Preface to the Second Edition

In the preface to the first edition of this book I remarked on the paucity of introductory texts devoted to the arithmetic of elliptic curves. That unfortunate state of affairs has long since been remedied with the publication of many volumes, among which may be mentioned books by Cassels [43], Cremona [54], Husemöller [118], Knapp [127], McKean et. al [167], Milne [178], and Schmitt et. al [222] that highlight the arithmetic and modular theory, and books by Blake et. al [22], Cohen et. al [51], Hankerson et. al [107], and Washington [304] that concentrate on the use of elliptic curves in cryptography. However, even among this cornucopia of literature, I hope that this updated version of the original text will continue to be useful.

The past two decades have witnessed tremendous progress in the study of elliptic curves. Among the many highlights are the proof by Merel [170] of uniform boundedness for torsion points on elliptic curves over number fields, results of Rubin [215] and Kolyvagin [130] on the finiteness of Shafarevich–Tate groups and on the conjecture of Birch and Swinnerton-Dyer, the work of Wiles [311] on the modularity of elliptic curves, and the proof by Elkies [77] that there exist infinitely many supersingular primes. Although this introductory volume is unable to include proofs of these deep results, it will guide the reader along the beginning of the trail that ultimately leads to these summits.

My primary goals in preparing this second edition, over and above the pedagogical aims of the first edition, are the following:

- Update and expand results and references, especially in Appendix C, which includes a new section on the variation of the trace of Frobenius.
- Add a chapter devoted to algorithmic aspects of elliptic curves, with an emphasis on those features that are used in cryptography.
- Add a section on Szpiro's conjecture and the ABC conjecture.
- Correct, clarify, and simplify the proofs of some results.
- Correct numerous typographical and minor mathematical errors. However, since this volume has been entirely retypeset, I beg the reader's indulgence for any new typos that have been introduced.
- Significantly expand the selection of exercises.

It has been gratifying to see the first edition of this book become a standard text and reference in the subject. In order to maintain backward compatibility of cross-references, I have taken some care to leave the numbering system unchanged. Thus Proposition III.8.1 in the first edition remains Proposition III.8.1 in the second edition, and similarly for Exercise 3.5. New material has been assigned new numbers, and although there are many new exercises, they have been appended to the exercises from the first edition.

Electronic Resources: There are many computer packages that perform computations on elliptic curves. Of particular note are two free packages, Sage [275] and Pari [202], each of which implements an extensive collection of elliptic curve algorithms. For additional links to online elliptic curve resources, and for other material, the reader is invited to visit the *Arithmetic of Elliptic Curves* home page at

`www.math.brown.edu/˜jhs/AECHome.html`

No book is ever free from error or incapable of being improved. I would be delighted to receive comments, positive or negative, and corrections from you, the reader. You can send mail to me at

`jhs@math.brown.edu`

Acknowledgments for the Second Edition

Many people have sent me extensive comments and corrections since the appearance of the first edition in 1986. To all of them, including in particular the following, my deepest thanks: Jeffrey Achter, Andrew Bremner, Frank Calegari, Jesse Elliott, Kirsten Eisenträger, Xander Faber, Joe Fendel, W. Fensch, Alexandru Ghitza, Grigor Grigorov, Robert Gross, Harald Helfgott, Franz Lemmermeyer, Dino Lorenzini, Ronald van Luijk, David Masser, Martin Olsson, Chol Park, Bjorn Poonen, Michael Reid, Michael Rosen, Jordan Risov, Robert Sarvis, Ed Schaefer, René Schoof, Nigel Smart, Jeroen Spandaw, Douglas Squirrel, Katherine Stange, Sinan Unver, John Voight, Jianqiang Zhao, Michael Zieve.

Providence, Rhode Island JOSEPH H. SILVERMAN
November, 2008

Preface to the First Edition

The preface to a textbook frequently contains the author's justification for offering the public "another book" on a given subject. For our chosen topic, the arithmetic of elliptic curves, there is little need for such an apologia. Considering the vast amount of research currently being done in this area, the paucity of introductory texts is somewhat surprising. Parts of the theory are contained in various books of Lang, especially [135] and [140], and there are books of Koblitz [129] and Robert [210] (the latter now out of print) that concentrate on the analytic and modular theory. In addition, there are survey articles by Cassels [41], which is really a short book, and Tate [289], which is beautifully written, but includes no proofs. Thus the author hopes that this volume fills a real need, both for the serious student who wishes to learn basic facts about the arithmetic of elliptic curves and for the research mathematician who needs a reference source for those same basic facts.

Our approach is more algebraic than that taken in, say, [135] or [140], where many of the basic theorems are derived using complex analytic methods and the Lefschetz principle. For this reason, we have had to rely somewhat more on techniques from algebraic geometry. However, the geometry of (smooth) curves, which is essentially all that we use, does not require a great deal of machinery. And the small price paid in learning a little bit of algebraic geometry is amply repaid in a unity of exposition that, to the author, seems to be lacking when one makes extensive use of either the Lefschetz principle or lengthy, albeit elementary, calculations with explicit polynomial equations.

This last point is worth amplifying. It has been the author's experience that "elementary" proofs requiring page after page of algebra tend to be quite uninstructive. A student may be able to verify such a proof, line by line, and at the end will agree that the proof is complete. But little true understanding results from such a procedure. In this book, our policy is always to state when a result can be proven by such an elementary calculation, indicate briefly how that calculation might be done, and then to give a more enlightening proof that is based on general principles.

The basic (global) theorems in the arithmetic of elliptic curves are the Mordell–Weil theorem, which is proven in Chapter VIII and analyzed more closely in Chapter X, and Siegel's theorem, which is proven in Chapter IX. The reader desiring to reach these results fairly rapidly might take the following path:

I and II (briefly review), III (§§1–8), IV (§§1–6), V (§1)
VII (§§1–5), VIII (§§1–6), IX (§§1–7), X (§§1–6).

This material also makes a good one-semester course, possibly with some time left at the end for special topics. The present volume is built around the notes for such a course, taught by the author at M.I.T. during the spring term of 1983. Of course, there are many other ways to structure a course. For example, one might include all of chapters V and VI, skipping IX and, if pressed for time, X. Other important topics in the arithmetic of elliptic curves, which do not appear in this volume due to time and space limitations, are briefly discussed in Appendix C.

It is certainly true that some of the deepest results in the subject, such as Mazur's theorem bounding torsion over $\mathbb{Q}$ and Faltings' proof of the isogeny conjecture, require many of the resources of modern "SGA-style" algebraic geometry. On the other hand, one needs no machinery at all to write down the equation of an elliptic curve and to do explicit computations with it; so there are many important theorems whose proof requires nothing more than cleverness and hard work. Whether your inclination leans toward heavy machinery or imaginative calculations, you will find much that remains to be discovered in the arithmetic theory of elliptic curves. Happy Hunting!

Acknowledgements

In writing this book, I have consulted a great many sources. Citations have been included for major theorems, but many results that are now considered "standard" have been presented as such. In any case, I can claim no originality for any of the unlabeled theorems in this book, and I apologize in advance to anyone who may feel slighted. The excellent survey articles of Cassels [41] and Tate [289] served as guidelines for organizing the material. (The reader is especially urged to peruse the latter.) In addition to [41] and [289], other sources that were extensively consulted include [135], [139], [186], [210], and [236].

It would not be possible to catalogue all of the mathematicians from whom I learned this beautiful subject, but to all of them, my deepest thanks. I would especially like to thank John Tate, Barry Mazur, Serge Lang, and the "Elliptic Curves Seminar" group at Harvard (1977–1982), whose help and inspiration set me on the road that led to this book. I would also like to thank David Rohrlich and Bill McCallum for their careful reading of the original draft, Gary Cornell and the editorial staff at Springer-Verlag for encouraging me to undertake this project in the first place, and Ann Clee for her meticulous preparation of the manuscript. Finally, I would like to thank my wife, Susan, for her patience and understanding through the turbulent times during which this book was written, and also Deborah and Daniel, for providing much of the turbulence.

Cambridge, Massachusetts JOSEPH H. SILVERMAN
September, 1985

Acknowledgments for the Second Printing

I would like to thank the following people, who kindly provided corrections that have been incorporated into this second revised printing: Andrew Baker, Arthur Baragar, Wah Keung Chan, Yen-Mei (Julia) Chen, Bob Colemen, Fred Diamond, David Fried, Dick Gross, Ron Jacobwitz, Kevin Keating, Masato Kuwata, Peter Landweber, H.W.

Lenstra Jr., San Ling, Bill McCallum, David Masser, Hwasin Park, Elisabeth Pyle, Ken Ribet, John Rhodes, David Rohrlich, Mike Rosen, Rene Schoof, Udi de Shalit, Alice Silverberg, Glenn Stevens, John Tate, Edlyn Teske, Jaap Top, Paul van Mulbregt, Larry Washington, Don Zagier.

It has unfortunately not been possible to include in this second printing the many important results proven during the past six years, such as the work of Kolyvagin and Rubin on the Birch and Swinnerton-Dyer conjectures (C.16.5) and the finiteness of the Shafarevich–Tate group (X.4.13), Ribet's proof that the conjecture of Shimuara–Taniyama–Weil (C.16.4) implies Fermat's Last Theorem, and recent work of Mestre on elliptic curves of high rank (C §20). The inclusion of such material (and more) will have to await an eventual second edition, so the reader should be aware that some of our general discussion, especially in Appendix C, is out of date. In spite of this obsolescence, it is our hope that this book will continue to provide a useful introduction to the study of the arithmetic of elliptic curves.

Providence, Rhode Island JOSEPH H. SILVERMAN
August, 1992

Contents

CHAPTER I

CHAPTER II

CHAPTER III

CHAPTER IV

CHAPTER V

CHAPTER VI

CHAPTER VII

Elliptic Curves over Local Fields

CHAPTER VIII

Elliptic Curves over Global Fields

CHAPTER IX

Integral Points on Elliptic Curves

CHAPTER X

CHAPTER XI

APPENDIX A

APPENDIX B

APPENDIX C

Introduction

The study of Diophantine equations, that is, the solution of polynomial equations in integers or rational numbers, has a history stretching back to ancient Greece and beyond. The term *Diophantine geometry* is of more recent origin and refers to the study of Diophantine equations through a combination of techniques from algebraic number theory and algebraic geometry. On the one hand, the problem of finding integer and rational solutions to polynomial equations calls into play the tools of algebraic number theory that describe the rings and fields wherein those solutions lie. On the other hand, such a system of polynomial equations describes an algebraic variety, which is a geometric object. It is the interplay between these two points of view that is the subject of Diophantine geometry.

The simplest sort of equation is linear:

$$aX + bY = c, \qquad a, b, c \in \mathbb{Z}, \qquad a \text{ or } b \neq 0.$$

Such an equation always has rational solutions. It has integer solutions if and only if the greatest common divisor of a and b divides c, and if this occurs, then we can find all solutions using the Euclidean algorithm.

Next in order of difficulty come quadratic equations:

$$aX^2 + bXY + cY^2 + dX + eY + f = 0, \qquad a, \dots, f \in \mathbb{Z}, \quad a, b \text{ or } c \neq 0.$$

They describe conic sections, and by a suitable change of coordinates *with rational coefficients*, we can transform a given equation into one of the following forms:

$$\begin{aligned} AX^2 + BY^2 &= C &\qquad& \text{ellipse,}\\ AX^2 - BY^2 &= C && \text{hyperbola,}\\ AX + BY^2 &= 0 && \text{parabola.}\end{aligned}$$

For quadratic equations we have the following powerful theorem that aids in their solution.

Hasse–Minkowski Theorem 0.1. ([232, IV Theorem 8]) *Let $f(X,Y) \in \mathbb{Q}[X,Y]$ be a quadratic polynomial. The equation $f(X,Y) = 0$ has a solution $(x,y) \in \mathbb{Q}^2$ if and only if it has a solution $(x,y) \in \mathbb{R}^2$ and a solution $(x,y) \in \mathbb{Q}_p^2$ for every prime p. (Here $\mathbb{Q}_p$ is the field of p-adic numbers.)*

In other words, a quadratic polynomial has a solution in $\mathbb{Q}$ if and only if it has a solution in every completion of $\mathbb{Q}$. Hensel's lemma says that checking for solutions in $\mathbb{Q}_p$ is more or less the same as checking for solutions in the finite field $\mathbb{Z}/p\mathbb{Z}$, and this is turn is easily accomplished using quadratic reciprocity. We summarize the steps that go into the Diophantine analysis of quadratic equations.

(1) Analyze the equations over finite fields [quadratic reciprocity].

(2) Use this information to study the equations over complete local fields $\mathbb{Q}_p$ [Hensel's lemma]. (We must also analyze them over $\mathbb{R}$.)

(3) Piece together the local information to obtain results for the global field $\mathbb{Q}$ [Hasse principle].

Where does the geometry appear? Linear and quadratic equations in two variables define curves of genus zero. The above discussion says that we have a fairly good understanding of the arithmetic of such curves. The next simplest case, namely the arithmetic properties of curves of genus one (which are given by cubic equations in two variables), is our object of study in this book. The arithmetic of these so-called *elliptic curves* already presents complexities on which much current research is centered. Further, they provide a standard testing ground for conjectures and techniques that can then be fruitfully applied to the study of curves of higher genus and (abelian) varieties of higher dimension.

Briefly, the organization of this book is as follows. After two introductory chapters giving basic material on algebraic geometry, we start by studying the geometry of elliptic curves over algebraically closed fields (Chapter III). We then follow the program outlined above and investigate the properties of elliptic curves over finite fields (Chapter V), local fields (Chapters VI, VII), and global (number) fields (Chapters VIII, IX, X). Our understanding of elliptic curves over finite and local fields will be fairly satisfactory. However, it turns out that the analogue of the Hasse–Minkowski theorem is false for polynomials of degree greater than 2. This means that the transition from local to global is far more tenuous than in the degree 2 case. We study this problem in some detail in Chapter X. Finally, in Chapter XI we investigate computational aspects of the theory of elliptic curves, especially those that have become important in the field of cryptography.

The theory of elliptic curves is rich, varied, and amazingly vast. The original aim of this book was to provide an essentially self-contained introduction to the basic arithmetic properties of elliptic curves. Even such a limited goal proved to be too ambitious. The material described above is approximately half of what the author had hoped to include. The reader will find a brief discussion and list of references for the omitted topics in Appendix C, about half of which are covered in the companion volume [266] to this book.

Our other goal, that of being self-contained, has been more successful. We have, of course, felt free to state results that every reader should know, even when the proofs are far beyond the scope of this book. However, we have endeavored not to use such results for making further deductions. There are three major exceptions to this general policy. First, we do not prove that every elliptic curve over $\mathbb{C}$ is uniformized by elliptic functions (VI.5.1). This result fits most naturally into a discussion of modular functions, which is one of the omitted topics; it is covered [266, I §4] in the companion volume. Second, we do not prove that over a complete local field, the "nonsingular" points sit with finite index inside the set of all points (VII.6.1). This can be proven by quite explicit polynomial computations (cf. [283]), but they are rather lengthy and have not been included for lack of space. (This result is proven in the companion volume [266, IV §§8, 9].) Finally, in the study of integral points on elliptic curves, we make use of Roth's theorem (IX.1.4) without giving a proof. We include a brief discussion of the proof in (IX §8), and the reader who wishes to see the myriad details can proceed to one of the references listed there.

The prerequisites for reading this book are fairly modest. We assume that the reader has had a first course in algebraic number theory, and thus is acquainted with number fields, rings of integers, prime ideals, ramification, absolute values, completions, etc. The contents of any basic text on algebraic number theory, such as [142, Part I] or [25], should more than suffice. Chapter VI, which deals with elliptic curves over $\mathbb{C}$, assumes a familiarity with the basic principles of complex analysis. In Chapter X, we use a little bit of group cohomology, but just H^0 and H^1. The reader will find in Appendix B the cohomological facts needed to read Chapter X. Finally, since our approach is mainly algebraic, there is the question of background material in algebraic geometry. On the one hand, since much of the theory of elliptic curves can be obtained through the use of explicit equations and calculations, we do not want to require that the reader already know a great deal of algebraic geometry. On the other hand, this being a book on number theory and not algebraic geometry, it would not be reasonable to spend half the book developing from first principles the algebro-geometric facts that we will use. As a compromise, the first two chapters give an introduction to the algebraic geometry of varieties and curves, stating all of the facts that we need, giving complete references, and providing enough proofs so that the reader can gain a flavor for some of the basic techniques used in algebraic geometry.

Numerous exercises have been included at the end of each chapter. The reader desiring to gain a real understanding of the subject is urged to attempt as many as possible. Some of these exercises are (special cases of) results that have appeared in the literature. A list of comments and citations for the exercises may be found on page 461. Exercises with a single asterisk are somewhat more difficult, while two asterisks signal an unsolved problem.

References

Bibliographical references are enclosed in square brackets, e.g., [289, Theorem 6]. Cross-references to theorems, propositions, lemmas, etc., are given in full with the chapter roman numeral or appendix letter, e.g., (IV.3.1) and (B.2.1). Reference to an exercise is given by the chapter number followed by the exercise number, e.g., Exercise 3.6.

Standard Notation

Throughout this book, we use the symbols

$$\mathbb{Z},\ \mathbb{Q},\ \mathbb{R},\ \mathbb{C},\ \mathbb{F}_q,\ \text{and } \mathbb{Z}_\ell$$

to denote the integers, rational numbers, real numbers, complex numbers, a field with q elements, and the ℓ-adic integers, respectively. Further, if R is any ring, then R^* denotes the group of invertible elements of R, and if A is an abelian group, then $A[m]$ denotes the subgroup of A consisting of elements of order dividing m. For a more complete list of notation, see page 467.