Approximating Volumes and Integrals in o-Minimal and p-Minimal Theories

Marek Karpinski (Bonn) [*] Angus Macintyre (Oxford) [†]

Abstract

Motivated by recent development in VC-dimension of neural networks and corresponding semi-Pfaffian sets, we introduce a new method of approximating volumes and integrals for a vast class of geometric and number theoretical problems.

[*]Dept. of Computer Science, University of Bonn, 53117 Bonn, and the International Computer Science Institute, Berkeley, California. Research supported by DFG Grant KA 673/4-1, and by the ESPRIT BR Grants 7097 and EC-US 030 and by DIMACS.
E-mail: marek@cs.uni-bonn.de

[†]Mathematical Institute, University of Oxford, Oxford OX1 3LB. Research supported in part by a Senior Research Fellowship of the EPSRC.
E-mail: ajm@maths.ox.ac.uk

1 Introduction

The locally compact fields of characteristic zero are $\mathbb{C}$, $\mathbb{R}$, and finite extensions of the p-adic fields $\mathbb{Q}_p$ (Pontrjagin – see [W67]). These fields are completions of number fields, and are linked in the adelic setting for arithmetic. It has become clear in the last thirty years that there are strong metamathematical resemblences too [M86]. Each such field is decidable, and has a natural quantifier-elimination relative to predicates with obvious topological significance [T51], [M76]. Van den Dries and Denef [DD88] revealed a striking analogy between $\mathbb{R}$ and the p-adic fields at the level of subanalytic sets. In this paper we consider analogies at the level of probability theory, mainly but not exclusively for Haar measure. The central topic is the Vapnik-Chervonenkis (cf. [L92], [GJ93], [KM97a]) dimension (VC-dimension) of families of definable sets, and we pay special attention to various ways of approximating the measure of definable sets. This research was also motivated by the problems of estimating the VC-dimension of neural networks.

The first of our results were obtained in October – November 1994, just prior to the Dagstuhl meeting on Neural Computing, 7 – 11 Nov., 1994. In the real case they overlap with those of Pascal Koiran [K95], see also [KM96], with whom we discussed at Dagstuhl approximate definition of volume via VC-dimension.

2 Basic Structures

2.1 Let K be a field, usually locally compact of characteristic zero. We use the standard first-order language for K, with $+, -, \cdot, 0, 1$, and, if need be, some other distinguished constants.

If $K = \mathbb{R}$ (or a real closed field) we can define order in the field language by

$$x \geq y \iff (\exists t)(t^2 = x - y) \qquad .$$

So, in a clear sense, the Euclidean topology and metric are interpretable in the field language.

If $K = \mathbb{Q}_p$, $p \neq 2$, we define:

$$v(x) \geq 0 \Longleftrightarrow (\exists y)(y^2 = 1 + px^2) \qquad .$$

There are similar devices for such a definition, for all finite extensions of all p-adic fields. So, again, the natural topology and metric are interpretable in the field language.

For $\mathbb{C}$, the situation is different. $\mathbb{C}$ has a lot of field automorphisms, but only two continuous ones (for the Euclidean topology), so the topology is not definable. To restore the analogy, one should extend the field language by an extra predicate for $\mathbb{R}$, and then of course

$$\begin{aligned} |z| \leq 1 \quad \Longleftrightarrow \quad & (\exists u, v, w, t)(u, v, w \in \mathbb{R} \\ & \wedge \quad t^2 + 1 = 0 \\ & \wedge \quad z = u + tv \\ & \wedge \quad u^2 + v^2 + w^2 = 1) \end{aligned}$$

Though it is not central to this paper, we will sometimes (for completeness) refer to $(\mathbb{C}, \mathbb{R})$ instead of $\mathbb{C}$ when we want to keep an analogy.

2.2 Order and p-adic analogues Though $<$ is definable in $\mathbb{R}$, there are compelling mathematical reasons (going back to Tarski) for taking it as primitive. Henceforward when we discuss any language for $\mathbb{R}$, $<$ will be a primitive.

Similarly, for the $\mathbb{Q}_p$ there are compelling reasons for taking as primitives each:

$$P_k(x) \Longleftrightarrow (\exists y)(y^k = x)$$

see [M76, D86]. Again, in any discussion of $\mathbb{Q}_p$, these are taken as primitive. For the finite extensions of $\mathbb{Q}_p$, see [PR84] for the appropriate modification.

For $\mathbb{C}$, as needed, we add $\mathbb{R}$, and $<$ on $\mathbb{R}$ as primitives.

2.3 The subanalytic settings Fix one of the locally compact K as above, with appropiate language. For every point $\alpha = (\alpha_1, \ldots, \alpha_r)$ in K^r (any r), and every closed *compact* polydisc Δ in K^r, centered at α and every formal power series $f(x_1, \ldots, x_r)$ such that $f(x_1 - \alpha_1, \ldots, x_r - \alpha_r)$ converges (in K's metric) on an open set properly containing Δ, we add a primitive $f_{\Delta,\alpha}$, which is to be interpreted as a function $K^r \to K$, 0 outside Δ, and defined by the sum of the above series on Δ.

The restriction to *compact* Δ is esssential for the results below. Note that there is huge redundancy, and that an entire analytic function need have no name, though its restrictions to compacta will.

In each setting (relative to a K) let L_{an} be the resulting language.

Rather remarkably, one has:

Theorem 1: ([DD88]) Let X be a compact subset of $\mathbb{R}^n$. A subset Y of X is subanalytic in X if and only if there is a formula $\Phi(v_1, \ldots, v_n)$ of L_{an} such that

$$Y = \{\bar{x} \in \mathbb{R}^n : \mathbb{R}^n \models \Phi(\bar{x})\} \cap X \qquad .$$

For the definition and history of the notion *subanalytic* see [GM87].

Van den Dries and Denef [DD88] then *defined* subanalytic for K a finite extension of $\mathbb{Q}_p$, and *derived* a theory with a very strong analogy to that in $\mathbb{R}$. It seems that no one has written down the analogous treatment for $(\mathbb{C}, \mathbb{R})$, but it is known to experts that this is straightforward.

2.4 $\mathbf{L}_{\text{an,exp}}$ (for $\mathbb{R}$ only) Here one adds to L_{an} a symbol exp for the exponential function on $\mathbb{R}$. The sets definable in this language have most of the properties of subanalytic sets, and have proved useful in cohomological investigations [SV96] where the usual subanalytic theory is inadequate. For a beautiful discussion of logical extensions of the subanalytic category, see [DM96]. For the fundamental results, see [W96a], [DM96], [DMM94].

3 A Unifying Feature — VC-Dimension

3.1 It has been proved that $\mathbb{R}$ in L, L_{an}, $L_{\text{an,exp}}$, L_{DS}, L_{S}, L_{Pfaff}, and the Pfaffian closure of any of the above, is o-minimal, and this yields a striking array of Finiteness Theorems [D97], though usually with little effective content. Similar results for $\mathbb{C}$ in L or L_{an} are easily obtained.

For the totally disconnected fields it is not clear what Finiteness Theorems (Cell Decompositions, etc. [D97]) to emphasize, and work is still in progress on the elusive notion of p-minimality [DHM96].

In this paper we concentrate on combinatorial and probabilistic notions which make sense for all the locally compact K.

3.2 VC-Dimension We assume familiarity with the basics about VC dimension ([BEHW90], [V82], [L92], [GJ93], [KM97a]). For a class $\mathcal{C}$ of subsets of a universe X, we say a subset S of X is *shattered* by $\mathcal{C}$ if

$$\{S \cap C : C \in \mathcal{C}\} = P(S) \qquad \text{(powerset of S)}.$$

If arbitrarily large finite subsets of X are shattered by $\mathcal{C}$, we say $\mathcal{C}$ has infinite VC-dimension. Otherwise, we say $\mathcal{C}$ has finite VC-dimension, and define the VC-dimension of $\mathcal{C}$ as the maximal cardinality of a shattered set. Before turning to examples there are two general results to emphasize, one combinatorial and one probabilistic.

3.3 Sauer's Lemma Suppose $\mathcal{C}$ has finite VC-dimension d. Fix an integer n, and consider $S \subseteq X$, $|S| = n$. Let $\Pi_{\mathcal{C}}(n)$ be the supremum, as S varies, of

$$\Big|\{S \cap C : C \in \mathcal{C}\}\Big| \qquad .$$

Of course $\Pi_{\mathcal{C}}(n) \leq 2^n$.

Sauer's Lemma (cf., e. g., [AB92]) says:

Lemma 2: If $n \geq d$, $\Pi_{\mathcal{C}}(n) < \left(\frac{e\,n}{d}\right)^d$

Thus $\Pi_{\mathcal{C}}(n)$ has polynomial growth.

Because of this there is clear interest in the infimum of all reals r such that $\Pi_{\mathcal{C}}(n)/n^r$ is bounded (as $n \to \infty$).

Dudley [D84] calls this r dens($\mathcal{C}$), the density of $\mathcal{C}$.

3.4 Probabilistic Considerations As above, $\mathcal{C}$ has finite VC-dimension d.

Let μ be a probability distribution on X, and, for $m \geq 1$, μ^m the product distribution on X^m (to be construed as the space of ordered m-samples from X).

One makes some mild assumptions on μ ("the VC assumptions").

These are:

i) Each $C \in \mathcal{C}$ is μ-measurable;

ii) (χ_C is the characteristic function of C).

Define $\mathrm{RF}_C(x_1, \cdots, x_n)$ as

$$\frac{1}{n} \sum_i \chi_C(x_i) \qquad .$$

We require, for each n, the functions

$$\sup_{C \in \mathcal{C}} \left| \mathrm{RF}_C(x_1, \cdots, x_n) - \mu(C) \right|$$

and

$$\sup_{C \in \mathcal{C}} \left| \mathrm{RF}_C(x_1, \cdots, x_n) - \mathrm{RF}_C(y_1, \cdots, y_n) \right|$$

to be respectively μ^n and μ^{2n} measurable.

Then one has the splendid theorems of Vapnik-Chervonenkis [BEHW90], [AB92], [V82]:

Theorem 3: ($X, \mathcal{C}, d, \mu$ as above, μ satisfying th VC-assumptions)

Fix $\delta, \epsilon > 0$. Suppose

$$n \geq \max\left(\frac{4}{\varepsilon} \log \frac{2}{\delta} , \frac{8d}{\varepsilon} \log \frac{13}{\varepsilon} \right)$$

Then

$$\mu^n\Big(\{(x_1, \ldots, x_n) : \sup_{C \in \mathcal{C}} \left| \mathrm{RF}_C(\bar{x} - \mu(C) \right| < \varepsilon \} \Big) > 1 - \delta \qquad .$$

For the significance of this for learning theory, see [V82], [BEHW90]. Of course the theorem gives a uniform version of Bernoulli's theorem that the frequency of occurrence of a certain event $\mathcal{A}$ in a sequence of independent trials converges (in probability) to the probability of this event. It allows many probabilities to be tested simultaneously by a single random sample.

An important corollary of the main theorem is

Corollary: (Same notation as Theorem 3.) If $m \geq \frac{8d}{\varepsilon} \log \frac{13}{\varepsilon}$ there exists a set $\{x_1, \ldots, x_m\} \subseteq X$ (called a ε-net) such that for every C

$$\mu(C) \geq \varepsilon \qquad C \cap \{x_1, \ldots, x_m\} \neq \emptyset \qquad .$$

For future reference, and to make it clear what advantages finite VC-dimension bestows, we recall the Chernoff and Hoeffding inequalities for arbitrary (X, μ) with no assumptions on VC-dimension.

The two Chernoff bounds are concerned with n independent trials each of which has probability p of success (in our uses $p = 1/2$, and the trials are for membership of a fixed C). Let $\mathrm{LE}(p, n, s)$ (resp. $\mathrm{GE}(p, n, s)$ be the probabilities that there at most (resp. at least) s successes in n independent trials. Then for $0 \leq \beta \leq 1$

i) $\mathrm{LE}(p, n, (1-\beta)np) \leq e^{-\beta^2 np/2}$,

ii) $\mathrm{LE}(p, n, (1-\beta)np) \leq e^{-\beta^2 np/3}$.

We can state Hoeffding's result in slightly greater generality. It says:

Let $Y_1, \ldots, Y_n$ be independent random variables with zero means and bounded ranges, $a_i \leq Y_i \leq b_i$. For each $\eta > 0$

$$\Pr\Big\{|Y_1 + \cdots + Y_n| \geq \eta\Big\} \leq 2\exp[-2\eta^2 \Big/ \sum_{i=1}^{n} (b_i - a_i)^2] \qquad .$$

For a proof, see [P84], [Appendix A].

4 VC in Logic

4.1 Let M be a structure for some first-order logic, and let $\varphi(v_1, \ldots, v_k, y_1, \ldots, y_\ell)$ be a formula. For $\tilde{\beta} \in M^\ell$ define a subset $\varphi_{\tilde{\beta}}$ of M^k by

$$\varphi_{\tilde{\beta}} = \{\bar{x} \in M^k : M \models \varphi(\bar{x}, \tilde{\beta})\} \qquad .$$

Let $\mathcal{C}_\varphi = \{\varphi_{\tilde{\beta}} : \tilde{\beta} \in M^\ell\}$, with $X = M^k$.

We consider examples where $\mathcal{C}_\varphi$ has finite VC-dimension d, and we try (but usually fail) to get bounds for d.

4.2 Shelah [S78] defined (for purposes of his Classification Theory) the notion: M does not have the Independence Property (IP).

There is also a local notion for formulas [S78], which turns out to be dual to finiteness of VC-dimension [L92, Def. 1.1]. The duality involves M^k and M^ℓ for $\varphi(v_1, \ldots, v_k, y_1, \ldots, y_\ell)$ as above. In particular one has:

Theorem 4: M does not have the Independence Property iff each $\mathcal{C}_\varphi$ has finite VC-dimension.

Since logicians have given many examples of failure of the Independence Property, one can simply list many examples of $\mathcal{C}_\varphi$ with finite VC-dimension. (These results have little effective content in general).

4.3 Examples

4.3.1 M any o-minimal structure [KPS86], [L92]. The bounds on VC-dimension in [L92] are in general astronomical, involving iterated Ramsey numbers. It is however of great interest that Wilkie has shown that the density of $\mathcal{C}_\phi$ is $\leq \ell$, and we give an independent proof in section 5 below. We have already tested numerous examples.

4.3.2 M any weakly o-minimal structure. The method of [L92] works. The most interesting example is a real closed ring [CD83], or $(\mathbb{R}^*, \mathbb{R})$, $\mathbb{R}^*$ a nonstandard model of the reals. See [DL95] for a vast generalization.

4.3.3 M any ordered abelian group [GS84].

4.3.4 M any henselian field with residue field not having the Independence Property, and satisfying some "Ax-Kochen-Ersov" restrictions [D81]. In particular $\mathbb{Q}_p$ is an example, or any finite extension.

4.3.5 Let M be a structure built on a field as in 4.3.4, such that for any $\varphi(\nu, y_1, \ldots, y_\ell)$ there are finitely many $\psi_i(\nu, w_1, \ldots, w_k)$ in the field language such that for any $\tilde{\beta} \in M^\ell$ there is an i and a $\tilde{\tilde{\gamma}}$ such that $\varphi_{\tilde{\beta}} = (\psi_i)_{\tilde{\tilde{\gamma}}}$. Then, by 4.3.4 and [L92] M does not have IP.

4.3.6 For $M = \mathbb{Q}_p$ (or a finite extension) in L_{an}, [DHM96] verifies the hypotheses of 4.3.5 so all $\mathcal{C}_\varphi$ have finite VC-dimension.

4.4 Effective Bounds These are unfortunately rare. The main one is in [KM97a]:

Theorem 5: Work in L_{Pfaff}. Fix q many $h(\nu_1, \ldots, \nu_k, w_1, \ldots, w_\ell)$ occuring in a Pfaffian chain of length q and degree D. Let $f_i(\bar{\nu}, \tilde{w})$, $1 \leq i \leq s$ be polynomials of degree $\leq \Delta$ in $\bar{\nu}$, $\tilde{w}$ and the q many h. Let $\Phi(\bar{\nu}, \tilde{w})$ be a Boolean combination of the $f_i(\bar{\nu}, \tilde{w}) > 0$.

Then the VC-dimension of $E_{\tilde{\beta}}$ is bounded by

$$2(q\ell)(q\ell - 1) \;+\; 2\ell \log\Delta \;+\; 2\ell \log(\ell\Delta + \ell D + 1) \;+ \\ 2q\ell \log\ell \;+\; 2q\ell \log(\ell\Delta + \ell D + 1) \qquad .$$

If one tracks the proof back to its roots one sees that it involves homology theory and more specifically Morse theory. For this see [K91], [H76], [KM97a].

For $q = 0$ one gets estimates for semi-algebraic sets (as was done by Goldberg and Jerrum [GJ93] if given in a quantifier-free way. In Tarski's original language, for arbitrary Φ one has to keep careful track of quantifier-elemination a la Renegar [R92], as was done by [GJ93]. The details are very unpleasant.

We wish to stress that in the p-adic "semi-algebraic" case one has *no* bounds.

It is worth noting that for Φ as in Theorem 5 our *method* gives for density of $\mathcal{C}_\Phi$ a bound ℓ, in agreement with Wilkie's [W96b].

One final remark is that we know how to get effective bounds for various ordered abelian groups, particularly $\mathbb{Z}$.

4.5 More examples via Interpretations We refer to [H93] for the logical notion of interpretation. Easily, if M_1 is interpretable in M_2, and M_2 does not have IP then M_1 does not have IP. Using the real and p-adic examples above, one can get examples in the setting of analytic or C^∞ manifolds, p-adic manifolds, real or p-adic Lie groups, etc.

Also, $(\mathbb{C}, \mathbb{R})$, even in its subanalytic versions, is interpretable in the analogous version of $\mathbb{R}$. So we get lots of complex examples too.

5 Probabilistic Considerations Again

5.1 Any of the above locally compact K carries a Haar measure μ, uniquely normalized to give measure 1 to the unit ball. Lack of IP of course descends to the unit ball, giving us the probabilistic consequences of finite VC-dimension.

We do not yet need, however, to confine ourselves to μ. We let ν be a Borel probability measure on some compact subset X of K^k. As usual $\varphi(\nu_1, \ldots, \nu_k, w_1, \ldots, w_\ell)$ is fixed, so $\mathcal{C}_\varphi$ has finite VC-dimension d. About φ we assume one of the following:

(a) $K = \mathbb{R}$ with any o-minimal structure including $+, \cdot$;

(b) $K = a$ finite extension of $\mathbb{Q}_p$, with subanalytic structure;

(c) $K = (\mathbb{C}, \mathbb{R})$ in L_{an}.

Lemma 6: (Above assumptions) Each $\varphi_{\tilde{\beta}}$ is Borel.

Proof: (a) This is immediate from cell-decomposition [D97]. In fact we prove more in Section 9 below. (b) This follows from the quantifier-elimination of [DD88]. (c) Exercise. $\diamond$

5.2 Now we turn to the VC-assumptions of 3.4, for any (a),(b),(c). That is, we have to consider

$$\sup_{C \in \mathcal{C}_\varphi} \left| \mathrm{RF}_C(x_1, \cdots, x_n) - \nu(C) \right|$$

and

$$\sup_{C \in \mathcal{C}_\varphi} \left| \mathrm{RF}_C(x_1, \cdots, x_n) - \mathrm{RF}_C(y_1, \cdots, y_n) \right|$$

Since $\mathcal{C}_\varphi$ is in general uncountable, it is not general nonsense to get these functions (ν^n, resp. ν^{2n}) measurable. See [D84] and [P84] for discussion of this kind of point, and methods of solution.

We will follow Pollard's Appendix C. Our setting involves $\varphi(\nu_1, \ldots, \nu_k, w_1 \ldots, w_\ell)$, and a measure ν on $\mathbb{R}^k$, about which we assume only that it is complete. Just to follow Pollard, we put μ^ℓ on $\mathbb{R}^\ell$, with μ Haar measure. Then we put $\nu \times \mu^k$ on $\mathbb{R}^{k+l}$.

Now, for φ coming from (a),(b),(c)

$$(\bar{\nu}, \tilde{w}) \longmapsto \begin{cases} 1 & \text{if } \varphi(\bar{\nu}, \tilde{w}) \\ 0 & \text{if } \neg\,\varphi(\bar{\nu}, \tilde{w}) \end{cases}$$

is measurable.

It follows by Fubini that

$$\tilde{\beta} \longmapsto \nu(\varphi_{\tilde{\beta}})$$

is μ^k measurable.

Now

$$(x_1, \ldots, x_n, \tilde{\beta}) \longmapsto \mathrm{RF}_{\varphi_{\tilde{\beta}}}(x_1, \ldots, x_n)$$

is evidently $\nu^n \times \mu^\ell$-measurable, and so

$$(x_1, \ldots, x_n, \tilde{\beta}) \longmapsto \left| \mathrm{RF}_{\varphi_{\tilde{\beta}}}(x_1, \ldots, x_n) - \nu(\varphi_{\tilde{\beta}}) \right|$$

and

$$(x_1, \ldots, x_n, y_1, \ldots, y_n) \longmapsto \left| \mathrm{RF}_{\varphi_{\tilde{\beta}}}(x_1, \ldots, x_n) - \mathrm{RF}_{\varphi_{\tilde{\beta}}}(y_1, \ldots, y_n) \right|$$

are respectively $\nu^n \times \mu^k$- and ν^{2n}-measurable.

Finally it follows by the analytic considerations of Pollard (p. 197) that

$$\sup_{\tilde{\beta}} \left| \mathrm{RF}_{\varphi_{\tilde{\beta}}}(x_1, \cdots, x_n) - \nu(\varphi_{\tilde{\beta}}) \right|$$

and

$$\sup_{\tilde{\beta}} \left| \mathrm{RF}_{\varphi_{\tilde{\beta}}}(x_1, \cdots, x_n) - \mathrm{RF}_{\varphi_{\tilde{\beta}}}(y_1, \cdots, y_n) \right|$$

are ν-measurable.

We have proved:

Lemma : (Hypotheses (a),(b),(c)) The VC-assumptions hold for $\mathcal{C}_\varphi$.

5.3 It is of course extremely remarkable that for the huge variety of sets coming under (a),(b),(c), the convergence of frequencies to measures is completely independent of ν. For example, the language of exponentiation interprets the absolute metric in hyperbolic geometry, as well as many topics about sigmoidal neural nets. The p-adic case, via Poincaré series, is of considerable interest in number theory and Lie theory [D86, S93].

6 Approximate Definitions of Haar measure

6.1 The results now to be described overlap significantly with those of Koiran [K95], but were independently obtained (see Section 1).

The goal was to see if $\mu(\varphi_{\tilde{\beta}})$ could be obtained "in first-order terms" .

Setting: $M = \mathbb{R}$ with L-structure including $+, -, \cdot, <$, and, as usual, $\varphi(\nu_1, \ldots, nu_k, w_1, \ldots, w_\ell)$ a formula in the language L. Assume $\varphi_{\tilde{\beta}}$ is μ^k-measurable, for all $\tilde{\beta}$. For convenience, assume each $\varphi_{\tilde{\beta}} \subseteq [0,1]^k$, so has $\mu^k(\varphi_{\tilde{\beta}}) \leq 1$.

Problem 1: Is $\tilde{\beta} \longmapsto \mu^k(\varphi_{\tilde{\beta}})$ L-definable?

Answer: No, for the semi-algebraic language. For consider $\varphi_{\tilde{\beta}}$ as the circle with radius $\frac{1}{2(1+\beta^2)}$. Then $\mu^2(\varphi_{\tilde{\beta}}) = \frac{\pi}{4} \cdot \frac{1}{1+\beta^2}$.

Since π is undefinable in the real field, the answer follows. This of course may seem like cheating as the function is definable from π.

A more convincing counterexample is easily provided. (Hint: Consider $\arctan(\beta)$ as a measure of a suitable φ_β. No constants can help here).

To formulate our second problem, we need a definition.

Definition 1: ($\varepsilon > 0$) $\Psi(w_1, \ldots, w_\ell, y)$ strongly defines an ε-approximate volume for φ if $\mathbb{R} \models \Psi(\tilde{\beta}, r) \Longleftrightarrow |r - \mu^k(\varphi_{\tilde{\beta}})| < \varepsilon$.

Problem 2: Is there Ψ in L (even using parameters) so that Ψ strongly defines an ε-approximate volume for φ?

Answer: No, essentially for the same reason as before. For, if ε is fixed, $\mu^k(\varphi_{\tilde{\beta}})$ is the centre of the interval $\{r : |r - \mu^k(\varphi_{\tilde{\beta}})| < \varepsilon\}$, and so could be defined from Ψ.

One may imagine that such examples exist only because the semi-algebraic sets are rather restricted. One may seek other o-minimal examples for which Problem 1 has a positive answer. For $L_{\text{an,exp}}$ one does no better, as one sees by considering the undefinability result in [DMM94].

There remains however the intriguing possibility of some huge o-minimal structure on $\mathbb{R}$ for which Problem 1 has a positive solution.

6.2 The right notion

We need another definition.

Definition 2: Ψ defines an ε-approximate volume for φ if

i) $\mathbb{R} \models \Psi(\tilde{\beta}, r) \longrightarrow |r - \mu^k(\varphi_{\tilde{\beta}})| < \varepsilon$;

ii) $|r - \mu^k(\varphi_{\tilde{\beta}})| < \frac{\varepsilon}{4} \longrightarrow \mathbb{R} \models \Psi(\tilde{\beta}, r)$.

Theorem 8: For each φ as in 5.1 (a) or (c) and each $\varepsilon > 0$, there exists Ψ which defines an ε-approximate volume for φ.

This is essentially Theorem 5 in Koiran's [K95], and his proof is essentially the same as ours. (There are some obscurities in his presentation, for example in Definition 3 where the notion is trivial unless E is replaced by a (definable) family). Also, in his Theorem 2, he seems to ignore the VC-assumptions. In fact, in his application they are trivially satisfied.

6.3 Our motivation in finding Ψ was the following. From finiteness of VC-dimension of $\varphi_{\tilde{\beta}}$ one knows that "generically" a small sample $\bar{x}$ will give, via RF$(\bar{x})$, $\mu^k(\varphi_{\tilde{\beta}})$ with error $< \varepsilon$ (provided the VC-assumptions are satisfied). Now the sample $\bar{x}$ is in $([0,1]^k)^n$, where n is the size of the sample, and $([0,1]^k)^n$ can be constructed as a compact group via

$$(\mathbb{R}^k)^n \longrightarrow \left((\mathbb{R}/\mathbb{Z})^k\right)^n \qquad ,$$

and of course Haar measure is induced. Now we use a model-theoretic intuition:

generic $\equiv$ large $\equiv$ a small number of translates cover the group [P87].

Koiran [K95] got the same proof, from a complexity-theoretic intuition.

6.4 The proof We interpret $\mathbb{R}/\mathbb{Z}$ in the L-structure, using representations in $[0,1[$. Let I_k be the group $\left(\mathbb{R}/\mathbb{Z}\right)^k$, and $\oplus_k, \ominus_k$ its group addition and subtraction. These are clearly interpretable too.

Since Koiran's paper is published [K95], it seems reasonable to sketch our proof following his notation.

Let $n \in \mathbb{N}$, $\nu, \alpha \in [0,1]$.

Define

$$S_{\tilde{\beta},\nu,\alpha} = \left\{(x_1,\ldots,x_n) \in \left([0,1]^k\right)^n : |\mathrm{RF}_{\varphi_{\tilde{\beta}}} - \nu| \le \alpha\right\}$$

Evidently, uniformly,

$$S_{\tilde{\beta},\nu,\alpha} = \Gamma_{\tilde{\beta},\nu,\alpha}$$

for a certain formula Γ, and an easy calculation gives

Lemma 9: If φ is $\sum_m$, Γ may be chosen as a Boolean combination of $\sum_m$ formulas. If φ is Δ_m, Γ may be chosen Δ_m too.

Corollary: For each φ as in 5.1 each $S_{\tilde{\beta},\nu,\alpha}$ is Borel, and the family of them has finite VC-dimension.

Proof: Clear. $\diamond$

Now, for $\gamma \in (I^k)^n$, define

$$T_{\tilde{\beta},\nu,\alpha,\gamma} \qquad \text{as} \qquad \gamma \ominus S_{\tilde{\beta},\nu,\alpha} \qquad .$$

As before, uniformly

$$T_{\tilde{\beta},\nu,\alpha,\gamma} = \Theta_{\tilde{\beta},\nu,\alpha,\gamma}$$

for a formula Θ, where:

Lemma 10: If φ is $\sum_m$, Θ may be chosen as a Boolean combination of $\sum_m$ formulas. If φ is Δ_m, Θ may be chosen Δ_m too.

Corollary: For each φ as in 5.1 each $T_{\tilde{\beta},\nu,\alpha,\gamma}$ is Borel, and the family of them has finite VC-dimension.

Proof: Clear. $\diamond$

The next lemma is almost proved in Koiran ([K95], Lemma 6).

Lemma 11: Let D be the VC-dimension of the family of all $\lambda \ominus \varphi_{\tilde{\beta}}$, $\lambda \in I^k, \tilde{\beta}$ as usual. (This is finite under the assumptions of 5.1). Then the VC-dimension of the family of all $T_{\tilde{\beta},\nu,\alpha,\gamma}$ is $\leq 4D \cdot n \log n$.

Proof: Just check [K95] to get the constant right. $\diamond$

Now let $\mathcal{C}$ be the collection of all $T_{\tilde{\beta},\nu,\alpha,\gamma}$. By chapter 4 and 5, $\mathcal{C}$ has finite VC-dimension and satisfies the VC-assumptions.

This allows us to get, given $\varepsilon > 0$, an ε-net for $\mathcal{C}$. Let $\{t_1, \ldots, t_m\}$ be such an ε-net. We know by the Corollary to Theorem 3 that there is such a net with $m \geq \frac{32Dn \log n}{\varepsilon} \cdot \log(13/\varepsilon)$, where D is as in Lemma 11.

What does it mean for $\{t_1, \ldots, t_m\}$ to be an ε-net for $\mathcal{C}$? Note that all elements of $\mathcal{C}$ have the same measure ρ (as in Lemma 13), so it says that if $\rho \geq \varepsilon$ then every member of $\mathcal{C}$ meets $\{t_1, \ldots, t_m\}$. But $T_{\tilde{\beta},\nu,\alpha,\gamma}$ meets $\{t_1, \ldots, t_m\}$ if and only if for some i

$$t_i \in \gamma \ominus S_{\tilde{\beta},\nu,\alpha} \qquad ,$$

so if and only if

$$\gamma \in t_i \oplus S_{\tilde{\beta},\nu,\alpha} \qquad .$$

Since γ was arbitrary, the sets $t_i \oplus S_{\tilde{\beta},\nu,\alpha}$, $1 \leq i \leq m$ cover $(I^k)^n$. Since they all have the same measure, we deduce that the measure of $S_{\tilde{\beta},\nu,\alpha}$ is $\geq 1/m$.

To recapitulate, we got this conclusion from the preceding covering statement, which depends on $m \geq \frac{32Dn \log n}{\varepsilon} \cdot \log(13/\varepsilon)$, *and* $\rho \geq \varepsilon$.

Next:

Lemma 12:

1. If $\left(\mu^k\right)^n (S_{\tilde{\beta},\nu,\varepsilon/2}) > 2e^{-n\varepsilon^2/2}$ then $|\nu - \mu^k(\varphi_{\tilde{\beta}})| < \varepsilon$;
2. If $|\nu - \mu^k(\varphi_{\tilde{\beta}})| \leq \varepsilon/4$ then $\left(\mu^k\right)^n (S_{\tilde{\beta},\nu,\varepsilon/2}) \geq 1 - 2e^{-n\varepsilon^2/8}$.

Proof: This is done in [K95], using Hoeffding's inequality. $\diamond$

Now we put this together with the lower bound $1/m$ for $(\mu^k)^n(S_{\tilde{\beta},\nu,\alpha})$. If $1/m > 2e^{-n\varepsilon^2/2}$, *and* $\alpha = \varepsilon/2$, *and* $\rho = \mu(S_{\tilde{\beta},\nu,\alpha}) \geq \varepsilon$, we can conclude $|\nu - \mu^k(\varphi_{\tilde{\beta}})| < \varepsilon$. And the conclusion depends only on the covering of $(I^k)^n$ by m translates of $S_{\tilde{\beta},\nu,\varepsilon/2}$.

Conversely, suppose $|\nu - \mu^k(\varphi_{\tilde{\beta}}) < \varepsilon/4$. By Lemma 14.2, $(\mu^k)^n(S_{\tilde{\beta},\nu,\varepsilon/2}) \geq 1 - 2e^{-n\varepsilon^2/8} \geq 1/2$ if $n \geq \frac{8 \ln 4}{\varepsilon^2}$. So $(\mu^k)^n(T_{\tilde{\beta},\nu,\varepsilon/2,\gamma}) \geq 1/2$ all γ. If $\varepsilon \leq 1/2$, we can conclude from the preceding that m translates if $S_{\tilde{\beta},\nu,\varepsilon/2}$ cover $(I^k)^n$, provided $m \geq \frac{32Dn \log n}{\varepsilon} \cdot \log(13/\varepsilon)$.

Putting the estimates from the two sides together, we identify the following conditions on ε, n, m:

i) $\varepsilon \leq 1/2$;

ii) $m > 1/2e^{n\epsilon^2/2}$; (#)

iii) $n \geq \frac{8 \ln 4}{\varepsilon^2}$;

iv) $m \geq \frac{32Dn \log n}{\varepsilon} \cdot \log(13/\varepsilon)$.

Evidently, given $\varepsilon \leq 1/2$ we may then pick n to satisfy (ii), and then m to satisfy (ii) and (iv).

We have proved Theorem 9, in the following form (with hypotheses as in original statement).

Theorem 8: Suppose ε, n, m satisfy (#), where D is the VC-dimension of all $\lambda \ominus \varphi_{\tilde{\beta}}$, $\lambda \in I^k$, $\tilde{\beta}$ as usual. Then the formula Ψ, saying that m translates if $S_{\tilde{\beta},\nu,\varepsilon/2}$ cover $(I^k)^n$, defines an ε-approximate volume for φ.

6.5 Remarks on the Theorem The following are worth recording.

a) Derandomization has been achieved via group-theoretical considerations, and the result is specific to Haar measure.

b) If φ is $\sum_j$, Ψ is $\sum_{j+2}$.

c) One can easily write down a bound on the length of Ψ in terms of that of φ and the number D, which is readily calculated in the quantifier-free Pfaffian case [KM97a].

d) There seems little doubt that one can do something similar for rotation-invariant measures on spheres, etc.

7 p-adic version

7.1 The argument in section 6 adapts easily enough to the p-adic situation. Replace I by the compact ring $\mathbb{Z}_p$. The usual $+, -$ replace the $\oplus, \ominus$ of section 6.

Exactly as in section 6, the property that m translates if $S_{\tilde{\beta},\nu,\varepsilon/2}$ cover $(\mathbb{Z}_p^k)^n$ "defines" an ε-approximate volume for φ.

However this does not literally give us a p-adic formula Ψ, since in $S_{\tilde{\beta},\nu,\varepsilon/2}$ the $\tilde{\beta}$ are p-adic and the ν and ε real. There are two remedies.

One is to take Ψ in a 2-sorted language, with a p-adic sort and a real sort.

The other is to stay in one sort, and remark that in the above $S_{\tilde{\beta},\nu,\varepsilon/2}$ $\mathrm{RF}_{\varphi_{\tilde{\beta}}}(x_i, \ldots, x_n)$ takes only values j/n, $0 \le j \le n$. So one might reasonably restict ν to lie in some finite set $\subseteq \mathbb{Q} \subseteq \mathbb{Q}_p$. Then for each such restriction on ν we get a version of Ψ doing the right ε-approximation.

8 Estimating Integrals

8.1 Again we wish to have an argument uniform for $\mathbb{R}$ and $\mathbb{Q}_p$. The problem is to estimate (or express approximately)

$$\int_{X_{\tilde{\beta}}} f_{\tilde{\beta}}\, d\bar{x} \quad \text{, where} \quad \langle X_{\tilde{\beta}}\rangle, \langle f_{\tilde{\beta}}\rangle$$ are definable families, and the integration is against Haar measure.

Since definable sets and functions are measurable in the above situation, we may as well take $X_{\tilde{\beta}}$ constant. We restrict to integration over a (power of) the standard compact $\{x : |x| \leq 1\}$. We shall also assum the $f_{\tilde{\beta}}$ are uniformly bounded on their common domain.

Note however an essential difference between the real and p-adic cases. In the real case we integrate real functions, but in the p-adic case we integrate also real functions, typically $|g(\bar{x})|^s$, where g is a p-adic valued function and s a real variable (as in p-adic Poincare' series [D84]).

8.2 In the real case, if $f \geq 0$ we can of course express $\int f$as the volume "under the graph of f", but for the VC-methods the probabilities would be against the measure $\int f d\mu$, and could be reduced to μ estimates using a nonzero lower bound for f, and an upper bound.

8.3 We follow the treatment in Chapter 7 of [V82]. We give prominence to the real o-minimal case, and sketch later the modifications needed for the p-adic case.

We work in $\mathbb{R}^k$, with a definable family of functions $f(x_1, \ldots, x_k, \tilde{\beta})$, where $\tilde{\beta}$ runs through $\mathbb{R}^\ell$. Let I be the standard $[-1,1]^k$, and we consider the integrals

$$\int_x f(\bar{x}, \tilde{\beta})\, d\mu \quad , \mu \text{ Lebesgue measure.}$$

In fact, just as in the discussions of volumes, for now μ can be an arbitrary probability measure on X.

We do have to make some assumptions now, however. Let us assume that we have that the $f(\bar{x}, \tilde{\beta})$ take values in a fixed compact subset Y of $\mathbb{R}$, and fix τ such that

$$\sup_{\substack{\bar{x}\in X\\ y\in Y\\ \tilde{\beta}}} \left(y - f(\bar{x},\tilde{\beta})\right)^2 \leq \tau \qquad . \tag{1}$$

Now we assume we are in an o-minimal situation, and let h be the VC-dimension of

$$\left\{(\bar{x},\tilde{y},t) \;:\; f(\bar{x},\tilde{y}) > t\right\} \qquad . \tag{2}$$

For points $\bar{x}_1,\ldots,\bar{x}_\ell$ in I^k, let $\sum_{\tilde{\beta}}(\bar{x}_1,\ldots,\bar{x}_\ell) = \frac{1}{\ell}\sum_i f(\bar{x}_i,\tilde{\beta})$. We want to relate these sums to the integral, and fortunately Vapnik [V82] proves:

Theorem 9: Suppose $\ell > h$. Then for $\eta > 0$

$$\mu^\ell\Bigg(\Big\{(\bar{x}_1,\ldots,\bar{x}_\ell) \;:\; \Big|\int_X f(\bar{x},\tilde{\beta})\,d\mu - \textstyle\sum_{\tilde{\beta}}(\bar{x}_1,\ldots,\bar{x}_\ell)\Big| < 2\tau\sqrt{\dfrac{h(\ln(\frac{2\ell}{h}+1)-\ln(\frac{\eta}{9}))}{\ell}}\Big\}\Bigg) \geq 1-\eta \qquad .$$

(Notice that the measurability assumptions concealed in the proof of the above are automatic in the o-minimal case).

8.4 The p-adic case The only real differences are that I is replaced by $\mathbb{Z}_p$, and we integrate $|f(\bar{x},\tilde{\beta})|^s$ for some real (large positive) s.

Now the values will be assumed to lie in some compact Y_s (and if all $|f| \leq 1$, we can take a fixed Y_s for all s).

We have to think now of s as a parameter and consider the relation

$$|f(\bar{x},\tilde{y})|^s > t \qquad ,$$

$\bar{x},\tilde{y}$ p-adic, s,t real.

This is a subset of a product of real and p-adic spaces, so the issue of VC-dimension is less clear. But a moment's thought shows no more shattering is possible than for the p-adic relation

$$|f(\bar{x},\tilde{y})| > |w| \qquad ,$$

and this has VC-dimension h say (2).

Now, for $\bar{x}_1, \ldots, \bar{x}_\ell$ in X, and $s \in \mathbb{R}$, let

$$\sum\nolimits_{\tilde{\beta},s}(\bar{x}_1, \ldots, \bar{x}_\ell) = \frac{1}{\ell} \cdot \sum |f(\bar{x}_i, \tilde{\beta})|^s \qquad .$$

Let

$$\tau_s = \sup_{\substack{\bar{x} \in X \\ y \in Y_s \\ \tilde{\beta}}} \left| (y - |f(\bar{x}, \tilde{\beta})|^s \right|^2 < \infty$$

Then Vapnik gives us again:

Theorem 10: Suppose $\ell > h$. Then for $\eta > 0$

$$\mu^\ell \Bigg(\Big\{ (\bar{x}_1, \ldots, \bar{x}_\ell) \; : \; \Big| \int\limits_X f(\bar{x}, \tilde{\beta}) \, d\mu - \sum_{\tilde{\beta},s} (\bar{x}_1, \ldots, \bar{x}_\ell) \Big| < 2\tau_s \sqrt{\frac{h(\ln(\frac{2\ell}{h} + 1) - \ln(\frac{\eta}{9}))}{\ell}} \Big\} \Bigg) \geq 1 - \eta \qquad .$$

(Normally one will have a uniform bound on τ_s as s varies).

8.5 ϵ-Approximation of Integrals Now the discussion in Section 6 goes through with only obvious changes. It seems to us unnecessary to spell out the results or the proofs. The notion of expressing an ε-approximate integral is the obvious analogue of that for volume. The translation argument goes through, and Hoeffding's Inequality (cf. Lemma 12) still applies.

In the p-adic case, too, there is no real change.

8.6 Remark about Poincaré series By [M76], [D84] the maps

$$s \longmapsto \int |f(\bar{x}, \tilde{\alpha})|^s \, d\mu$$

are, uniformly in $\tilde{\alpha}$, rational functions of p^{-s}. It is, however, not at all obvious how to bound the complexity of such rational functions. If one knew this, one could interpolate the functions in the style of [GKS94]. This observations makes it obvious that in the p-adic case VC-dimensions are closely connected to interpolation of Poincaré series.

9 Linear Bound on Density

9.1 We are going to prove an unpublished result of Wilkie (from 1990). We have not seen his proof, which we suspect is more combinatorial than ours. In the case of quantifier-free Pfaffian formulas, our method in [KM97a] gives a very sharp result. We only recently realized, in writing [KM97b], that our axiomatic development there is almost good enough to handle the general case.

The result is:

Theorem 10: Let $\Phi(\nu_1, \ldots, \nu_k, y_1, \ldots, y_\ell)$ be a formula in an o-minimal theory on $\mathbb{R}$. Then $\text{dens}(\mathcal{C}_\Phi) \leq \ell$.

(It is obvious that this bound is optimal).

The proof goes via several ideas of independent interest.

Lemma 11: In an o-minimal theory on $\mathbb{R}$, any definable set $\mathcal{A}$ in $\mathbb{R}^n$ is a Boolean combination of definable closed sets.

Proof: This goes by induction on the dimension of $\mathcal{A}$ (see 4.2 of [DM96]), using $\mathcal{A} = \bar{\mathcal{A}} \setminus \text{Fr}(\mathcal{A})$, and the basic $\dim \text{Fr}(\mathcal{A}) < dim(\mathcal{A})$ if $\mathcal{A} \neq \emptyset$ ($\text{Fr}(\mathcal{A}) =$ Frontier of$\mathcal{A}$). $\diamond$

Corollary: In an o-minimal theory on $\mathbb{R}$, if $\mathcal{A}$ is a definable set in $\mathbb{R}^n$, and m is arbitrary then there are finitely many definable total C^m functions f_i such that $\mathcal{A}$ is a Boolean combination of the $\text{Zer}(f_i)$.

Proof: By the Zeroset Property [DM96]. $\diamond$

Corollary: For any $k \geq 1$ any o-minimal system on $\mathbb{R}$ admits quantifier-elimination in terms of total $\mathbb{C}^k$ functions.

Proof: Clear. $\diamond$

This allows us to make full use of a method from [KM97a]. There $n = k + \ell$, and we had a formula $\Phi(x_1, \ldots, x_k, y_1, \ldots, y_\ell)$ which was a Boolean combination of finitely many equations $f_i(\bar{x}, \tilde{y}) = 0$, f C^∞. The C^∞ assumption was to guarantee applicability of Sard, and can be avoided. Let us quickly review our argument.

Functions f_i, $0 \leq i < s$, are involved. Pick $\bar{a}_1, \ldots, \bar{a}_n \in \mathbb{R}^k$, and try to calculate the cardinality of $\mathrm{P}_{\mathcal{C}_\Phi}(\{\bar{a}_1, \ldots, \bar{a}_n\})$. Consider the functions $f_i(\bar{a}_j, \tilde{y})$ on $\mathbb{R}^\ell$. There are ns of them at most. As in [KM97a] one easily shows that there is $\varepsilon > 0$ such that

$$\left| \mathrm{P}_{\mathcal{C}_\Phi}(\{\bar{a}_1, \ldots, \bar{a}_n\}) \right| \quad \text{is bounded,}$$

for any $0 < \varepsilon_{ij} < \varepsilon, \; 0 \leq i < s, \; 1 \leq j \leq n$, by the number of connected components of the complement on $\mathbb{R}^\ell$ of the union of the sets

$$\{\tilde{y} : \; f_i(\bar{a}_j, \tilde{y}) = \varepsilon_{ij}\} \cup \{\tilde{y} : \; f_i(\bar{a}_j, \tilde{y}) = -\varepsilon_{ij}\} \quad 0 \leq i < s, \; 1 \leq j \leq n.$$

Now the argument in [KM97a] uses only that the f_i are $C^{\ell+1}$, and bounds the required number of connected components by $\mathrm{B} \cdot (\frac{2sn\ell}{\ell})^\ell$, where B is the maximum possible number of connected components of any set in $\mathbb{R}^\ell$ defined as

$$\{\tilde{y} : \; \mathrm{F}_1(\bar{\alpha}_1, \tilde{y}) = \lambda_1 \wedge \cdots \wedge \mathrm{F}_r(\bar{\alpha}_r, \tilde{y}) = \lambda_r\}$$

where $r \leq \ell$, the $\bar{\alpha}_a$ and λ_a are parameters, and each F is an f_i. (B is of course finite, by o-minimality.)

B does not depend on n, so $\mathrm{dens}(\mathcal{C}_\Phi) \leq \ell$.

Thus:

Theorem 12: For any $\Phi(\nu_1, \ldots, \nu_k, y_1, \ldots, y_\ell)$ in an o-minimal theory on $\mathbb{R}^\ell$ $\mathrm{dens}(\mathcal{C}_\Phi) \leq \ell$.

10 Concluding Remarks

This work suggests some new problems.

Problem 1: (cf. 6.1) Extend the result of [KM97a], bounding VC-dimension in the quantifier-free Pfaffian case, to a relative Pfaffian case of the type analyzed in [KM97b].

Abstractly, we know how to do this. The $q(q-1)/2$ term in [K91] is the degree of a certain iteration of the Jacobian, and in any given relativization we could work this out. But we are, rather, interested in finding cases, where a similar quadratic bound would hold.

Problem 2: Is there a p-adic analogue of Theorem 10? We would not expect the bound ℓ for the density, but see no reason why there should not be some linear bound.

Finally, there are significant problems about finite fields. It is known [D80] that nonprincipal ultraproducts of finite fields have the independence property, so there is no hope of strong (independent of the field) bounds on VC-dimension of $\mathcal{C}_\Phi$. But one may reasonably hope to discover the dependence on p, both for VC and for density.

References

[AB92] M. Anthony, N. Biggs, *Computational Learning Theory: An Introduction*, Cambridge University Press, 1992.

[BEHW90] A. Blumer, A. Ehrenfeucht, D. Haussler, M. Warmuth, *Learnability and the Vapnik-Chervonenkis Dimensions*, J. ACM **36** (1990), pp. 929–965.

[CD83] G. Cherlin, M. Dickmann, *Real closed rings II: Model Theory*, Annals of Pure and Applied Logic 25 (1983), pp. 213-231.

[D81] F. Delon, *Types sur C((X)) Théories Stables*, 2° année, IHP, Paris (1981).

[D84] J. Denef, *The Rationality of the Poincaré Series associated to the p-adic Points on a Variety*, Invent. Math. 77 (1984).

[D86] J. Denef, *p-adic semi-algebraic sets and cell decomposition*, J. Reine Angew. Math. 369 (1986), pp. 154–166.

[DD88] J. Denef, L. van den Dries, *p-adic and real subanalytic sets*, Annals of Mathematics 128 (1988), pp. 70–138.

[D97] L. van den Dries, *Tame Topology and o-Minimal Structures*, to appear in Cambridge University Press.

[DM96] L. van den Dries, C. Miller, *Geometric categories and o-minimal structures*, Duke Journal **84** (1996), pp. 497-540.

[DHM96] L. van den Dries, D. Haskall and D. Macpherson, *One-dimensional p-adic subanalytic sets*, to appear in J. London Math. Soc.

[DL95] L. van den Dries and A.H. Lewenberg, *T-convexity and tame extensions*, J. Symbolic Logic **60** (1995), pp. 74–102.

[DMM94] L. van den Dries, A. Macintyre and D. Marker, *The elementary theory of restricted analytic fields with exponentiation*, Annals of Mathematics 140 (1994), pp. 183–205.

[D84] R.M. Dudley, *A course on emperical processes*, pp. 2-142, in Lecture Notes in Mathematics 1097, (ed. P.L. Hennequeri), 1984.

[D80] J.L. Duret, *Les corps faiblement algebriquement clos non separablement clos ont la proprieté d'cridépendance*, in Model Theory of Algebra and Arithmetic, Lecture Notes in Mathematics **83** (1980), pp. 136-162.

[GM87] M. Goresky, R. Macpherson, *Stratified Morse Theory*, Springer-Verlag (1987).

[GJ93] P. Goldberg, M. Jerrum, *Bounding the Vapnik Chervonenkis Dimension of Concept Classes Parametrized by Real Numbers*, Machine Learning **18** (1995), pp. 131-148. A preliminary version appeared in Proc. 6th ACM Workshop on Computational Learning Theory, pp. 361–369.

[GKS94] D. Grigoriev, M. Karpinski, M. Singer, *Computational Complexity of Sparse Rational Interpolation*, SIAM J. Comput. 23 (1994), pp. 1–11.

[GS84] Y. Gurevich, P. Schmitt, *The Theory of Ordered Abelian Groups does not have the Independence Property*, in Transactions of the American Math. Soc., Vol. 284 (1984), pp. 171-182.

[H76] M.W. Hirsch, *Differential Topology*, Springer-Verlag (1976).

[H93] W. Hodges, *Model Theory*, Cambridge University Press (1993).

[KM97a] M. Karpinski, A. Macintyre, *Polynomial Bounds for VC Dimension of Sigmoidal and General Pfaffian Neural Networks*, J. Comput. Syst. Sci. **54** (1997), pp. 169–176 (a preliminary version appeared in Proc. 27th ACM STOC (1995), pp. 200–208).

[KM97b] M. Karpinski, A. Macintyre, *o-Minimal Expansions of the real Field: A Characterization, and an Application to Pfaffian Closure*, Research Report No. 85173-CS, Univ. Bonn, 1997, to be submitted.

[KM97c] M. Karpinski, A. Macintyre, *Closing o-Minimal Structures Under Definite Integrals*, in preparation.

[KM96] M. Karpinski, A. Macintyre, *Approximating the Volume of General Pfaffian Bodies*, to appear in LNCS, Springer Verlag, 1997.

[KW93] M. Karpinski and T. Werther, *VC Dimension and Uniform Learnability of Sparse Polynomials and Rational Functions*, SIAM J. Computing **22** (1993), pp. 1276–1285.

[K91] A.G. Khovanski, *Fewnomials*, American Mathematical Society, Providence, R.I., 1991

[KPS86] J. Knight, A. Pillay, C. Steinhorn, *Definable Sets and Ordered Structures II*, Trans. Amer. Math. Soc. **295** (1986), pp. 593-605.

[K95] P. Koiran, *Approximating the Volume of Definable Sets*, Proc. 36th IEEE FOCS (1995), pp. 258–265.

[L92] M.C. Laskowski, *Vapnik-Chervonenkis Classes of Definable Sets*, J. London Math. Society **45** (1992), pp. 377–384.

[M76] A. Macintyre, *On definable subsets of p-adic fields*, J. Symbolic Logic 41, 1976, pp. 605–610.

[M86] A. Macintyre, *Twenty years of p-adic model theory*, Proc. Logic Colloquium 84, North Holland 1986, pp. 121–153.

[P87] B. Poizat, *Groupes Stables*, Nur al Mantiq Wal-Márifah, Paris (1987).

[P84] D. Pollard, *Convergence of Stochastic Processes*, Springer 1984.

[P87] B. Poizat, *Groupes Stables* Nur al Mantiq Wal-Márifah, Paris (1987).

[PR84] A. Prestel, P. Roquette, *Formally p-adic Fields*, Lecture Notes in Mathematics Vol. 1050, Springer-Verlag, 1984.

[R92] J. Renegar, *On the computational Complexity and Geometry of the First-Order Theory of the Reals*, Parts I, II, III, j. of Symb. Comp. 3 (1992), pp. 255-352.

[S93] M. du Sautoy, *Finitely Generated Groups, p-Adic Analytic Groups and Poincaré Series*, Anals of Mathematics **137** (1993), pp. 639-670.

[SV96] W. Schmid, K. Vilonen, *Characteristic cycles for constructive sheaves*, Inventions Math **124** (1996), pp. 451-502.

[S78] S. Shelah, *Classification Theory*, North Holland (1978).

[T51] A. Tarski, *A Decision Method for Elementary Algebra and Geometry*, 2nd Edition, University of California Press, Berkeley, 1951.

[V82] V. Vapnik, *Estimation of Dependencies Based on Empirical Data*, Springer Series in Statistics, 1982.

[W67] A. Weil, *Basic Number Theory*, Springer-Verlag, 1967.

[W95] A.J. Wilkie, *Some Remarks on a Paper of Charbonel*, Preprint, Oxford University, 1995.

[W96a] A.J. Wilkie, *Model Completeness Results for Expansions of the Ordered Field of Real Numbers by Restricted Pfaffian Functions and the Exponential Function*, J. AMS **9** (1996), pp. 1051-1094.

[W96b] A.J. Wilkie, *A General Theorem of the Complement and Some New o-Minimal Structures* , preprint, Oxford University, 1996.