The Riemann Hypothesis in Characteristic p in Historical Perspective

Peter Roquette

Version of 3.7.2018

Preface

This book is the result of many years' work. I am telling the story of the Riemann hypothesis for function fields, or curves, of characteristic p starting with Artin's thesis in the year 1921, covering Hasse's work in the 1930s on elliptic fields and more, until Weil's final proof in 1948. The main sources are letters which were exchanged among the protagonists during that time which I found in various archives, mostly in the University Library in Göttingen but also at other places in the world. I am trying to show how the ideas formed, and how the proper notions and proofs were found. This is a good illustration, fortunately well documented, of how mathematics develops in general.

Some of the chapters have already been pre-published in the "*Mitteilungen der Mathematischen Gesellschaft in Hamburg*". But before including them into this book they have been thoroughly reworked, extended and polished. I have written this book for mathematicians but essentially it does not require any special knowledge of particular mathematical fields. I have tried to explain whatever is needed in the particular situation, even if this may seem to be superfluous to the specialist.

Every chapter is written such that it can be read independently of the other chapters. Sometimes this entails a repetition of information which has already been given in another chapter. The chapters are accompanied by "summaries". Perhaps it may be expedient first to look at these summaries in order to get an overview of what is going on.

The preparation of this book had been supported in part by the *Deutsche Forschungsgemeinschaft* and the *Marga und Kurt Möllgaard Stiftung*. A number of colleagues and friends have shown interest and helped me with

their critical comments; it is impossible for me to mention all of them here. Nevertheless I would like to express my special thanks to Sigrid Böge, Karin Reich, Franz-Viktor Kuhlmann and Franz Lemmermeyer.

References: Most of the letters and documents cited in this book are contained in the *Handschriftenabteilung* of the University Library of Göttingen (except if another source is mentioned), most but not all of them in the Hasse-files. The letters from Hasse to Davenport are contained in the archive of Trinity College, Cambridge. The letters from Hasse to Mordell are contained in the archive of King's College, Cambridge. The letters from Hasse to Fraenkel are in the archive of the Hebrew University, Jerusalem.

I have translated the cited text into English for easier reading. Transcriptions of the full letters (and more) in their original language are available at my homepage:

https://www.mathi.uni-heidelberg.de/~roquette/

Contents

Chapter 1

Overture

1.1 Why history of mathematics ?

Mathematics is, on the one hand, a *cumulative science*. Once a mathematical theorem has been proved to be true then it remains true forever: it is added to the stock of mathematical discoveries which has piled up through the centuries and it can be used to proceed still further in our pursuit of knowledge.

On the other hand, the mere proof of validity of a theorem is in general not satisfactory to mathematicians. We also wish to know "why" the theorem is true, we strive to gain a better understanding of the situation than was possible for previous generations. Consequently, although a mathematical theorem never changes its content we can observe, in the history of our science, a continuous change of the *form of its presentation*. Sometimes a result appears to be better understood if it is generalized and freed from unnecessary assumptions, or if it is embedded into a general theory which opens analogies to other fields of mathematics. Also, in order to make further progress possible it is often convenient and sometimes necessary to develop a framework, conceptual and notational, in which the known results become trivial and almost self-evident, at least from a formalistic point of view. So when we look at the history of mathematics we indeed observe changes, not in the nature of mathematical truth, but in the attitude of mathematicians towards it. It may well be that sometimes a new proof is but a response to a current fashion, and sometimes it may be mere fun to derive a result by unconventional means. But mostly the changes in attitude reflect a serious effort towards a better understanding of the mathematical universe.

I believe it is worthwhile to observe such trends in the past and see how they have led to the picture of today's mathematics. Here I am telling the fascinating story of the emergence of the Riemann Hypothesis in characteristic p and of its proof. Initiated by Artin in analogy to algebraic number theory, further developed by Hasse and Deuring in the framework of function fields, and later embedded into the new algebraic geometry by A. Weil, this development exhibits all the features of mathematical research mentioned above.

Our story covers roughly the years from 1921 to 1948 (with an appendix devoted to Bombieri's proof in 1976). In this period Hasse and his team did not yet reach a proof for arbitrary function fields of higher genus. But all the prerequisites of a proof had been available already in the year 1936. In order to provide evidence for this I have included a chapter with a short "virtual" proof which, indeed, could have been given in 1936 by Hasse or Deuring or any one of their team.

This story is a good example of what we often observe in the history of mathematics:[1] that much effort and time had to be spent by the pioneers to explore their way into unknown territory. Thereby they paved the way for the next generations who now can travel comfortably along their smooth tracks. I have extensively used original letters and documents since this provides a glimpse into the communication channels of mathematicians of that time. One can see the difficulties which had to be overcome although nowadays they are not any more considered difficult at all. This book may be regarded as just a commentary to those letters which had been exchanged at that time.

I have met in my life most of the people from whose letters I am citing but I have to admit that in the past I did not ask them about their own opinions and remembrances of those times. Today I regret that I missed the opportunity to get to know more about the personal memories of these people. Nevertheless, my interpretation of their letters and notes may show

[1] And not only of mathematics.

a certain personal touch in as much as it could possibly depend on my own impressions of their personalities.

1.2 Artin and the intervention by Hilbert

Our story starts in Göttingen in the early 1920s. The mathematical scene in Göttingen had recovered from the difficult years of World War I and the immediate post-war period. Göttingen was again considered as "the" center of mathematics in the world. It was to become an attractive place in particular for young mathematicians from Germany and from abroad, who were eager to learn about the latest developments. The mathematical atmosphere was bristling with new ideas – an ideal background for progress.

But not all the young people arriving in Göttingen found the atmosphere congenial to their expectations. Let me cite from a letter written by a young Austrian postdoc to his academic teacher. The letter is dated 30. November 1921. He had recently arrived in Göttingen and had already introduced himself to the big shots, i.e., to Courant, Hilbert, Klein and Landau. (Emmy Noether was not yet considered an important figure in the Göttingen scene whom any newcomer had to formally introduce himself – although she was the one who cared personally for the young people arriving in Göttingen, including our man.) He had been friendly received, and Hilbert had invited him to give a talk at the Mathematical Society of Göttingen in order to report on his thesis and further work. The talk had been scheduled for 25. November 1921.

However, in his letter five days later he reports that he is deeply disappointed:

> *"... I have now given my talk but with Hilbert I had no luck. Landau and the other number theorists were quite pleased with it, they said so even during my talk when Hilbert often interrupted me. But he kept interrupting frequently – finally I could not speak any more at all – and he said that from the start he did not even listen since he had the impression that everything was trivial. But then he changed his mind when I mentioned the said decomposi-*

> *tion of prime numbers.*[2] *I had to do this out of the proper context since I could not speak and hence could not present the latest results of my thesis and of my recent investigations. But anyhow this talk had not been successful and Hilbert, through his criticism, has killed my enthusiasm for this work. By the way, in my opinion (and in the opinion of the others) his criticism is not justified. I do not know your opinion about this but as to myself, the delight for these results is gone.*"

The name of this young man was Emil Artin, he was 23 years of age and wrote this letter to his Ph.D. advisor Gustav Herglotz in Leipzig. Hilbert's conduct came as a shock to Artin. In fact, Artin closed his letter with the words:

> "*Now please excuse me, Herr Professor, from again having bothered you with such a long letter but this will probably not happen any more with this subject since I intend to drop it.*"

Artin's letter had 5 pages. We have cited the last page only, the earlier pages contain a report on certain supplements to his thesis. These pointed in the direction of the RHp[3] but Artin was not able to present them in his talk, due to Hilbert's interruptions. (See section 3.2.)

Artin's thesis contained the foundations of the arithmetic theory of quadratic function fields over a finite base field, in analogy to quadratic number fields. It included the definition and study of the zeta functions of those fields, with the aim of applying this to class number formulas, density results etc. Artin had added a list of about 30 numeric cases where he computed the class numbers and, on the way, he observed the validity of the Riemann hypothesis in characteristic p for these examples (see section 2.2).

Young Artin's disappointment about Hilbert's unfounded criticism went quite deep. Indeed he did what he had announced in his letter to Herglotz, namely he turned to other problems and in his publications never took up the problem of the RHp. Also, he left Göttingen and accepted a position at Hamburg

[2] See formulas (3.15b) and (3.16b) in section 3.2.2.

[3] Here and in the following I am using the abbreviation "RHp" for "Riemann Hypothesis in function fields of characteristic p".

University which then witnessed the rapid rise of Artin to one of the top mathematicians of his time.

To vindicate Hilbert it should be said that some days after the talk he admitted that now he appreciated Artin's work and did not any more consider it trivial. He offered Artin the publication of his new results in the *Mathematische Annalen.* But Artin did not accept the offer.

Artin's thesis appeared in print three years later in the *Mathematische Zeitschrift* [Art24]. It is the only one of all his publications in which the RHp is mentioned. In the *Nachlass* of Artin there was found a print-ready manuscript dated November 1921, containing his subsequent investigations which he had reported to Herglotz. But Artin had never submitted this for publication. See [Art00].

1.3 Hasse's project

It took quite a while until the RHp was taken up by another mathematician. His name was Helmut Hasse. He had refereed Artin's thesis in the "*Jahrbuch für die Fortschritte der Mathematik*"[4] but had not given any sign that he was particularly interested in the RHp – until he had occasion to talk to Artin in November 1932 when Hasse visited Artin in Hamburg and gave a colloquium talk. At that time, not only Artin but also Hasse was an established mathematician in the top ranks. They had exchanged letters since many years, see [FLR14].

In his Hamburg talk 1932 Hasse spoke about the problem of estimating the number of solutions of diophantine congruences, a problem in which he had recently become interested through his friend Harold Davenport. In the ensuing discussion after the talk Artin pointed out that Hasse's problem was equivalent to the RHp – as a consequence of his (Artin's) unpublished results which he had reported in his letter to Herglotz but never published. Thus Hasse's problem on diophantine congruences mutated into the RHp. With this information at hand he and Davenport succeeded quickly with the proof of RHp for generalized Fermat function fields and for Davenport-Hasse fields.

[4] Today the reviews in this journal are incorporated in the database of "zbMATH".

Parallel to this Hasse also considered the problem of the RHp for elliptic function fields. There he could draw on the classic theory of complex multiplication of elliptic functions which he was familiar with. The date of his success in the elliptic case is documented by a letter of Hasse to Mordell of 6. March 1933 where he reports that he has just obtained the proof. But Hasse's second (final) proof for elliptic fields appeared in print three years later only [Has36c]. Hasse's result in the elliptic case received much attention at that time in the mathematical community. It led to an invitation for a special 1-hour lecture at the next conference of the International Mathematical Union which was scheduled for 1936 at Oslo.

Thus Artin, although not any more active in this direction, had contributed essentially to the further development by informing Hasse about his results which he had found in 1921 but Hilbert did not wish to take notice of.

By the way, Hasse like Artin had also not found the mathematical atmosphere in Göttingen congenial to his expectations. He had entered Göttingen University at the end of 1918 when he was 20 years of age. There, in his first semester he attended Hecke's course on algebraic numbers and analytic functions (see [Hec87]). This was quite fascinating as he later recalled. But unfortunately Hecke soon left Göttingen and went to Hamburg. Hasse then became interested in Hensel's theory of p-adic numbers but he was told (by Courant) that in Göttingen this was considered but an unimportant side track and hence not worthwhile to study. So Hasse left Göttingen in 1920 and went to Hensel in Marburg where he wanted to learn more about p-adics. (Thus when Artin arrived in November of 1921 Hasse was not any more in Göttingen, the two would meet first in September of 1922 at the annual meeting of the German Mathematical Society (DMV) in Leipzig.)

In Marburg, Hasse soon got his Ph.D. with a thesis on the Local-Global Principle for quadratic forms. In consequence of Hasse's further work the p-adic numbers became an important, indispensable tool of algebraic number theory – contrary to Courant's prophecy which at that time seems to have been the general opinion in the circle around Hilbert.

After his success in the case of elliptic function fields, Hasse pushed towards a proof of RHp for arbitrary function fields with finite base fields. This turned out to become a larger project since it became necessary first to develop generally the algebraic theory of function fields (over arbitrary base fields) including their Jacobians. At that time this was not yet sufficiently established. From today's geometric viewpoint this is seen as part of the transfer of classical algebraic geometry of curves from characteristic 0 to the case when the base field is of arbitrary characteristic p.

In this story we shall meet the names of quite a number of mathematicians who at least for some time joined Hasse in this work. Some of them are:[5]

- *F. K. Schmidt* (1901-1977) who provided the proper definition of the zeta function of an arbitrary function field over a finite base field. Moreover, he proved the Riemann-Roch Theorem for function fields and showed that this is essentially equivalent with the functional equation of the zeta function [Sch31a].
- *Harold Davenport* (1907-1969) who had introduced Hasse to the problem of counting solutions of diophantine congruences. Moreover, jointly with Hasse he showed that for generalized Fermat function fields and related fields (the so-called Davenport-Hasse fields) the zeros of the zeta function are given explicitly by means of Gauss sums. This solved the RHp for these fields [DH34].
- *Ernst Witt* (1911-1991) who gave the first proof of the functional equation for the L-series of function fields over a finite base field. (But he never published it.) Moreover, among quite a number of other important contributions, he determined jointly with Hasse the structure of the maximal unramified abelian extension of exponent p (the characteristic of the field) by means of the so-called Hasse-Witt matrix [HW36].
- *Max Deuring* (1907-1984) who provided an algebraic theory of correspondences and so constructed the endomorphism ring of the Jacobian of a function field of arbitrary genus [Deu37], generalizing what Hasse had done in the elliptic case. Moreover he continued Hasse's work on elliptic function fields by determining completely the structure of their endomorphism rings.
- *André Weil* (1906-1998) who had shown interest in Hasse's project right from the beginning. He had visited Hasse in Marburg in the summer

[5] Most mathematicians which are mentioned in this book have a biographic article in "Wikipedia" or in "Mac Tutor History of Mathematics Archive" or in other openly accessible places; hence I believe it is not necessary here to always include biographical information – except in a few cases when some such information may be of interest in the present context.

> of 1933 and they exchanged letters thereafter. With the outbreak of World War II their contact broke down. Weil could escape from the Nazi terror from France to USA. There in the year 1941 he announced a proof of the Riemann hypothesis for arbitrary function fields with finite base field [Wei41].

Thus Weil was able to complete Hasse's project ten years after it had been started.

1.4 Weil's contribution

There is a letter of Weil to Artin dated 10. July 1942 in which he informed Artin about the main ideas and some details of his proof. At that time both Weil and Artin resided as refugee immigrants in the USA. This letter, which Weil has included and commented in his Collected Works [Wei80], is more detailed than his announcement in 1941. The fact that the letter was sent to Artin shows that Artin still was regarded as the ultimate expert for the RHp – although he had nothing published on this subject except what was contained in his thesis more than 20 years ago. The final version of Weil's proof appeared six years later [Wei48a].

But Weil did much more. In his proof he stressed the analogy of RHp to problems which had been treated within the framework of classical Italian algebraic geometry. But there, algebraic geometry was studied over the complex base field $\mathbb{C}$ only, i.e., in characteristic 0. So Weil had first to make sure that those results which he needed remain valid in characteristic $p > 0$. To this end he wrote a book containing a complete foundation of algebraic geometry in arbitrary characteristic [Wei46]. This was a formidable task. Weil did not only establish the theory of curves in characteristic p (this would have been sufficient for the RHp) but he covered varieties of arbitrary dimension. This enabled him to develop the theory of abelian varieties over any characteristic. In particular his treatment included Jacobian varieties which led to another proof of the RHp [Wei48b].

Weil's "Foundations of Algebraic Geometry" [Wei46] had an enormous influence. In the course of time it led to what today is generally accepted under the name of "arithmetic geometry", i.e., to the use of the language of algebraic geometry in algebra and number theory, and in other branches of mathematics.

After Weil's success there appeared several papers putting into evidence that the RHp could also have been proved within the theory of algebraic function fields, in which it had been started by Artin in 1921 and continued by Hasse. The last(?) word in this endeavor was given by Bombieri in 1972 [Bom74]. He perfected Hasse's idea to search for a proof which works in the function field directly. He concentrated his proof on the Frobenius map, without caring for the general theory of correspondences of curves. Still he relied heavily on Artin's unpublished results of 1921.

Summary

After having received his degree in Leipzig, Artin spent a year 1921/22 as post-doc in Göttingen. In several letters from there to his academic teacher Herglotz, Artin developed some important further results for zeta functions of quadratic function fields beyond his thesis. But these results were never published. The reason was that when Artin reported about it in Göttingen in the presence of Hilbert, the latter critized his work heavily. Although Hilbert later changed his mind and offered Artin publication of his new results, Artin did not accept. He left Göttingen and went to Hamburg where he turned to other problems. He continued to be interested in the RHp but did not publish anything more in this direction.

Ten years later the RHp was taken up by Hasse. In the year 1932 he gave a colloquium talk in Hamburg about the problem of estimating the number of solutions of diophantine equations, a problem in which he had recently become interested through his friend Harold Davenport. In the discussion with Artin the latter pointed out that the Hasse-Davenport problem was equivalent to the RHp for the function fields in question – as a consequence of Artin's unpublished results which he had reported in his letters to Herglotz but never published. Hasse, stimulated by results of Davenport, succeeded quickly with the proof of RHp for what today are called the generalized Fermat function fields and for related fields. Parallel to this Hasse was able to prove the RHp for elliptic function fields. There he could draw on the classic theory of *complex multiplication of elliptic functions which he was familiar with. While preparing his proof for publication Hasse found that it is possible to develop much of "complex multiplication" directly in the case of characteristic $p > 0$. He decided to write a new proof on this basis. This appeared in print three years later, in the year 1936. Although it was still limited to the elliptic case, it received much attention at that time in the mathematical community. It led to an invitation for a special 1-hour lecture at the next conference of the International Mathematical Union at Oslo in the year 1936.*

After his success in the elliptic case Hasse and his collaborators started a project towards a proof of RHp for function fields of arbitrary genus. A number of new results were reached in this direction, in particular by Deuring. Today we can see that these results were well sufficient to compose a proof of the RHp. However this goal was not reached at the time.

André Weil had shown interest in Hasse's project right from the beginning. During the 1930s they exchanged a number of letters, among others about the envisaged proof of the RHp for function fields of higher genus. Weil discovered that there is a close connection of the problem to results of Severi in classical algebraic geometry. He managed to translate the foundations of algebraic geometry to characteristic $p > 0$ and on this basis solve the RHp for general function fields in the year 1941. His final proof appeared 1948.