and $k \leq n$ a possible dimension for A. Then* $\dim(X) \geq k$ *if and only if there is a coordinate-projection, $\pi: M^n \to M^k$ such that the projection $\pi(X)$ has non-empty interior.*

Remark 6.2 The maximum such possible k is sometimes referred to as the *topological dimension* of the set – although (and as we have bemoaned before) this term is also used to refer to a variety of inequivalent concepts. It is also perhaps worth noting that, in our setting, this topological dimension agrees with the definable (or cellular) dimension, just as we are about to show so too does the algebraic dimension we have been discussing here.

Corollary 6.1 *For $X \subseteq M^n$ a definable set, the notion of* $\dim(X)$ *as described in this chapter (Definition 6.4) agrees with that of the previous chapter (Definition 5.1).*

We will come back to prove this corollary in a moment, for now turning our attention back to a proof of the neglected proposition.

Proof (of Proposition 6.2) We start with the forward direction. Suppose $\dim(X) \geq k$. Then for our $a \in X$ of interest, there is a k-tuple that, without loss of generality, can be taken to be $a_1, \ldots, a_k$ that are independent over A. Then, should we take the projection onto the first k coordinates, we know that $(a_1, \ldots, a_k)$ lies in that projection, and so by the lemma we just proved, that projection has non-empty interior.

Now for the other direction, we suppose such a projection, $\pi\colon M^n \to M^k$ exists such that $\pi(X)$ has non-empty interior. Again, we suppose for convenience that this occurs for the first k coordinates. Then, $\pi(X)$ is definitionally A-definable, and having non-empty interior, must contain an open A-definable box, $U = I_1 \times \ldots \times I_k$. Using saturation, we inductively find $(a_1, \ldots, a_k) \in U$ independent over A. Then, taking $a \in X$ such that $\pi(a) = (a_1, \ldots, a_k)$, we have that this witnesses the dimension of X is no less than k. So, $\dim(X) \geq k$. □

We now prove the capstone corollary to these last two chapters.

Proof (of Corollary 6.1) Ummmmmmm □

Chapter 7
A Slew of Results on Smooth Cell Decompositions

Abstract To wrap up now what may have felt like a very long first section, we state a number of results on *smooth cell decompositions* – in particular starting with a definition of what makes a cell decomposition smooth in any case. For the beleaguered reader, just champing at the bit for discussion of Pila-Wilkie to start properly: we promise, this is the last bit between you and that carrot that's been dangling in front of you for that past while. For the reader, however, feeling that more detail should be wrung out before moving on, it is with great disappointment that we inform you that you have chosen the wrong text to follow, and that it took this many chapters for you to come to that conclusion. Either way, make room somewhere in your mind for a few more definitions, lemmata, propositions, and theorems, and we will quickly be on our way to a not *too* involved proof of the theorem it feels we set out to prove so long ago.

This chapter in particular really is more a survey of useful results and ideas than the preceding chapters, with much less attention paid to proving the claimed statements either at all, or with much degree of rigour. For those interested in such detail, they are encouraged to reference (as one can for most that we have done and will do here) *Tame topology and O-minimal structures* [**?**].

7.1 Smoothness

Recall our setting; we are in an o-minimal expansion of an ordered field, $\mathcal{M} = (M, <, +, \cdot, 0, 1, \ldots)$, and so much of what you'll recall from your classes in calculus and analysis apply – for example, the notion of differentiability of a univariate function on a point in the interior of its domain can be well-defined. As a result, we use this ordered-field structure to define continuous differentiability in almost exactly this way.

Note that we hand-wave quite a lot of the calculus-centric content here, as is it (reasonably) presumed to be relatively obvious how one would extend or make any arguments presented either more rigorous or work in higher dimensions than we discuss. We hope, for those ardent fans of the basics of calculus, that this is not too much a of a disappointment.

Theorem 7.1 *Suppose $f\colon (a,b) \to M$ is definable, and $r \in \mathbb{N}$. Then f is C^r except at no more than finitely-many points (where C^r is r continuous differentiability)*

Proof (or rather, a bit of reasoning about the idea) Much like in the calculus one should find themselves familiar with, we can define C^r broadly as usual, and we will heuristic our way through how one does this when $r = 1$. We have the limit definition of the derivative of f, and view the limit at each point from below and above due to the ordering of our field. We should note, of course, this holds necessarily only on some subset of the domain, rather than its entirety by necessity. By Monotonicity theorem, we can assume continuity of f, and then if we can show the limit from above exists and is positive on some interval, then f is strictly increasing (vice versa for decreasing). Then, after a slight bit of hand-waving, if we can show the two functions (limits from above and below) are M-valued (non $\pm\infty$), then f is actually differentiable, and can only be $\pm\infty$ at finitely-many points. Again, this is not claimed to be rigorous by any means. □

7.2 Smooth Cells and Decompositions

Of course, to get to our usual definition (or rather, its analogue), we could follow all the standard methods (mostly) one sees in their first calculus course, but for our purposes, we move right along to talking about *cells* again.

Definition 7.1 (C^r-cells) A C^r-cell is defined in almost precisely the same way as we did a cell originally, save for the requirement that all functions involved (as in, as part of definitions and proceeding results) to be C^r themselves. For brevity, we will not re-enumerate the definition here.

Very quickly then we move on to *smooth* cell decomposition, but won't prove it – again in our excitement and closeness of what is to come.

Theorem 7.2 (Smooth Cell Decomposition) *Suppose $r \in \mathbb{N}$, then (just as in the Cell Decomposition theorem there are two parts)*

1. *Suppose $X_1, \ldots, X_k \subseteq M^n$ are definable. Then, there exists a cell decomposition, $\mathfrak{D}$, compatible with $X_1, \ldots, X_{k]}$ such that each $C \in \mathfrak{D}$ is C^4.*
2. *(In analogy to piecewise-continuity) If $f\colon X \to M$ is definable, $X \subseteq M^n$, then there is a cell decomposition, $\mathfrak{D}$ into C^r-cells such that the restriction of f to C is C^r for each $C \in \mathfrak{D}$.*

Proof (or rather, how one might prove Theorem 7.2) No proof is provided, although it doesn't vary too much from that of the proof of Cell Decomposition *without* smoothness. As before, one proceeds by induction on n (we know, small promise broken), although now, similarly to how in the proof of the Cell Decomposition theorem we proved two inductive statements back and forth, here we do a similar thing where we prove that, for the given function in 1, we prove the set of points at which that function is continuous must be dense in X. Again, we proved something very similar before – just now this proves the existence of derivatives of certain type. In actuality, we get to use the Cell Decomposition theorem proper to make things a bit easier on ourselves, making this proof (arguably) a bit less difficult. □

And now finally (which must feel shockingly early given the length of this chapter), we are going to finish off with a statement that we will come to use, but do nothing here to prove or really discuss at much length.

Recalling that our expansion is over an ordered real field, we have been given quite a bit of specificity already – so we may ask ourselves whether we can actually improve upon the r in our C^r specification (as in, taking it to be ∞ rather than just finite), or even better, just analytic. We may have set up the question a bit too cheerfully there then, because the answer is actually **no**, we may not. While the original result is attributable to Pila and Wilkie, the results we present here is of Le Gal and Rolin [**?**]

Proposition 7.1 (Le Gal et Rolin [?]) *Given an o-minimal structure* $\tilde{\mathbb{R}} = (\overline{\mathbb{R}}, \ldots)$ *which* does not *have* C^∞ *cell decomposition* ***as defined in the usual way****, then the structure does not have it for any consistent definition of the concept.*

We bold that second part of the (very informally given) statement in particular, to emphasize that there is an *normal* way of doing things, as inspired by the calculus we know on the reals. It is to that **usual way** that is being referred.

Funnily enough, however, one bumps into and finds themselves interested in structures that *do* have C^∞ cell decomposition quite often, which is quite a nice thing indeed. And it is with this we come to what may seem an abrupt conclusion (of Part 1).

We covered a lot here, much of each piece building on the last – so one would be forgiven (praised even) for giving the previous several chapters a quick once, twice, or thrice-over just to make sure everything is solidly in place in their mind map. In the next section, we start on the Pila-Wilkie theorem proper, and the two broad constituents that make it up – each of them a piece of meaningful machinery in its own right. This first section should have been a (hopefully) reasonably good primer on what one needs to know about o-minimal, cell decompositions, smooth cell decompositions, and dimensionality to fully understand the results we are about to come upon, and the proofs we use to ensure their correctness.

Part II
The Pila-Wilkie Theorem

Now we start properly on the promise we set out to make good on from the start: an honest proof of the Pila-Wilkie theorem. The previous part, in no means brief and right to the point, was hopefully either a good refresher or first introduction to the necessary machinery we are about to put to use to prove this eponymous theorem. As mentioned before, there are two main parts to this proof: thing 1, and thing 2. We now have everything we need to put those together in full gruesome (meant in a good way, of course) detail – albeit using more modern proofs than originally used by Pila and Wilkie.

We use idk the source on the 1 for thing 1 first, using blank or something or the other instead of blank for the first part. Then, we use the similarly recent proof of the result of fuck me i forget but it was 2021 and was something-Gromov to get Resulto numero 2. From there, Pila-Wilkie itself is far from a stretch, and we will soon have built it up in (almost) all its full spectacle.

Chapter 8
Finally: A New Beginning

Abstract One can imagine the collective sighs of all that read through the previous section to get to this, the *exciting* part of the module, almost-deafening. It is here we will first start our proper discussion of the Pila-Wilkie theorem, along with some bits and bobs as to how we will go about actually proving it, and some previous of what it entails. There are two broad *bits* to the proof we will discuss – although each is deserving of a chapter (or perhaps even more) on its own, and so there proofs are outsourced to later to come parts. For now, we just start by dipping our toes into what we've been building to through the previous entire first part.

8.1 How We're Going to Prove It

As mentioned before – but bears repeating – our approach to proving Pila-Wilkie theorem (in our special little case) broadly follows that of Bhardwaj and van den Dries [**?**]; a recent paper that only came out this year (should you be reading this in 2022). Their major contribution includes the use of semialgebraic cell decompositionsto simplify several arguments of the original proof by Pila and Wilkie. They also provide a full treatment of a recent variant [**?**] of the (classical) Yomdin-Gromov theorem [**?**] (again, assuming your readership takes places not too soon after these words reach the page) in 2021, and a result of Bombieri and Pila [**?**], which is also made use of. In our case, we are going to be following the (recent) treatment of what we will come to colloquially call *parameterization* by [**?**].

[1]

[1] Please note that this horizontal line is *not* intended to cause any calming effects, as we have used a *similar* (but distinct) line for previously. Delineation is not only all we intend it be used for, but it is all that the reader is permitted to take from it. Observe with caution and integrity.

8.2 Some Recollections and Definitions

Let us now begin with a recollection of our setting. We are no longer going to speak to the generality of any ordered group, and instead an o-minimal expansion $\tilde{\mathbb{R}} = (\overline{\mathbb{R}}, \ldots)$ of the ordered field of $\mathbb{R}$. Ultimately, the Pila-Wilkie theorem is about rational points on definable sets – so taking definable sets over this expansion (supposing points have n dimensions) and intersecting with $\mathbb{Q}^n$, then Pila-Wilkie tells us something about the number of such points we can expect; in short, how *few* of them we should expect up to a certain *height*. This concept was briefly touched upon early in the first part, but just to recap, we can define certain functions that characterize the height of a number (à la multiplicative height for rationals), and then bound the size of this set from above as a function of that height parameter. Without use of a height, there could of course be infinitely-many points, so this concept is crucial (and broadly the *point* of the theorem).

Let's begin by again defining the multiplicative height on $\mathbb{Q}$.

Definition 8.1 (Multiplicative Height on $\mathbb{Q}$) Suppose some $q = {}^a/_b$ for a, b coprime, $a, b \in \mathbb{Z}$, and $b \neq 0$. Then, the height function is denoted by $H\colon \mathbb{Q} \to \mathbb{Z}_{\geq 0}$, and defined

$$H(q) = \max\{\, |a|,\ |b| \,\}$$

for $|\cdot|$ simply the absolute value on $\mathbb{Q}$.

8.2.1 A Little Idiosyncrasy

For reasons of preserving syntax used consistently (rather than some off-handed syntax we have been replacing) throughout this part of the course, we overload the use of H as both a function on the ration numbers that maps some $q \in \mathbb{Q}^n$ to its height, *as well as* the integral height itself. You will often find sentences of the form 'fixing some $H \geq 1$, we take $X(S, H)$ to be the set $\{s \in S : H(s) \leq H\}$'. While a bit confusing, especially at first blush, this choice for syntactic overloading was made due to its persistence throughout Dr Jones' lectures.

Corollary 8.1 *Fixing some height, H, there are only finitely many $q \in \mathbb{Q}$ with height H.*

Proof This should be clear. □

As usual, we now address the multidimensional-case

Definition 8.2 (Multiplicative Height on $\mathbb{Q}^n$) Suppose some $q = (q_1, \ldots, q_n) \in \mathbb{Q}^n$ for $n \in \mathbb{N}$, all in lowest terms as before, then we define the height function to be the coordinate-wise maximum:

$$H(q) = \max \left\{ H(q_j) : 1 \leq j \leq n \right\}.$$

With that in mind, we have the following definition.

Definition 8.3 Let $X \subseteq \mathbb{R}^n$ and $H \in \mathbb{N}$. Denote the set

$$X(\mathbb{Q},\ H) = \{ q \in X \cap \mathbb{Q}^n\ :\ H(q) \leq H \}$$

and further, refer to the algebraic and transcendental parts of X and their corresponding sets such that we have

$$X(\mathbb{Q},\ H) \setminus X^{\text{alg}}(\mathbb{Q},\ H) = X^{\text{tr}}(\mathbb{Q},\ H),$$

where we trust the superscript notations to be clear.

To rephrase what we said above, the Pila-Wilkie theorem gives us an upper bound on

$$|X^{\text{tr}}(\mathbb{Q},\ X)|,$$

with H as a parameter controlling the cardinality of this set. Ideally, of course, we want good bounds on this, especially as $H \to \infty$ – which quite nicely, Pila-Wilkie gives us.

8.3 Building to Pila-Wilkie

A special case of this coming from Pila (building on an earlier paper he had worked on with Bombieri) is the following:

Theorem 8.1 (A Special Case by Pila) *Suppose $f : [0,1] \to \mathbb{R}$ is analytic and transcendental, and further that $X = \text{graph}\,(f)$. Then, for all $\varepsilon > 0$, there exists some $c > 0$ such that for any $H \in \mathbb{N}$,*

$$|X(\mathbb{Q},\ H)| \leq c \cdot H^{\varepsilon}$$

In general, a definable set may contain a rational line (with interior, that is), in which case the number of points is enormous. Here is where the idea of algebraic and transcendental parts of X, again denoted x^{alg} and X^{tr}. However, we now do you the service of definition, such that anything useful can be done with them.

Definition 8.4 (Algebraic Part of X) The *algebraic part* of some $X \subseteq \mathbb{R}^n$, denoted x^{alg}, is given by the union of all connected, infinite, $\overline{\mathbb{R}}$-definable subsets of X. That is, definable in the real field (with parameters).

By our definition a few above, it is clear then that the set of interest to us is $X^{\mathrm{tr}}(\mathbb{Q}, H)$, where we take all of X and throw away all the 'junk' bits we identify with X^{alg}. Notice that we assume our X to be definable, but these partitions into algebraic and transcendental sets can be *far* from definable, especially x^{alg}. Indeed, understanding each of these two sets are different problems from one another, requiring different approaches to say meaningful things about. For now, we focus on X^{tr}.

8.4 The Pila-Wilkie theorem Proper and the Two Ingredients

We stated it in the Introduction so as to give something to look forward to, but we now do so again – without further ado, the Pila-Wilkie theorem:

Theorem 8.2 (The Pila-Wilkie Theorem) *Suppose $X \subseteq \mathbb{R}^n$ is a definable set (in our o-minimal structure). Then for all $\varepsilon > 0$, there exists some $c > 0$ such that for all $H \in \mathbb{N}$,*

$$|X^{tr}(\mathbb{Q}, H)| \leq cH^{\varepsilon}.$$

Now this right-hand side might be (should be) looking very familiar from just a moment ago – but recall that that was for an example on *all* points in the set, and in a special case. Here, there result is much more general, and only pays attention to the transcendental parts of X.

Nice though it would be to just put a quick proof underneath and sate ourselves of this theorem, it does unfortunately require a bit more work than that. We are going to go through in the next few chapters and prove the necessary constituents to a proof of Theorem 8.2. The two main bits are as follows:

1. **The o-minimal bit**: Suppose we have a definable set $X \subset (0, 1)^n$ (by normalization) and some $r \in \mathbb{N}$. Then there exists a set of functions, $\{\varphi_1, \ldots, \varphi_k\}$ where $\varphi_j : (0, 1)^{\dim(X)} \to X$ that are C^r definable maps such that all derivatives of each map up to degree r are bounded (in modulus) by 1 and whose union is X.
2. **The number theory bit**: Let $k, n \in \mathbb{N}$, $k < n$. Then for each $d \in \mathbb{N}$, there exists some $r \in \mathbb{N}$ and $c, \varepsilon > 0$ with the property that if $\varphi : (0, 1)^k \to (0, 1)^n$ is a C^r map with derivatives bounded by 1 up to order r (note: not necessarily definable), and $X = \text{image}\,(\varphi)$, then for $G \in \mathbb{N}$, the set $X(\mathbb{Q}, H)$ is contained in at most $c \cdot H^{\varepsilon}$ algebraic hypersurfaces of degree at most d. Further, this $\varepsilon \to 0$ as $d \to \infty$.

In essence, this first point means that definable sets possess a certain type of parameterization, we will come to know as *cellular r-parameterizations*. In reference to the work we follow, the Yomdin-Gromov theorem was mentioned (which was not actually a joint work, but work continued by Gromov, inspired by Yomdin), which essentially shows this result over particular special cases. We use a newer variant as mentioned before, which will allow us to go on to show that Pila-Wilkie generalizes the previously given – first to more general o-minimal structures, and then to further generalizations we will discuss later on.

Just as a point of note, it is this second part of the proof that we have (almost *ad nauseam*) been mentioning that has its roots in work between Bombieri and Pila [**?**].

Unfortunately, unlike we have seen with a couple other multi-part results in the first part, the Pila-Wilkie theorem does *not* immediately and obviously follow, even once we have these two results – so that will take a bit of doing in and of itself. So, for the next three odd chapters, we will preoccupy ourselves with proving first the *o-minimal bit*, then the *number theory bit*, and finally, pulling it all together. The third and final part of this set of notes will then discuss, without too much detail, some extensions, generalizations, and applications of the Pila-Wilkie theorem, and how we have seen it being used since its inception.

8.5 Just a Couple of Remarks Before we Begin

To not faff on too long, here we just enumerate a few remarks on what we've just stated, and a bit of what is to come. These remarks will not be crucial to what follows, but are of significant interest, nonetheless.

Remark 8.1 (A Series of Remarks)

1. The bound given in the Pila-Wilkie theorem of cH^{ε} is the *best* possible bound. It is already the best in the special case of Pila on curves given above – and so this implies that Pila-Wilkie cannot be improved over the reals. In general, the proof involves a bit more.
2. However, we often can do better in 'effective' and common examples. The first point merely points out that, in general, we can come up with structures that can do no better than the general case – but in many particular restrictions, we can do better. A good example is in $\mathbb{R}_{\text{exp}}$, where Pila conjectured that we can bound by $c \log(H)^k$. As well, Binyamini and Novikov [**?**] prove such a bound for $(\overline{\mathbb{R}}, \exp_{[0,1]}, \sin_{[0,1]})$, the restricted exp and sin functions. These sorts of (improved) bounds are known are many classes of sets, though very often they are *not* the definable sets of some o-minimal structure, requiring more restriction than we assume here.
3. The c in the bound cH^{ε} is *not* effective – which is to say, not easily or reasonably describable or calculable. Given that the nature of this work is often *not* computational, this is not necessarily a terrible thing – although should it be effective, it would be the odd mathematician to be unhappy about it.
4. There are various stronger results (eg. bounds on number field degrees over $\mathbb{Q}$, rather than rational points), where the height function is no longer rational but algebraic. For example, consider

$$| X^{\mathrm{tr}}(d, H) | = | \alpha \in X : [\mathbb{Q}(\alpha) : \mathbb{Q}] \leq d, \; H(\alpha) \leq H |$$
$$\leq cH^{\varepsilon} \text{ when } c = c(d).$$

Here what we hope for (in nice structures, not in general), something of the form

$$c \cdot d^k \cdot \log (H)^{\ell}.$$

Such bounds also have applications that generally differ to those of Pila-Wilkie (some of which we will come to see later on).

5. Finally, there has been a good bit of work put in (and still being put in) to the *parameterization result*. This was the original result by Yomdin that was later worked on by Gromov. As it turned out, of course, Pila and Wilkie were able to use that result (in conjunction with a good bit else) to prove the theorem on which this course is primarily interested. Further improvements to these parameterization results have effective application such as, for example, the *number* of $\varphi_1, \ldots, \varphi_k$ needed, as a function of r, as stated in part (ingredient) 1 in our proof of the Pila-Wilkie theorem. Or, for example, when can we do better than having these function be C^r – such as being C^{∞} with some bound (not necessarily 1, as we have) on derivatives of all orders.

If any of these topics seemed of particular interest to you, then it will be a bit of a disappointment that we do *not* discuss any of them in great detail here – and those we do, we do so at more of a surface-level than the other content presented. That said of course, the references provided (and still currently worked on literature in the area) is a great place to look for further information on these topics.

Chapter 9
The O-Minimality Bit (Parameterization)

Abstract We initially and have for a bit now been calling this section 'the o-minimal bit' in order to differentiate it from the number theoretic part that is to come next. However, going forward, and as was alluded to a bit in the last chapter, this 'o-minimal bit' is about parameterization, and we shall generally go on to refer to this section and result as the 'parameterization result', or simply 'parameterization'. SAY MORE.

9.1 Setting

In this section, we actually step back a bit from the specificity of $\mathbb{R}$, and work over a more general o-minimal expansion of an ordered field, $\mathcal{M} = (M, <, +, \cdot, 0, 1, \ldots)$. We let $I = (0, 1)$ the open unit interval, interpreted in our model, $\mathcal{M}$. For the next little while, we have Binyamini and Novikov [**?**] to thank for what is to come. We begin with some introduction to the world in which we will find ourselves for the next little bit.

Definition 9.1 A *basic cell*, $C \subseteq I^n$ is a set given by a product of copies of I and singletons containing 0, such that there are in total n factors of this set product. On such a cell, C, a continuous map, $\varphi\colon C \to I^n$ is called *cellular* if

1. For $\varphi = (\varphi_1, \ldots, \varphi_n)$ with each φ_j only dependent on the first j coordinates.
2. For each j, and each $(x_1, \ldots, x_{j-1})$ in the projection of C onto the the first $j-1$ coordinates, the function $x_j \mapsto \varphi_j(x_1, \ldots, x_j)$ is strictly increasing.

This next remark is one to remember, as it is a fact that we will be using and making reference too often.

Remark 9.1 Suppose $\varphi\colon C \to I^n$, $\psi\colon C' \to C$ are both cellular. Then so too is their composition, $\varphi \circ \psi\colon C' \to I^n$.

Checking that the two necessary conditions holds is rather trivial here, and so not included. This next bit is mostly just notational.

Definition 9.2 (and notation for the norm) Given $U \subseteq \mathbb{R}^n$ and $f : U \to \mathbb{R}^m$ a C^r map, we set the r-norm of f to be

$$|| f ||_r = \left(\max_{|\alpha| \leq r} \sup | (D^{(\alpha)} f_j(x) | \right) \cdot \frac{1}{\alpha!}$$

Note of course that the scaling by factorial α is not strictly relevant, but more of a convenience. Ultimately, just so it is entirely clear what the point of all this is, what we are aiming to find here is parameterizations into the cellular maps on which these norms are bounded. To that end, consider:

Definition 9.3 Let $r \in \mathbb{N}$ and $x \subseteq I^n$ definable. Then a *cellular r-parameterization* (CRP) is a finite set, Φ, of definable C^r cellular maps such that each

$$||\varphi||_r \leq 1$$

for each $\varphi \in \Phi$, and that the union of their images covers X – that is,

$$\bigcup_{\varphi \in \Phi} \text{image}\,(\varphi) = X.$$

Suppose now we want a cellular r-parameterization between definable sets – what does that require?

Definition 9.4 Suppose we have some $Y \subseteq I^m$, and $f : X \to Y$ definable. Then. cellular r-parameterization of f is a cellular r-parameterization Φ of X, with the *extra* property that

$$||f \circ \varphi||_r \leq 1$$

for all φ in Φ.

9.2 Do These Even Exist?

As is often amusing to those outside of mathematics, and even those *in* maths, (at least from the perspective of a graduate student in the area), so often we find ourselves setting up a series of definitions and such defining objects that we wish to study – with a capstone theorem being that they exist at all. While this may seem the completely backwards way to go about doing things, it is often one that ends up working quite well – but not always. We're sure, of course, that we don't need to recite the (likely apocryphal) story of the aspiring mathematician who worked for quite a while in this way on Hölder continuous maps with $\alpha > 1$ – having only before seen the case with alpha in the unit interval. It was only at his defence that he was asked to show an example of such a *non-constant* map – which, of course, do not exist. Of course, similar such apocrypha exist about anti-metric spaces, and other such seemingly fun ideas. Luckily, or perhaps even more accurately, *skillfully*, it turns out that our setup *does* end up working out for us as desired. We now go on to state and prove the theorem that these objects we have described above do, in fact, exist, and are non-trivial.

9.2.1 Yes – And Here's a Proof

Theorem 9.1 (Parameterization (by Binyamini & Novikov)) *Let $n \in \mathbb{N}$. Just like with the proof of Cell Decomposition, we are going to have two inductive towers, where we again employ the back-and-forth method to prove the theorem for all n. Note as another similarity that the first statement is one about sets, and the second of functions. These are given:*

$(I)_n$ *If $r \in \mathbb{N}$ and $X \subseteq I^n$ is definable, then X has a cellular r-parameterization.*
$(II)_n$ *If $r,\ m \in \mathbb{N}$ and $X \subseteq I^n$, $Y \subseteq I^m$, and $f: X \to Y$ definable, then f has a cellular r-parameterization.*

Of course, it is clear that $(II)_n$ is the more difficult of the two to prove, and in fact, $(II)_n$ can then be used to prove $(I)_n$ (so we have perhaps presented them in a bit of a disordered way).

Before we start of the proof, we give provide a quick remark on these cellular r-parameterizations– and that is that we can *almost* compose these cellular r-parameterizations. This is to say essentially the following:

Remark 9.2 Suppose Φ is a cellular r-parameterization of some $X \subseteq I^m$ and that for each $\varphi \in \Phi$, given $\varphi: C \to X$, we have a cellular r-parameterization Ψ_φ of C. What we then would like to do is take all the φ and compose with all corresponding Ψ_φ, and receive cellular maps. We *almost* get this result. For example, taking some $\psi \in \Psi_\varphi$, We can compute to show

$$||\varphi \circ \psi||_r \leq c_{n,\ r},$$

which, although not necessarily 1, is finite and so bounded. Of course, we can *make* this constant one by doing a bit of extra parameterization. Start by covering $C' = \text{dom}\,(\psi)$ with $(c+1)^k$ boxes, where we take $k = \dim(\psi)$, we let $c = c_{n,\ r}$, and each box is itself a translate of $(0, {}^1\!/_c)^k$. This is justified due to the fact that on each such box, we have a natural linear map from $(0, 1)^k$ to $(0, {}^1\!/_c)^k$ – and in so applying these natural maps, we end up with $c_{n,\ r} = 1$ after the computation, for each φ and ψ. So now, at the expense of ending up with a whole bunch more maps, we end up with a cellular r-parameterization. What we just described – the process of going from a constant of n and r to 1 using linear maps will hence be referred to as **linear substitution**. Going on, and in the next proof, we are going to be justifying parts *by linear substitution* quite a bit, so what may have seemed like a brief detour is actually quite important

Remark 9.3 This may seem completely apropos of nothing, but now and going on for the rest of these notes, we assume our model, $\mathcal{M}$, to be $\aleph_0$-saturated. If this idea is familiar to you: great. Otherwise, don't worry too much about it. Essentially this assumption allows that whatever we prove here to hold for models that are less than $\aleph_0$-saturated.

Not to keep the eager reader too antsy about getting to the proof of the above theorem, but we are first going to start with a lemma (and then another, which some may even say is starting with two lemmata) that will make the proof of $(\mathrm{II})_n$ *quite* a bit nicer for ourselves.

9.2.2 A Few Lemmata to Get us Going

Lemma 9.1 *Let $n, r \in \mathbb{N}$, and suppose that every $F\colon X' \to I$ definable on a definable X' of dimension n has a cellular r-parameterization. Then for every definable $G\colon X \to I^m$ with $m \in \mathbb{N}$ and such that* $\dim(X) = n$ *has a cellular r-parameterization.*

Proof (of Lemma 9.1) Fix some $G = (G_1, \ldots, G_m)$ and let Φ_1 be a cellular r-parameterization of G_1 (which we know we can find by assumption). It is then enough to find a cellular r-parameterization of $G \circ \varphi$ for $\varphi \in \Phi_1$, and then use *linear substitution* as described it before. So, we can suppose $||G_1||_r \leq 1$. Recall of course that by our assumption, G_1 depended only on its first indeterminate – and so we can similarly get a cellular r-parameterization, Φ_2 of G_2. Taking some $\varphi \in \Phi_2$ then evidently (as before) $G_2 \circ \varphi$ is as desired – but now we need to worry about $G_1 \circ \varphi$. Generically, we get that $||G_1 \circ \varphi||_r \leq c_{r,\,n}$. We can then keep doing this for all $3 \leq j \leq m$, and then for each have that $||G_j||_r \leq C_{r,\,n}$. At this point, we can do our linear substitutions to get a bound of 1 for each j. □

Remark 9.4 It may not be immediately obvious why the above lemma is useful to us. Principally, its use will be in proving $(\mathrm{II})_n$ in the theorem, where it essentially says that we get to treat our $f\colon X \to Y$ for $Y \subseteq I^n$ as if it were a *function*, $\hat{f}\colon X \to \hat{Y}$ for $\hat{Y} \subseteq I$, rather than a *map* as it is given.

Proof (of Theorem 9.1) It should be clear that we induct on n to prove this. The statement $(\mathrm{I})_1$ is immediately obvious from what we've said earlier, so to finish the base case, we just need to prove $(\mathrm{II})_1$. Unfortunately, this is nowhere near as immediate – so much so that we actually appeal to a lemma.

Lemma 9.2 (Gromov) *Suppose $r \geq 2$, $f\colon I \to I$ definable and C^{r-1} with $||f||_{r-1} \leq 1$. Then f has a cellular r-parameterization.*

Proof (of Lemma 9.2) First, by applying the Smooth Monotonicity theorem, we can break up into intervals on each of which we assume that our f is r-times differentiable – and further that its r-th derivative is strictly monotonic (if the derivative is 0, then there's nothing further needed). Without loss of generality, we assume one of these intervals is I itself. We take $f^{(r)}$ strictly decreasing on I, and $f^{(r)} > 0$. Then, by applying the mean value theorem (which we took for granted as part of the clarity of implementing calculus-type concepts in our model), for all $x \in I$ there is some $\xi_n \in (0, x)$ such that

$$\frac{c_r}{x} \geq \frac{f^{(r-1)}(x) - f^{(r-1)}(0)}{x} \qquad = f^{(r)}(\xi_n)$$
$$\geq f^{(r)}(x)$$

for c_r some constant depending on r. Let $g : x \mapsto f(x^2)$. Then we already have that $||g||_{r-1}$ is bounded, and so it suffices to check for r. Computing this, we get

$$g^{(r)}(x) = a + c_r \cdot f^{(r)}(x^2) \cdot x^r,$$

for a simply the sum of some bounded terms (about which needn't worry). But, by the above,

$$f^{(r)}(x^2) \cdot x^r \leq c_r \cdot x^{r-2}$$

and since r was presumed to be at least 2, then this is bounded as desired. So, we get that $||g||_r < c_r$, which we can again make 1 by linear substitution. □

Gromov's lemma gave us the case where r was at least 2, so now we handle the case where r is 1.

Lemma 9.3 *Suppose $f : I \longrightarrow I$ is a definable function, and $r \in \mathbb{N}$. Then there exist definable C^r functions, $\varphi_1, \ldots, \varphi_k$ with each given $\varphi_j : I \longrightarrow I^2$, whose images cover* graph (f) *such that that $||\varphi_j||_r \leq 1$, and each coordinate of each φ_j is either constant or strictly monotonic.*

Proof (of Lemma 9.3) By smooth monotonicity, we can assume that f is C^1 on I and either constant or strictly monotonic with derivative f' sitting in one of the following intervals over I:

- $(-\infty, 1)$
- $[-1, 0)$
- $(0, 1]$
- $(1, \infty)$. □

We can then assume that $0 < f' \leq 1$ by using f^{-1} in place of f if necessary, or reversing the orientation of I (depending on which of the follow above intervals f' sits in). Then, we simply apply the previous lemma $r - 1$ times. to get a parameterization, Φ of I such that $||f \circ \varphi||_r \leq 1$, and so we take $\{ (\varphi, f \circ \varphi) : \varphi \in \Phi \}$ as our parameterization of the graph. □

So, we've now proved (through the previous two lemmata) the parameterization of the graph of f, and what we want to prove now is the existence of a cellular r-parameterization. We are now equipped to do so (recall that we are currently proving $(\mathrm{II})_1$). It suffices to do this for a definable $f: I \to I$. We apply the previous lemma to get a parameterization, Φ of the graph of f. We address each φ in the parameterization; if $\varphi \in \Phi$ has strictly decreasing first coordinate, we compose with $1 - x$ (making it increasing), and otherwise leave it be; so when we take the set

$$\{\ \varphi_1\ :\ \varphi = (\varphi_1,\ \varphi_2) \in \Phi\}$$

we are ensured that it is a cellular r-parameterization of I. Now all we need to check is composition with f holding as expected. Recall that $||f \circ \varphi_1||_r = ||\varphi_2||_r \leq 1$, and so we do have a cellular r-parameterization of f.

That may have felt like quite a lot, especially for just the base case, and so again I introduce our old friend the horizontal line, in hopes that it brings us respite in what may be an overwhelming time. If not at all distressed, feel free to ignore.

Now isn't that just *pleasant*. Try not to spend too long, and we will now move on to the inductive stage of this proof.

We assume that $(\mathrm{I})_m$ and $(\mathrm{II})_m$ hold for $m \leq n$. We start by proving the (much) easier of the two parts, $(\mathrm{I})_{n+1}$. By cell decomposition, we can suppose $X \subseteq I^{n+1}$ is a cell, and so either the graph of a function or the space between two function.

If the former, that is, $X = \operatorname{graph}(f)$ for $f: X' \to I$ a continuous definable function on some $X' \subseteq I^n$, then applying $(\mathrm{II})_n$ to this function, we get a cellular r-parameterization of f, Φ – and so taking some $\varphi \in \Phi$, we have that $\varphi: C' \to I^n$ for C' some basic cell in I^n. Simply set $C = C' \times \{\ 0\ \}$ and define

$$\begin{aligned}&\psi: C \to I^{n+1}\\ &\psi: (x, 0) \mapsto (\varphi(x),\ f \circ \varphi(x)).\end{aligned}$$

Since $||f \circ \varphi||_r \leq 1$, taking all such ψ gives a cellular r-parameterization of X as desired.

Supposing now instead that $X = (f,\ g)_{X'}$ is the space between two graphs, we apply $(\mathrm{II})_n$ to the map $(f,\ g)$ (note the perhaps confusing notation; the non-subscripted (f, g) is a *map*, and not the space between graphs) we get a cellular r-parameterization, Φ of $(f,\ g)$. Again, taking some $\varphi \in \Phi$ we have that $\varphi: C' \to I^n$ for $C' \subseteq I^n$ a basic cell. Similarly to the last case, set $C = C' \times I$ and define

$$\begin{aligned}&\psi: C \to I^{n+1}\\ &\psi: (x', x_{n+1}) \mapsto (\varphi(x'),\ g \circ \varphi(x') \cdot x_{n+1} + (1 - x_{n+1}) \cdot f \circ \varphi(x')).\end{aligned}$$

Since we have already that $||g \circ \varphi||_r \leq 1$ and $||f \circ \varphi||_r \leq 1$, we easily see that $||\psi||_r \leq c$, which can be changed to a bound of 1 by linear substitution. Thus again, we get a cellular r-parameterization of X – and now with all cases exhausted, we have proven $(\mathrm{I})_{n+1}$. We hate to remind the reader that this was the easy one, and not just by a little bit. Just because it is this author's contention that we all deserve it, *please* once, and perhaps finally, enjoy the following line.

We assume $(\mathrm{I})_m$ for $m \leq n$, and $(\mathrm{II})_m$ for $m < n$, and show that $(\mathrm{II})_n$ holds.

9.2.2.1 A Uniform Version of One of Our Statements

The following was presented as an exercise to the viewer, but we will simply be presenting it as a proposition in these notes, following our own (hopefully correct) proof.

Proposition 9.1 (Uniform Version of $(\mathrm{II})_{n-1}$) *Supposing $(II)_{n-1}$ holds, and that $f : I \times I^{n-1} \longrightarrow I$ is definable, there is a partition, $I_1, \ldots, I_k$ of I into definable sets such that for each $j = 1, \ldots, k$, there is a uniform parameterization of f, (Φ_j a set of definable maps given $\varphi : I_j \times C \longrightarrow I^{n-1}$ for C a basic cell in I^{n-1} depending on φ) such that if $a \in I_j$, then $\Phi_{j,a} = \left\{ \varphi(a, \ldots) : \varphi \in \Phi_j \right\}$ is a cellular r-parameterization of $f(a, \ldots)$.*

Proof (of Proposition 9.1) UMMMMMMM

Using $\aleph_0$-saturation and definable choice, prove the following. See ON THE PILA-WILKIE THEOREM NEER BHARDWAJ AND LOU VAN DEN DRIES □

9.2.3 Finally Inducting

With that under our belt, we are now ready to prove $(\mathrm{II})_n$. We do this by a series of reductions that will make what is a seemingly more difficult result much more within reach. First, suppose $f : X \longrightarrow I$, with $X \subseteq I^n$ definable. What we want then is a cellular r-parameterization for f. By $(\mathrm{I})_n$, we can reparameterize X, and work with each chart separately. As such, we can assume that X is a basic cell. If this has any 0 coordinates, we can use $(\mathrm{II})_m$ for some $m < n$, and we are done. So, we assume that all coordinates of X are non-zero. This of course then means that, because it is a basic cell, it *must* be I^n itself. So, we are now considering the function $f : I^n \longrightarrow I$. This is our first reduction. Our next reduction is going to be showing that we can assume f to be C^r and for each a in our interval, evaluating with a as our first coordinate gives a map (on I^{n-1}) with bounded derivative (that is, $||f(a, \ldots)||_r \leq 1$).

This is where the *uniform* version of $(\mathrm{II})_{n-1}$ becomes useful. We apply the proposition to $f : I \times I^{n-1} \longrightarrow I$ What the *uniform* version told us was that we get *finitely-many* of these parameterizing families, Φ_j as above, and a partition $(I_1, \ldots, I_k)$. Going even one level deeper, for each such I_j, we can reparameterize (as we did earlier) to assume that each I_j is just I. Thus, partitioning and rescaling, we can take each $I_j = I$ (after which point we now drop the j from the notation). We then fix some $\varphi\colon I \times C \to I^{n-1}$ for $C = I^{n-1}$. Notice we ignore the immediately solvable case (as before) where would could just appeal to one of the assumed statements for $(\mathrm{II})_m$, $m < n$ in our strong induction. Setting now

$$G(x_1, \ldots, x_n) = f(x_1, \varphi(x_1, \ldots, x_n))$$

by the parameterization from (a good bit) earlier, we get that for each $a \in I$, we have that $||G(a, \ldots)||_r \leq 1$. By smooth cell decomposition, there is a definable $Z \subseteq I^n$ having dimension less than n, and such that the function restriction $G|_{I^n \setminus Z}$ is C^r. Using $(\mathrm{I})_m$ for $m < n$, we find a cellular r-parameterization, Φ_Z for Z (which we must be able to since Z has dimension less than n). Since $\dim(Z) < n$, if $\varphi \in \Phi_Z$ has domain C and C has a 0 coordinate, then we can apply $(\mathrm{II})_{n-1}$ to each $G \circ \varphi$, for each φ. To summarize, in particular because it is easy to get lost in this, what we have done is reparameterize f on Z – and now we go on to worry about the complement of Z in I^n.

That is all to say, we are left again with the restricted function $G|_{I^n \setminus Z}$ – which the keen amongst you should notice (the restriction of) is a set in $(\mathrm{I})_n$. Even for the less keen, we know from $(\mathrm{I})_n$ that we can parameterize such sets – and so the end is (perhaps) within sight! Suppose we have a cellular r-parameterization, Ψ of $I^n \setminus Z$. Fixing some $\psi in \Psi$, we can define

$$t\colon C \to I^n \setminus Z,$$

and we presume that C is *not* I^n for reasons of non-triviality.

We know that, by definition, ψ maps into the space where G is C^r, and so $G \circ \psi$ is C^r. For $\alpha \in (\mathbb{N})^{n-1}$ a multi-index, with order $|\alpha| \leq r$, for any $a \in I$, we have to check that, being cellular (and so depending only on the first-coordinate, a),

$$|D^\alpha (G \circ \psi)(a, \ldots)| = |D^\alpha ((G(\psi_1(a), \ldots) \circ (\psi_2, \ldots, \psi_n))|$$

since ψ is cellular. And so, we can conclude that

$$|D^\alpha ((G(\psi_1(a), \ldots) \circ (\psi_2, \ldots, \psi_n))| \leq c_{n,\, r}$$

since $||G(b, \ldots)||_r \leq 1$ for each $b \in Z$ (in particular where $b = \varphi_1(a)$) and $||\psi||_r \leq 1$ – and so as ever, by linear substitution, we can make our $c_{n,\, r}$ equal to 1.

So, we can assume that $||G \circ \psi(a, \ldots)||_r \leq 1$ for each $a \in I$. We are now so close to *almost* done.

So, should we replace f with $G \circ \psi$ for φ_1, we may then assume that $f\colon I^n \to I$ is C^r, and that for each $a \in I^{n-1}$, we have $||f(a, \ldots)||_r \leq 1$.

9.2.4 A Final Lemma

We now need one quick lemma to finish off (though we will not be proving it until a bit later).

Lemma 9.4 *Suppose $f: I^n \to I$ is definable, and C^1, and suppose as well that*

$$\left| \frac{\partial f}{\partial x_j}(x) \right| \leq 1$$

for each $x \in I^n$, and $j = 2, \ldots, n$. Then, the set of $a \in I$ such that $\frac{\partial f}{\partial x_1}(a, \ldots)$ is unbounded is finite

The proof of this lemma is to come later, and for now we will just take this fact for granted.

Recalling where we were in our proof, we can take $\mathbb{N}^n$ and order it first by degree and then lexicographically. Taking then, $\alpha \in \mathbb{N}^n$ be at least in this ordering such that

$$|D^\alpha(f)| > 1$$

– that is, its derivatives are 'big'. If there is no such α, then we are finished! Otherwise, we parameterize f to bound *this* derivative and such the resulting (next) function has increased α, and continue this inductively. We now show that this is in fact something we *can* do – that is, ensure that we increase α – and then finish by induction.

First note that our multi-index, α, living in $\mathbb{N}^n$ must have $\alpha_1 \geq 1$. By the above (unproven) lemma, there are only finitely-many $a \in I$ such that $D^\alpha f(a, \ldots)$ is unbounded. As such, we can use $(\mathrm{II})_{n-1}$ to handle these cases, and so assume that for each $a \in I$, we do have $D^\alpha f(a, \ldots)$ bounded.

Now, we define

$$S = \left\{ x \in \mathrm{i}^n \ : \ |D^\alpha f(x)| \geq \frac{1}{2} \sup_{x' \in I} (D^\alpha f(x, x')) \right\}.$$

By our assumption, we know this supremum to be bounded, and so this set is actually definable (not in the o-minimal sense, just in that it makes sense to define). By definable (this time in the o-minimal sense) choice, there is a definable $\gamma: I \to S$ with $\gamma_1(t) = t$ simply the identity. Consider the map

$$t \mapsto (\gamma(t), D^{\alpha'}) f(\gamma(t))$$

where $\alpha' = (\alpha_1 - 1, \alpha_2, \ldots, \alpha_n)$. By $(\mathrm{II})_1$, we get a cellular r-parameterization, Φ, of this map. Take some $\varphi \in \Phi$, and consider

$$G(x_1, \ldots, x_n) = f(\varphi(x_1), x_2, \ldots, x_n).$$

This of course makes sense because, as part of a cellular r-parameterization of a map on the unit interval, φ maps into the unit interval itself.

A pithy little aside that is just *too* good to not mention here is Dr Jones' comment about how "one should never differentiate in public" – which we find quite amusing.

Computing for $\beta < \alpha$, we get

$$\left| D^{\beta}\, G \right| < c_{n,\, r}$$

for $c_{n,\, r}$ a constant depending on n and r as usual – so this is fine. What we need to worry about is α. When we consider $D^{\alpha}\, G$, we are differentiating f to some order *other* than specified by α (but smaller, which is okay), but we may run into problem cases where we take the α derivative of f, and then a bunch of φs come out. So, computing $D^{\alpha}\, G$, we get some number of terms dependent on n and r, and bounded by $c_{n,\, r}$ plus some term

$$\varphi'(x_1)^{\alpha_1} \;\cdot\; (D^{\alpha} f)\;\; (\varphi_1(x_1),\; x_2,\; \ldots,\; x_n)$$

with the left-most term coming from taking the derivative of the expression above in phi. By our definition of S and γ, we have that

$$\begin{aligned}
\mid \varphi'(x_1)^{\alpha_1} \;\cdot\; (D^{\alpha} f)\; (\varphi(x_1),\; x_2,\; \cdots, x_n) \mid &\leq 2 \cdot \mid \varphi'(x_1) \mid^{\alpha_1} \;\cdot \mid (D^{\alpha} f)\; (\gamma(\varphi(x_1))) \mid \\
&\leq 2 \cdot \mid \varphi'(x_1) \mid \;\cdot \mid (D^{\alpha} f)\; (\gamma(\varphi(x_1))) \mid
\end{aligned}$$

since $\alpha_1 \geq 1$, and $|\varphi'| \leq 1$. Thus, we need to bound this right-hand side of the inequality. To do so, we compute

$$\frac{\partial}{\partial x_1}\left(\left(D^{\alpha'} f\right)(\gamma(\varphi(x_1)))\right) = \varphi'(x_1) \cdot (D^{\alpha} f)\,(\gamma(\varphi(x_1))) \tag{9.1}$$

$$+\, \varphi'(x_1) \cdot \sum_{j=2}^{n} \frac{\partial \gamma_j}{\partial t}\,(\varphi(x_1)) \;\cdot\; \left(D^{\alpha^{(j)}} f\right)\;(\gamma(\varphi(t))) \tag{9.2}$$

note that we ignore the γ_1' since it is just the identity map. We have that $a^{(j)} = \alpha' + (0,\; \ldots,\; 0,\; 1,\; 0,\; \ldots,\; 0)$ with 1 at the j-th coordinate. We claim that this is sufficient to control (prove the bound) as claimed.

Recall that Φ is a cellular r-parameterization of the map taking t to $(\gamma(t),\; D^{\alpha'} f(\gamma(t)))$, so substituting in φs, we get bounded derivatives. That is, the left-hand side of Equation 9.1 is bounded by Φ a cellular r-parameterization of $\left(D^{\alpha'} f\right)(\gamma(t))$. Similarly, we have that each partial derivative in the sum in Equation 9.2 is also bounded, again by Φ a cellular r-parameterization of γ – and the other incidence of $\varphi'(x_1)$ in Equation 9.2 is bounded for the same reason. The final factor in Equation 9.2 is finally also bounded because $\alpha^{(j)} < \alpha$, and α was assumed to be minimal such that the derivative exceeded 1. Hence, putting this all together, we get that the right-hand side of Equation 9.1 *must be bounded*; to be clear, that is

$$\varphi'(x_1) \cdot (D^{\alpha} f)\,(\gamma(\varphi(x_1)))$$

is bounded.

Now what to do with all this? Well, we have that (skipping right past the linear substitution step), that

$$|\ \varphi'(x_1) \cdot (D^{\alpha} f)\,(\gamma(\varphi(x_1)))\ | \leq 1,$$

and finally we are finished with the proof. □

That is of course just a very funny joke we cruelly taunt the reader with – as if you'll remember correctly, we never actually went on to prove that last lemma, whose proof we left for some poor sob down the road. It's perhaps just unfortunate then that that poor sob is, in fact, ourselves. Before we prove that lemma directly, however, we first prove another lemma, in what is becoming an increasingly ridiculous chain of lemmata each feeding into one other, almost and perhaps with no end in sight. But that's mathematics for you.

Lemma 9.5 *Suppose $f: M \times I \to I$ a family of functions mapping from the unit interval into itself a definable family of C^1 functions. Then, there is a $c > 0$ such that for any $a \in M$, and $B > 0$, the set*

$$T = \left\{\, t \in I : \ \left| f_a'(t) > B \right| \right\}$$

has $\mu(T) < \dfrac{c}{B}$, for μ the sum of lengths of intervals function.

Proof (of Lemma 9.5) Consider the two sets whose union compose that in the original statement,

$$\begin{aligned} T_+ &= \left\{\, t \in I : \ f_a'(t) > B \,\right\} \\ T_- &= \left\{\, t \in I : \ f_a'(t) < -B \,\right\}. \end{aligned}$$

These are both given by finite unions of open intervals, and that number of intervals, by o-minimal, is bounded uniformly in a and B. So, each of these intervals has length not exceeding $\frac{1}{B}$ given we have that $f_a: I \to I$ (by mean value theorem as before). Of course then, the result almost immediately follows, because we may just take our constant to be the number of intervals in these two sets. □

We *now* use this last lemma to prove the lemma we took for granted during the more arduous proof of parameterization. Because it's been a bit, we recall the theorem (mostly in full) before providing a proof.

Lemma 9.6 (Really Lemma 9.4) *Suppose $f: I^n \to I$ is definable, and C^1, and suppose as well that*

$$\left| \frac{\partial f}{\partial x_j}(x) \right| \leq 1$$

for each $x \in I^n$, and $j = 2, \ldots, n$. Then, the set of $a \in I$ such that $\dfrac{\partial f}{\partial x_1}(a, \ldots)$ is unbounded is finite

Proof (of Lemma 9.4) Supposing otherwise, the this set must contain an interval, and we can assume that interval to be I – that is,

$$\frac{\partial f}{\partial x_1}(a, \ldots)$$

is unbounded for all $a \in I$. By choice and smooth monotonicity, we can assume that we get a C^1 $\gamma\colon (0, \infty) \times I \to I^{n-1}$ such that for all $B > 0$ and $t \in I$, we have

$$\left| \frac{\partial f}{\partial x_j}(t, \gamma_B(t)) \right| > B,$$

and is definable. Applying the previous lemma (Lemma 9.5) to the coordinates of γ_B and to the family $f(t, \gamma_B(t))$, there is a constant, C, such that outside a set of measure $\leq \frac{C}{B}$, we have

$$\left| \frac{d}{dt} f(t, \gamma_B(t)) \right| \leq \frac{B}{3}$$

and coordinate maximum,

$$\max_{j=1, \ldots, n-1} \left| \gamma'_{B,\, j}(t) \right| \leq \frac{B}{3n}.$$

Outside of this set (that is, where we have these bounds), we have, by computation, that

$$\frac{d}{dt} f(t, \gamma_B(t)) = \frac{\partial f}{\partial x_1}(t, \gamma_B(t)) + \sum_{j=2}^{n} \frac{\partial f}{\partial x_j}(t, \gamma_B(t)) \cdot \gamma'_{B,\, j-1}(t),$$

which is claimed to be less than $\dfrac{B}{3}$. In the sum, we have that the $\gamma'_{B,\, j-1}(t)$ are each less than $\dfrac{B}{3n}$, and the derivatives in that same product are each less than 1 – making the entire sum at most $\dfrac{B}{3}$ – but the first summand is greater than B regardless, and so we have a contradiction. That is, as soon as this set of small measure isn't everything, this contradiction occurs. Thus, once $B > C$, we run into trouble. □

Now with that we truly are done (this section)! Depending on your particular interests, this rather technical proof may have seemed either garish or quite snazzy – but either way, we are now left with a powerful result, and the first of the major two we will be using to prove the Pila-Wilkie theorem. If this sort of technicality is not exactly your wheelhouse – well then perhaps this subject area isn't for you in general – but you might be a bit disappointed to hear that the next section isn't much less so; however, is it quite different, and it would not be inconceivable that those not so fond of this material find themselves quite enjoying what's coming up: the *number theory bit*. Depending on the sort of number theory you study/have studied, some of how we go about things may strike you as a bit odd approach to number theoretic proofs, but only time will tell. Much like we went on to refer to the *o-minimal bit* as **parameterization**, we will similarly start calling this number theoretic bit the **diophantine part** – which we note that we *do not* capitalize simply to follow the convention of the lectures.

Chapter 10
The Number Theory Bit (Diophantine Part)

Abstract As we near the end (though not too closely, mind) of our journey through a proof of Pila-Wilkie, we find ourselves now at the part that the naïve amongst us may have expected to come much sooner – and that's the number theoretic part. As mentioned in just the last chapter, this is also called the *diophantine part*, the treatment of which we will be following from Habegger [**?**]. As introduced there (though not the inception of the concept), this didn't investigate the rational points on definable curves as we have been throughout, but rather points that are *very close* to them. **To be clear, this is not what we are going to be doing** – but using several of the ideas from the proof in that paper, we get to 'skip' our way along in the proof of Pila-Wilkie to looking at points of bounded degree, rather than just rational points.

> Please again note that 'diophantine' is purposefully not capitalized as a stylistic choice to follow the material as presented in the course, and there is nothing more to be read into it than that.

We start with a proof sketch, before getting into all the minutiae proper, and broadly describe how we're going to be proceeding throughout this chapter.

Now, as we are investigating points of bounded height, we naturally need a height function (just as we did for the rationals, as discussed earlier). We define one as follows:

$$Suppose q \in \overline{\mathbb{Q}}$$

Chapter 11
Sticking it all Together

Abstract Sticking it all Together