Mitt. Math. Ges. Hamburg **22** (2004), 1–68

The Riemann hypothesis in characteristic p, its origin and development

Part 2. The first steps by Davenport and Hasse

Von *Peter Roquette (Heidelberg)*

Abstract

This paper is a continuation of Part 1. We report on how Hasse met Davenport in 1931, how the latter introduced his friend Hasse to the problem on diophantine congruences, and how Artin identified this problem with the Riemann hypothesis for function fields. We discuss Hasse's first proof for elliptic fields which used classic uniformization and complex multiplication; the idea for this developed during a discussion with Mordell about Siegel's great paper on binary diophantine equations.

Contents

1 Introduction

This paper is the second part of a larger work on the Riemann hypothesis for function fields; more parts will follow in due course. Part 1 has appeared in [Rq:2002a].

We have described in Part 1 how Artin's thesis in 1921 had triggered a remarkable development, viz., the systematic investigation of algebraic function fields with finite base fields. The aim of those investigations was to transfer the main structure theorems of number fields to the case of function fields with finite base fields, thus exhibiting the close analogy between these two classes of fields. Artin in his thesis had done this for quadratic function fields, and now arbitrary function fields were considered. This happened in the 1920s.

Today it is common usage to treat both cases, number fields and function fields with finite base fields, simultaneously under the name of "global fields". These global fields can be characterized by a set of axioms referring to their valuations. [1] We should be aware that these axioms resulted from the findings obtained during the said development in the 1920s.

[1] Such axioms were given by Hasse in his book "*Zahlentheorie*" (English translation [H:2002]) and also by Artin-Whaples in their well-known paper [A-Wh:1945]. The essential axiom is the product formula for valuations, combined with some natural finiteness conditions.

But not only *algebraic* number theory, also some *analytic* number theory was transferred to the function field case. Artin in his thesis had introduced the zeta function in the case of quadratic function fields. We have seen in Part 1 that F. K. Schmidt [FK:1926] had generalized Artin's theory for *arbitrary* function fields F with finite base fields, and he established a birational invariant theory of the zeta function

$$\zeta_F(s) = \prod_{\mathfrak{p}} \frac{1}{1 - |\mathfrak{p}|^{-s}} = \sum_{\mathfrak{a} \geq 0} |\mathfrak{a}|^{-s} \tag{1}$$

where $\mathfrak{p}$ ranges over the primes (or valuations) of the field F, and where $|\mathfrak{p}|$ denotes the absolute norm, i.e., the order of the residue field $F\mathfrak{p}$. On the right hand side, $\mathfrak{a}$ ranges over the positive divisors and $|\mathfrak{a}\,|$ is defined multiplicatively in terms of $|\mathfrak{p}|$. F. K. Schmidt obtained, among other results, the functional equation of $\zeta_F(s)$, and also class number formulas, similar to those which Artin had given in the quadratic case.

The "*Riemann hypothesis*" for a function field F with finite base field predicts that all zeros s of F. K. Schmidt's zeta function have real part $\Re(s) = \frac{1}{2}$. Artin in his thesis had verified this by numerical computation in a number of examples; he had discussed about 40 function fields over base fields with small primes, most of them elliptic (i.e., of genus 1) and some of genus 2.

In the following, when we talk about the "Riemann hypothesis" then we mean the "Riemann hypothesis for F.K.Schmidt's zeta function for function fields over finite base fields" in the sense we have just explained – except if we explicitly say otherwise.

However, as we have pointed out in Part 1, this Riemann hypothesis did not play a dominant role in the development of the 1920s, not in Artin's thesis and not in the papers by F. K. Schmidt or others. Instead, one of the main motivation for the transfer of analytic number theory to the function field case, was the establishment of *class field theory* for function fields. (At that time class field theory was based on analytic number theory, viz., L-functions and zeta functions. See [Rq:2001].)

Herglotz, who had been Artin's thesis advisor, spoke of a "curious fact" only when he reported that Artin had verified the Riemann hypothesis in some numerical examples. It seems that this opinion was shared by others too. There were only few voices which expressed some more explicit interest in a proof of the Riemann hypothesis. One of them was Artin himself, not in his thesis [2] but in some letters which he wrote in November 1921 to Herglotz (see section 2.2.3 of Part 1). Another one was Hasse who in his *Jahrbuch*-review of Artin's thesis [H:1924a] explicitly points to the Riemann hypothesis. But in

[2] In Artin's report [A:1921] on his thesis he did not even mention that he had verified the Riemann hypothesis in some numerical examples.

general it seems that in the 1920s, the proof of the Riemann hypothesis was not considered to be a pressing problem. In any case, nobody at that time had a definite idea how to approach it.

This situation changed dramatically in the early 1930s when Hasse took up the question and succeeded in proving the Riemann hypothesis for elliptic function fields. After that the Riemann hypothesis for arbitrary function fields moved into the center of mathematical research interest. Let us cite from a letter of Mordell to Hasse of March 9, 1933. Hasse had informed him that he (Hasse) had just succeeded in proving the Riemann hypothesis in the elliptic case. Mordell replied:

> "*. . . I was exceedingly interested in your mathematical news and was very glad to hear that you had completely knocked down the bottom out of* $y^2 \equiv f_4(x) \mod p$. *It is a wonderful achievement and I shall look forward with the greatest interest to seeing your paper in print. I hope you will make it as easy as possible for the reader to understand, without reference to all the theorems on Klassenkörpertheorie etc. For as this is the first case of any exact result for zeros of Zeta-functions on* $\Re(s) = \frac{1}{2}$, *the paper is sure to attract an enormous amount of attention. . .*"

In Mordell's notation, $f_4(x)$ denotes a monic polynomial of degree 4. Whenever necessary he assumes it to be non-degenerate, i.e., without multiple roots. Thus the equation $y^2 = f_4(x)$ defines an elliptic function field $F = K(x, y)$ over any field of characteristic $\neq 2$ which contains the coefficients of $f_4(x)$. And conversely, every elliptic function field $F|K$ of characteristic $\neq 2$ is generated by such equation – at least if the base field is finite since then there exists a prime divisor of degree 1, as F. K. Schmidt had shown. [3] Thus when Mordell speaks of the congruence $y^2 \equiv f_4(x) \mod p$ then he is referring to elliptic function fields over the prime field $\mathbb{F}_p$ of characteristic $p > 2$.

His prediction about the "enormous amount of attention" turned out not to be overdone.

Hasse had become interested into questions of this kind through his friend Harold Davenport who had been introduced to him by Mordell in 1931. At that time Davenport was working on the estimate of the number of solutions

[3] There is a slight discrepancy between the notion of "elliptic function field" as used by Hasse, and that of today's use. Hasse required that the function field $F|K$ is of genus 1 whereas today it is often required in addition that $F|K$ admits a prime of degree 1. If the base field K is finite then the two notions coincide, due to F. K. Schmidt's theorem. But for a general base field K, a function field $F|K$ of genus 1 need not have a K-rational prime. In fact, Shafarevich [Sh:1957] showed that over an algebraic number field there exist function fields of genus 1 whose minimal prime divisor degree is arbitrary large; this had been a long time open problem.

of certain diophantine congruences. Hasse tried to improve Davenport's and Mordell's results by putting them into the framework of modern algebraic number theory. It was Artin who suggested to Hasse that the Davenport-Mordell problem was in fact identical with the Riemann hypothesis for function fields. This happened in November 1932 and is well documented by the letters from Hasse to Davenport and by Hasse's own notes. (See secion 3.4 .) Thus Artin, who had not published anything about the Riemann hypothesis since his thesis, was instrumental in directing Hasse to this problem.

The connection between the Riemann hypothesis and Davenport's problem on diophantine congruences was a completely new aspect; it had not been considered before. It motivated Hasse to work on the Riemann hypothesis for the next 10 years.

In the present Part 2 we will report on how Hasse met Davenport in 1931, how the latter introduced his friend Hasse to his problem on diophantine congruences, and how Artin identified this problem with the Riemann hypothesis. We will discuss Hasse's first proof for elliptic function fields, which used classic uniformization and complex multiplication; the idea for this developed during a discussion with Mordell about Siegel's great paper on binary diophantine equations. Hasse's second and final proof, working directly in characteristic p, will be discussed in the next Part.

Remark 1: Nowadays it is common to use geometric language and to speak of "curves" instead of "function fields". Accordingly one speaks of the "Riemann hypothesis for curves". Both terminologies are essentially equivalent since any irreducible curve defines a function field as its field of rational functions and, conversely, every function field (of dimension 1) is the field of rational functions on some irreducible curve. This curve can be chosen to be smooth [4] and projective, and then it is essentially uniquely determined by its function field.

In the early days of Hasse and Artin, the terminology of "fields" was prevalent because the main motivation came from the comparison of number fields and function fields. Later it was observed that some important special features of function fields have no immediate analogue in the number field case and could be better understood through the analogy to the geometric situation with curves. This led to the gradual adoption of the geometric language which today is in general use because of its flexibility. We can even pinpoint the date when the geometric language has first been considered in this context in a relevant manner. This was July 16, 1937 when A. Weil in a letter to Hasse pointed out that the Italian school and Severi's "*Trattato*" [Sev:1927] had results which, when transferred to the algebraic situation, would be of relevance. This will be discussed in detail in one of the future Parts.

In this Part 2 we shall use the terminology of "function fields", like Artin and

[4] at least if the base field is perfect.

Hasse did. But it should be understood that everything can easily be translated into the geometric language.

Remark 2: In the preparation of this paper we have used not only published material but also the information contained in personal documents like letters, manuscripts etc. All those documents which we cite are contained in the *Handschriftenabteilung* of Göttingen University Library, except when we explicitly mention another source. As a general rule, letters which were addressed to Hasse can be found in Göttingen, whereas letters which Hasse wrote to other people are preserved at other places (if preserved at all). Letters from Hasse to Mordell we have found in the archives of King's College, and those from Hasse to Davenport at Trinity College, both in Cambridge, England.

Although quite a number of letters from the Hasse correspondence is preserved, the reader should be aware that, on the other hand, quite another number of letters seems to be lost. What we have found does not constitute a complete set of the Hasse correspondence.

Remark 3: Since publication of Part 1 [Rq:2002a] we have found some further documents which led us to make some changes in that manuscript. These changes concern certain details only, mostly the timing of events, and will be mentioned here in due course. In our homepage we show the manuscript of Part 1 with those changes already performed.

Acknowledgement: I would like to thank Keith Conrad for careful proofreading and many valuable comments, not only for this manuscript but also for several others from my homepage.

2 Biographic Notes

Let us start with some notes on the biographies of the three main actors in the scene of this Part 2. We do not attempt complete biographies here. We have included those biographical informations only which seem to be relevant to our topic. For more we refer to the standard literature.

Besides of the three names mentioned in the next three sections there are also E. Artin and F. K. Schmidt who were involved in the development. Biographic notes about them are to be found in Part 1.

2.1 Hasse 1930

Helmut Hasse was born 1898 [5] in Kassel, an old medium sized town in Hessen not far south from the university town of Göttingen. When he was 14 his

[5] Thus Hasse was of the same age as Artin. In their joint paper [A-H:1925] they mention that the write-up of the paper was done by the "younger" of the two. Since both were born

family moved to Berlin. In 1915 he obtained his "*Abitur*" and volunteered for service in the navy. In 1917 he got permission to study at the University of Kiel, with Toeplitz. One year later he moved to Göttingen where he studied mainly with Hecke. When the latter left Göttingen in 1920, Hasse changed again universities and went to Marburg to study with Kurt Hensel.

In May 1921 he received his doctorate. In his thesis he formulated the famous "Local-Global Principle" for quadratic forms over the rational number field $\mathbb{Q}$. His papers on quadratic forms culminated 1924 in the (nontrivial) proof that the Local-Global Principle for quadratic forms holds over an arbitrary algebraic number field [H:1924]. With this paper he solved, at least partially [6], the 13*th* problem of Hilbert which reads: [7]

> "*...to solve a given quadratic equation with algebraic numerical coefficients in any number of variables by integral or fractional numbers belonging to the algebraic realm of rationality determined by the coefficients.*"

(We remark that Artin too had solved one of Hilbert's problems, the 17*th*, on positive definite rational functions [A:1927]. Thus both Hasse and Artin belong to the "honors class of the mathematical community" in the sense of Hermann Weyl. See [Weyl:1944]. [8])

In the meantime Hasse had obtained a position as *Privatdozent* at the University of Kiel. The mathematicians in Kiel had close contacts to their colleagues in Hamburg. And so Hasse and Artin met frequently, and a long lasting friendly relationship began, documented in a number of letters between Artin and Hasse. [9] In 1925 Hasse accepted a professorship at the University of Halle. In the summer semester 1930 he returned to Marburg, as the successor of his academic teacher and "fatherly friend" (*väterlicher Freund*) Kurt Hensel.

in the same year it takes a precise knowledge of their birth dates to decide who indeed was the younger one. Fact is that Hasse was 175 days younger than Artin.

[6] Hilbert's wording admits two interpretations. One of them is to regard the phrase "integer or fractional numbers" as denoting arbitrary numbers of the number field in question. In this interpretation Hasse could be said to have solved the problem completely. The other interpretation is that Hilbert actually meant two different problems: One is to solve the quadratic equation in integers of the field, and the second problem requires solutions with arbitrary numbers of the field. In this interpretation, which is usually accepted by the mathematical community because it would generalize Minkowski's work for the rationals to arbitrary number fields, Hasse would have solved only one of the two problems. The other problem (solution in integers) has been studied by Siegel and others.

[7] Hilbert in his address of 1900 used German language. The following is a free translation. On this occasion we would like to state that whenever we cite a source which is originally written in German, or in French, then we have translated it into English.

[8] The expression "honors class" has been taken over from Weyl by Yandell in his book [Yan:2002] about the people who worked on Hilbert's problems.

[9] We plan to publish the commented Artin-Hasse correspondence, jointly with G. Frei.

Following up his work on quadratic forms, Hasse got interested into higher reciprocity laws and class field theory. He wrote his seminal 3-part "class field report" [H:1926], [H:1927a], [H:1930] which influenced a whole generation of mathematicians. He worked on the theory of complex multiplication; his two great papers [H:1927], [H:1931] on this topic are still of current interest. Starting from 1928 there developed a cooperation with Richard Brauer and Emmy Noether which culminated 1931 in their famous paper on the Local-Global Principle for algebras over number fields [BHN:1932]. This was not only an important achievement in the structure theory of algebras and their representations, but it also had far-reaching consequences for class field theory, preparing the way for the introduction of algebraic cohomology.

From this brief vita we see that by 1930 Hasse, like Artin, had become a successful mathematician and one of the leaders in number theory research. [10] But until then, nothing in his vita points to a particular interest in the Riemann hypothesis, although he kept himself well informed about the advances in the theory of function fields. Hasse's special interest in the Riemann hypothesis arose when he met Davenport. This happened after Hasse had written a letter to Mordell in November 1930.

2.2 Mordell 1930

Louis Joel Mordell was born in Philadelphia, Pennsylvania, in the year 1888. Thus he was 10 years older than Hasse. At the age of 19 he went to Cambridge, England and obtained a scholarship at St. John's College. His main interest in mathematics was number theory, in the direction of diophantine equations. In particular he investigated the equation $y^2 = x^3 + k$ over and over again from different angles, so that this equation sometimes was called "Mordell's equation" (although it had been studied much earlier already, e.g., by Fermat). As he himself reports, in Cambridge at that time there was not much interest in such problems. He considered himself as "self taught". In 1922 he became Reader at the University in Manchester, 1923 Professor there.

Mordell's name became widely known through his paper [Mor:1922] where he proved the finite basis theorem for the group of points of an elliptic curve over the rationals. His theorem was later generalized by A. Weil to abelian varieties over arbitrary number fields of finite degree [W:1928]. This theorem is nowadays known as "Mordell-Weil theorem". [11] At the end of the same

[10] For a more complete biography of Hasse see [Fr:1985] and [Fr-Rq:2002].

[11] Mordell himself never accepted this terminology. He insisted that these are two different theorems, one to be called "Mordell's theorem" and the other "Weil's theorem". But somehow the mathematical community did not heed to his wish. – In a similar situation, however, the naming of a theorem by the mathematical community turned out to be different. For, Hasse proved the Riemann hypothesis for elliptic curves, and his theorem was later generalized by A. Weil to curves of arbitrary genus (over finite base fields). Nevertheless the theorem is *not*

paper [Mor:1922] we find also the famous "Mordell conjecture", raising the question whether absolutely irreducible curves of genus > 1 over the rationals should have only finitely many rational points. For a long time this conjecture withstood all attempts for proof until finally Faltings [Fa:1983] succeeded.

In the "Dictionary of Scientific Biography" Mordell is characterized as "problem solver, not a system builder". In fact he had acquired an enormous amount of knowledge about solutions of special diophantine equations. His book [Mor:1969] is a treasure of interesting examples. He has never made a secret of his dislike of pure "high brow" theories (this was his standard expression); in his eyes such theories could be justified, if at all, only if they do contribute to solving interesting problems.

In Manchester Mordell was able to build a strong school of mathematicians, but this was later, in the last part of the thirties. During the years 1924–1927 he had a brilliant student, Harold Davenport. [12]

We do not know when Mordell and Hasse met for the first time. The first letter from Hasse to Mordell which is preserved, dates from Nov. 26, 1928. We know that on July 16, 1930 Mordell had visited Marburg for a colloquium talk on the invitation of Hasse.

Hasse's letter to Mordell which we mentioned at the end of the foregoing section is dated Nov. 25, 1930. Mordell had asked Hasse for his opinion about Maclagan Wedderburn who had been proposed for election into the Royal Society. Now Hasse responded to this request. At the end of the letter we read: [13]

> "*. . . I enclose, as you wanted, a few words about the work of Wedderburn. I am much enjoyed to have such an occasion for doing something for the glory of this big man.*
>
> *I hope to agree with you that it was better to write that acknowledgement in German. It would be better, and easier for me too, to write this letter in German. But I am happy to have got an opportunity for practice my knowledge in English. You may be interested to hear that I have continued my zealous studies in your language this summer. . .*"

And Hasse continues, apparently in a rather quaint [14] English:

> "*In order to have further occasion for applying and enriching my knowledges I would much like to get a young English fellow at home. It would be very kind of you, if you could send me one of your*

called "Hasse-Weil theorem".

[12] More biographical information about Mordell can be found in Cassels' article [Ca:1974].

[13] Hasse's letter is written in English.

[14] This is the expression which Davenport used in a letter to Mordell.

> *students during next summer term (April-July). We would invite that student to dwell and eat with us. He would be obliged to speak English with us at any time we are together (at breakfast, dinner, tea, lunch etc.)...From my point of view it would be best, if he were student of pure mathematics out of an advanced course of yours...I would much like to hear from you, whether you know a clever and handsome fellow for this purpose.*"

Thus Hasse wished to polish up his English. He had found his English not sufficient to write a letter of opinion on Wedderburn for the Royal Society. At those times, English had not yet become the *lingua franca* for science and not, in particular, for mathematics. Usually mathematicians wrote their papers (or books, or letters) in one of the major European languages, viz. English, French, German or Italian. It was tacitly assumed that their mathematical colleagues everywhere were able to read any of these languages. But, of course, it was another thing to be able to *write* in a foreign language. [15]

The correspondence with Mordell was perhaps not the only motivation for Hasse to update his English. Quite generally he wished to read English texts more easily than he was able to at the time. [16] Also, it may be that already at this time he contemplated to write a paper in English language. In fact, we know that half a year later, in May 1931, Hasse submitted a long paper in English to the Transactions of the American Mathematical Society [H:1932b]. That paper provided a thorough introduction to the new results and methods on cyclic algebras which Hasse had obtained recently, jointly with R. Brauer and Emmy Noether. Although for that paper he could enlist linguistic advice elsewhere [17], Hasse wished that in the future he would be able to write English without having to rely on external help; so he turned to Mordell with his request.

Already two days after Hasse had sent the letter, on Nov. 27, 1930, Mordell replied to him as follows:

> "*...I can suggest the very person you want to go to Marburg. Mr. Harold Davenport, Trinity College, Cambrigde. He was formerly one of our students, the best we have had for many years. He is*

[15] Even Hilbert had to admit, in a letter to Felix Klein of May 23, 1893, that he was not able to write a paper in English. See [Fr:1985a]. We note that at those times, in a German *Gymnasium* (secondary school) the teaching of modern languages was somewhat a matter of second importance as compared with the teaching of classical Greek and Latin, the latter being considered to carry *Bildungswert*. And among the modern languages taught in school, French was dominant.

[16] Hasse himself had told this to Günther Frei when asked about the case. I am grateful to Professor Frei for providing me with this information.

[17] He got it from H. T. Engstrom, a young American postdoc who studied in Göttingen with Emmy Noether at that time.

> *now doing research, and lately he has proved some such result as* $\sum_{n=0}^{p-1}\left(\frac{n^4+an^2+bn+c}{p}\right)=\mathcal{O}(p^{\frac{3}{4}})$ *where the left hand ( ) is the symbol of quadratic reciprocity. I think Hopf in the* Zeitschrift *a year or two ago showed the right hand side* $<\frac{p}{6}$ *!!* [18] *He is interested in certain aspects of number theory and I believe he would be free to go. I have written to him and asked him to write direct to you...*"

2.3 Davenport 1930

Davenport needed some time to think it over. On Nov. 30, 1930 he wrote to Mordell thanking him for passing Hasse's request on to him. He regarded it as a great compliment, especially, he wrote, in view of the phrase "handsome fellow". And he was very interested by Prof. Hasse's scheme. The only possible objections which he could see were the following:

1) "*...there may be nobody at Marburg interested in the analytical theory of numbers,*"
2) "*there may be too many distractions there for me to get much work done; and I must write a fellowship thesis by next August...*"

By Dec. 7, 1930 Davenport had obviously waived these objections since he wrote to Hasse:

> "*Dear Prof. Hasse,*
>
> *Prof. Mordell has told me of your letter to him, in which you say that you would like to know of an advanced English student of pure mathematics, whom you could invite to Marburg next summer term. May I offer you my services?*
>
> *I used to be a student of Mordell's at Manchester, but for the last three years I have been studying here. I am particularly interested in the analytical theory of numbers – Gitterpunktprobleme, ζ-function, etc. Are you interested in these subjects, or is there anyone else at Marburg who is? So far I have only written two short papers, which will appear soon in the Journal of the London Mathematical Society; one on the distribution of quadratic residues* mod (p), *the other on Dirichlet's L-functions.*
>
> *I am 23 years old, and not at all 'handsome' (as you required in your letter). Also I do not swim or drink beer – and I understand that these are the principal recreations in Germany...*"

[18] Mordell seems to have forgotten a square root sign, for Hopf had proved the right hand side $<\frac{p}{\sqrt{6}}$. See section 3.1.

Hasse seems not to have minded the “shortcomings” with which Davenport had advertised himself and he sent Davenport a definite invitation. And in this letter he said that he is very interested in the analytical theory of numbers and that the Seminar at Marburg is inviting several distinguished German mathematicians to give lectures there next summer. [19]

So in the next summer (1931) Davenport stayed as “language teacher” with the Hasses in Marburg. There developed a friendship for many years between the Hasse family and the younger Davenport. Certainly this had an effect on Hasse's proficiency in English, but at the same time Davenport succeeded to raise Hasse's interest in English history, English literature and quite generally in everything which was considered as “typically English”. Hasse kept this interest throughout his life. On the other hand, Davenport also profited from this contact; later on he was fluent in German. [20]

This friendship between Hasse and Davenport had a very remarkable consequence for the work of Hasse in the 1930s. For, the conversation between the two was not confined to the English language and literature but, of course, it soon included mathematics. It seems quite natural that one of the first questions of Hasse to his younger colleague was about Davenport's results in the two papers which he had mentioned in his letter. In particular the paper on the distribution of quadratic residues mod p got the attention of Hasse. We will report about it and more in the next sections.

The article for Davenport in the “Dictionary of Scientific Biography” says that he became a “*natural academic leader*”, and one of the “*most influential mathematicians of his time*”. The following citation from Rogers' biography [Rog:1972] will explain this in some more detail:

> “*... the extent to which he helped others can only be guessed, he was probably responsible for encouraging work at least as extensive as his own. ... He made his collaborators and colleagues his friends, and gave them generously of his humour and wisdom. He made a practice of writing helpful letters to all who approached him on mathematical matters whether they were professionals, students,*

[19] This letter not preserved, but we know about it since Davenport reports on it to Mordell on Dec 13, 1930.

[20] Excerpt from a letter of Davenport to Mordell in September 1931, from Bad Elster where he was staying with Hasse during the annual meeting of the DMV (=*Deutsche Mathematiker Vereinigung*):

“*The Hasses and I have been on a motor tour during the past 12 days, in which we have visited the Black Forest, Switzerland, the Italian Lakes, and Tyrol, with very much pleasure and edification. Hasse is now taking an active part in the D. M. V. congress here, I a more passive part. I can never be sufficiently grateful to you for passing on Hasse's invitation to me: I have had an excellent time in Marburg.*”

By the way, the car of that motor tour was Davenport's; at that time Hasse did not own a car.

> *amateurs or even cranks. By correspondence and by direct contact he stimulated and encouraged many mathematicians to do much of their best mathematics. . .*"

Two disciples of Davenport were awarded the Fields Medal. [21]

But all this came much later. When Davenport was visiting Hasse in 1931 he was 23, a young mathematician who had studied from 1924 to 1927 in Manchester with Mordell, and thereafter in Cambridge at Trinity College. He had distinguished himself in the examinations at both Manchester and Cambridge. Now he was preparing a thesis for his fellowship award (which he would obtain in 1932) and had just published his first papers. [22]

His results in these papers were considered to be important, as Mordell had mentioned in his letter to Hasse cited above. But in Marburg he was exposed for the first time with what was called "Modern Algebra", which means thinking in terms of algebraic structures like fields, rings, ideals etc., as it was propagated by Emmy Noether. Hasse had adopted Emmy Noether's ideas to a large degree. [23] In their correspondence we can read that Hasse, upon request from Davenport, explained to him the fundamentals about finite fields and their multiplicative characters. Also, Davenport learned the theory of algebraic numbers from the classical work of Dirichlet-Dedekind which Hasse had recommended to him. [24]

Halberstam [25] reports that in later years, Davenport would say that although he had learned a great deal from Hasse, he had not learned nearly as much as he would have done if he had been "less pig-headed". Rogers [Rog:1972] interprets this term as "more receptive". I am inclined to interpret this "pig-headedness" as a sign of intellectual independence, not absorbing Hasse's ideas and Hasse's style without critical scrutiny, and standing on his own.

In any case, it seems to me quite remarkable that a close friendship between Hasse and the nine years younger Davenport arose, both being (or becoming) quite dominant characters, of different mathematical tastes and different academic backgrounds. Hasse seems to have recognized the outstanding mathematical stature of Davenport, and accordingly he treated Davenport not as a young student to be educated but as a colleague on equal terms. From their correspondence we see that Hasse often reported to his young friend about his own (Hasse's) work, his problems and results. And he asked Davenport for his

[21] K. F. Roth 1958 and A. Baker 1970.

[22] For more details of Davenport's biography see [Mor:1971], [Rog:1972].

[23] See Hasse's 1929 lecture in Prague on "*Die moderne algebraische Methode*" (The modern algebraic method), one year before he met Davenport [H:1930a].

[24] Frei [Fr:1977] reports that Hasse himself, as a young man of 18 during his service in the Navy, had read the book of Dirichlet-Dedekind on the recommendation of his school teacher, Dr. Herrmann Wolff.

[25] In his comments to the Collected Works of Davenport [Da:1977].

opinion which he highly valued although he did not always share this opinion.

3 Estimating character sums and exponential sums.

3.1 Distribution of quadratic residues

In Mordell's letter to Hasse as cited in section 2.2 , he had mentioned Davenport's recent result that for a biquadratic polynomial

$$f(x) = x^4 + ax^2 + bx + c$$

with integer coefficients, one has

$$S(f) := \sum_{x \bmod p} \left(\frac{f(x)}{p}\right) = \mathcal{O}(p^{\frac{3}{4}}) \qquad \text{for} \quad p \to \infty\,. \tag{2}$$

Here, p is an odd prime number and $\left(\frac{z}{p}\right)$ denotes the quadratic residue character modulo p , defined for $z \in \mathbb{Z}$, which assumes the value $1, -1$ or 0 according to whether z is a quadratic residue, or z is a quadratic non-residue, or $z \equiv 0 \bmod p$ respectively. On the right hand side of (2) the symbol $\mathcal{O}$ is the so-called Landau symbol; the relation (2) means that there is a constant C, not depending on p and not on a, b, c, such that $|S(f)| < Cp^{\frac{3}{4}}$.

On first sight, Davenport's result (2) looks like a rather special technical lemma. But then, why did Mordell call this a "really significant result" ? (He did so in [Mor:1971] where he gave a survey of Davenport's results.)

The importance of statement (2) stems from its connection to various number theoretic problems. Davenport's motivation in his paper [Da:1930] is the distribution of quadratic residues. [26] Let $n \in \mathbb{N}$. For a prime $p > n$ consider sequences of n consecutive quadratic residues modulo p. Do there exist such sequences and if so, how many (modulo p) ? This is an old problem going back to Gauss. Davenport cites Jacobsthal [Ja:1906], [Ja:1910] for $n = 2, 3$. [27]

For arbitrary n, it was expected that the number R_n of such sequences is about $\frac{p}{2^n}$. More precisely, it was expected that

$$R_n = \frac{p}{2^n} + \mathcal{O}(\sqrt{p}) \qquad \text{for} \qquad p \to \infty\,. \tag{3}$$

[26] Another, perhaps more important problem connected with estimation problems like (2), is the number of solutions of diophantine congruences; this we shall discuss in the next sections.

[27] We had mentioned these papers in Part 1 already, in connection with Artin's work. See section 2.3.2 of Part 1.

In particular it follows that $R_n \to \infty$ for $p \to \infty$. Similarly, one may also consider the number N_n of sequences of n consecutive non-residues modulo p or, more generally, of sequences $a, a+1, a+2, \ldots, a+n-1$ with prescribed quadratic residue characters $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ where each $\varepsilon_i = \pm 1$. For brevity we will discuss here sequences of consecutive residues only although Davenport includes also these other problems.

Davenport considers the cases $n = 4$ or 5 and he shows that

$$R_4 = \frac{p}{16} + \mathcal{O}(p^{\frac{3}{4}}) \qquad \text{and} \qquad R_5 = \frac{p}{32} + \mathcal{O}(p^{\frac{3}{4}})\,. \tag{4}$$

We see that this is a partial result towards the expected (3).

The relation of this problem to the problem of estimating sums like $S(f)$ is seen if the definition of the quadratic residue symbol is put into the form

$$1+\left(\frac{z}{p}\right) = \begin{cases} 2 & \text{if } z \text{ is a quadratic residue mod } p \\ 0 & \text{if } z \text{ is a quadratic non-residue mod } p \\ 1 & \text{if } z \equiv 0 \bmod p\,. \end{cases} \tag{5}$$

In any case, $1+\left(\frac{z}{p}\right)$ is the number of solutions x modulo p of the congruence $x^2 \equiv z \bmod p$. From this it is almost immediate that

$$R_n = \frac{1}{2^n} \sum_{1 \le x \le p-n} \prod_{0 \le i < n} 1+\left(\frac{x+i}{p}\right)\,.$$

Expanding the product it is seen that R_n differs from $\frac{p}{2^n}$ by a number, not greater than 2^n, of sums of the type

$$S(f) = \sum_{x \bmod p} \left(\frac{f(x)}{p}\right) \tag{6}$$

for polynomials $f(x)$ of the form

$$f(x) = (x+a_1)(x+a_2)\cdots(x+a_r) \quad \text{with} \quad a_i \not\equiv a_j \bmod p \quad (i \ne j) \tag{7}$$

where $r \le n$. Thus we are led to the estimation of sums (6) which are similar to (2) but for polynomials of the form (7).

Davenport's proof of (4) consists of showing that $S(f) = \mathcal{O}(p^{\frac{3}{4}})$ for degrees $r = 3$ and $r = 4$. The cases $r = 1$ and $r = 2$ were easy and well known, while in the case $r = 5$ only the special polynomial $f(x) = x(x+1)(x+2)(x+3)(x+4)$ had to be considered which Davenport could handle too.

How did Davenport become interested in this problem of distribution of quadratic residues ? Mordell [Mor:1971a] writes that the problem had been given to Davenport by Littlewood, who supervised the work of young Davenport. On the other hand, Rogers [Rog:1972] reports that Littlewood regarded his supervision of Davenport as "nominal", saying that Davenport thought of his own problems and that he (Littlewood) just read his work and made encouraging noises. Thus it appears that Davenport got interested in the problem after reading papers.

Some time before Davenport wrote his paper, but in the same year, there had appeared a paper by Heinz Hopf [Ho:1930] who had proved $|S(f)| < \frac{p}{\sqrt{6}}$ for degree $r = 4$, if p is sufficiently large. We have seen above that Mordell had mentioned Hopf's paper in his letter to Hasse and had put two exclamation signs behind it, in order to stress that Davenport's result was much stronger. Davenport too cites Hopf's result in the introduction to his paper. We can imagine that Davenport, when reading Hopf's paper, found that he himself could do better and thus became interested in the subject. And when he reported to Littlewood about it then he was encouraged by Littlewood's "noises".

It seems quite remarkable that the brief and unpretentious note by Hopf [28] on the distribution problem for quadratic residues has caused, via Davenport–Mordell–Artin–Hasse, the tremendous step forward to the proof of the Riemann hypothesis, and more. "*Problems*", says Mordell, "*are the lifeblood of mathematics*". Here we see an outstanding example. Let us cite Mordell's own words from [Mor:1971a]:

> "*It is well known what an important part has been played by problems, even of the simplest character, in furthering research, discovery and the advancement of mathematics... The solution of a problem frequently requires new ideas and new methods. The generalization it suggests, its consideration from a different point of view or its rephrasing may lead to a new problem of far greater significance than the original one which may turn out to be only a very special case of a general theorem. Sometimes it seems almost incredible what striking and far-reaching fundamental developments have arisen in directions which seem very remote indeed from the problem from which they arose. Problems are the lifeblood of mathematics.*"

When Hasse wrote his textbook "*Vorlesungen über Zahlentheorie*" [H:1950] he included the distribution problem for quadratic residues into his book – as

[28] Most of the work of Heinz Hopf belongs to algebraic topology. In the course of time he rose to one of the leading algebraic topologists. The paper in question seems to be an outgrowth of his student years in Berlin where he had been close to I. Schur, scientifically and personally. See [Fr-St:1999].

an example of an elementary number theoretic problem which leads to the Riemann hypothesis for function fields. Since his book is meant to be an elementary textbook, Hasse does not present there a proof of the Riemann hypothesis. But he devotes the whole chapter §10 of his book to the distribution problem for quadratic residues. He discusses the cases $n = 2, 3$ very thoroughly and gives some kind of preview for higher n. From all we know about Hasse and his relation to Davenport it seems to me quite evident that the inclusion of this problem into Hasse's book can be traced back to the summer of 1931 when he was confronted with it by Davenport.

3.2 Diophantine congruences

Now, the distribution problem for quadratic residues is certainly an interesting problem of number theory but when Mordell called Davenport's result "really significant" then he was not referring to that problem only. There is a more fundamental problem of number theory connected with the sums of type (2), namely counting the number of solutions of diophantine congruences.

Let $f(x, y)$ be a polynomial of two variables with integer coefficients; in this context $f(x, y)$ is assumed to be absolutely irreducible. For a prime number p we consider the "diophantine congruence"

$$f(x, y) \equiv 0 \bmod p\,. \tag{8}$$

Let N denote the number of solutions modulo p; we also write $N[f(x, y) \equiv 0 \bmod p\,]$. There arises the question whether such solutions exist, i.e., whether $N > 0$, at least for large prime numbers p. If so then what can be said about the growth of N for $p \to \infty$? From heuristic arguments it was expected that N is about p. More precisely, it was expected that

$$N[f(x, y) \equiv 0 \bmod p\,] = p + \mathcal{O}(\sqrt{p}) \qquad \text{for} \qquad p \to \infty\,. \tag{9}$$

At the time of Davenport's paper this was regarded as an important unsolved problem of number theory. Any result of the form

$$N[f(x, y) \equiv 0 \bmod p\,] = p + \mathcal{O}(p^{\gamma}) \qquad \text{for} \qquad p \to \infty \tag{10}$$

for some $\gamma < 1$ was considered as a step towards the expected (9).

Davenport's result (2) can be regarded a special case of this. For, assume that $f(x, y)$ is of the special form $y^2 - f(x)$ with a polynomial $f(x)$ of one variable and integral coefficients. [29] Then (8) can be written in the form

$$y^2 \equiv f(x) \bmod p\,. \tag{11}$$

[29] We take the liberty of using the same symbol f once for denoting a polynomial $f(x, y)$ of two variables, and then for a polynomial $f(x)$ of one variable. We hope this will not create confusion, and that it will always be clear from the context what is meant in the particular situation. By the way, this notation is also used in the original papers of Mordell and Davenport.

Without restriction we assume that $f(x)$ has no multiple roots. From (5) it is almost immediate that

$$N[\,y^2 \equiv f(x) \bmod p\,] = \sum_{x \bmod p} \left(1 + \left(\frac{f(x)}{p}\right)\right) = p + S(f) \qquad (12)$$

where $S(f)$ is the sum as in (6) for the polynomial $f(x)$ appearing in (11). We see that Davenport's result (2) yields (10) in the special case (11) with $f(x)$ of degree 4. In this case, according to Davenport one can take $\gamma = \frac{3}{4}$.

The problem (9) for diophantine congruences is closely related to the Riemann hypothesis. We have explained this in section 6.4 of Part 1, where we have discussed Hasse's survey article [H:1934]. There it is shown that the Riemann hypothesis for the function field $F = \mathbb{F}_p(x, y)$ defined by (8) would imply (9). Also, any result that the zeros s of the zeta function of F have real part $\mathfrak{R}(s) \leq \gamma < 1$ for some γ would imply (10).

Seen in this way, Davenport's result was closely linked with the Riemann hypothesis for quadratic function fields in the sense of Artin's thesis. Of course, Hasse's survey article [H:1934] had not yet appeared when Davenport wrote his article. But Davenport was concerned with the special case $y^2 \equiv f(x) \bmod p$ which means that the function field $F = \mathbb{F}_p(x, y)$ is quadratic. These quadratic fields had been extensively discussed in Artin's thesis.

Nevertheless, Artin was not mentioned in Davenport's paper [Da:1930] . It seems that in 1930 when Davenport wrote the paper, he did not yet know Artin's thesis. He learned about it from Hasse during his visit 1931 in Marburg, as reported by Halberstam in [Da:1977].

3.3 Reducing exponents

Artin's thesis *was* mentioned, however, in Mordell's subsequent paper [Mor:1933] where quite a number of new results about diophantine congruences were presented. Mordell had become interested in the subject after reading Davenport's paper, and he seemed to be intrigued by Davenport's method. By refining that method Mordell was able to obtain a better exponent in certain cases, and also to cover quite a number of new cases, i.e., other diophantine congruences in two variables.

Mordell's paper appeared in 1933 but he had obtained the results already in 1931. For, on Nov. 8, 1931 he wrote to Hasse:

> "*. . . during the last three weeks I became very interested in Davenport's note on the distribution of quadratic residues and I could not do anything else. I have only within the last few days proved that the number of solutions of* $y^2 \equiv ax^3 + bx^2 + cx + d \bmod p$ *is*

> $p + \mathcal{O}(p^{\frac{3}{4}})$ *& more generally when* y^2 *is replaced by* y^m *except in one trivial case. Davenport has also found the theorem & proof of a different case about the same time. If I remember any German, I might speak on this to your students etc. as the method is very elementary.*"

In the last sentence Mordell refers to his future visit in Marburg which was planned in early 1932. One month later, on Dec 14, 1931 Mordell wrote again:

> "*. . . You may also be interested in knowing that I have made further progress with congruences. The cubic congruence* $f(x,y) \equiv 0$ *has in general* $p + \mathcal{O}(p^{\frac{2}{3}})$ *solutions. Also* $y^m \equiv a_1 x^n + \cdots + a_{n+1}$ *has in general* $p + \mathcal{O}(p^{\gamma(m,n)})$ *solutions where*
>
> $$\begin{aligned} \gamma(m,n) &= \frac{2}{3} \qquad \text{if} \quad n=4,\ m=2 \\ &= \frac{7}{8} \qquad \text{if} \quad n=6,\ m=2 \\ &= \frac{5}{6} \qquad \text{if} \quad n=4,\ m=4 \\ &= \frac{3}{4} \qquad \text{if} \quad n=3,\ (m=2 \ \text{ included above})^{\,30} \\ &= \frac{1}{2} \qquad \text{if} \quad n=3,\ m=3\,. \end{aligned}$$
>
> *Davenport has also found results of this kind; and I saw him three days ago. . .*"

As to Davenport's results, he had obtained in addition

$$\gamma(4,4) = \frac{2}{3}\,, \gamma(m,4) = \frac{5}{6}\,, \gamma(5,5) = \frac{5}{6}\,, \gamma(3,6) = \gamma(6,6) = \frac{7}{8}\,, \gamma(2,7) = \frac{19}{20}$$

in his papers [Da:1932] and [Da:1933a].

In the printed version of Mordell's paper we find also the following result which he did not mention in his letter to Hasse:

$$N[ax^m + by^n + c \equiv 0 \bmod p\,] = p + \mathcal{O}(p^{\frac{1}{2}})\,. \tag{13}$$

This congruence is today called "Davenport-Hasse" congruence. We shall return to this in Part 3.

[30] Indeed, if $m = 2$ and $n = 3$ then the above result for "cubic congruence in general" applies. Of course, the same applies if $m = n = 3$ but in this case the next line gives a better result.

We see that now the investigations went beyond congruences of the form (11) which define *quadratic* function fields over $\mathbb{F}_p$ in the sense of Artin. Those with higher $m > 2$ lead to *cyclic* extensions of degree $\leq m$ of $\mathbb{F}_p(x)$ – at least if $p \equiv 1 \bmod m$ which is a reasonable assumption in this context and will be imposed tacitly in the sequel in this context. [31]

If $m > 2$ then the sums like (2) have to be replaced by so-called *character sums* which are defined as follows. Let χ denote a character of order m of the multiplicative group $\mathbb{F}_p^\times$. The values of χ are the m-th roots of unity in the complex number field. Put $\chi(0) = 0$. For any polynomial $f(x)$ with coefficients in $\mathbb{F}_p$ we put

$$S_\chi(f) := \sum_{x \bmod p} \chi(f(x))\,. \tag{14}$$

In the case $m = 2$ there is only one character of order 2, namely the quadratic residue symbol $\chi(z) = \left(\dfrac{z}{p}\right)$ and then the character sum (14) coincides with (2).

An element $t \in \mathbb{F}_p^\times$ is an m-th power residue if and only if $\chi(t) = 1$, and we have

$$1+\chi(t)+\chi^2(t)+\cdots+\chi^{m-1}(t) = \begin{cases} m & \text{if } t \text{ is an } m\text{-th power residue mod } p \\ 0 & \text{if } t \text{ is not an } m\text{-th power residue mod } p \\ 1 & \text{if } t \equiv 0 \bmod p\,, \end{cases} \tag{15}$$

in generalization of statement (5). The right hand side gives the number of solutions $y \in \mathbb{F}_p$ of the equation $y^m = t$. From this it is immediate that

$$\begin{aligned} N[\,y^m \equiv f(x) \bmod p\,] &= \sum_{x \bmod p} \bigl(1 + \chi(f(x)) + \chi^2(f(x)) + \cdots + \chi^{m-1}(f(x))\bigr) \\ &= p + S_\chi(f) + S_{\chi^2}(f) + \cdots + S_{\chi^{m-1}}(f) \end{aligned}$$

We see that any result of the form

$$S_{\chi^\mu}(f) = \mathcal{O}(p^\gamma)$$

for the non-trivial powers of χ implies

$$N[y^m \equiv f(x) \bmod p\,] = p + \mathcal{O}(p^\gamma)\,.$$

In this way the above relations were proved by Mordell and Davenport.

But Mordell also says very clearly in his paper that the same problem arises for the number of solutions of any congruence of the form (8). He considered

[31] For otherwise, m could be replaced by the greatest common divisor of m and $p-1$ without changing the number N of solutions.

the "general cubic congruence" (for which he had obtained the exponent $\frac{2}{3}$) as a starter only, as well as (13).

In the introduction of his paper, Mordell gives some historical account of the problem of counting solutions of diophantine congruences of the form (8). The oldest reference is a paper of 1832 by Libri [Lib:1832] in which the cubic congruence

$$x^3 + y^3 + 1 \equiv 0 \bmod p\,, \tag{16}$$

is investigated and where explicit expressions for the number N of solutions are found; from this one can see that $N = p + \mathcal{O}(p^{\frac{1}{2}})$. This is in fact the best possible exponent, and Mordell was able to generalize it for the case $n = m = 3$ in the last line of his list. [32]

As said above already, Mordell also mentions Artin in his paper. After discussing Libri's result with the best possible exponent $\frac{1}{2}$, he remarks:

> "*Perhaps some of the other exponents are too great. Prof. Artin informs me that the case $m = 2$, $n = 3$ arises in some of his investigations and that he thinks the best possible result then is*
>
> $$\gamma(2,3) = \frac{1}{2}\,.$$
>
> *In fact, he gives the empirical result for $p > 3$,*
>
> $$-2\sqrt{p} < S < 2\sqrt{p}$$
>
> *where*
>
> $$S = \sum_{x \bmod p} \left(\frac{x^3 + ax + b}{p}\right)\,.\ \text{''}$$

After that Mordell reports about Artin's line of arguments with his ζ-function.

I am somewhat puzzled by Mordell's remark. First of all, by "empirical result" he obviously means Artin's tables of class numbers in his thesis [A:1924] (see Part 1, section 2.1.2). There, Artin presented the result of his computations for a number of explicitly given quadratic function fields, each of them verifiying the validity of the Riemann hypothesis. The majority of Artin's examples have characteristic $p = 3$ (only some have $p = 5$ or 7). So why did Mordell write $p > 3$? Was this a misprint only and he meant $p > 2$? More puzzling, why should Artin tell Mordell that he expects $\gamma(2,3)$ only, i.e., for elliptic function fields? We know from Artin's letters to Herglotz that Artin expected the Riemann hypothesis to be true in all quadratic cases. We would therefore

[32] It has been pointed out to me by Franz Lemmermeyer that the congruence (16) had been discussed in Gauss' *Disquisitiones Arithmeticae* already, in connection with the problem of counting pairs of consecutive cubes modulo p. See chapter 7, section 358 of D.A.

expect Artin to tell Mordell that he thinks $\gamma(2,n)=\frac{1}{2}$ to be true for all n and not only for $n=3$.

We know that Mordell had shown Hasse the raw manuscript of his paper (this was in January 1932). Certainly, Hasse knew about Artin's thesis. Why did not Hasse tell him ?

Perhaps Mordell's remark about Artin's work was not yet included in the raw manuscript which Hasse had seen, and it was added later. Mordell (as well as Hasse) attended Artin's famous Göttingen lectures on class field theory which took place from February 29 to March 2, 1932. [33] But Mordell had submitted his manuscript [Mor:1933] shortly before already; it was received by the editors of *Mathematische Zeitschrift* on Feb 25, 1932. We can imagine that Mordell talked about his already submitted paper to Artin in Göttingen, and that the latter informed him on his (Artin's) thesis, in particular on his numerical computations with some elliptic function fields, verifying the Riemann hypothesis in a number of cases. Upon this Mordell may have hastily written a short remark on Artin's work which he then added to his manuscript.

In any case, Mordell and Davenport continued their work; they used their methods to treat also the so-called *exponential sums*, which are expressions of the form

$$S_e(f) := \sum_{x \bmod p} e(f(x)) \qquad \text{with} \quad e(z) = e^{\frac{2\pi i\, z}{p}} \tag{17}$$

where $f(x)$ is a polynomial mod p. Mordell [Mor:1932] showed that for such sums

$$S_e(f) = \mathcal{O}(p^{1-\frac{1}{n}}) \tag{18}$$

where $n=\deg f(x)$. If $n=3$ then $1-\frac{1}{n}=\frac{2}{3}$. Davenport later [Da:1933] reduced this exponent to $\frac{5}{8}$. Moreover, for $n\geq 4$ Davenport obtained the exponent $1-\frac{1}{m}$ where m is the largest integer of either of the forms 2^r , $3\cdot 2^r$ not exceeding n.

Exponential sums like (17) can also be formed for a rational function instead of a polynomial, where in the summation the poles of the rational function are to be omitted. For rational functions of the form

$$f(x) = ax^n + bx^{-n} \qquad \text{with} \quad n \neq 0$$

Davenport showed that $S_e(f)=\mathcal{O}(p^{\frac{2}{3}})$, which was an improvement of some result of Mordell. For $n=1$ these are the so-called Kloosterman sums. [34]

The estimation of such exponential sums is closely related to the Riemann hypothesis for function fields of the form $F=K(x,y)$ with $y^p-y=f(x)$ over

[33] Notes of these lectures were taken by Olga Taussky. An English translation of her notes appeared in H. Cohn's book [Co:1978].

[34] For such Kloosterman sums, the same exponent $\frac{2}{3}$ had been obtained, at the same time but independent of Davenport, by H. Salié [Sal:1931].

$\mathbb{F}_p$. But in 1932 this was not yet known to Mordell or Davenport; we shall return to this question later.

So we see that in 1931/32, shortly after Davenport had stayed with Hasse in Marburg, there started quite an activity in the direction of estimating character sums and exponential sums. At first, Hasse himself was not actively participating [35] but he appears to have shown keen interest in this work. He was always kept informed about the newest results, by mail and also personally because he met Davenport as well as Mordell several times in 1932. [36]

Nevertheless, although Hasse held these results and their authors in high regard he was not much impressed by the *methods* used, and not by the *general attitude* of the authors towards these problems. Hasse tended to think about diophantine congruences not as problems *per se*, but as manifestations of *mathematical structures*. Thus, he said to Davenport, the results so far are obtained only through clever computations, manipulating and estimating algebraic and analytic expressions. Hasse acknowledged that the methods used may be nontrivial [37] but they did not seem to him adequate since they lead to many different exponents in the remainder term in so many special cases, whereas in every case the exponent $\frac{1}{2}$ was expected. Perhaps it would be possible to reduce some of the exponents a little further by refining those methods. But instead of "reducing exponents" [38] the proper thing to do would be to find out the *structure* behind this, and to see how and why the optimal exponent $\frac{1}{2}$ is connected with that structure.

Davenport may have replied that he does not believe that those abstract methods can do much better than the very explicit methods which he and Mordell had developed and used. After all, what only counts are the results and not the methods. And when Hasse still insisted on his view, Davenport challenged him to solve the problem with his abstract structural methods. Finally Hasse accepted this challenge and started to work on the problem.

The above description of a possible dispute between Hasse and Davenport is not purely fictional. Hasse used to tell us this story along these lines when asked about his first steps towards the proof of the Riemann hypothesis. At

[35] This may have been because he was busy with quite a number of other mathematical research activities. In the years 1931-1933 there appeared 25 papers of Hasse, the highlights among them belonging to complex multiplication, local and global theory of algebras, class field theory and norm residues, including his famous Marburg lecture notes. The late twenties and early thirties can be considered as the most productive years in Hasse's life.

[36] Besides of several meetings of Davenport and of Mordell with Hasse in Germany and Switzerland, the latter visited them in England in October 1932.

[37] Mordell in [Mor:1933] even said that Davenport's method was "*very ingenious*".

[38] The terminology "reducing exponents" seems to have been established between Hasse and Davenport, somewhat ironically on the side of Hasse (you should do better than reducing exponents) and in a sense provocative on the side of Davenport (I can at least reduce exponents, and what can you do with your abstract methods?). In one letter of Feb. 25, 1932 Davenport writes "*I haven't reduced any exponents recently, I regret to say.*"

the other side, Davenport told the story in the same spirit to Mordell to whom he was quite close. Mordell reports in [Mor:1971a] :

> "*Davenport was staying with Hasse at Marburg in the earlier thirties and challenged him to find a concrete illustration of abstract algebra. This led Hasse to his theory of elliptic function fields...*"

3.4 The Hamburg colloquium and Artin's comment

Probably in the fall of 1932 Hasse started to seriously think about possible strategies to find the structure behind the problem of diophantine congruences.

Consider a diophantine congruence of the form (8). This can be regarded as an equation $f(x, y) = 0$ over the finite field $\mathbb{F}_p$ and as such it defines a function field F over $\mathbb{F}_p$ if p is sufficiently large. [39] As we have already said, Hasse knew that the Riemann hypothesis for this function field would imply (9), at least in the case of quadratic congruences (11) which had been treated by Artin in his thesis [A:1924]. Nevertheless it seems that at first Hasse hesitated to attack the Riemann hypothesis directly. After all, in the number field case all the attempts to prove the classic Riemann hypothesis had not succeeded (this is so even today, at the time of writing this article). Artin in his thesis had said that the general proof of the Riemann hypothesis for function fields will probably "*have to deal with difficulties of similar type as with Riemann's zeta function.*" This did not sound very encouraging. [40]

In the number field case there are several theorems which were proved without use of the classic Riemann hypothesis, although they are known to be consequences of it. And so Hasse at first tried to attack the problem of diophantine congruences without the Riemann hypothesis for function fields. He studied the proofs in Davenport's and Mordell's papers in detail and, as a first step, tried to simplify them in his sense, i.e., to do them more systematically, hoping that finally he would be able to cover more situations than just those special examples of Davenport and Mordell.

But soon, as we shall see, he changed his viewpoint as a result of a discussion with Artin on the occasion of a colloquium talk in Hamburg.

During all his life, whenever Hasse gave a talk at a colloquium or at another occasion, he carefully prepared a manuscript for it. Since many of those manuscripts are preserved we know today what he was talking about and how.

[39] Note that if $f(x, y) \in \mathbb{Z}[x, y]$ is absolutely irreducible then for all sufficiently large primes p the reduced polynomial mod p is absolutely irreducible too. This is an old theorem. I do not know who had first formulated and proved it. F. K. Schmidt in a letter to Hasse writes that he had learned this theorem from a paper by Ostrowski [Os:1919].

[40] Although Artin had added that in the function field case "*the situation is clearer and more lucid because it essentially concerns polynomials*".

There were several talks in 1932/33 where Hasse reported about his work on the Davenport-Mordell results, and so from his own lecture notes we can follow the progress of his thoughts. These talks were:

November 1932	Kiel
some days later 1932	Hamburg
January 1933	Göttingen
February 1933	Marburg
May 1933	Marburg again
September 1933	Würzburg
January 1934	Hamburg again

Kiel was the place where Hasse had started his mathematical studies in 1917, and also where he had taught as a *Privatdozent* during 1922-1925. In 1928 there had been strong attempts by Adolf Fraenkel who at that time still was in Kiel, to draw Hasse back again to Kiel with an attractive offer. Although this was not successful, Hasse kept always good relations not only to Fraenkel but also to his other colleagues and his former teachers in Kiel. [41] Therefore it is not surprising that it was Kiel where Hasse presented the results of his newly acquired interest for the first time. According to Hasse's lecture notes he covered in this colloquium the most striking recent results of Mordell and Davenport, those which we have discussed in section 3.3 above. Hasse tried to present the proofs in a more systematic manner. Mordell's paper had not yet appeared at that time. [42]

After his November colloquium 1932 in Kiel, Hasse planned to visit Hamburg with "*the only purpose to be together with the Artins*" as he wrote to Davenport. It seems that he wished to consult Artin about several mathematical questions which had come up in their correspondence. But somehow Artin had talked Hasse into delivering his Kiel lecture a second time in Hamburg. This lecture had the same title as that in Kiel, namely:

[41] During his visit to Kiel, Hasse stayed with the Fraenkels.

[42] In Hasse's legacy we have found two manuscripts carrying the note: "Kiel, November 1932". Besides of the one which we have reported about here, there is another one about the structure theorems on algebras over number fields, containing a brief survey on the Local-Global-Principle and cyclicity of algebras. We do not know whether Hasse had given two talks in Kiel. But this seems to us rather unlikely since he did not mention a word on it in his letter to Davenport. Rather, it seems that the talk about algebras had been originally proposed by Hasse when he received the invitation to Kiel. This was the same talk which he had presented in September 1932 in Zürich at the International Congress. We know from the correspondence Hasse-Fraenkel that Hasse's colloquium lecture in Kiel had been discussed in Zürich among the two, and at that time it seemed natural that Hasse would propose the same subject as he had talked in Zürich. But in the meantime, in October 1932 Hasse had visited England and met Davenport and Mordell; there he became interested in the problem of congruences modulo p and, hence, changed the subject of his talk on short notice.

> "*On the asymptotic behavior of numbers of solutions of congruences modulo p .*"

There is no hint in the title, nor in Hasse's own lecture notes, of a connection of this topic to function fields and the Riemann hypothesis.

But in Hamburg, Hasse's lecture contained more than the Kiel lecture several days earlier. For, in a letter of Dec 7, 1932 Hasse reported to Davenport: [43]

> "*My lectures found much interest with the Hamburg and Kiel mathematicians. In Hamburg, I was able to produce a couple of new results, which I had found during my journey back from Kiel in a* Personenzug. *I have proved that*
>
> $$S_e(f) = \sum_{x\in\mathbb{F}_q} e(f(x)) \; = \mathcal{O}(q^{1-\frac{1}{n}}) \tag{19}$$
>
> *where n is the degree of f(x) (a polynomial with coefficients also in* $\mathbb{F}_q$*) and*
>
> $$e(z) = e^{\frac{2\pi i}{p}\,\mathrm{tr}(z)} \qquad (z \in \mathbb{F}_q) \tag{20}$$
>
> *where* tr *denotes the trace:*
>
> $$\mathrm{tr}(z) = z + z^p + \cdots + z^{p^{r-1}} \qquad (q = p^r). \tag{21}$$
>
> *I have also applied my (i. e. Mordell's) method to the character sums and found that in the elliptic case*
>
> $$N[\,y^2 = f(x)] = \; q + \mathcal{O}(q^{1-\frac{1}{6}}) \qquad (x,\, y \in \mathbb{F}_q). \tag{22}$$
>
> *My method has been very rough, and I am pretty sure I can improve this to* $\mathcal{O}(q^{1-\frac{1}{4}})$ *corresponding to your best–known result. . .* [44]

It seems that Hasse's original motivation for this generalization was to obtain the Davenport-Mordell results for diophantine congruences in an arbitrary algebraic number field of finite degree. To do this, one has to consider polynomials with integer coefficients in that number field, and to count solutions of

[43] Hasse's letters to Davenport are generally written in English, hence translation has not been necessary. It seems that as a result of Davenport's instructions, Hasse's English was not that "quaint" any more than it used to be in his first letter to Davenport.

[44] For the convenience of the reader we have changed Hasse's notation somewhat. So we have written $\mathbb{F}_q$ where Hasse wrote E_q (with "E" for "*endlich*"). Secondly, we have incorporated the trace into the definition of $e(z)$ in (20) which Hasse does not do. (In his letter he writes $e(\mathrm{tr}\, z)$ instead. But later he switched to the notation which we have given here.) Finally, Hasse defines the exponential sums as a mean, involving the denominator q which we do not write; accordingly his estimate is $\mathcal{O}(q^{-\frac{1}{n}})$ instead of $\mathcal{O}(q^{1-\frac{1}{n}})$. Hasse calls the mean $\frac{1}{q}S_e(f)$ the "distribution function" (*Verteilungsfunktion*).

congruences like $y^m \equiv f(x) \bmod \mathfrak{p}$ for a prime ideal $\mathfrak{p}$. But instead of congruences modulo $\mathfrak{p}$ one may write equations in the finite residue field. This then led Hasse to consider the generalization of Davenport's and Mordell's results to arbitrary finite base field $K = \mathbb{F}_q$. In fact, Hasse explains on a postcard (dated Feb. 7, 1933) to Davenport: "*There is no need of representing the abstract Galois field as a residue class field for a prime ideal $\mathfrak{p}$ in a (suitably chosen) algebraic number field.*"

In this general setting the error term is again given by character sums of the form (14), this time referring to an arbitrary non-trivial character χ of the multiplicative group of the finite field K, and the summation extended over $x \in K$ (instead of $\mathbb{F}_p$). If m is the order of χ then necessarily $q \equiv 1 \bmod m$ and the function $z \mapsto \chi(z)$ is a homomorphism of $K^\times$ onto the group of m-th roots of unity in the complex number field. Putting $\chi(0) = 0$ the summation in (14) extends over $x \in K$. For the problem (22) one has to take for χ the uniquely determined quadratic character of $K^\times$.

As to exponential sums, the exponential function $e(z) = e^{\frac{2\pi i z}{p}}$ in (17) is to be re-defined including the trace as in (20). Then $z \mapsto e(z)$ is a homomorphism of the additive group of the finite field K onto the group of p-th complex roots of unity.

It turned out that this generalization, namely from $\mathbb{F}_p$ to an arbitrary finite field K, was a straightforward routine – we note that Hasse could do it while travelling in a "*Personenzug*" from Kiel to Hamburg. [45] Then why did Hasse report it to Davenport with such apparent emphasis? I believe that he wished to inform Davenport about Artin's comment on this. For, Hasse's letter continues as follows, referring to statement (22) in the elliptic case:

> "*As Artin pointed out, this means that the zeros of his congruence ζ–function lie all in $\Re(s) \leq 1 - \frac{1}{6}$, or $1 - \frac{1}{4}$ respectively.*"

In other words, Artin had said: "If you can prove the validity of the estimate (22) not only over $\mathbb{F}_p$ but also over any finite extension $\mathbb{F}_q$ of $\mathbb{F}_p$ (with p replaced by q and with the constant implied by the $\mathcal{O}$-symbol being independent of q) then this has consequences for the nontrivial zeros of my ζ-function of the function field defined by $y^2 = f(x)$ over $\mathbb{F}_p$. Namely, those zeros will have real part $\Re(s) \leq 1 - \frac{1}{6}$. Similar argument if in (22) the exponent $1 - \frac{1}{6}$ is replaced by $1 - \frac{1}{4}$."

We may add that the same holds if $1 - \frac{1}{6}$ is replaced by arbitrary $\gamma < 1$. If $\gamma = \frac{1}{2}$ then the Riemann hypothesis will follow.

[45] In the hierarchy of trains within the German railway system at that time, the "Personenzug" ranged as a slow train, stopping at each station. Today the travelling time from Kiel to Hamburg in a slow train is about 75 minutes. In 1932 it may have been roughly the same.

It appears that Hasse had not known this fact before, since otherwise he would not have reported it to Davenport in the form he did. We have said in Part 1 (section 2.3.1) already that Artin had never published this although he had mentioned it in a letter to Herglotz written in November 1921 already. I do not know whether Artin had communicated it to other people in the meantime. In any case he told it to Hasse after the colloquium talk in November 1932.

This comment by Artin seems to have decidedly changed the viewpoint of Hasse. He now realized that the Davenport-Mordell results could be regarded as giving information on the zeros of the respective zeta functions – provided these results could be obtained for an arbitrary finite base field K instead of $\mathbb{F}_p$ only. And he realized that not only the Riemann hypothesis would imply the Davenport-Mordell estimates with the best possible exponent $\frac{1}{2}$ but also the converse: the latter estimates are indeed *equivalent* to the Riemann hypothesis, again under the provision that they could be obtained over any finite base field K instead of $\mathbb{F}_p$. In this spirit Hasse wrote his last sentence of his above mentioned letter to Davenport:

> "*I hope to be able to extend all your and Mordell's results to Galois fields.*"

It may well be that Artin's comment was referring to quadratic function fields only, which he had discussed in his thesis and in his letters to Herglotz. In any case we shall see that under Hasse's hand, with the help of F. K. Schmidt, the same reasoning became valid for arbitrary function fields with finite base field.

Thus Hasse, after his discussion with Artin, had now found the algebraic structure behind the various estimates of Davenport and Mordell, namely *function fields over finite base fields and the zeros of their zeta functions.* These were the structures which were first investigated by Artin in his thesis, followed in the 1920s by F. K. Schmidt and others (see Part 1). Hasse was well informed about the theory of function fields, and now he realized that it is closely related to the estimation problems of Davenport and Mordell.

There arose the task to study these structures with the aim to find out more about the region which contains the zeros of the zeta functions. Hasse wished to obtain estimates about the real parts of the zeros s, in the form $\mathfrak{R}(s) \leq \gamma$ with γ as small as possible. ($\gamma = \frac{1}{2}$ would be the Riemann hypothesis.) It appears that at the time of his Hamburg colloquium lecture, Hasse did not yet believe in the general validity of the Riemann hypothesis. We conclude this from an account of S. Iyanaga who was present at Hasse's colloquium talk in Hamburg and narrates the following in a letter to me dated July 29, 1998: [46]

[46] I am indebted to Prof. Iyanaga for letting me share his recollections. The cited text is a translation from his letter. – REMARK: In Part 1 [Rq:2002a] I have said that S. Iyanaga had been present at Hasse's Hamburg lecture in 1934 (see section 3.6 of Part 1). In the meantime,

> *"...I was reminded that I had been there when Hasse gave a talk in Hamburg 1932. At that occasion Artin voiced his strong opinion that the same results would hold in arbitrary function fields, to which Hasse replied: "I am not so optimistic, I have studied the problem." Now we see that Artin had been right, but I still remember with admiration the great progress which Hasse had achieved with his work..."*

It is a common observation in the history of mathematics that a problem can be more easily solved if it can be put into a structural framework which seems "adequate" to it – at least in the eyes of those people who work on that problem. In any case, after his Hamburg lecture Hasse realized that the framework of function fields and their zeta functions was adequate to the Davenport-Mordell problem. And three months later, at the end of February 1933, Hasse succeeded to prove the Riemann hypothesis for elliptic function fields. (See section 5.1 below). Another 11 months later, when Hasse again adressed the Hamburg colloquium, he was already convinced that the Riemann hypothesis holds for function fields of arbitrary genus. (See Part 3.)

3.5 Summary

In 1931 Hasse met the young student Davenport who had been recommended to him by Mordell. There developed a friendship between the two, and Hasse became interested in Davenport's work which was concerned with estimating the number of solutions of diophantine congruences. Some months later Mordell extended and generalized Davenport's results. Although Hasse appreciated the high value and the ingenuity of Mordell's and Davenport's methods he voiced his opinion that the structural methods of "modern algebra" would lead to better estimates. Davenport challenged him to solve the problem with abstract structural methods. Hasse accepted this challenge and started to work on the problem, looking for the adequate algebraic structure connected with it.

In November 1932 Hasse delivered a colloquium lecture in Hamburg and met Artin there. From Artin's comments Hasse learned that the Riemann hypothesis is indeed equivalent *to the Davenport-Mordell problem for diophantine congruences, provided the latter is extended to arbitrary finite base fields instead of the prime field $\mathbb{F}_p$ only. And so he saw the structure behind the Davenport-Mordell problem, namely the theory of algebraic function fields and their zeta functions according to F. K. Schmidt.*

however, after I have learned that Hasse had talked in Hamburg twice on elliptic function fields – once in 1932 and another time in 1934 – Professor Iyanaga has confirmed (in a letter dated April 2, 2003) that it was 1932 and not 1934 when he attended Hasse's lecture. Hence, in this respect I have to revise my corresponding statement in Part 1 .

4 Hasse's new point of view

4.1 The Göttingen colloquium

On January 10, 1933, not much more than one month after his Hamburg talk, Hasse delivered another talk on the same topic, this time in Göttingen at the *Mathematische Gesellschaft* (Mathematical association).

A visit to the *Mathematische Gesellschaft* had been proposed by Emmy Noether about a year ago already; originally she had wished Hasse to report on their joint paper about the structure of algebras [BHN:1932]. But somehow the plans were changed; in the summer of 1932 Hasse had talked in Noether's seminar and not in the *Mathematische Gesellschaft.* Nevertheless Hasse had been "put on the list" which meant somewhat like a standing invitation to the *Gesellschaft.* Emmy Noether then proposed that he should come in the winter semester 1932/1933; she wrote to Hasse that

> "*Courant will be here in the winter, and some of the newest number theory will be good for him.*"

On November 11, 1932 she wrote again to Hasse and asked for a date for his visit. [47]

It seems that Hasse in his reply mentioned what he had learned from Artin in Hamburg, and that he now was interested in function fields and their zeta functions. For in Noether's next letter, dated Dec 9, 1932, she asked for the precise title of Hasse's talk and wrote:

> "*You may mention zeta functions as a lure*",

meaning that such title would attract more mathematicians, also those (including Courant) whose field of interest was somewhat distant from number theory. Hasse chose the title:

> *On the zeros of Artin's congruence zeta functions.*

Through this title Hasse advertised his new view point. In Kiel and Hamburg the title had been "On asymptotic behavior of numbers of solutions of congruences". Now he wished to point out that those asymptotic estimates are

[47] She invited him to stay at her home: "*But this time you should really stay with me. My guest room has been initiated already by Alexandroff who stayed for 4 weeks.*" – By the way, one day after his talk at the *Gesellschaft* Hasse gave a second talk, upon Noether's request in her seminar where she had prepared the participants, as she wrote, for the Local-Global Principle. Hasse talked about his Local-Global Principle for quadratic forms.

of interest mainly because they give information about the position of zeros of the zeta function. It is evident that this change of paradigm had come about as a consequence of Artin's comment. Looking through Hasse's notes we find that he did not present any essential new results when compared to Hamburg. The difference was just his viewpoint. [48]

According to the title, Hasse discussed mainly *quadratic* function fields and therefore *Artin's* zeta functions only. But we can see from his notes that Hasse briefly mentioned also F. K. Schmidt's zeta functions for arbitrary function fields.

Let Hasse himself explain his new vision in his own words, to be found in his *Zentralblatt* review of Mordell's paper [Mor:1933], the same paper which we have discussed above in section 3.3. The following text is an excerpt from this review. [49]

> "*The paper is concerned with special cases of the following general problem: Let $f(x,y)$ be a polynomial with integer coefficients which is absolutely irreducible over the finite field $\mathbb{F}_p$ of p elements, and N the number of solutions of $f(x,y)=0$ in $\mathbb{F}_p$. One should find an estimate of the form*
>
> $$\text{(A):} \qquad |N - p\,| \leq Cp^{\gamma}$$
>
> *where the exponent $\gamma < 1$ is as small as possible, and C is a positive constant. Both γ and C should not depend on p, and also not on the special choice of the coefficients of f, but only on the algebraic invariants of the function field F defined by the equation $f(x,y)=0$ over $\mathbb{F}_p$.*
>
> *I would like to remark in advance that the final solution of this general problem is closely related to the analogue of the Riemann hypothesis for F. K. Schmidt's zeta function $\zeta_F(s)$ for F. If the infinite solutions are correctly included into the count then N becomes the number of prime divisors of degree 1 of F and the theory of $\zeta_F(s)$ shows that γ can be chosen as the maximal real part θ of the zeros of $\zeta_F(s)$. In addition, one can choose $C = 2g$ where g is the genus of F. It is known that $\theta < 1$ but a bound which is independent of p is not yet known. The analogue of the Riemann hypothesis, $\theta = \frac{1}{2}$, would imply that one could choose $\gamma = \frac{1}{2}$.*

[48] Remark: In Part 1, section 3 I have said that Hasse in this Göttingen lecture talked about his proof of the Riemann hypothesis in the elliptic case. Now I have to correct that statement in view of Hasse's own lecture notes. From his correspondence with Davenport it is clear that Hasse had obtained the proof for the elliptic case at the end of February 1933 only. See section 5.1 below.

[49] The notation is ours, not always coinciding with that of Hasse in the review.

> *Conversely, the statement* (A) *for F and for all constant field extensions of F (where on both sides p is to be replaced by the order $q = p^r$ of the field of constants, and γ, C are independent of r too) would imply that $\theta \leq \gamma$, hence for $\gamma = \frac{1}{2}$ the analogue of the Riemann hypothesis for F would follow. – The author explains this connection in the hyperelliptic cases $f(x,y) = y^2 - f(x)$ only, with the special congruence zeta functions of Artin.*"

Only after this introduction Hasse proceeds to review the results of Mordell's paper in more detail; these are essentially the same as Mordell had stated in his letter to Hasse of Dec 14, 1931, where he obtained the values $\gamma = \frac{2}{3}, \frac{7}{8}, \frac{5}{6}, \frac{3}{4}, \frac{1}{2}$ in various situations; see section 3.3.

The above review text shows clearly Hasse's change of viewpoint. Whereas Mordell regarded the theory of the zeta function as a means to obtain good estimates of the form (A), Hasse now proposed to look in the other direction. Namely, any estimate of the form (A) would lead to a result about the real parts of the zeros of the zeta function, *provided that the estimate* (A) *can be proved over all finite field extensions of $\mathbb{F}_p$ too.* In this spirit Hasse added a last paragraph to his review, confirming what he had already announced in his letter to Davenport:

> "*I would like to add that the results of the author can be transferred almost word for word to the case of an arbitrary finite field K instead of $\mathbb{F}_p$ as field of coefficients. Hence, as said at the beginning, they lead to a bound of the maximal real part of zeros θ by the respective γ and, if $\gamma = \frac{1}{2}$, to the Riemann hypothesis for the respective zeta function of F. K. Schmidt.*"

In this review text we see those ideas emerge which later appeared in detail in Hasse's survey paper [H:1934].

Some of Hasse's statements in the above review need explanation. Let us discuss this in the context of an arbitrary finite field K instead of $\mathbb{F}_p$; let q be the order of K. Every absolutely irreducible polynomial $f(x,y)$ over K determines an affine curve Γ; let $F = K(x,y)$ be its function field over K, where (x,y) is a generic point of Γ over K. The solutions in K of the equation $f = 0$ are precisely the K-rational points of Γ. Let N_Γ denote their number.

The first thing to do, Hasse says in his review, is to "*include the infinite solutions correctly into the count*". What does Hasse mean by this?

If the curve Γ is smooth then the local ring of each point of Γ is a valuation ring of F. The point is K-rational if and only if the residue field of that valuation ring coincides with K. In this way the K-rational points of Γ correspond bijectively to certain prime divisors of F of degree 1. But these are not all prime divisors of degree 1; there are also the poles of either x or y. These poles, according to Hasse, are to be viewed as the "infinite solutions". [50] Of course, not every pole of x or y is necessarily of degree 1. Anyhow, there are at most $n_x := [F : K(x)]$ poles of x and $n_y := [F : K(y)]$ poles of y of degree 1. Therefore, if N denotes the number of all prime divisors of degree 1 of F, we have $|N - N_\Gamma| \leq n_x + n_y$. Now, these numbers n_x and n_y are not altered if F is replaced by any base field extension $F^{(m)} = FK^{(m)}$, with $|K^{(m)}| = q^m$. That is, we have $n_x = [F^{(m)} : K^{(m)}(x)]$ and similarly for n_y. Thus, if we put $c = n_x + n_y$ then

$$|N_\Gamma - N| \leq c \tag{23}$$

where the bound c remains unaltered under base field extension.

When Hasse said that "the infinite solutions should be correctly included into the count" then he means that N_Γ should be replaced by N; the inequality (23) shows that this is permitted if one wishes to obtain an estimate like (A).

However, if the curve Γ is not smooth then the above bound is not correct. We have reasons to believe that Hasse, in his review of Mordell's paper, had in mind Artin's case of quadratic function fields only, given by an equation of the form $y^2 = f(x)$ where $f(x)$ has no multiple roots – or, more generally, $y^m = f(x)$ with $m \not\equiv 0 \bmod p$. In these cases the curve Γ is indeed smooth and Hasse's wording is correct.

But in general, if Γ has a singular point then there may be several prime divisors of F lying above it; these are called the "branches" of that singularity. In general not all of those branches are of degree 1. But in any case it is necessary to obtain a bound (independent of q) about the possible number of branches over singular points.

It seems that Hasse, when he finally discovered this necessity, was not quite sure about such a bound, and so he asked F. K. Schmidt to confirm it. Hasse's letters to F. K. Schmidt are lost but the replies of F. K. Schmidt are preserved and contained in Hasse's legacy. We have found 6 letters and postcards from F. K. Schmidt to Hasse between Jan 18 and Feb 6, 1933 where F. K. Schmidt explained that the relation (23) was true in general with a suitable modification of c. He exhibited a bound d for the number of branches lying over some singularity of Γ. The main point is that d is essentially defined by the singularity degree of Γ and hence is a geometric invariant, which means it is not altered by any extension of the base field. And so, (23) remains valid with the (very

[50] In general these are *not* identical with the points at infinity of the projectivization of Γ. For instance, let Γ be the lemniscate $x^2+x^2y^2+y^2 = 1$. Its projectivization has two points on the line at ∞, but these are singular points, each having two branches. These four branches are the "infinite solutions" which Hasse has in mind, and they have to be counted together with the finite solutions of the congruence. This had already been observed by Gauss in his "last entry". See Part 1, section 3.

rough) bound $c = n_x + n_y + d$. [51]

Thus, instead of only mentioning the "*infinite solutions*" which have to be taken into account, Hasse should also have mentioned the "*branches of the singular solutions*" in case the curve Γ has singularities. (Later in his survey paper [H:1934] he correctly used F. K. Schmidt's bound.)

Consequently, also for singular curves, N_Γ can be replaced by the number N of prime divisors of degree 1 of $F|K$. Geometrically speaking, the problem to obtain an estimate like (A) is of birational nature, and hence Γ may be replaced by the smooth projective curve (not necessarily planar) which is birationally equivalent to Γ over K.

Now, the connection with the Riemann hypothesis stems from the fact that this number N appears in the theory of F. K. Schmidt's zeta function $\zeta_F(s)$ defined as in (1). After introducing the new variable $t = q^{-s}$ it turns out that $\zeta_F(s)$ becomes a rational function in t, of the form [52]

$$\zeta_F(s) = \frac{L(t)}{(1-t)(1-qt)} \qquad (t = q^{-s}) \tag{24}$$

where

$$L(t) = 1 + (N - q - 1)t + \cdots + q^g t^{2g} \tag{25}$$

is a polynomial whose degree is twice the genus g of the function field F. The coefficient of t is $N - q - 1$. This formula is stated in one of the letters of F. K. Schmidt to Hasse which we mentioned above; it can be extracted in a straightforward manner from F. K. Schmidt's original paper [FK:1926]. If we write

$$L(t) = \prod_{1 \le i \le 2g} (1 - \omega_i t) \tag{26}$$

then the numbers ω_i are algebraic integers, and they are the inverses of the roots of ζ_F when regarded as a function of t. Thus the Riemann hypothesis says that $|\omega_i| = q^{\frac{1}{2}}$ for $i = 1, \ldots 2g$. From (25) and (26) we obtain

$$N - q - 1 = -(\omega_1 + \cdots + \omega_{2g})\,. \tag{27}$$

In this way the number N is connected with the inverse roots ω_i of the L-polynomial $L(t)$ of F.

Let θ denote the maximal real part of the zeros of $\zeta_F(s)$; then q^θ is the maximum of the $|\omega_i|$ and it follows from (27)

$$|N - q - 1| \le 2g \cdot q^\theta. \tag{28}$$

[51] F. K. Schmidt does not explicitly use the singularity degree. He works with the discriminant degree of $f(x, y)$ with respect to y, assuming y to be separable and integral over $K[x]$.

[52] See Part 1, section 6.1.

This is essentially the relation which Hasse mentions in his review of Mordell's paper. But here we see the term $N - q - 1$ appearing whereas Hasse speaks about $N - q$ (or rather $N - p$). Of course, the relation (28) implies $|N-q| < C \cdot q^{\theta}$ for *some* constant C, as in formula (A) of Hasse's review, but in order to take $C = 2g$ the relation (28) is not quite sufficient. Again we have the impression that Hasse, in his review of Mordell's paper, had in mind quadratic function fields only and Artin's zeta function.

Let us close this section by citing from a letter which Hasse had sent to Davenport on July 23, 1933.

It seems that Davenport had criticized Hasse's review of Mordell's paper because it contained too much of Hasse's view instead of a description of Mordell's methods. In his letter Hasse offered some apology but also gave a broad explanation of his motivation and his new viewpoint. Therefore we believe it worthwhile to be cited here.

> "*My dear Harold, many thanks for your letter...I quite agree with you that I ought to have mentioned something about Mordell's method instead of laying the main stress upon my own point of view. I most certainly appreciate the high value and the ingenuity of his line of attack and I do not in the least shut my eyes to the fact that his argument is at present the only one leading to definite results with the overwhelming lot of all these problems. On the other hand, the difference between us is that I do not consider the asymptotic questions as the original problem, particularly not when* p *is considered variable, perhaps a little more when* r *in* $q = p^r$ *is variable for fixed* p. *From my present point of view the analogue to Riemann's hypothesis lies in the center of interest, and the asymptotic, or rather non–asymptotic, behaviour of certain numbers of congruence solutions is the rational expression for this problem. From this point of view the question whether the F. K. Schmidt function contributes by itself to the solution of the congruence questions or not is quite unimportant. The line of idea is:*
>
> *F.K. Schmidt's function*
>
> $\longrightarrow$ *its zeros and their Riemann hypothesis*
> $\longrightarrow$ *connection with character sums or numbers of solutions*
> $\longrightarrow$ *sizing up of them by Mordell's method (or, if possible, exact determination by uniformization or arithmetical argument*[53]

[53] We shall see later in sections 5 and Part 3 what Hasse means when he speaks of "uniformization" or "arithmetical argument" in connection with the zeros of F. K. Schmidt's zeta function.

> *I will not say that, by putting F. K. Schmidt's function at the beginning, I confess myself as a decided analyst. On the contrary: F. K. Schmidt's function again is only a formal expression for the arithmetic and algebraic properties of the field F of algebraic functions, and it is the study of this field, which I consider as the original problem. In particular, the number of solutions, slightly filled up by the "infinite solutions"* [54]*, appears from here as the number of prime divisors of degree one, i. e., the analogue to the well–known densities in the common algebraic number theory.* [55] *The analogy of the algebra and arithmetic in a ... function field to the common algebraic number theory is perhaps the deepest reason for my own interest in all those questions as well as for their permanent significance altogether.*
>
> *That is exactly the line of ideas which I am going to follow in my great paper* [56] *on fields F of genus 1. ... you must allow me the right of putting my discoveries my own way ... I think, however, to serve you by this in the long run. For, while it is certain that Mordell's and your publications will find due interest with mathematicians of your own tendency, they must certainly run the risk of being overlooked or even regarded as uninteresting... by a great school of mathematicians that undoubtably forms an integrating and most active part of contemporary mathematics altogether.*
>
> *That is the reason why I dared bringing my own point of view even in a review on Mordell's paper. It seemed to me far more important to review for those who are liable to overlook the golden core in his paper ... than for those who, as you, already know the essence of it. I do not agree with you that a Zentralblatt review ought to save the reader reading the original paper, at any rate not in general. It ought to show the interested reader that there is something which deserves his particular interest. It should give therefore the ideas (in words) more than the details (in formulae), and of course no proofs at all. It should show where a result belongs in the system of knowledge. And it should be written with the intention to interest as far a circle of mathematicians as possible for the thing, provided that the thing deserves interest altogether... "*

We observe that when Hasse wrote this letter he was already in the possession of his proof of the Riemann hypothesis in the elliptic case. (See section 5.) He had already published a short preliminary announcement in the "*Göttinger*

[54] and the branches of the singular solutions, we may add. See our discussion above.

[55] We cannot quite follow Hasse here. In our opinion the exact analogue to the "well known densities" in algebraic number theory would be something different. Nevertheless it seems to be of historical interest that Hasse was motivated by this "analogy" as invoked by him here.

[56] Hasse means "long" paper.

Nachrichten" [H:1933]. In contrast to that short paper he was preparing a long paper where he intended to present all the details of proof. This long paper however did never appear in print because during its preparation Hasse changed his proof; the final proof appeared in three parts in 1936 only. Compare section 5.

By the way, upon Davenport's reminder Hasse decided to add a sentence to his review describing roughly Mordell's method.

4.2 The GF-method

As we have seen, Hasse in his letter to Davenport was quite outspoken about his new view. From Davenport we have not found any statement of similar general principle, not in his letters and not elsewhere. But we have got the impression that in contrast to Hasse, he continued to be predominantly interested in good estimates of diophantine congruences; he accepted the theory of zeta functions as a powerful method to this end only. In his discussions with Hasse he created a name for this method, namely "GF-method" where "GF" stands for Galois field, i.e. finite field.

The name "GF-method" appears in a letter of Davenport to Hasse dated Feb. 21, 1933. It seems that Hasse had asked him to explain Mordell's proof for the congruence $y^3 \equiv f_3(x) \bmod p$ where $f_3(x)$ is a polynomial of degree 3 without multiple roots. This congruence is the last entry in the list which Mordell had sent to Hasse in his letter of Dec 14, 1931. (See section 3.3.) In this particular case Mordell had already obtained the best possible exponent $\frac{1}{2}$ and Hasse wished to find out the main ideas behind Mordell's proof. Davenport wrote:

> "*. . . One way of seeing that $y^3 \equiv f_3(x)$ has $p+\mathcal{O}(\sqrt{p})$ solns. is your G.F. method. In $\mathbb{F}_{p^{3r}}$ $f_3(x)$ splits up, and the problem is the same as that for $y^3 \equiv f_2(x)$, and I suppose there will be no difficulty in showing that this has $p^r + \mathcal{O}(p^{\frac{1}{2}r})$ solutions. (In particular this follows from the corresponding result for $ax^2 + by^3 + c \equiv 0$, a case of $ax^m + by^n + c \equiv 0$). The result for p follows from that for $\mathbb{F}_{p^{3r}}$, although $f_3(x)$ does not split up $\bmod p$. Have I understood your + Artin's method correctly ? Has any account of it appeared in print, by the way ?. . .*"

We see that he indeed had understood the "GF-method" correctly. This consists of the following steps:

1. Find a suitable finite extension K of $\mathbb{F}_p$ (a GF=Galois field) and a curve

Γ birationally equivalent over K to the curve $y^3 = f_3(x)$, which is suitable for the problem in the following sense:

2. The number N_Γ of K-rational points should admit an estimate of the form $|N_\Gamma - q| \leq C\, q^{\frac{1}{2}}$ where q denotes the order of K. Moreover, a similar estimate should be valid for every finite extension $K^{(m)}$ of K, of degree m, with the same constant C. In other words: $|N_\Gamma^{(m)} - q^m| \leq C\, q^{\frac{m}{2}}$ for $m \to \infty$.
3. If this can be achieved then the Riemann hypothesis holds for the function field $F = K(\Gamma)$ over K.
4. But then the Riemann hypothesis holds also for the original field $\mathbb{F}_p(x, y)$ with $y^2 = f_3(x)$; for F is a base field extension of it.
5. This implies that the number of $\mathbb{F}_p$-rational solutions of $y^3 = f_3(x)$ is $p + \mathcal{O}(\sqrt{p})$.

Clearly, this "method" can be applied not only for $y^3 = f_3(x)$ but for any absolutely irreducible equation $f(x, y) = 0$ over $\mathbb{F}_p$. Moreover, instead of $\mathbb{F}_p$ one can work over any arbitrary finite field instead.

Davenport in his letter applies this method in the following way: Let K be any finite field over which $f_3(x)$ has a linear factor; this is certainly the case over $\mathbb{F}_{p^3}$. After a linear transformation over K one can assume that this factor is just x. After replacing x, y by $\dfrac{1}{x}, \dfrac{y}{x}$ the equation attains the form $y^3 = f_2(x)$, where $f_2(x)$ is a quadratic polynomial with different roots. By a linear transformation $f_2(x)$ can be brought into the form $f_2(x) = ax^2 + b$. Now we have an equation of the form $y^3 = ax^2 + b$ which, Davenport believes, will present no difficulties. He refers to the more general equation $ax^m + by^n + c = 0$ about which there had been an exchange of letters between Davenport and Hasse before.

I do not know Hasse's reply to Davenport's letter. Certainly Davenport had understood the GF-method correctly but Hasse, perhaps, would have omitted the last point 5. because his main interest was in the Riemann hypothesis which is already reached in point 4.

In the spirit of this GF-method, we remark that Davenport could have applied the method further to obtain additional simplification in this example. For, the curve $y^3 = f_3(x)$ is an elliptic curve with invariant $j = 0$ and therefore, over the algebraic closure of $\mathbb{F}_p$ it can be transformed into Weierstrass normal form $y^2 = x^3 - 1$. This is the curve for which Artin had settled the Riemann hypothesis in his letter to Herglotz. Alternatively, the curve is birationally equivalent to $x^3 + y^3 = 1$; this is the equation which Libri had already discussed in 1832, and also Gauss in his Disquisitiones.

We observe that in this "GF-method", p denotes a fixed prime number and thus the limit $p \to \infty$ is no longer of interest – as Hasse had said in his letter to Davenport. Instead, if one works over a field of q elements then the GF-method implies that one has to consider also its finite extensions $K^{(m)}$ with q^m elements; in this sense the limit $m \to \infty$ acquires relevance.

4.3 The Marburg seminar and Hasse's survey article

In Hasse's legacy we find notes for lectures delivered in February 1933 in Marburg. These notes cover more details than the notes for his Göttingen lecture. It seems to us that all this could not have been covered in one or two hours, so it is likely that Hasse gave a series of talks in his special algebra seminar in Marburg.

Looking through these notes we have found that the whole arrangement and the general formulas are similar to the published text in Hasse's survey article [H:1934] which appeared in the proceedings of the Berlin Academy of Sciences. A detailed discussion of that survey article has already been given in Part 1, section 6. There, I said that the article was written in 1934, but in the meantime I have come to the conclusion that this article may have been composed earlier already, perhaps even in February 1933 parallel or shortly after his seminar talks in Marburg where he presented this material.

Let us review the indications which seem to support the timing in favor of February 1933:

1. The article carries the subtitle: *Based in part on information obtained by Prof. Dr. F. K. Schmidt and Prof. Dr. E. Artin.* Now we have seen in the foregoing section that Hasse's correspondence with F. K. Schmidt about the zeta function had been in February 1933 and shortly before, more precisely: between Jan 20 and Feb 6, 1933. And the meeting with Artin had been at the end of November 1932, also not long before February 1933. It is not unreasonable to assume that Hasse wrote down the information which he had thus obtained as soon as possible.

2. In a letter to Davenport dated February 2, 1933, Hasse's remarks:

> "*After carefully studying F. K. Schmidt I am very much satisfied. This paper is really far better than Artin's, for the simple reason that all his formulae and notions are birationally invariant.*"

Here, Hasse refers on the one hand to Artin's thesis [A:1924], and on the other hand to F. K. Schmidt's paper [FK:1931] [57]. The latter had appeared in 1931 after a long delay. Much earlier, Hasse had already been informed about the content of this paper through F. K. Schmidt's letters but now we see that in January 1933 Hasse has read the paper in detail. It appears that his aim was

[57] See Part 1, sections 2 and 5.

to obtain a solid foundation for his survey article which, indeed, is based on the birationally invariant zeta function of F. K. Schmidt.

3. We note that Hasse does not mention in his article [H:1934] that he had already succeeded to prove the Riemann hypothesis for elliptic function fields. But he had obtained that proof, at least its early stages, at the end of February 1933 (as we shall see in the next section). This again points towards the conclusion that Hasse had prepared his manuscript for the article along with his seminar talks, hence before the end of February 1933.

4. We have seen in the foregoing section that Davenport, in a letter to Hasse dated Feb. 21, 1933, used what he called the "GF-method" and then asked: "*Has any account of it appeared in print, by the way ?*" This might have caused Hasse to sit down and write up the relevant facts, which he just had presented in his seminar, for publication.

Now, if indeed Hasse had prepared his survey article [H:1934] in February 1933 already, there arises the question why it appeared 1934 only. The publication date given in the paper is July 12, 1934 but there is no date of receipt mentioned. Perhaps, when Hasse wrote the manuscript in 1933 he did not intend to publish it but only wished to collect all information known at the time about zeta functions for function fields, and to arrange them in a form which is in line with his new point of view, just for himself and for Davenport. After all, the results which Hasse presents in this paper are for the most part not his own but due to Artin and F. K. Schmidt. And perhaps it was at a later stage only that Hasse decided to submit this survey article to publication, at a time when Hasse wished to have a convenient place for later reference.

Another reason for the delay may have been the political turbulences which had swept Germany after the Nazi party had come to power on Jan. 30, 1933. This may have had consequences for the operation of the Berlin Academy of Sciences, where Hasse's survey article appeared, resulting in a delay of its editorial dealings.

But all this is only speculation; we do not have evidence for this since, as already said, the article does not show a date of receipt.

4.4 Summary

From Hasse's own lecture notes for various colloquium lectures we can reconstruct the evolvement of his ideas on the Riemann hypothesis. After his meeting with Artin at the Hamburg colloquium (November 1932) the emphasis of his work was directed towards the study of the structure of function fields and their zeta functions. He voiced his new viewpoint in a Göttingen lecture (January 1933) to which he had been invited by Emmy Noether. Also, he publicized his view on the occasion of a Zentralblatt *review of a paper by Mordell. He pre-*

sented the details in several seminar lectures in Marburg (Feb 1933). On this occasion Hasse may have drawn up his survey paper on F. K. Schmidt's zeta functions which, however, appeared in 1934 only. In that paper he gathered all results known at that time about zeta functions of function fields, including unpublished results of Artin and F. K. Schmidt. That paper had been discussed in detail in Part 1.

5 The elliptic case: Hasse's first proof

5.1 The breakthrough

We all know that a good way to study a mathematical subject is to give a lecture course about it. The necessity to arrange the theory in a systematic way and to explain to the audience the various connections between the different results, often leads to new insights and, in consequence, to new results.

This happened also with Hasse on the occasion of his lectures in the Marburg seminar in February 1933 (see section 4.3). For in this same month Hasse succeeded to prove the Riemann hypothesis for elliptic function fields, i.e., for function fields of genus 1. This was three months after Hasse's notable conversation with Artin in Hamburg which had helped him to find his new viewpoint, i.e., directing his interest to the abstract algebraic theory of function fields over finite base fields.

But the structure theory of function fields in characteristic p was not yet sufficiently developed at the time, and so Hasse could not draw much from it for his problem. In fact, in the next years we can see Hasse himself busy pushing forward that theory, in order to find adequate structures which would explain his success. For the time being, Hasse's proof was based on classic complex multiplication and class field theory – techniques which hitherto nobody had seen to have any connection with the theory of zeta functions and the Riemann hypothesis in characteristic p. [58] In this respect Hasse's proof constituted a definite breakthrough although he could only deal with elliptic fields at the time. But he realized what should be done with fields of higher genus, namely using the analytic theory of abelian functions (where, however, the corresponding theory of CM-fields was not yet sufficiently developed).

From Hasse's correspondence with Davenport and with Mordell we can fairly accurately determine the date when this breakthrough happened. It seems that Davenport was the first one whom Hasse informed about it. That letter is lost, but the reply of Davenport is preserved, dated February 21, 1933. There we read:

[58] Including Herglotz [Her:1921] in his discussion of the lemniscate $x^2y^2 + x^2 + y^2 = 1$ over $\mathbb{F}_p$ (see Part 1, section 3). Herglotz does not even mention the Riemann hypothesis.

> "*. . . I am much excited as to whether your new idea for* $y^2 \equiv f_3(x)$ *comes off. The result for* $f_3(x,\, y) \equiv 0$ *follows from it without further work. Are you going to get new automorphisms or birational transformations from your method, or what ? . . .*"[59]

When Davenport writes that $f_3(x, y) \equiv 0$ follows from the former case "*without further work*" then this shows that he realizes the problem to be of birational nature; this he had learned from Hasse.

From Davenport's wording it is not absolutely certain that Hasse at that time had already completed his proof. His words could also be interpreted such that Hasse had given him a rough outline of his ideas without having them worked out already. This interpretation seems not to be unrealistic because in the next letter of Davenport we read:

> "*. . . I am waiting with great eagerness to hear what the final result of your work will be. It will be a marvellous achievement, and should lead to the solution of other problems, i.e., Kloosterman sums, which are closely connected to* $y^2 \equiv f_3(x)$. *I re-read your letter in which you explained your method the other day. . . I hope in a few days I shall be able to congratulate you on a final solution of the problem. . .*"

In a letter dated March 6, 1933 Hasse reported his new result to Mordell. That letter is preserved. It leaves no doubt that at this date Hasse was in possession of the proof:

> " *Dear Prof. Mordell, I succeeded recently in proving that the number of solutions of* $y^2 \equiv f_4(x) \bmod p$ *is* $p+$ *term which is less than or equal to* $2\sqrt{p}$ *. Moreover, the same holds for any Galois–field instead of rational Galois–field mod.* p *, that is, the analogue to Riemann's hypothesis is true for the corresponding Artin Zetafunction.*"

And Hasse continues:

> "*It is a curious fact that the leading idea of my proof may be considered as the fruit from our reading Siegel's great paper last year, or rather of my learning your method in the elliptic case. For, as there the* <u>equation</u> $y^2 = f_4(x)$ *is treated by uniformizing it through*

[59] It is amusing that Davenport always writes $y^2 = f_3(x)$ when he is talking about elliptic function fields, whereas Mordell prefers $y^2 = f_4(x)$. Of course, every elliptic function field (with finite base field of characteristic > 2) can be generated by both of these normal forms. Hasse, in his correspondence with Davenport and with Mordell, used $f_3(x)$ or $f_4(x)$ respectively, according to his adressee.

> *elliptic functions, so I now treat the* <u>*congruence*</u> $y^2 \equiv f_4(x)$ *mod.p by uniformizing it the same way...*"

Here, Hasse refers to Mordell's visit one year earlier, during the Easter vacations 1932. The Mordells had stayed with the Hasses in Marburg. It happened that Hasse at that time had to write a *Jahrbuch review* of Siegel's paper [Si:1929], and he used Mordell's presence to read Siegel's paper together. [60] It seems that on that occasion Mordell had explained to Hasse his use of the uniformization of elliptic curves in his old paper [Mor:1922] which contained Mordell's part of the "Mordell-Weil theorem". – We shall explain in section 5.3.2 what Hasse had in mind when he said he is treating the elliptic congruence by "uniformizing through elliptic functions".

Mordell's reply to Hasse's letter is dated March 9; we have cited it partly in the introduction already. After the text which we have cited there, Mordell writes:

> "*What a tremendous vindication (for those who need it and have not appreciated the K.k.theory* [61] *that the proof should depend upon such a comparatively high brow theory. I feel rather relieved to think I did not spend too much time on further results of this kind with my method, and very pleased that my old paper should have supplied even an amount* ε *of usefulness. We must read another paper some other time.*"

The first sentence reflects what we have already mentioned in section 2.2, that Mordell was no friend of "high brow theory". In this case, however, he seems to have accepted it; to him the proof of the Riemann hypothesis seems to carry sufficient "vindication".

Perhaps it was the use of class field theory in Hasse's first proof which later, in October 1933, induced Mordell to ask Hasse about a possible English translation of Hasse's Marburg Lecture Notes [H:1932] on class field theory. [62] Those notes contained the foundations of general class field theory according to the state of the art at the time, they had been mimeographed and distributed among interested mathematicians (including Davenport and Mordell). Perhaps Mordell wanted his British colleagues and students to learn class field theory in order that they would be able to follow Hasse's first proof of the Riemann

[60] The review appeared in volume 56 of the *Jahrbuch für die Fortschritte der Mathematik.* It has the unusual size of more than 5 pages.

[61] K.k.theory = *Klassenkörpertheorie* = class field theory

[62] It may well be that also Emmy Noether, when she was in England in the fall of 1933, had strongly proposed this plan for translation. She indicates this in a letter to Hasse of March 6, 1934 from Bryn Mawr where she reports that she had insisted to produce more copies of the translation because of strong demand.

hypothesis.[63] Hasse consented to the translation and recommended Reinhold Baer as someone who would well be able to do it.[64] Moreover, Baer was to edit and add the second part of Hasse's lecture, covering Hasse's new theory of norm residues; this part had not yet been included in the published lecture notes. In the end Baer and Mahler were designated for the translation. But in spring 1934 this plan was abandoned for reasons not known to me. One reason may have been that Hasse had changed his proof of the Riemann hypothesis and the second proof did not use class field theory any more.

In the following sections we will discuss Hasse's first proof in more detail.

5.2 The problem

Let F be an elliptic function field with finite base field K, and let $q = p^r$ be the order of K. The genus of F is $g = 1$, and hence the L-polynomial (25) of F is quadratic:

$$L(t) = 1 + (N - q - 1)t + qt^2 = (1 - \omega_1 t)(1 - \omega_2 t)\,.$$

The Riemann hypothesis claims that $|\omega_1| = |\omega_2| = q^{\frac{1}{2}}$. It is convenient to work with the reciprocal polynomial:

$$L^*(t) = t^2 + (N - q - 1)t + q = (t - \omega_1)(t - \omega_2) \tag{29}$$

which has ω_1 and ω_2 as its roots. If $L^*(t)$ is irreducible over $\mathbb{Q}$ then ω_1, ω_2 are conjugate integers in some quadratic number field Ω. If this Ω is an *imaginary* quadratic field then $|\omega_1| = |\omega_2|$, and since their product is q the Riemann hypothesis follows. We see that it suffices to solve the following problem (independent of whether $L^*(t)$ is irreducible or not):

> *Given an elliptic function field $F|K$ with finite base field K with q elements, let N denote the number of prime divisors of degree 1 in $F|K$. Find an imaginary quadratic number field Ω and an element $\pi \in \Omega$ such that*
>
> $$\mathcal{N}(\pi) = q \qquad \text{and} \qquad \mathcal{S}(\pi) = -(N - q - 1) \tag{30}$$
>
> *where $\mathcal{N}$, $\mathcal{S}$ denote norm and trace for $\Omega|\mathbb{Q}$.*

For, it follows from this that π and its conjugate π' are the roots of $L^*(t)$ and hence ω_1, ω_2 can be identified with π, π'.

[63] Mordell himself did not need a translation since he had fairly good knowledge of German.

[64] Reinhold Baer had to leave Germany due to the antisemitic policy of the German Nazi government, and he had come to England after Hasse had recommended him to Mordell. Hasse held a friendship with Baer from their time as colleagues in Halle – a friendship which lasted lifelong. Hasse hoped that Baer would get some additional financial support through the translating job.

Note that $\mathcal{N}(\pi-1)=\mathcal{N}(\pi)-\mathcal{S}(\pi)+1$, hence the relations (30) are equivalent to

$$\mathcal{N}(\pi)=q \qquad \text{and} \quad \mathcal{N}(\pi-1)=N\,. \qquad {}^{65}) \tag{31}$$

Today we can readily solve this problem, as follows:

F is the function field of an elliptic curve $\mathcal{E}$ defined over K, i.e., $F=K(\mathcal{E})$. We may regard $\mathcal{E}$ as an abelian variety of dimension 1. The K-rational points of $\mathcal{E}$ form a finite subgroup $\mathcal{E}(K)$, and its order equals the number N of prime divisors of degree 1 of F. Consider the endomorphism ring $\mathrm{End}(\mathcal{E})$. It is known that (up to a few exceptional cases) this endomorphism ring is a subring of an imaginary quadratic field Ω.

Every endomorphism $\mu\in\mathrm{End}(\mathcal{E})$ has a degree, and this degree equals the norm $\mathcal{N}(\mu)$ from Ω to $\mathbb{Q}$. If μ is separable then the degree equals the order of the kernel of μ.

Now, the *q-Frobenius endomorphism* $\pi\in\mathrm{End}(\mathcal{E})$ is defined over K, and it raises the coordinates of any point of $\mathcal{E}$ into their q-th power. Applied to a generic point, we see that its degree is $[F:F^q]=q$, which gives the first relation in (31). On the other hand, by definition π fixes precisely the K-rational points; hence $\mathcal{E}(K)$ is the kernel of $\pi-1$. This gives the second relation in (31) since $\pi-1$ is separable.

In other words: *The polynomial $L^*(t)$ is the characteristic polynomial of the Frobenius endomorphism within the endomorphism ring* $\mathrm{End}(\mathcal{E})$*, and the latter is a subring of an imaginary quadratic field* Ω.

The "few exceptional cases" mentioned above are those where $\mathcal{E}$ is "supersingular" , i.e., where $\mathrm{End}(\mathcal{E})$ is not commutative [66]. Then $\mathrm{End}(\mathcal{E})$ is contained in the quaternion algebra which is ramified at p and ∞ only. Every maximal commutative subfield of this quaternion algebra is imaginary quadratic. Since the Frobenius endomorphism π is contained in such a maximal subfield, the same argument as above can be applied in this case. ($\mathcal{N}$ denotes the reduced norm from this quaternion algebra.)

But this line of argument was not yet available when Hasse started his work in 1933. The standard notions and results about elliptic curves in characteristic p and their endomorphism rings which we have used here, were unknown. In fact, what we have sketched above constitutes Hasse's second proof which we will discuss in more detail in the next part. It was quite a formidable task to develop the theory of elliptic function fields and the structure of their endomorphism rings from scratch, sufficiently far so that the above reasoning could be applied.

[65]) Compare this with the last entry in Gauss' diary which we have discussed in Part 1, section 3.

[66] Those elliptic curves had been discovered by Hasse [H:1934d], [H:1936c] . The terminology "supersingular" seems to have been introduced by Deuring [Deu:1941].

Nevertheless, in his first proof Hasse succeeded to construct, starting from $F|K$, a certain imaginary quadratic field Ω and $\pi \in \Omega$ satisfying (31). He found Ω and π on a detour via characteristic zero, lifting the given function field $F|K$ suitably to an elliptic function field over an algebraic number field, and then applying the classic theory of complex multiplication. It was only after having completed his first proof, while preparing the manuscript, that Hasse discovered he was able to transfer the relevant results from classic complex multiplication to the case of characteristic p – thus arriving at his second proof we have sketched above already.

5.3 The proof

Hasse has never published his first proof completely. But there is enough material from which we can quite well extract his main ideas and are able to reconstruct the proof. For the discussion in this section we have used the following sources:

1. Hasse's letter to Mordell dated March 6, 1933 – the same letter from which we have cited above already. It contains not only an announcement of Hasse's results but also a clear description of his ideas for his first proof, without details however.
2. Hasse's preliminary announcement which appeared in the proceedings of the "*Gesellschaft der Wissenschaften in Göttingen*" [H:1933]. The paper is dated April 14, 1933 and was communicated by E. Landau in the meeting of the society of April 28, 1933. Again, this is a decription of the method only, without details of proof. It is still encumbered with certain restrictive assumptions on the degree f of the invariant of the given elliptic function field. In Hasse's letter to Mordell these restrictions had not been mentioned, from which we may perhaps conclude that he was pretty sure to be able to overcome this difficulty in due course. But in his Göttingen preliminary announcement Hasse was cautious, and so he worded the title of this paper as:

    > "*Proof of the analogue of the Riemann hypothesis for Artin's and F. K. Schmidt's congruence zeta functions in certain elliptic cases.*"

    Thus he does not yet claim to have a proof for "all" but only for "certain" elliptic cases.

    It is in this paper that Hasse acknowledged fully the role of Mordell and Davenport; he wrote:

    > "*. . . and anyway, for my whole investigation I got the main motivation in many discussions with him [Davenport] and Mordell*

> *about numbers of solutions of congruences and related things.*"[67]

3. Hasse's own lecture notes for a lecture in Marburg, May 1933. This lecture was meant as a continuation of his lecture in the foregoing semester, in February of the same year (see section 4.3). Now he wished to report about his progress, i.e., his (first) proof for the elliptic case.
4. Another manuscript of Hasse's lecture notes, this time for a talk at the annual DMV-meeting in Würzburg, September 1933. These notes are written with typewriter, and they contain more details, though not the full proof. [68] It seems improbable that Hasse could have covered all this in a short talk at the DMV-meeting. We have the impression that this manuscript was conceived as a publication somewhere, probably in the *Jahresbericht der DMV* where it was common to report about the talks at the DMV-meetings, and at the same time add some additional information which could not be covered in the talk. But when Hasse actually wrote his report [H:1934b] in the *Jahresbericht* then this turned out to be brief, containing the following announcement:

   > "*In the meantime I succeeded to carry the proof for all elliptic cases with characteristic not* 2 *with purely algebraic methods... Since this new proof seems to me to be more suited to the problem, I will refrain from the detailed publication of the first analytic proof. The publication of the second algebraic proof will shortly be given as a sketch in the "Abhandlungen des Mathematischen Seminars Hamburg", and in detail in Crelle's Journal.*"

   The announced sketch in the "*Hamburger Abhandlungen*" appeared in [H:1934a], even without the restriction that the characteristic $p \neq 2$. The publication in Crelle's Journal appeared in three parts: [H:1936a], [H:1936b], [H:1936c].
5. An extensive manuscript, comprising 94 pages, from Hasse's own hand where he presents all the details. [69] It seems to us that this was conceived as part of the manuscript for Crelle's Journal where Hasse planned to publish his proof. Consider what he said later in [H:1966], in accordance with what we have cited above in 4.:

   > " *I had the intention of letting detailed proofs follow in Crelle's Journal. I never made that true however. For shortly afterwards I discovered an entirely new proof based on the theory of*

[67] It was this remark which induced me to search for more details about the cooperation between Hasse, Mordell and Davenport, with the result as presented in the foregoing sections.

[68] At the end of this manuscript Hasse had added a chapter about his work with Davenport on the Riemann hypothesis in the case $ax^m + by^n = 1$, solving this with Gauss sums. Hasse had worked on this during the summer of 1933 jointly with Davenport, who stayed that semester in Göttingen but often came over to Marburg. We shall discuss it in Part 3.

[69] I am indebted to Reinhard Schertz for providing me with a copy of this manuscript.

> *abstract elliptic function fields, and free from the restrictions to $p \neq 2, 3$ and f odd.*"

The restriction "*f odd*" refers to the degree restriction mentioned above in 2. already.

6. Hasse's proof, many years later in [H:1966], of an important lemma which he had to use in his first proof. Hasse says there:

   > "*I outlined all this* [the proof of the Riemann hypothesis in the elliptic case] *briefly in a preliminary communication. When today I come back to my original proof, it is because I think that the fundamental lemma on which my uniformization was based has sufficient interest in itself to be published with full proof.*"

   The "fundamental lemma" is the same lemma which we have called "invariant-lifting lemma." See section 5.3.1 below.
7. Shiratani's proof of the invariant-lifting lemma [Shi:1967]. While Hasse [H:1966] still had to impose the said degree restrictions, Shiratani was able to free the lemma from these restrictions. However he had to use results which had not yet been available in 1933.

We will discuss Hasse's first proof in the next three sections. In this proof Hasse assumed $p > 3$. The reason was that for $p = 2$ or 3 the explicit formulas for generating the function field, for the computation of the absolute invariant, for addition of points and for complex multiplication, are different from the general case. Hasse was confident that if a proof could be achieved for $p > 3$ then it would be a matter of routine only to include $p = 2$ and $p = 3$ too.

5.3.1 Lifting of the absolute invariant

In characteristic $p > 3$ the given elliptic function field $F|K$ of characteristic p can be generated by two elements x, y satisfying an equation in Weierstrass normal form:

$$y^2 = 4x^3 - ax - b \quad \text{with} \quad a, b \in K\,. \tag{32}$$

The discriminant

$$\Delta := a^3 - 27b^2$$

does not vanish, and the absolute invariant

$$j := 12^3\,\frac{a^3}{\Delta} \tag{33}$$

characterizes the elliptic function field $F|K$ up to base field extensions. More precisely: If $F'|K$ is another elliptic function field having the same invariant j then F and F' become isomorphic under some finite base field extension.

Hasse excludes the cases $j = 0, 12^3$ since, he says, these cases can easily be dealt with directly. He refers to a forthcoming joint paper with Davenport. [70]

The first step of Hasse's "uniformization" of the function field F, or rather of the projective curve defined by (32), is to lift the invariant $j \in K$ to characteristic 0. Hasse does it by means of the analytically defined modular function $j(\mathfrak{w})$. As follows.

An elliptic curve Γ over the complex field $\mathbb{C}$ can be regarded as a Riemann surface of genus 1. Every such surface is analytically uniformized by the factor space $\mathcal{E}_{\mathfrak{w}} := \mathbb{C}/\mathfrak{w}$ where $\mathfrak{w}$ is the lattice of the periods. We write $\Gamma_{\mathfrak{w}}$ for the elliptic curve uniformized by $\mathfrak{w}$. The function field $F_{\mathfrak{w}}$ of $\mathcal{E}_{\mathfrak{w}}$ over $\mathbb{C}$ is generated by the Weierstrass $\wp$-function and its derivative:

$$F_{\mathfrak{w}} = \mathbb{C}(\wp, \wp')$$

where

$$\wp(z|\mathfrak{w}) = \frac{1}{z^2} + \sum_{0 \neq w \in \mathfrak{w}} \left(\frac{1}{(z-w)^2} - \frac{1}{w^2} \right) , \tag{34}$$

z being a complex variable. $\wp(z|\mathfrak{w})$ as a function of z is meromorphic with a pole of order 2 at 0, and it is periodic with respect to $\mathfrak{w}$, i.e., it is a meromorphic function of $\mathbb{C}/\mathfrak{w} = \mathcal{E}_{\mathfrak{w}}$. This function satisfies the differential equation [71]

$$\wp'^{\,2} = 4\wp^3 - g_2\wp - g_3 \tag{35}$$

where

$$g_2(\mathfrak{w}) = 60 \sum_{0 \neq w \in \mathfrak{w}} \frac{1}{w^4} , \qquad g_3(\mathfrak{w}) = 140 \sum_{0 \neq w \in \mathfrak{w}} \frac{1}{w^6} . \tag{36}$$

The points of $\Gamma_{\mathfrak{w}}$ (in suitable coordinates) are then given by the pairs

$$\Big(\wp(z|\mathfrak{w}) \, , \ \wp'(z|\mathfrak{w})\Big) \qquad \text{for} \quad z \in \mathcal{E}_{\mathfrak{w}} \, .$$

The point at ∞ of $\Gamma_{\mathfrak{w}}$ corresponds to $z = 0$.

The invariant $j(\mathfrak{w})$ of the lattice (or of the function field) is defined by

$$j(\mathfrak{w}) = 12^3 \, \frac{g_2(\mathfrak{w})^3}{\Delta(\mathfrak{w})} \qquad \text{where} \qquad \Delta(\mathfrak{w}) = g_2(\mathfrak{w})^3 - 27 g_3(\mathfrak{w})^2 \, . \tag{37}$$

These are the classical formulas for the uniformization of the elliptic curves (or function fields) over $\mathbb{C}$. Two elliptic curves $\Gamma_{\mathfrak{w}}$ and $\Gamma_{\mathfrak{w}'}$ are isomorphic if and

[70] It is not clear whether he meant the paper [Da-H:1934] since he also planned another joint paper with Davenport which, however, was never completed. But in [Da-H:1934] the cases $j = 0, 12^3$ are contained as special cases.

[71] Differentiation is to be understood with respect to the complex variable z.

only if their lattices $\mathfrak{w}$ and $\mathfrak{w}'$ are proportional, which means that there exists a complex number λ such that $\lambda\mathfrak{w} = \mathfrak{w}'$. We have:

$$g_2(\lambda\mathfrak{w}) = \lambda^{-4} g_2(\mathfrak{w}) \,, \qquad g_3(\lambda\mathfrak{w}) = \lambda^{-6} g_3(\mathfrak{w}) \,, \qquad j(\lambda\mathfrak{w}) = j(\mathfrak{w}) \,. \qquad (38)$$

If $\mathfrak{w}$ and $\mathfrak{w}'$ are not proportional then $j(\mathfrak{w}) \neq j(\mathfrak{w}')$. Every complex number is the invariant $j(\mathfrak{w})$ of some lattice $\mathfrak{w}$.

A *multiplier* of $\mathfrak{w}$ is defined to be a complex number μ such that $\mu\mathfrak{w} \subset \mathfrak{w}$. This defines an endomorphism $\mu : \mathcal{E}_{\mathfrak{w}} \to \mathcal{E}_{\mathfrak{w}}$. The ring M of all those multipliers contains $\mathbb{Z}$. If there are other multipliers in M then these are not real numbers; in this case the lattice $\mathfrak{w}$ (or the elliptic curve $\Gamma_{\mathfrak{w}}$) is said to admit *complex multiplication.* From this the whole theory derives its name.

If the multiplier ring M is complex then it is a subring of some imaginary quadratic number field Ω. [72] In this case the lattice $\mathfrak{w}$ is proportional to some lattice $\mathfrak{a} \subset \Omega$, and $\mathfrak{a}$ is an ideal (integral or fractional) of M. [73] The ring M is determined by its *conductor* $m \in \mathbb{N}$, in such a way that M consists of all algebraic integers in Ω which are congruent to some rational integer modulo m. [74] Let us write M_m if we wish to indicate the conductor m of M.

Now, the first observation in the theory of complex multiplication is that the invariant $j(\mathfrak{a})$ of such a "singular" lattice is an algebraic number. Moreover, the field extension $\Omega_m := \Omega(j(\mathfrak{a}))$ depends on m only, and it is abelian over Ω. Its Galois group is canonically isomorphic to a certain group of divisor classes of Ω, namely the factor group D_m/H_m where D_m is the group of all divisors of Ω relatively prime to m, and H_m the group of the principal divisors of the form (ab^{-1}) where $0 \neq a, b \in \mathsf{M}_m$ are both prime to m. This isomorphism is an example of Artin's general reciprocity law.

Ω_m is called the m-th *ring class field* over Ω. According to general class field theory, Ω_m can be characterized by the splitting properties of the prime divisors $\mathfrak{p}$ of Ω which are prime to m. Such a prime divisor splits completely in Ω_m if and only if $\mathfrak{p} \in H_m$. More generally, let $\mathfrak{p}^f$ be the smallest power of $\mathfrak{p}$ contained in H_m; then f is the relative degree of a prime divisor $\mathfrak{P}|\mathfrak{p}$ of Ω_m.

All these facts were well known from the classical theory of complex multipli-

[72] In today's algebraic theory the terminology "*endomorphism ring*" is used instead of "*multiplier ring*".

[73] In the following, the letter $\mathfrak{a}$ will denote a lattice contained in some imaginary quadratic field. In the older literature these are called the "singular" lattices. Perhaps it is not superfluous to point out that this terminology of "singular" does not have anything to do with "singular points" etc. in algebraic geometry. In the old terminology "singular" just means "special" (and "general" often means "except perhaps some singular cases"). Nowadays, in order to avoid misunderstandings, the terminology "singular" for a lattice which admits complex multiplication is not in use any more, but in Hasse's time it still was. Same with "supersingular".

[74] Hasse says "index" instead of "conductor", following the classic terminology in the theory of complex multiplication.

cation. Hasse was quite familiar with the theory of complex multiplication. Early in his youth, when he was in Kiel as *Privatdozent*, he was confronted with the theory of complex multiplication through the study of Takagi's paper. In Part I of his class field theory report [H:1926] he had discussed the Weber-Takagi class field theory, in particular the so-called *Jugendtraum* of Kronecker which claimed that every abelian field over Ω is contained in the field generated by the singular values $j(\mathfrak{a})$ and the roots of unity. Since Takagi and Fueter we know that this is not quite true because one has to add certain division values of elliptic functions. Already in his report Hasse announced that he himself will give a more adequate treatment, in as much it is possible to work with Weber's function τ only.

And he did so in his papers [H:1927], [H:1931] whose aim was to systematize the theory of complex mutiplication, once from the viewpoint of modern class field theory, and a second time to derive class field theory for imaginary quadratic fields with analytic methods. [75] Concerning these papers, Hasse had had an exchange of letters with Hecke and with Artin; both had read and commented his manuscripts. And in 1931 there followed three more papers on complex multiplication [H:1931a], [H:1931b], [H:1931c]. In Hasse's legacy we have found an extensive manuscript on complex multiplication, going beyond his published papers. [76]

This familiarity with complex multiplication was quite decisive for Hasse's success in solving the above mentioned problem (31). Without this knowledge Hasse would not have found the motivation to look for the construction of Ω and π within the realm of complex multiplication.

Hasse's **invariant-lifting lemma** now reads as follows:

> *Let K be a finite field of order $q = p^r$. Let j be an element of K and f its degree, so that $\mathbb{F}_p(j) = \mathbb{F}_{p^f}$ and $f \mid r$.*
> *There exists an imaginary quadratic number field Ω and a lattice $\mathfrak{a} \subset \Omega$ with multiplier ring M such that:*
>
> (i) *$p = \mathfrak{p}\mathfrak{p}'$ splits in Ω into prime divisors (which may be equal or not).*
> (ii) *$\mathfrak{p}$ (or equivalently, p) does not divide the conductor m of M.*
> (iii) *Each prime divisor $\mathfrak{P}$ of p in the ring class field $\Omega_m = \Omega(j(\mathfrak{a}))$ is of degree f, hence its residue field $\Omega_m/\mathfrak{P}$ is isomorphic to $\mathbb{F}_{p^f}$.*

[75] An excellent survey of the theory in the same style which Hasse had used is contained in Deuring's article [Deu:1958]. Today, the theory of complex multiplication can be derived without the use of analytic functions; see [Deu:1949], [BCHIS:1966].

[76] Prof. Schertz has pointed out to me that this manuscript is modeled after Weber's treatment of complex multiplication in [Web:1908].

(iv) *After suitable identification of* $\Omega_m/\mathfrak{P}$ *with* $\mathbb{F}_{p^f}$ *we have*

$$j(\mathfrak{a}) \equiv j \bmod \mathfrak{P}\,. \tag{39}$$

"In general" the lattice $\mathfrak{a}$ *is uniquely determined, up to a proportionality factor which has to be contained in* Ω*; exceptions can arise only when* p *is ramified in* Ω*, and* $f = 1$ *or* $f = 2$*.*

In other words: The value $j(\mathfrak{a}) \in \mathbb{C}$ of the modular j-function at the "singular" lattice $\mathfrak{a}$ is a lift of the given element $j \in K$, with the properties as announced.

Actually, Hasse in [H:1933] states this theorem under the already mentioned extra condition that the degree f is *odd*. [77] This is indeed necessary for Hasse's method [H:1966] to work. For he had to choose the invariant $j(\mathfrak{a})$ (and hence $\mathfrak{a}$) among the roots of the so-called "invariants equation" of transformation degree p^f; the behavior of that equation modulo p for $f > 1$ is somewhat involved and yields the result only if f is odd.

Hasse said that this extra condition is of "technical" nature, and he was sure that this obstacle could be overcome in due course. As we remarked above already, Shiratani [Shi:1967] indeed showed that Hasse's invariant-lifting lemma holds also if the degree f is even, but for this he used the later work of Hasse and Deuring which had not yet been available in 1933. [78]

From today's knowledge, the "technical" obstacle which Hasse encountered is not completely technical but it is partly due to the inherent structure. This becomes clearer if we consider the uniqueness assertion in the above lemma which holds "in general" only. In fact, Hasse found exceptions of that uniqueness, but for odd f he showed that these can occur only if $f = 1$ and p is ramified in Ω. (In such exceptional case, there are two essentially different solutions $\mathfrak{a}$, $\mathfrak{a}'$.) Today we know that uniqueness fails to hold if and only if the endomorphism ring of F is non-commutative, i.e., F is "supersingular". And that those supersingular cases can occur only if $f = 1$ or $f = 2$.

At the time when Hasse conceived his proof he had not yet discovered the existence of those "supersingular" cases; this happened later. See [H:1934d] and [H:1936c]. It may well be that this discovery occurred while he attempted to free his proof from his extra condition. As it turns out, this extra condition

[77] Therefore the case $f = 2$ in the last sentence of the lemma does not show up in Hasse's manuscript.

[78] Among other results, Shiratani used Deuring's lemma that every elliptic function field in characteristic p equipped with a multiplier can be obtained as a good reduction of some elliptic function field in characteristic 0 with a multiplier. See [Deu:1941]. This lemma is in fact the essential ingredient of Shiratani's proof, and Hasse's invariant-lifting lemma is a more or less direct consequence of it. It should be noted that Deuring's paper became possible only *after* Hasse had developed the general theory of the endomorphism rings of elliptic curves over arbitrary characteristic.

avoided the supersingular cases for $f = 2$. In the case $f = 1$, Hasse was able to deal with the supersingular cases because the "invariants equation" of transformation degree p, as is well known in complex multiplication, is of particular simple form modulo p, namely

$$J_p(t,t) \equiv -(t^p - t)^2 \bmod p\,.$$

Now, in the situation of the invariant-lifting lemma let $\mathfrak{p}$ denote the prime divisor of Ω induced by $\mathfrak{P}$. Since $\mathfrak{P}$ is of degree f it follows from class field theory that $\mathfrak{p}^f$ is a principal divisor of some element in the multiplier ring M relatively prime to m. Since f divides r it follows that $\mathfrak{p}^r \cong \pi$ is principal too. Similarly for $\mathfrak{p}'$. Consequently we have

$$q = p^{\,r} = \pi\pi' \tag{40}$$

with some conjugate elements $\pi, \pi' \in \mathsf{M}$ relatively prime to m.

This takes care of the first relation in (31).

5.3.2 Uniformization

Before discussing the second relation in (31) let us explain what Hasse, in his letter to Mordell of March 6, 1933, describes as "uniformization of an elliptic congruence". (See section 5.1.)

Instead of the Weierstrass functions $\wp(z|\mathfrak{w})$, $\wp'(z|\mathfrak{w})$ Hasse uses the modified functions [79]

$$\widetilde{\wp}(z|\mathfrak{w}) = 12\frac{\wp(z|\mathfrak{w})}{\sqrt[6]{\Delta(\mathfrak{w})}}\,, \qquad \widetilde{\wp'}(z|\mathfrak{w}) = \frac{\wp'(z|\mathfrak{w})}{\sqrt[4]{\Delta(\mathfrak{w})}}\,. \tag{41}$$

Accordingly the Weierstrass equation (35) is to be replaced by the modified Weierstrass equation

$$12^3\,\widetilde{\wp'}^2 = 4\widetilde{\wp}^3 - 12\gamma_2\,\widetilde{\wp} - 8\gamma_3 \tag{42}$$

where $\gamma_2(\mathfrak{w})$ and $\gamma_3(\mathfrak{w})$ are modular functions which can be expressed by $j(\mathfrak{w})$:

$$\gamma_2(\mathfrak{w}) = \sqrt[3]{j(\mathfrak{w})} \qquad \text{and} \quad \gamma_3(\mathfrak{w}) = \sqrt[2]{j(\mathfrak{w}) - 12^3}\,. \tag{43}$$

[79] The notation is ours. Classically, Hasse writes π, $\hat{\pi}$ instead of $\widetilde{\wp}$, $\widetilde{\wp}'$. Note that $\widetilde{\wp}'$ is the modification of the derivative $\wp'$ and not the derivative of the modification $\widetilde{\wp}$. – Observe that the modular functions $\sqrt[6]{\Delta(\mathfrak{w})}$ and $\sqrt[4]{\Delta(\mathfrak{w})}$ are multi-valued. The formulas below are to be understood that one branch of $\sqrt[12]{\Delta(\mathfrak{w})}$ has to be chosen, and then in (41) its respective powers are to be taken. Similarly the third and second roots in (44) below have to be interpreted coherently with $\sqrt[12]{\Delta(\mathfrak{w})}$.

One of the reasons for this modification is that the coefficients $\gamma_2(\mathfrak{w})$,$\gamma_3(\mathfrak{w})$ of the algebraic equation (42) are directly expressible (although not rationally) by the absolute invariant $j(\mathfrak{w})$.

Now take $\mathfrak{w} = \mathfrak{a}$ to be the "singular" lattice of the invariant-lifting lemma. Then the coefficients of the corresponding modified Weierstrass equation are

$$\gamma_2(\mathfrak{a}) = \sqrt[3]{j(\mathfrak{a})} \qquad \text{and} \quad \gamma_3(\mathfrak{a}) = \sqrt[2]{j(\mathfrak{a}) - 12^3}\,. \tag{44}$$

and hence algebraic over $\Omega(j(\mathfrak{a}))$. In fact, it is known that together with $j(\mathfrak{a})$ they generate an abelian extension of Ω which can also be described as a certain ring class field, namely Ω_{tm} with certain t dividing 6. The elliptic curve $\Gamma_\mathfrak{a}$ with the equation

$$\Gamma_\mathfrak{a}: \qquad 12^3\,y^2 = 4x^3 - 12\gamma_2(\mathfrak{a})x - 8\gamma_3(\mathfrak{a}) \tag{45}$$

is defined over Ω_{tm} , and its absolute invariant is $j(\mathfrak{a})$. Thus $\Gamma_\mathfrak{a}$ admits $\mathcal{E}_\mathfrak{a} = \mathbb{C}/\mathfrak{a}$ as its uniformization. This uniformization is given by the modified Weierstrass functions, in such a way that the points $(x, y) \in \Gamma_\mathfrak{a}$ are given by

$$x = \widetilde{\wp}(z\,|\mathfrak{a})\,, \quad y = \widetilde{\wp}'(z\,|\mathfrak{a}) \qquad \text{(for} \quad z \in \mathcal{E}_\mathfrak{a})\,. \tag{46}$$

Now consider the prime divisor $\mathfrak{P}$ of Ω_m as in the invariant-lifting lemma. We extend $\mathfrak{P}$ to a prime divisor (valuation) first of Ω_{tm} and then of the complex field $\mathbb{C}$; for simplicity the extended prime will also be denoted by $\mathfrak{P}$. Let γ_2, γ_3 be the residue classes of $\gamma_2(\mathfrak{a})$, $\gamma_3(\mathfrak{a})$ modulo $\mathfrak{P}$; they are elements in the algebraic closure of K. Then the "reduced curve"

$$\Gamma: \qquad 12^3 y^2 = 4x^3 - 12\gamma_2 x - 8\gamma_3 \tag{47}$$

is defined over some finite overfield of K, and its absolute invariant is j in view of (39). Consequently, Γ and the given curve (32) are birationally equivalent over some finite overfield of K. Since a base extension is permitted (GF-method!) we may assume from the start that γ_2, γ_3 are already in K and that the two curves are birationally equivalent over K. [80]

In other words, we have Hasse's **Uniformization theorem**:

> *After a suitable finite base field extension we may assume that $F = K(\Gamma)$ where Γ denotes an elliptic curve defined by an equation of modified Weierstrass form* (47) *with coefficients $\gamma_2\,, \gamma_3 \in K$ and*
>
> $$\gamma_2 = \sqrt[3]{j}\,, \qquad \gamma_3 = \sqrt[2]{j - 12^3}\,, \tag{48}$$

[80] Observe that by performing a base field extension, q is replaced by a power of q. Accordingly, π and its conjugate π' which are defined by (40) are also to be replaced by their respective powers.

j being the absolute invariant of F. This curve Γ is the reduction modulo $\mathfrak{P}$ of the curve $\Gamma_{\mathfrak{a}}$ in (45)*, i.e., we have*

$$\gamma_2 \equiv \gamma_2(\mathfrak{a})\,, \quad \gamma_3 \equiv \gamma_3(\mathfrak{a}) \quad \bmod \mathfrak{P} \tag{49}$$

in addition to (39) *. By definition, $\Gamma_{\mathfrak{a}}$ is uniformized by $\mathcal{E}_{\mathfrak{a}} = \mathbb{C}/\mathfrak{a}$ via the modified Weierstrass functions $\widetilde{\wp}(z|\mathfrak{a})$, $\widetilde{\wp'}(z|\mathfrak{a})$. Consequently, given any point $z \in \mathcal{E}_{\mathfrak{a}}$ then the congruences*

$$x \equiv \widetilde{\wp}(z|\mathfrak{a})\,, \quad y \equiv \widetilde{\wp'}(z|\mathfrak{a}) \quad \bmod \mathfrak{P} \tag{50}$$

define a point $(x, y) \in \Gamma$.

Here, the points at infinity of $\Gamma_{\mathfrak{a}}$ and Γ have to be taken into account in the usual manner.

5.3.3 Lifting of rational points

The curve Γ is smooth and therefore, the number of K-rational points of Γ (including the one point at infinity) equals the number N of prime divisors of degree 1 of the function field $F|K$ of Γ. Hence in order to prove the second relation in (31) one has to count the number of K-rational points of Γ. For this purpose, Hasse proves that the K-rational points of Γ are lifted bijectively to the $\pi - 1$-division points of $\Gamma_{\mathfrak{a}}$. Let us first explain the notion of division point.

Let $0 \neq \mu \in \mathsf{M}$ be any multiplier $\mathfrak{a}$. The "μ-division points" of $\mathcal{E}_{\mathfrak{a}}$ are defined to be those points $z \in \mathcal{E}_{\mathfrak{a}} = \mathbb{C}/\mathfrak{a}$ which are annihilated by μ, i.e. $\mu z \equiv 0 \bmod \mathfrak{a}$. These points form a finite additive group $\frac{1}{\mu}\mathfrak{a}/\mathfrak{a}$ which is isomorphic to $\mathfrak{a}/\mu\mathfrak{a}$. To determine the order of this group, take a $\mathbb{Z}$-basis a_1, a_2 of $\mathfrak{a}$ and express $\mu a_1, \mu a_2$ by this basis. The determinant of the ensuing 2×2-matrix is the order of $\mathfrak{a}/\mu\mathfrak{a}$. On the other hand, this determinant equals the norm $\mathcal{N}(\mu)$ of μ within the imaginary quadratic field Ω.

Hasse considers the division points for $\mu = \pi - 1$. Thus $\mathcal{N}(\pi - 1)$ equals the number of those division points.

The $\pi - 1$-*division field* $\Omega_{\pi-1}$ of $\Gamma_{\mathfrak{a}}$ is defined by adjoining to Ω_{tm} the division values of the modified Weierstrass functions:

$$\Omega_{\pi-1} = \Omega_{tm}\left(\widetilde{\wp}(z|\mathfrak{a})\,, \widetilde{\wp}'(z|\mathfrak{a}) : \; z \in \frac{\mathfrak{a}}{\pi - 1}\right) \tag{51}$$

It is known, and Hasse shows it, that $\Omega_{\pi-1}$ is abelian over Ω, and it is a certain ray class field.

Hasse's **lifting theorem** says:

> *Let π be as in* (40)*. After replacing π by $-\pi$ if necessary, we have: The residue field of the $\pi - 1$-division field is K, i.e., $\Omega_{\pi-1}/\mathfrak{P} = K$. Consequently, for every division point $z \in \mathcal{E}_{\mathfrak{a}}$ the corresponding point $(x, y) \in \Gamma$ defined by* (50) *is K-rational. Conversely, every K-rational point $(x, y) \in \Gamma$ can be lifted to a $\pi - 1$-division point z , and this is unique. Hence the number N of K-rational points of Γ equals the number of $\pi - 1$-division points of $\mathcal{E}_{\mathfrak{a}}$, hence $N = \mathcal{N}(\pi - 1)$.*

This takes care of the second relation in (31) .

Since Γ is a good reduction of $\Gamma_{\mathfrak{a}}$ modulo $\mathfrak{P}$, today we would consider the statement of this lifting theorem almost self evident, provided the ± sign in front of π is chosen such that π induces in Γ the Frobenius endomorphism. But we should remember, once again, that in the year 1933 no such systematic theory of good reduction existed. In fact, the latter had been introduced by Deuring [Deu:1941] in order to algebraize the theory of complex multiplication by means of Hasse's algebraic theory of endomorphisms.

Hasse himself used in his proof the analytic and arithmetic properties of the elliptic functions (including Weber's τ-function). By looking through Hasse's manuscript [81] we have found quite an amount of detailed information (based on his former papers [H:1927] and [H:1931]) which he had to collect in order to come to the final conclusion. One of the main problems for him was to decide which of the two signs $\pm\pi$ has to be chosen. To this end, Hasse starts to consider both the $\pi - 1$- and the $\pi + 1$-division points (which are the same as the $(-\pi) - 1$-division points). He considers the ray class fields obtained by adjoining the values of $\widetilde{\wp}$ and $\widetilde{\wp}'$ for both these kinds of division points, and carefully studies the splitting of $\mathfrak{P}$ in these fields, using the decomposition laws of class field theory which he derives from the arithmetic properties of the so-called q-expansions (in the sense of classical theory of modular functions) of the elliptic functions involved.

Hasse sketches this method in his letter to Mordell which we cited in 1. of section 5.3, and he adds:

> "*... There is of course much detail which I could not give here, and my proof will be pretty long and difficult in print. On the other hand I am glad that this is so. For it shows that the theory of singular moduli and singular values of elliptic functions, and even the general Klassenkörpertheorie which I need for certain conclusions, are far from being very abstract castles built into the air but concrete enough to yield non-trivial results on rational numbers...*"

And then he adds:

[81] the one mentioned in 5. of section 5.3

> "*Perhaps my argument will become clearer to you if I expose it in the trivial case* $\quad y^2 \equiv x^2 - 1 \bmod p \quad \ldots$" (52)

Of course this congruence is not elliptic but of genus 0. This was precisely the reason why Hasse thought that this example was suitable for the explanation of the method because instead of class fields over imaginary quadratic fields, only cyclotomic fields over $\mathbb{Q}$ appear, but the line of proof is quite similar to the elliptic case.

The equation $y^2 = x^2 - 1$ is uniformized by

$$x = \cos 2\pi z\,, \quad y = i \sin 2\pi z\,.$$

with $z \in \mathbb{R}/\mathbb{Z}$ (the real numbers modulo 1). For any $n \in \mathbb{N}$, the n-th division points are given by those z for which $nz \equiv 0 \bmod \mathbb{Z}$. The division values are then given by

$$\cos\frac{2\pi\nu}{n}\,, \quad i\sin\frac{2\pi\nu}{n} \qquad (0 \le \nu < n)\,.$$

The n-th division field is obtained by adjoining these division values to $\mathbb{Q}$; this is precisely the n-th cyclotomic field $\mathbb{Q}(\sqrt[n]{1})$.

The prime p splits completely in the $p-1$-th cyclotomic field, let $\mathfrak{P}$ be one of its extensions [82]. Then the congruences

$$x_\nu \equiv \ \cos\frac{2\pi\nu}{p-1}\,, \qquad y_\nu \equiv \ i\sin\frac{2\pi\nu}{p-1} \qquad \bmod \mathfrak{P}$$

define $p-1$ solutions (x_ν, y_ν) of the given congruence in the prime field $\mathbb{F}_p$, and it is easy to verify that these are all different mod p. The analogue of Hasse's lifting theorem in this situation would be to show that *every* solution of the congruences

$$y^2 \equiv x^2 - 1 \bmod p \qquad \text{with} \quad x, y \in \mathbb{F}_p \tag{53}$$

is obtained in this way. (In this context there are no solutions at ∞ to be taken into consideration.) This is not difficult to verify but Hasse wishes to present this fact in a similar way which he had used in his proof of the above lifting theorem for elliptic congruences. Thus he compares the $p-1$-division values with the $p+1$-division values. He writes:

> "*For any n, the division values* $\cos\frac{2\pi\nu}{n}$ $(\nu = 0, 1, \ldots n-1)$ *generate the real subfield of the n^{th} cyclotomic field. If n divides $p \pm 1$ then p splits into prime divisors of degree* 1 *in that field. Take the equation with all those $p-1+p+1 = 2p$ division values as roots. By a very easy calculation one finds its left hand side* $\equiv (x^p - x)^2 \bmod p$. *This means that those division values form twice a complete set of residues modulo* $\mathfrak{P}$."

[82] $\mathfrak{P}$ can be extended to the algebraic closure of $\mathbb{Q}$. We will denote this extension also by $\mathfrak{P}$.

Consequently, if (x,y) is a solution in $\mathbb{F}_p$ of the congruence (53) then $x \equiv \cos\frac{2\pi\nu}{p-1} \bmod \mathfrak{P}$ for some ν, or $x \equiv \cos\frac{2\pi\mu}{p+1} \bmod \mathfrak{P}$ for some μ. Suppose the latter. Then we have $y \equiv i\sin\frac{2\pi\mu}{p+1} \bmod \mathfrak{P}$ (after replacing μ by $-\mu$ if necessary). Observe that $\mathbb{Q}(\cos\frac{2\pi\mu}{p+1}, i\sin\frac{2\pi\mu}{p+1})$ is the n-th cyclotomic field for $n = \frac{p+1}{\gcd(\mu,p+1)}$. Also observe that if $n>1$ then p does *not* split completely in this cyclotomic field, and accordingly $i\sin\frac{2\pi\mu}{p+1} \bmod \mathfrak{P}$ is *not* in the prime field $\mathbb{F}_p$. Hence necessarily $n=0$ or $n=1$ and therefore $y \equiv 0 \bmod \mathfrak{P}$ which implies $x \equiv \pm 1 \bmod \mathfrak{P}$. Thus $(x,y) = (x_\nu, y_\nu)$ with $\nu = 0$ or $\nu = \frac{p-1}{2}$.

It seems that Hasse wrote this not only to Mordell but also to Davenport. For the latter replied, under the date of March 17, 1933, the following:

> "*. . . I re-read your letter in which you explained your method the other day, and can now follow it more or less in so far as it relates to* $y^2 \equiv x^2 - 1$. *But I do not see how you* <u>*discovered*</u> *the fact about* $\cos(\frac{\nu}{p-1})$, $\cos(\frac{\mu}{p+1})$. *What is the connection between the solutions as they arise in your method, and the parametric solution* $x \equiv \frac{1}{2}(t+t^{-1})$, $y \equiv \frac{1}{2}(t-t^{-1})$ *? . . .*"

Here, the "fact" about those division values is that p splits completely in the field generated by those values, i.e. in the real parts of the cyclotomic fields $\mathbb{Q}(\sqrt[n]{1})$ for $n \mid p \pm 1$. Davenport's question how Hasse "discovered" that fact is, it seems to us, not quite to the point. Hasse did not "discover" this fact for the cyclotomic fields; this was well known. Of course, Davenport's suggestion to use directly the parametric solution makes the arguments much easier. But Hasse wished to *explain the main ideas* of his method by transferring his arguments from complex multiplication to the case of cyclotomy – in the hope that in this way also those people who were not too familiar with complex multiplication, would be able to appreciate his ideas. Hasse was able to "discover" that fact *in the complex multiplication case* because, as we mentioned above already, he was quite familiar with complex multiplication.

At this point let us compare Hasse's method with the method of Herglotz [Her:1921] who had discussed the lemniscate congruence

$$x^2 + x^2y^2 + y^2 \equiv 1 \bmod p \qquad \text{for} \quad p \equiv 1 \bmod 4\,. \tag{54}$$

(See Part 1, section 3.) In this case the invariant is $j \equiv 12^3 \bmod p$ which Hasse had excluded, nevertheless it may be worthwhile to compare Hasse's with Herglotz' method. The lifting of the invariant yields $j(\mathfrak{a}) = 12^3$ with multiplier ring $\mathsf{M} = \mathbb{Z}[i]$, and one can take $\mathfrak{a} = \mathsf{M}$ since M is a principal ideal ring. The uniformization of the lemniscate equation which Herglotz used was given by the lemniscate functions

$$x = \sin\,\mathrm{lemn}(u)\,, \qquad y = \cos\,\mathrm{lemn}(u)$$

instead of Hasse's $\widetilde{\wp}$, $\widetilde{\wp}'$. Herglotz showed that

> "*the solutions of* (54) *coincide precisely with the congruence solutions modulo π of the division equation, by $\pi - 1$, for the lemniscate functions.*"

Here, π denotes a prime number of M dividing p. We see that this is precisely the lifting theorem of Hasse in the case of the lemniscate congruence. The problem which one of the $\pm\pi$, $\pm i\pi$ to take did not occur since Gauss had already stated the normalization condition: $\pi \equiv 1 \bmod (1-i)^3$. [83] The use of class field theory is hidden in Herglotz' paper because he explicitly discusses the division equation which, when taken modulo p, shows directly that multiplication by π acts as Frobenius operator on the division values, provided π is taken in the normalization as prescribed by Gauss. Hence the prime ideal $\mathfrak{p} \cong \pi$ splits completely in the field generated by the division values of the lemniscate functions.

Thus Hasse's proof of his lifting theorem can be regarded as a direct generalization and adaption of Herglotz' proof for the lemniscate.

As we have already said in Part 1, all the available evidence points towards the fact that Hasse did not know Herglotz' paper, nor did Davenport or Mordell.

5.3.4 Normalization

Davenport in his letter of March 17, 1933 asks Hasse:

> "*What do you think the form of the ordinates of the zeros of Artin's ζ–function will be ? ... There must be an enormous amount of ingenuity in your method when it comes to $y^2 \equiv f_3(x)$. Best wishes for its success...*"

We do not know Hasse's reply to this question. But we can imagine that in Hasse's opinion this question was not well put. Hasse wished to have the zeros s of the zeta function, or rather the numbers $\omega = q^s$, characterized not by their rectangular or polar coordinates but by their *arithmetic structure*. This means, firstly, the prime decomposition of ω. In addition there may be other properties, e.g., congruence conditions which serve to determine the ω completely.

In the elliptic case (with the technical restrictions as mentioned) Hasse's result was that $\omega = \pi$ with $\pi \cong \mathfrak{p}^r$ where $\mathfrak{p}$ is a prime divisor of p in Ω, of degree 1, and where r is the degree of K over $\mathbb{F}_p$. See (40), and the statement (i) of the

[83] Note that in the case of the lemniscate, the multiplier ring is $\mathsf{M} = \mathbb{Z}[i]$ which contains four units, namely ± 1, $\pm i$.

invariant-lifting lemma. This takes care of the prime decomposition of ω. But the prime decomposition determines ω up to a unit only, i.e., up to a factor ± 1. [84]

Therefore, Hasse [H:1933] looked for a normalization condition which determines π completely, with no uncertainty which of the $\pm\pi$ to take – similarly as in the case of the lemniscate where the normalization was given by Gauss, namely $\pi \equiv 1 \bmod (1-i)^3$.

Let $D < 0$ denote the discriminant of the imaginary quadratic field Ω. Then Hasse's normalization condition works only if $q \equiv 1 \bmod 4$ and it reads:

$$\begin{aligned} \pi &\equiv 1 \bmod 4 \qquad \text{if} \quad \left(\frac{D}{2}\right) = +1 \\ \pi^3 &\equiv 1 \bmod 4 \qquad \text{if} \quad \left(\frac{D}{2}\right) = -1 \end{aligned}$$

where $2 \nmid D$ and $\left(\frac{D}{2}\right)$ is the quadratic residue symbol in Ω. There is also some condition if $\left(\frac{D}{2}\right) = 0$, i.e., if 2 is ramified in Ω. But this is rather technical and so we abstain from giving it here.

Hasse says: "*In the case* $q \equiv -1 \bmod 4$ *the normalization seems to be more difficult.*" As to our knowledge, the problem is not solved up to this day.

5.4 The case of arbitrary genus

Already in Hasse's preliminary announcement [H:1933] Hasse says he expects the Riemann hypothesis to be true also for "general binary congruences", which is to say for function fields of arbitrary genus. Compare the date of the announcement, namely April 28, 1933, with the date of Hasse's Hamburg colloquium lecture which was end of November 1932. At the latter colloquium Hasse was not yet convinced about the general validity of the Riemann hypothesis – contrary to Artin who at that occasion had voiced his strong opinion in favor of it. (We know this from Iyanaga's report about the Hamburg colloquium; see section 3.4.)

Thus it took less than 5 months for Hasse to change his outlook concerning the Riemann hypothesis. It seems that Artin's comment in Hamburg had been instrumental in this, for it induced Hasse to look at the problem from a different point of view. Certainly, Hasse's success in the elliptic case made him more confident concerning the general case.

[84] Recall that Hasse had excluded the cases $j = 0$ and $j = 12^3$ which implies that $j(\mathfrak{a}) \neq 0, 12^3$ in $\mathbb{C}$. Therefore, as is well known from complex multiplication, the multiplier ring M of $\mathfrak{a}$ contains ± 1 as the only units.

In the preliminary announcement [H:1933] Hasse points out the formal analogy of his method, to the method of Siegel [Si:1929] developed in the theory of diophantine approximations. (This he had said already in his letter to Mordell of March 6, 1933; see section 5.1). Now Hasse mentions

> "*. . . Siegel's principle that algebraic relations between values of power series are in general based on algebraic relations between the corresponding functions*"

and says that this corresponds to his (Hasse's) principle

> "*that solutions of a congruence are based on a corresponding equation between algebraic numbers.*"

Hasse continues:

> "*And like Siegel has solved the problem of finiteness of solutions of a general binary diophantine equation by uniformizing with general abelian functions, it is to be expected that similarly the problem of number of solutions for general diophantine congruences can be solved by uniformizing with general abelian functions – and that in this way the analogue of the Riemann hypothesis can be solved for the general zeta functions of F. K. Schmidt.*"

We see that even in this early stage, Hasse expects that abelian functions have to be used in the case of higher genus, in place of elliptic functions. But it seems that he did not yet have detailed ideas how to achieve this aim, and that he was not quite sure whether he would be able to do it. Anyway, in his letter to Mordell of March 6, 1933 which we have already cited several times, and in which he explained the main ideas of his proof by elliptic functions, he added a postscript as follows.

> "*P. S. Obviously the general congruence $f(x, y)$ may be treated the same way, "only" with the "slight" generalisation of the elliptic functions into abelian functions quite analogous to Siegel! Do it!* "

Thus Hasse tried to transfer the problem to Mordell. The latter, however, did not keep the ball but threw it back. In his reply on March 9, 1933 Mordell too put a postscript, and this read:

> "*P. S. I think the results for $y^2 \equiv f_n(x)$ etc. should follow without infinite difficulty, but the zeta fn. theory will not be so simple now. Obviously you are now the one to try it.*"

Today we know that the later development went a somewhat different course, because the use of analytic abelian functions in characteristic 0 can be avoided by working directly in characteristic $p > 0$. But in some way Hasse's idea was indeed used, in as much as the algebraic theory of abelian function fields (or abelian varieties) in characteristic p have played an important role in the context of the Riemann hypothesis. We shall report this in more detail in one of the future Parts.

At the end of his preliminary announcement [H:1933] Hasse himself says that the use of analytic functions (which would require to extend the theory of complex multiplication) may perhaps be replaced by algebraic-arithmetic methods. For those he cites the thesis of André Weil [W:1928]. [85] We shall see in Parts 3 and 4 that indeed Weil was to play an important role in the further development of the Riemann hypothesis for curves. But already in 1933 Weil seems to have become interested in Hasse's work. We conclude this from a remark in a letter of Hasse to Davenport of July 24, 1933. There, Hasse informed Davenport that Weil had come over from Frankfurt for a day. We can imagine that Weil, who visited Siegel in Frankfurt, had heard about Hasse's ideas and thus came over to Marburg for a day in order to get details.

One year later, on June 18, 1934, Weil wrote to Hasse as follows: [86]

> "*I have again thought about your problem. According to my experiences in this subject it seems not to be expected that one can perform useful computations with abelian functions if one does not have theta functions at one's disposal…In my opinion there is no choice other than to operate with theta functions in the usual way, by defining them algebraically. For, the theta function is nothing else than a divisor on the Jacobian variety…*" [87]

We do not know whether Hasse and Weil had met in the meantime and discussed the problem of higher genus further, or whether this letter was referring to their discussion in July 1933. Nor do we know the reply of Hasse to Weil's letter. We have cited Weil's letter here in order to point out that as early as 1933/34 Hasse had the idea of an algebraic theory of abelian functions, and

[85] It seems that he had in mind Weil's theory of "distributions" which Weil had developed and used in his thesis.

[86] I am indebted to Günther Frei for granting me access to the correspondence file Hasse-Weil. – The cited letter seems to be the first one with mathematical content. There is one earlier letter, dated Aug 4, 1931, where Weil informs Hasse about the tragic death of Jacques Herbrand who some days earlier had an accident in the Alps. Herbrand had spent some time in Germany with a Rockefeller grant, and he had close contacts (among others) to the people working on the foundation of class field theory, around Artin, Hasse and Emmy Noether.

[87] Weil wrote in German, we have translated it into English. – Weil was able to write and speak German perfectly, without any accent. Hasse in a later letter called him a "*Sprachgenie*".

that he had discussed it with André Weil. For a more detailed discussion of the developments in the following years we refer to the forthcoming Parts 3 and 4 of our report.

Let us close this section with a citation from a letter of Emmy Noether to Hasse, dated March 3, 1933 [88]:

> "*First of all my congratulation to the "Riemann hypothesis". You have done unbelievably many things lately! I assume that now you will be able to get at the general Artin-Schmidt zeta function since you already use general class field theory...*"

When she mentions the "general Artin-Schmidt zeta function" then of course she means the zeta function for function fields of arbitrary genus.

5.5 Summary

At the end of February 1933, during or shortly after his seminar lectures in Marburg, Hasse obtained his first proof of the Riemann hypothesis for elliptic function fields.

The proof proceeded by lifting an elliptic curve which is defined over a finite field K, to a suitable elliptic curve which is defined over an algebraic number field and admits complex multiplication. This lifting process turns out to be essentially unique (up to some exceptional cases called "supersingular"). Let Ω denote the imaginary quadratic field containing the complex multipliers of the lifted curve. The order q of the finite base field K splits in Ω into a product $q = \pi\pi'$ of conjugate elements which, after suitable normalization by unit factors, turn out to be the inverse zeros of the zeta function of the given elliptic curve over K (if the zeta function is considered as rational function of the variable $t = q^{-s}$). To show this, the decomposition laws of class field theory are applied to the $\pi - 1$-division field over Ω. By the reduction theory of curves (which however was not yet established at that time) these decomposition laws imply that after reduction, π is the Frobenius endomorphism of the original elliptic curve.

Hasse's first proof was never published but preliminary announcements appeared. From several documents of Hasse's legacy, including his lecture notes for the DMV-congress in Würzburg (Sep 1933) the proof can be recovered in all details. It is, however, not quite complete since for technical reasons Hasse had to exclude several cases which he could not cover in the first attempt. While

[88] This was a few weeks before Emmy Noether was dismissed as a university professor and was forced to emigrate, due to the antisemitic policy of the Nazi regime. The friendly correspondence between Hasse and Emmy Noether continued until her untimely death in Bryn Mawr, 1935.

writing up the manuscript for publication he tried to eliminate those restrictions and during this activity he found a simpler proof, valid quite generally, and working without the detour over characteristic zero and class field theory. The final published proof will be discussed in one of the following Parts.

Starting from about February or March 1933, Hasse became convinced that the Riemann hypothesis holds for function fields of arbitrary genus over finite base fields. Already in his first preliminary announcement he said that the theory of abelian functions should give the result in a similar way as the theory of elliptic functions did in the case of genus 1 *. He also mentioned the methods of A. Weil's thesis as possibly useful for this aim, and he met Weil to discuss this idea.*

Additional Comment:

This first proof of Hasse is of historical interest not only because it shows us the genesis of the ideas leading to the proof of the Riemann hypothesis in characteristic p . It also carries the first instance of a non-trivial application of good reduction. Later this was systematically developed by Deuring who, with this tool, established an algebraic theory of the class fields of complex multiplication in characteristic 0*, without using complex analysis. Moreover, Hasse's lifting lemmas became the basis for Deuring's later results on the so-called Hasse-Weil zeta functions of elliptic curves defined over number fields. It is no surprise that Deuring's theory works for CM-curves, i.e., curves with complex multiplication. Here in Hasse's first proof we see some of the main ideas of that further development in a nutshell already.*

References

[A:1921] E. Artin, *Quadratische Körper im Gebiete der höheren Kongruenzen.* Dissertationsauszug. Jahrb. phil. Fak. Leipzig 1921, II. Halbjahr (1921) 157-165 3

[A:1924] —— *Quadratische Körper im Gebiete der höheren Kongruenzen I, II.* Math. Zeitschr. 19 (1924) 153–246 21, 24, 39, 65

[A:1927] —— *Über die Zerlegung definiter Funktionen in Quadrate.* Abh. Math. Sem. Univ. Hamburg 5 (1927) 100–115 7

[A-H:1925] E. Artin, H. Hasse, *Über den zweiten Ergänzungssatz zum Reziprozitätsgesetz der ℓ-ten Potenzreste im Körper k_ℓ der ℓ-ten Einheitswurzeln und in Oberkörpern von k_ℓ.* J. Reine Angew. Math. 154 (1925) 199–214 6

[A-Wh:1945] E. Artin, G. Whaples, *Axiomatic characterization of fields by the product formula for valuations.* Bull. Amer. Math. Soc. 51 (1945) 469–492 2

[BCHIS:1966] A. Borel, S. Chowla, C.S. Herz, K. Iwasawa, J.P. Serre, *Seminar on complex multiplication 1957-1958.* Springer Lecture Notes No. 21 (1966) 51

[BHN:1932] R. Brauer, H. Hasse, E. Noether, *Beweis eines Hauptsatzes in der Theorie der Algebren.* J. Reine Angew. Math. 167 (1932) 399–404 8, 30

[Ca:1974] J. W. S. Cassels, *L. J. Mordell.* Bulletin of the London Math. Soc. 6 (1974) 69-96 9

[Co:1978] H. Cohn, *A classical invitation to algebraic numbers and class fields.* Universitext, Springer 1978 22

[Da:1930] H. Davenport, *On the distribution of quadratic residues (mod p).* J. London Math. Soc 6 (1931) 49–54 14, 18

[Da:1932] —– *On the distribution of the ℓ-th power residues (mod p).* J. London Math. Soc., 7 (1932) 117–121 19

[Da:1933] —– *On certain exponential sums.* J. Reine Angew. Math. 169 (1933) 158–176 22

[Da:1933a] —– *On the distribution of quadratic residues (mod p). (Second paper.)* J. London Math. Soc., 8 (1933) 46–52 19

[Da:1977] —– *The Collected Works I-IV.* Ed. by B. J. Birch, H. Halberstam, C. A. Rogers. Academic Press 1977 13, 18

[Da-H:1934] H. Davenport, H.Hasse, *Die Nullstellen der Kongruenzzetafunktionen in gewissen zyklischen Fällen.* J. Reine Angew. Math. 172 (1934) 151–182 49

[Deu:1941] M. Deuring, *Die Typen der Multiplikatorenringe elliptischer Funktionenkörper*. Abh. Math. Sem. Univ. Hamburg 14 (1941) 197–292 45, 52, 56

[Deu:1949] —– *Algebraische Begruendung der komplexen Multiplikation.* Abh. Math. Sem. Univ. Hamburg 16 (1949) 32–47 51

[Deu:1958] —– *Die Klassenkörper der komplexen Multiplikation.* Enzyklopädie d. Math. Wissenschaften, 2. Aufl. Band I, 2.. Teil, Heft 10, Art. 23 (Teubner 1958) 51

[Fa:1983] G. Faltings, *Endlichkeitssätze für abelsche Varietäten über Zahlkörpern.* Invent. Math. 73 (1983) 349–366 9

[Fr:1977] G. Frei, *Leben und Werk von Helmut Hasse. 1. Teil: Der Lebensgang.* Collection Mathématique, Dépt. de Mathématiques, Université Laval, No. 28 (1977) 13

[Fr:1985] —– , *Helmut Hasse (1898-1979). A biographical sketch dealing with Hasse's fundamental contributions to mathematics, with explicit references to the relevant mathematical literature.* Expositiones Math. 3 (1985) 55-69 8

[Fr:1985a] —– , *Der Briefwechsel David Hilbert - Felix Klein (1986–1918).* Arbeiten aus der Niedersächsischen Staats- und Universitätsbibliothek Göttingen, Band 19 (1985) 10

[Fr-Rq:2002] G. Frei, P. Roquette, *Helmut Hasse in Halle.* In: Manfred Goebel et al. (Ed), Aspekte der Mathematikgeschichte in Halle. Fachbereich Mathematik und Informatik der Martin-Luther-Universität Halle-Wittenberg, Report Nr.19 (2002) 83–98 8

[Fr-St:1999] G. Frei, U. Stammbach, *Heinz Hopf.* In: History of Topology, ed. by I.M. James. Elsevier Amsterdam (1999) 991-1008 16

[H:1924] H. Hasse, *Darstellbarkeit von Zahlen durch quadratische Formen in einem beliebigen algebraischen Zahlkörper.* J. Reine Angew. Math. 153 (1924) 113–130 7

[H:1924a] —– *Review of Artin's paper* [A:1924]. Jahrbuch über die Fortschritte der Mathematik vol. 50 (1924) 3

[H:1926] —– *Bericht über neuere Untersuchungen und Probleme aus der Theorie der algebraischen Zahlkörper. I. Klassenkörpertheorie*. Jber. Deutsch. Math. Verein. 35 (1926) 1–55 8, 51

[H:1927] —– *Neue Begründung der komplexen Multiplikation I. Einordnung in die allgemeine Klassenkörpertheorie.* J. Reine Angew. Math. 157 (1927) 115–139 8, 51, 56

[H:1927a] —— *Bericht über neuere Untersuchungen und Probleme aus der Theorie der algebraischen Zahlkörper. Ia. Beweise zu I.* Jber. Deutsch. Math. Verein. 36 (1927) 233–311 8

[H:1930] —— *Bericht über neuere Untersuchungen und Probleme aus der Theorie der algebraischen Zahlkörper. II. Reziprozitätsgesetze.* Jber. Deutsch. Math. Verein. Ergänzungsband 6 (1930) 1–204 8

[H:1930a] —— *Die moderne algebraische Methode.* Jber. Deutsch. Math. Verein. 39 (1930) 22–34. English translation in: The Mathematical Intelligencer 8 (1986) 18–25 13

[H:1931] —— *Neue Begründung der komplexen Multiplikation II. Aufbau ohne die Benutzung der allgemeinen Klassenkörpertheorie.* J. Reine Angew. Math. 165 (1931) 64–88 8, 51, 56

[H:1931a] —— *Zum Hauptidealsatz der komplexen Multiplikation.* Monatshefte f. Math. 38 (1931) 315-322 51

[H:1931b] —— *Ein Satz über die Ringklassenkörper der komplexen Multiplikation.* Monatshefte f. Math. 38 (1931) 323-330 51

[H:1931c] —— *Das Zerlegungsgesetz für die Teiler des Moduls in den Ringklassenkörpern der komplexen Multiplikation.* Monatshefte f. Math. 38 (1931) 331-344 51

[H:1932] —— *Vorlesungen über Klassenkörpertheorie.* Mimeographed Notes, Marburg 1932. Reprinted as a book: Physica Verlag, Würzburg 1967 43

[H:1932b] —— *Theory of cyclic algebras over an algebraic number field.* Transactions of the American Mathematical Society 34 (1932) 171–214 10

[H:1933] —— *Beweis des Analogons der Riemannschen Vermutung für die Artinschen und F.K.Schmidtschen Kongruenzzetafunktionen in gewissen elliptischen Fällen. Vorläufige Mitteilung.* Nachr. Ges. Wiss. Göttingen I. Math.-Phys. Kl. Fachgr. I Math. Nr.42 (1933) 253–262 37, 46, 52, 60, 61, 62

[H:1934] —— *Über die Kongruenzzetafunktionen. Unter Benutzung von Mitteilungen von Prof. Dr. F.K. Schmidt und Prof. Dr. E. Artin.* S.-Ber. Preuß. Akad. Wiss. H. 17 (1934) 250–263 18, 32, 34, 39, 40

[H:1934a] —— *Abstrakte Begründung der komplexen Multiplikation und Riemannsche Vermutung in Funktionenkörpern.* Abh. Math. Sem. Univ. Hamburg 10 (1934) 250–263 47

[H:1934b] —— *Riemannsche Vermutung bei den F.K. Schmidtschen Kongruenzzetafunktionen.* Jber. Deutsch. Math. Verein. 44 (1934) 44 (kursiv) 47

[H:1934c] —— *Theorie der relativ-zyklischen algebraischen Funktionenkörper, insbesondere bei endlichem Konstantenkörper.* J. Reine Angew. Math. 172 (1934) 37–54

[H:1934d] —— *Existenz separabler zyklischer unverzweigter Erweiterungskörper vom Primzahlgrad p über elliptischen Funktionenkörpern der Charakteristik p .* J. Reine Angew. Math. 172 (1934) 77–85 45, 52

[H:1935] —— *Zur Theorie der abstrakten elliptischen Funktionenkörper.* Nachr. Ges. Wiss. Göttingen I. Math.-Phys. Kl. Fachgr. I Math. Neue Folge Band 1 Nr.7 (1935) 119–129

[H:1936a] —— *Zur Theorie der abstrakten elliptischen Funktionenkörper. I. Die Struktur der Gruppe der Divisorenklassen endlicher Ordnung.* J. Reine Angew. Math. 175 (1936) 55–62 47

[H:1936b] —— *Zur Theorie der abstrakten elliptischen Funktionenkörper. II. Automorphismen und Meromorphismen. Das Additionstheorem.* J. Reine Angew. Math. 175 (1936) 69–88 47

[H:1936c] —— *Zur Theorie der abstrakten elliptischen Funktionenkörper. III. Die Struktur des Meromorphismenrings; die Riemannsche Vermutung.* J. Reine Angew. Math. 175 (1936) 193–208 45, 47, 52

[H:1950] —— *Vorlesungen über Zahlentheorie.* Springer Verlag (1950) 16

[H:1966] —— *Modular functions and elliptic curves over finite fields.* Rend. di Mat. Appl. V ser. 25 (1966) 248–266 47, 48, 52

[H:1975] —— *Mathematische Abhandlungen.* Ed. by H. W. Leopoldt and P. Roquette. (Berlin 1975), Bd.1–3

[H:2002] —— *Number theory.* Transl. from the 3rd German edition, edited and with a preface by Horst Günter Zimmer. Reprint of the 1980 edition. Classics in Mathematics. Berlin: Springer (2002) 2

[Her:1921] G. Herglotz, *Zur letzten Eintragung im Gaußschen Tagebuch*. Ber. Verhandl. Sächs. Akad. Wiss. Math.-Phys. Kl. 73 (1921) 271–276 41, 58

[Hil:1900] D. Hilbert, *Mathematische Probleme. Vortrag, gehalten auf dem internationalen Mathematiker-Congress zu Paris 1900.* Nachr. Ges. Wiss. Göttingen 1900, 253-297.

[Ho:1930] H. Hopf, Über die Verteilung quadratischer Reste. Mathematische Zeitschrift 32 (1930) 222–231 16

[Ja:1906] E. Jacobsthal, *Anwendungen einer Formel aus der Theorie der quadratischen Reste.* Dissertation Berlin 1906. 14

[Ja:1910] —— *Über die Darstellung der Primzahlen der Form* $4n + 1$ *als Summe zweier Quadrate.* J. Reine Angew. Math. 137 (1910) 167–309 14

[Ka:1979] E. Kani, *Nonstandard diophantine geometry.* Proc. Queen's Number Theory Conf. 1979, Queen's Pap. Pure Appl. Math. 54 (1980) 129-172

[Lib:1832] G. Libri, *Mémoire sur la théorie des nombres.* J. Reine Angew. Math. 9 (1832) 54–80, 169–188, 261–276 21

[Mor:1922] L.J. Mordell, *On the rational solutions of the indeterminate equations of the 3rd and 4th degrees.* Proc. Cambridge Phil. Soc. 21 (1922) 179–192 8, 9, 43

[Mor:1932] —— *On a sum analogous to a Gauss sum.* Quart. J. Math., Oxford Ser. 3, 161-167 (1932). 22

[Mor:1933] —— *The number of solutions of some congruences in two variables.* Math. Zeitschr. 37 (1933) 193–209 18, 22, 23, 31

[Mor:1969] —— *Diophantine Equations*. Academic Press (1969) 312 pp. 9

[Mor:1971] —— *Harold Davenport (1907–1969).* Acta Arithmetica 18 (1971) 1–4 13, 14

[Mor:1971a] —— *Some aspects of Davenport's work.* Acta Arithmetica 18 (1971) 5–11 16, 24

[Os:1919] A. Ostrowski, *Zur arithmetischen Theorie der algebraischen Grössen.* Nachr. Ges. Wiss. Göttingen 1919, 279–298 24

[Rog:1972] C. A. Rogers, *Harold Davenport.* Bull. London Mathematical Society 4 (1972) 66–99 12, 13, 16

[Rq:2001] P. Roquette,*Class field theory in characteristic p, its origin and development.* In: K. Miyake (ed.) Class Field Theory – Its Centenary and Prospect. Advanced Studies in Pure Mathematics 30 (2001) 549–631 3

[Rq:2002] —— *History of Valuation Theory, Part I.* In: Fields Institute communication series vol. 32 (2002) 66pp.

[Rq:2002a] ——, *The Riemann hypothesis in characteristic p, its origin and development. Part 1. The formation of the zeta functions of Artin and F. K. Schmidt.* Mitteilungen der Mathematischen Gesellschaft in Hamburg, Band XXI/2 (2002) 79–157 2, 6, 28

[Sal:1931] H. Salié, *Über die Kloostermanschen Summen* $S(u, v; q)$. Math. Zeitschr. 34 (1931) 91–103 22

[FK:1926] F. K. Schmidt, *Zur Zahlentheorie in Körpern der Charakteristik p. (Vorläufige Mitteilung.)* Sitz.-Ber. phys. med. Soz. Erlangen 58/59 (1926/27) 159–172 3, 34

[FK:1931] ——, *Analytische Zahlentheorie in Körpern der Charakteristik* p. Math. Zeitschr. 33 (1931) 1–32 39

[Sev:1927] F. Severi, *Trattato die geometria algebrica. Vol. I, 1: Geometria delle serie lineari.* (Italian) Bologna (1926) 5

[Sh:1957] I. R. Shafarevich, *Exponents of elliptic curves.* (Russian) Dokl. Akad. Nauk SSSR 114 (1957) 714-716 4

[Shi:1967] K. Shiratani, *Über singuläre Invarianten elliptischer Funktionenkörper.* J. Reine Angew. Math. 226 (1967) 108-115 48, 52

[Si:1929] C.L. Siegel, *Über einige Anwendungen diophantischer Approximationen.* Abh. Preuß. Akad. Wissensch. Phys. Math. Kl. Nr.1 (1929) 43, 61

[Web:1908] H. Weber, *Lehrbuch der Algebra.* 2nd ed. vol. 3. Vieweg & Sohn. Braunschweig (1908) 51

[W:1928] A. Weil, *L'arithmétique sur les courbes algébriques.* Acta Math. 52 (1928) 281–315 8, 62

[Weyl:1944] H. Weyl, *David Hilbert and his mathematical work.* Bulletin of the American Math. Soc. 50 (1944) 612-654 7

[Yan:2002] B. H. Yandell, *The honors class. Hilbert's problems and their solvers.* Natick, MA: A K Peters. 486 p. (2002) 7

Eingegangen am 27.09.03

Peter Roquette
Mathematisches Institut
Universität Heidelberg
D-69120 Heidelberg, Germany
e-mail: roquette@uni-hd.de