# Hilbert's Tenth Problem and Elliptic Curves

Barry Mazur

February 1, 2010

*(Notes for my Basic Notions talk—Feb. 1 2010)*

## Part I: Hilbert's Tenth Problem

Write down a random polynomial equation in two or more variables with coefficients in the ring of integers, e.g.,

$$3X^3 + 4Y^3 + 5Z^3 = 0$$

and—chances are— it will be very tricky to find all its solutions; often you will be quite challenged by the question of whether or not it *has* solutions.

Hilbert spurred mathematicians to systematically investigate the general question:

*How solvable are such Diophantine equations?*

I will talk about this, and its relevance to specific number theoretic projects, and then aim towards some recent work, joint with Karl Rubin.

Here is a close translation of Hilbert's formulation of the problem:

> Given a Diophantine equation with any number of unknown quantities and with rational integral numerical coefficients: To devise a process according to which it can be determined in a finite number of operations whether the equation is solvable in rational integers.

.

I wonder what Hilbert meant by *a Diophantine equation* and *process*. There is a vagueness in the quantification: are we allowed a different algorithm for each equation, are we expected to find single processes that work even allowing for variation of the exponents involved? (As in Catalan's Problem, or Fermat's Last Theorem). There is a specificity, though, in the type of solutions (rational integers). Nowadays one has a large number of different *processes* in our experience (i.e., *successes*). From algorithms to find the maxima of functions on convex polytopes (e.g.: Linear programming) to procedures for factoring numbers into product of primes. The basic questions we tend to ask about these have to do with running time.

We also have quite a number of guaranteed non-successes:

- There is no finite algorithm to determine, given a finite presentation of a group, whether or not the group is trivial. Or whether two finite presentations present isomorphic groups.

- The *recognition problem for manifolds in dimension four or higher* is unsolvable (it being related directly to the recognition problem for finitely presented groups).

And even when one looks for interesting Diophantine examples, they often come in formats somewhat different from the way Hilbert's Problem is posed. For example,

- we have a (deep) decision procedure to determine whether any given elliptic curve over the rational field $\mathbf{Q}$ has *finitely many* or *infinitely many* solutions. But this distinction

  $$\textit{finitely many} \leftrightarrow \textit{infinitely many}$$

  is not a distinction that Hilbert formulates.

- And, sometimes, we're interested not in answering this question for any single elliptic curve but, for whole families of them.

  For example, the *congruent number problem* is the problem of determining those positive integers $n$ that can be expressed as the area of a right triangle with three rational number sides. This turns out to be *equivalent* to asking that the elliptic curve

  $$y^2 = x^3 - n^2x$$

  have infinitely many rational points.

- And, as mentioned above, we sometimes try to find single processes that work even allowing for variation of the exponents involved. As in:

  1. *Catalan-type Problems*

     For a given integer $k$ find all *perfect powers* that differ by $k$.

     $$Y^n - X^m \quad = \quad k$$

     **Example:** *the only two consecutive perfect powers are:*

     $$8 = 2^3 \text{ and } 9 = 3^2,$$

     or as in:

  2. *Fermat's Last Theorem.*

So you might ask why—except for historical reasons—might one be interested in pursuing the question as Hilbert posed it. The answer (which is already enough to spark my interest) is that it is a problem that has led to the most magnificent developments in mathematical logic, and in the intersection of mathematical logic and number theory. But also, thanks to relatively recent work (of Denef, Denef-Lipschitz, Pheidas, Shalapentokh and Poonen) Hilbert's Problem calls for the answers to new *kinds of* questions in number theory, and specifically in the arithmetic of elliptic curves.

*So, back to Hilbert's Tenth Problem!*

In considering "Hilbert's 10th Problem" we often specifically interpret *Diophantine equation, process* and sometimes generalize the type of solutions being considered. We then end up with a question roughly of the following form:

> Let $A$ be a commutative ring. Does there exist a finite algorithm to determine whether any finite system of polynomial equations in finitely many variables with coefficients in $A$ has a solution in $A$ or not?

One standard way of refining the above question is to "set" it as a problem related to *listable* sets and *Diophantine sets*. The next two sections discuss this.

# 1 Listable sets of integers

I'll start with some examples of sets that are easy to "list"

- $$2, 3, 5, 7, 11, 13, 17, 19, 23, \ldots$$

- $$2!, 3!, 4!, 5!, \ldots$$

  Discuss what is meant by *easy*.

  See [M] and [Sh100b] for good expository accounts of this notion *listable*. Naively,

- a *listable* subset of $\mathbf{Z}$ (synonyms: *recursively enumerable, computably enumerable*) is a subset $\mathcal{L} \subset \mathbf{Z}$ for which there exists a finite computer program whose output gives a sequence $\alpha_1, \alpha_2, \alpha_3, \ldots$ of integers such that the set $\mathcal{L}$ is precisely this collection of numbers; i.e.,

  $$\mathcal{L} = \{\alpha_1, \alpha_2, \alpha_3 \ldots\}.$$

  A computer algorithm that does job this will be called a computer algorithm that "lists $\mathcal{L}$."

  Note, though, that the ordering in which the integers in $\mathcal{L}$ come via the computer's list may be helter-skelter in terms of absolute values. Therefore if you suspect that a given number, say 2, is *not* in $\mathcal{L}$ and need to have a definite guarantee of the truth of your suspicion, well (... if you are right!) running the helter-skelter computer algorithm for any finite length of time will be of no help to you.

- A more useful finite computer program might be, for example, a program that for each integer $N$ will, after some guaranteed time (e.g., no greater than $N^{N^{\cdots^N}}$ hours[1]) actually produces a *complete* list of *all* integers of absolute value $\leq N$ that are in $\mathcal{L}$. (Call such a program a *deluxe program*.)

- Somewhat intermediary to the above two types of computer programs (helter-skelter, and deluxe) would be a *pair* of computer programs, one of which spits out the elements of $\mathcal{L}$ and the other spits of the elements of the complement of $\mathcal{L}$. Supplied with such a pair of programs you might, at the very least, run the first program by day, and the second by night, for then you are guaranteed to know—in some (perhaps unspecified, but) finite time whether or not 2 is in your set $\mathcal{L}$.

# 2 The Halting Problem

As mentioned, a set $\mathcal{L}$ that is listable by a finite computer algorithm will be referred to as *listable* or *recursively enumerable*. And a set $\mathcal{L}$ that has the property that it and its complement are both listable is called *recursive*. What we will be using below is the fact (cf. [Sm]) that there exist recursively enumerable sets that are *not* recursive. (The computer algorithms that list such sets are necessarily quite helter-skelter!) The existence of recursively enumerable sets that are not recursive is a consequence of the famous 1936 theorem of Alan Turing that was phrased in terms of the *halting problem for algorithms*. Turing showed that there exists no universal algorithm to tell you whether or not any finite computer algorithm will terminate finitely, when run. More specifically, the so-called *halting set*

$$\mathbf{H} := \{\text{The set of couples } (P, x) \text{ where } P \text{ is a program and } x \text{ is a possible input to program } P \text{ and}$$

[1] $N$ successive exponentials, or choose any recursively formulatable estimate you like

such that *Program* $P$ will eventually *halt* if run with input $x$}

is recursively enumerable, (i.e., there is fairly evidently a computable function that lists all of the pairs $(P, x)$ it contains) but the complement of this set is not recursively enumerable.

# 3 Diophantine sets

Roughly, a *Diophantine subset* of integers (or of natural numbers) is a subset that can be defined using the *seemingly very restricted* vocabulary of polynomials. For the classical notion of Diophantine subset of a commutative ring see **[DL78], [Den80]**. Here is one way of formulating this concept. Let $A$ be a commutative noetherian integral domain, the main example being $A = \mathbf{Z}$.

**Definition 3.1** *Let $\mathcal{D} \subset A$ be a subset of the ring $A$.*
*Say that $\mathcal{D}$ is* **Diophantine in** *$A$ if there exists a finite set of polynomials with coefficients in $A$, in finitely many variables*

$$f_i(T; X_1, X_2, \ldots, X_n) \in A[T, X_1, X_2, \ldots, X_n]$$

*$(i = 1, 2, \ldots, m)$ such that when specializing to some value $T = \alpha \in A$ we have that the system of polynomial equations*

$$f_i(\alpha; X_1, X_2, \ldots, X_n) = 0$$

*(for $i = 1, 2, \ldots, m$) has a simultaneous solution*

$$(X_1, X_2, \ldots, X_n) = (a_1, a_2, \ldots, a_n) \in A^n$$

*if and only if*

$$\alpha \in \mathcal{D} \subset A.$$

*If this happens say that the set of polynomials* **cut out** *$\mathcal{D}$.*

Notice the evident proposition:

**Proposition 3.2** *If $A = \mathbf{Z}$ (or, more, generally, a countable ring) and $\mathcal{D} \subset A$ is Diophantine, then $\mathcal{D}$ is listable. Moreover, any set of polynomials $\{f_i(\alpha; X_1, X_2, \ldots, X_n)\}_i$ (for $i = 1, 2, \ldots, m$) that "cut out" $\mathcal{L}$ leads to a computer algorithm that lists $\mathcal{L}$.*

**Proof:** Choose some ordering of $A^{n+1}$ (e.g., lexicographical based on an ordering of $A$; if $A = \mathbf{Z}$ my preference is for the evident ordering of $\mathbf{Z}$: that is, $-n$ precedes $+n$ and otherwise it is nondecreasing in the absolute value of $n$) and run through the $n+1$-tuples $(\alpha; a_1, a_2, \ldots, a_n) \in A^{n+1}$ computing $f_i(\alpha; a_1, a_2, \ldots, a_n)$ for $i = 1, 2, \ldots$ : every time you get a *hit*—i.e., every time that $f_i(\alpha; a_1, a_2, \ldots, a_n) = 0$ for $i = 1, 2, \ldots$ you record the "$\alpha$" if it hasn't been previously recorded giving a (possibly empty, of course) sequence $\alpha_1, \alpha_2, \ldots$ listing $\mathcal{D}$.

**Remarks: (1)** The collection of Diophantine subsets of an integral domain $A$ is closed under finite union and intersection.

**Proof:** It suffices to do this for two Diophantine sets $D, E \subset A$: if the systems of polynomials $\{f_i(t; X_1, \ldots)\}_i$ and $\{g_j(t; Y_1, \ldots)\}_j$ cut out $D$ and $E$ respectively, then the "union" of the two systems, (viewed as polynomials in $t$ and the independent variables $X_\mu$ and $Y_\nu$) cuts out $D \cap E$ while the system given by $\{f_i(t; X_1, \ldots) \cdot g_j(t; Y_1, \ldots)\}_{i,j}$ cuts out $DS \cup E$.

**(2)** A more general (and, perhaps, algebro-geometrically more natural) way of thinking of Diophantine set is the following:

Let $S$ be an integral noetherian scheme—say an affine scheme $S = \mathrm{Spec}(A)$ where $A$ is a noetherian integral domain— and $T$ an $S$-scheme of finite type. Let $\mathcal{T} = T(S)$ the set of $S$-valued points of the $S$-scheme $T$. A subset $\mathcal{D} \subset \mathcal{T}$ is **Diophantine** if there is a morphism of $S$-schemes of finite type $f : X \to T$ such that

$$\mathcal{D} = f(X(S)) \subset T^(S) = \mathcal{T}.$$

To relate the above to the previous definition let $S = \mathrm{Spec}(A)$ and let $T = \mathrm{Spec}(A[t])$ denote the affine line over $\mathrm{Spec}(A)$. So the set $\mathcal{T}$ of $A$-rational points of $T$, i.e.,

$$\mathcal{T} = T(A) = \mathrm{Hom}_A(A[t], A).$$

is simply the set $A$. Diophantine subsets of the ring $A$ are nothing more than the images of the sets of $A$-rational points,

$$X(A) \longrightarrow T(A) = A,$$

where $X \to T$ range through all morphisms of finite type of $A$-schemes (of finite type) $X$.

A vague, but general question, then for any scheme $T$ of finite type over such a base $S$ would be:

*To give a useful algorithmic characterization of the subsets $\mathcal{D} \subset \mathcal{T}$ that are Diophantine.*

**(3)** For us the most important ring is $A = \mathbf{Z}$, and scheme $T$ is the affine line. In this context you can replace any finite system of polynomials $\{f_i(t; X_1, \dots)\}_i$ that "cut out" a set $D$ by a single polynomial

$$\Sigma_i f_i(t; X_1, \dots)^2.$$

Here is a list of subsets of $\mathbf{Z}$ that are Diophantine (and easily proven to be).

1. Lagrange proved that any positive whole number is expressible as a sum of four squares.

   E.g
   $$401 = (20)^2 + 1^2 + 0^2 + 0^2$$

   well ... that might have been too easy an example ...

   Lagrange's Theorem says, in our vocabulary, that the polynomial

   $$f(t; X_1, X_2, X_3, X_4) := \quad t - \Sigma_{i=1}^4 X_j^2$$

   cuts out the set of positive integers; so the set of positive numbers is Diophantine.

2. Therefore it follows, by easy steps, that these sets are too:

   - the set of numbers $\geq a$ for any given $a \in \mathbf{Z}$,
   - the set of numbers $\leq b$ for any given $b \in \mathbf{Z}$,
   - any finite subset of $\mathbf{Z}$,
   - the complement of any finite subset of $\mathbf{Z}$.

3. So, if $D$ is Diophantine, then any set obtained from $D$ by removing and adding finite sets is also Diophantine.

4. Arithmetic progressions are Diophantine; as are the set of all squares, all cubes, all $n$-th powers for any given $n$.

5. Composite numbers.

6. For any fixed (say, nonsquare, positive) integer $d$, consider the set of integers $t$ that come in solutions of the Pell equation
$$t^2 - ds^2 = 1$$
(this being a set that grows roughly exponentially).

# 4 Davis/(Julia) Robinson/Putnam/ Matiyasevich

Here, in a nutshell, is the general status of this question we inherited from Hilbert and from "classical work" of Martin Davis, Julia Robinson, Hillary Putnam and Yuri Matiyasevich. The culminating theorem is Matiyasevich's:

**Theorem 4.1** *Every listable subset of* $\mathbf{Z}$ *is Diophantine.*

Thus *listable* and *Diophantine* are equivalent conditions for subsets of $\mathbf{Z}$. Since there exist listable subsets of $\mathbf{Z}$ that are not recursive—i.., such that their complements are *not listable*, Theorem 4.1 gives a negative answer to Hilbert's question above, but does far more than just that.

For example:

1. It certainly shows that there are systems of polynomials over $\mathbf{Z}$ that admit no "deluxe computer program" as described.

2. The result also implies that relatively benign subsets of $\mathbf{Z}$ can be Diophantinely described, as well. This is not as clear as one might think even for the most familiar subsets. For example:

   - There is a system of polynomials that cut out the set of factorials $1!, 2!, 3! \ldots$ The fact that this set is Diophantine played a big role in the development of the subject[2].
     The *factorial* operation has quite a powerful effect if one allows it to be used as a piece of equipment to generate recursively enumerable sets. For example, the set of positive numbers $\alpha$ such that the expression
     $$(\alpha + 1) \cdot X_1 + \alpha! \cdot X_2 = 1$$
     has a zero for integers $X_1, X_2$ is precisely the set of prime numbers. But regarding prime numbers, more relevant for our story is the fact that...
   - There is a polynomial over $\mathbf{Z}$ whose set of positive values is the set of *exactly all* prime numbers for integral substitution of its variables. A specific such polynomial[3] is given in **[JSWW76]**:

     $(k+2)\{1-[wz+h+jq]^2-[(gk+2g+k+1)(h+j)+hz]^2-[2n+p+q+ze]^2[16(k+1)^3(k+2)(n+1)^2+1f^2]^2-[e^3(e+2)(a+1)^2+1o^2]^2-[(a^2 1)y^2+1-x^2]^2 x-[16r^2y^4(a^2-1)+1-u^2]^2-[((a+u^2(u^2-a))^2-1)(n+4dy)^2+1-(x+cu)^2]^2-[n+l+vy]^2-[(a^2-1)l^2+1-m^2]^2-[ai+k+1-l-i]^2-[p+l(a-n-1)+b(2an+2a-n^2-2n-2)-m]^2-[q+y(a-p-1)+s(2ap+2a-p^2-2p-2)-x]^2-[z+pl(a-p)+t(2ap-p^2-1)-pm]^2\}$

[2]To get such a polynomial one starts by finding a Diophantine way of expressing the binomial coefficients $\binom{n}{m}$ and then dealing with the—to me surprisingly unpromising—formula
$$m! = \lim_{n\to\infty} \frac{n^m}{\binom{n}{m}}.$$

[3]As you'll see from its equation, this is not the most efficacious way of finding prime numbers, but ...

# 5 Comments on the History

For a historical (and basic mathematical) account of this work it would be difficult, I think, to do better than the very informative wikipedia entry on *Hilbert's tenth Problem* which has a chart listing work ranging from 1944 when Emil Post said that Hilbert's tenth problem "begs for an unsolvability proof" to 1970 when Matijasevic clinched the theorem.

On the way to the final formulation of the theorem there is Martin Davis's formulation of what I'll, for reference, call **Davis Sets**, these being sets of natural numbers $\Delta$ such that there exists a polynomial with integral coefficients

$$P(T, K, Y, X_1, X_2, \ldots, X_n)$$

in independent variables $T, K, Y, X_1, X_2, \ldots, X_n$ for some $n$, such that $a \in \mathcal{A}$ if and only if there is a non-negative integer $y$ such that for all nonnegative integers $k < y$ the polynomial $P(a, k, y, X_1, X_2, \ldots, X_n)$ has a solution in natural numbers $X_1 = a_1, X_2 = a_2, \ldots, X_n = a_n$. Now *Davis sets* are fairly clearly recursively enumerable. In 1949 Davis proved the converse: that every recursively enumerable subset of the set of natural numbers has the above form; i.e., is *Davis*.

A year later, working independently, Julia Robinson formulated her hypothesis that asserts that—roughly speaking—there exists some function

$$\text{“}Exp:\text{”}\mathbf{N} \quad \longrightarrow \quad \mathbf{N}$$

that behaves at least vaguely like an exponential function and whose graph is Diophantine (a *sloppy exponential* would be enough). Her hypothesis that came to be known as "J.R." and explicitly is:

*Hypothesis J.R.:* There exists a Diophantine set $\mathcal{F}$ of couples $(a, b)$ in $\mathbf{N} \times \mathbf{N}$ with two properties:

(a) If $(a, b) \in \mathcal{F}$ then $a < b^b$.

(b) For each positive integer $k$ there is an $(a, b) \in \mathcal{D}$ with $b > a^k$.

Using hypothesis J.R., Robinson shows that the set EXP of triples $(a, b, c)$ with $a = b^c$ is Diophantine, and from this that the set of primes, and the set of factorials is Diophantine as well.

In 1959 Martin Davis and Hillary Putnam showed—assuming that there were arbitrarily long arithmetic progressions of prime numbers—that Hypothesis J.R. implies the equivalence of Diophantine and recursively enumerable, and thereby conditionally establishing a solution to Hilbert's Tenth Problem (the "conditions" being *the existence of arbitrarily long arithmetic progressions of primes,* and *J.R.*).

A year later, Robinson showed how to avoid the use of the hypothesis that arbitrarily long arithmetic progressions of primes exist, thereby showing that J.R. alone implies a solution to Hilbert's Tenth Problem.

In 1970, Matiyasevich provides a Diophantine definition of a set $\mathcal{F}$ as required by J.R.: he defined his $\mathcal{F}$ to be the collection of pairs $(a, b)$ such that

$$b = F_{2a}$$

where $F_n$ is the $n$th Fibonacci number, thereby completing the proof that all recursively enumerable sets are Diophantine and establishing the fact that Hilbert's tenth problem (over $\mathbf{Z}$) is unsolvable.

# 6 Some comments by Matiyasevich

(I find this quotation of Matiyasevich illuminating:)

> "The idea was as follows. A universal computer science tool for representing information uses words rather than numbers. However, there are many ways to represent words by numbers. One such method is naturally related to Diophantine equations. Namely, it is not difficult to show that every $2 \times 2$ matrix
>
> $$\begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}$$
>
> with the $m$'s being non-negative integers and the determinant $m_{11}m_{22} - m_{12}m_{21}$ equal to 1 can be represented, in a unique way, as a product of matrices
>
> $$M_0 := \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$
>
> and
>
> $$M_1 := \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$
>
> It is evident that any product of such matrices has non-negative integer elements and the determinant equals 1. This implies that we can uniquely represent a word in the two-letter alphabet $M_0, M_1$ by the four-tuple
>
> $$(m_{11}, m_{12}, m_{21}, m_{22})$$
>
> such that the numbers evidently satisfy the Diophantine equation
>
> $$m_{11}m_{22} - m_{12}m_{21} \;=\; 1.$$
>
> Under this representation of words by matrices, the operation of concatenation-of-words corresponds to matrix multiplication and thus can be easily expressed as a system of Diophantine equations, opening up a way of transforming an arbitrary system of word equations into "equivalent" Diophantine equations. Many decision problems about words had been shown undecidable, so it was quite natural to try to attack Hilbert's tenth problem by proving the undecidability of systems of word equations.
>
> My next attempt was to consider a broader class of word equations with additional predicates. Since the ultimate goal was always Hilbert's tenth problem, I could consider only such predicates, which (under suitable coding) would be represented by Diophantine equations. In this way I came to what I have called *equations in words and length*. Reduction of such equations was based on the celebrated Fibonacci numbers. It is well known that every natural number can be represented, in an almost unique way, as the sum of different Fibonacci numbers, none of which are consecutive[4] (this is the so called *Zeckendorf representation*). Thus we can look at natural numbers as words in a two-letter alphabet $\{0, 1\}$ with the additional constraint that there cannot be two consecutive 1's. I managed to show that under this representation of words by numbers both the concatenation of words and the equality of the length of two words can be expressed by Diophantine equations."

[4] E.g., $30 = 1 + 8 + 21$.

# 7 Unsolvability over the rings of integers of number fields

More recent work (Denef/Denef-Lipschitz/Pheidas/Shalpentokh/Poonen) developed ideas that culminated in the following result:

**Theorem 7.1** *Let $K$ be a number field and let $\mathcal{O}_K$ be the ring of integers in $K$.*

*If a certain stability result in the arithmetic of elliptic curves that we'll be discussing in detail in the next section holds*[5] *over $K$, then every recursively enumerable subset of $\mathcal{O}_K$ is Diophantine in $\mathcal{O}_K$ (in the sense of Definition 3.1 above).*

Karl Rubin and I have recently shown that this stability result holds *if you assume the* 2*-primary part of the classical Shafarevich-Tate Conjecture* [**MR09**]. As a consequence we have shown that, conditional on the 2-primary part of the Shafarevich-Tate Conjecture, Hilbert's Tenth problem has a negative answer for the ring of integers in *any* number field.

Since Kirsten Eisenträger has, in her thesis, related Hilbert's Tenth Problem over rings of integers in number fields to a much more general class of rings, one gets—thanks to her work:

**Theorem 7.2** *Conditional on the* 2*-primary part of the Shafarevich-Tate Conjecture, Hilbert's Tenth problem has a negative answer for any commutative ring $A$ that is of infinite cardinality, and is finitely generated over* $\mathbf{Z}$.

# 8 Integral points on plane curves of genus one

Here one has a striking explicit result, thanks to Baker's method. Let $f(X_1, X_2) \in \mathbf{Z}[X_1, X_2]$ be an absolutely irreducible polynomial such that the associated projective curve $f = 0$ has genus one. Let $n :=$ the (total) degree of $f(X_1, X_2)$ and let $H :=$ the maximum of the (ordinary) absolute values of the coefficients of $f(X_1, X_2)$. Then there are finitely many integral solutions $(a_1, a_2)$ of the equation $f(X_1, X_2) = 0$ and they are bounded explicitly by the inequality

$$\max\{|a_1|, |a_2|\} \ < \ \exp\exp\exp\{2H^{10n^{10}}\}.$$

For discussion about this, see section 4.4 of [**B75**].

# 9 The focus on cubics!

Here is a list of contexts for which there *are* algorithms to decide whether or not solutions exist over $\mathbf{Z}$:

- Systems of arbitrarily many polynomials in arbitrary degrees, but in *one* variable
- Systems of arbitrarily many linear polynomials in *many* variables
- A *single* quadratic polynomial in many variables

Here are contexts for which we know there is *no* algorithm to decide whether or not solutions exist over $\mathbf{Z}$.There is no algorithm that works for every:

[5] Specifically the *stability result* asserts that for every prime degree Galois extension of number fields $L/K$ there exists an elliptic curve $E$ over $K$ with

$$\mathrm{rank}E(K) = \mathrm{rank}E(L) > 0.$$

- system of arbitrarily many polynomials in many variables of degree $\leq 2$
- *single* polynomial of degree 4.

Some of the big unresolved questions for systems of polynomials with coefficients in $\mathbf{Z}$, then, focus on degree three,[6] and—for example—include:

1. Is there an algorithm to determine whether any given single degree three polynomial in many variables has a solution or not?

2. Is there an algorithm to determine *all* solutions of a single cubic homogenous polynomial in two variables?

I know what I "want the answer to be" (i.e., I want there to be such an algorithm: why not?) but I can't give any compelling reason for my optimism at least for the first of these questions. For the second question, the answer would be *yes* if some standard conjectures are true.

A proof of the Shafarevich-Tate Conjecture[7] would provide a proof that a certain algorithm works for general third degree polynomials $F(X, Y)$ and—more generally—finds rational points on curves of genus one. The algorithm itself is currently known, and used quite extensively. If it (always) works, then it gives an answer to the question posed above, and indeed allows us to find the rational points. But we don't know whether it will always terminate (in finite time) to provide us with an answer. The Shafarevich-Tate Conjecture would guarantee termination in finite time. This is a huge subject (the arithmetic theory of elliptic curves) and it would be good to understand it as well as we can possibly understand it. Note the curious irony in the formulation of Theorem 7.2:

> If we have a proof of the (2-primary part of the) Shafarevich-Tate conjecture
>
> *—i.e., colloquially speaking: if the algorithm that enables us to deal with arithmetic of cubic plane curves can be* **proved** *to work—*
>
> then we have a proof of the **non**-existence of a general algorithm for the ring of integers over any number field.

# 10 An Open Question

Gerald Sacks once suggested that solvability or unsolvability may be only one of a number of different ways of framing questions regarding Diophantine algorithms. I agree, and have always liked Serge Lang's attitude towards these matters, who—in effect—focussed much attention to the question of determining whether

[6] It is interesting how our lack of understanding of cubics seems to color lots of mathematics, from the ancient concerns in the "one-variable case" having to do with "two mean proportionals," and Archimedes' Prop.4 of Book II of *The Sphere and Cylinder* and Eutocius' commentaries on this, and–of course—the Italian 16c early algebraists.

[7] In fact, just, the $p$-primary part of the Shafarevich-Tate Conjecture, for any single prime number $p$ will do it.

there are finitely many, as opposed to infinitely many, solutions [8] and asked algebro-geometric questions about structure of the infinitely many solutions when they exist.

In this spirit allow me to formulate a question—without prejudice—that seems worth contemplating even if it is a bit premature to try to make much headway with it.

If $V$ is an algebraic variety over $\mathbf{Q}$ let $X(V;\mathbf{Q}) \subset V$ be the Zariski closure in $V$ of the set $V(\mathbf{Q})$ of $\mathbf{Q}$-rational points of $V$.

Define $D(V) = D(V;\mathbf{Q}) :=$ the number of irreducible components of $X(V;\mathbf{Q})$.

Suppose that we set out to find upper bounds for this function from algebraic varieties to natural numbers:

$$V \quad \mapsto \quad D(V).$$

Consider, for example, the case where $V$ is an irreducible curve.

- If $V$ is of genus 0, then $D(V)$ is either 0 or 1 depending upon whether $V$ has a rational point or not.

- If $V$ is of genus one, then
  - $D(V)$ is 0 if $V$ has no rational points,
  - $D(V)$ is 1 if $V$ has infinitely many rational points, and
  - $D(V)$ is the order of the Mordell-Weil group of $V$ over $\mathbf{Q}$, if that group is finite.

  In all cases for $V$ of genus one, then, (using **[M77]**) we get that $D(V) \leq 16$.

- If $V$ is of genus $> 1$, by Faltings' Theorem $D(V)$ is the (finite) number of rational points of $V$. Conditional on a conjecture of Lang, Caparoso, Harris and I have shown that $D(V)$ is bounded by a function that depends only on the genus of $V$.

In sum, we have that (conditional on a conjecture of Lang) for all algebraic varieties $V$ of dimension one,

> $D(V)$ is bounded from above by a function $F(|V_{\mathbf{C}}|)$ that depends only on the homotopy type $|V_{\mathbf{C}}|$ of the complex analytic space associated to $V$.

Is the above statement true or false for algebraic surfaces? Or, more generally, for algebraic varieties of arbitrary dimension?

# 11 Unsolvability in Algebraic Geometry and some implications

One is not yet finished mining this theorem or concrete versions of "unsolvable problems" but it clearly will give us a wealth of such problems. See, for example, recent postings of Harvey Friedman; these have possible relations to Mnëv's (1988) result that any scheme over $\mathbf{Z}$ can be expressed as a moduli space classifying configurations[9] of finite points in $\mathbf{P}^2$. Harvey Friedman poses nine different "Families of Problems" regarding configurations of rational lines in the Euclidean plane, These problems ask for existence or nonexistence of integral intersections (with various properties) of linear configurations. Friedman discusses whether the

[8] rather than existence or nonexistence of solutions

[9] By a *configuration type* let us mean a number $N$ and a collection of subsets $S_1, S_2, \ldots S_n$ of the set $[1, 2, \ldots, N]$. The configuration space associated to such a type is the space of all ordered sets of $N$ points in $\mathbf{P}^2$ subject to the requirement that the points corresponding to $S_1$ are collinear, and ditto for $S_2, \ldots, S_n$.

problems in each of these families can be done in ZFC or whether there are examples of problems in that family that cannot: apparently three of Friedman's problem-families can be solved in ZFC, three cannot, and for the remaining three—if Hilbert's Tenth Problem (over $\mathbf{Q}$) is undecidable—then these cannot be done in ZFC.

**Part II: Elliptic Curves**

I will be discussing joint work with Karl Rubin[10]. One of the goals of our project is (assuming the Shafarevich-Tate Conjecture) to construct examples of elliptic curves $E$ over arbitrary number fields $K$ with specifically prescribed behavior of their Mordell-Weil rank[11] this being useful for the applications discussed in Part I.

# 12 A brief introduction to the arithmetic of elliptic curves

Take a typical elliptic curve with rational coefficients such as:

$$\mathcal{E}: \quad Y^2 + XY + Y \;=\; X^3 - X^2 -$$

$$2006776241557552658503320820933854275093023031217895650 2X$$

$$+$$

$$3448161179503055646703298569039072037485 5$$

$$9443593191803612660082962919394487322434 29$$

We call the (smooth) projective curve over $\mathbf{Q}$ cut out by this equation or by any of its more general companions, which in usual notation are written:

$$Y^2 + a_1XY + a_3Y \;=\; X^3 + a_2X^2 a_4X + a_6$$

with $a_i \in \mathbf{Z}$ and having discriminant nonzero, *elliptic curves*. They are, in a natural way, algebraic groups: the classical *chord-and-tangent-process* endows the set of (rational) points on such an elliptic curve $E$ (over any number field $K$) with the structure of an abelian group (called the Mordell-Weil group of $E$ over $K$ and denoted $E(K)$. This group is known to be finitely generated (a theorem almost a century old due to Mordell in the case of $K = \mathbf{Q}$ and generalized by Weil for arbitrary global fields). So

$$E(K) \approx \mathbf{Z}^r \oplus \mathrm{T(E, K)}$$

where $r = r(E, K)$ is called the **Mordell-Weil rank of $E$ over $K$** and $T(E, K)$ is the (finite) **torsion subgroup of the Mordell-Weil group**.

The structure of the torsion subgroup part of the Mordell-Weil group is relatively more understood than the rank; for example, any $T(E, K)$ is isomorphic to a subgroup of $\mathbf{Q}/\mathbf{Z} \times \mathbf{Q}/\mathbf{Z}$; and for any given $K$ there are only finitely possible isomorphism types for the $T(E, K)$'s that can occur; and for $K = \mathbf{Q}$ the list of these finitely possible isomorphism types is explicitly known.

In contrast, very little is known of a general nature regarding the behavior of $r(E, K)$, for varying $E$ and $K$. I don't know, for example, whether

$$\sup_{E,K} \frac{r(E, K)}{[K : \mathbf{Q}]}$$

[10] regarding the 2-Selmer rank in families of quadratic twists of elliptic curves over arbitrary number fields

[11]
$$\text{rank } E(K) := \dim_{\mathbf{Q}} E(K) \otimes \mathbf{Q}.$$

is finite when $E/K$ runs over all elliptic curves over all number fields. (I bet no!).

In the case of the elliptic curve $\mathcal{E}$ displayed at the beginning of this section, it was shown (by Noam Elkies, 2006) that (it has trivial torsion over $\mathbf{Q}$) and its Mordell-Weil rank over $\mathbf{Q}$ is at least 28— this being, to my knowledge, the current world's record for Mordell-Weil ranks over $\mathbf{Q}$. If you are curious to see 28 linear independent rational points on this curve, here they are:

**Independent points of infinite order:**

P1 = [-21241500912543810732921374 63, 259854492051899599030515511070780628911531]
P2 = [23345098660347017568847545 37, 188720041954944691808683165528036279315 31]
P3 = [-16717360540623690638790386 63, 25170937726114428780850694724131912604913 1]
P4 = [21391302601391566664929821 37, 36639509171439729202421459692941297527531]
P5 = [15347067644671207238854773 37, 85429585346017694289021032862781072799531]
P6 = [-27310794878756770333415750 63, 26252181548433219164128407262390214338753 1]
P7 = [27757262668445716497054585 37, 12845755474014060248869487699082640369931]
P8 = [14943857293271889575418338 17, 88486605527733405986116494514049233411451]
P9 = [18684382286208873585090652 57, 59237403214437708712725140393059358589131]
P10 = [20089451088257437748665425 37, 47690677880125552882151750781541424711531]
P11 = [23483605409180251696516329 37, 17492930006200557857340332476448804363531]
P12 = [-14720840070904811744700086 63, 24664345065350371419994744154975979846913 1]
P13 = [29241286077080612133632889 37, 28350264431488878501488356474767375899531]
P14 = [53749938910660618932939345 37, 28618890842726338645117503191647989373153 1]
P15 = [17096907682333545233340085 57, 71898834974686089466159700529215980921631]
P16 = [24509540113535931440725951 87, 44452281735326343570492625506107147365 31]
P17 = [29692547092735591674646749 37, 32766893075366270801333682543160469687531]
P18 = [27119149349416926013328829 37, 20684366127783816986504139815065906135 31]
P19 = [20078586077996854528778328 937, 27796085411378066046560517256246240300915 31]
P20 = [21580824502407347743178106 97, 34994373401964026809969662241800901254731]
P21 = [20046454582470590224032249 37, 48049329780704645522439866999888475467531]
P22 = [29757494509479962649470913 37, 33398989826075322320208934410104857869131]
P23 = [-21024904676862851501473478 63, 25957639145987578957167739317168720322753 1]
P24 = [31158317991506303490219453 7, 16810438522998060354010947291566015347393 1]
P25 = [27739310083418652314437718 17, 12632162834649921002414116273769275813451]
P26 = [21565811881437684093634613 87, 35125092964022908897004150516375178087331]
P27 = [38663304998724125088156591 37, 12119775565594422629303692671502584732253 1]
P28 = [22308682897735760237786787 37, 28558760030597485663387020600768640028531]

# 13 Descent

The standard method—perhaps the only method—of finding upper bounds for $r(E, K)$ for specific elliptic curves $E$ over specific fields $K$ is called the *method of descent* that seems to have been already present in some arguments due to Fermat and has been elaborated and refined ever since. Rather than say, in any specificity what this method is (for which you can consult the undergraduate text [S-T] and the graduate text [Silv]) in a few words, here is the "shape" of the descent method as it has evolved in present times.

Fix the elliptic curve $E$ and the number field $K$. For each integer $N > 1$ one constructs a certain finite abelian group of exponent $N$ called the *$N$-Selmer group* $S_N(E, K)$ (this is given by a construction related to Galois cohomology) which has two key properties:

- $S_N(E,K)$ is c̈omputable in theory; i.e., there is indeed a finite algorithm that computes it. Also for $N = 2$ or for certain choice triples $(N, E, K)$, there is an elaborate technology for actually managing to compute it[12].
- There is a canonical *injective homomorphism*

$$E(K)/NE(K) \hookrightarrow S_N(E,K)$$

A consequence is that we may obtain—thanks to the above two items—a computable upper bound for the Mordell-Weil rank $r(E,K)$ and indeed, a computation of the $N$-Selmer group $S_N(E,K)$ for *any* $N \geq 2$ provides us with some finite upper bound for $r(E,K)$; specifically:

$$r(E,K) \leq \log_N |S_N(E,K)|.$$

Let

$$\mathcal{N} \subset \mathbf{N}$$

be any unbounded set of natural numbers.

**Conjecture 13.1** *By the* $\mathcal{N}$**-Shafarevich-Tate Conjecture**—*for* $(E,K)$—*we mean the claim that the cardinalities of the cokernel of the injective homomorphisms*

$$E(K)/NE(K) \hookrightarrow S_N(E,K)$$

*admit a (finite) upper bound for all* $N \in \mathcal{N}$*. It is evidently equivalent to ask that the numbers*

$$\frac{|S_N(E,K)|}{|E(K)/NE(K)|}$$

*admit such a finite upper bound. Or—also equivalently—*

$$|S_N(E,K)| = N^{r(E,K)} \times O(1)$$

*for all* $N \in \mathcal{N}$*.*

Clearly if for *any single* unbounded set of natural numbers $\mathcal{N}$ the $\mathcal{N}$-Shafarevich-Tate Conjecture held for $(E,K)$ we would have something better than a mere *upper bound* for the Mordell-Weil rank of $E$ over $K$: the above formula gives us a precise expression for $r(E,K)$, and it would even be nicely computable[13] if one could—in addition—bound the "$O(1)$" occurring in this formula.

By the *Shafarevich-Tate Conjecture for* $(E,K)$ with no prefix we'll mean the $\mathbf{N}$-Shafarevich-Tate Conjecture. This is indeed eponymously conjectured and this conjecture visibly implies the $\mathcal{N}$- Shafarevich-Tate Conjecture for any $\mathcal{N}$. By the *p-adic Shafarevich-Tate Conjecture for* $(E,K)$ we mean the $\mathcal{P}$-Shafarevich-Tate Conjecture for $(E,K)$ where $\mathcal{P}$ is the set of all powers of the prime number $p$.

We do have other methods for determining $r(E,K)$ either in special situations (via Heegner points and similar constructions) and we have other expressions for $r(E,K)$ conditional on various standard conjectures (as the order of vanishing of the $L$-function $L(E,K;s)$ at $s = 1$) or (conditional, again) formulas for it in terms of more intricate relationships (as in *mean values of biases*).

But a perfectly general method for determining $r(E,K)$ is to

[12]There was just a SAGE DAYS workshop at the Clay Mathematics Institute in Cambridge, where some fast methods were featured.

[13]I.e.,

$$r(E,K) = \lim_{N \in \mathcal{N}} \lfloor \log_N |S_N(E,K)| \rfloor.$$

- **by day:** search for rational points, say by increasing heights, and work out their possible dependence relations, getting a monotone increasing sequence of lower bounds for $r(E, K)$ and
- **by night:** compute $\log_N |S_N(E, K)|$ for a chosen succession of increase values of $N$ getting a monotone decreasing sequence of upper bounds for $r(E, K)$, and
- **hope:** that one day or one night the upper and lower bounds will be equal, giving you a perfectly unconditional computation of $r(E, K)$.

The effect of the Shafarevich-Tate Conjecture (for any single unbounded set $\mathcal{N}$) is to make the claim that *the above strategy will always succeed* if one lives long enough. It would be very very interesting to quantify—conjecturally—how long one has to live to see this conjectured algorithm terminate—as an elementary function of the basic numerical data of the input $(E, K)$.

# 14 Mordell-Weil Stability

The following types of questions, interesting in their own right, seem never to have been focused on by number theorists until it was shown that their answers bear interesting consequences for mathematical logic.

**Definition 14.1** *Let $L/K$ be an extension (of finite degree) of number fields and let $E$ be an elliptic curve over $K$. Say that $E$ has* **stable Mordell-Weil rank** *$r$* **for** *$L/K$ if (it does; i.e., if:)*

$$r = r(E, K) = r(E, L).$$

**Question 14.2** *Let $L/K$ be any extension of number fields and $r \geq 0$. Does there exist an elliptic curve $E$ defined over $K$ with stable rank $r$ for $L/K$?*

It isn't entirely unnatural that such a question might be put to useful purposes in attempting to negotiate between recursively enumerable sets expressed in a Diophantine way over alternatively $\mathcal{O}_L$ and $\mathcal{O}_K$, although —to my mind—it is still quite a surprise (thanks to important work of Poonen, Shlapentokh, and others) that an affirmative answer to this question for positive $r$ can be made to imply that $\mathcal{O}_K$ is Diophantine in $\mathcal{O}_L$ and therefore that a negative solution to Hilbert's Tenth Problem for $\mathcal{O}_K$ implies a similarly negative solution for $\mathcal{O}_L$. We shall discuss this in the next section.

# 15 Hilbert's Tenth problem relative to $L/K$

Let, as above, $L/K$ denote a (finite degree) extension of number fields and $\mathcal{O}_L/\mathcal{O}_K$ the induced extension of rings of integers. It would be pretty elegant if we could just put our finger on a simple finite system of polynomial equations over $\mathcal{O}_L$ that cleanly cut out $\mathcal{O}_K$ as a Diophantine subset of $\mathcal{O}_L$. The actual constructions, though, are somewhat more complicated, and more sophisticated. The constructions (in the work of Matiyasevich, Denef, Denef-Lipschitz, Pheidas) start with a particularly manageable system (not exactly "stable" in the above sense) of Pell-type equations (i.e., systems built with equations of the shape $X^2 - dY^2 = \pm 1$) to construct their Diophantine language to relate $\mathcal{O}_L$ to $\mathcal{O}_K$, but only for da special class of number fields $K$). The work of Poonen and Shlapentokh works more generally over any number field $K$ but requires a Mordell-Weil stability result:

**Theorem 15.1** *(Poonen, Shlapentokh) Assume there exists an elliptic curve $E$ defined over $K$ with stable rank $r > 0$ for $L/K$. Then if every recursively enumerable subset of $\mathcal{O}_K$ is Diophantine, then every recursively enumerable subset of $\mathcal{O}_L$ is Diophantine.*

A fairly elementary exercise, discussed in the appendix (we are indebted to Poonen for pointing this out to us) allows us to reduce the question of (integral!!) ) Diophantine-ness of recursively enumerable sets over arbitrary rings of integers to the following question in the arithmetic theory of elliptic curves (which Karl Rubin and I find wonderful):

**Theorem 15.2** *If for every Galois (cyclic) extension of number fields $L/K$ of prime degree there exists an $L/K$-stable elliptic curve, then for any number field $K$ any recursively enumerable subset of $\mathcal{O}_K$ is Diophantine over $\mathcal{O}_K$.*

It would follow that Hilbert's Tenth Problem would have a negative solution for the ring of integers in any number field.

As mentioned, then, in a previous section, Karl Rubin and I assume a classical conjecture (or at least a part of it) called the 2-adic Shafarevich-Tate Conjecture and—conditional on that conjecture—we prove, for every prime degree Galois extension $L/K$ of number fields there is an $L/K$-stable elliptic curve. Hence—again "modulo 2-adic Shafarevich-Tate"—we establish the negative result for Hilbert's Tenth Problem over the ring of integers of an arbitrary number field. A fairly elementary exercise, discussed in the appendix (we are indebted to Poonen for pointing this out to us) allows us to reduce the question of (integral!!) ) Diophantine-ness of recursively enumerable sets over arbitrary rings of integers to the following question in the arithmetic theory of elliptic curves (which Karl Rubin and I find wonderful):

**Theorem 15.3** *If for every Galois (cyclic) extension of number fields $L/K$ of prime degree there exists an $L/K$-stable elliptic curve, then for any number field $K$ any recursively enumerable subset of $\mathcal{O}_K$ is Diophantine over $\mathcal{O}_K$.*

It would follow that Hilbert's Tenth Problem would have a negative solution for the ring of integers in any number field.

**Theorem 15.4** *(Conditional on the 2-adic Shafarevich-Tate conjecture[14]) For any cyclic extension of number fields $L/K$ of prime degree, there exists an elliptic curve $E$ over $K$ such that*

$$0 \;<\; \text{rank } E(K) \;=\; \text{rank } E(L) \;=\; 1.$$

Slightly more specifically,

**Theorem 15.5** *Suppose that the kernel of multiplication by 2 in the Shafarevich-Tate group of any elliptic curve over any number field has even dimension over $\mathbf{F}_2$ (as would be implied if the 2-primary part of the Shafarevich-Tate conjecture were true).*

*Then for any number field $K$ there is no finite decision procedure that allows as its input any finite system of polynomial equations with many variables, and with coefficients in $\mathcal{O}_K$ and that has as output the answer to the question: does this finite system of equations have a simultaneous solution over $\mathcal{O}_K$?*

In order to prove the results described, Rubin and I were led to develop sharp methods to control 2-Selmer rank, and these methods have some side consequences that answer—also raise—some natural questions. Here is an example of two such results that we prove unconditionally:

**Theorem 15.6** *Given any number field $K$ there is an elliptic curve $E$ over $K$ such that there are many quadratic characters $\chi$ of $K$ such that the twisted elliptic curve $E^\chi$ has no nontrivial rational points over $K$.*

**Note:** We can even take $E$ to be the base change to $K$ of an elliptic curve over $\mathbf{Q}$.

**Theorem 15.7** *If $K$ is not totally imaginary, and $E$ is any elliptic curve over $K$ with no nontrivial $K$-rational 2-torsion, there are many quadratic characters $\chi$ of $K$ such that the twisted elliptic curve $E^\chi$ has no nontrivial rational points over $K$.*

**Notes:**

- By *many quadratic characters* we only mean to say that the number of characters $\chi$ with $N_{K/\mathbf{Q}}\text{cond}_K(\chi) < X$ that gives the conclusion of the theorem is $>> X/\log^a(X)$ for some positive number[15] $a$.

[14] Really we only use the following consequence of that conjecture: the 2-rank of the Shafarevich-Tate group is even.

[15] This number $a = 2/3$ or $a = 1/3$ according as the action of the Galois group on $E[2]$ is $\text{GL}_2(\mathbf{F}_2)$ or cyclic of order three.

- Over the field $K = \mathbf{Q}$ Ono-Skinner have proven (using Kolyvagin et al) a similar result to Theorem 15.7 by showing the analogous statement for central values of $L$ functions.
- The lower bound "$>> X/\log^a(X)$" is unfortunately weak: it is expected that the set of $\chi$'s satisfying the conclusion of Theorem 15.7 is of positive density, and—in fact—of density 1/2.

As hinted above, our method is to study the 2-Selmer group of an elliptic curve, and more specifically to see how tightly we can control changes in the rank (i.e., dimension over $\mathbf{F}_2$) of the 2-Selmer group as one twists the original elliptic curve by quadratic characters. (In the Number Theory Seminar–Feb. 3—I'll discuss these methods in slightly more technical detail.

But allow me to end with two questions we still don't know the answer to:

Given any elliptic curve $E$ over a number field $K$ and *any* finite extension $L/K$. Is there a quadratic twist of $E$ that has Mordell-Weil stability for $L/K$ and Mordell-Weil rank $\leq 1$?

Which hyperelliptic curves over a number field $K$ have the property that for any finite extension $L/K$ they possess quadratic twists $C$ over $K$ that

$$C(L) = C(K) \neq \emptyset?$$

**APPENDIX**

# 16 Collections of field extensions

Let $K$ be a field and $\bar{K}/K$ an algebraic closure of $K$. By a *small subextension* of $\bar{K}/K$ let us mean a field extension $M/L$ where $K \subset L \subset M \subset \bar{K}$ and such that $[M : K]$ is finite.

Let $\mathcal{D}$ be a collection of small subextensions of $\bar{K}/K$. Suppose that $\mathcal{D}$ is closed under these three operations.

1. *intersection :* if $M/L_1, M/L_2$ are in $\mathcal{D}$ then $M/L_1 \cap L_2$ is in $\mathcal{D}$.
2. *smaller sub-extensions with the same base:* If $L \subset N \subset M$ and $M/L$ is in $\mathcal{D}$ then so is $N/L$.
3. *composition:* If $M/N$ and $N/L$ are in $\mathcal{D}$, so is $M/L$.

**Lemma 16.1** *If $\mathcal{D}$ is a collection of small subextensions of $\bar{K}/K$ closed under the above three operations, then $\mathcal{D}$ consists of all small subextensions if and only if $\mathcal{D}$ contains all small subextensions $M/L$ that are cyclic of prime degree.*

**Proof.** Note first that if $\mathcal{D}$ contains all small subextensions $M/L$ that are cyclic of prime degree—i.e., if $\mathcal{D}$ satisfies the hypotheses of Lemma 16.1— and if $\mathcal{D}$ is closed under composition (item (3)) above) then $\mathcal{D}$ contains all cyclic extensions. Assume this.

Now let $M/L$ be an arbitrary Galois small subextension. Since $\mathcal{D}$ contains all cyclic small subextensions, for every element $\sigma \in G := \mathrm{Gal}(M/L)$ we have that $M/M^\sigma$ is in $\mathcal{D}$, (here $M^\sigma$ is the fixed subfield of $\sigma$). Since $L = \cap_{\sigma \in G} M^\sigma$ and since $\mathcal{D}$ is closed under intersection (item (1)) above) we have that $M/L$ is in $\mathcal{D}$. It then follows that for any small subextension $M/L$ its Galois closure over $L$ is in $\mathcal{D}$. Consequently since $\mathcal{D}$ is closed under smaller sub-extensions with the same base (item (2)) above) $M/L$ is also in $\mathcal{D}$.

## 16.1 Diophantine subsets

Recall the discussion in section 3 above. Here is one way of formulating this concept. If $A$ is a commutative ring with unit, consider the category $\mathcal{C}_A$ of pairs $(B, b)$ where $B$ is an $A$-algebra of finite presentation and $b \in B$ is an element. Call these **pointed $A$-algebras** where the *finite presentation* condition will be assumed as a running hypothesis. The initial object in this category is the polynomial ring $(A[t], t)$. The category $\mathcal{C}_A$ is closed under direct sum and tensor product. Given a pointed $A$-algebra $(B, b)$ we have a natural set-theoretic mapping

$$\eta_{(B,b)} : \mathrm{Hom}_{A_{\mathrm{alg}}}(B, A) \longrightarrow A$$

that sends $h \in \mathrm{Hom}_{A_{\mathrm{alg}}}(B, A)$ to $h(b) \in A$. Let $D(B, b) \subset A$ be the image of $\eta_{(B,b)}$. By a **Diophantine subset** of $A$ we mean a subset $D \subset A$ that is equal to $D(B, b)$ for some pointed $A$-algebra $(B, b)$. If this is the case we will say that $(B, b)$ *defines* $D$. If $(B_i, b_i)$ defines $D_i$ for $i = 1, 2, \ldots, s$ then the direct sum and the tensor product (both taken in the category $\mathcal{C}_A$) of the $(B_i, b_i)$ define $\cup_i D_i$ and $\cap_i D_i$ respectively.

It follows from this that the collection of Diophantine subsets of $A$ is closed under finite union and finite intersection.

**Proposition 16.2** *Let $\phi : A_o \to A$ be a ring homomorphism with these properties:*

- *$A$ is an $A_o$-algebra of finite presentation,*
- *$\phi$ is injective, so $A_o \subset A$,*
- *$A_o$ is a Diophantine subset of $A$.*

*Then every Diophantine subset of $A_o$ is a Diophantine subset of $A$.*

[... **proof to be included**]

**Definition 16.3** *If the conclusion of Proposition 16.2 holds for an injection of rings $A_o \hookrightarrow A$ we will say that $A_o$ is* **Diophantine** *in $A$.*

## 16.2 Rings of algebraic integers

Put $K = \mathbf{Q}$. From now on, by *small subextension* with no further qualifier we mean small subextension of $\bar{\mathbf{Q}}/\mathbf{Q}$. For $L/K$ let $\mathcal{O}_L$ denote the ring of integers in $L$. Say that a small subextension $M/L$ is **integrally Diophantine** if $\mathcal{O}_L$ is Diophantine in $\mathcal{O}_M$. Let $\mathcal{D}$ be the collection of *integrally Diophantine* small subextensions.

**Proposition 16.4** *The collection $\mathcal{D}$ of integrally Diophantine small subextensions is closed under the operations (1), (2), (3) listed in subsection 16 above.*

**Proof.** Closure under the operation (1) follows since an intersection of Diophantine sets is Diophantine. As for closure under (2) and (3) [... **proof to be included**].

*Bibliography*

[**B75**] A. Baker, *Transcendental Number Theory*, Cambridge Mathematical Library, Cambridge University Press (1975).
[**E03**] A.K. Eisenträger, Hilbert's Tenth Problem and Arithmetic Geometry, PhD Thesis, University of California, Berkeley, 2003.
[**Den80**] J. Denef, Diophantine sets over algebraic integer rings. II, Trans. Amer. Math. Soc. 257 (1980), no. 1, 227-236.
[**DL78**] J. Denef and L. Lipshitz, Diophantine sets over some rings of algebraic integers, J. London Math.

Soc. (2) 18 (1978), no. 3, 385-391.
[**H96**] E. Hrushovski, The Mordell-Lang conjecture for function fields. Journal of the American Mathematical Society **9** (1996), no. 3, 667-690.
[**H98**] E. Hrushovski, Proof of Manin's theorem by reduction to positive characteristic, *Model theory and algebraic geometry,* Lecture Notes in Math.,**1696**, Springer, Berlin, (1998) 197-205.
[**H00**] E. Hrushovski, Anand Pillay, Effective bounds for the number of transcendental points on subvarieties of semi-abelian varieties. American Journal of Mathematics **122** (2000), no. 3, 439-450.
[ **JSWW76**] J. P. Jones, D. Sato, H. Wada and Do. Wiens, Diophantine Representation of the Set of Prime Numbers, The American Mathematical Monthly, **83**, No. 6 (Jun. - Jul., 1976), pp. 449-464.
[**M77**] B. Mazur, Modular Curves and the Eisenstein ideal, Publ. IHES **47** (1977) 33-186.
[**MR09**] B. Mazur and K. Rubin, Ranks of twists of elliptic curves and Hilbert's Tenth Problem, (http://abel.math.harvard.ed and also on Archiv (arXiv:0904.3709v2 [math.NT] 25 Apr 2009)
[**Phe88**] Thanases Pheidas, Hilbert's tenth problem for a class of rings of algebraic integers, Proc. Amer. Math. Soc. **104** (1988), no. 2, 611-620.
[**PR06**] R. Pink, D. Roessler, On Hrushovski's proof of the Manin-Mumford conjecture (preprint).
[**Po**] Bjorn Poonen, Using elliptic curves of rank one towards the undecidability of Hilbert's Tenth Problem over rings of algebraic integers (. . . )
[**PW07**] J. Pila, A.J. Wilkie, The rational points of a definable set, MIMS Eprint 2007.198.
[**Shl89**] Alexandra Shlapentokh, Extension of Hilbert's tenth problem to some algebraic number fields, Comm. Pure Appl. Math. **42** (1989), no. 7, 939- 962.
[**Shl00b**] Alexandra Shlapentokh, Hilbert's tenth problem over number fields, a survey, *Hilbert's tenth problem: relations with arithmetic and algebraic geometry* (Ghent, 1999), Amer. Math. Soc., Providence, RI, 2000, pp. 107- 137.
[**Sm**] Raymond M. Smullyan, *Theory of formal systems*, **47** Annals of Mathematics Series, Princeton University Press.