Nonstandard Number Theory

Angus Macintyre*

0. Introduction. This paper is intended as an outline of a model-theoretic method in diophantine geometry. The foundations, and the most advanced development, are due to Robinson and Roquette [RR], [Roq 1]. The key idea is to relate the algebraic geometry of a number field $K$ to the nonstandard arithmetic of an enlargement $K^*$. For example, generic points acquire arithmetical structure. The main success of the method till now is a new proof of the Siegel–Mahler–Lang Theorem [RR], and new insights into the basic problem of effectiveness in that theorem [T]. The role of Hilbert’s Irreducibility Theorem in diophantine geometry is clarified [Roq 1].

I concentrate on the Robinson—Roquette formulation of Weil’s theory of distributions [W]. My only original contribution is a “covering theorem” relating geometric and arithmetical idèles. This gives another formulation of Weil’s theory, based on the heuristic principle that “adèles of $K$ become principal in $K^*$.”

1. Foundations.

1.1. Let Sets be the category of sets and functions. Let $I$ be any set, and $\text{Sets}^I$ the product (or functor) category. We have the “diagonal functor” $\Delta: \text{Sets} \rightarrow \text{Sets}^I$. Let $D$ be an ultrafilter on $I$. We have the “collapsing functor” $[D]: \text{Sets}^I \rightarrow \text{Sets}^I/D$.

---

* Partially supported by grant MCS77—07731 from the National Science Foundation of the United States. The paper was completed at the Mathematics Institute of Warsaw University. The author thanks the administration of that institution for their hospitality and financial support. Most of all, he thanks his Warsaw colleagues, especially Dr. Cecylia Rauszer, for their unfailing kindness.
Let \( \text{Sets}^* \) be the image of the collapsing functor. \( \text{Sets}^* \) is the nonstandard universe corresponding to \( D \). Let \( * : \text{Sets} \rightarrow \text{Sets}^* \) be \( \lceil [D] \rceil \circ \Delta \). I write \( x^* \) for \( *\langle x \rangle \).

Sets, \( \text{Sets}^I \) and \( \text{Sets}^* \) are topoi. I say this not to annoy, but simply because I believe that it is mathematically natural to have categorical foundations for nonstandard analysis. The tradition, deriving from Robinson, of converting \( \text{Sets}^* \) locally, to an \( \varepsilon \)-structure, leads to a gruesome formalism, is quite unnecessary in practice, and can even obscure correct arguments.

With no assumptions on \( D \), \( * \) is an elementary map [CK]. If \( D \) is suitably good [CK], \( \text{Sets}^* \) has the all-important property of \( \varkappa \)-saturation for some \( \varkappa \) (usually \( \varkappa = \omega_1 \) suffices). To say \( \text{Sets}^* \) is \( \varkappa \)-saturated means the following:

If \( A_\lambda \rightarrow A \) \((\lambda \in A)\) are morphisms, where \( \text{card} (A) < \varkappa \), and if for each finite subset \( A_0 \) of \( A \) there are morphisms \( 1 \rightarrow \psi_\lambda : A_0 \rightarrow A \) \((\lambda \in A_0)\) so that

\[
\begin{array}{ccc}
1 & \xrightarrow{\psi_\lambda, A_0} & A \\
\downarrow & & \downarrow \\
\psi_\mu, A_\mu & \xrightarrow{\psi_\mu, A_0} & A \\
\end{array}
\]

commutes for \( \lambda, \mu \) in \( A_0 \), then there are morphisms \( 1 \rightarrow A_\lambda \) \((\lambda \in A)\) so that

\[
\begin{array}{ccc}
1 & \xrightarrow{A_\lambda} & A \\
\downarrow & & \downarrow \\
A_\mu & \xrightarrow{A_\mu} & A \\
\end{array}
\]

commutes for \( \lambda, \mu \) in \( A \). (1 is the terminal object of \( \text{Sets}^* \).)

1.2. For reasons I do not clearly understand, it is often useful to work in the double enlargement \( \text{Sets}^{**} \). This is obtained as follows. The objects and morphisms of \( \text{Sets}^* \) are sets, so there is a well-defined category \( (\text{Sets}^*)^I \), obtained by forming \( (\text{Sets}^*)^I \) and then collapsing modulo \( D \). \( \text{Sets}^{**} \) is defined as \( (\text{Sets}^*)^* \), and \( ** : \text{Sets} \rightarrow \text{Sets}^{**} \) is the obvious composite

\[
\text{Set} \xrightarrow{*} \text{Sets}^* \xrightarrow{\text{new}*} (\text{Sets}^*)^*.
\]

2. Universal domains. 2.1. Let \( K \) be a number field, and \( K^* \) its enlargement. All one needs of \( * \) is \( \omega_1 \)-saturation. I now show \( K^* \) encodes many important infinitistic constructions over \( K \). In this sense \( K^* \) is a universal domain for \( K \). However, varieties over \( K \) with no points in \( K \) have no points in \( K^* \), so my sense of universal domain is not that of algebraic geometry.
2.2. Let $S_K$ be the set of those topologies $T$ on $K$ coming from a nontrivial absolute value $\| \cdot \| : K \to \mathbb{R}$. For $T \in S_K$, $T^*$ is a (base for a) topology on $K^*$, coming from a generalized absolute value $\| \cdot \|^* : K^* \to \mathbb{R}^*$. $(S_K)^*$ is a set of (bases for) topologies on $K^*$, coming from generalized absolute values $\| \| : K^* \to \mathbb{R}^*$. The standard elements of $(S_K)^*$ are those of the form $T^*$, $T \in S_K$.

(a) $B_T = \{ x \in \mathbb{R}^* : |x|^* < r \text{ some } r \text{ in } \mathbb{R} \}$;
(b) $I_T = \{ x \in \mathbb{R}^* : |x|^* < r \text{ all } r > 0 \text{ in } \mathbb{R} \}$.

The local compactness of $R$ implies that $B/I$ is canonically isomorphic to $R$. One has a retraction $R \to B/I \to R$ where the leftmost map is the natural inclusion, and the other is the fundamental standard part map $st_R$.

Now let $T \in S_K$. Consider

(a) $B_T = \{ x \in K^* : \| x \|^* \in B \}$;
(b) $I_T = \{ x \in K^* : \| x \|^* \in I \}$.

($\| \cdot \|$ is any absolute value for $T$.) Define $\| \cdot \|$ on $B_T/I_T$ by

$$\| x + I_T \| = st_R (\| x \|^*) .$$

Then

**Theorem 1** $B_T/I_T$ with $\| \cdot \|$ is canonically isometrically isomorphic to $K_T$, the completion of $K$ at $T$.

Next consider:

(a) $B_\infty = \{ x \in K^* : x \in B_T, \text{ all standard } T, \text{ and } x \text{ is in the unit ball of } T \text{ for all but finitely many standard } T \}$;
(b) $I_\infty = \bigcap_{T \text{ standard}} I_T$.

Then

**Theorem 2** $B_\infty/I_\infty$ is canonically $K$-isomorphic to $A_K$, the ring of $K$-adèles.

It is entirely routine to describe nonstandardly the adèlic topology and supplement Theorem 2. Since $A_K/K$ is compact [C] there is an adelic standard part map $st_{\text{ad}}$ giving a retraction

$$(A_K/K) \to (A_K/K)^* \xrightarrow{\text{st}_{\text{ad}}} (A_K/K)$$

$$(A_K)^*/K^*$$

3. **Product formula, idèles.** 3.1. Since Artín—Whaples one knows that the product formula is the axiom for number theory. To formulate it one needs a canonical choice $\| \cdot \|_T$ of absolute value for each $T$ in $S_K$. That choice is explained
measure-theoretically [C] and may also be explained in terms of nonstandard counting.

Let the canonical choice be made as in [C].

**Product formula.** If \( x \in K^K, \prod_{T \in S_K} \|x\|_T = 1. \)

This presupposes:

1st discreteness property. If \( x \in K^K, \) then \( \|x\|_T = 1 \) for all but finitely many \( T \) in \( S_K. \)

Progressively more subtle are:

2nd discreteness property. For each \( c > 0 \) in \( R \) there are only finitely many \( x \) in \( K \) such that \( \|x\|_T \leq c \) for all \( T \) in \( S_K. \)

3rd discreteness property (Roth's Theorem). Let \( S \) be a finite subset of \( S_K \) and let \( a_T (T \in S) \) be elements of \( K. \) Let \( k \in R, k > 2. \) Then there are only finitely many elements \( x \) in \( K \) satisfying

\[
\prod_{T \in S} \|x - a_T\|_T \leq \frac{1}{H(x)^k}
\]

(where \( H(x) = \prod_{T \in S_K} \max(1, \|x\|_T) \)).

These properties are exploited in Sets*, via the principle that a set finite in Sets has no nonstandard "members" in Sets*.

3.2. **Idèles.** The group \( J_K \) of idèles is the group of invertible elements of \( A_K \) topologized to make inversion continuous [C]. Given our discussion of \( A_K, \) it 'i routine to obtain \( J_K \) nonstandardly. Note that all \( \hat{A} \)-adèles become principal in \( K^* \)

For \( f \in J_K \) (classical definition) one defines

\[
c(f) = \prod_{T \in S_K} \|f(T)\|_T,
\]

and

\[
J^0_K = \{ f \in J_K : c(f) = 1 \}.
\]

One has:

**Compactness property:** \( K^K \) is discrete in \( J^0_K, \) and \( J^0_K/K^K \) is compact.

Whence, there is an idelic standard part map

\[
(J^0_K/K^K)^* \xrightarrow{\text{st}_{id}} J^0_K/K^K,
\]

so that \( \text{st}_{id}(\alpha) \) is idèlically-infinitesimally close to \( \alpha. \) Let

\[
F_K = \{ \alpha \in (J_K)^* : \|\alpha(T)\|_T \in B \setminus I, \text{ all } T \}.
\]

By unravelling the compactness conditions, one has

**Theorem 3 (Robinson—Roquette).**

\[
(J^0_K)^* = K^K \cap (J^0_K)^*,
\]

That is, elements of \( (J^0_K)^* \) are principal, modulo \( F_K. \)
Theorem 3 is a key result of nonstandard number theory. Let
\[ \hat{J}_{K^*} = (J_K)^*/F_K. \]
I refer to \( \hat{J}_{K^*} \) as the collapsed \( K^*-\)idèles.

3.3. The formulation in \( V^{**} \). In my approach to Theorem 3 I freely mixed standard and nonstandard methods. But there is a definite gain in being systematically nonstandard.

\( A_K \), an object in Sets, has been identified as \( B_{I^{I^J}} \). The latter is not an element of \((\text{Sets})^*\) (i.e. it is external) but it is an element of Sets, as is the ultrafilter \( D \) inducing \( * \). That is, \( B_{I^{I^J}} \) is an element of Sets defined in terms of \( D \). So one considers \( (B_{I^{I^J}})^* \), which is naturally isomorphic to \( B_{I^{I^J}}^*/I_{I^{I^J}}^* \), and I hope it is obvious that the latter is naturally isomorphic to \((A_K)^*\) (which I may write as \( A_{K^*} \) or \( A_K^* \)).

The appropriate picture is:

\[
\begin{array}{ccc}
\text{Sets} & \xrightarrow{\ast} & \text{Sets}^* \\
\uparrow & & \uparrow \\
K & \xrightarrow{\ast} & K^* \\
\downarrow & & \downarrow \\
B_{\infty} & \xrightarrow{\text{new}^*} & (B_{\infty})^*
\end{array}
\]

I must explain \((\ast)^*\) and \(\text{new}^*\). For the latter, see 1.2. \((\ast)^*\) is got by applying the functor \( \ast \) to the map \( \ast: K \rightarrow K^* \). It is to be stressed that \(\text{new}^* \neq (\ast)^*\). However, both maps are elementary.

4. Nonstandard theory of distributions. 4.1. Embedding function fields in \( K^* \). Let \( A \) be a variety (in affine \( n \)-space) defined over \( K \).

**Lemma 4.** Let \( M \) be any set. Then \( A \) has infinitely many points in \( K^n \cap M \) iff \( A \) has a nonstandard point in \( M^* \).

**Proof.** Trivial, by \( \omega_1 \)-saturation.

So we come to the (at first glance tenuous) connection between nonstandard analysis and diophantine geometry. For example, let \( S \) be a finite subset of \( S_K \), and define \( O_S \), the set of \( S \)-integers, as the set of those \( y \) in \( K \) such that \( \|y\|_T < 1 \) for all \( T \in S \). Let \( A \) be a curve over \( K \). Then:

**Theorem 5 (Siegel–Mahler–Lang).** Suppose \( A \) has genus \( \geq 1 \). Then \( A \) has only finitely many points in \( O_S^n \).

For a proof using algebraic geometry and Roth’s Theorem, see [L] or [S]–[D].

The connection with Lemma 7 is made by taking \( M=O_S^n \). One then has the immediate reformulation:
THEOREM 5*. Suppose $A$ has genus $\geq 1$. Then $A$ has no nonstandard point in $O^*_5$.

The nonstandard analysis of Theorem 5 looks at an arbitrary curve $A$ over $K$, and examines the consequences of the

Assumption. $A$ has a nonstandard point $\eta$ in $O^*_5$.

An important consequence, using the fact that $A$ has dimension 1, is that $\eta$ is a generic point of $A$. But this is a generic point with arithmetical structure. One seeks to prove Theorem 5 by confronting the geometry and the arithmetic of $\eta$.

Define $K(\eta)$ as the subfield of $K^*$ generated by the coordinates of $\eta$. Then of course $K(\eta)$ is the function field of $A$ over $K$. One now considers the general case

$$K \xrightarrow{\alpha} L \xrightarrow{\beta} K^*$$

where $L$ is a function field of one variable over $K$ and the diagram is a $K$-embedding.

$L$ has thus two divisor theories. The first is the classical "geometric" theory, as expounded in [Chev], which studies those absolute values $\| \cdot \|$ on $L$ which take the constant value 1 on $K^*$. The second is the restriction of the divisor theory on $K^*$ (based on $(S_K)^*$).

Let $S_L$ be the set of the above "geometric topologies" on $L$. I refer to [Chev] and [S] for all details on the product formula for $S_K$, and for the $L$-adèles $A_L$ and the $L$-idèles $J_L$.

4.2. Covering $S_L$ by $(S_K)^*$. Suppose $T \in (S_K)^*$. $\| \cdot \|_T$ restricts to a "generalized absolute value" $L \rightarrow R^*$. If $T$ is nonarchimedean, then the induced topology is actually given by a (Krull) valuation. If the valuation is trivial on $K^*$ then it is either trivial on $L$ or discrete on $L$ [Chev]. In the latter case the induced topology is in $S_L$. Our problem is how to handle those $T$ for which the valuation is not trivial on $K^*$. (There is also the problem of handling the finitely many archimedean $T$. This will be solved in passing.)

Let $T \in (S_K)^*$. Define

$$W_T(x) = -\log \|x\|_T \text{ for } x \in K^{**}.$$ 

Now collapse the additive group $R^*$ by the convex subgroup $B$, and consider

$$v_{T,B}(x) = W_T(x) + B \in R^*/B.$$

If we put $v_{T,B}(0) = \rightarrow R^*/B$, then we have:

**Lemma 6.** $v_{T,B}$ is a valuation on $K^*$, trivial on $K$.

$v_{T,B}$ induces a topology $T|B$ on $L$. Let

$$R_L = \{T|B : T \in (S_K)^*, T|B \text{ nontrivial on } L\}.$$

The next lemma is vital. It depends on the 2nd discreteness property, and the Riemann–Roch Theorem.

**Lemma 7.** $R_L = S_L$. 
This is the most that can be said in general. But if genus \((L) \gg 1\), much more is true, and this is precisely the Siegel–Mahler–Lang Theorem. Define
\[ R^*_L = \{T/B \in R_L : T \text{ nonstandard} \}. \]
Then Theorem 5 is equivalent to:

**Theorem 5**. \( R^*_L = S_L \) if genus \((L) \gg 1\).

Apparently all known proofs of Theorem 5 involve considering finite algebraic extensions \( K_1 \) of \( K \) and the corresponding extensions \( L_1 = K_1 \otimes_K L \) of \( L \). Each element of \( S_L \) lifts to finitely many elements of \( S_{L_1} \) (all conjugate if \( K_1 \) is normal over \( K \)). Of central importance is the induced conorm embedding \( A_L \to A_{L_1} \) (cf. [C]).

So it seems appropriate to construct a link
\[ A_L \to (A_K)^* \]
since we already have Lemma 7.

Further motivation, and the key in obtaining a correct version of (2), comes from:

**Lemma 8**. Let \( T_0 \in S_L \). Then \( \{T \in (S_K)^* : T/B = T_0\} \) is contained in a \(^*\)-finite set.

To get (2), one chases round:

Make the obvious definition of \( B_T, I_T \) for \( T \) in \( S_L \), and then put:
(a) \( B_{\infty,L} \) = \{\( x \in L^* : x \in B_T \), all standard \( T \), and \( x \) is in the unit ball of \( T \) for all but finitely many standard \( T \}\);
(b) \( I_{\infty,L} = \bigcap_{T \text{ standard}} I_T \).

There is a natural ring embedding (not an isomorphism!)
\[ A_L \to B_{\infty,L}/I_{\infty,L} \quad \text{(cf. 2)}. \]

The philosophy is that elements of \( A_L \) are made principal in \( L^* \).

Now embed \( B_{\infty,L} \) in \( K^{**} \) via \( \beta^* \). Make the following definitions:
(c) For \( T \) in \( (S_K)^{**} \),
\[ T^* = \{\|\beta^*(y^*)\|_T : y \in L\} \]
(i.e. set of values of \( \| \cdot \|_T \) on \( L \));
(d) For \( T \) in \( (S_K)^{**} \),
\[ B^U_T = \{x \in K^{**} : \|x\|_T \text{ is bounded above by an element of } B^* \cdot T\} \];
(e) $B_T = \{ x \in K^* : x \in B_T \}$, all $T$ in $(S_K)^*$, and $x \in B_T$ for all but $\ast$-finitely many $T$ in $(S_K)^*$;
(f) For $T$ in $(S_K)^*$, $I_T = \text{set of non units of } B_T$;
(g) $I_\infty = \bigcap_{T \in (S_K)^*} I_T$.

The following is obtained by elementary considerations, unpacking definitions in the covering mechanism:

**Lemma 9.** Via $\beta^*$, $B_{\infty,L}$ is sent to $B_{\infty}^{(L)}$, and $I_{\infty,L}$ is sent to $I_{\infty}^{(L)}$. So $\beta^*$ induces an embedding

$$B_{\infty,L}/I_{\infty,L} \rightarrow B_{\infty}^{(L)}/I_{\infty}^{(L)}.$$ 

This is the best we can do towards (2). $B_{\infty,L}/I_{\infty,L}$ naturally contains $A_{L}$. And, clearly, $B_{\infty}^{(L)}/I_{\infty}^{(L)}$ is similar to $(A_K)^*$, with a distortion factor specific to $L$.

It will be shown in detail in another publication that Lemma 9 covers the Robinson–Roquette version (in terms of nonstandard divisors) of Weil’s theory of distributions.

One formulates an idèlic version of Lemma 9, and obtains a commuting diagram

$$\begin{array}{ccc}
J_{\infty,L} & \longrightarrow & J_{\infty}^{(L)} \\
\downarrow & & \downarrow \\
D_L & \longrightarrow & (D_K)^* \text{ (finite divisors)}
\end{array}$$

The left map is the natural divisor map.
The top map is induced by $\beta^*$.
The bottom map is that of Robinson–Roquette.
The right map is difficult to construct except by strict use of functoriality.

**5. Concluding the proof of Theorem 5**. 5.1. The main idea is to exploit the strange analogy between $L \rightarrow K^*$ and a finite algebraic extension $L \rightarrow L_1$. One considers possible counterexamples to Theorem 5 (so called exceptional primes) and exceptional divisors which are sums of distinct exceptional primes. Using the above analogy, together with Roth’s Theorem, one proves

$$\text{deg}(A) < 2 \cdot [L : K(x)]$$

for $A$ exceptional and $x \in L \setminus K$.

5.2. One then deduces that $\deg(A) = 0$, if genus $(L) > 1$. The strategy of Robinson–Roquette was to find finite unramified extensions $L'$ of $L$ with commuting diagram

$$\begin{array}{ccc}
L & \longrightarrow & L' \\
\beta & \swarrow & \searrow \\
& K^* & \\
& \beta &
\end{array}$$

For a function field $M$ over $K$, define

$$d_M = \min_{x \in M \setminus K} [M : K(x)].$$
Then by elementary functorial arguments one shows that $\deg (A) \leq 2d_L[L': L]$. So one wants to find $L'$ with $d_L \leq (1/2)[L': L]$. By an elementary, but lengthy, argument in [Roq 2] Roquette showed that if genus $(L) \geq 1$ then for sufficiently large $n$ the maximal unramified semiabelian extension of $L$ of exponent $n$ satisfies this inequality. Finally, by using Theorem 3, one shows that these semiabelian extensions are $L$-embeddable into $K^*$. This concludes the proof.

6. The connection with Hilbert's irreducibility theorem. In the last part of their proof Robinson and Roquette are treating special cases of the following.

**Problem.** Given $L \to K^*$ as above, what is the structure of $\text{Alg} (L, K^*)$, the lattice of intermediate fields finite-dimensional over $K$?

By a result of Gilmore–Robinson [Roq 1] Hilbert's Irreducibility Theorem for any $K$ (not necessarily a number field) is equivalent to the existence of a pure transcendental $L=K(x)$ with $\text{Alg} (L, K^*)$ empty. In a beautiful paper [Roq 1] Roquette shows that this is a useful result. Yet again one admires Robinson's vision, that glimpsed these deep ideas over twenty years ago.

Another interesting possibility for $\text{Alg} (L, K^*)$, $L=K(x)$, is given by Roquette [Roq 1]. Namely, there may be exactly one extension of each degree $n$. This is closely connected with the nonstandard analysis of curves of genus 0 [RR].

The general problem must be of extreme difficulty. For, if there is an $L$ so that $L$ has an unramified abelian extension of exponent 2 and dimension 8 in $K^*$, then Mordell's Conjecture is false, and conversely.

7. Effective estimates. An outstanding problem of number theory is to find effective bounds for the $S$-integral points on a curve $C$ of genus $\geq 1$. It is widely conjectured that such a method exists. For integer points, and curves over $\mathbb{Q}$ of genus 1, Baker [B] found such a method.

Inspection of the classical proof [L] reveals two nodes of noneffectiveness. One is the use of Roth's Theorem, where no effective bounds are presently known. The second is the use of generators for the Mordell–Weil group of the Jacobian. Despite extensive scrutiny of the proof, and advice from leading authorities, I do not see how to eliminate either ineffectiveness.

An astonishing feature of the Robinson–Roquette proof is that there is no use of the generators of the Mordell–Weil group (although there are manoeuvres in the proof reminiscent of the proof of the Mordell–Weil Theorem). Robinson's last mathematical achievement was to see that the new proof reveals that there is an effective bound in the Siegel–Mahler–Lang Theorem, relative to effective bounds in Roth's Theorem. This is by no means evident from the Robinson–Roquette proof (still less from mine), but a careful axiomatization of what is used in the proof will yield the result. Takeuti[T] verified this claim of Robinson, by showing that the proof could be done in a suitable fragment of nonstandard analysis.

[The use of model theory to give effective estimates in algebra is well established. See [R] for the first ideas, and [D] for some nice new ones.]
As a matter of fact, number theorists tend to dismiss this relative recursive estimate. I do not understand why. If they are to be taken seriously in conjecturing an absolute effective estimate, why do they not provide, by approved methods, an elimination of the Mordell–Weil Theorem from the classical proof of Siegel–Mahler–Lang?

Postscript. This paper was completed in a mood of deep dismay, after the tragic death of my friend and former student George Loullis. I dedicate this work to his memory, in fond remembrance of his cheerful presence at work and at play.

References


YALE UNIVERSITY
NEW HAVEN, CONNECTICUT 06520, U.S.A.

UNIVERSITY OF WARSAW
WARSAW, POLAND