MAX-PLANCK-INSTITUT
FÜR
INFORMATIK

Quantifier Elimination in p–adic Fields

D. Dubhashi

MPI–I–93–155 November 1993

INFORMATIK

Im Stadtwald
66123 Saarbrücken
Germany

Quantifier Elimination in p–adic Fields

D. Dubhashi

MPI–I–93–155 November 1993

Quantifier Elimination in p-adic Fields

Devdatt P. Dubhashi
Max-Planck-Institut für Informatik
Im Stadtwald
W6600 Saarbrücken

November 23, 1993

Abstract

We present a tutorial survey of quantifier-elimination and decision procedures in p-adic fields. The p-adic fields are studied in the (so-called) P_n–formalism of Angus Macintyre, for which motivation is provided through a rich body of analogies with real-closed fields. Quantifier-elimination and decision procedures are described proceeding via a Cylindrical Algebraic Decomposition of affine p-adic space. Effective complexity analyses are also provided.

1 Introduction

The field of p-adic numbers is obtained from the rationals by a completion process completely analogous to that by which the reals are constructed from the rationals. The p-adic numbers occupy a central position in Algebraic Number Theory, [7, 6, 20]. From a computational standpoint however, the p-adics have been much less studied when compared to the extensive work on the reals. In this article, we present a survey of the p-adics, focussing especially on algorithms for quantifier elimination over the p-adics.

This survey is organised as follows: in § 2, we outline the construction and basic properties of the field of p-adic numbers. In § 3, we give a thumbnail sketch of the history of work in quantifier elimination over the p-adics. § 4 gives a introductory guided tour of the so-called P_n–formalism of Macintyre. It is in this formalism that the correspondences between the reals and the p-adics stand out most clearly, and we offer a sampler of these in § 4.4. In § 5, we describe effective and quantitative quantifier elimination and decision procedures for the p-adics. We also give complexity analyses and matching lower bounds.

2 The p-adic Numbers

2.1 Valuations and Completions

A central concept in Field Theory [25] and in Algebraic Number Theory, [20] is that of a *valuation* of a field. A standard reference for valuation theory is [16].

Definition .1 [Valuation] A **valuation** of a field k is a map, $v : k \to \mathsf{R} \cup \{\infty\}$ such that

1. $v(x) = \infty$ iff $x = 0$.
2. $v(x + y) \geq \min(v(x), v(y))$ (**Ultrametric Inequality**).

□

The field of rational numbers, Q admits the so-called "p-adic valuation", v_p:

Definition .2 [p-adic Valuation on Q] Let $0 \neq x \in \mathsf{Q}$. The **p-adic valuation** of x, $v_p(x)$ is the unique integer n such that

$$x := p^n \cdot \frac{r}{s}$$

where p does not divide r, s. (Set $v_p(0) := \infty$.) □

From the p-adic valuation v_p, we obtain the **p-adic metric on Q**, $|\cdot|_p$:

Definition .3 [p-adic Metric on Q] For $x, y \in \mathsf{Q}$, set:

$$|x - y|_p := p^{-v_p(x-y)}.$$

(By convention, $p^{-\infty} := 0$.) □

The *p-adic norm* of $x \in \mathsf{Q}$ is thus $|x|_p := p^{-v_p(x)}$. Intuitively, the p-adic norm measures how divisible x is by p – the higher the divisibility, the smaller the norm.

The p-adic norm has many properties quite different from the usual Euclidean norm. They all stem from the fact that the p-adic metric satisfies not merely the triangle inequality but also the curious **ultrametric inequality** (compare with Definition .1):

$$|x+y|_p \leq \max(|x|_p, |y|_p).$$

The p-adic metric is therefore distinguished from the usual absolute value by referring to it as a *non-archimedean* metric. By a famous theorem of Ostrowski, [16], the absolute value together with the p-adic metrics $|\cdot|_p$ for each prime p, constitute all the possible metrics on Q (upto equivalence of metrics).

Given a metric space, one can pass to the *completion* via the standard Cauchy construction. For instance, the field of real numbers, R is obtained from the rationals by forming the completion with respect to the (usual) absolute value metric. When we perform the exactly analogous process with the p-adic metric $|\cdot|_p$, we obtain the field of *p-adic numbers*, Q_p. The metric $|\cdot|_p$ on Q extends to a metric on the completion, Q_p, and we will also use the notation $|\cdot|_p$ for the extended metric. By the theorem of Ostrowski, R and the fields Q_p for each prime p, together comprise all the completions of Q.

Henceforth we will fix a particular prime p, and omit the subscript on valuations and metrics.

2.2 Some properties of Q_p

A useful alternative description of the p-adic numbers is obtained via an *inverse limit* construction from finite residue rings. The subring $Z_p \subseteq Q_p$defined by

$$Z_p := \{a \in Q_p : v(a) \geq 0\}$$

is called the ring of *p-adic integers*.

Proposition .4 (p-adics via Inverse Limits) *1.* $Z_p = \lim_{\leftarrow} Z/p^n Z$

2. There are canonical projection homomorphisms, $res_k : Z_p \rightarrow Z/p^k Z, k \geq 1$.

3. Q_p *is the field of fractions of* Z_p.

As a consequence, we get a nice representation of p-adics which resembles the usual decimal representation of reals and is the source of formal similarities to *Laurent* series :

Proposition .5 (p-adic Expansions) *Every non-zero p-adic $a \in \mathbb{Q}_p$, can be expressed uniquely in the form*

$$\sum_{n \geq N} a_n \cdot p^n$$

where $0 \leq a_n < p$ are integers and $a_N \neq 0$. The integer N is determined by the condition $v_p(a) = -N$. (So every $a \in \mathbb{Z}_p$ can be expressed uniquely in the form $\sum_{n \geq 0} a_n \cdot p^n$.)

A curious consequence of the ultrametric inequality that is crucially responsible for many properties of the p-adics is the following

Proposition .6 (Ultrametric Equality)

1. *If $v(x) < v(y)$, then $v(x+y) = v(x)$.*
2. *More generally, if $v(x_k) < v(x_i)$ for some k and all i, $1 \leq i, k \leq n$, then $v(\sum_{1 \leq i \leq n} x_i) = v(x_k)$.*

The key technical tool for handling the p-adics is the following algebraic-topological criterion for roots of polynomials in $\mathbb{Z}_p[x]$. It relates roots in the finite residue rings to roots in the inverse limit.

Lemma .7 (Hensel's Lemma)

1. *Let $f \in \mathbb{Z}[x]$ and let a_0 be a root of f in $\mathbb{Z}/p\mathbb{Z}$. If $f'(a_0) \neq 0$ in $\mathbb{Z}/p\mathbb{Z}$, then a_0 can be lifted to a unique root $\xi \in \mathbb{Z}_p$ of f such that $res_1(\xi) = a_0$.*
2. *More generally, if α is a root of f in a finite ring $\mathbb{Z}/p^n\mathbb{Z}$ and if $n > 2r$ where $r = v(f'(\alpha))$, then α can be lifted to a unique root $\xi \in \mathbb{Z}_p$ of f such that $res_{n-r}(\xi) = res_{n-r}(\alpha)$.*

3 History of Quantifier Elimination over the p-adics

In an award winning series of papers, James Ax and Simon Kochen, [1, 2], and independently, Ju.L. Ersöv, [17] showed that the first-order theory of p-adic fields was decidable. Their proofs employed the model theoretic technique of *ultra-products*, and hence, via Gödel's Completeness Theorem, yielded only the *existence* of general-recursive procedures. Using ingenious, but elementary arguments, Paul J. Cohen[1] gave the first effective decision procedure for the p-adics, [9]. His proof consisted of a procedure to eliminate field quantifiers in favour of quantifiers ranging over a finite family of finite residue rings. Many of the key ideas in subsequent work can be traced back to this paper[2]. Subsequently, Volker Weispfenning refined and extended Cohen's work in a series of papers, [33, 34].

All the previous work took place in the setting of valued fields. The p-adics were studied in a two-sorted language, one for field elements and one for elements of the *value group* [3]. This was a somewhat unnatural formulation requiring, to yield the quantifier-elimination property, the presence of an awkward "cross-section" predicate to navigate between the two sorts.

In an insightful paper [23], Angus Macintyre introduced the so-called P_n–formalism. This made it possible to study the p-adics in a smooth way employing a one-sorted language. Furthermore, the new formalism brought to light some very deep parallels between the p-adics and the reals (see § 4.4 below). Macintyre went on to demonstrate model-theoretically, the existence of quantifier-elimination in the P_n–formalism. An explicit quantifier-elimination procedure in this formalism was given by [34]. Subsequently, Denef [12, 13] obtained an explicit algebraic cell decomposition of p-adic affine space, and gave another proof of quantifier elimination [4]. An algorithmic version of Denef's cell decomposition and a complexity-theoretic analysis appears in the author's Ph.D. dissertation, [15]. In § 5, we will outline the main ideas behind these results.

4 The P_n-formalism

4.1 Motivation

What is the analogue in the p-adic case to the order structure on the reals? This is the key question to be answered if we wished to keep the develop-

[1] Of the *Continuum Hypothesis* fame!

[2] However, this is rather hard reading, due partly to idiosyncratic notations. For instance Macintyre, [24] says " ...it is not clear what Cohen is trying to prove..."!

[3] Given the valuation v over a field k, $v(k) := \{v(x) : x \in k\}$ is the value-group.

[4] Also [12] contains a beautiful application of the decomposition to a problem in semi-algebraic geometry over the p-adics, namely to count p-adic points on a rational curve.

ment of the p-adic theory similar to that of the reals. Initial attempts were directed to making the valuation structure the p-adic counterpart to order: the relation $x \geq 0$ in the reals was sought to be paralleled by $v(x) \geq 0$ in the p-adics. For various reasons, this is not entirely satisfactory, see [24].

A more natural replacement for the order relation comes from Artin's theory of real-closed fields developed in order to solve *Hilbert's 17th Problem*, [21]. This asks whether a polynomial $f \in \mathsf{R}[x_1, \cdots, x_n]$ which is *positive definite*[5] can be represented as the sum of squares of *rational functions* over the reals, that is, as the sum of squares of elements in $\mathsf{R}(x_1, \cdots, x_n)$. In his celebrated affirmative solution, [22], Artin brought to the forefront, the crucial role of the subgroup, $\mathsf{R}^{*(2)}$ of the multiplicative group, R^*, consisting of those elements that are *squares*. Seen in this way, the signs of real numbers have the following interpretation independent of the order structure: $[\mathsf{R}^* : \mathsf{R}^{*(2)}] = 2$, we have coset representatives $+1, -1$ and the natural homomorphism $sgn : \mathsf{R}^* \rightarrow \mathsf{R}^*/\mathsf{R}^{*(2)}$ giving the coset representative is the familiar and all-important *sign* homomorphism.

One can now hope to carry out precisely the same programme in the p-adic case. For certain technical reasons though, see [24], we are forced to consider *all* n–th powers in the p-adic case, not merely the squares.

4.2 Structure of n–th powers in Q_p

Let $\mathsf{Q}_p^{*(n)}$ denote the subgroup of the multiplicative group Q_p^* consisting of n–th powers. A well known structure theorem [29, 10] then yields:

Proposition .8 (Structure of n-th powers in Q_p) *1.* $\mathsf{Q}_p^* \approx \mathsf{Z} \times \mathsf{F}_p^* \times \mathsf{Z}_p$.

2. $\mathsf{Q}_p^{*(n)} \approx n\mathsf{Z} \times \mathsf{F}_p^{*(n)} \times n\mathsf{Z}_p$.

3. $\mathsf{Q}_p^*/\mathsf{Q}_p^{*(n)} \approx \mathsf{Z}/n\mathsf{Z} \times \mathsf{F}_p^*/\mathsf{F}_p^{*(n)} \times 1 + p\mathsf{Z}/p^{v(n)}\mathsf{Z}$.

Thus $\mathsf{Q}_p^{*(n)}$ has finite index in Q_p^* (which can in fact be explicitly computed, see [15]). We will denote the canonical projection homomorphism by $\rho_n : \mathsf{Q}_p^* \rightarrow \mathsf{Q}_p^*/\mathsf{Q}_p^{*(n)}$. For each $x \in \mathsf{Q}_p^*$, $\rho_n(x)$ gives the canonical coset representative for x just as, in the real case, $sgn(x)$, for $x \in \mathsf{R}^*$, gives the canonical coset representative (with respect to the subgroup of sqaures).

[5] That is, $f(a_1, \cdots, a_n) \geq 0$ for all $(a_1, \cdots, a_n) \in \mathsf{R}^n$

4.3 Valuation and n-th powers

Macintyre, [23], first suggested that the p-adics be studied in the "P_n-formalism" to pursue the analogy to the reals, with the coset representatives taking the place of sign conditions. In this formalism, there are unary predicates, P_n for each $n \geq 2$ standing for the n-th powers, so

$$P_n(x) \leftrightarrow \exists y(y^n = x)$$

(Here x and y refer to field elements.)

How does this new formalism relate to the underlying valuation structure? An outstanding merit of the P_n–formalism is that the valuation structure is *definable* in it:

Proposition .9 (Definability of valuation) *1. If $p \neq 2$, then for any $a, b \in k$,*

$$v(a) \leq v(b) \leftrightarrow P_2(a^2 + p \cdot b^2)$$

2. If $p = 2$, then for any $a, b \in k$,

$$v(a) \leq v(b) \leftrightarrow P_3(a^3 + p \cdot b^3)$$

Conversely, to see what kind of sets the P_n predicates define in the valuation topology, we need a key lemma apparently due to Robinson, [28] which asserts that two elements which are sufficiently close together in the (p-adic) metric topology are in the same coset of n–th powers:

Lemma .10 *If $v(y - x) > 2(v(x) + v(n))$ for $x, y \in k$, then $\rho_n(x) = \rho_n(y)$.*

Corollary .11 *There is a (integer) function $\lambda(n)$, such that if $v(x) > \lambda(n)$, then $P_n(1 + x)$.*

This corollary actually enables one to weaken somewhat the assumptions of the original lemma and should be compared with the Ultrametric equality, Proposition .6.

Proposition .12 *We have*

1. *If $v(y - x) > v(x) + \lambda(n)$, then $\rho_n(x) = \rho_n(y)$.*
2. *If $v(x_k) + \lambda(n) < v(x_i)$ for some k and all $i \neq k$, $1 \leq i, k \leq m$, then*

$$\rho_n\left(\sum_{1 \leq i \leq m} x_i\right) = \rho(x_k)$$

Finally, we can say how sets defined by the P_n predicates cohere with the valuation topology.

Proposition .13 *For each $n \geq 2$, the sets*

$$\{x \in k^* : P_n(x)\}$$

are closed and open in the valuation topology.

4.4 The R-Q_p Analogy

A deep and attractive parallel exists between the reals and the p-adics when the latter are treated in the P_n formalism[6]. In this subsection, we briefly sketch some of these similarities. See [24] for an extensive comparison.

Below, we list many assertions in two parts – one labelled R applicable to the reals, and the other labelled Q_p applicable to the p-adics. In this way, the correspondence will be transparent.

4.4.1 Signs and Cosets

Evidence that the coset representatives of n–th powers play a role for the p-adics analogous to that of signs in the reals is presented in the next two propositions from [15].

Proposition .14 (Signs and cosets) R *Let $x \in R, f(x) \neq 0$. Then in all sufficiently small neighborhoods around x, f takes the same sign as $f(x)$.*

Q_p *Let $x \in Q_p, f(x) \neq 0$. Then in all sufficiently small neighborhoods around x, f takes values in the same coset as $f(x)$.*

What about neighborhoods of a zero ? For differentiable f, the values of f in a neighborhood of a zero lie in a coset determined by that of the first non-zero derivative and that of the arbitrarily small increment :

Proposition .15 (Signs and cosets, cont'd) *Let f be differentiable n times and suppose $f^n(\xi) \neq 0$ while $f^i(\xi) = 0$ for all $i < n$.(In particular ξ is a zero of f.)*

[6] Note that via the relation $x \geq 0 \leftrightarrow P_2(x)$, the reals also can be presented in the P_n formalism (in fact only P_2 suffices).

R *For all sufficiently small ϵ, we have*

$$sgn(f(\xi+\epsilon)) = (sgn(\epsilon)^n) \cdot sgn(f^n(\xi))$$

In particular, if $f'(\xi) \neq 0$, then

$$sgn(f(\xi+\epsilon)) = sgn(\epsilon) \cdot sgn(f'(\xi))$$

Q_p *Let $m \geq 2$. For all ϵ such that $|\epsilon| < p^{-2(v(f^n(\xi))+v(m))}$, we have*

$$\rho_m(f(\xi+\epsilon)) = (\rho_m(\epsilon))^n \cdot \rho_m(f^n(\xi))$$

In particular if $f'(\xi) \neq 0$, then

$$\rho_m(f(\xi+\epsilon)) = \rho_m(\epsilon) \cdot \rho_m(f'(\xi))$$

4.4.2 Algebra

A real-closed field is a field *elementarily equivalent* to R; namely a field which satisfies exactly the same sentences of the language of ordered fields (or equivalently, the language of rings augmented with the P_2 predicate), as R, see for example [14]. Correspondingly, a **p-adically-closed field** is one which is elementarily equivalent to Q_p in the language of rings augmented with all the predicates P_n, for $n \geq 2$, see [27].

There is in fact an intrinsic characterization of real-closed and p-adically closed fields given by certain canonical *completeness schemas.*

Proposition .16 (Real-closed fields and p-adically-closed fields) *1. A real field is real-closed iff (a) every odd degree polynomial over the field has a root in the field and (b) every element or its negative is a square in the field.*

2. A p-adic field is p-adically closed iff (a) Hensel's lemma holds, (see § 2.2) and (b) for each $n \geq 2$, there exist a set of integers, $b_1, \cdots, b_{k(n)}$ such that for each x in the field, $P_n(b_i \cdot x)$ for some i.

Proposition .17 (Closures) *1.* (**Artin-Schreier**) *Every real field has a real-closure which is unique upto isomorphism.*

2. (**Prestel-Rocquette, Robinson, Belair**) *Every p-adic field has a p-adic closure which is unique upto isomorphism.*

4.4.3 Model Theory

There are striking parallels in the model theory of the reals and the p-adics in the P_n–formalism. For definitions and concepts from Model Theory, we refer to the classic work of Chang and Kiesler, [8] or to the chapter by Kiesler in [3].

Proposition .18 (Model-Completeness) *1.* R *The theory of real-closed fields is model-complete. That is, let $K \subseteq L$ be real-closed. Then a first-order sentence in the language of ordered fields with parameters from K holds in L iff it holds in K. The same is true in the language of fields augmented with the P_2-predicate.*

2. Q_p *The theory of p-adically closed fields is model-complete. That is, let $K \subseteq L$ be p-adically closed. Then a first-order sentence in the language of valued fields with parameters from K is true in L iff it is true in L. The same holds in the language of fields augmented with all the predicates $P_n, n \geq 2$.*

A consequence of this result are the following *transfer* principles:

Proposition .19 (Transfer Principles) *1. (***Tarski's transfer principle for** R*) A first-order sentence is true in all real-closed fields iff it is true in* R.

*2. (***Ax-Kochen-Ersov transfer principle for** Q_p*) A first-order sentence is true in all p-adically closed fields iff it is true in* Q_p.

This can also be deduced from the following important property of the two theories:

Theorem .20 (Quantifier-Elimination) *1.* R *(Tarski, [31], also [11, 5]) The theory of real-closed fields admits elimination of quantifiers in the language of ordered fields, or in the language of field theory augmented with the predicate P_2.*

2. Q_p *(Macintyre, [23], also [13]) The theory of p-adically closed fields admits elimination of quantifiers in the language of fields augmented with all the predicates $P_n, n \geq 2$. We do* not *get elimination of quantifiers in the pure language of valued fields.*

The last statement is a notable advantage of the P_n–formalism over the traditional valuation formalism. We will describe effective and quantitative versions of this theorem in § 5.

A nice converse to the above was obtained by Macintyre, McKenna and van den Dries, [26].

Proposition .21 *In the P_n-formalism:*

1. R *An ordered field that admits elimination of quantifiers is real-closed.*
2. Q_p *A "Belair-Robinson" p-adic field ([24]) that admits elimination of quantifiers is p-adically closed.*

4.4.4 Semi-algebraic Geometry

A *semi-algebraic* set in R^n is one that is specified by a set of polynomial equations and inequalities. Thus a semi-algebraic set in R^n has the form:

$$\{\mathbf{x} \in R^n : f(\mathbf{x}) = 0, g_1(\mathbf{x}) \sim 0, \cdots, g_k(\mathbf{x}) \sim 0\}$$

where $f, g_1, \cdots, g_k \in R[x]$ and $\sim$ is one of the two *order* relations, $<, >$. *Real semi-algebraic geometry* comprises the study of these semi-algebraic sets and their morphisms.

The p-adic counterpart to this is a **p-adic semi-algebraic set,** which is one of the form

$$\{\mathbf{x} \in Q_p^n : P_n(f_1(\mathbf{x})) \wedge \cdots \wedge P_n(f_k(\mathbf{x}))\}$$

In Real semi-algebraic geometry, one has a strengthened form of the quantifier-elimination theorem called the "finiteness theorem", [32]: a definable *open* subset of R^n can be defined using only *positive* boolean operations[7] and *strict* polynomial inequalities. Introduce the notation, as in [28], $R_n(x) \leftrightarrow x \neq 0 \wedge P_n(x)$. One can of course express the condition $f(x_1, \cdots, x_n) > 0$ by $R_2(f(x_1, \cdots, x_n))$ in the reals. In [28], an exact p-adic analogue is obtained:

Theorem .22 ("Finiteness" Theorem) *($k :=$ R or Q_p) Any definable open subset of k^n can be defined using only $\vee, \wedge$ and the $R_n, n \geq 2$.*

One also obtains a sensible notion of dimension which has the following properties, [30]

[7] i.e. no negations.

Proposition .23 (Dimension) *($k :=$ R or Q_p)*

1. $\dim(k^m) = m$.

2. *If X, Y are definable subsets of k^m, then*

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

and $\dim(X) \leq \dim(Y)$ *if* $X \subseteq Y$.

3. *If X is definable, then* $\dim(X)$ *is equal to the algebro-geometric dimension of the Zariski closure of X.*

5 Quantifier Elimination and Decision Procedures via Algebraic Cell Decomposition

The basic intent behind a Cell Decomposition is to partition affine space into "cells" in each of which a given set of polynomials "behaves well". In the well-known Cylindrical Algebraic Decomposition algorithm over the reals, for instance [11], real affine space is partitioned into cells delineated by polynomial inequalities such that in each cell a given set of polynomials maintains constant sign. The natural analogue over the p-adics would be to partition p-adic affine space into cells in each of which a given set of polynomials maintains fixed coset representatives of certain nth powers. In this section, we describe and refine a Cell Decomposition lemma due to Denef, [12, 13], that meets these requirements. In order to obtain this decomposition, an auxiliary decomposition is needed that partitions p-adic affine space into cells in each of which given polynomials are well behaved with respect to their valuations, in a sense which is made precise below.

First we need some preliminary definitions. Recall the following definition motivated in § 4.4.4.

Definition .24 A subset of $k^m, m \geq 1$ is *semi-algebraic* if it is a boolean combination of subsets of the form

$$\{\mathbf{x} \in k^m : P_n(f(\mathbf{x}))\}$$

where $f \in k[x_1, \ldots, x_m]$ and $n \geq 2$. □

We extend this notion to functions as follows, [13]:

Definition .25 A function $f : k^m \to k, m \geq 1$ is semi-algebraic if for every semi-algebraic subset $S \subseteq k \times k^r, r \geq 1$, the set

$$\{(x, y) \in k^{m+r} : (f(x), y) \in S\}$$

is semi-algebraic. □

Remarks:

1. The graph of a semi-algebraic function is semi-algebraic.
2. A polynomial is a semi-algebraic function.
3. The class of semi-algebraic functions is closed under composition, addition and multiplication.

Sets defined by polynomial equalities and inequalities between valuations of polynomials are semi-algebraic. In [13], it is proved that sets defined by congruences are also semi-algebraic.

Now we can define the p-adic analog of a cell:

Definition .26 [p-adic cell] A *cell* in $k^m \times k$ is a set of the form

$$\{(\mathbf{x}, t) : \mathbf{x} \in C \quad \wedge \quad v(a_1(\mathbf{x}))\square_1 v(t - c(\mathbf{x}))\square_2 v(a_2(\mathbf{x}))\}$$

where $C \subseteq k^m$ is semi-algebraic, a_1, a_2 and c are semi-algebraic functions and $\square_1, \square_2$ denotes either $\leq, <$ or no condition. The cell is said to have *center* $c(\mathbf{x})$. □

5.1 Valuation Decomposition

The next proposition gives a decomposition of p-adic affine space into cells such that in each cell, the valuation of a given polynomial is bounded near the valuation of an individual term. The idea of this decomposition goes back to Cohen, [9] and the key tools used are the Ultrametric Equality and Hensel's Lemma. We use the construction as in Denef, [13], but supplement it with quantitative information which bounds its size.

Proposition .27 (Cohen, Denef, [9, 13], Algebraic Cell Decomposition for valuation *Let $f(\mathbf{x}, t)$ be a polynomial in t with coefficients which are semi-algebraic functions of $\mathbf{x} \in k^m$. Then, there exists a finite partition of $k^m \times k$ into* *cells A such that each cell has associated with it a center, $c(x)$ (where $c(x)$ is semi-algebraic) and a bound $e \in \mathbb{N}$ such that if we write*

$$f(\mathbf{x},t) := \sum_{i \geq 0} a_i(\mathbf{x})(t - c(\mathbf{x}))^i$$

then

$$v(f(\mathbf{x},t)) \leq \min_i v(a_i(\mathbf{x})(t - c(\mathbf{x}))^i + e$$

In particular, let $f \in \mathbb{Z}[x_1,\ldots,x_n]$ and suppose the degrees of f in $x_1 \ldots x_n$ are bounded by $d_1,\ldots,d_n$ and that the coefficients are bounded in size by L. (Denote $\sum_i d_i$ by D.) Then there exists a partition of size $O(2^{c \cdot 2^{D+L+n}})$ and the constant $e = O(2^{D+L+n})$ for each such cell.

5.2 Decomposition for n-th powers

The next proposition is a p-adic analogue of Cylindrical Algebraic Decomposition for the reals [8]. The key idea is to use Proposition .12 to convert the valuation decomposition into a decomposition for nth powers.

Proposition .28 (Denef, [13], Algebraic Cell Decomposition Lemma for nth powers *Let $f_i(\mathbf{x},t), 1 \leq i \leq r$ be polynomials in t with coefficients which are semi-algebraic functions of $\mathbf{x} \in k^m$, and let $n > 1$ be fixed. Then there exists a finite partition of $k^m \times k$ into cells A, such that each cell has a center $c(\mathbf{x})$ (which is semi-algebraic), such that for all $(\mathbf{x},t) \in A$, we have,*

$$f_i(\mathbf{x},t) = u_i(\mathbf{x},t)^n \cdot h_i(\mathbf{x}) \cdot (t - c(\mathbf{x}))^{\nu_i}, \quad 1 \leq i \leq r$$

with $v(u_i(\mathbf{x},t)) = 0$, $h_i(\mathbf{x})$ a semi-algebraic function of $\mathbf{x}$ and $\nu_i \in \mathbb{N}$ for each $1 \leq i \leq r$.

In particular, if applied to r polynomials in $\mathbb{Z}[x_1,\cdots,x_m]$ with degrees bounded by $d_1,\cdots,d_m$ (with $D := d_1 + \cdots + d_m$) and coefficient size bounded by L, it yields a decomposition into at most $O(2^{c \cdot r \cdot L \cdot 2^{D+n}})$ such cells for some constant c.

5.3 A Decision Procedure

In this subsection, we use the quantitative version of the Cell Decomposition lemma, Proposition .28 to give a decision procedure for the full theory,

[8] As stated, it is not a *cylindrical* decomposition as usually understood, but can be made such, [30]

$\mathrm{Th}(\mathbf{Q}_p, +, \times, 0, 1, \{P_n\}_{n\geq 2})$ in the form of an alternating Turing machine algorithm running in exponential time. This also yields a deterministic decision procedure running in exponential space or in double exponential time.

We describe an alternating Turing machine algorithm to decide sentences of the theory of p-adically-closed fields. At any point in the computation, a processor is attempting to verify a statement of the form

$$\begin{aligned}\mathbf{Q}_p \models (Qx_1)\cdots(Qx_k)&\\ &\varphi(f_1(x_1,\cdots,x_k),\cdots,f_r(x_1,\cdots,x_k))\end{aligned}$$

where φ is a boolean combination of sentences of the form $P_n(f_i(x_1,\cdots,x_k))$. (Assume, without loss of generality, that the last quantifier is $\exists$.) By computing the Cell Decomposition (Nth powers version) for the polynomials involved, this amounts to verifying several condition of the form

$$\begin{aligned}&\mathbf{Q}_p \models\\ &\quad (Qx_1)\cdots(\exists x_k)\\ &\quad\quad \varphi'(g_1(x_1,\cdots,x_{k-1}),\cdots,g_s(x_1,\cdots,x_{k-1}))\\ &\quad\quad \wedge \quad v(a_1(x_1,\cdots,x_{k-1}))\square_1\\ &\quad\quad\quad\quad v(t-c(x_1,\cdots,x_{k-1}))\square_2\\ &\quad\quad\quad\quad v(a_2(x_1,\cdots,x_{k-1}))\\ &\quad \wedge \quad P_n(\rho\cdot(t-c(x_1,\cdots,x_{k-1})))\end{aligned}$$

The processor activates several child processors, one for each cell, each attempting to verify such a condition. If the quantifier is $\exists$, these are generated using $\vee$-branching, if it is $\forall$, $\wedge$-branching is used. As in [13], or otherwise, one can eliminate the x_k variable at this stage and so each child processor is reduced to verifying a condition of the form

$$\mathbf{Q}_p \models (Qx_1)\cdots(Qx_{k-1}) \quad \varphi(g_1(\mathbf{x}),\cdots g_s(\mathbf{x}))$$

(where $\mathbf{x} := (x_1,\cdots,x_k)$).

Finally a (super-exponential) number of processors, each attempting to verify a quantifier-free statement can all use the criterion in Proposition .8.

To analyze the complexity of the decision procedure, suppose we start with the sentence

$$\begin{aligned}\mathbf{Q}_p \models (Qx_1)\cdots(Qx_m)&\\ &\varphi(f_1(x_1,\cdots,x_m),\cdots,f_r(x_1,\cdots,x_m))\end{aligned}$$

Applying Proposition .28 to the set of polynomials $f_1, \cdots, f_r$, and using the notation from there, we see that there are approximately $2^{c \cdot r \cdot L \cdot 2^{D+n}}$ cells that need to be generated. However, the alternating machine can generate them in time $O(r \cdot L \cdot 2^{D+n})$ by its parallel branching capability. So there will be a double exponential number of processes running concurrently to verify their respective conditions. At the bottom it is clear that the criterion of Proposition .8 can be used to verify the quantifier-free condition in exponential time. Thus the overall alternating algorithm runs in time $O(r \cdot L \cdot 2^{D+n})$ i.e. in double exponential time. Moreover, the machine clearly makes at most m alternations.

This analysis combined with results from the last section yield:

Theorem .29 (The Decision Problem for Q_p) *The decision problem for linear sentences over Q_p is complete for the Berman complexity class $\cup_\gamma \mathcal{STA}(*, 2^{\gamma n}, n)$. In particular it can be solved in $\mathcal{EXPSPACE}$ and in double exponential time.*

5.4 A Quantifier-Elimination Procedure

The decision problem of the last section can actually be converted into a deterministic quantifier-elimination algorithm in the straightforward way, eliminating one variable at a time by taking a disjunction over all the cells produced in the decomposition. To analyze the complexity of the algorithm, we merely note that the size of the constants involved in the formulas can at most increase by a constant, and that the number of polynomials produced to replace a given polynomial in a cell is at most four times the original. This yields

Theorem .30 (Quantifier-Elimination for Q_p) *There is a quantifier elimination procedure for the theory of linear sentences in Q_p. Given a formula F, this produces a quantifier-free formula F' equivalent to F in double exponential time. Moreover, $l(F') \leq 2^{2^{\gamma l(F)}}$ for some constant γ.*

5.5 Lower Bounds

Berman [4], showed that $\cup_c \mathcal{STA}(*, 2^{cn}, n)$ is polynomial-time reducible to the theory of R. It was observed by, *inter alia*, Weispfenning [35] and Fürer [18],that this reduction (and in fact the original one of [19]) makes no essential use of the order relation. In fact, it holds for any theory in the language $L_{G1} := (0, 1, +)$ of abelian groups, such that all its models are groups, and in one model, an element (for instance, 1) of infinite order exists. More precisely,

Proposition .31 (Fischer-Rabin, Berman) *Let G be a torsion-free abelian group with distinguished element $0 \neq 1$. G may carry additional structure for a language L extending L_{G1}. Then, the decision problem for first order sentences in the L-theory of G is hard for $\cup_c STA(*, 2^{cn}, n)$ under polynomial time reductions.*

Combining this with the result of § 5.3, we obtain

Corollary .32 *The decision problem for the theory of p-adically closed fields is complete for $\cup_c STA(*, 2^{cn}, n)$ under polynomial time reductions.*

For the lower bound on explicit quantifier elimination, we combine two ingredients. The first is a construction of Fischer and Rabin ([19], Th.8, Cor.9), slightly extended as in [35]:

Lemma .33 (Fischer-Rabin) *There is a positive constant c, and a sequence $\mu_n, n \geq 1$ of L_{G1}-formulas with one free variable x, such that*

1. *If G is an abelian group in which 1 is an element of infinite order, then*
$$\mu_n^G := \{a \in G : G \models \mu_n(a)\} = \{0, 1, \cdots, 2^{2^n} - 1\}$$
2. *If G is an abelian group with distinguished element $1 \neq 0$, then*
$$\mu_n^G \supseteq \{0, 1, \cdots, 2^{2^n} - 1\}$$
3. *$l(\mu_n) \leq c \cdot n$.*

The second ingredient is a general technical lemma due to Weispfenning, [35], relating the geometry of definable sets to their sizes. Using this, we obtain, [15, 35]:

Proposition .34 *Any quantifier-free formula σ_n equivalent to μ_n over Q_p has $l(\sigma_n) \geq 2^{2^n}$ for $n \geq 1$. Hence any quantifier elimination procedure for Q_p requires at least double exponential space.*

Acknowledgements

I would like to express my heartfelt thanks to my advisor Dexter C. Kozen for initiating my interest in the p-adics, and for his constant support and guidance throughout. Thanks also to Moss Sweedler and to the Cornell Computer Algebra Group ("Thursday Lunch Seminar" attendees) for patiently listening to various stages of this research. I would like to express my appreciation for the wonderful atmosphere of academic freedom at the Max-Planck-Institut für Informatik, for which credit must go largely to the Direktor, Kurt Mehlhorn. Thanks are due to the efforts of an anonymous referee for several important corrections and improvements to this paper. For financial and resource support, I would like to thank Cornell University, National Science Foundation Grants CCR 8806096 and CCR 8901061 and the Max-Planck-Institut für Informatik.

References

[1] J. Ax and S. Kochen. (1965) Diophantine problems over local fields, I,II. *Amer. J. Math.*, 87:605–648.

[2] J. Ax and S. Kochen. (1966) Diophantine problems over local fields III. *Ann. Math*, 83(2):437–456.

[3] J. Barwise, editor. (1977) *Handbook of Mathematical Logic.* North Holland, Amsterdam.

[4] L. Berman. (1980) The complexity of logical theories. *Theoret. Comp. Sci.*, 11:71–77.

[5] M. Ben-Or, D. Kozen, and J.H. Reif. (1986) The complexity of elementary algebra and geometry. *J. Comp. Sys. Sci.*, 32:251–264.

[6] J.W.S. Cassels. (1986) *Local fields*, volume 3 of *London Mathematical Society.* Cambridge University Press, London, New York, Melbourne, Sydney.

[7] J.W.S. Cassels and A. Frohlich, editors. (1967) *Algebraic Number Theory.* Thompson Book Company.

[8] C.C. Chang and J.H. Kiesler. (1973) *Model Theory.* North Holland, Amsterdam.

[9] P. J Cohen. (1969) Decision procedures for real and p-adic fields. *Comm. Pure Appl. Math.*, 22:131–151.

[10] P.M. Cohn. (1989) *Algebra*, volume 2. John Wiley.

[11] G.E. Collins. (1975) *Quantifier elimination for Real Closed Fields by Cylindrical Algebraic Decomposition*, volume 33 of *Lecture Notes in Computer Science*, pages 134–183. Springer-Verlag.

[12] J Denef. (1984) The rationality of the poincare series associated to the p-adic points on a variety. *Invent. Math.*, 77:1–23.

[13] J. Denef. (1986) p-adic semialgebraic sets and cell decomposition. *J. Reine Angewandte Math.*, 369:154–166.

[14] M.A. Dickmann. (1985) Applications of model theory to real algebraic geometry: A survey. In *Methods in Mathematical Logic Proceedings, 1983*, Lecture Notes in Mathematics, 1130, pages 76–150, Springer-Verlag.

[15] D.P. Dubhashi. (1992) *Algorithmic Investigations in p-adic Fields.* PhD thesis, Cornell University, Ithaca, N.Y., U.S.A., August 1992.

[16] O. Endler. (1972) *Valuation Theory.* Universitext. Springer-Verlag.

[17] Ju. L. Ersov. (1965) On elementary theories of local fields. *Algebra in Logika*, 4:5–30.

[18] M. Fürer. (1982) The complexity of presburger arithmetic with bounded quantifier alternation depth. *Theoret. Comp. Sci.*, 18:105–111.

[19] M.J. Fischer and M.O. Rabin. (1974) Super-exponential complexity of Presburger Arithmetic. *SIAM-AMS Proc.*, 7:27–41.

[20] H. Hasse. (1980) *Number Theory*, volume 229 of *Grundlehren der mathematischen Wissenschaften.* Springer-Verlag, Berlin-Heidelberg-New York, third edition, 1980. Corrected and enlarged translation of *Zahlentheorie*.

[21] I. Kaplansky. (1977) *Hilbert's Problems.* Department of Mathematics, Univ. of Chicago.

[22] T.Y. Lam. (1984) An introduction to real algebra. *Rocky Mountain J. Math.*, 14(4):767–814.

[23] A. Macintyre. (1976) On definable subsets of p-adic fields. *J. Symb. Logic*, 41:605–610.

[24] A. Macintyre. (1984) Twenty years of p-adic model theory. In J.B. Paris, Wilkie A. J., and Wilmers G. M., editors, *Logic Colloquim, '84*, pages 121–153, Amsterdam, 1984. Elsevier Science Publishers B.V. North Holland.

[25] P.J. McCarthy. (1976) *Algebraic Extensions of Fields.* Dover Publications, Inc.

[26] A. Macintyre, K. McKenna, and L. van den Dries. (1983) Elimination of quantifiers in algebraic structures. *Advances in Mathematics*, 47:74–87.

[27] A. Prestel and Rocquette P. (1984) *Lectures on formally p-adic fields*, volume 1050 of *Lecture Notes in Mathematics.* Springer-Verlag, Berlin-Heildelberg-New York.

[28] E. Robinson. (1987) The geometric theory of p-adic fields. *J. Algebra*, 110:158–172.

[29] J.P. Serre. (1973) *A Course in Arithmetic.* Graduate Texts in Mathematics. Springer-Verlag, New York, Heidelberg, Berlin.

[30] P. Scowcroft and L. van den Dries. (1988) On the structure of semialgebraic sets over p-adic fields. *J. Symbolic Logic*, 53(4):1138–1164.

[31] A. Tarski. (1951) *A Decision Method for Elementary Algebra and Geometry.* Univ. of California Press, 2 edition.

[32] L. van den Dries. (1982) Some applications of a model-theoretic fact to (semi-) algebraic geometry. *Indag. Math.*, 44.

[33] V. Weispfenning. (1976) On the elemntary theory of Hensel fields. *Annals Math. Logic*, 10:59–93.

[34] V. Weispfenning. (1983) Quantifier elimination and decision procedures for valued fields. In G.H. Muller and M.M. Richter, editors, *Models and Sets : Logic Colloquim '83*, pages 419–472, Springer-Verlag, Berlin.

[35] V. Weispfenning. (1988) The complexity of linear problems in fields. *J. Symbolic Computation*, 5:3–27.