SOME THEOREMS AND CONJECTURES IN DIOPHANTINE EQUATIONS

BY SERGE LANG

The theory of diophantine equations may be regarded as the natural continuation of algebraic geometry proper: Having once obtained a general theory of algebraic equations in several variables over essentially arbitrary ground fields (or rings), one tries to get statements depending on the special arithmetic structure of the coefficient domain. By definition, this becomes diophantine analysis. We shall list a few of the theorems and conjectures which arise in this direction.

Let k be a field, and $f(X_1, \cdots, X_n)$ a polynomial, also written $f(X)$, with coefficients in k. The equation $f=0$ defines an algebraic set, i.e. the set of all n-tuples $(x_1, \cdots, x_n)$ in some algebraically closed field containing k, such that $f(x)=0$. Such a point (x) in n-space is said to be a zero of f. It is said to be a rational point in k if all x_i lie in k. If f is a form (i.e. a homogeneous polynomial) then one views f as defining an algebraic set in projective space, and one considers nontrivial zeros, that is zeros such that not all x_i are 0. A nontrivial zero then defines a point in projective space, which is rational over k if again the coordinates can be chosen in k.

More generally, one considers systems of equations, or varieties (meaning an absolutely irreducible algebraic set). If V is a variety defined over a field k, then a point in it is rational over k if it has a set of coordinates in k.

The basic coefficient domain is that of the rational numbers $\mathbf{Q}$ or the integers $\mathbf{Z}$. It is but a step from this to a finite extension k of $\mathbf{Q}$ (called a number field) or the ring of integers I_k of k instead of $\mathbf{Z}$. We have primes $\mathfrak{p}$ associated with such fields: They are the absolute values which either induce the ordinary absolute value on $\mathbf{Q}$ (called archimedean primes) or the p-adic absolute value, defined by a prime number p:

$$|p^r m/n|_{\mathfrak{p}} = 1/p^r$$

if $m, n \in \mathbf{Z}$, $mn \neq 0$, and $p \nmid mn$. The latter are called finite primes. One can then form the completion $k_{\mathfrak{p}}$ under the prime $\mathfrak{p}$, which is called a $\mathfrak{p}$-adic field, and is the field of real or complex numbers if $\mathfrak{p}$ is

An address delivered before the February Meeting of the Society in New York on February 27, 1960 by invitation of the Committee to Select Hour Speakers for Eastern Sectional Meetings; received by the editors March 16, 1960.

archimedean. (It is also becoming standard to consider field and ring extensions of finite type.)

Because of the topology and the completeness, a $\mathfrak{p}$-adic field gives rise to simpler diophantine problems than a number field, and one tries to reduce certain classes of diophantine problems to $\mathfrak{p}$-adic ones. Let us observe right away that a variety (say affine) defined over a number field k has a rational point in all but a finite number of $\mathfrak{p}$-adic fields $k_{\mathfrak{p}}$. (Remember V is absolutely irreducible.) This is easily seen, for instance as follows: By cutting V with sufficiently general hyperplane sections, one reduces the question to the case where V is a curve. One then reduces mod $\mathfrak{p}$. By the Riemann hypothesis in function fields (i.e. Weil's theorem [25]) the curve has a simple point mod $\mathfrak{p}$ for all but a finite number of primes, and this simple point can be refined to a $\mathfrak{p}$-adic point, say by Hensel's lemma.

1. **Hasse's theorem.** Let us begin with essentially the simplest type of variety, that defined by a quadratic equation. The main result here is Hasse's theorem (for an exposition, see for instance [26]). *Let k be a number field, f a quadratic form with coefficients in k. Then f has a nontrivial zero in k if and only if f has a nontrivial zero in each $k_{\mathfrak{p}}$.*

This is supplemented by a useful $\mathfrak{p}$-adic criterion: If $\mathfrak{p}$ is a finite prime, then *every quadratic form in* 5 *variables over the $\mathfrak{p}$-adic field $k_{\mathfrak{p}}$ has a nontrivial zero in $k_{\mathfrak{p}}$.*

These two theorems are typical examples of the following general principles: To get a global theorem from local ones, and to get solutions if the number of variables is large. As a corollary, we see that every quadratic form in 5 variables over a number field k, which is indefinite for every real embedding of the number field, has a nontrivial zero in k.

One may try to embed the above statements in theories concerning either forms of higher degree than 2, or concerning principal homogeneous spaces. Let us discuss the first.

2. **Quasi-algebraic closure.** There is a theorem of Peck [17] that *a form over a number field k of degree d, in n variables has a nontrivial zero in k if n is sufficiently large compared to d, and if k is totally imaginary (i.e. has no embedding into the reals).*

The condition that k be totally imaginary is essential: A sum of squares in a real field never has a nontrivial zero, and in fact, Artin gives an example of a form which is indefinite, has a $\mathfrak{p}$-adic zero for all $\mathfrak{p}$, has arbitrarily many variables, and still no zero in k, namely

$$(X_1^2 + \cdots + X_n^2)^2 - 2(Y_1^2 + \cdots + Y_m^2)^2 = 0$$

over the rational numbers (substituting sums of squares in $U^2 - 2T^2$).

Artin's substitution method is also used to reduce a system of simultaneous equations to one equation (especially with reference to quasi-algebraic closure, see below). One also knows how to reduce a system over a finite extension E of k to a system of equations in k (linearizing by means of a basis). For both of these, cf. [6; 7].

The above reductions are " multiplicative." This is important, because one may ask whether by restricting one's attention to forms of odd degree one does not recover the desired conclusion even when the field is real. Taking function fields over the reals as ground fields, I gave precise criteria under which such forms have nontrivial zeros [7]. The analogue of Peck's theorem for forms of odd degree and any number field was proved by Birch [1].

The Birch-Peck theorem holds for a large number of variables. Except for quadratic forms, one has no precise bound in number fields, but Artin has at least made conjectures concerning it. He defines a field K to be *quasi-algebraically closed* (QAC) if every form with coefficients in K, of degree d, in n variables, with $n > d$, has a nontrivial zero in K. A field which is quasi-algebraically closed does not admit division algebras of finite degree above it. Tsen proved that a function field in one variable over an algebraically closed constant field has no such division algebras. Analysing Tsen's proof, Artin was led to make the above remark, to define quasi-algebraic closure, to realize that Tsen's proof actually showed that such a function field was quasi-algebraically closed, and in view of Wedderburn's theorem, to conjecture that *finite fields are* QAC. This was proved by Chevalley [5]. Furthermore, it is known from class field theory that the field Ω obtained by adjoining all roots of unity to the rationals admits no finite division algebra above it. This and the analogy with function fields (a function field in one variable over a finite field to which one adjoins all roots of unity becomes a function field over an algebraically closed constant field) led him to *conjecture that Ω is* QAC. Thus, for instance, every form of degree d in n variables over the rationals with $n > d$ would have a nontrivial zero in some cyclotomic field.

How about number fields proper? In this case, Artin suggested that probably the condition $n > d$ has to be replaced by $n > d^2$ (always provided the field is totally imaginary). This is true in the analogous case of function fields over finite fields. At any rate, Artin conjectured that *a cubic form in* 10 *variables over the rationals* $\mathbb{Q}$ *has a nontrivial*

zero in $\mathbf{Q}$. Artin also made the analogous local conjectures. The one concerning the roots of unity is proved in my thesis [6], and the case of cubic forms is settled by Lewis [12]. The $\mathfrak{p}$-adic case proper remains open, in spite of the fact that the analogue for power series in one variable over a finite field is easily taken care of [6] with $n>d^2$. By the way, in each case the condition $n>d$ or $n>d^2$ is easily seen to be best possible. None of the global conjectures has yet been proved.

One can consider function fields over number fields or $\mathfrak{p}$-adic fields as ground fields themselves, and extend to those and to power series fields the same type of result. The condition $n>d$ (or $n>d^2$) has to be replaced by $n>d^i$ where i goes up with the number of variables. Although one can settle the function field case [6], the case of power series in several variables also remains open.

Finally, to go back to number fields, it seems to me reasonable to expect that a form with $n>d$ at least has a nontrivial zero in all but a finite number of $\mathfrak{p}$-adic fields. As pointed out previously, there is a problem here only if the form is not absolutely irreducible.

3. **Principal homogeneous spaces.** Let us return to quadratic forms. Let f, g be two quadratic forms over a number field k, in the same number of variables. They are equivalent over k if there exists a matrix T with coefficients in k such that $Tf=g$. It follows immediately from Hasse's theorem that if f, g are equivalent over every $k_\mathfrak{p}$ then they are equivalent over k (see Witt [26]).

Observe that the set of transformations T such that $Tf=g$ is a principal homogeneous space over the orthogonal group of f, which operates simply transitively on this set. More generally, let G be a group variety defined over a field k. A variety V is said to be a principal homogeneous space of G over k if V is defined over k, and we are given over k an everywhere defined rational map of $G\times V$ into V such that for every point $v\in V$, the map $x\to xv$ of G into V establishes an isomorphism of G onto V (for the structure of algebraic variety). Cf. Weil [22], who was the first to call attention to principal homogeneous spaces in relation to diophantine analysis.

This is precisely the situation we have with our quadratic forms f, g and they are equivalent over k if and only if the principal homogeneous space has a rational point. This aspect of Hasse's theorem may therefore be formulated by saying that *a principal homogeneous space over the orthogonal group has a rational point in* k *if and only if it has a rational point in every* $k_\mathfrak{p}$. Serre has suggested that this may remain true for any semi-simple group G, not only the orthogonal group.

The conclusion that the existence of a $\mathfrak{p}$-adic point for all $\mathfrak{p}$ implies the existence of a rational point in the number field k does not hold when one considers other types of group varieties, for instance an abelian variety, or for concreteness an elliptic curve. Selmer [18] has given examples of elliptic curves over the rationals, namely $3X^3+4Y^3+5Z^3=0$, which have a point in every $\mathbf{Q}_p$ but not a rational point in $\mathbf{Q}$. Every such curve can be regarded as a principal homogeneous space over its Jacobian. (See also Cassels [2].)

Here again, before dealing with the global theory, one studies the local one, over a $\mathfrak{p}$-adic field. Over the reals, this is the way one can interpret a paper of Witt [27]. Over $\mathfrak{p}$-adic fields, Shafarevic [19] and Tate [21] have considered the question, and obtained a classification theorem in the case of elliptic curves.

In many cases, the principal homogeneous spaces and the existence of birational correspondences between curves had been studied by Chatelet [4], who pointed out their connection with cohomology. Let G be a group variety and V a principal homogeneous space defined over a field k. Let K be a Galois extension of k, with Galois group $\mathfrak{g}=\mathfrak{g}_{K/k}$, in which V has a rational point v_0. For each $\sigma\in\mathfrak{g}$, the point σv_0 lies in V, and hence there exists a unique element x_σ in G, rational over K, such that $x_\sigma\sigma v_0=v_0$. One verifies that (x_σ) is a 1-cocycle of $\mathfrak{g}$ in G_K, i.e. that $x_\sigma\sigma x_\tau=x_{\sigma\tau}$. Defining coboundaries in the obvious way, one obtains a cohomology set $H^1(\mathfrak{g}, G_K)$. One sees immediately that V has a rational point in k if and only if its associated cohomology class in $H^1(\mathfrak{g}, G_K)$ is trivial. Going to the injective limit to the separable algebraic closure of k, one is led to study the set $H^1(k, G)$, limit of the $H^1(\mathfrak{g}_{K/k}, G_K)$, and whose elements are in bijective correspondence with the isomorphism classes of principal homogeneous spaces of G over k. (Cf. [11]). If G is commutative, then $H^1(k, G)$ is of course a group, called the first cohomology group, and one can define also the higher dimensional ones.

When the ground field is a $\mathfrak{p}$-adic field ($\mathfrak{p}$ finite) Tate [21] has obtained a duality theorem: *Let A be an abelian variety defined over $k_\mathfrak{p}$. Then $H^1(k_\mathfrak{p}, A)$ is dual to the compact group of rational points in $k_\mathfrak{p}$ of the Picard variety of A.* He has also obtained a complete analysis of the cohomology involved for a coefficient module which arises from the points of a commutative group variety which is of multiplicative type (i.e. becomes a product of multiplicative groups over the algebraic closure), and for abelian varieties, both for the limit cohomology and in finite layers. The former, in finite layers, are dual to the modules arising in the Nakayama-Tate theorem of class field theory [15]. For *an abelian variety A over $k_\mathfrak{p}$, he shows that $H^r(k_\mathfrak{p}, A)=0$ if*

$r>1$, and if K is finite Galois over $k_\mathfrak{p}$, then $\boldsymbol{H}^r(\mathfrak{g}_{K/k_\mathfrak{p}}, A_K)$ is dual to $\boldsymbol{H}^{1-r}(\mathfrak{g}_{K/k_\mathfrak{p}}, \hat{A}_K)$ where $\hat{A}$ is the Picard variety, and $\boldsymbol{H}$ now denotes Tate's cohomology functor with $-\infty<r<\infty$.

One may consider more generally the limit cohomology (say in dimension 1) with arbitrary coefficients: Let F be a group on which the Galois group Γ_k of the algebraic closure of k over k acts continuously (regarding Γ_k as compact with Krull topology, and F as discrete). (Cf. Bourbaki seminar, 1959, Exposé on Tate's work.) One can build $H^1(\Gamma_k, F)$ as a limit set (group if F is commutative) just as with the connected group varieties. If k is a $\mathfrak{p}$-adic field, and F is commutative and finitely generated (over $\boldsymbol{Z}$) then Tate has shown that this $H^1(\Gamma_k, F)$ is finite. If F is a finite group (not necessarily commutative) it is actually easy to prove the finiteness statement directly, using the fact that a $\mathfrak{p}$-adic field has only a finite number of extensions of given degree. If F is finitely presented (i.e. given by a finite number of generators and relations) the answer is not known.

Going over to the global case, one sees that $H^1(k, A)$ is a large group. Shafarevic has given examples of elliptic curves A over the rationals $\boldsymbol{Q}$ such that $H^1(k, A)$ has elements of arbitrarily high period [19]. Over a suitable number field, this can be done more easily [11]. On the other hand, Shafarevic and Tate have been led to conjecture that *if A is an abelian variety defined over a number field k, then the subgroup of $H^1(k, A)$ consisting of those elements which split over every $k_\mathfrak{p}$ is finite.* Stated in geometric terms, this means that the set of isomorphism classes of principal homogeneous spaces over A defined over k, which have a rational point in every $\mathfrak{p}$-adic field is finite.

Tate has proved the analogous statment for the cohomology arising from groups of multiplicative type, or from coefficient modules which are finitely generated (over $\boldsymbol{Z}$). As in the local case, if F is a finite group (not necessarily commutative), one can give a direct proof of the analogous fact, without using class field theory, and with all but a finite number of $\mathfrak{p}$, instead of all $\mathfrak{p}$. (One uses that in a Galois extension of degree >1, infinitely many primes do not split completely.)

As Chatelet perceived [3; 4], the noncommutative cohomology also arises when one asks for conditions under which two varieties become isomorphic to each other. For instance, let V, W be two projective nonsingular varieties defined over k. Let T: $V\to W$ be an isomorphism defined over the Galois extension K of k, with group $\mathfrak{g}$. Then $T^{-1}T^\sigma$ is a cocycle of $\mathfrak{g}$ with coefficients in the group of automorphisms of V defined over K, and V is isomorphic to W over k if and only if this cocycle splits (cf. Weil [24]). This situation is

similar to that which arose from Hasse's theorem. Chatelet saw that from class field theory, one can deduce the theorem that *if a variety over a number field becomes isomorphic to projective space over every* $\mathfrak{p}$*-adic field* $k_{\mathfrak{p}}$*, then it is isomorphic to projective space over* k. The cocycle one gets is in the projective group, connected with the multiplicative group (handled in class field theory) through the exact sequence with the full linear group, whose Galois cohomology in dimension 1 is trivial.

The group of automorphisms of a variety (projective nonsingular) is still somewhat of a mystery in general, although one knows that it is a group extension of an algebraic group by a discrete group (Matsusaka). One may ask whether this discrete group (which in general is not commutative) is finitely presented. For curves of genus $\geqq 2$, it is classical that it is finite, and hence *if* V *is a curve of genus* $\geqq 2$ *defined over a number field* k*, then the set of curves (up to* k*-isomorphism) which are defined over* k*, and are isomorphic to* V *over every* $k_{\mathfrak{p}}$*, is a finite set.*

For all the above groups, one can look at the Galois cohomology for varieties defined over number fields, and ask in each case whether that part of the first cohomology set which splits at all $\mathfrak{p}$ is finite. Supposing for instance that the conjecture of Shafarevic-Tate is true for abelian varieties, can it be extended to all algebraic groups defined over a number field? Can the preceding theorem be extended to all varieties (projective nonsingular), beginning with elliptic curves, or analogously, can it be extended to forms of arbitrary degree, considering only the group of all linear automorphisms?

I would like to conclude this discussion of principal homogeneous spaces by pointing out that *over a finite field, every homogeneous space of a group variety has a rational point* [**8**].

4. **Curves.** We have seen in Hasse's theorem, and the theory of quasi-algebraic closure, that equations with many variables have a tendency to have solutions. In the opposite direction, equations with few variables have a tendency not to have any, or at any rate rather few. In this connection, one has again some theorems and some conjectures.

Foremost among the theorems is the following one of Siegel's [**20**]: *A curve* $f(X, Y)=0$ *defined over a number field has only a finite number of integral points (i.e. points* (x, y) *whose coordinates are integers of that field) if its genus is* $\geqq 1$. For curves of genus 1 over the rationals, Mahler has extended this to points having only a finite number of prime numbers in their denominators [**13**], and actually

I can extend the theorem to a curve of genus $\geqq 1$ defined over a field of finite type over $\boldsymbol{Q}$, with points having their coordinates in a subring of finite type over $\boldsymbol{Z}$ [9].

Mordell [14] has conjectured that actually *the curve will have only a finite number of rational points (in a number field) if its genus is at least* 2.

A curve of genus 1 is an elliptic curve, and if it has a rational point its rational points form a group (coming from the addition formula for elliptic functions). If this group contains one element of infinite order, then it can not be finite, of course. For curves of genus 1 over the rationals $\boldsymbol{Q}$, Mordell proved that *this group is finitely generated* [14]. This was extended by Weil [23] to the group of rational points of an abelian variety over a number field, and this Mordell-Weil theorem has applications to geometric problems of algebraic geometry (Néron [16], see also [10]).

A curve of genus $\geqq 2$ can always be embedded in its Jacobian J, an abelian variety, over a field in which it has a rational point. The group of rational points J_k being finitely generated if k is a number field, one sees Mordell's conjecture in the following light: The intersection of this finitely generated group with the curve should be finite, the curve being of lower dimension than J (if its genus is $\geqq 2$) and thus rather thinly distributed in J. One may even ask whether it might not be true that Mordell's conjecture could be extended to any subvariety of an abelian variety, which does not contain the translation of an abelian subvariety.

Thus, for curves of genus $\geqq 2$ and abelian varieties, the conjecture (resp. theorem) asserts that there are as few rational points as is compatible with the obvious structure of the variety under consideration. The Fermat curve $X^n+Y^n=1$ has genus $\geqq 2$ if $n\geqq 4$, and thus falls under Mordell's conjecture.

In this connection, I would conjecture that Siegel's finiteness statement concerning integral points should in fact be true for affine subsets of abelian varieties, or at least an affine subset which is the complement of a hyperplane section in some projective embedding. The theorem for curves should then be obtainable by pull-back from the Jacobian.

Finally, there is a remarkable conjecture of Siegel, at the end of [20], which I quote: "Die Untersuchungen der vorangehenden Paragraphen geben die Möglichkeit, eine Schranke für die Anzahl der Lösungen der diophantischen Gleichung $f(X, Y)=0$ als Funktion der Koeffizienten von f explicit aufzustellen, falls diese Gleichung nur endlich viele Lösungen besitzt. Man kann nun vermuten, dass

sich sogar eine Schranke finden lässt, die nur von der Anzahl der Koeffizienten abhängt; doch dürfte dies recht schwer zu beweisen sein." Siegel then goes on to give evidence for this conjecture, by treating "allerdings sehr spezielle Resultate."

BIBLIOGRAPHY

1. B. J. Birch, *Homogeneous forms of odd degree in a large number of variables*, Mathematika vol. 4 (1957) pp. 102–105.

2. J. W. Cassels, *Arithmetic on curves of genus* 1, J. Reine Angew. Math. vol. 202 (1959) pp. 52–99.

3. F. Chatelet, *Méthode Galoisienne et courbes de genre* 1, Ann. Univ. Lyon, vol. 9 (1946) pp. 40–49.

4. ———, *Variations sur un thème de Poincaré*, Ann. Sci. Ecole Norm. Sup. vol. 59 (1944) pp. 249–300.

5. C. Chevalley, *Démonstration d'une hypothèse de M. Artin*, Abh. Math. Sem. Univ. Hamburg vol. 11 (1935) pp. 73–75.

6. S. Lang, *On quasi algebraic closure*, Ann. of Math. vol. 55 (1952) pp. 373–390.

7. ———, *The theory of real places*, Ann. of Math. vol. 57 (1953) pp. 380–391.

8. ———, *Algebraic groups over finite fields*, Amer. J. Math. vol. 78 no. 3 (1956) pp. 555–563.

9. ———, *Integral points on curves*, to appear.

10. S. Lang and A. Néron, *Rational points of abelian varieties in function fields*, Amer. J. Math. vol. 81, no. 1 (1959) pp. 95–118.

11. S. Lang and J. Tate, *Principal homogeneous spaces over Abelian varieties*, Amer. J. Math. vol. 80, no. 3 (1958) pp. 659–684.

12. D. J. Lewis, *Cubic homogeneous polynomials over p-adic number fields*, Ann. of Math. vol. 56, no. 3 (1952) pp. 473–478.

13. K. Mahler, *Über die rationalen Punkte auf Kurven vom Geschlecht Eins*, J. Reine Angew. Math. vol. 170 (1934) pp. 168–178.

14. L. J. Mordell, *On the rational solutions of the indeterminate equation of the third and fourth degrees*, Proc. Cambridge Philos. Soc. vol. 21 (1922) pp. 179–192.

15. T. Nakayama, *Cohomology of class field theory and tensor product modules*, Ann. of Math. vol. 65 (1957) pp. 255–267.

16. A. Néron, *Problèmes arithmétiques et géometriques rattachés à la notion de rang d'une courbe algébrique dans un corps*, Bull. Soc. Math. France vol. 80 (1952) pp. 101–166.

17. L. J. Peck, *Diophantine equations in algebraic number fields*, Amer. J. Math. vol. 71 (1949) pp. 387–402.

18. E. Selmer, *The diophantine equation* $ax^3+by^3+cz^3=0$, Acta Math. vol. 85 (1951) pp. 203–362.

19. I. Shafarevic, *Birational equivalence of elliptic curves* and *Exponents of elliptic curves*, Doklady Akad. Nauk SSSR vol. 114 (1957) pp. 267–270 and pp. 714–716.

20. C. L. Siegel, *Über einige Anwendungen Diophantischer Approximationen*, Abh. Preuss. Akad. Wiss. Phys. Math. Kl. (1929) pp. 41–69.

21. J. Tate, *Galois cohomology of abelian varieties over p-adic fields*, to appear. Cf. Bourbaki seminar, 1957.

22. A. Weil, *Algebraic groups and homogeneous spaces*, Amer. J. Math. vol. 77, no. 3 (1955) pp. 493–512.

23. ———, *L'arithmétique sur les courbes algébriques*, Acta Math. vol. 52 (1928) pp. 281–315.

24. ———, *The field of definition of a variety*, Amer. J. Math. vol. 78, no. 3 (1956) pp. 509–524.

25. ———, *Sur les courbes algébriques et les variétés qui s'en déduisent*, Paris, Hermann, 1948.

26. E. Witt, *Theorie der quadratischen Formen in beliebigen Körper*, J. Reine Angrew. Math. vol. 176 (1937) pp. 31–44.

27. ———, *Zerlegung reeller algebraische Funktionen in Quadrate*, J. Reine Angew. Math. vol. 171 (1934) pp. 4–11.

COLUMBIA UNIVERSITY