Diophantine geometry as Galois theory in the mathematics of Serge Lang

Minhyong Kim

March 4, 2006

Lang's conception of Diophantine geometry is rather compactly modelled by the following celebrated conjecture [29]:

> Let V be a subvariety of a semi-abelian variety A, $G \subset A$ a finitely generated subgroup, and $Div(G)$ the subgroup of A consisting of the division points of G. Then $V \cap Div(G)$ is contained in a finite union of subvarieties of V of the form $B_i + x_i$, where each B_i is a semi-abelian subvariety of A and $x_i \in A$.

There is a wealth of literature at this point surveying the various ideas and techniques employed in its resolution, making it unnecessary to review them here in any detail [10, 39]. However, it is still worth taking note of the valuable *generality* of the formulation, evidently arising from a profound instinct for the plausible structures of mathematics. To this end, we remark merely that it was exactly this generality that made possible the astounding interaction with geometric model theory in the 90's [3]. That is to say, analogies to model-theoretic conjectures and structure theorems would have been far harder to detect if attention were restricted, for example, to situations where the intersection is expected to be finite. Nevertheless, in view of the sparse subset of the complex net of ideas surrounding this conjecture that we wish to highlight in the present article, our intention is to focus exactly on the case where A is compact and V does *not* contain any translate of a connected non-trivial subgroup. The motivating example, of course, is a compact hyperbolic curve embedded in its Jacobian. Compare then the two simple cases of the conjecture that are amalgamated into the general formulation:

(1) $V \cap A[\infty]$, the intersection between V and the torsion points of A, is finite.
(2) $V \cap G$ is finite.

Lang expected conjecture (1) to be resolved using Galois theory alone. This insight was based upon work of Ihara, Serre, and Tate ([32], VIII.6), dealing with the analogous problem for a torus, and comes down to the conjecture, still unresolved, that the image of the Galois representation in $Aut(A[\infty]) \simeq GL_{2g}(\hat{\mathbb{Z}})$ contains an open subgroup of the homotheties $\hat{\mathbb{Z}}^*$. Even while assertion (1) is already a theorem of Raynaud [44], significant progress along the lines originally envisioned by Lang was made in [2] by replacing $A[\infty]$ with $A[p^\infty]$, the points of p-power torsion, and making crucial use of p-adic Hodge theory.

It is perhaps useful to reflect briefly on the overall context of Galois-theoretic methods in Diophantine geometry, of course without attempting to do justice to the full range of interactions and implications. Initially, that Galois theory is relevant to the study of Diophantine problems should surprise no one. After all, if we are interested in $X(F)$, the set of rational points of a variety X over a number field F, what is more natural than to observe that $X(F)$ is merely the fixed point set of $\Gamma := Gal(\bar{F}/F)$ acting on $X(\bar{F})$? Since the latter is an object of classical geometry, such an expression might be expected to nicely circumscribe the subset $X(F)$ of interest. This view is of course very naive and the action on Γ on $X(\bar{F})$ is notoriously difficult to use in any direct fashion. The action on *torsion points* of commutative group varieties on the other hand, while still difficult, is considerably more tractable, partly because a finite abelian subgroup behaves relatively well under specialization.

Such an arithmetic variation exerts tight control on the fields generated by the torsion points, shaping Galois theory into a powerful a tool for investigations surrounding (1).

For conjecture (2), where the points to be studied are not torsion, it is not at all clear that Galois theory can be as useful. In fact, my impression is that Lang expected *analytic geometry* of some sort to be the main input to conjectures of type (2). This is indicated, for example, by the absence of any reference to arithmetic in the formulation. We might even say that implicit in the conjecture is an important idea that we will refer to as the *analytic strategy*:

(a) replace the difficult Diophantine set $V(F)$ by the geometric intersection $V \cap G$;
(b) try to prove this intersection finite by analytic means.

In this form, the strategy appears to have been extraordinarily efficient over function fields, as in the work of Buium [4]. Even Hrushovski's [15] proof can interpreted in a similar light where passage to the completion of suitable theories is analogous to the move from algebra to analysis (since the theory of fields is not good enough). These examples should already suffice to convince us that it is best left open as to what kind of analytic means are most appropriate in a given situation. The proof over number fields by Faltings [9], as well as the curve case by Vojta [47], utilizes rather heavy Archimedean analytic geometry. Naturally, the work of Vojta and Faltings draws us away from the realm of traditional Galois theory. However, in Chabauty's theorem [5], where V is a curve and the rank of G is strictly less than the dimension of A, it is elementary non-Archimedean analysis, more specifically p-adic abelian integrals, that completes the proof. Lang makes clear in several different places ([28], [32], notes to chapter 8, [36], I.6) that Chabauty's theorem was a definite factor in the formulation of his conjecture. This then invites a return to our main theme, as we remind ourselves that non-Archimedean analysis has come to be viewed profitably over the last several decades as a projection of *analysis on Galois groups*, a perspective of which Lang was well aware ([33], chapter 4). As such, it has something quite substantial to say about non-torsion points, at least on elliptic curves ([18], for example). Hodge theory is again a key ingredient, this time as the medium in which to realize such a projection [43].

It is by now known even to the general public that a careful study of Galois actions underlies the theorem of Wiles [49], and roughly one-half of the difficulties in the theorem of Faltings [8]. There, Galois representations must be studied in conjunction with an array of intricate auxiliary constructions. However, the most basic step in the Galois-theoretic description of non-torsion points, remarkable in its simplicity, goes through the Kummer exact sequence

$$0 \to A[m](F) \to A(F) \to A(F) \xrightarrow{\delta} H^1(\Gamma, A[m]) \to H^1(\Gamma, A) \to .$$

In this case, an easy study of specialization allows us to locate the image of δ inside a subgroup $H^1(\Gamma_S, A[m])$ of cohomology classes with restricted ramification, which then form a finite group. We deduce thereby the finiteness of $A(F)/mA(F)$, the weak Mordell-Weil theorem. Apparently, a streamlined presentation of this proof, systematically emphasizing the role of Galois cohomology, first appears in Lang's paper with Tate [38]. There, they also emphasize the interpretation of Galois cohomology groups as classifying spaces for *torsors*, in this case, for A and $A[m]$. (We recall that a torsor for a group U in some category is an object corresponding to a set with simply transitive U-action, where the extra structure of the category, such as Galois actions, prevent them from being trivial. See, for example, [40], III.4.) This construction has been generalized in one direction to study non-torsion algebraic cycles by associating to them extensions of motives [7]. More pertinent to the present discussion, however, is a version of the Kummer map that avoids any attempt to abelianize, taking values, in fact, in *non-abelian* cohomology classes.

In the course of preparing this article, I looked into Lang's *magnum opus* [32] for the first time in many years and was a bit surprised to find a section entitled 'non-abelian Kummer theory.' What is non-abelian there is the Galois group that needs to be considered if one does not assume a priori that the torsion points of the group variety are rational over the ground field. The field of m-division points of the rational points will then have a Galois group H of the form

$$0 \to A[m] \to H \to M \to 0$$

where $M \subset GL_{2g}(\mathbb{Z}/m)$. Thus, 'non-abelian' in this context is used in the same sense as in the reference to non-abelian Iwasawa theory. But what is necessary for hyperbolic curves is yet another layer of non-commutativity, this time in the coefficients of the action. Given a variety X with a rational point b, we can certainly consider the étale fundamental group $\hat{\pi}_1(\bar{X}, b)$ classifying finite étale covers of $\bar{X}$. But the same category associates to any other point $x \in X(F)$ the set of étale paths

$$\hat{\pi}_1(\bar{X}; b, x)$$

from b to x which is naturally a torsor for $\hat{\pi}_1(\bar{X}, b)$. All these live inside the category of pro-finite sets with Galois action. There is then a non-abelian continuous cohomology set $H^1(\Gamma, \hat{\pi}_1(\bar{X}, b))$ that classifies torsors, and a non-abelian Kummer map

$$\delta^{na} : X(F) \to H^1(\Gamma, \hat{\pi}_1(\bar{X}, b))$$

sending a point x to the class of the torsor $\hat{\pi}_1(\bar{X}; b, x)$. This is obviously a basic construction whose importance, however, has begun to emerge only in the last twenty or so years. It relies very much on the flexible use of varying base-points in Grothendieck's theory of the fundamental group, and it appears to have taken some time after the inception of the arithmetic π_1 theory [45] for the importance of such a variation to be properly appreciated [12, 6, 16]. In fact, the impetus for taking it seriously came also for the most part from Hodge theory [13, 14]. As far as Diophantine problems are concerned, in a letter to Faltings [12] written shortly after the proof of the Mordell conjecture, Grothendieck proposed the remarkable conjecture that δ^{na} should be a bijection for compact hyperbolic curves. He expected such a statement to be directly relevant to the Mordell problem and probably its variants like conjecture (2). This expectation appears still to be rather reasonable. For one thing, it is evident that the conjecture is a hyperbolic analogue of the finiteness conjecture for Tate-Shafarevich groups. And then, profound progress is represented by the work of Nakamura, Tamagawa, and Mochizuki [42, 46, 41], where a statement of this sort is proved when points in the number field are replaced by dominant maps from other varieties. Some marginal insight might also be gleaned from [22] and [23] where a unipotent analogue of the Kummer map is related to Diophantine finiteness theorems. There, the ambient space inside which the analysis takes place is a classifying variety $H^1_f(\Gamma_v, U_n^{et})$ of torsors for the local unipotent étale fundamental group (rather than the Jacobian), while the finitely-generated group G is replaced by the image of a map

$$H^1_f(\Gamma_S, U_n^{et}) \to H^1_f(\Gamma_v, U_n^{et})$$

coming from a space of global torsors. Thereby, one obtains a new manifestation of the analytic strategy proving $X \cap Im[H^1_f(\Gamma_S, U_n^{et})]$ to be finite in some very special circumstances, and in general for a hyperbolic curve over $\mathbb{Q}$ if one admits standard conjectures from the theory of mixed motives (for example, the Fontaine-Mazur conjecture on geometric Galois representations). Fortunately, Chabauty's original method fits naturally into this setting as the technical foundation of the analytic part now becomes non-abelian p-adic Hodge theory and iterated integrals. Incidentally, some sense of the Diophantine content of these ideas can already be gained by deriving the *injectivity* of δ^{na} from the Mordell-Weil theorem.

It should be clear at this point that the Galois theory of the title refers in general to the theory of the fundamental group. Serge Lang was profoundly concerned with the fundamental group for a good part of his mathematical life. A rather haphazard list of evidence might be comprised of:

-his foundational work on unramified class field theory for varieties over finite fields, where he proves the surjectivity of the reciprocity map among many other things [25, 26];

-his study of the ubiquitous 'Lang torsor' [27];

-his work with Serre on fundamental groups of proper varieties in arbitrary characteristic [37];

-his extensive study with Kubert of the modular function field [24];

-his work with Katz [21] on finiteness theorems for relative π_1's that made possible the subsequent proof by Bloch [1], and then Kato and Saito [19, 20] of the finiteness of CH_0 for arithmetic schemes.

Besides these influential papers, the reader is referred to his beautiful AMS colloquium lectures [31] for a global perspective on the role of covering spaces in arithmetic.

Even towards the end of his life when his published work went in an increasingly analytic direction, he had a keen interest both in fundamental groups and in the analogy between hyperbolic manifolds and number fields wherein fundamental groups play a central role. In my last year of graduate school, he urged me strongly to study the work of Kato and Saito (and apply it to Arakelov theory!) even though it had been years since he had himself been involved with such questions. From the Spring of 2004, I recall a characteristically animated exchange in the course of which he explained to me a theorem of Geyer [11] stating that abelian subgroups of $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ are pro-cyclic. It was clear that he perceived this fact to fit nicely into his vivid ideas about the heat kernel [17], but in a manner that I failed (and still fail) to comprehend properly. (He was unfortunately secretive with his deeper reflections on the arithmetic significance of his later work, allowing only informal glimpses here and there. It is tempting but probably premature to speculate about a Galois theory that encompasses even Archimedean analysis.) The preoccupation with hyperbolic geometry that was evident even from the 70's ([30, 34, 35] and [36], chapters 8 and 9) could rather generally be construed as reflecting a persistent intuition about the relevance of fundamental groups to Diophantine problems. (An intuition that was shared by Grothendieck [12] and even Weil [48].)

As for the direct application of non-abelian fundamental groups to Diophantine geometry that we have outlined here, one can convincingly place it into the general framework of Lang's inquiries. He is discussing the theorem of Siegel in the following paragraph from the notes to chapter 8 of [32]:

> The general version used here was presented in [28] following Siegel's (and Mahler's) method. The Jacobian replaces the theta function, as usual, and the mechanism of the covering already used by Siegel appears here in its full formal clarity. It is striking to observe that in [25], I used the Jacobian in a formally analogous way to deal with the class field theory in function fields. In that case, Artin's reciprocity law was reduced to a formal computation in the isogeny $u \mapsto u^{(q)} - u$ of the Jacobian. In the present case, the heart of the proof is reduced to a formal computation of heights in the isogeny $u \mapsto mu + a$.

We have emphasized above the importance of the Kummer map

$$x \mapsto [\hat{\pi}_1(\bar{X}; b, x)] \in H^1(\Gamma, \hat{\pi}_1(\bar{X}; b)).$$

When X is defined over a finite field $\mathbf{F}_q$ and we replace $\hat{\pi}_1(\bar{X}, b)$ by its abelian quotient $H_1(\bar{X}, \hat{\mathbb{Z}})$, the map takes values in

$$H^1(\mathrm{Gal}(\bar{\mathbf{F}}_q/\mathbf{F}_q), H_1(\bar{X}, \hat{\mathbb{Z}})) = H_1(\bar{X}, \hat{\mathbb{Z}})/[(Fr - 1)H_1(\bar{X}, \hat{\mathbb{Z}})],$$

$Fr \in Gal(\bar{\mathbf{F}}_q/\mathbf{F}_q)$ being the Frobenius element. But this last group is nothing but the kernel

$$\hat{\pi}_1^{ab}(X)^0$$

of the structure map

$$\hat{\pi}_1^{ab}(X) \to \hat{\pi}_1(\mathrm{Spec}(\mathbf{F}_q)).$$

Thus the abelian quotient of the Kummer map becomes identified with the *reciprocity* map [19]

$$CH_0(X)^0 \to \hat{\pi}_1^{ab}(X)^0$$

of unramified class field theory evaluated on the cycle $(x) - (b)$. In other words, the reciprocity map is merely an 'abelianized' Kummer map in this situation. There is no choice but to interprete the reciprocity law [19, 20] as an 'abelianized Grothendieck conjecture' over finite fields.

Of course it is hard to imagine exactly what Lang himself found striking in the analogy when he wrote the lines quoted above. What is not hard to imagine is that he would have been very much at home with the ideas surrounding Grothendieck's conjecture and the non-abelian Kummer map.

References

[1] Bloch, Spencer Algebraic K-theory and classfield theory for arithmetic surfaces. Ann. of Math. (2) 114 (1981), no. 2, 229–265.

[2] Bogomolov, Fedor Alekseivich Sur l'algèbricité des représentations l-adiques. C. R. Acad. Sci. Paris Sér. A-B 290 (1980), no. 15, A701–A703.

[3] Model theory and algebraic geometry. An introduction to E. Hrushovski's proof of the geometric Mordell-Lang conjecture. Edited by Elisabeth Bouscaren. Lecture Notes in Mathematics, 1696. Springer-Verlag, Berlin, 1998. xvi+211 pp.

[4] Buium, A. Intersections in jet spaces and a conjecture of S. Lang. Ann. of Math. (2) 136 (1992), no. 3, 557–567.

[5] Chabauty, Claude Sur les points rationnels des variétés algèbriques dont l'irrégularité est supérieure à la dimension. C. R. Acad. Sci. Paris 212, (1941). 1022–1024.

[6] Deligne, P. Le groupe fondamental de la droite projective moins trois points. Galois groups over Q (Berkeley, CA, 1987), 79–297, Math. Sci. Res. Inst. Publ., 16, Springer, New York, 1989.

[7] Deninger, Christopher; Scholl, Anthony J. The Beĭlinson conjectures. L-functions and arithmetic (Durham, 1989), 173–209, London Math. Soc. Lecture Note Ser., 153, Cambridge Univ. Press, Cambridge, 1991. (

[8] Faltings, G. Endlichkeitssätze für abelsche Varietäten über Zahlkörpern. Invent. Math. 73 (1983), no. 3, 349–366.

[9] Faltings, Gerd Diophantine approximation on abelian varieties. Ann. of Math. (2) 133 (1991), no. 3, 549–576.

[10] Faltings, Gerd The general case of S. Lang's conjecture. Barsotti Symposium in Algebraic Geometry (Abano Terme, 1991), 175–182, Perspect. Math., 15, Academic Press, San Diego, CA, 1994.

[11] Geyer, Wulf-Dieter The automorphism group of the field of all algebraic numbers. Proceedings of the 5th School of Algebra (Rio de Janeiro, 1978), pp. 167–199, Soc. Brasil. Mat., Rio de Janeiro, 1978

[12] Grothendieck, Alexander Brief an G. Faltings. London Math. Soc. Lecture Note Ser., 242, Geometric Galois actions, 1, 49–58, Cambridge Univ. Press, Cambridge, 1997.

[13] Hain, Richard M. The geometry of the mixed Hodge structure on the fundamental group. Algebraic geometry, Bowdoin, 1985 (Brunswick, Maine, 1985), 247–282, Proc. Sympos. Pure Math., 46, Part 2, Amer. Math. Soc., Providence, RI, 1987.

[14] Hain, Richard M.; Zucker, Steven Unipotent variations of mixed Hodge structure. Invent. Math. 88 (1987), no. 1, 83–124.

[15] Hrushovski, Ehud The Mordell-Lang conjecture for function fields. J. Amer. Math. Soc. 9 (1996), no. 3, 667–690.

[16] Ihara, Yasutaka Braids, Galois groups, and some arithmetic functions. Proceedings of the International Congress of Mathematicians, Vol. I, II (Kyoto, 1990), 99–120, Math. Soc. Japan, Tokyo, 1991.

[17] Jorgenson, Jay; Lang, Serge The ubiquitous heat kernel. Mathematics unlimited—2001 and beyond, 655–683, Springer, Berlin, 2001.

[18] Kato, Kazuya p-adic Hodge theory and values of zeta functions of modular forms. Cohomologies p-adiques et applications arithmétiques. III. Astérisque No. 295 (2004), ix, 117–290.

[19] Kato, Kazuya; Saito, Shuji Unramified class field theory of arithmetical surfaces. Ann. of Math. (2) 118 (1983), no. 2, 241–275.

[20] Kato, Kazuya; Saito, Shuji Global class field theory of arithmetic schemes. Applications of algebraic K-theory to algebraic geometry and number theory, Part I, II (Boulder, Colo., 1983), 255–331, Contemp. Math., 55, Amer. Math. Soc., Providence, RI, 1986.

[21] Katz, Nicholas M.; Lang, Serge Finiteness theorems in geometric classfield theory. With an appendix by Kenneth A. Ribet. Enseign. Math. (2) 27 (1981), no. 3-4, 285–319 (1982).

[22] Kim, Minhyong The motivic fundamental group of $\mathbf{P}^1 \setminus \{0, 1, \infty\}$ and the theorem of Siegel. Invent. Math. 161 (2005), no. 3, 629–656.

[23] Kim, Minhyong The unipotent Albanese map and Selmer varieties for curves. Arxiv math.NT/0510441.

[24] Kubert, Daniel S.; Lang, Serge Modular units. Grundlehren der Mathematischen Wissenschaften, 244. Springer-Verlag, New York-Berlin, 1981. xiii+358 pp.

[25] Lang, Serge Unramified class field theory over function fields in several variables. Ann. of Math. (2) 64 (1956), 285–325.

[26] Lang, Serge Sur les séries L d'une variété algèbrique. Bull. Soc. Math. France 84 (1956), 385–407.

[27] Lang, Serge Algebraic groups over finite fields. Amer. J. Math. 78 (1956), 555–563.

[28] Lang, Serge Integral points on curves. Inst. Hautes Etudes Sci. Publ. Math. No. 6 1960 27–43.

[29] Lang, Serge Division points on curves. Ann. Mat. Pura Appl. (4) 70 1965 229–234.

[30] Lang, Serge Higher dimensional diophantine problems. Bull. Amer. Math. Soc. 80 (1974), 779–787.

[31] Lang, Serge Units and class groups in number theory and algebraic geometry. Bull. Amer. Math. Soc. (N.S.) 6 (1982), no. 3, 253–316.

[32] Lang, Serge Fundamentals of Diophantine geometry. Springer-Verlag, New York, 1983. xviii+370 pp.

[33] Lang, Serge Cyclotomic fields I and II. Combined second edition. With an appendix by Karl Rubin. Graduate Texts in Mathematics, 121. Springer-Verlag, New York, 1990. xviii+433 pp.

[34] Lang, Serge Hyperbolic and Diophantine analysis. Bull. Amer. Math. Soc. (N.S.) 14 (1986), no. 2, 159–205.

[35] Lang, Serge Introduction to complex hyperbolic spaces. Springer-Verlag, New York, 1987. viii+271 pp.

[36] Lang, Serge Number theory. III. Diophantine geometry. Encyclopaedia of Mathematical Sciences, 60. Springer-Verlag, Berlin, 1991. xiv+296 pp.

[37] Lang, Serge; Serre, Jean-Pierre Sur les revètements non ramifiés des variétés algèbriques. Amer. J. Math. 79 (1957), 319–330.

[38] Lang, Serge; Tate, John Principal homogeneous spaces over abelian varieties. Amer. J. Math. 80 1958 659–684.

[39] McQuillan, Michael Division points on semi-abelian varieties. Invent. Math. 120 (1995), no. 1, 143–159.

[40] Milne, James S. Étale cohomology. Princeton Mathematical Series, 33. Princeton University Press, Princeton, N.J., 1980. xiii+323 pp.

[41] Mochizuki, Shinichi The profinite Grothendieck conjecture for closed hyperbolic curves over number fields. J. Math. Sci. Univ. Tokyo 3 (1996), no. 3, 571–627.

[42] Nakamura, Hiroaki Galois rigidity of the étale fundamental groups of punctured projective lines. J. Reine Angew. Math. 411 (1990), 205–216.

[43] Perrin-Riou, Bernadette Fonctions L p-adiques. Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994), 400–410, Birkhäuser, Basel, 1995.

[44] Raynaud, M. Sous-variétés d'une variété abélienne et points de torsion. Arithmetic and geometry, Vol. I, 327–352, Progr. Math., 35, Birkhäuser Boston, Boston, MA, 1983.

[45] Revêtements étales et groupe fondamental. Séminaire de Géométrie Algébrique du Bois Marie 1960–1961 (SGA 1). Dirigé par Alexandre Grothendieck. Augmenté de deux exposés de M. Raynaud. Lecture Notes in Mathematics, Vol. 224. Springer-Verlag, Berlin-New York, 1971. xxii+447 pp.

[46] Tamagawa, Akio The Grothendieck conjecture for affine curves. Compositio Math. 109 (1997), no. 2, 135–194

[47] Vojta, Paul Siegel's theorem in the compact case. Ann. of Math. (2) 133 (1991), no. 3, 509–548.

[48] Weil, André Généralisations des fonctions abéliennes. J. de Math. P. et App. (IX) 17, 47–87.

[49] Wiles, Andrew Modular elliptic curves and Fermat's last theorem. Ann. of Math. (2) 141 (1995), no. 3, 443–551.