The Riemann hypothesis in characteristic p, its origin and development

Part 1. The formation of the zeta-functions of Artin and of F. K. Schmidt

Peter Roquette (Heidelberg)

Revised July 23, 2003

Contents

Abstract

This paper is the first part of a larger project which will give a comprehensive view of the history around the Riemann hypothesis for function fields. (A preliminary version has appeared 1997/98.) This Part 1 is dealing with the development before Hasse's contributions to the Riemann hypothesis. We are trying to explain what he could build upon. The time interval covered will be between 1921 and 1934. We start with Artin's thesis of 1921 where the Riemann hypothesis for function fields was spelled out and discussed for the first time, in the case of quadratic function fields. We will describe the activities which followed Artin's thesis up to F. K. Schmidt's classical paper 1931 on the Riemann-Roch theorem and the zeta function for arbitrary function fields. Finally we will review Hasse's paper of 1934 where he gives a summary of all what was known at that time about zeta functions of function fields.

1 Introduction

This work is an outgrow of our not yet completed project, jointly with Günther Frei, to publish a commented edition of the Artin-Hasse correspondence. There are 49 letters from Emil Artin to Helmut Hasse preserved, mostly of mathematical content, written between 1923 and 1953. The present work can be regarded as a comment to one of these letters, which is the only one where the Riemann hypothesis for function fields is mentioned and which reads as follows: [1]

Hamburg, Jan 17, 1934

"*Dear Mr. Hasse!*

Would you like to visit us in this semester for guest lectures? You could speak about whatever you like. Perhaps the beautiful results on the Riemann hypothesis? This is the most beautiful thing that has been done in decades. My audience would be very much interested in it. If you write whether and when you will come then I shall manage everything with our administration and you will receive the official

[1] The letter is written in German language. The English text is our own free translation – here and wherever in this article we will cite German text in English language.

invitation. It would be nice if you could come for one week but if not then something less will also do. Looking forward to obtaining your consent, with many kind regards from ourselves to you all

your Artin"

When Artin referred to the "Riemann hypothesis" then he meant the "Riemann hypothesis for functions fields with finite base fields".

We use the term "function field" in the sense as Artin and Hasse did. Thus a function field $F|K$ is a finitely generated extension of transcendence degree 1 over its base field K. In other words, F is the function field of an absolutely irreducible curve defined over K. If K is finite then one can associate to F a zeta function

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

where $\mathfrak{p}$ ranges over the primes (or prime divisors) 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 $|\mathfrak{p}|$.

The above definition of $\zeta_F(s)$ as a function of the complex variable s has been given by F. K. Schmidt in [FK:1926]. The infinite product and sum on the right hand side of (1) converge if the real part of s is $\Re(s) > 1$. Using the Riemann-Roch theorem, $\zeta_F(s)$ appears as a rational function of the variable $t = q^{-s}$ where $q = |K|$ denotes the order of K. Hence $\zeta_F(s)$ extends uniquely to a meromorphic function on the whole complex plane.

The *Riemann hypothesis* for F says that all roots of $\zeta_F(s)$ have their real part $\Re(s) = \frac{1}{2}$. Equivalently, if ζ_F is considered as a rational function of $t = q^{-s}$ then all roots have absolute value $|t| = \sqrt{q^{-1}}$.

Of course, Riemann himself did not consider zeta functions of function fields; he was interested in the ordinary "Riemann" zeta function $\zeta(s)$, and he conjectured that the so-called *non-trivial* zeros of $\zeta(s)$ have real part $\Re(s) = \frac{1}{2}$. The validity of this conjecture would have far-reaching consequences in the theory of prime number distribution and other arithmetic questions.

Nowadays it has become customary to speak of "Riemann hypothesis" also in the case of zeta functions of other kind, for instance the Dedekind zeta function of a number field or the F. K. Schmidt zeta function of a function field. Similarly as for the Riemann zeta function, the validity of the Riemann hypothesis for $\zeta_F(s)$ in the function field case has far-reaching consequences for the arithmetic of function fields.

In the following, "Riemann hypothesis" will always refer to the zeta function of function fields, if nothing is said to the contrary.

In February 1933, about 11 months before Artin's letter, Hasse had succeeded in proving the Riemann hypothesis for elliptic function fields, i.e., for function fields of genus 1. In his original proof [H:1933] Hasse had used the classical theory of complex multiplication. But later he discovered that he could do without it, since he was able to develop some kind of abstract complex multiplication in characteristic p. It seems that Artin had heard about it, and it was this new theory of Hasse that Artin had in mind when he talked about the "*most beautiful thing that has been done since decades*".

Hasse, who held a position in Marburg at that time, accepted Artin's invitation. In the first week of February 1934 he delivered three lectures at the University of Hamburg, giving an outline of his new proof. The lectures were published in the same year in the "*Hamburger Abhandlungen*" [H:1934a]. This initiated a tremendous development, not only among the mathematicians around Hasse but also in wider circles. It culminated in the proof of the Riemann hypothesis for arbitrary function fields by A. Weil in [W:1948a], based on Hasse's and Deuring's ideas, and finally in the extremely short and elegant proof by Stepanov-Bombieri in [Bo:1974]. Moreover, in this process it was necessary to develop in some detail the arithmetic of function fields, or curves, over fields of arbitrary characteristic, and this can be viewed as one of the origins of today's image of abstract algebraic and diophantine geometry.

For a description of the contents of this paper we refer to the table of contents above. The reader will observe that our discussion is not limited exclusively to the zeta functions but we have also to take into account the development of the arithmetic of function fields which, as we have just pointed out, is intimately connected with zeta functions.

As said in the abstract, this paper is Part 1 of a larger project about the history of the Riemann hypothesis for function fields. In Part 2 we shall report about how Hasse had been introduced by Davenport to the problem of the Riemann hypothesis, how he was able to solve it for the case of genus 1 using the classical theory of complex multiplication and its class field theory, and about his joint work with Davenport on a class of function fields of higher genus (today called Davenport-Hasse fields) for which they could prove the Riemann hypothesis by identifying the zeros of the zeta function as certain Gauss and Jacobi sums. The story will be continued in Parts n for n>2.

We will use not only published material but also the information that is contained in personal documents like letters, manuscripts etc. of the persons involved. Mathematics does not develop just by itself but it is shaped by human people. It is my opinion that if we know more about the personalities of the acting people we can understand better the development of our science.

Remark. All the letters and other documents which we cite are contained in the library and the archive of Göttingen university, except when we explicitly mention another source. At this occasion I would like to thank the colleagues in departments, libraries and archives for their help.

Acknowledgement: I would like to thank Günther Frei, Franz Lemmermeyer and Peter Ullrich for valuable comments.

2 The beginning: Artin's thesis

Emil Artin was born 1898 in Vienna. He was brought up in Reichenberg, a German speaking town in Northern Bohemia belonging to the Austro-Hungarian empire. [2] In 1916 he enrolled at the University of Vienna where, among others, he attended a lecture course by Ph. Furtwängler. After one semester of study he was drafted to the army. In January 1919 he entered the University of Leipzig. [3]

[2] The town is now called Liberec, in the Czech Republic.

[3] This information is taken from Artin's own hand-written vita that he submitted together with his thesis to the Faculty at Leipzig University.– The documents for Artin's Ph.D. are kept in the archive of the University of Leipzig.

In June 1921 he obtained his Ph.D. with Herglotz as his thesis advisor.

Gustav Herglotz (1881-1953), also of Austro-Bohemian origin, was an all-round mathematician whose work covered astronomy, mathematical physics, geometry, applied mathematics, differential equations, potential theory, and also number theory. His five number theoretical papers were published within the period of 1921-1923. It seems that Artin came to Leipzig just during the time when Herglotz was interested in number theory, and so he inherited this interest from his academic teacher. Or was it the other way round, that Herglotz got interested in number theory through his brilliant student Artin?

In the preface of Artin's Collected Papers [A:1965] the editors remark that Artin kept a heartfelt appreciation towards Herglotz throughout his life. Herglotz was the only person whom Artin recognized as his academic teacher. Ullrich [Ul:2000] points out that in Artin's early letters to Herglotz in 1921/22, he always signed with the words "*your grateful disciple*" ("*Ihr dankbarer Schüler*").

2.1 Quadratic function fields

In his thesis [A:1924] Artin considers quadratic function fields, i.e., quadratic extensions of the rational function field $K(x)$. Artin restricts his investigation to characteristic $\neq 2$. [4] Thus a quadratic function field is of the form $F = K(x, \sqrt{D})$ with $D \in K[x]$ square free. Such fields are also called "hyperelliptic". But strictly speaking, this applies only if the degree $\deg D > 4$. If $\deg D = 3$ or 4 then F is "elliptic". If $\deg D = 1$ or 2 then F is of "rational type", and it is rational if K is finite. In order to avoid discussion of trivial cases, let us assume here that $\deg D > 0$ although Artin in his discussion carries the case $\deg D = 0$ as far as possible. (In that case $F = K'(x)$ where $[K' : K] = 2$, i.e., F is a quadratic constant field extension of $K(x)$.)

The base field K is assumed to be finite. Actually, Artin in his published thesis works over the prime field $K = \mathbb{F}_p$ only. But in a letter to Herglotz dated Nov 13, 1921 he says: [5]

> *"It should be observed that the theory remains valid word by word over an arbitrary Galois field, if p is understood to be not a prime number but the corresponding prime power whose exponent is not important. Of course this is self-evident and not new."*

This letter was written one month after Artin had sent his manuscript to the *Mathematische Zeitschrift*, which was Oct 14, 1921. Thus very probably Artin knew about this generalization before sending it. Then why did he not include this more general case? Perhaps his above letter contains the answer: Artin did not regard this as necessary because the said generalization was self-evident to him.

Accordingly we will discuss Artin's thesis in his spirit, i.e., as if it would refer to an arbitrary finite base field K, of characteristic $p \neq 2$. We denote the

[4] The case of characteristic 2 can be discussed in a similar manner by using the Artin-Schreier generators of quadratic field extensions. But the theory of Artin-Schreier extensions in characteristic p did not yet exist at the time of Artin's thesis; it was published [A-S:1927] only. The use of Artin-Schreier generators in the arithmetic theory of function fields was initiated by Hasse in [H:1934c].

[5] All letters from Artin to Herglotz cited in this article are contained in the Herglotz legacy, which is kept at the Göttingen University Library.

order of K by q. This is a power of p whose exponent, as Artin said in his letter, may be arbitrary.

Artin's aim is to develop the theory of quadratic function fields, including their zeta functions, in complete analogy to the theory of quadratic number fields. It was Herglotz who had suggested this topic to Artin; this has been stated by Artin himself in his vita that he submitted to the Faculty in Leipzig. Frei [Fr:2001] forwards the opinion that Herglotz had proposed this topic because he had read a paper by Kornblum [Kor:1919] that deals with L-series in rational function fields. This opinion is confirmed by Herglotz' report about Artin's thesis to the Faculty in Leizig: It starts with a brief description of Kornblum's paper. See section 2.2 for the relation of Artin's thesis to Kornblum's result. [6]

We would like to point out that the Riemann hypothesis for function fields was *not* the main theme of Artin's thesis. In fact, Herglotz says in his report:

> "*The author develops the complete "number theory" in quadratic function fields – to the same extent as it is known today for quadratic number fields.*"

And then Herglotz proceeds to state the main points of the arithmetic of quadratic number fields, which Artin had transferred to the function field case:

- foundation of ideal theory,
- theory of units,
- the zeta function and its functional equation,
- number of ideal classes,
- existence of the genera.

And only in a side remark Herglotz mentions that Artin had obtained certain evidence of the curious fact that the non-trivial roots of the zeta function have real part $\frac{1}{2}$.

Artin assumes the theory of quadratic number fields to be well known, but he does not cite any literature for this. When and where did Artin himself learn about quadratic fields? It seems probable that he had learned it through a lecture by Herglotz on this topic. For, in the summer semester 1919 Herglotz had announced a lecture course "Elementary Number Theory", and in the following semester "Number Theory (Quadratic Number Fields)". And in the summer 1920 Herglotz offered three courses, "Algebraic equations", "Geometry of numbers", and "Problem sessions on number theory", altogether 9 hours weekly. [7] We cannot help thinking that this amassment of algebraic and number theoretic courses in Leipzig was done to satisfy the thirst for knowledge of his eager student Artin. Herglotz' lectures were generally regarded as "pieces of art" ("*Kunstwerke*") according to Constance Reid [Rei:1976].

A short, 9-page preview of Artin's thesis was published in [A:1921] already, in the *Jahrbuch* of the Philosophical Faculty of Leipzig University. There, Artin says that his proof of the finiteness of class number is analogous to the proof as presented in the second volume of Weber's Algebra book [Web:1895]. So we may suppose that Artin had read Weber, as had probably every young mathematician in Germany at that time who was working in algebra or number theory. At one

[6] Herglotz' report is kept at the archive of Leipzig University.

[7] I am obliged to Frau Dr. Peter from the University of Leipzig for sending me the lecture announcements of the years 1918–1921.

point in his thesis he refers in some detail to Landau's book "Introduction to the elementary and analytic theory of algebraic numbers and ideals" [Lan:1918].

On the function field side, Artin refers to Dedekind's paper [De:1857]. There, Dedekind had developed the arithmetic theory of the polynomial ring $\mathbb{F}_p[x]$ over the prime field $\mathbb{F}_p$, including the quadratic reciprocity law in this ring – in complete analogy to the arithmetic of $\mathbb{Z}$ and Gauss' quadratic reciprocity law there. Artin says that this suggests to extend Dedekind's theory by adjoining a quadratic irrationality $\sqrt{D(x)}$ to the field of rational functions modulo p. Thus he wants his work to be considered as a continuation of Dedekind's paper more than 65 years ago. The title of that old paper of Dedekind reads:

> *Outline of a theory of higher congruences with respect to a real prime number modulus.*

This title may sound somewhat strange to us, but it makes sense if we recall that in those times a "real prime number" was understood to be an ordinary prime number $p \in \mathbb{Z}$, in contrast to "imaginary" primes which may be primes in $\mathbb{Z}[\sqrt{-1}]$ or in other rings of algebraic numbers. And "higher congruences" in $\mathbb{Z}[x]$ meant simultaneous congruences modulo p and modulo some polynomials $f(x) \in \mathbb{Z}[x]$; this is essentially equivalent to congruences in $\mathbb{F}_p[x]$ modulo some polynomials $\overline{f}(x) \in \mathbb{F}_p[x]$. Dedekind regarded $\mathbb{F}_p[x]$ not as a mathematical structure in its own right, but as the result of reduction of $\mathbb{Z}[x]$ modulo p.

On first sight Artin seems to adopt the same view point. For the title of his thesis is:

> *Quadratic fields in the domain of higher congruences.*

Since Artin uses Dedekind's terminology of "higher congruences", one may be tempted to conclude that he too wants to regard his fields as obtained by reduction mod p from a function field in characteristic 0. But in the very beginning of his paper he says: "*We will call functions and numbers to be equal if they are congruent modulo p in the sense of Dedekind.*" This makes clear that, although Artin wishes his work to be regarded as a follow up of Dedekind's, he immediately switches to the viewpoint, "modern" at the time, that he is working in fields of characteristic p in the sense of Steinitz' great paper [Ste:1910]. [8]

Artin's thesis is divided into two parts: an arithmetic part and an analytic one. The adjectives "arithmetic" and "analytic" are Artin's. But what is their meaning in this context?

Artin does not give any explanation; obviously he assumes that the reader will know how these words were used at the time. But the usage of the word "arithmetic" has changed through the times, and even at the time of Artin's thesis it was not used uniformly. Emmy Noether, for instance, used "arithmetic" in connection with anything referring to ideal theory and the decomposition into prime ideals, and in an extended way also to module theory. Today this would be regarded as belonging to "algebra", or sometimes to "algebraic geometry" but not to "arithmetics". Some authors, also in Artin's time, would reserve the word "arithmetic" for those topics that refer explicitly to properties of the

[8] For some time function fields over finite base fields were called "congruence function fields" (*Kongruenzfunktionenkörper*) and their zeta functions were called "congruence zeta functions" (*Kongruenzzetafunktionen*); see, e.g., [H:1934b], [Rq:1953], [We-Zi:1991]. This terminology can be understood as a remnant of Dedekind's "higher congruences".

natural numbers. But do function fields in charactersistic p fit into this scheme? Similar comments can be given for the use of "analytic". If this should imply that analytic functions are involved then one could point out that the zeta function of a function field is essentially a rational function. The Riemann hypothesis refers to the zeros of a polynomial. Is this "analytic" or "algebraic"?

So let us refrain from the attempt to describe in detail the meaning of "arithmetic" versus "analytic" in the context of Artin's thesis. Artin himself seems to have used these adjectives in some intuitive way, hoping that the reader will be able to follow his intentions; let us try to do this too. Within the theory of function fields, "analytic" will be used for those topics which refer to the zeta function, L-series and similar objects of function fields, and "arithmetic" for those which do not, preferably if the base field is finite and thus the analogy to number theory is quite close.

2.1.1 The arithmetic part

As said above, K denotes a finite field of characteristic $\neq 2$. Let $F = K(x, \sqrt{D})$ where $D \in K[x]$ is assumed to be square free. Let $R = K[x, \sqrt{D}]$ denote the integral closure of $K[x]$ in F. [9] In this first part Artin sets out to develop the "arithmetics" of the ring R. His main results are the following statements.

(i) **Prime ideal decomposition:** *R is a Dedekind ring.*

(ii) **Ramification:** *Let $P \in K[x]$ be a prime polynomial. P is ramified in R if and only if P divides D.*

(iii) **Unit theorem:** *The unit group $R^\times$ is either a torsion group, consisting of the constants, or there exists a fundamental unit ε which generates $R^\times$ modulo torsion.*

(iv) **Class number:** *The ideal class group of R is finite.*

(v) **Decomposition law:** *Suppose P does not divide D. Then P splits in R if and only if D is a quadratic residue modulo P. If this is the case then $P \cong \mathfrak{p}\mathfrak{p}'$, the product of two different prime ideals of R. Otherwise P remains prime in R, i.e., $P \cong \mathfrak{p}$.*

(vi) **Reciprocity:** *In $K[x]$ there holds a quadratic reciprocity law, in analogy to Gauss' quadratic reciprocity law in $\mathbb{Z}$.*

Artin does not use the terminology of "Dedekind ring"; this was introduced much later. [10] Artin just proves that every proper ideal of R admits a unique representation as a product of prime ideals. In fact, he does not even present a complete proof. After having developed the relevant facts about ideals and their norms in R he says: "*It is seen that the arguments are completely parallel to those in the number field case. Hence in the following it will suffice to state the definitions and theorems.*" We have said earlier already that Artin does not cite any source where the reader could find in detail those arguments for number fields. Obviously he could assume that the material was common knowledge among the prospective readers of his papers.

[9] The notation is ours, not Artin's. Quite generally, for the convenience of the reader we use consistently our own notation, which often is different from the notations used in the various papers which we will report on.

[10] Ullrich [Ul:1999] reports that I.S. Cohen [Coh:1950] used the word "Dedekind domain". Lemmermeyer in his *Zentralblatt*-review of Ullrich's paper points out that Dieudonné [Dieu:1947] already used this terminology.

It is worth while to observe how Artin handles the place at infinity. He realizes that in $K(x)$ the place $x \mapsto \infty$ plays a role similar to the ordinary absolute value in the number field case. Now in the latter case, the completion of $\mathbb{Q}$ with respect to the ordinary absolute value is $\mathbb{R}$, the field of real numbers. The theory of quadratic number fields $\mathbb{Q}(\sqrt{D})$ looks quite different according to whether $\sqrt{D}$ is real or not. Here, "real" means that D is a square in $\mathbb{R}$, and then $\mathbb{Q}(\sqrt{D})$ can be regarded as a subfield of $\mathbb{R}$.

In analogy to this, Artin in the function field case considers the valuation of $K(x)$ at infinity and its completion. He does not use the terminology of valuation or completion, he just says that he will extend $K[x]$ by considering not only polynomials but arbitrary power series of the form $f(x) = \sum_{\nu=-\infty}^{n} a_\nu x^\nu$ with $a_\nu \in K$, which is to say Laurent series with respect to x^{-1}. If $a_n \neq 0$ then Artin defines n to be the degree and $|f(x)| = q^n$ as the size of that Laurent series. (q denotes the order of K.) In fact, this is the valuation at infinity. And the field of those Laurent series is the completion of $K(x)$ with respect to this valuation.

The use of this valuation in the function field case is quite the same as that of the ordinary valuation in the number field case. It is used to estimate ideals and functions in order to prove the finiteness of the class number, and it is also used to define continued fraction expansions in order to obtain the fundamental unit in the "real" case. Here, $\sqrt{D}$ is called "real" if it can be represented by a Laurent series of the above type, and "imaginary" if not. Artin shows that $\sqrt{D}$ is real if and only if the degree of D is even and, in addition, the highest coefficient of D is a square in K. Of course this is a consequence of Hensel's Lemma but Artin gives explicitly the Laurent expansion of $\sqrt{D}$ in terms of the binomial expansion for the exponent $\frac{1}{2}$.

Concerning the quadratic reciprocity law (vi): If the prime function P does not divide D then Artin introduces the Legendre symbol $(\frac{D}{P})$ which assumes the value 1 or -1 according to whether D is a quadratic residue modulo P or not. [11] Artin points out that Dedekind [De:1857] had stated the quadratic reciprocity law for this symbol without proof, and therefore he will now present a proof in detail. Well, Dedekind had said the following, after stating the reciprocity law: That he had transferred to the function field case all ingredients of Gauss' fifth proof, and therefore it would not be necessary to repeat the proof in every detail.

It seems that this did not satisfy Artin. So he presented his own proof of the quadratic reciprocity law:

$$\left(\frac{Q}{P}\right)\left(\frac{P}{Q}\right) = (-1)^{nm\frac{q-1}{2}} \tag{2}$$

where P, Q are different monic prime polynomials in $K[x]$, of degrees n and m respectively. Artin's proof does not follow the lines of Gauss' fifth proof. Frei [Fr:2001] points out that Artin's proof can be regarded as the transfer from a proof of Kummer [Kum:1861].

Apparently Artin did not realize that the quadratic reciprocity law in $K[x]$ for finite K can be proved quite elementarily. In fact, this is so even for the ℓ-th power reciprocity law for any number ℓ dividing $q-1$. This had been observed by Kühne [Kue:1902] already. It seems that Artin did not know Kühne's

[11] Artin writes $[\frac{D}{P}]$ and reserves parentheses $(\frac{d}{p})$ for the quadratic residue symbol in $\mathbb{Z}$.

paper. Later, F. K. Schmidt in his thesis [FK:1925] rediscovered Kühne's power reciprocity law; see section 4.4.2.

Artin also defines the Jacobi symbol $(\frac{M}{N})$ for two *arbitrary* polynomials M, N in $K[x]$ without common divisor, in such a way that it becomes bi-multiplicative in each of the two variables M amd N – quite the same as for the ordinary Jacobi symbol in $\mathbb{Z}$. Formula (2) remains valid for this extended symbol if M and N are monic. If they are not and if a, b are their highest coefficients respectively, then the inversion factor on the right hand side of (2) has to be modified by a factor which depends on the quadratic character of a and b in K.

The importance of the reciprocity law for Artin's investigation is its application to the decomposition law in statement (v) above. Namely, (v) shows that the decomposition type in R of a monic prime polynomial P, not dividing D, is governed by the value

$$\chi_D(P) = \left(\frac{D}{P}\right) \tag{3}$$

which is either 1 or -1. On the other hand, $\chi_D(M) = \left(\frac{D}{M}\right)$ is defined as a Jacobi symbol for an *arbitrary* polynomial $M \in K[x]$ relatively prime to D. It follows from the reciprocity law that $\chi_D(M)$, as a function of M, is a *quadratic character* which differs from the residue character $M \mapsto \left(\frac{M}{D}\right)$ modulo D by a factor depending only on the degree and on the highest coefficient of M (if D is considered to be fixed). From this it follows in a straightforward manner the following statement:

Let $n > 0$ denote the degree of the discriminant D. If M ranges over all monic polynomials, relatively prime to D and of fixed degree m then

$$\sum_{\deg M = m} \chi_D(M) = 0 \qquad \text{if} \qquad m \geq n\,. \tag{4}$$

The condition that M is relatively prime to D can be omitted by putting $\chi_D(M) = 0$ if M has a common divisor with D.

The statement (4) becomes a key result in the "analytic" part of Artin's thesis.

2.1.2 The analytic part

In the second part of his thesis Artin introduces the zeta function

$$\zeta(s) = \prod_{\mathfrak{p}} \frac{1}{1 - |\mathfrak{p}|^{-s}} = \sum_{\mathfrak{a} \neq 0} |\mathfrak{a}|^{-s} \tag{5}$$

where s is a complex variable. Here, $\mathfrak{p}$ ranges over all prime ideals $\neq 0$ of R and $|\mathfrak{p}| = q^{\deg \mathfrak{p}}$ is the order of the residue field $R/\mathfrak{p}$. Accordingly $\mathfrak{a}$ ranges over all integral ideals $\neq 0$ of R and $|\mathfrak{a}|$ is the order of $R/\mathfrak{a}$. Since R is a Dedekind ring, every nonzero prime ideal $\mathfrak{p}$ of R defines a unique place of the field F which may also be denoted by $\mathfrak{p}$, and $R/\mathfrak{p} = F\mathfrak{p}$. But in this way one does not obtain all the places of F; the places at infinity (where $x(\mathfrak{p}) = \infty$) do not correspond to prime ideals of R. (There are one or two such places of F.) Hence Artin's zeta function $\zeta(s)$ does not coincide with F. K. Schmidt's zeta function $\zeta_F(s)$ as defined in (1); they differ by the Euler factors belonging to the (one or two) infinite places. Because of this, Artin's formulas look somewhat different from the corresponding formulas which we are used today when referring to F. K. Schmidt's zeta function. Also, the Riemann hypothesis for Artin's zeta function concerns the "nontrivial" roots of $\zeta(s)$ only: these are supposed to have real part $\Re(s) = \frac{1}{2}$, whereas there may be some "trivial" zeros of $\zeta(s)$. We shall see below which roots have to be regarded as "trivial" in this sense.

Artin himself in his thesis is not concerned with the difference between $\zeta(s)$ and $\zeta_F(s)$; at that time the latter was not yet defined and Artin's zeta function was for him the first and only zeta object to study in quadratic function fields. [12]

The investigation of the properties of $\zeta(s)$ is the main objective in the second part of Artin's thesis. The first and essential observation is that $\zeta(s)$ is a rational function if considered as a function of the variable $t = q^{-s}$. If this would have been known for the F. K. Schmidt zeta function already then it would follow immediately for Artin's zeta function too, because the (one or two) additional Euler factors are also rational functions of t. F. K. Schmidt in [FK:1931a] proves the rationality of his zeta function by means of the theorem of Riemann-Roch. But as we have just said, F. K. Schmidt's result was much later, and also the Riemann-Roch theorem had not yet been established for function fields with finite base field. Thus Artin had to use another strategy. He represented the zeta function of the quadratic function field by means of an L-series – following the procedure in the number field case.

Let $P \in K[x]$ denote a monic prime polynomial. The decomposition type of P into prime ideals in R is governed by the value of the character $\chi_D(P)$, as Artin had shown in the first part of his thesis (see (3)). Let us write this information in a table as follows:

$\chi_D(P)$	1	-1	0
decomposition:	$P \cong \mathfrak{p}\mathfrak{p}'$	$P \cong \mathfrak{p}$	$P \cong \mathfrak{p}^2$
degree:	$\deg \mathfrak{p} = \deg \mathfrak{p}' = \deg P$	$\deg \mathfrak{p} = 2 \deg P$	$\deg \mathfrak{p} = \deg P$

That is, if $\chi_D(P) = 1$ then P splits into two different prime ideals $\mathfrak{p}$ and $\mathfrak{p}'$, each of the same degree as P, etc. In the Euler product (5) we look at those factors which belong to prime ideals $\mathfrak{p}$ dividing P; there are two such factors if $\chi_D(P) = 1$ and one factor in the other cases. In every case, an inspection shows that one can write (putting $t = q^{-s}$):

$$\zeta(s) = \prod_P \left(\frac{1}{1 - t^{\deg P}} \right) \left(\frac{1}{1 - \chi_D(P) t^{\deg P}} \right) \tag{6}$$

Here P ranges over the monic prime polynomials of $K[x]$. The product of the first factors is

$$\prod_P \frac{1}{1 - t^{\deg P}} = \sum_M t^{\deg M} = \sum_{m \geq 0} q^m t^m = \frac{1}{1 - qt}$$

where M ranges over all monic polynomials $\neq 0$ of $K[x]$. This function can be identified with the zeta function of $K[x]$. The product of the second factors on the right hand side of (6) is

$$\prod_P \frac{1}{1 - \chi_D(P) t^{\deg P}} = \sum_M \chi_D(M) t^{\deg M} = \sum_{m \geq 0} \sigma_m t^m \tag{7}$$

[12] But in a letter to Herglotz written on Nov 30, 1921, Artin already considered the change of his zeta function by birational transformations. See section 2.3.3 below.

where

$$\sigma_m = \sum_{\deg M = m} \chi_D(M)\,, \tag{8}$$

and we see that this is the L-series with respect to the character χ_D. It appears as a power series in $t = q^{-s}$. [13] But now, using the relation (4) Artin concludes that this power series is in fact a polynomial of degree $< n$; recall that $n = \deg D$. This gives:

The function $\zeta(s)$ can be written in the following form, with $t = q^{-s}$:

$$\zeta(s) = \frac{L(t)}{1 - qt} \tag{9}$$

with $L(t)$ a polynomial of degree $\leq n - 1$:

$$L(t) = \sum_{m=0}^{n-1} \sigma_m t^m \tag{10}$$

where the coefficients σ_m are given by (8)*.* (Actually, it turns out with the help of the functional equation that the degree of $L(t)$ is precisely $n - 1$.)

This result is central in Artin's thesis. It shows quite explicitly that the zeta function of a function field behaves quite differently from the zeta function in the number field case. $\zeta(s)$ is rational if regarded as a function in the variable $t = q^{-s}$, and it is even a polynomial apart from the trivial factor $\frac{1}{1-qt}$. In the introduction Artin had said:

> *It appears that a general proof* [of the Riemann hypothesis for quadratic function fields] *will have to deal with problems of similar type as with Riemann's $\zeta(s)$, although here* [in the function field case] *the situation is clearer and more lucid because it essentially concerns polynomials.*

From the second part of this statement we see that he was fully aware of the fact that in the function field case the situation is different from the classical case.

Artin was not the first to have observed this. There is a paper by H. Kornblum [Kor:1919] where L-series $L(t|\chi)$ are considered, for arbitrary characters χ in $K[x]$ modulo D; we have mentioned Kornblum's paper already. For non-trivial characters Kornblum had shown that $L(t|\chi)$ is a polynomial of degree $< n$, and he had done it in the same way as Artin does in his thesis by means of (4). (For more about Kornblum's paper see section 2.2.)

But Artin observes, beyond Kornblum's results, the significance of $L(t)$ for the arithmetic of the domain R, in particular the Riemann hypothesis. This can be regarded as the main achievement in Artin's thesis.

Artin immediately starts to draw consequences from (9). Among other things he arrives at the following results:

1. Trivial and non-trivial zeros: If $\sqrt{D}$ is "real" then $L(1) = 0$. If $\sqrt{D}$ is "imaginary" and the degree of D is even then $L(-1) = 0$. These zeros are

[13] Artin does not introduce the variable t and keeps writing q^{-s} instead; in this way the power series in t appears as a Dirichlet series in s. He also does not use the notation L at this point.

called “trivial”. All the other “nontrivial” zeros of $L(t)$ are contained in the region $\frac{1}{q} < |t| < 1$ which means $0 < \Re(s) < 1$.

The Riemann hypothesis says that the *non-trivial roots* have $\Re(s) = \frac{1}{2}$. Artin cannot prove this in general but he observed that the weaker result in statement **1.** has already important consequences. (See **4.** and also the next section 2.2.) It is during the proof of **1.** that Artin, as we have said above already, refers to Landau's textbook [Lan:1918]; at some point he is content with saying: “*The rest of the proof runs completely analogous to the ordinary proof, so much that it may be allowed to suppress it.*” This happens when Artin wishes to show, by analytic means, that there are no zeros on the line $\Re(s) = 1$.

2. Class number formula: Formulas for the class number h of R are obtained by computing the residue of $\zeta(s)$ at $s = 1$. Artin obtains the following formulas:

$$\lim_{s\to 1}(s-1)\zeta(s) = \begin{cases} \dfrac{\sqrt{q}\cdot h}{\sqrt{|D|}\cdot \log q} & \text{if } \sqrt{D} \text{ is “imaginary” and } \deg D \text{ odd} \\ \dfrac{(q+1)h}{2\sqrt{|D|}\cdot \log q} & \text{if } \sqrt{D} \text{ is “imaginary” and } \deg D \text{ even} \\ \dfrac{(q-1)\mathcal{R}h}{\sqrt{|D|}\cdot \log q} & \text{if } \sqrt{D} \text{ is “real” .} \end{cases} \tag{11}$$

In the “real” case, $\mathcal{R}$ denotes the regulator of R which is the integer defined by the formula $|\varepsilon| = q^{\mathcal{R}}$ involving the fundamental unit ε. Thus $\mathcal{R} = \frac{\log|\varepsilon|}{\log q}$. By means of the functional equation (see below) this can be transferred to $s = 0$, i.e., $t = 1$. Artin arrives at the following formulas involving the coefficients σ_m of the polynomial $L(t)$:

$$h = \begin{cases} \sigma_0 + \sigma_1 + \cdots + \sigma_{n-1} & \text{if } \sqrt{D} \text{ is “imaginary”} \\ -\dfrac{1}{\mathcal{R}}\left(\sigma_1 + 2\sigma_2 + \cdots + (n-1)\sigma_{n-1}\right) & \text{if } \sqrt{D} \text{ is “real” .} \end{cases} \tag{12}$$

3. Functional equation: We do not write down the explicit form of the functional equation for Artin's zeta function. Let us be content by saying that $\zeta(1-s)$ is expressed in terms of $\zeta(s)$, and hence $L(\frac{1}{qt})$ in terms of $L(t)$. This yields certain symmetry properties of the coefficients σ_m which Artin uses to reduce computations of the class number h by means of (12). If D is of odd degree $n = 2m+1$ (and hence there are no nontrivial zeros of $L(t)$) then these read:

$$\sigma_{2m-\nu} = q^{m-\nu}\sigma_\nu\,. \tag{13}$$

4. Prime ideal distribution: Let S_n denote the number of prime ideals $\mathfrak{p}$ of R with $|\mathfrak{p}| = q^n$, then

$$S_n = \frac{q^n}{n} + O\left(\frac{q^{n\theta}}{n}\right) \qquad \text{with} \qquad \frac{1}{2} \le \theta < 1\,. \tag{14}$$

The exponent θ in the error term is the maximum of the real parts $\Re(s)$ of the non-trivial zeros s of $L(s)$. Today, since the Riemann hypothesis is known to hold, we can conclude that $\theta = \frac{1}{2}$.

5. Table of class numbers: Artin computes numerically tables of class numbers, for small prime numbers $q = p \le 7$ and small degrees 3 or 4. (These fields are all elliptic.) Using the above mentioned symmetry properties it suffices to compute σ_1. In each of Artin's cases, about 40 in number, he verifies numerically the Riemann hypothesis, by computing the coefficient σ_1 and discussing the zeros of the corresponding L-polynomial. Furthermore, Artin remarks (without showing tables) that for $q = p = 3$ he has extended his computations to discriminants of degree 5, except some prime discriminants, and again, the Riemann hypothesis could be verified. (These fields are hyperelliptic of genus 2.)

This last remark about discriminants of degree 5 had been inserted into the manuscript shortly before the paper was sent to print, which was Oct 13, 1921. Hence between the first version, which was submitted to the Leipzig Faculty in May 1921, and the final version in October, Artin had done quite a number of additional computations in order to verify the Riemann hypothesis. But now, in a letter to Herglotz on the same day, Oct 13, he writes:

> "*I believe that more computations are of no use, for they would not lead to a decision. (If so, then in the negative sense.)*"

From this statement we can see Artin's own attitude towards the question of whether the Riemann hypothesis is valid or not. Although formally he leaves the question open, admitting that the final decision may turn out to be negative, he is putting that possibility into parentheses. And he decided to stop his computations. So we may infer that indeed he expects the Riemann hypothesis to hold generally.

Sometimes in the literature it is said that the Riemann hypothesis for hyperelliptic function fields "was first conjectured by Artin". But in Artin's thesis we do not find any statement which looks like such conjecture. It is true that Artin was the first to state the problem and to open the scene. But as far as we know, he never came out with the conjecture that the Riemann hypothesis would be true. The only indication about his personal attitude is contained in the above passage in his letter to Herglotz. In this sense, perhaps, Artin can be said to have "conjectured" the Riemann hypothesis. But it would be more precise to say that in his thesis he had "verified" it in examples.

Although Artin cannot prove the Riemann hypothesis in general he proves the following conditional result:

6. Finitely many imaginary fields with given class number: *If the validity of the Riemann hypothesis is assumed for quadratic function fields, the following holds: For all $q > 2$ and $n = \deg D > 2$ there are only finitely many "imaginary" quadratic function fields $F = \mathbb{F}_q(x, \sqrt{D})$ with given class number h. Fields with $h = 1$ are possible only for $q = 3$. (Probably only one field).*

Artin mentions this result for $h = 1$ because in the number field case, the determination of all imaginary quadratic fields with class number 1 was a classical problem which at the time of Artin's thesis was not solved. Nine imaginary quadratic number fields were known but it was not yet known whether that list was complete (which it is). Thus Artin wished to put into evidence that in the case of function fields the situation is easier than for number fields.

The "one field" with class number $h = 1$ for $q = 3$ which Artin mentions, is listed in his class number tables; it belongs to the discriminant $D = t^3 - t - 1$. Artin does not say how he derived at his prediction that this is the only imaginary quadratic function field with $h = 1$ and $q = 3$, neither in the published thesis nor in his letter to Herglotz. Has he just used computational evidence? Artin's prediction was later verified by MacRae [MacR:1971]. Besides of the Riemann hypothesis MacRae used the Riemann-Roch theorem. This, as we know from F. K. Schmidt, is equivalent to the functional equation of the zeta function, and the latter was available to Artin for quadratic function fields. Thus in principle, Artin could have done himself the (elementary) computation which MacRae did. But it is not straightforward. [14]

2.2 Prime polynomials in arithmetic progression

In the last two sections of the analytic part, Artin considers the problem of primes in an arithmetic progression of $K[x]$, where “prime” means “monic prime polynomial”. He refers to the paper by Kornblum [Kor:1919] which we have mentioned above already, and he announces that he succeeded to obtain a much stronger result.

Heinrich Kornblum had been a young student in Göttingen who had written a manuscript on primes in arithmetic progressions in $K[x]$, under the supervision of Landau. He was almost 24 when World War I. started. He volunteered for the army and died soon thereafter, in the fighting in France. In 1919 Landau edited Kornblum's manuscript and finished it for publication.

Kornblum had proved that there are infinitely many primes $P \in K[x]$ satisfying $P \equiv A \mod C$ where A, C are given polynomials without common divisor. [15] For this purpose he had used the classical idea of Dirichlet (1837) and transferred it from $\mathbb{Z}$ to $K[x]$. He had introduced the L-series $L(s|\chi)$ for arbitrary characters χ of $K[x]$ modulo C:

$$L(s|\chi) = \prod_{\gcd(P,C)=1} \frac{1}{1 - \chi(P)t^{\deg P}} = \sum_{\gcd(M,C)=1} \chi(M)t^{\deg M} \tag{15}$$

where P ranges over monic prime polynomials and M over arbitrary monic polynomials $\neq 0$. Kornblum proved that $L(1|\chi) \neq 0$ for every character $\chi \neq 1$. From this it is standard to deduce by Dirichlet's method that in any given residue class A modulo C with $\gcd(A, C) = 1$, not only there exist infinitely many primes but also that the set of these primes has (Dirichlet) density $\frac{1}{\Phi(C)}$, where $\Phi(C)$ denotes the number of relatively prime residue classes in $K[x]$ modulo C. It turned out that in the function field case the analytic situation is much easier to handle than in the classical number field case since now, as Kornblum observed, $L(s|\chi)$ is a rational function in the variable $t = q^{-s}$, and it is a polynomial if $\chi \neq 1$. We had already mentioned this in the foregoing section.

In Artin's thesis Kornblum's result is improved by showing:

> *In each relatively prime residue class A modulo C there exist primes P of arbitrary given degree n, provided n is sufficiently large. And these P are asymptotically equally distributed among those residue classes.*

The main point in the proof is to show for $\chi \neq 1$ not only that $L(1|\chi) \neq 0$ (like Kornblum) but the stronger result that $L(s|\chi)$ has no zero on the line $\Re(s) = 1$.

[14] Note that Artin works in characteristic > 2. At the time of MacRae's paper [MacR:1971] it was possible to handle also the case of characteristic 2. MacRae found that in characteristic 2 there are precisely three “imaginary” quadratic fields with $h = 1$. See also footnote 69.

[15] Kornblum considers this problem over $\mathbb{F}_p$ only, similar as Artin does in his thesis. But it is “self-evident” that the same results hold over an arbitrary finite base field K.

(Since $L(s|\chi)$ is a periodic function this implies that there exists a whole strip $1-\epsilon < \Re(s) < 1+\epsilon$ which does not contain a zero of $L(s|\chi)$.) If χ is of order > 2 then χ is different from its complex conjugate, and it was possible just to copy the known proof presented in Landau's book from the number field case; Artin cites Landau's "*Handbuch*" [Lan:1909]. The main difficulty arises, as in the number field case, if χ is a quadratic character. In this case Artin shows that there is a square free discriminant D such that $\chi = \chi_D$ in our above notation. It follows that the Kornblum L-series $L(s|\chi)$ coincides (up to finitely many Euler factors) with the L-series which was identified by Artin as a factor of the zeta function of the quadratic field $K(x, \sqrt{D})$; see (9). Hence Artin could apply his result **1.** above.

Finally, the proof of Artin permits to estimate the error term in the asymptotic. Let $S_n(A, C)$ denote the number of primes P of degree n with $P \equiv A \mod C$. Then Artin shows

$$S_n(A, C) = \frac{1}{\Phi(C)} \frac{q^n}{n} + O\left(\frac{q^{n\theta}}{n}\right) \qquad \text{with} \qquad \frac{1}{2} \leq \theta < 1 \tag{16}$$

for $n \to \infty$. Compare this with (14). Today, as a consequence of the Riemann hypothesis, we can take $\theta = \frac{1}{2}$.

A presentation of this material from today's viewpoint can be found in the book of Rosen [Ros:2002].

2.3 Artin's letters to Herglotz

The legacy of Gustav Herglotz is kept at the university library in Göttingen. It contains several letters from Artin. In two of those letters, written in November 1921, Artin is explaining some further ideas and results in connection with his thesis. These letters became known to the mathematical community thanks to Ullrich's paper [Ul:2000]. It seems not without interest to discuss some of the contents here again, because we can observe in a nutshell already several ideas which later became important for the proof of the Riemann hypothesis.

The legacy of Emil Artin is kept at the university library in Hamburg; see [Ul:2000a]. Recently Ullrich has found therein two versions of a manuscript by Artin in which he included his ideas mentioned in the letters to Herglotz. One version is written in a form suited for publication but Artin did not publish this manuscript. It has been published recently [A:2000], edited by Ullrich. Thus our comments to Artin's letters can also be read as comments to Artin's paper [A:2000].

The letters are sent from Göttingen. Artin had obtained his Ph.D. at Leipzig on June 20, 1921 and now has moved to Göttingen for further studies. In his letters he reports to his former academic teacher about his impressions in the new environment. He had made his visits to Courant, Hilbert, Klein and Landau [16], and Hilbert had invited him to give a talk at the *Mathematische Gesellschaft*. Artin reports about the seminars which he is attending. But otherwise he feels quite lonely:

> "*Unfortunately, I have here very little contact to the lecturers, and therefore I am missing the personal stimulus which in Leipzig I have*

[16] We note that Emmy Noether is not mentioned. It seems that she was not considered to be one of the dignitaries whom a young newcomer had to pay a formal visit.

had to such high degree by yourself, Herr Professor. *For this I will always be grateful to you.*" [17]

But then he proceeds to report about his work, following the lines set in his thesis.

2.3.1 Extension of the base field

Artin finds it self-evident that the whole theory in his thesis, although presented over $\mathbb{F}_p$ as the base field, remains valid over every finite base field K. This we have mentioned already. [18] Now, in a letter dated Nov 13, 1921, he studies base field extensions in a systematic way.

Let $F = K(x, \sqrt{D})$ be a quadratic function field with a finite base field K, and $\deg D > 0$. Consider the extension $K^{(m)}$ of K of degree m and let $F^{(m)} = K^{(m)}(x, \sqrt{D})$ be the corresponding constant field extension of F. Let $R, R^{(m)}$ be the integral closures of $K[x]$ in F, and of $K^{(m)}[x]$ in $F^{(m)}$ respectively. The corresponding zeta functions are denoted by $\zeta(s)$ and $\zeta^{(m)}(s)$. Artin had found a connection between $\zeta(s)$ and $\zeta^{(m)}(s)$. His formula is:

$$\zeta^{(m)}(s) = \prod_{0 \le \mu \le m-1} \zeta\left(s + \frac{2\pi i \mu}{m \log q}\right) \tag{17}$$

For the corresponding L-polynomials it follows that the roots of $L^{(m)}(t)$ are precisely the m-th powers of the roots of $L(t)$. Artin concludes:

> " *1.) If the Riemann hypothesis holds for $\zeta(s)$ then also for $\zeta^{(m)}(s)$. Hence with every zeta function there is a whole bunch of infinitely many other zeta functions for which the Riemann hypothesis holds. (In every $F^{(m)}$ with arbitrary m.)*
>
> *2.) Conversely: If the Riemann hypothesis is proved for $\zeta^{(m)}(s)$ then also for $\zeta(s)$. Hence it is sufficient to prove the Riemann hypothesis over all finite fields K for those discriminants only which split into linear factors. Over every such field there are only finitely many cases to be dealt with.*"

In the next letter he goes one step further and obtains the following conclusion. Recall the notation in (9) where σ_1 denotes the first coefficient of the polynomial $L(t) = 1 + \sigma_1 t + \cdots$. Artin derives from (17):

> "*. . . for the proof of the Riemann hypothesis it is only necessary to have the following 'raw' estimate of σ_1, but over all Galois fields:*
>
> $$|\sigma_1| < A\sqrt{q}$$
>
> *where A depends only on the degree of the discriminant.*"

[17] Constance Reid reports in her book [Rei:1976] that Herglotz had little contact with his students. If this was the case then it seems that his relation to Artin was an exception.

[18] A sizable part of Artin's unpublished manuscript [A:2000] is devoted to develop the theory of finite fields from scratch. It seems that Artin considered the generalization to arbitrary finite base fields not as evident for others as he did for himself.

It is immediate from the definition in (9) that

$$N_R = q + \sigma_1 \tag{18}$$

where N_R is the number of prime ideals of degree 1 in the ring R. Thus we may write the above inequality as

$$|N_R - q| < A\sqrt{q}\,. \tag{19}$$

We see that now Artin begins to think about a general proof of the Riemann hypothesis, more than just numerical verification in examples. In fact, the inequality (19) has become crucial in all later setups for the proof of the Riemann hypothesis. It is remarkable that Artin at this early stage is fully aware of the importance of (19).

Many years later, when Hasse started to work on the Riemann hypothesis, then he consulted Artin and obtained from him the information about (17), perhaps also (19) . It was clear to both of them that this was not restricted to Artin's zeta functions of quadratic function fields, but could also be applied to the F. K. Schmidt zeta functions of arbitrary function fields. In this way the results of Artin's work became part of the foundations for the proof of the Riemann hypothesis in general. See section 6.3.

REMARK: When I wrote my paper [Rq:1997] I did not yet know of the letters from Artin to Herglotz. In that paper I had contributed to Hasse the idea, that for the proof of the Riemann hypothesis one should consider not only a fixed base field K, but at the same time all finite extensions of K. Now we see that this idea had already appeared earlier in a letter from Artin to Herglotz.

2.3.2 Complex multiplication

Suppose the discriminant $D = D(x)$ is cubic. Then all the zeros of the polynomial $L(t) = 1 + \sigma_1 t + qt^2$ are non-trivial. The "inverse" polynomial

$$L^*(t) = t^2 + \sigma_1 t + q \tag{20}$$

admits the Frobenius endomorphism π as its root. The Riemann hypothesis is equivalent to the inequality

$$\sigma_1^2 - 4q < 0 \tag{21}$$

which is to say that the ring $\mathbb{Z}[\pi]$ is imaginary.

Now, Artin does not talk about the Frobenius endomorphism and not about endomorphisms at all. But he puts himself the question how to describe the connection between the elliptic function field $F = K(x, \sqrt{D})$ over a finite field K of order q, and the square free kernel $d \in \mathbb{Z}$ of $\sigma_1^2 - 4q$, so that $\sigma_1^2 - 4q = u^2 d$ with $u \in \mathbb{Z}$. It appears that he hoped to find some way to deduce that $d < 0$ and, hence, to prove the Riemann hypothesis at least in the case of a cubic discriminant D [19]. Later, when Hasse was going to prove the Riemann hypothesis for elliptic function fields (i.e., for cubic discriminants), one of the essential steps in his proof was to show generally that the endomorphism ring of F is imaginary quadratic. [20]

[19] Artin writes $\sigma_1^2 - 4q = -u^2 d$ and so he is looking for a proof of $d > 0$.

[20] Hasse discovered that in characteristic p there are some exceptional cases where this ring is non-commutative; then every maximal commutative subring is imaginary quadratic.

It seems remarkable that Artin in 1921, without referring to complex multiplication, poses the problem in this way. Now we can understand that in 1934, when he invited Hasse for a lecture course in Hamburg, he was so enthusiastic about Hasse's work: for Hasse could do what Artin had seen in 1921 but was not able to prove in general. Probably Hasse did not know about Artin's 1921 letter to Herglotz. In that letter, Artin had written:

> *The big question is, which d belongs to the given discriminant D? In general I know nothing about this. But in two cases I can decide this and will now report on it.*

And he proceeds to discuss the following two cases:

$$\begin{aligned} D &= x^3 - x \qquad \text{if} \qquad q \equiv 1 \mod 4 \\ D &= x^3 - 1 \qquad \text{if} \qquad q \equiv 1 \mod 6\,. \end{aligned}$$

For the proof of the Riemann hypothesis, the congruence conditions for q are not essential because of what Artin has just mentioned: One may replace the base field K by its extension of order 2, hence q by q^2, which then satisfies the congruence condition. (The powers of $p = 3$ have to be excluded in the second case). Also, one may replace $x^3 - x$ by $x^3 - bx$ with $b \in K$, because after a suitable extension of K this can be transformed easily into the form as given. Similarly, $x^3 - 1$ may be replaced by $x^3 - b$.

In those two cases Artin succeeds to determine $d = -1$ and $d = -3$ respectively. And so he can proudly announce:

> *Accordingly one has the Riemann hypothesis for all q, and all $D = x^3 - bx$ and $D = x^3 - b$.*

Note that this result is of quite different nature from the numerical verification which Artin gave in his published thesis, because it is valid for *all* q whereas in the thesis only small prime numbers are involved.

We observe that the two cases which Artin had chosen are fields of absolute invariant $j = 0$ and $j = 1$, and these are known to have endomorphism ring $\mathbb{Z}[i]$ and $\mathbb{Z}[\varrho]$ respectively, where i is a primitive 4-th and ϱ a primitive 3-rd root of unity. In these cases the endomorphism ring has units of order > 2. But it seems that Artin at that time was not aware of the connection to complex multiplication since he did not mention it in his letter. But then, how did he proceed to obtain $d = -1$ and $d = -3$ in those two cases?

He did it with what he called a trick. Consider $D = x^3 - x$ with $q \equiv 1 \mod 4$. Artin compares the field $F = K(x, \sqrt{D})$ with the field $\overline{F} = K(x, \sqrt{\overline{D}})$ where $\overline{D} = x^3 - bx$, and b is not a square in K. These fields are not isomorphic but they become isomorphic after base field extension from K to $K^{(4)}$, the extension of degree 4 over K. Let $\overline{\sigma}_1$ be the first coefficient of the L-polynomial for $\overline{F}$, in the same way as σ_1 is defined for F. Then Artin compares $F^{(2)}$ and $\overline{F}^{(2)}$, the base field extensions of degree 2. Using (17) he finds the following relation:

$$\sigma_1^2 + \overline{\sigma}_1^2 = 4q \tag{22}$$

which shows that, indeed, $\sigma_1^2 - 4q = -\overline{\sigma}_1^2$ and thus $d = -1$. [21]

[21] A closer look at this "trick", which today could be interpreted in the framework of Galois cohomology, will reveal that it depends on the existence of automorphisms of F of order > 2. Thus the two examples which Artin used are essentially the only ones for which it works.

By the very defintion of σ_1 and $\overline{\sigma}_1$ Artin obtains the following cute side result in the case $q = p$ is a prime number: Let $p \equiv 1 \mod 4$. Then: *If b is a non-square modulo p and if we put*

$$\alpha = \sum_{0 \le \nu \le \frac{p-1}{2}} \left(\frac{\nu^3 - \nu}{p} \right) \qquad \text{and} \qquad \beta = \sum_{0 \le \nu \le \frac{p-1}{2}} \left(\frac{\nu^3 - b\nu}{p} \right) \tag{23a}$$

then

$$p = \alpha^2 + \beta^2 \,. \tag{23b}$$

Here, the brackets denote the ordinary quadratic residue symbol in $\mathbb{Z}$.

Thus Artin's arguments lead to an explicit algorithm to write any $p \equiv 1 \mod 4$ as a sum of two squares. He writes, however:

> "*... unfortunately, this result is not new, it had been found by Jacobsthal using pure computation, as I have been informed some days ago.*" [22]

But, he added, by doing a similar trick for the discriminant $D = x^3 - 1$ he had obtained the following which he believed was new: Let $p \equiv 1 \mod 6$. Then: *If b is not a cubic residue modulo p and if we put*

$$\alpha = \sum_{0 \le \nu \le p-1} \left(\frac{\nu^3 - 1}{p} \right) \qquad \text{and} \qquad \beta = \sum_{0 \le \nu \le p-1} \left(\frac{\nu^3 - b}{p} \right) \tag{24a}$$

then

$$3p = \alpha^2 - \alpha\beta + \beta^2 \,. \tag{24b}$$

From this one can easily obtain a representation of p by the quadratic form $X^2 - XY + Y^2$.

2.3.3 Birational transformation

Artin considers applying a linear fractional transformation to x. He replaces x by some $x' = \dfrac{ax+b}{x-c}$ with $a, b, c \in K$. Let R' denote the integral closure of $K[x']$ in F. Artin gives a formula expressing the ζ-function belonging to R' in terms of the ζ-function belonging to R. We do not write down this formula because it is quite obvious, exchanging the Euler factors belonging to the infinite primes with respect to x, with the Euler factors belonging to the infinite primes of x'. Artin remarks:

> "*If D is divisible by $x - a$ and of even degree then the degree will decrease. Since D will obtain a linear factor after suitable increase of the base field K, one may assume the degree to be odd. This decrease of degree looks very similar to hyperelliptic integrals.*"

By "degree" he means the degree of the discrimimant D. At the same time, the degree of the corresponding L-polynomial, which is one less than the degree of D, will also decrease.

[22] Artin does not give any reference. A search in the *Jahrbuch* database reveals that the result is contained in Jacobsthal's Berlin thesis [Ja:1906]. The publication appeared 4 years later in Crelle's Journal [Ja:1910].

By considering such transformations, Artin gets rid of the limitations which are set by the attempt to get complete analogy to the number field case. For, in the number field case the ring $\mathbb{Z}$ is given and cannot be changed; if one insists on complete analogy then in the function field case, the polynomial ring $K[x]$ should be fixed and should not be changed. By leaving this restriction and considering also linear fractional transformations, Artin starts to use fully the possibilities which function fields offer, which are not available in number fields.

In this way Artin can simplify the problem by reducing it to the case where the degree of D is odd. Then $\sqrt{D}$ is "imaginary" and remains so after every base field extension; thus for the proof of the Riemann hypothesis it is not necessary to consider the "real" case any more. Although Artin does not mention it, he certainly knew that if $\deg D$ is odd then all the zeros of the polynomial $L(t)$ are nontrivial, and they are birational invariants of the field F.

This is the first step towards a birational invariant definition of the zeta function, which later was given by F. K. Schmidt.

2.4 Hilbert and the consequences

By all what Artin does and says in his thesis and subsequently in his letters to Herglotz, it is evident that he considers the Riemann hypothesis to be important for the theory of arithmetic of function fields in general. Surely, there is a long way to go from the knowledge that (19) would be sufficient, to an actual proof of the Riemann hypothesis. We shall see that it required an enormous effort until this goal was reached. But certainly (19) is to be considered as a first step.

Why did Artin not proceed further in this direction? We would expect him to continue studying function fields and their arithmetic, with the aim of further approaching the Riemann hypothesis. However, none of his later papers are dealing with this topic.

Certainly, Artin continued to be interested in the subject and observed keenly the further development; this can be inferred, e.g., from his letter to Hasse cited in the introduction of this article. In his lectures on number theory he tried to include function fields. He was striving for a unified theory for number fields and function fields, which later was established as the theory of "global fields". But in his publications Artin kept silent on the topic of function fields and their zeta functions. This is evidenced by the editors of his "Collected Papers" [A:1965] who put his thesis quite singularly into a separate chapter, thus indicating that the thesis is somewhat isolated among the other papers of Artin, none of which can be regarded to be closely related to the thesis.

Can we find an explanation for this sudden change in Artin's work?

One of the reasons may perhaps be traced back to Artin's experience when he gave a talk, on the invitation of Hilbert, in the *Mathematische Gesellschaft* in Göttingen. The subject of the talk was his thesis and the additional ideas set forth in his letters to Herglotz. Hilbert had severely criticized Artin during this talk.

Ullrich in [Ul:2000] has narrated this story. Artin wrote to Herglotz that Hilbert had spoiled all his enthusiasm for his work by this criticism. Artin does not want to stay with his topic, and he will chuck it in. Although Hilbert, some days later, changed his opinion and offered publication of Artin's manuscript in the *Mathematische Annalen*, Artin did not accept.

As Ullrich [Ul:2000] reports, Herglotz was able to arrange that Artin could move from Göttingen to Hamburg where he was offered by Blaschke a position as assistant. [23] There, in the very neighborhood of Hecke, he found the mathematical atmosphere suitable for him. Soon he rose to one of the leading figures in algebra and number theory. But for the Riemann hypothesis in function fields, his thesis remained his only published contribution.

Summary

Artin's thesis is concerned with quadratic function fields over finite base fields. Artin goes about to transfer the arithmetic theory of quadratic number fields to the case of quadratic function fields. This includes, first the arithmetic properties like decomposition into prime ideals, theory of units, class number etc., and secondly the analytic theory of the zeta function of a quadratic field. Artin observes that the zeta function is a rational function of $t = q^{-s}$, and it is a polynomial in t up to a trivial factor. Artin derives the class number formula in terms of the zeros of this polynomial, as well as the formula for the distribution of prime ideals. He formulates the Riemann hypothesis for the non-trivial zeros of that polynomial, and he verifies it numerically in a number of cases.

After having received his degree, Artin spent a year as post-doc in Göttingen. In several letters from there to his academic teacher Herglotz, Artin develops some further results in which we can observe the nuclei of what later will become essential features in the general proof of the Riemann hypothesis. But these results were never published. The reason may be 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. But his thesis remained to be a landmark for the development of the theory of function fields which followed. He continued to be keenly interested in this development, but did not publish anything more in the direction of the Riemann hypothesis.

3 Herglotz: Gauss' last entry

There is one paper [Her:1921] of Herglotz which directly concerns the Riemann hypothesis for function fields and therefore has to be mentioned here. That paper appeared in the same year when Artin had submitted his thesis, and its title is: "*On the last entry in Gauss' diary.*" This concerns the integer solutions of the congruence

$$x^2y^2 + x^2 + y^2 \equiv 1 \mod p \tag{25}$$

for a prime number $p \equiv 1 \mod 4$. Gauss, in his diary dated July 9, 1814, had noted the number of solutions as follows:

[23] We would like to use this opportunity to point out that it was Blaschke, the first mathematician at the newly founded University of Hamburg, who succeeded to raise this place within a few years to one of the leading mathematical centers in Germany. He did this through a careful *Berufungspolitik*. The Hamburg Mathematical Seminar in its first decades is a good example that mathematical excellence cannot be created by more money or more positions only, but that the decisive point is to attract excellent people.

In the ring $\mathbb{Z}[i]$ of Gaussian integers decompose $p = \pi\overline{\pi}$ as a product of an "imaginary" prime number π and its conjugate $\overline{\pi}$. We may identify the factor rings $\mathbb{Z}[i]/\pi = \mathbb{Z}/p$, and hence the solutions of (25) in $\mathbb{Z}$ (mod p) are the same as the solutions in $\mathbb{Z}[i]$ (mod π) of

$$x^2y^2 + x^2 + y^2 \equiv 1 \mod \pi\,. \tag{26}$$

Now, Gauss had found that the number of solutions of (26) equals the norm

$$N(\pi - 1) = (\pi - 1)(\overline{\pi} - 1) = p - S(\pi) + 1\,, \tag{27}$$

under the following specifications:

1. The 4 infinite solutions $(x, y) \mapsto (\infty, \pm i)$, $(x, y) \mapsto (\pm i, \infty)$ have to be included in the counting.
2. π has to be normalized such that $\pi \equiv 1 \mod (1 - i)^3$.

This normalization can be achieved by multiplying π with a suitable unit $\varepsilon \in \{\pm 1, \pm i\}$. After such normalization, π is called "primary". Note that $1 - i$ is a Gaussian prime number dividing 2.

Gauss did not actually give a proof, he had observed this fact "by induction" which according to the terminology at his time meant either heuristically, or experimentally for several p.

The story of this " last entry" is well known; see, e.g., the presentation in Lemmermeyer's book [Lem:2000] and the literature cited there. Herglotz' presents a proof of Gauss' statement. Before discussing this, let us first point out the connection of Gauss' statement to the Riemann hypothesis for the function field $F = K(x, y)$ where where x, y satisfy the lemniscate equation

$$x^2y^2 + x^2 + y^2 = 1\,. \tag{28}$$

Thereby $K = \mathbb{F}_p$ is assumed to be the prime field.

F is a quadratic extension of the rational function field $K(x)$, and from (28) we obtain

$$y^2 = \frac{1 - x^2}{1 + x^2}\,. \tag{29}$$

This relation is satisfied by the 4 infinite solutions mentioned above in 1. This explains why these are also counted as solutions of (26). Seen in this way, Gauss' statement says that $N(\pi - 1)$ is *the number of all places of degree* 1 *of* F; this number is independent of the choice of generators x, y of F.

If we multiply the above relation with $(1 + x^2)^2$ we find

$$F = K(x, \sqrt{1 - x^4}) \tag{30}$$

which is of the form Artin had discussed in his thesis, with $D = 1 - x^4$. We have seen in section 2.3.3 that Artin admits also linear fractional transformations in order to reduce the degree of the discriminant D. Such a degree reduction is obtained by putting $x = \frac{1+ix'}{1-ix'}$. We compute

$$F = K(x', \sqrt{x'^3 - x'}) \tag{31}$$

This is precisely the form which Artin had discussed in his letter to Herglotz on Nov 30, 1921; see section 2.3.2. In that letter Artin had presented his proof of the Riemann hypothesis for this field F.

But the validity of the Riemann hypothesis for this field is also immediate from Gauss' statement, as follows:

As Artin had pointed out in his letter to Herglotz, the Riemann hypothesis is equivalent with the the inequality (21) which, according to (18), can be written as

$$|N' - p| < 2\sqrt{p}\,. \tag{32}$$

Here, N' denotes the number of prime ideals of degree 1 in the integral closure R' of $K[x']$ in F. Every such prime ideal defines a place of F of degree 1. There is precisely one more place of F of degree 1, and this belongs to $x' \mapsto \infty$. (This is so since the degree of $D' = x'^3 - x'$ is odd.) Thus the number of all places of F of degree 1 is $N' + 1$. Hence, using Gauss' statement, we obtain $N' + 1 = N(\pi - 1) = p - S(\pi) + 1$. We conclude that $N' - p = -S(\pi)$. Writing $\pi = a + bi$ we have $S(\pi) = 2a$ and since $p = a^2 + b^2$ we see that $S(\pi)^2 = 4a^2 < 4p$ which gives (32).

Conversely, from Artin's results as explained in his letter to Herglotz, it is straightforward to deduce Gauss' statement (27) for a Gaussian prime π dividing p, without however the normalization in 2. which is a more subtle affair.

In any case, we see that between the results of Herglotz and those of his disciple Artin there is a very close connection indeed. The question arises why none of them cites the work of the other.

As pointed out above, Herglotz' paper [Her:1921] appeared in the same year as when Artin submitted his thesis. It is conceivable that Herglotz had completed his work before Artin started with his thesis. And that Herglotz, because he had seen the connection with the Riemann hypothesis for one particular quadratic function field, had proposed to Artin to look into this question quite generally, for arbitrary quadratic function fields. But then Artin should have cited Herglotz' paper as the one where his work started. Since he did not, and since he explained the case in such detail in his letter to Herglotz, we conclude that Artin did not know Herglotz' paper at that time. Why not? Why did Herglotz not show his paper to Artin? The theory of complex multiplication which Herglotz used, was to become in the future, in the hands of Hasse, a powerful tool for the proof of the Riemann hypothesis for arbitrary elliptic function fields. If Herglotz would have foreseen the importance of complex multiplication for the Riemann hypothesis then he would probably have proposed to Artin to study that theory, with the aim to apply this to the proof of the Riemann hypothesis. But apparently he did not.

Another possibility would be that Herglotz wrote his paper later, after he had seen Artin's thesis. Inspired by Artin's work he may have realized that Gauss' last entry would imply the validity of the Riemann hypothesis for the function field of the lemniscate, and so he set to work on it. In this case Herglotz would have cited Artin and shown the connection between his paper and the work of his disciple Artin. But he did not.

We have no obvious explanation for the fact that none of them, Herglotz and Artin, cited the other.

As to the contents of Herglotz paper [Her:1921]: He first presents the statement (27) of Gauss' last diary entry, and then reports (without giving any reference) that Dedekind had verified Gauss' statement for all $p < 100$, and that Fricke had pointed out the coincidence of the equation (28) with the equation which is satisfied by the lemniscate functions of Gauss:

$$x = \varphi(u) = \sin \operatorname{lemn}(u) \qquad \text{and} \qquad y = \psi(u) = \cos \operatorname{lemn}(u)\,.$$ [24]

Herglotz continues:

> *In addition we shall remark here that the solutions of* (25) *coincide precisely with the congruence solutions modulo* π *of the division equations for*
> $$x = \varphi\left(\alpha \frac{\omega_3}{\pi - 1}\right) \qquad \text{and} \qquad y = \psi\left(\alpha \frac{\omega_3}{\pi - 1}\right)\,. \tag{33}$$ [25]

After this introduction Herglotz starts with his calculations. The division equation is used in a form which Weierstrass had given, and the greater part in Herglotz' calculations seems to be catching the connection to the notations which Weierstrass had used. [26] The essential feature in these calculations is the observation that at a certain point there appears an Eisenstein equation with respect to the prime π and therefore, modulo π, the result follows.

The normalization $\pi \equiv 1 \bmod (1-i)^3$ implies that the multiplication with π of the arguments in (33) induces the map $(x, y) \to (x^p, y^p) \bmod \pi$, which is to say the Frobenius map of the reduced curve. This is implicitly used in Herglotz' computations.

We observe that Herglotz never says explicitly that Gauss' statement follows from his results. He seems to consider it as self-evident that the number of (x, y) in (33) is $N(\pi - 1)$ when α ranges over $\mathbb{Z}[i]/(\pi - 1)$. Indeed this belongs to the basics of complex multiplication and we have no doubt that Herglotz regarded it as such. He was a very knowledgable mathematician, well acquainted with the old masters of the science. When he not explicitly claimed to have proved the statement in Gauss' last entry, then because in his eyes this was self-evident from what he really proved. His style used to be very concise – or maybe we should say "minimal" if it comes to explanations for the reader.

This may have been the reason why the paper of Herglotz was not properly appreciated by the mathematical public of the time. The paper was reviewed by Gábor Szegö in the "*Fortschritte der Mathematik*", and he just repeats the author's claim that that the solutions of (25) coincide precisely with the congruence solutions modulo π of the division equations for the Gauss' lemniscate functions. No mention of the fact that this implies the validity of Gauss' statement. The same we can observe in the commentaries in Herglotz' "Collected works" [Her:1979]: The article on Gauss' last entry was commented on by Theodor Schneider, and again he does not mention that the validity of Gauss' entry follows from Herglotz' result.

When in the year 1933, Hasse used complex multiplication to prove the Riemann hypothesis for all elliptic function fields, he did not cite Herglotz' paper although Herglotz had used the same method, namely complex multiplication, in the special case of the lemniscate function field. And in the joint paper of

---

[24]) The function sin lemn is defined as the inverse function of $\int_0^x \frac{dx}{\sqrt{1-x^4}}$, and accordingly cos lemn by means of (29).

[25]) Herglotz uses the Weierstrass notation, namely: $\omega_3 = -(1+i)\omega$ where ω is the primitive real period of $\varphi(u)$. In the above equation, $\alpha \in Z[i]$ is arbitrary.

[26] We have not checked these calculations. Schappacher [Sch:1997] writes that "*Herglotz uses the Weierstrass theory, albeit with notation that has not quite survived to the present day.*"

Hasse and Davenport [Da-H:1934] we find an extra section where a particularly simple proof of the Riemann hypothesis for the field $K(x, \sqrt{1-x^4})$ is given; in view of (30) this is precisely the field of the lemniscate. Again, Herglotz' paper was not cited. Probably Hasse (and Davenport too) did not know it.

So we see that Herglotz' paper was not widely known among the mathematicians of the time, at least it was not realized that it yields a proof of Gauss' last entry. After many years only this consequence of Herglotz' paper was rediscovered by Deuring [Deu:1941]. That paper was dedicated to Herglotz on his 60th birthday 1941. Deuring mentions Hasse's work on the Riemann hypothesis for elliptic fields and then writes:

> "*Apparently the mathematicians who work on these problems did not notice that this theorem about congruences of genus* 1 *had been known to Gauss already, at least in the case of lemniscatic functions. . . Finally Herglotz has proved Gauss' statement. His method, namely to use the division of elliptic functions by* $\pi - 1$*, is the same, in principle, which Hasse had used in his first paper on the Riemann hypothesis.*"

Of course, Deuring knew the theory of complex multiplication well; after all he was the one who, following Hasse's lines, completely remodelled the theory replacing the analytic framework with a purely algebraic one. [27] Thus for him it was obvious that Herglotz' results imply the validity of Gauss' last entry. And so, Herglotz' paper came to be appreciated belatedly as a forerunner of Hasse's general theory of elliptic function fields and the Riemann hypothesis.

But even so, we have no explanation why Herglotz himself did not mention his own paper, not to Hasse and not to anyone around Hasse.

On January 10, 1933 Hasse delivered a colloquium talk in Göttingen at the *Mathematische Gesellschaft*. At that time Herglotz was in Göttingen already, having accepted in 1925 an offer from Göttingen as the successor of the applied mathematician C. Runge. It seems to us very probable that Herglotz, being a member of the *Mathematische Gesellschaft*, attended Hasse's talk where Hasse presented his new view of the Riemann hypothesis in connection with the work of Davenport and Mordell. [28] Would it not be natural that in the ensuing discussion, Herglotz would mention that he had proved the Riemann hypothesis for the lemniscate, in his paper on Gauss' last entry? Later in the same year, on December 10, 1933, Hasse gave another talk in Göttingen; at this time he had already obtained his proof of the Riemann hypothesis for arbitrary elliptic curves. This time we know for certain that Herglotz attended the talk for he, together with F. K. Schmidt, had invited Hasse to Göttingen. [29] And again he did not mention his own paper. If he would have then Hasse, who generally

[27] See the papers [Deu:1949] and [Deu:1952] which appeared in 1949 and 1952 respectively. From a letter of Deuring to Hasse dated June 10, 1937 we infer that already in 1937 Deuring was in the possession of these results. And in February, 1939 he gave a talk on this at the University of Hamburg.

[28] REMARK: In an earlier version of this Part 1 we said that Hasse presented in this talk his proof of the Riemann hypothesis in the elliptic case. But now we know from the Hasse-Davenport correspondence that he obtained his proof at the end of February only. We will discuss this in Part 2.

[29] Herglotz had offered that Hasse may stay in his house during the time of his visit. We do not know whether Hasse accepted this offer.

was very careful in his citations, would certainly have mentioned it in one of his papers.

In the summer of 1934 Hasse accepted a position in Göttingen and from then on the two, Hasse and Herglotz, worked as colleagues at the same institution. It is inconceivable that Herglotz was not informed about Hasse's work at that time. Then, why did he keep silent about his own work, realizing that Hasse did not know it?

We have no explanation, except perhaps that the years of Herglotz' activities in number theory had passed, and his interest had shifted to partial differential equations and differential geometry.

Summary

In 1921, the same year when Artin submitted his thesis, Herglotz published a note containing a proof of the statement in the last entry of Gauss' diary of 1814. This concerned the number of solutions modulo p of the lemniscate equation, for $p \equiv 1 \mod 4$. *It implies the validity of the Riemann hypothesis for the function field of the lemniscate. But Artin does not cite Herglotz' paper in his thesis, and it appears that he did not know it.*

The method Herglotz used belongs to the theory of complex multiplication. Later in 1933, when Hasse obtained the first proof of the Riemann hypothesis for elliptic function fields, Hasse used the same method as Herglotz had in the special case of the lemniscate function field. But Hasse did not cite Herglotz' paper and it appears he did not know it either. And again, Herglotz did not tell him about it. Quite generally, Herglotz' paper seems not to have been widely known at that time, and it was 1941 only that Deuring brought Herglotz' paper to the attention of the mathematical public.

4 Building the foundations: The arithmetic part

After Artin's thesis a new development started: the systematic investigation of algebraic function fields with finite base fields, exploiting the analogy to algebraic number fields. Artin had dealt with quadratic function fields only, and now there was rising interest to deal with arbitrary function fields as well.

Of course the deep analogy between number fields and function fields had been observed since a long time already, and had inspired a number of the leading mathematicians in the 19th century. [30] But before Artin, algebraic function fields were mostly considered over the field $\mathbb{C}$ of complex numbers. In that case, function fields could be envisaged as consisting of "functions" in a sense which people were used to, i.e., functions on a compact Riemann surface or multi-valued functions over the complex plane, with values in $\mathbb{C}$ (including ∞). The transition from this to the abstract setting, where fields were just mathematical structures satisfying certain axioms, started after the seminal paper of Steinitz on field theory, published in 1910. After that it became gradually standard to define function fields in an abstract way over any field K as base field. Finite base fields and the real and complex base fields thus appear as special cases

[30] See the paper [Ul:1999] by Ullrich for a discussion of the history of this analogy.

of a general theory. But it took some time until this viewpoint was generally accepted. [31]

Certainly, there had been earlier instances where function fields over finite fields were investigated, like Dedekind's paper [De:1857] which we already mentioned in section 2.1. [32] But then finite fields were considered as being constructed by means of "higher congruences", not in an abstract way. It was Artin's thesis which changed this in respect to the theory of function fields. He adopted the abstract point of view in the sense of Steinitz. This provided the mathematician with greater freedom and flexibility in dealing with the theory of function fields. There followed an extended activity, in order to better understand the close analogy between number fields and function fields with finite base fields.

These activities had a twofold aim: First, there was to be built the *arithmetic theory* of function fields with finite base fields. Secondly, resting on the arithmetic theory, the *analytic theory* was to be created, dealing with L-functions and zeta functions in analogy to Dirichlet and Dedekind. These two parts correspond to the two parts of Artin's thesis.

In this section 4 we will discuss the arithmetic theory. The analytic theory will follow in section 5.

4.1 The fundamental theorems

Consider the following situation:

K a finite fields with q elements,
F a finite field extension of the rational function field $K(x)$,
R the integral closure of the polynomial ring $K[x]$ in F.

Then the following statements are regarded as the fundamental theorems of arithmetic in F.

(i) **Prime ideal decomposition:** *R is a Dedekind ring.*
(ii) **Ramification:** *Let $P \in K[x]$ be a prime polynomial. P is ramified in R if and only if P divides the discriminant D of R over $K[x]$ – provided F is separable over $K(x)$.*
(iii) **Unit theorem:** *The unit group $R^\times$ modulo its torsion group is a free abelian group with $r-1$ generators, where r denotes the number of poles of x in F.*
(iv) **Class number:** *The ideal class group of R is finite.*

These statements are generalizations of the corresponding theorems of Artin for quadratic function fields, see section 2.1.1. (There, we had seen that Artin's thesis contained two more theorems, concerning the decomposition law of primes and the quadratic reciprocity law. These do not have a direct counterpart in the

[31] Artin [A:1950] has pointed out, however, that there was a difference in attitude between American and European authors. From the very beginning, he said, the abstract point of view was dominant in American publications on algebra.

[32] Recently Günther Frei has discovered that the theory of function fields with finite base fields had already been started by Gauss in his unpublished "Chapter 8" of the *Disquisitiones arithmeticae*. See [Fr:2001a].

general situation here. We shall discuss their generalization to cyclic extensions in section 4.4.2, in connection with F. K. Schmidt's thesis.)

Within a few years after Artin's thesis there appeared three papers, independent of each other, dealing with the above theorems. The young authors were

P. Sengenhorst:	Ph.D. 1923 in Göttingen with Landau,
F. K. Schmidt:	Ph.D. 1925 in Freiburg with Krull,
H. Rauter:	Ph.D. 1926 in Halle with Hasse.

Before discussing their papers in detail, let us make a few general remarks on the four statements (i)-(iv) above.

4.1.1 Ideal theory

Today, for the Dedekind property (i) we would just refer to Emmy Noether's seminal paper [Noe:1926] which characterized Dedekind rings by axioms and, in addition, showed that these axioms are preserved under finite field extensions. Since $K[x]$ is a principal ideal ring and hence a Dedekind ring, theorem (i) follows. It is true that Noether had proved the preservation property for separable extensions only. This was necessary to copy the usual proofs for the "finiteness property" which says that the Dedekind ring in the extension field is a finite module over the Dedekind ring in the original field. But in the same year, Artin and van der Waerden [A-vdW:1926] extended this finiteness theorem to inseparable extensions, provided the fields have finite p-rank which is the case in our situation. [33]

Noether's paper [Noe:1926] carries the title (in English translation):

> "*Abstract structure of the ideal theory in algebraic number and function fields.*"

Compare this with the title of her preliminary announcement [Noe:1924], after she had given a talk on this subject at the DMV-meeting [34] in Innsbruck:

> "*Abstract structure of the ideal theory in algebraic number fields.*"

We see that in the meantime, between 1924 and 1926, Emmy Noether had added "function fields" to the title. This reflects that she had become aware of the ongoing activities concerning function fields, and she wished to point out explicitly that her approach works as well in the function field case. In this way she had an important share in developing the structure theorems in function fields. We should notice, however, that Noether's arguments work for function fields over an arbitrary base field K; the special property that K is finite is not relevant for the validity of theorem (i).

The above mentioned authors Sengenhorst, F. K. Schmidt and Rauter did not refer to Noether's result. Probably they did not know it: Compare the publication date 1926 of Noether's paper with the dates we have given above

[33] Much later, Grell [Gre:1935] proved Noether's theorem (but not the finiteness theorem of Artin-van der Waerden) for arbitrary finite field extensions. F. K. Schmidt in [FK:1936a] says that he (Schmidt) had presented Grell's result already in 1927 in his *Habilitationsvortrag* in Erlangen.

[34] DMV=*Deutsche Mathematiker Vereinigung.*

for their work. But even if they knew [35] we should take into account that Noether's axiomatic approach was considered in many quarters as being very abstract and difficult. For many it may have seemed more natural to fall back on the classical approach to theorem (i) for number fields, and to verify that the arguments there can be transferred to the function field case.

Thus each of the three authors above gives a proof of theorem (i) along the classical lines. For the corresponding proofs in the number field case, Sengenhorst and F. K. Schmidt are citing Landau's text book [Lan:1918] which also Artin had used. Rauter cites Hecke's text book [He:1923] which had appeared in 1923.

But there has to be some extra reasoning dealing with inseparability. To this end the following lemma was proved by Sengenhorst as well as by F. K. Schmidt (Rauter assumed that the extension $F|K(x)$ is separable):

If F over $K(x)$ has inseparability degree p^s then F contains $K(x^{p^{-s}})$ and is separable over this subfield.

This allows to reduce the above mentioned "finiteness theorem" to the separable case. In the present situation this says that R admits a $K[x]$-basis. For the proof of the lemma it is only needed that K is perfect, which of course is true for finite fields. An immediate consequence of the lemma is that every function field F over a finite field K can be separably generated.

REMARK: In my report [Rq:2001] I have forwarded the opinion that F. K. Schmidt was the first to prove that function fields over a perfect base field are separably generated. Now we see that this fact is already contained in Sengenhorst's thesis of 1923, two years before F. K. Schmidt's.

Concerning the discriminant theorem (ii): Here we have a similar situation as for theorem (i), namely: There is a paper by Noether [Noe:1927] in which she proves this theorem in an abstract, much more general setting. Again, Noether's result could not be known to our authors. Hence their method of proof was to check the classical arguments which had been used for number fields, and then transfer these arguments to the function field case. F. K. Schmidt used the definition of the discriminant from Hilbert's "*Zahlbericht*", whereas Rauter used Dedekind's definition. Sengenhorst did not discuss the discriminant theorem (ii).

4.1.2 Global fields

The unit theorem (iii) and the class number theorem (iv) are of different kind than the general ideal theoretic theorems (i) and (ii). Whereas the latter are valid for arbitrary base fields K, the two former rely heavily on the fact that the base field K is finite.

In the corresponding number field case it is necessary to use the archimedean valuations, besides of the $\mathfrak{p}$-adic valuations belonging to the prime ideals $\mathfrak{p}$ of R. In the function field case, the analogue of the archimedean valuations are the "valuations at ∞", which are defined as those valuations of F which induce in $K(x)$ the degree valuation. We can also describe them as the "poles" of x in K. The role of those valuations is quite analogous to the role of the archimedean valuations in number theory: they are used to estimate the "size" of ideals and

[35] Remember that Emmy Noether had given a talk on this topic 1924 at the DMV-meeting in Innsbruck, and that a preliminary announcement had been published in the DMV-*Jahresbericht* the same year already [Noe:1924].

elements of R. This had already been observed and used by Artin in his thesis; see section 2.1.1.

All valuations of $F|K$, whether they are poles of x or not, are related by means of the *product formula*:

$$\prod_{\mathfrak{p}} \|a\|_{\mathfrak{p}} = 1 \qquad \text{for} \qquad 0 \neq a \in F\,. \tag{34}$$

Here, $\|a\|_{\mathfrak{p}}$ denotes a valuation, in suitable normalization, belonging either to a prime ideal of R or to a pole of x.

It was Hasse who developed the idea that this product relation of the valuations for a function field, together with the finiteness of the base field, is the main source for the fundamental theorems (i)-(iv), and similarly for algebraic number fields. In other words: Hasse's concept was to give a foundation for the theory of "global fields" by means of that product formula. By definition, a "global field" is either an algebraic number field of finite degree, or an algebraic function field over a finite base field. [36] Hasse's classic book "*Zahlentheorie*" [H:2002] was written just for this purpose, i.e., to build the theory of global fields upon the notion of valuation or "prime divisor", together with the product formula (34). [37] The theorems (i)-(iv) for global fields are covered in Hasse's book.

Artin and Whaples [A-Wh:1945] published an axiomatic foundation for global fields which is based on a similar idea. It is remarkable that Artin-Whaples too included a proof of the unit theorem (iii) and the class number theorem (iv) in their paper, as almost immediate consequences of their axioms. In the introduction they said they wished to demonstrate that these theorems can be formulated and proved without ideal theory, which means without relying on the Dedekind ideal theorem (i). The beauty and simplicity of their proofs are apparent.

Thus from today's point of view, for the proof of (iii) and (iv) we would just refer to Artin-Whaples [A-Wh:1945] or to Hasse's book [H:2002].

This was not possible for Sengenhorst, and not for F. K. Schmidt. So they had to introduce "valuations at ∞" and to transfer the classical arguments from number fields etc. to the function field case. This required a certain amount of valuation theory. At the time of Sengenhorst and F. K. Schmidt this was already available, but still valuation theory was not widely known and accepted. In the next sections we shall discuss in more detail how Sengenhorst and F. K. Schmidt dealt with valuation theory. – Rauter did not include theorems (iii) and (iv) into his paper.

4.2 Sengenhorst

Paul Sengenhorst (1894-1968) studied in Göttingen 1913–1916 and 1918–1920, the interruption being due to military service in the war. In 1920 he passed the state examination and accepted a position as a gymnasium teacher in East Prussia. In this position he continued to work on his Ph.D. thesis, on a problem which Landau had suggested.

[36] Hasse did not use the word "global field"; this was introduced much later.

[37] The original idea for such book can be found in the late 1920's already in Hasse's correspondence with Hensel. The manuscript was completed in November 1938. But due to external circumstances it appeared in 1948 only.

This was in the same year when Artin was working on his thesis. It seems that Sengenhorst's work was overlapping with Artin's. Recall that the topic of Artin's thesis had been suggested by Herglotz, relating to the paper by Kornblum [Kor:1919]. Kornblum had been a student of Landau, and Landau had edited his paper. (See section 2.2.) Thus it is not surprising that both Herglotz and Landau had suggested to their Ph.D. students similar problems, both linking to Kornblum's paper.

In Behnke's article [Be:1968] about Sengenhorst we read:

> "*It must have been disturbing for Sengenhorst when he heard about the work of Artin. The two met and Artin reported to Sengenhorst about his results. This must have been depressing for Sengenhorst who lived quite isolated in a small town in East Prussia. Without doubt he also realized the genius of Artin. But then he got advice from Landau. The topic of his thesis was changed. In 1922 he moved back to Göttingen and there he wrote his paper on function fields in characteristic p.*"

The new topic which Landau had proposed to Sengenhorst was to develop the arithmetics for arbitrary function fields with finite base fields. In early 1923 Sengenhorst submitted his thesis to the faculty in Göttingen and got his Ph.D. [38]

4.2.1 The thesis

Sengenhorst's thesis appeared in the *Mathematische Zeitschrift* [Se:1925] with the title:

> *On fields of characteristic p.*

In the handling of the valuations at ∞, Sengenhorst follows Artin and identifies the corresponding completion of $K(x)$ with the field of all Laurent series in x^{-1} over K. For the study of the prolongations of this valuation from $K(x)$ to F, it is clear that it becomes necessary to apply something like Hensel's Lemma – whereas Artin could get away with the alternative of $\sqrt{D}$ being "real" or not. Sengenhorst knows that he has to use some amount of valuation theory for this. Let us cite from his introduction:

> *Apart from the arguments of classic number- and ideal theory, we shall use the results of the algebraic theory of abstract fields of Mr. Steinitz* [Ste:1910]*, as well as the methods which had been developed by Mr. Hensel* [Hen:1908] *when dealing with p-adic numbers. The presentation will be much simplified by using the notions of valuation theory introduced by Mr. Kürschák* [Ku:1913].

The very fact that this had to be stated explicitly shows that this was not quite the usual procedure at that time.

But Sengenhorst does not keep his promise of simplification by using Steinitz and Kürschák. He does not use Kürschák's result that the valuation extends

[38] Thereafter Sengenhorst accepted a position as gymnasium teacher in Berlin-Charlottenburg and later in Münster, but he kept contact to university circles (Schur in Berlin and Behnke in Münster; there is also some amount of correspondence in the legacy of Hasse). For more biographic information see Behnke's obituary for Sengenhorst [Be:1968].

uniquely, in the well-known manner, from a complete field to every algebraic extension. He does not consider, after Steinitz, the algebraic closure of the complete field but restricts himself to stay in certain finite extensions, big enough to contain the given F and all its conjugates over $K(x)$; such extension is then declared as the analogue to the complex number field. If the author would indeed have used Kürschák and Steinitz then, as he had promised, this would have meant simplification and, hence, concentration on the essentials. But instead, Sengenhorst spends 24 pages (out of 39) just to derive all these known facts from general valuation theory, for the special field $K(x)$ and its completion. In doing so he copies the arguments for p-adic numbers from Hensel's book [Hen:1908] and verifies that these arguments apply also to his Laurent series, including a proof of Hensel's Lemma. He does not realize that the axiomatic foundation by Kürschák [Ku:1913] covers all the results that he needs.

This blows up Sengenhorst's paper unproportionally and makes it clumsy to read. It seems that Landau as his Ph.D. advisor is responsible for this. For, in his report on Sengenhorst's paper Landau writes:

> "*On my special request the manuscript was written somewhat long-winded, more than the author himself would print it; this was done in order that it could be read without having to look up too much in the literature.*"

It is a curious fact that Landau, whose own style is characterized as being extremely brief and of "striking simplicity", required from his Ph.D. a long-winded exposition. We get the impression that Landau was not used to such "abstract" structures like function fields. In his report he repeatedly points out that the paper investigates abstract fields, not number fields, that integral algebraic entities in this paper are not algebraic integers, etc.

Perhaps this is the reason why Sengenhorst's paper was not widely read and did not have a great influence on further development. But it was read by the Czech mathematician K. Rychlík in Prague. Rychlík wrote a letter to Sengenhorst telling him that all the constructions and results concerning valuations in Sengenhorst's paper are contained as special cases in his (Rychlík's) paper [Ry:1923]. [39] In 1927 Sengenhorst published a one-page supplement [Se:1927] to his paper, saying that Rychlík's [Ry:1923] was not known to him at the time when his manuscript went to print.

Remark: If Rychlík's paper was not known to Sengenhorst, the paper of Kühne [Kue:1903] certainly was known to him because he cited it in his introduction. But he only said that "*some results*" of his paper had already been obtained by Kühne. A closer look at Kühne's paper reveals that "some results" is precisely the *unit theorem* as stated above in (iii). For historical reasons it seems to be of interest that the unit theorem for function fields was treated as early as 1903 but that this had gone unnoticed by the mathematical community except Sengenhorst; but he said not more than one sentence about it.

[39] That paper is usually cited as having appeared in 1924. But volume 153 of Crelle's Journal consisted of two issues and the first issue, containing Rychlík's paper, appeared on Aug 27, 1923 already, hence before the appearence of Sengenhorst's paper. By the way, Rychlík's paper in Czech language had already appeared four years ago in 1919 [Ry:1919] already, and it had been refereed in the "*Jahrbuch der Fortschritte der Mathematik*", vol. 47. – For more about Rychlík and his work on valuation theory see [Rq:2002].

Kühne's paper also contained a proof of Hensel's Lemma for the valuation at ∞ in $K(x)$. [40]

4.3 Rauter

Herbert Rauter (1899–?) had a similar mathematical background as had Sengenhorst. He started his university studies in 1920 after military service during the war (including a British prison camp in Palestine). He studied at the university of Königsberg, with three interim semesters in Berlin. In 1924 he passed the state examination and accepted a position as gymnasium teacher in the town of Tilsit in East Prussia. In this position he worked 1925-1926 on his Ph.D. thesis about quadratic forms in function fields, under the supervision of Helmut Hasse. At that time Hasse held a position as professor in Halle.

We do not know when and how Rauter had established contact with Hasse. It may have been in September 1925 during the Danzig meeting of the DMV. At that time Danzig was a German town but had been politically separated from Prussia after the first world war. But there were still many ties, cultural, economical and personal, between East Prussia and Danzig. The home town of Rauter, where he grew up, was Osterode, quite close to Danzig. [41] And so it does not seem improbable that Rauter, a gymnasium teacher who was mathematically interested and active, went to Danzig to participate in the DMV meeting. He had just joined the DMV, in the same year 1925. We can imagine that Rauter attended Hasse's lecture on class field theory in Danzig, and that he was fascinated by these new visions of number theory. – But this is only speculative. In fact we have no evidence that Rauter had attended the Danzig meeting except that his correspondence with Hasse started shortly thereafter.

4.3.1 The thesis

Since Rauter and Hasse lived in different towns, the mathematical discussion between them proceeded by letters. Rauter's letters to Hasse are preserved in the Hasse legacy. There are quite a number of such letters from the years 1926-1927 and later.

The first of these letters is dated January 14, 1926. Rauter responds to a "long" letter of Hasse where, it seems, Hasse had proposed the direction of study for Rauter and given comments. Now Rauter reports that he had started with his work. It appears that Hasse had proposed to transfer the results of his own thesis to the case of function fields. Hasse in his thesis [H:1923] had considered quadratic forms over the rational number field $\mathbb{Q}$ and established the Local-Global Principle for isotropy of these forms, as well as local criteria for isotropy. This, Hasse now proposed, should be transferred to quadratic forms over the rational function field $K(x)$ where K is a finite field. At first Rauter should take K to be $\mathbb{F}_p$, the prime field of characteristic p, as Artin had done in his thesis. Of course $p > 2$.

[40] I am indebted to Franz Lemmermeyer for information about Kühne's work. Kühne (1867-1907) had obtained his Ph.D. 1892 at the University of Berlin. He became teacher at the technical school in Dortmund. It would be desirable to obtain more biographic information about this interesting mathematician.

[41] Today the towns of Danzig and Osterode are called *Gdansk* and *Ostroda* respectively; Königsberg and Tilsit are *Kaliningrad* and *Sovjetsk*.

In Hasse's thesis, the quadratic Hilbert symbol plays a dominant role, and its properties are derived from the quadratic reciprocity law in $\mathbb{Q}$. Now, Rauter reports in his letter that he is going to define the quadratic Hilbert symbol in the function field case, and to establish its properties by means of the quadratic reciprocity law of Artin's thesis. We see that it was the results of Artin's thesis which had induced Hasse to propose the treatment of quadratic forms over $K(x)$.

Thus Rauter's thesis is to be regarded as a corollary to both Artin's and Hasse's theses.

In some later letter, dated February 5, 1926, Rauter responded to a postcard of Hasse. It seems that Rauter had successfully transferred the main results of Hasse's thesis to the function field case, but Hasse had now proposed to extend his investigations such as to include arbitrary finite base fields K instead of $\mathbb{F}_p$ only. Rauter writes that he would rather not do this because it would require a complete rewriting of his paper and perhaps the results could not be achieved so easily. Rauter wished to complete his thesis as soon as possible.

We see that here the same problem came up as we had described with Artin's thesis in section 2.1, the difference being that Artin considered the extension to arbitrary finite base fields as "self-evident" whereas Rauter feared there may arise difficulties.

In the first days of June, Rauter travelled to Halle and there he passed his Ph.D. examination on June 6, 1926. [42]

Rauter's thesis was never published in a mathematical journal. It seems to us that Hasse wished to wait with a publication until Rauter had achieved more results. A copy of the thesis is kept at the library of the University of Halle. A look at it reveals that it is structured completely after Hasse's thesis, transferring step by step Hasse's arguments from the number field case to the function field case. The crucial idea is to define the Hilbert symbol $\left(\frac{a,b}{\mathfrak{p}}\right)$ locally for every prime $\mathfrak{p}$ of $K(x)$, and to verify that it satifies the product formula

$$\prod_{\mathfrak{p}} \left(\frac{a,b}{\mathfrak{p}}\right) = 1$$

for $a, b \in K^\times$. Here, Rauter uses the quadratic reciprocity law of Dedekind-Artin. After that, it turned out that everything works precisely as in Hasse's thesis. One difference to the number field case is perhaps worth mentioning, namely that in dimension $n \geq 5$ every quadratic form over $K(x)$ is isotropic. This is a consequence of the fact that the prime at ∞ represents a non-archimedean valuation, whereas over $\mathbb{Q}$ it is archimedean and therefore the local condition for isotropy is non-trivial for every dimension.

[42] Originally Rauter had planned to submit his thesis to the Faculty of Königsberg University, although it was written under the supervision of Hasse in Halle. In fact, he had presented the manuscript already to Knopp who at that time held a professorship in Königsberg. But it turned out that, on the one hand, Knopp had just accepted an offer from the University of Tübingen and hence did not wish to have another Ph.D. candidate shortly before he left Königsberg. On the other hand, Rauter writes, Knopp told him that he was not acquainted with the theory of p-adic numbers. Hasse and Hensel, in the opinion of Knopp, were the only mathematicians in the world who were able to deal with them. Certainly this was not the case but we conclude that in 1926 the theory of p-adics was still not universally appreciated by the mathematical community. – Finally Hasse had offered Rauter to get his Ph.D. from the University of Halle.

4.3.2 More work

After Rauter had finished his thesis, Hasse had apparently proposed to him to extend his results to quadratic forms over an arbitrary function field F with finite base field, not necessarily rational. This would have transferred, to the function field case, the results of Hasse's paper [H:1924] where he established the Local-Global Principle over an arbitrary algebraic number field.

But first there had to be developed the arithmetic in arbitrary function fields with finite base fields, and Rauter now started to do this. It seems that originally he did not know of Sengenhorst's paper [Se:1925]. When Hasse informed him about Sengenhorst, Rauter wrote in a letter of September 6, 1926: "*I can say that my results were obtained independently of Sengenhorst*". In Rauter's paper [Rau:1928] he mentions in a footnote that some of his results had already been obtained earlier by Sengenhorst. He does not mention F. K. Schmidt's thesis – because that was not published and probably he did not know it. [43]

From the results (i)-(iv) as stated in section 4.1, Rauter's paper contains only (i) and (ii) which is to say the Dedekind property and the ramification theorem. But in addition Rauter gives a treatment of Hilbert's theory of higher ramification groups in the case that $F|K(x)$ is a Galois extension – with the result that everything goes as in the case of number fields.

In contrast to Sengenhorst's paper, Rauter's is written in concise, clear style. [44] However it does not get near the goal which Hasse seems to have had in mind, namely the quadratic reciprocity law for arbitrary (not necessarily rational) function fields with finite base fields. This would have paved the way for the transfer of Hasse's Local-Global Principle for quadratic forms, from arbitrary number fields to arbitrary function fields.

Nowadays the theory of quadratic forms over global fields can be found in the monograph by O'Meara [OMe:1963]. O'Meara had been a student of Artin in Princeton.

One of the first readers of Rauter's paper was Emmy Noether. In a letter of May 14, 1928 she writes to Hasse: [45]

> *"I have had a look at the paper of Rauter in the recent Crelle issue; he forgot to say that he assumes F to be separable over $F_0 = K(x)$ although he uses this several times, mainly in arguments which he did not present in detail.* [46] *By the way, the ideal theory is preserved under inseparable extensions too, as has been shown by F. K. Schmidt and in more general cases by Artin-van der Waerden;* [47] *but in this*

[43] From which fact one could conclude that Rauter did not attend F. K. Schmidt's talk in Danzig – if he was in Danzig at all.

[44] This is certainly due to Hasse who returned the manuscript several times to the author requiring a better exposition. Hasse used to do this with all of his students; in his opinion authors should not neglect to care about a clear and adequate presentation of their results. Quite often Hasse went to great length in order to explain to his students his idea of how to write a mathematical paper.

[45] Actually, this letter is not dated by Emmy Noether. She wrote it on the back side of a bill which she had just obtained for reprints of papers which she and Richard Brauer, together with Hasse, had recently published [Br-Noe:1927], [H:1927b]. Although these were two papers, they had been written in close collaboration and the reprints were bound together. May 14, 1928 is the date of that bill.

[46] This remark of Noether gave rise to a one-page commentary [Rau:1928a] of Rauter to his paper, where he admits this omission.

[47] See footnote 33.

case the theorems for the different and discriminant will change. [48]

The fact that everything goes like it does in number fields is due to the following assumptions which are solely used also in number fields:

1) *F is separable over F_0.*
2) *The integers in F_0 form a principal ideal ring.*
3) *The residue class ring with respect to every integral ideal in F_0 (and hence also in F) consists of finitely many elements.*

1) and 2) lead to ideal theory; 3) gives ramification theory; in particular since it refers to the theory of the same residue class rings...

Hence the essential differences appear in the different behavior of the infinite places only."

We see that Noether did here what she always did, namely in order to clarify a theory she puts it into a more general, abstract setting defined by axioms. [49]

Emmy Noether wrote that letter in the year 1928, two years after her paper [Noe:1926] had appeared where she characterized Dedekind rings. Since Rauter had completed his paper in 1926, he could perhaps not be blamed for not knowing Noether's paper and to work her result into his paper. It looks like her letter is to be read as containing some criticism towards Hasse who certainly knew Noether's result and nevertheless had accepted Rauter's paper for publication in Crelle's Journal.

In fact, Emmy Noether had announced her results already in September 1924 at the annual meeting of the DMV in Innsbruck; see [Noe:1924]. Hasse had attended the Innsbruck meeting and therefore he was informed about Noether's result. In several letters of 1925-26 between Hasse and Noether her axioms for Dedekind rings had been discussed. [50] Specifically, in her letter of Nov 3, 1926 (hence before her paper appeared) she had explained to Hasse the implication of her 5th axiom (which was integrally closedness), and on her letter we have found notes of Hasse's hand elaborating on her arguments. So we know that Hasse was informed and also interested.

Thus Noether's letter commenting Rauter's paper seems implicitly to ask Hasse: "Why didn't you tell Rauter that this has been done already, once and for all? After all you knew how to do it!"

We do not know the reason why Hasse had accepted Rauter's paper without insisting that at least a reference to Noether's was included. Perhaps he hoped

[48] Noether did not specify which change she had in mind. Did she anticipate Tate's theory [Ta:1952] of a modified different for inseparable extensions?

[49] Noether seems to claim that for number fields and for function fields one has "the same residue class rings" and therefore the same ramification theory. Of course this is not true, and accordingly the structure of the ramification groups in characteristic p is different from that in characteristc 0. It is not clear what Noether could have meant. We note that some years later one of her students, Deuring, got the task to develop the Hilbert ramification theory for an arbitrary valued Galois extension [Deu:1931]. Perhaps Emmy Noether has had second thoughts and wished the matter to be cleared up.

[50] In modern terminology her axioms are: The ring should be Noetherian, integrally closed, and every proper factor ring Artinian. Noether formulated this in 5 axioms; in the correspondence with Hasse she called them "Innsbruck axioms" because she had presented them in her Innsbruck talk. And the rings satisfying those axioms were called "5-axiom rings" – before they got baptized as Dedekind rings.

that in due course Rauter would produce more and better results. From the Hasse-Rauter correspondence we have got the impression that Hasse for some time overestimated the mathematical abilities of Rauter. He had sent Rauter a copy of his class field report [H:1926], obviously with the hope that Rauter would be able to go some way towards transferring class field theory to the function field case. But in Rauter's papers [51] there is no sign that he was able to follow Hasse's proposals.

4.4 F. K. Schmidt

Friedrich Karl Schmidt (1901–1977) studied 1920-1925, mainly in Freiburg. [52] In those times it was common among German students to change their university for one or two semesters and return thereafter to their home university to finish their studies. F. K. Schmidt spent one semester at Marburg University (summer 1923). In 1924-25 he wrote his Ph.D. thesis. (Thus F. K. Schmidt's thesis was written earlier than Rauter's.)

Formally, his advisor was A. Loewy. However, as F. K. Schmidt mentions in the preface of his thesis, the actual proposal for this topic was given by W. Krull who supervised the work. At that time Krull held a position as *Privatdozent* and assistant to Loewy.

Krull, during his student years, had spent two years 1920-21 in Göttingen and there he had been strongly influenced by Emmy Noether. From then on he always kept close scientific contact to her. Krull brought to the small place of Freiburg the new ideas in algebra which were discussed in the Noether circle in Göttingen. It appears that the choice of the subject of F. K. Schmidt's thesis was directly influenced by those ideas. For biographical information about Krull see [Kru:1999].

4.4.1 The thesis

F. K. Schmidt's thesis carries the title:

Arbitrary fields in the domain of higher congruences.

This wording signifies that he wished his work to be regarded as continuation of Artin's thesis. Like Artin he used the terminology "*domain of higher congruences*". But, again like Artin, he used it only in the title and the introduction, in order to put his paper into the proper historical perspective; otherwise he adhered to the "modern" terminology and notions of Steinitz and Noether. He says he is going to generalize Artin's theory from quadratic to a arbitrary function fields $F|K$.

However in his thesis he considered only the generalization of the first, arithmetic, part of Artin's thesis. But it may well be that originally he had planned

[51] Besides of the paper [Rau:1928] under discussion Rauter managed to publish some more papers which, however, are "trivial and contain nothing new", as Hasse in a later letter, of March 2, 1930, conceded.

[52] In Germany the name "Schmidt" is quite common. There are several known mathematicians with this name. In order to identify them it is common to use their first names, or first name initials. We shall follow this habit here too; this is the reason why we always use the initials when mentioning F. K. Schmidt, whereas with other mathematicians the initials are not used in general.

to deal with the second, analytic, part as well, for in a letter to Hasse dated May 6, 1926 he says:

> "*Regarding the limit formula for the ζ-function in fields of characteristic p, I had not been able to transfer it in my thesis.*"

Here, "limit formula" means the computation of the residue of the ζ-function at its pole $s = 1$. In the number field case, this limit formula for the residue contains important arithmetic invariants of the field. Obviously, Hasse had asked him whether he had been able to transfer such formula to the function field case, in generalization of Artin's formulas (11). We shall see in section 5 that later, F. K. Schmidt succeeded to transfer the limit formula and more. But here, in his thesis, the zeta function is not investigated and not even mentioned.

When F. K. Schmidt wrote his thesis he did not know Sengenhorst's paper. Recall that Sengenhorst's appeared in print in 1924 only. So it is not surprising that young F. K. Schmidt was ignorant of it when he started to write his thesis in 1924. It is surprising, though, that his supervisor Krull did not know it. Krull had close contact to Emmy Noether in Göttingen. Perhaps Emmy Noether did not know it either? This is hard to believe, for Emmy Noether was interested in the subject and, moreover, Sengenhorst's thesis was submitted to the Faculty in Göttingen in 1923 already. But it may well have been that Noether was not too happy with the way Sengenhorst had handled the subject and had expressed her wish that the subject should be taken up again. Recall that Sengenhorst had written his thesis not with Noether but with Landau as his supervisor. The style of Sengenhorst's paper does not suggest that he belonged to the circle around Noether. [53]

F. K. Schmidt's thesis contains proofs of the theorems (i)-(iv) which we stated in section 4.1. In the handling of the places of F at ∞, F. K. Schmidt's presentation is a significant progress. He says:

> "*Obviously, the notion of the imaginary, which plays a role in the determination of units in number fields, cannot be transferred immediately* [to function fields]. *However the theory of units can be developed in a simple manner by using ideas from complex analysis... It is possible to define formally the infinite in such a way that it is susceptible to algebraic treatment.*"

And then he refers to the classical paper of Dedekind-Weber [De-Web:1882]. There, the authors had worked in the context of complex analysis, which means that they had considered a function field F with base field $K = \mathbb{C}$, the complex numbers. In that situation, a place ("*Punkt*") of $F|K$ was defined as a K-homomorphic mapping of F into K including ∞, with the usual rules concerning ∞. Dedekind and Weber already remarked that their theory is of algebraic nature and thus remains valid if $\mathbb{C}$ is replaced by the field of all algebraic numbers. (We may add that it remains valid for any algebraically closed base field.) Now, F. K. Schmidt is going to use this idea in the case when K is not necessarily algebraically closed.

[53] Landau in his report about Sengenhorst's thesis wrote: "*The candidate is a sympathetic, quite shy person who during his student time and later evaded timidly all stimulations, and generally all mathematical conversations.*" Landau did not mean to criticize the candidate, he wrote this to point out that the candidate had written his thesis completely on his own without outside help.

He realizes that it does not suffice to consider mappings of F to K (including ∞). Instead he defines a place of $F|K$ as a K-homomorphic mapping of F into some algebraic overfield of K (including ∞). Two such places are said to determine the same "point" of $F|K$ if they differ only by a K-isomorphism of their image fields. [54] The elements f of the function field F now appear indeed as functions; they are defined on the set of the points (or places) $\mathfrak{p}$ of $F|K$, the value $f(\mathfrak{p})$ of that function being in the abstract image field which we denote by $F\mathfrak{p}$. This is a finite extension of K; its degree $d = [F\mathfrak{p} : K]$ is called the "degree" of $\mathfrak{p}$ and denoted $\deg(\mathfrak{p})$. [55] F. K. Schmidt says that $\mathfrak{p}$ is a "d-fold" point, in view of the fact that over the algebraic closure $\widetilde{K}$ of K this point will decompose into d points of the constant field extension $F\widetilde{K}|\widetilde{K}$.

In the process of building a new mathematical theory, it is of great importance to find the proper notions and constructions which, on the one hand carry enough potential to handle the upcoming problems, and on the other hand correspond to the underlying ideas and imagination about the structures involved. For function fields over the complex field, Dedekind and Weber had proposed their notion of "point" which had proved to be very fruitful and, consequently, had been the model for similar notions in other branches of mathematics, e.g., in functional analysis where also the "points" are constructed as mappings of the algebras involved. Today we are quite used to the idea that an algebraic structure like field, ring, algebra etc. carries its spectrum, i.e., its space of points, which is inherently connected with the original structure. In the time of Dedekind and Weber, this was a "revolutionary idea". [56]

In retrospective we can say that F. K. Schmidt found the proper extension of the Dedekind-Weber notion of point to the case when the base field is not algebraically closed. This constituted a major advance; we should not underestimate the conceptual difficulties which many mathematicians of those times had to cope with in front of this abstract construction. We shall see in the next chapter that this new idea provides the key to the birationally invariant theory of the zeta function of a function field.

In a footnote F. K. Schmidt tells us that this idea was based on an oral communication by W. Krull. It seems interesting that already in 1925 Krull had developed the idea of "place" of a field, which he expounded later in his seminal paper [Kru:1932] on general valuation theory for arbitrary fields.

If $f \in F$ and $\mathfrak{p}$ is a place (or "point") of $F|K$ then the "order" $\mathrm{ord}_{\mathfrak{p}}(f)$ of f at $\mathfrak{p}$ is defined in the usual manner, following Dedekind-Weber. This gives the $\mathfrak{p}$-adic valuation of F, in its additive form. F. K. Schmidt normalizes this valuation such that for a prime element t at $\mathfrak{p}$ we have $\mathrm{ord}_{\mathfrak{p}}(t) = \deg \mathfrak{p}$, the degree of $\mathfrak{p}$. With this normalization we have the sum formula

$$\sum_{\mathfrak{p}} \mathrm{ord}_{\mathfrak{p}}(a) = 0 \qquad \text{for} \qquad 0 \neq a \in F \tag{35}$$

[54] The terminology in valuation theory is not uniform. The notion of "point", "prime divisor", "valuation" are used synonymously, depending on the associations which the author wishes to evoke. The same is true for "place", whereby two places are considered to be equal if they determine the same "point" in the sense of F. K. Schmidt. The terminology "point", which F. K. Schmidt uses, was coined by Dedekind-Weber. – We shall use freely any of these words, whatever seems appropriate in the situation at hand.

[55] F. K. Schmidt uses the word "order" (*Ordnung*) instead of "degree" . It was Hasse in [H:1934] who proposed to use "degree" (*Grad*).

[56] See the assessment of the Dedekind-Weber paper by D. Geyer in [Ge:1981].

which expresses the fact that every $a \in F^\times$ has the same number of zeros and of poles, if counted with the proper multiplicity. This relation is equivalent with the product formula (34) if we normalize the multiplicative valuation such that

$$\|a\|_\mathfrak{p} = q^{-\mathrm{ord}_\mathfrak{p}(a)}\,.$$

The completion of F at some point $\mathfrak{p}$ is identified with the field of "Laurent series" in a prime element t for $\mathfrak{p}$. [57]

F. K. Schmidt shows that the places of $F|K$ with $x(\mathfrak{p}) \neq \infty$ correspond bijectively to the prime ideals of R. (In view of this we use the same symbol $\mathfrak{p}$ to denote a prime ideal of R and its corresponding place.) The finitely many points with $x(\mathfrak{p}) = \infty$, i.e., the poles of x, are the analogues to the archimedean valuations in the number field case. Let S denote the set of these poles and r their number.

Now, the units in $R^\times$ are characterized as those elements $u \in F$ all of whose poles and zeros are contained in S. In other words, the principal divisor (u) is contained in the divisor group generated by the $\mathfrak{p} \in S$. The latter is a free abelian group of rank r. The sum formula (35) gives a linear relation which is satisfied by the principal divisors of the units $u \in R^\times$. It follows that $R^\times/K^\times$, as a free abelian group, has rank $\leq r-1$. *Thus the essential statement of the unit theorem* (iii) *is that this rank is precisely* $r-1$.

To show this, F. K. Schmidt uses what he calls the classical Dirichlet argument about the existence of independent units. He points out that the logarithms which appear in Dirichlet's proof are to be replaced in the function field case by the order functions, i.e., the additive valuations at the poles of x.

This approach of F. K. Schmidt to the unit theorem puts into evidence that the theorem depends solely on the set of poles of x. In fact, F. K. Schmidt's proof yields the following theorem:

> Unit Theorem. *Let S be any finite non-empty set of points of F, and r the number of points in S. Let U_S denote the group of S-units in F, i.e., those elements $u \in F$ all of whose zeros and poles are contained in S. Then the torsion group of U_S is cyclic, and the torsion factor group of U_S is a free abelian group of rank $r-1$.*

In this theorem and its proof there is no mention of the ring R and its ideal theory.

In later years Hasse had pointed out that the same theorem holds in algebraic number fields if the "points" are interpreted as the prime divisors of the number field, including the primes at infinity which correspond to the archimedean valuations of the field. There is no requirement that S should contain the archimedean valuations, and hence there is in general no ring of which U_S is the unit group in the ring theoretic sense. This generalized unit theorem is to be found in the mimeographed notes of Hasse's Marburg lectures on class field theory [H:1932]. It has been included in Hasse's book "Number Theory" [H:2002] which we have mentioned above already. Artin and Whaples [A-Wh:1945] have

[57] F. K. Schmidt takes the coefficents of those Laurent expansions from a fixed set of representatives of the image field $F\mathfrak{p}$ of the place. It seems that he did not yet realize that $F\mathfrak{p}$ is naturally contained in the completion, as a consequence of Hensel's Lemma (note that the base field is finite, hence perfect). Many years later, in a joint paper with Hasse [H-FK:1933], he proved this in a much more general situation. Compare our comments in [Rq:2002] to that paper.

included this theorem, being "fundamental to class field theory", into their paper about the characterization of global fields.

We see that the nucleus of this later development can be found already in the 1925 thesis of F. K. Schmidt.

4.4.2 The power reciprocity law

We have already mentioned in section 2.1.1 that F. K. Schmidt's thesis contains a proof of the power reciprocity law. Let $\ell > 1$ be an integer not divisible by the characteristic p and assume the finite field K contains the ℓ-th roots of unity, which means that the order q of K satisfies

$$q \equiv 1 \mod \ell\,. \tag{36}$$

Let $P \in K[x]$ be a prime polynomial. The ℓ-th power residue symbol $(\frac{M}{P})_\ell$ is defined for polynomials $M \in K[x]$ which are not divisible by P. By definition, $(\frac{M}{P})_\ell$ is the ℓ-th root of unity in K which satisfies

$$\left(\frac{M}{P}\right)_\ell \equiv M^{\frac{|P|-1}{\ell}} \mod P \tag{37}$$

where $|P| = q^{\deg P}$ is the order of the residue field modulo P. We have $(\frac{M}{P})_\ell = 1$ if and only if M is an ℓ-th power modulo P.

Actually, F. K. Schmidt defines the power residue symbol not as an ℓ-th root of unity in K, but as an ℓ-th root of unity in the complex number field $\mathbb{C}$. He chooses a fixed isomorphism of the group of ℓ-th roots of unity in K to the group of ℓ-th roots of unity in $\mathbb{C}$ and then defines the power residue symbol as the image of our $(\frac{M}{P})_\ell$ under this isomorphism. But it is more convenient to us if we regard $(\frac{M}{P})_\ell$ as an element of K.

If P, Q are two different monic prime polynomials of degree n, m respectively then the reciprocity law reads as follows:

$$\left(\frac{Q}{P}\right)_\ell \left(\frac{P}{Q}\right)_\ell^{-1} = (-1)^{mn\frac{q-1}{\ell}} \tag{38}$$

in generalization of Dedekind's law (2). F. K. Schmidt says proudly:

> *This general reciprocity law is of special interest because for numbers* [instead of functions] *it is completely unknown.*

This of course was not quite so, even at the time of F. K. Schmidt's thesis. There had been a great activity to establish explicit reciprocity laws in various situations of number fields. An overview can be obtained from the book [Lem:2000] of Lemmermeyer. But there does not exist a reciprocity formula in number fields of the same simplicity and generality as (38) in function fields, and this is obviously what F. K. Schmidt meant to say.

The proof which F. K. Schmidt presents in his thesis is superseded by his second proof which he gives in [FK:1926]. There, he shows that the ℓ-th power residue symbol for monic prime polynomials $P = \prod_i (X - a_i)$, $Q = \prod_i (X - b_j)$ can be written as

$$\left(\frac{Q}{P}\right) = \prod_{\substack{1 \le i \le n \\ 1 \le j \le m}} (a_i - b_j)^{\frac{q-1}{\ell}} \tag{39}$$

from which (38) is seen immediately.

The story of the power reciprocity law for $K[x]$ is somewhat curious. First of all, the formula (38) had already been stated and proved by Kühne [Kue:1902]. It seems that neither F. K. Schmidt nor Artin knew Kühne's paper. In fact, Kühne's paper seems to have been forgotten by the mathematical community at large, similar as Kühne's paper [Kue:1903] which we have mentioned in section 4.2. In 1932 Carlitz [Car:1932] rediscovered the power reciprocity law (38), without referring to Kühne or to F. K. Schmidt. Two years later O. Ore [Ore:1934] pointed out that Carlitz seems to have overlooked the paper [FK:1926] of F. K. Schmidt of 1926 – but he himself had overlooked Kühne's paper. Ore says he will give a "new and very simple proof" in its most general form. However, after a closer look it turns out that Ore's proof is essentially the same as the simple proof of F. K. Schmidt in [FK:1926]. The only difference is that F. K. Schmidt considers only monic prime polynomials $P(x), Q(x)$ whereas Ore admits arbritrary polynomials which are relatively prime to each other, but which are not necessarily prime or monic. This generalization is almost immediate. [58]

Apparently F. K. Schmidt had included the general power reciprocity law because Artin had used the quadratic law in the case of quadratic function fields, for the investigation of the zeta function. And F. K. Schmidt may have thought that in the general case, the general power reciprocity is needed for the zeta function. This turned out not to be the case since later in his paper [FK:1931] F. K. Schmidt himself was able to study the zeta function on the basis of the Riemann-Roch theorem, without recurrence to the power reciprocity law.

But the ℓ-th power reciprocity law is necessary for the investigation of the cyclic field extensions of $K(x)$ of degree ℓ, in particular for the decomposition law of primes in such cyclic extension. Accordingly, in the last part of his thesis F. K. Schmidt derives this decomposition law. Still assuming that $q \equiv 1 \mod \ell$ he considers the cyclic Kummer extension $F = K(x, \sqrt[\ell]{D})$ with $D \in K[x]$. He assumes that $[F : K(x)] = \ell$ which means that D is not an m-th power for any divisor $m > 1$ of ℓ. F. K. Schmidt's result is a direct generalization of Artin's result in the case $\ell = 2$ (see statement (v) in section 2.1.1):

(v) *Suppose P does not divide D. Then P splits completely in R if and only if $\left(\frac{D}{P}\right)_\ell = 1$. If this is the case then $P \cong \mathfrak{p}_1\mathfrak{p}_2 \cdots \mathfrak{p}_\ell$, the product of ℓ different prime ideals of R, of relative degree 1. In general, if $\left(\frac{D}{P}\right)_\ell$ is of order $m|\ell$ then P splits in F into $\frac{\ell}{m}$ different prime ideals, each of relative degree m over P.*

This theorem and the power reciprocity law can be used to exhibit cyclic extensions of $K(x)$ as class fields. Later Hasse [H:1934c] developed systematically the theory of cyclic extensions of function fields over finite fields, and in particular their class field theory. In that connection he explicitly cited and used F. K. Schmidt's reciprocity law (38). Compare [Rq:2001]. It is interesting that this information was used by Davenport and Hasse [Da-H:1934] for the proof of the Riemann hypothesis for certain cyclic function fields, the so-called Davenport-Hasse fields. This will be discussed in Part 2.

[58] I am indebted to Franz Lemmermeyer for pointing out to me the papers by Kühne, Carlitz and Ore.

4.4.3 The Danzig meeting

F. K. Schmidt had passed his Ph.D. examination on May 22, 1925, at the University of Freiburg. In September that same year the young Ph.D. attended the annual meeting of the DMV at the town of Danzig. He had announced a talk of 20 minutes about the results of his thesis. The speakers had to submit in advance a short abstract of their talk, which would be published in the "*Jahresbericht*" of the DMV. At the end of his abstract F. K. Schmidt wrote: "*An extended version of the talk will appear in this Jahresbericht*". However, this extended version never appeared, neither in the *Jahresbericht* of the DMV nor elsewhere. F. K. Schmidt's thesis was never published. [59]

The reason for this was that F. K. Schmidt had heard about Sengenhorst's paper. We do not know whether Sengenhorst too attended the Danzig meeting. [60] If so then F. K. Schmidt and Sengenhorst would have met there personally. In any case, F. K. Schmidt met Hasse and Emmy Noether in Danzig, and they most probably were informed about Sengenhorst's paper. The talk of F. K. Schmidt was scheduled for the same afternoon session as Hasse's and Emmy Noether's. [61] Maybe in the discussion after F. K. Schmidt's talk one of them mentioned that the topic had already been covered to a large extent by Sengenhorst.

In any case, these news seem to have induced F. K. Schmidt not to publish his thesis although, as we have seen, it contained more material than Sengenhorst's, and in several respects the material appears to be presented more adequately.

4.5 Artin

There are certain indications that Artin, although he did not publish other papers on function fields after his thesis [A:1921], was early in the possession of the fundamental theorems (i)-(iv) stated in section 4.1. We have not found a letter or document confirming this but the following indications point towards this conclusion. These are:

1. Van der Waerden in his review of F. K. Schmidt's paper [FK:1931] in "*Zentralblatt für Mathematik*" (vol. 1, 1931) writes:

> "*If K is a Galois field of characteristic p, x an indeterminate and F a field of finite degree over $K(x)$ then, due to Artin, there is in F an ideal theory, a theory of units and an analytic theory, whose results coincide to a large degree with the corresponding results for number fields. The properties of the ζ-function, however, had been obtained by Artin only for quadratic fields by means of special computations...*"

[59] It was however mimeographed from the hand-written manuscript and distributed to some university libraries, probably also sent as preprint to several mathematicians. A copy of it is preserved in the library of the university of Freiburg.

[60] According to the membership list of the DMV, Sengenhorst joined the DMV in the same year 1925.

[61] Hasse delivered an invited lecture of 60 min., a report on Takagi's class field theory. This lecture had an enormous influence; it led to Hasse's "class field report" which appeared in three parts [H:1926], [H:1927a], [H:1930]. See our paper [Rq:2001]. – Emmy Noether talked about her ideas on doing representation theory in the setting of algebras and ideal theory; this culminated finally in her famous paper [Noe:1929].

Speaking of an "ideal theory" and a "theory of units" he obviously includes theorem (i)-(iv) in some form. The very fact that he attributes these results to Artin seems to imply that he had learned them from Artin, perhaps in lectures or in personal conversations. In fact, we read in [vdW:1975] that van der Waerden had attended Artin's lectures in Hamburg 1926, and was in close contact with him. Note that van der Waerden's "*Modern Algebra*" resulted from this lecture. [62]

F. K. Schmidt cites Sengenhorst in the preface of his paper [FK:1931], as well as his own (unpublished) thesis. Van der Waerden must have read this before reviewing F. K. Schmidt's paper. But he neglects those citations and insists in his review that the fundamental results (i)-(iv) are due to Artin.

2. In the same year 1926 van der Waerden and Artin published a joint paper [A-vdW:1926] on the preservation of the maximum and minimum condition in finite field extensions, in particular for the inseparable case. Although this topic is not confined to function fields of one variable over finite fields, it has nevertheless some points in common with it; we have mentioned this already in section 4.1.1. We can imagine that Artin and van der Waerden obtained their result while discussing the arithmetic of function fields. And in those discussions van der Waerden may have learned that Artin was in possession of the fundamtental theorems (i)-(iv).

3. In a letter to Hasse dated August 19, 1927 Artin writes:

> "*... To any Galois field K we adjoin an indeterminate x and denote the ensuing field F. Hence the field of my Ph.D. thesis. Of course the whole class field theory is valid here, but with modifications ...*"

Thus Artin puts it as evident that class field theory holds in the function field situation. If this was his opinion, then it would have been even more evident to him that the fundamental theorems (i)-(iv) are valid. It is true that Artin in his letter speaks of the rational function field $K(x)$ only, and the theorems (i)-(iv) are trivial in that case. However, when Artin considers finite abelian extensions over $K(x)$ and their properties as class fields, then naturally he has also to take into account the arithmetic in those abelian extensions, i.e., the validity of the fundamental theorems there. [63]

4. [64] In 1924 Artin wrote a review of Jung's book "*Einführung in die Theorie der algebraischen Funktionen einer Veränderlichen*". Jung's book was written in the style of the well known classic "*Theorie der algebraischen Funktionen*" of Hensel and Landsberg, which means combining analytic and algebraic meth-

[62]Van der Waerden reports in [vdW:1975] that Artin had promised to write a book on algebra for Springer. Van der Waerden was to take notes, and Artin proposed that they would write the book together. Artin's lectures were marvellous. Artin was perfectly satisfied with van der Waerden's manuscript and said, "*Why don't you write the whole book?*".

[63]The "modifications" which Artin mentions in his letter are connected with the fact that in the function field case, the maximal unramified abelian extension is infinite. (It contains all constant field extensions.) – We observe that in the same year 1927 when Artin wrote this letter, he published together with Schreier the theory of cyclic field extensions of degree p in characteristic p – which later became known as "Artin-Schreier extensions". In order to develop class field theory for Artin-Schreier extensions one has to use quite different methods as one is used from ordinary class field theory for number fields; this was shown by Hasse in 1934. It is doubtful whether Artin, when he mentioned "modifications", already had in mind the class field theory of these Artin-Schreier extensions.

[64]This paragraph has been added in January 2003.

ods. [65] Artin expresses his opinion that for an algebraist, it would have been more satisfactory

> "...*to use more extensively the methods of Weber-Dedekind or, better still, to build the theory simultaneously for number fields and algebraic function fields. This is well possible without the presentation lacking conciseness and comprehensibility.*"

Artin's review appeared 1924 in vol. 3 of the "*Hamburger Abhandlungen*". It is true that on this occasion Artin did not explicitly mention that, in his envisaged theory, finite base fields should be admitted. (Jung's book covers only the theory over the complex number field as base field.) But the reference to a unified theory of number fields and function fields seems to point in this direction.

We conclude from **1.-4.** that Artin too has to be mentioned when it comes to describing the formation of the arithmetic in arbitrary function fields. It has often been pointed out that Artin's mathematical achievements, and his influence on many mathematicians, rested not only on his published papers but to a high degree on his excellent teaching, in class room and in personal conversations. Here we observe another instance.

Summary

After Artin's thesis, there was rising interest to generalize Artin's results from quadratic to arbitrary function fields with finite base fields. The structure theory of general algebraic number fields was to be transferred to general function fields as far as possible. This had to be done in two steps: First the arithmetic theory of function fields, and secondly the analytic theory of function fields. In the years 1923-1926 there were three Ph.D. theses written which dealt with the first step: Sengenhorst (1923, Landau), F. K. Schmidt (1925, Krull) and Rauter (1926, Hasse). The most interesting one was the thesis of F. K. Schmidt but this was not published because he became aware of the papers of Sengenhorst and Rauter. But F. K. Schmidt was the only one among the three who pursued the theory of function fields further, and his follow-up work on zeta functions (which we will discuss in the next section) became very influential in the later development.

Emmy Noether was very interested in this development, and she contributed significantly through her paper on the axiomatic description of Dedekind rings, and through another one about different and discriminant. Artin was very interested too and he seems to have included this topic in his Hamburg lectures, maybe 1926. Hasse gave valuable impulses not only to his student Rauter but also to F. K. Schmidt.

5 Building the foundations: the analytic part

Sengenhorst and Rauter did not continue to work on function fields with finite base fields, but F. K. Schmidt did. It seems that he was excited about the news on class field theory which he had heard in Hasse's Danzig lecture. In those times, class field theory of number fields depended on the analytic theory of the zeta function and the L-functions. So F. K. Schmidt started to develop the

[65] Artin says "arithmetic" methods.

same for function fields, with the aim of transferring the theorems on class field theory to function fields over finite base fields. Apparently this had been suggested to him by Hasse. This resulted in the classic paper by F. K. Schmidt on the analytic theory for function fields [FK:1931], which we shall discuss here.

5.1 F. K. Schmidt's preliminary announcement

In 1926 F. K. Schmidt accepted a position in Erlangen as assistant to Professor Haupt. From Erlangen he started his correspondence with Hasse. We have already mentioned F. K. Schmidt's letter of May 6, 1926 to Hasse, in which he said that in his thesis he had not been able to transfer the limit formula of the ζ-function to the function field case. (See section 4.4.1.) Some months later, on August 8, 1926, he wrote again and could report that he had been successful:

> "*After taking up the subject anew, I soon succeeded in transferring completely Dedekind's well known results to fields of characteristic p.*"

And then he went on to state his result. Consider the situation as we have described in section 4.1:

- K a finite fields with q elements,
- F a function field with base field K,
- x a transcendental element in F,
- R_x the integral closure of the polynomial ring $K[x]$ in F.

Compared with section 4.1 we have changed the notation somewhat by writing R_x instead of R in order to point out that the ring $R_x \subset F$ depends on the choice of the transcendental element $x \in F$. Similar change of notation will be done also with other objects depending on x.

In this situation F. K. Schmidt defines the zeta function by the same formula as did Artin in (5), namely

$$\zeta_x(s) = \prod_{\mathfrak{p}} \frac{1}{1 - |\mathfrak{p}|^{-s}} = \sum_{\mathfrak{a} \neq 0} |\mathfrak{a}|^{-s} \tag{40}$$

where $\mathfrak{p}$ ranges over the prime ideals of R_x and $|\mathfrak{p}|$ is the order of the residue field $R_x/\mathfrak{p}$; similarly $|\mathfrak{a}|$ is the order of the residue class ring $R_x/\mathfrak{a}$ for any nonzero ideal $\mathfrak{a} \subset R_x$. The only difference to (5) is that now the field F is not assumed to be quadratic over the rational function field $K(x)$. In his letter, F. K. Schmidt states the limit formula under the simplifying assumption that the infinite place of $K(x)$ splits completely in F; this is the analogue of a totally real number field. This implies that the number r_x of infinite places in F equals the degree $[F : K(x)]$, and that F is separable over $K(x)$. F. K. Schmidt's limit formula is:

$$\lim_{s \to 1} (s-1)\zeta_x(s) = \frac{(q-1)^{r_x - 1} \cdot \mathcal{R}_x h_x}{\sqrt{|d_x|} \cdot \log q} \tag{41}$$

which is the same formula as Artin's (12) in the "real" quadratic case. In our situation here, $d_x \in K[x]$ denotes the discriminant of R_x over $K[x]$ and $|d_x| = q^{\deg d_x}$. Moreover, $\mathcal{R}_x$ is the "regulator" which, F. K. Schmidt says, is defined completely analogous to the number field case, namely as the absolute value of a certain determinant, the only difference being that the logarithms appearing in the number field case have to be replaced here by the suitably normalized order functions for the places at infinity. This last observation we have found already in F. K. Schmidt's thesis, in the discussion of Dirichlet's unit theorem for function fields.

F. K. Schmidt does not give any proof in his letter to Hasse. Instead he says:

> "*On the suggestion of Prof. Haupt, a preliminary announcement about my results and methods is to appear shortly in the Erlangen Reports. Since there the printing will be very fast I will soon have reprints available, and then I will venture to send you a copy.*"

The "Erlangen Reports" (*Erlanger Berichte*) is a local journal of Erlangen University. It contains not only mathematics but also articles from other fields including medical sciences, hence it would perhaps not be read by many mathematicians. [66] But it had the advantage that it could publish very quickly, and this may have induced Haupt to recommend this journal to F. K. Schmidt.

Otto Haupt held a position as full professor at Erlangen University. Although his primary interests were in real analysis and geometry, he was also keenly interested in the modern developments of algebra and number theory. Haupt kept contact with Emmy Noether who whenever she visited her home town Erlangen, was heartily welcomed in the Haupt residence. [67] From the remarks in F. K. Schmidt's letter we infer that Haupt was impressed by F. K. Schmidt's work and therefore wished to secure priority for him in publication, in particular in view of F. K. Schmidt's earlier experiences with his thesis.

Already in November that year there appeared F. K. Schmidt's note [FK:1926], under the title:

On number theory in fields of characteristic p.
(Preliminary announcement.)

It is signed by the author with "August 1926", which is the month he wrote his letter to Hasse.

In this preliminary announcement a proof of the limit formula (41) is sketched, proceeding parallel to the proof of the limit formula in the number field case. Moreover, F. K. Schmidt presents a very simple proof of the power reciprocity law (38) which we already discussed in section 4.4.2. But the most important item of this note is the 3-page "Remark added in proof". There, F. K. Schmidt announces that after some deliberation he has now decided to change his viewpoint.

He announces that now he prefers to use the birationally invariant zeta function $\zeta_F(s)$ whose definition we have stated in formula (1) already; this function contains contributions from *all* places $\mathfrak{p}$ of F, including the poles of

[66] But it had been read by Oystein Ore [Ore:1934] in far away New Haven, as we have reported in section 4.4.2 above. Perhaps F. K. Schmidt had sent him a reprint? Ore regularly exchanged reprints with Emmy Noether and she might have turned his attention to F. K. Schmidt's paper.

[67] Inspired by the discussions with Emmy Noether, Otto Haupt wrote a textbook on the then "modern" algebra [Hau:1929] which was the first such textbook, before van der Waerden's appeared. Haupt's book covered more material than van der Waerden's; the fact that the latter became more widely known than the former seems to be due to the style of writing.

x. He refers to the paper of Dedekind-Weber [De-Web:1882] who first had developed a birational invariant theory of function fields. He mentions that already in his thesis [FK:1925] he had pointed out how to generalize Dedekind-Weber's notion of point, or prime divisor, for function fields with finite base field. The limit formula for his new function $\zeta_F(s)$ is of a particularly simple form:

$$\lim_{s\to 1}(s-1)\zeta_F(s) = \frac{q^{1-g}\cdot h}{(q-1)\log q} \tag{42}$$

where $h = h_F$ is the class number of the function field F, defined to be the number of divisor classes (modulo principal divisors) of degree 0. Moreover, $g = g_F$ is a new invariant which F. K. Schmidt says is quite analogous to the *genus* of a field of algebraic functions. Until then, the genus had been defined for fields of algebraic functions over the complex numbers only or, what amounts to the same, for compact Riemann surfaces.

The zeta function $\zeta_x(s)$ which F. K. Schmidt had defined in the first part of his note, and the new zeta function $\zeta_F(s)$ differ only by the finitely many Euler factors belonging to the poles of x, i.e.,

$$\zeta_F(s) = \zeta_x(s)\cdot \prod_{x(\mathfrak{p})=\infty} \frac{1}{1-|\mathfrak{p}|^{-s}}\,. \tag{43}$$

Hence it is easy to interpret any statement about one of these zeta functions in terms of the other. For instance, the limit formula (42) for the new zeta function $\zeta_F(s)$ can be used to obtain the limit formula (41) for $\zeta_x(s)$, if one uses the following facts:

$$2g-2 = \deg(\mathfrak{D}_x) - [F:K(x)] \tag{44}$$

where $\mathfrak{D}_x$ denotes the different of $F|K(x)$ (including the contributions of the primes at ∞). Moreover,

$$h = \frac{h_x\mathcal{R}_x}{u_x} \tag{45}$$

where on the right hand side the numbers h_x, $\mathcal{R}_x$ are those which appear in (41). The number u_x is defined to be the greatest common divisor of the degrees of the primes at ∞; it does not appear in (41) because there it was assumed that the infinite prime of $K(x)$ splits completely in F which implies $u_x = 1$. Thus, as F. K. Schmidt points out, the simple limit formula (42) is the source for all of Artin's formulas (11). [68] It is not necessary, as Artin did, to compute the residues in each of those cases separately. [69]

F. K. Schmidt refers to Hilbert's famous Paris lecture [Hil:1900], where Hilbert had mentioned the class number h and the genus g as being analogous notions

[68] In those formulas, the index x is to be added at the proper places. In the "imaginary" case, $\mathcal{R}_x = 1$.

[69] Remark to formula (45): Artin in his thesis was concerned with the class number h_x in quadratic function fields $F = K(x,\sqrt{D})$. He had determined all "imaginary" quadratic fields with $h_x = 1$ and characteristic > 3; see section 2.1.2, **6.** For "imaginary" quadratic fields we have $\mathcal{R}_x = 1$ and, in order that $h_x = 1$ we have necessary $u_x = 1$ and hence $h = 1$. The determination of all function fields (quadratic or not) with $h = 1$ has been carried out by M. L. Madan and his collaborators; see [Mad:1975].

in the theories of algebraic numbers and of algebraic functions. (Hilbert mentioned this in the discussion of his 12th problem). In the present context, where number theory and function theory are joining, both h and g appear in the same formula (42).

One important fact is missing in F. K. Schmidt's preliminary note [FK:1926], namely the functional equation of his zeta function. This, and all the proofs including the Riemann-Roch theorem, are contained in the final version of the paper which constitutes F. K. Schmidt's main contribution.

5.2 The main contribution

The final version [FK:1931] appeared five years later only. But it was already completed in the summer of 1927 when F. K. Schmidt presented to the Faculty in Erlangen his *Habilitationsschrift* with the title:

Abelian fields in the domain of higher congruences.

That paper consisted of two parts:

Part I: *Analytic number theory in fields of characteristic p.*
Part II: *Class field theory for an algebraic function field of one variable with finite base field.*

On inspection, Part I turns out to be identical with the paper [FK:1931] published under the same title. The second part is the same as [FK:1931a]. We have not been able to find out why F. K. Schmidt had delayed the publication for several years.

One of the reasons for the delay may be that F. K. Schmidt's original aim was to transfer class field theory to function fields. This had been suggested to him by Hasse. In fact, already in the preliminary announcement [FK:1926] he had said that he was heading towards this aim, and for this he needed the analytic theory of the zeta function and the L-functions. However, F. K. Schmidt was not able to overcome the essential difficulties which are connected with abelian fields of p-power degree where p is the characteristic. Those fields cannot be treated by just transferring the classical methods. Accordingly, in Part II of his *Habilitationsschrift* where he does class field theory, he had to restrict his investigations to those abelian extensions whose degree is not divisible by p.

So F. K. Schmidt may have wished to wait with the publication of his *Habilitationsschrift* until he could overcome those difficulties. Finally, when he realized that he was not able to overcome the difficulties with p-power degrees, then he may have decided to send his papers to publication in the form as they had been completed in 1927 already.

In any case, the first part of F. K. Schmidt's *Habilitationsschrift*, i.e., the paper [FK:1931], became fundamental for the development of the theory of function fields.

5.2.1 The Riemann-Roch theorem

As stated in the preliminary announcement, F. K. Schmidt in [FK:1931] works in a birational setting. But his former results in his thesis [FK:1925] depended on the choice of a transcendental element $x \in F$ and hence they were not of birational nature. For this reason he now in [FK:1931] decided to develop the arithmetic of function fields anew. Accordingly the notion of *divisor* instead of *ideal* is put into the focus of attention. A divisor is a formal power product of prime divisors, or points, in the sense of Dedekind-Weber. [70]

The first main achievement of F. K. Schmidt is the discovery that the classical theorem of Riemann-Roch [71] on compact Riemann surfaces can be transferred to function fields with finite base field. Actually, his proof of the Riemann-Roch theorem works for arbitrary perfect base fields, not necessarily finite.

In order to formulate the Riemann-Roch theorem, he first defines the genus g of the function field F. To do this, F. K. Schmidt chooses a separating element $x \in F$ and defines g by means of the formula (44). It is easy to see that g is an integer ≥ 0; the main point is to show that it does not depend on the choice of the separating element x. To this end he shows that the principal divisor of the differential quotient $\frac{dy}{dx}$ for two separating elements $x, y \in F$ is as follows:

$$\frac{dy}{dx} \cong \frac{\mathfrak{D}_y \mathfrak{n}_x^2}{\mathfrak{D}_x \mathfrak{n}_y^2} \tag{46}$$

where $\mathfrak{D}_x$ denotes, as above, the different of $F|K(x)$ and $\mathfrak{n}_x$ is the pole divisor of x in F. Since principal divisors have degree 0 we conclude

$$\deg(\mathfrak{D}_x) - 2\deg(\mathfrak{n}_x) = \deg(\mathfrak{D}_y) - 2\deg(\mathfrak{n}_y)\,.$$

On the left hand side we have $2g-2$ in view of the definition (44) and so indeed, the number g does not depend on the choice of the separating element $x \in F$.

F. K. Schmidt's proof of (46) requires to use the formal properties of the derivation $y \to \frac{dy}{dx}$ including the chain rule. [72] He does not prove those properties but says that they can be proved "as usual" if the base field K is algebraically closed, and if this is not the case then one can do base field extension.

From (46) we see that the divisors $\frac{\mathfrak{D}_x}{\mathfrak{n}_x}$ and $\frac{\mathfrak{D}_y}{\mathfrak{n}_y}$ are equivalent, for they differ only by the principal divisor of $\frac{dy}{dx}$. Hence the divisor class W of $\frac{\mathfrak{D}_x}{\mathfrak{n}_x}$ does not depend on the choice of the separating element x; it is called the *differential class* of F and any of its divisors $\mathfrak{w}$ is called a *differential divisor*. [73] Sometimes the terminology *canonical class* is used, in particular in a geometric setting. Its degree is $2g-2$.

[70] Dedekind-Weber [De-Web:1882] used the terminology "polygon" instead of "divisor". More precisely, "polygon" is used by Dedekind-Weber for "integral divisor" whereas an arbitrary divisor is called "polygon quotient". The word "divisor" is used by Hensel-Landsberg [Hen-La:1902].

[71] F.K. Schmidt always writes "Roche" in his paper, instead of "Roch". This could possibly lead to confusion because the mathematicians Roch and Roche are not identical. In a postcard to Hasse dated January 4, 1934 he apologizes for his mistake of constantly appending an *'e'* to the name of Roch. And he adds somewhat jokingly: "*Unfortunately, this constant 'e' is recently appended also by other people, probably following my example. Thus again we observe that evil begets evil.*"

[72] F. K. Schmidt does not speak of "derivation" but just says "differential quotient".

[73] E. Witt, who had attended F. K. Schmidt's lectures on function fields in the winter semester 1933/34 at Göttingen, presented in [Wi:1935] what he calls a simplification of this invariance proof. He explicitly refers to §4 of F.K. Schmidt's paper [FK:1931] and proposes to replace that section by his (Witt's) proof. The "simplification" of Witt consists essentially of proving, in the algebraic setting including characteristic p, the well known explicit formula for the divisor of a differential, whereas F.K. Schmidt works with derivations only, not with differentials. – Independently of Witt and at the same time, Hasse [H:1935a] gave the same proof in the framework of his general theory of differentials.

This being said, F. K. Schmidt is now ready to state the Riemann-Roch theorem. Every divisor $\mathfrak{a}$ of F determines a module $\mathcal{L}(\mathfrak{a})$, consisting of all functions whose principal divisor is a multiple of $\mathfrak{a}^{-1}$. Its K-dimension is called the dimension of $\mathfrak{a}$, denoted by $\dim(\mathfrak{a})$. This dimension depends only on the class of $\mathfrak{a}$ (modulo principal divisors). The Riemann-Roch theorem gives a formula for $\dim(\mathfrak{a})$ in terms of the degree:

$$\dim(\mathfrak{a}) = \deg(\mathfrak{a}) - g + 1 + \dim\left(\frac{\mathfrak{w}}{\mathfrak{a}}\right) \tag{47}$$

where g is the genus of the function field F and $\mathfrak{w}$ is any differential divisor.

In the case of an algebraic function field over the complex numbers, an algebraic proof of the Riemann-Roch theorem had been given in the book of Hensel-Landsberg [Hen-La:1902]. The main idea was to work with so-called "normal bases" which had already been used by Dedekind-Weber. They can be defined as follows: Let x be a separating element of F. Let $\mathcal{L}_x(\mathfrak{a})$ consist of those functions in F which are multiples of $\mathfrak{a}^{-1}$ up to a divisor which contains only poles of x. This is a (fractional) R_x-ideal. Similarly for $\frac{1}{x}$ instead of x. Then

$$\mathcal{L}(\mathfrak{a}) = \mathcal{L}_x(\mathfrak{a}) \cap \mathcal{L}_{1/x}(\mathfrak{a})\,.$$

Now, there exists a $K[x]$-basis $u_1, \dots u_n$ of $\mathcal{L}_x(\mathfrak{a})$ and a $K[1/x]$-basis $v_1, \dots, v_n$ of $\mathcal{L}_{1/x}(\mathfrak{a})$ such that

$$x^{e_i} u_i = v_i \qquad\qquad (1 \le i \le n)\,.$$

with certain integers $e_i \in \mathbb{Z}$, some of which may be < 0. The u_i are called a *normal basis* of $\mathcal{L}_x(\mathfrak{a})$. And then

$$\dim(\mathfrak{a}) = \sum_{e_i \ge 0} (e_i + 1)\,.$$

In the book of Hensel-Landsberg, this formula is the key for the algebraic proof of the Riemann-Roch theorem.

F. K. Schmidt discovered that this proof can be transferred to function fields with arbitrary perfect base field. For this he had to overcome two obstacles. The first one was that not every prime divisor or point of F is necessarily of degree 1. Today we are quite used to work with a situation where points may have higher degrees, but in the time of F. K. Schmidt this created problems sometimes. In particular it was not possible to copy the proof of Hensel-Landsberg word for word. F. K. Schmidt solved this problem by extending the base field suitably whenever this seemed necessary to him, so that the prime divisors under discussion split into primes of degree 1. In order that this method would work he had to verify that dimension and degree of a divisor do not change after extending the base field.

In today's terminology, he proved that a function field with perfect base field is *conservative*. This is a fact which is of interest independent of the proof of the Riemann-Roch theorem. Moreover, we see that F. K. Schmidt's proof of the Riemann-Roch theorem works for every conservative function field $F|K$, regardless of whether the base field K is perfect or not.

The second obstacle was connected with the structure of the different $\mathfrak{D}_x$ of $F|K(x)$. In the case of Hensel-Landsberg the characteristic is 0 and, hence, the ramification of $F|K(x)$ is tame; in this case the structure of the different is easily described by the ramification degrees. This had been essentially used in Hensel-Landsberg [Hen-La:1902]; in fact it had been used there already in the *definition* of the different. In characteristic $p > 0$ however, there may be wild ramification. F. K. Schmidt defines the different divisor $\mathfrak{D}_x$, following Dedekind, by means of the trace property. This takes care of wild ramification. But in order to be able to transfer the arguments of Hensel-Landsberg to the case at hand, F. K. Schmidt chooses the separating element x suitably, so that the prime divisors under discussion are tamely ramified in $F|K(x)$. Today this seems to us quite unnecessary but for F. K. Schmidt this was the key for success.

F. K. Schmidt's proof, freed from those unnecessary artificial detours, has later been included into Hasse's book "Number Theory" [H:2002]. Nowadays there are other, more direct proofs which do not require the use of a separating element x; this development was initiated by F. K. Schmidt himself in [FK:1936], followed by A. Weil in [W:1938]. See also [Rq:1958].

One important consequence of the Riemann-Roch theorem is its so-called Riemann part which says that for divisors of sufficiently large degree their dimension is directly computable by their degree, more precisely:

$$\dim(\mathfrak{a}) = \deg(\mathfrak{a}) - g + 1 \qquad \text{if} \qquad \deg(\mathfrak{a}) > 2g - 2\,. \tag{48}$$

This follows immediately from the Riemann-Roch theorem (47) since the differential divisor $\mathfrak{w}$ is of degree $2g - 2$; hence $\frac{\mathfrak{w}}{\mathfrak{a}}$ is of degree < 0 and thus its dimension is 0.

F. K. Schmidt does not state this formula explicitly but he uses it while deriving the properties of his zeta function. Erroneously he says that (48) holds for $\deg(\mathfrak{a}) \geq 2g - 2$ which obviously is not true. (Take $\mathfrak{a} = \mathfrak{w}$.) This error is perpetuated through the whole paper and also through F. K. Schmidt's other paper [FK:1931]. Hasse [H:1934] calls this an "annoying misprint" (*ein störender Druckfehler*).

If we compare this paper of F. K. Schmidt [FK:1931] with the other papers which we have discussed before, i.e., with the papers of Artin, Sengenhorst and Rauter, then we observe that here quite different ideas and notions are at work. Whereas the ideas in the former papers had been transferred from number theory, here we observe a transfer from the theory of complex valued functions. This had already started in F. K. Schmidt's thesis. Much of the later work of Hasse and others uses this analogy to the theory of complex valued functions. Hasse used to call this the "algebraic theory of functions" (*Algebraische Funktionentheorie*). Let us cite from Hasse's report [H:1942b]:

> "*Number Theory owes to the Algebraic Theory of Functions a methodology which has been formed by analogy, and which in the past decades has led to a remarkable enrichment and refinement of the arithmetic theories; in addition it has led to a number of new, interesting results whose proof has its roots in the Algebraic Theory of Functions.*"

Today it has become customary to speak of "Algebraic Geometry of Curves" instead of "Algebraic Theory of Functions".

5.2.2 Theory of the zeta function

F. K. Schmidt's second main achievment in [FK:1931] is the discovery that over a finite base field, the Riemann-Roch theorem is intimately connected with the properties of his zeta function (1). In the second part of the paper he gives the following fundamental results as almost immediate consequences of the Riemann-Roch theorem.

1. Rationality. *$\zeta_F(s)$, as defined by (1), becomes a rational function with respect to the variable $t = q^{-s}$, with poles of order 1 for $t = 1$ and $t = q^{-1}$ of order 1.*

To show this, F. K. Schmidt writes $|\mathfrak{a}|^{-s} = t^{\deg(\mathfrak{a})}$ and so $\zeta_F(s) = \sum_{n\geq 0} c_n t^n$ where c_n denotes the number of integral divisors of degree n. These divisors belong to h divisor classes, and for each divisor class C the number of integral divisors in C is $\frac{q^{\dim(C)}-1}{q-1}$. Now, if $\deg(C) = n > 2g-2$ then by the Riemann part (48) of the Riemann-Roch theorem we have $\dim(C) = n - g + 1$. It follows

$$\zeta_F(s) = \sum_{0\leq n\leq 2g-2} c_n t^n + h \cdot \sum_{n>2g-2} \frac{q^{n-g+1}-1}{q-1}\, t^n \tag{49}$$

$$= \sum_{0\leq n\leq 2g-2} c_n t^n + \frac{h \cdot t^{2g-1}}{q-1}\left(\frac{q^g}{1-qt} - \frac{1}{1-t}\right) \tag{50}$$

where the last formula (50) is valid for $g > 0$ only; in the trivial case $g = 0$ the appropriate modifications have to be done.

There is some minor but important detail to be cleared up. Namely, it is not clear *a priori* that there exist divisors in $F|K$ of every degree n. Hence one has to work provisionally with the smallest positive divisor degree $m \geq 1$ of $F|K$. The summation index n in formula (49) ranges over the multiples of m, and so $\zeta_F(s)$ becomes a rational function of t^m, with denominator $(1-t^m)(1-q^m t^m)$. But still, every pole is of order 1. In order to show that $m = 1$, F. K. Schmidt uses the following "trick" (as Artin would have called it). He considers the unique base field extension $F^{(m)}$ of degree m. Then every prime $\mathfrak{p}$ of F splits in $F^{(m)}$ in exactly m different primes: $\mathfrak{p} = \mathfrak{P}_1 \cdots \mathfrak{P}_m$. Therefore the product expansion of the zeta function in (1) gives $\zeta_{F^{(m)}}(s) = \zeta_F(s)^m$. Comparing the pole orders of both sides yields $m = 1$. Hence the following important

2. Theorem of F. K. Schmidt. *Every function field F with finite base field K admits a divisor of degree 1.*

If $g = 0$ then this implies that $F = K(x)$ is a rational function field over K. The fact that this algebraic statement had been proved by analytical means, namely with the help of the zeta function, was felt to be quite remarkable at that time. In a letter to Hasse of January 21, 1933 F. K. Schmidt wrote:

> "*...I show that the Riemann hypothesis always holds in case of genus 0. Of course this is immediate as soon as F, in case of genus 0, has been identified as a field of rational functions in one indeterminate... Known fact in case of algebraic functions with complex coefficients! But here, where the base field is finite, by no means trivial, so far not even provable with algebraic methods but only with transcendental methods...*"

By "transcendental" methods he means methods that involve the zeta function of the field. Some months later a non-transcendental proof of F. K. Schmidt's theorem was found by Witt. In August 7, 1933 F. K. Schmidt wrote to Hasse:

> "*...Some days ago Witt wrote that for an algebraic function field with finite base field he could establish a divisor of degree* 1 *with purely arithmetic methods. But his proof is longer than my analytic proof. Unfortunately he did not include his proof in the letter...*"

Witt included his proof in his paper [Wi:1934b]. From today's point of view Witt used Galois cohomology for cyclic Galois groups. With the proper prerequisites from cohomology Witt's proof can be made even shorter than F. K. Schmidt's.

3. Limit formula. Inspection of (50) leads straightforwardly to the limit formula (42) which F. K. Schmidt had already announced in his preliminary announcement [FK:1926].

4. Functional equation. $\zeta_F(s)$ *satisfies the functional equation*

$$\zeta_F(1-s) = q^{(g-1)(2s-1)}\zeta_F(s)\,. \tag{51}$$

Again, F. K. Schmidt proves this by inspection of the expansion (50), this time using the full Riemann-Roch theorem (47), not only the Riemann part as in the proof of **1.** It does not seem necessary to go into details here. [74]

5. Non-invariant zeta functions. Let $x \in F$ be transcendental and R_x the integral closure of $K[x]$ in F. The zeta function $\zeta_x(s)$ of R_x is defined by F. K. Schmidt as in his preliminary announcement; see (40). This zeta function is connected with $\zeta_F(s)$ of the field by means of (43). Hence the properties of $\zeta_x(s)$ can be deduced from those of $\zeta_F(s)$. F. K. Schmidt had already announced this in [FK:1926]; here he presents the proofs which are straightforward.

Summary

The theory of zeta functions for arbitrary function fields was started by F. K. Schmidt, in generalization of the second part of Artin's thesis. After a preliminary announcement in 1926 F. K. Schmidt decided to build the theory from a birational point of view, in the spirit of the classical work of Dedekind-Weber of 1882. The basic notions are those of "prime divisor" or "point" of a function field, and of "divisor". F. K. Schmidt proved the Riemann-Roch theorem for divisors of algebraic function fields with perfect base fields. If the base field is finite then he exhibited a close relationship between the Riemann-Roch theorem and the properties of his zeta function, in particular the functional equation. This is to be regarded as his main achievement. His paper became a classic and is generally considered as the beginning of what Hasse later called the "Algebraic Theory of Function Fields."

6 Hasse's Survey

The general objective of F. K. Schmidt had been to transfer class field theory from number fields to function fields over finite base fields. The Riemann hypothesis was not mentioned, neither in F. K. Schmidt's paper [FK:1931] nor in his correspondence with Hasse during those years.

[74] In our report [Rq:2001] we have presented the very elegant method of Witt for arranging this computation.

Nevertheless, when Hasse started in 1932/33 to work on the Riemann hypothesis then he could fall back on F. K. Schmidt's theory as a basis. In the present section we will discuss Hasse's paper [H:1934] entitled

"*On congruence zeta functions. Based in part on informations by Prof. Dr. F. K. Schmidt und Prof. Dr. E. Artin.*"

We do not know precisely when this paper was written. In an earlier version of this Part 1 we had said that Hasse wrote this paper in the year 1934. But the evidence which has become available in the meantime now admits also the possibility that it was written in February 1933 or shortly after. (See Part 2 .) In any case, Hasse wished to collect as much information as possible about zeta functions of function fields, and therefore he had asked F. K. Schmidt and Artin what they knew about it. And then, in order to have a solid foundation for later reference he wrote the survey paper [H:1934] where he presented the theory of F. K. Schmidt's zeta functions *ab ovo* in a systematic way, including (with proper citations) the additional informations he had obtained from Artin and F. K. Schmidt.

There are only few proofs in the paper. Nevertheless it can be regarded as a good reference, offering a survey of all what was known at the time about zeta functions of function fields. This is the reason why we have decided to include the discussion of this paper into the present Part 1 of our work where we wish to report about the formation of the zeta functions.

6.1 The L-polynomial.

Hasse starts with a brief but fairly complete review of F. K. Schmidt's paper [FK:1931] which we have just discussed in section 5.2. And then he introduces the L-polynomial for F. K. Schmidt's ζ-function by the formulas

$$\zeta_F(s) = \frac{L(t)}{(1-qt)(1-t)} \qquad \text{with} \qquad t = q^{-s} \tag{52}$$

where $L(t)$ is a polynomial of degree $2g$ of the form

$$L(t) = 1 + (N - q - 1)t + \cdots + q^g t^{2g}\,, \tag{53}$$

N denoting the number of points of F of degree 1. (Recall that g is the genus of F.) Hasse says that these formulas had been communicated to him by F. K. Schmidt.

We have already mentioned the letter of F. K. Schmidt to Hasse dated Jan 21, 1933; see section 5.2.2, no.2. That letter was accompanied by a manuscript where F. K. Schmidt gave detailed information on what he knew about the ζ-function of a function field $F = K(x, y)$ over a finite field K. That manuscript seems to be lost but it is very probable that the formulas (52),(53) were contained therein. They are straightforward consequences of (50) from F. K. Schmidt's paper [FK:1931].

F. K. Schmidt's formulas (52),(53) should be compared with Artin's formulas (9), (10) in his thesis. Note that Artin did not use the birational invariant zeta function and, hence, what had been called $\zeta(s)$ there should now be called $\zeta_x(s)$. But F. K. Schmidt had already given general rules (43)–(45) how to rewrite formulas for $\zeta_x(s)$ in terms of $\zeta_F(s)$. If one applies this to Artin's formulas, there will appear F. K. Schmidt's (52), (53). Thereby one has to take into account the trivial roots of Artin's L-polynomial $L_x(t)$, and the genus formula for the quadratic function field $F = K(x, \sqrt{D})$ with discriminant $D \in K[x]$ of degree n :

$$2g = \begin{cases} n-1 & \text{if } n \equiv 1 \mod 2 \\ n-2 & \text{if } n \equiv 0 \mod 2 \,. \end{cases}$$

In this way Artin's formulas (9), (10) for quadratic function fields appear as very special cases of F. K. Schmidt's (52), (53) which are valid for arbitrary function fields.

Note that the formula (52) implies *F. K. Schmidt's Theorem* on the existence of a divisor of degree 1 (see section 5.2.2, **2.**).

The functional equation (51) for $\zeta_F(s)$ relates to the substitution $s \mapsto 1-s$. In terms of t this means $t \mapsto \frac{1}{qt}$. Thus (51) yields the functional equation for $L(t)$ as follows:

$$L\left(\frac{1}{qt}\right) = q^{-g}t^{-2g}L(t)\,. \tag{54}$$

For the coefficients σ_ν of $L(t)$ this leads to

$$\sigma_{2g-\nu} = q^{g-\nu}\sigma_\nu\,. \tag{55}$$

These are precisely the symmetry relations (13) which Artin had established in the case of quadratic fields and used to compute his class number tables.

The limit formula (42) relates to the limit $s \to 1$ which means $t \to q^{-1}$. By means of the functional equation this transforms in a formula for $t \to 1$. One obtains the class number formula

$$h = L(1)\,. \tag{56}$$

6.2 The inverse roots

Hasse writes the L-polynomial $L(t)$ in the product form

$$L(t) = \prod_{1 \le i \le 2g} (1 - \omega_i t) \tag{57}$$

where the ω_i are the inverse roots of $L(t)$. Thus

$$h = \prod_{1 \le i \le 2g} (1 - \omega_i) \tag{58}$$

which expresses the class number by means of the inverse roots. (The Riemann hypothesis is equivalent with the fact that $|\omega_i| = q^{1/2}$ for $1 \le i \le 2g$.)

Hasse states the estimate

$$1 < |\omega_i| < q \qquad \text{for} \qquad 1 \le i \le 2g\,. \tag{59}$$

The weaker inequality $|\omega_i| \le q$ follows immediately from the fact that the product in the defining formula (1) converges for $\Re(s) > 1$ and hence $\zeta_F(s)$ has no zero in that region. For the stronger inequality $|\omega_i| < 1$ one has to verify that

$\zeta_F(s)$ admits no zero on the line $\Re(s) = 1$. Hasse says that this can be proved with analytic methods in precisely the same way as for the Riemann zeta function or the Dedekind zeta function in the number field case. We are reminded that Artin in his thesis had forwarded the same argument for his zeta function, thereby referring to Landau's book [Lan:1918]; see section 2.1.2. The inequality $1 < |\omega_i|$ is then obtained via the functional equation.

Let q^θ denote the maximum of the absolute values $|\omega_i|$ $(1 \leq i \leq 2g)$. From (59) it follows $\theta < 1$. From this and by the functional equation we conclude $\theta \geq \frac{1}{2}$. From (57) it follows for the first coefficient of $L(t)$ that

$$N - q - 1 = -(\omega_1 + \cdots + \omega_{2g}) \,. \tag{60}$$

and hence

$$|N - q - 1| \leq 2gq^\theta \qquad \text{with} \qquad \frac{1}{2} \leq \theta < 1 \,. \tag{61}$$

(The Riemann hypothesis is equivalent with $\theta = \frac{1}{2}$.)

6.3 Base field extension

Hasse goes on to state the formula for the behaviour of the zeta function under base field extension. Let $K^{(m)}$ denote the extension of K of degree m and $F^{(m)} = FK^{(m)}$ the corresponding base field extension of the function field F. Then

$$\zeta_{F^{(m)}}(s) = \prod_{0 \leq \mu \leq m-1} \zeta_F\left(s + \frac{2\pi i \mu}{m \log q}\right) \tag{62}$$

which is for the F. K. Schmidt zeta function the same formula as Artin's formula (17) for his zeta function.

Hasse states that (62) had been communicated to him by Artin. Recall that the corresponding formula (17) for Artin's zeta function had never been published; it had been stated by Artin in a letter to Herglotz. Among the letters of Artin to Hasse we find none which deals with this formula. We conclude that Artin had orally informed Hasse about it, at one of the frequent occasions when they met. In fact, there is a letter of Hasse to Davenport dated Dec 7, 1932 in which he reported about a colloquium lecture which he had recently given in Hamburg, and about Artin's comment to Hasse's results. We conclude from this that it was on this occasion that Artin informed Hasse about his formula (17). See Part 2.

The proof of (62) is obtained by comparing $\zeta_{F^{(m)}}(s)$ and $\zeta_F(s)$ in their Euler products, and applying the known splitting behavior of the prime divisors $\mathfrak{p}$ of F in the base field extension $F^{(m)}$.

Let $L^{(m)}(t)$ denote the L-polynomial for the field $F^{(m)}$. From (62) it follows that the $2g$ inverse roots of $L^{(m)}(t)$ are the m-th powers ω_i^m of the inverse roots ω_i of $L(t)$. Hence

$$L^{(m)}(t) = \prod_{1 \leq i \leq 2g} (1 - \omega_i^m t) \tag{63}$$

and for the number $N^{(m)}$ of prime divisors of degree 1 in $F^{(m)}$ we obtain the estimate

$$|N^{(m)} - q^m - 1| = |\omega_1^m + \cdots + \omega_{2g}^m| \leq 2gq^{m\theta} \tag{64}$$

with the same θ as in (61). Hasse states and proves the following lemma, also due to Artin:

Suppose it is known that for $m \to \infty$ there is an estimate

$$|N^{(m)} - q^m - 1| \leq A \cdot q^{m\gamma} \tag{65}$$

with constants $A > 0$, $\gamma < 1$ which do not depend on m. Then we have $\theta \leq \gamma$, which is to say that $|\omega_i| \leq q^\gamma$ for $1 \leq i \leq 2g$. And A can be replaced by $2g$.

In one of his letters to Herglotz, Artin had written something similar for his zeta functions in the quadratic case. Compare (19). Artin's proof of this lemma is applicable also to F. K. Schmidt's zeta function and proceeds by formal power series expansion of the logarithm, starting from (63):

$$\begin{aligned} -t\,\frac{d\,\log L(t)}{dt} &= \sum_{1\leq i\leq 2g}\sum_{0\leq m}\omega_i^m t^m \\ &= -\sum_{0\leq m}(N^{(m)} - q^m - 1)t^m\,. \end{aligned}$$

Consequently if (64) holds then we conclude that the power series on the right hand side is convergent for $q^\gamma|t| < 1$. In particular, this region does not contain any zero of $L(t)$, i.e. $q^\gamma|\omega_i^{-1}| \geq 1$ which means $|\omega_i| \leq q^\gamma$ $(i = 1, \ldots 2g)$. Hence $\theta \leq \gamma$ as contended. [75]

Clearly, this lemma aims at the Riemann hypothesis. For, when the hypothesis of the above lemma can be verified for $\gamma = \frac{1}{2}$ then $\theta = \gamma = \frac{1}{2}$.

But Hasse does not say anything explicit about the Riemann hypothesis in this paper. This supports our conclusion, already mentioned above, that Hasse had completed this survey paper *before* he had discovered his first proof in the elliptic case; that event can be dated at the end of February 1933. *After* that discovery, in his publications Hasse expressed his confidence that the Riemann hypothesis is valid generally. For instance, in his first preliminary announcement [H:1933] he wrote:

> "*... it is to be expected that ... the analogue of the Riemann hypothesis for the general congruence zeta functions of F. K. Schmidt can be proved ...*"

And one year later, in the introduction of his manuscript for Hasse's Hamburg lectures [H:1934a], he said that he regards the elliptic case as the simplest nontrivial case of a general theory.

6.4 Diophantine congruences

In the last section of his survey [H:1934] Hasse leaves the view point of birational invariance. Let x, y be elements which generate the function field:

$$F = K(x, y) \qquad \text{with} \qquad f(x, y) = 0$$

where $f(x, y)$ is an absolutely irreducible polynomial with coefficients in K. Let N_f denote the number of solutions (a, b) of $f(a, b) = 0$ which are rational in

[75] Hasse, with a slightly different argument, proved the lemma under the weaker assumption that (65) holds for an arbitrary given sequence of numbers m tending to ∞.

K, i.e., $a, b \in K$. This number N_f is compared with the number N of prime divisors of degree 1 of F. It is shown that

$$0 \le N - N_f \le C \tag{66}$$

with a constant C depending on the degree and the singularity degree of $f(x,y)$ only; in particular it is independent of the base field K. This had been verified by F. K. Schmidt on Hasse's request. The corresponding letters of F. K. Schmidt to Hasse are dated January and February 1933 and are preserved.

It follows from (66): *If an estimate of the form* (65) *is proved for N then such estimate holds also for N_f, with the same γ but possibly with another constant A. This implies that*

$$N_f = q + O(q^\gamma) \qquad \text{for} \qquad q \to \infty\,. \tag{67}$$

(In particular: If the Riemann hypothesis is proved for the zeta function of F then one can take $\gamma = 1/2$.)

The very fact that Hasse includes a result like this in his paper shows that he has become interested in counting the solutions of congruences $f(x,y) \equiv 0 \bmod p$ for polynomials $f(x,y) \in \mathbb{Z}[x,y]$ which are absolutely irreducible. And this shows the influence of Davenport who had worked on such diophantine congruences for elliptic equations $f(x,y)$ but reached only some exponent $\gamma = \frac{5}{8} > \frac{1}{2}$. [76] Hasse wished to state clearly that quite generally, the Riemann hypothesis (or weaker estimates of the roots of F. K. Schmidt's zeta functions) implies asymptotic estimates of the the number of solutions of diophantine congruences.

One can say that this discovery, namely the connection of the Riemann hypothesis with the problem of asymptotic estimate of diophantine congruences, had put the Riemann hypothesis into the focus of attention of number theorists. Before that, as we have seen, the theory of the zeta function was developed by F. K. Schmidt not with any intention towards the Riemann hypothesis but as a tool for establishing class field theory for function fields. The validity of the Riemann hypothesis which Artin had verified in the quadratic case in a number of examples, was just seen as a "curious fact", as Herglotz had expressed it in his report about Artin's thesis. It is true that Artin had already stated some conditional results, i.e., he had drawn consequences under the assumption that the Riemann hypothesis holds for quadratic number fields. But as we have seen in section 2.1.2, this went well along classical lines, like the asymptotic estimate of density of primes in arithmetic progressions, or finiteness of the number of "imaginary" quadratic fields with fixed class number. But now, a host of arithmetic problems which hitherto could be dealt with by insufficient means only, could be solved in one stroke if the Riemann hypothesis could be proved.

Let us cite from the introduction of the manuscript for Hasse's Hamburg lecture [H:1934a]. After he has explained the problem of estimating the number of solutions of binary diophantine congruences and mentioned several authors [77] who had contributed to this problem, he says:

> "*These problems will all be solved in the best possible way if the analogue to the Riemann hypothesis in the algebraic theory of function*

[76] More on the cooperation of Hasse with Davenport will be found in Part 2.

[77] He mentions A. Brauer, Davenport, Dörge, H. Hopf, Jacobsthal, Kloosterman, Mordell, Salié, I. Schur, Sterneck.

fields $F = K(x, y)$ with finite base field is proved. In these lectures this is done for the simplest non-trivial special case – with a method which can be generalized, as will not be doubted by those who are familiar with the algebraic theory of functions and in particular with the algebraic structure of the theory of abelian functions."

Today, since the language of algebraic geometry is generally in use, we would speak of the theory of "abelian varieties" instead of "abelian functions".

As said above already, the "simplest non-trivial special case" which Hasse has in mind, is the case of function fields of genus 1. We shall report on this in Part 2.

Summary

In 1934, after Hasse had succeeded to prove the Riemann hypothesis in the elliptic case, he published a survey on all what was known at the time about F. K. Schmidt's zeta functions for function fields of arbitrary genus. He included hitherto unpublished results of Artin and F. K. Schmidt. For the first time, this paper exhibits the connection of the problem of estimating the zeros of F. K. Schmidt's zeta functions, with the problem of an asymptotic estimate of the number of diophantine congruences. Although the Riemann hypothesis is not explicitly mentioned in this paper, it is clear from the context that the Riemann hypothesis had served as the motivation for it. The paper was written at about the same time when Hasse delivered his Hamburg lectures on the invitation of Artin. In those lectures Hasse talked about his new proof of the Riemann hypothesis for elliptic function fields. But he stated clearly that the elliptic case should be considered as a very special case only. Those who are acquainted with the theory of abelian functions [i.e., rational functions on abelian varieties]*, Hasse said, will be convinced that his methods can be generalized to yield a proof of the Riemann hypothesis in the general case. He compared the situation with that of Hilbert in 1898, when Hilbert developed the theory of quadratic extensions of number fields but had in mind the theory of general abelian field extensions as class fields.*

References

[A:1921] E. Artin, *Quadratische Körper im Gebiete der höheren Kongruenzen.* Dissertationsauszug. Jahrb. phil. Fak. Leipzig 1921, II. Halbjahr (1921) 157-165 6, 44

[A:1924] —— *Quadratische Körper im Gebiete der höheren Kongruenzen I, II.* Math. Zeitschr. 19 (1924) 153–246 5

[A:1950] —— *The influence of J. H. M. Wedderburn on the development of modern algebra.* Bull. Amer. Math. Soc. 56 (1950) 65–72 28

[A:1965] —— *Collected Papers.* Addison-Wesley (1965) 5, 21

[A:2000] —— *Quadratische Körper über Polynombereichen Galois'scher Felder und ihre Zetafunktionen.* (Manuscript, edited by P. Ullrich.) Abh. Math. Sem. Univ. Hamburg 70 (2000) 3–30 16, 17

[A-S:1927] E. Artin, O. Schreier, *Eine Kennzeichnung der reell abgeschlossenen Körper.* Abh. Math. Sem. Univ. Hamburg 5 (1927) 225–231 5

[A-vdW:1926] E. Artin, B.L. van der Waerden, *Die Erhaltung der Kettensätze der Idealtheorie bei beliebigen endlichen Körpererweiterungen.* Nachr. Ges. Wiss. Göttingen (1926) 23–27 29, 45

[A-Wh:1945] E. Artin, G. Whaples, *Axiomatic characterization of fields ba the product formula for valuations*. Bull. Amer. Math. Soc. 51 (1945) 469–492 31, 41

[Be:1968] H. Behnke, *Paul Sengenhorst 1894–1968.* Math.-Phys. Semesterber., n. F. 15 (1968) 235–239 32

[Bo:1974] E. Bombieri, *Counting points on curves over inite fields (d'apres S. A. Stepanov).* Sem. Bourbaki 1972/73, Expose No.430, Lecture Notes Math. 383 (1974) 234-241 4

[Br-Noe:1927] R. Brauer, E. Noether, *Über minimale Zerfällungskörper irreduzibler Darstellungen.* Sitzungsber. Preuss. Akad. Wiss. Berlin (1927) 221–228 36

[Car:1932] L. Carlitz, *The arithmetic of polynomials in a Galois field.* American Journ. of Math. 54 (1932) 39-50 43

[Coh:1950] I. S. Cohen, *Commutative rings with restricted minimum condition.* Duke Math. J. 17 (1950) 27–42 8

[Da-H:1934] H. Davenport, H.Hasse, *Die Nullstellen der Kongruenzzetafunktionen in gewissen zyklischen Fällen.* J. Reine Angew. Math. 172 (1934) 151–182 26, 43

[De:1857] R. Dedekind, *Abriss einer Theorie der höheren Congruenzen in Bezug auf einen reellen Primzahl-Modulus*. J. Reine Angew. Math. 54 (1857) 1–26 7, 9, 28

[De-Web:1882] R. Dedekind, H. Weber, *Theorie der algebraischen Funktionen einer Veränderlichen*. J. Reine Angew. Math. 92 (1882) 181–290 39, 49, 51

[Deu:1931] M. Deuring, *Verzweigungstheorie bewerteter Körper.* Math. Annalen 105 (1931) 277–307 37

[Deu:1941] —— *Die Typen der Multiplikatorenringe elliptischer Funktionenkörper*. Abh. Math. Sem. Univ. Hamburg 14 (1941) 197–292 26

[Deu:1949] —— *Algebraische Begruendung der komplexen Multiplikation.* Abh. Math. Sem. Univ. Hamburg 16 (1949) 32–47 26

[Deu:1952] —— *Die Struktur der elliptischen Funktionenkörper und die Klassenkörper der komplexen Multiplikation.* Mathematische Annalen 124 (1952) 393–426 26

[Dieu:1947] J. Dieudonné, *Sur les produits tensoriels.* Ann. Sci. Éc. Norm. Supér., III. Sér. 64 (1947) 101–117 8

[Fr:1985] G. Frei, *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

[Fr:2001] —— *On the history of the Artin reciprocity law in abelian extensions of algebraic number fields: How Artin was led to his reciprocity law* In: The legacy of Niels Henrik Abel. The Abel bicentennial, Oslo 2002. Edited by O. A. Laudal. Springer (to appear) 6, 9

[Fr:2001a] —— *Gauss' unpublished Section Eight of the Disquisitiones Arithmeticae: The beginning of the theory of function fields over a finite field.* In: The shaping of arithmetic: Number theory after Carl Friedrich Gauss' Disquisitiones Arithmeticae. Edited by C. Goldstein, N. Schappacher, J. Schwermer. Springer (to appear) 28

[Ge:1981] W.-D. Geyer, *Die Theorie der algebraischen Funktionen einer Veränderlichen nach Dedekind und Weber*. In: W. Scharlau (Ed.) *Richard Dedekind 1831–1981*. (Braunschweig 1981) 109–133 40

[Gre:1935] H. Grell, *Über die Gültigkeit der gewöhnlichen Idealtheorie in endlichen algebraischen Erweiterungen erster und zweiter Art.* Math. Zeitschr. 40 (1935) 503–505 29

[H:1923] H. Hasse, *Über die Darstellbarkeit von Zahlen durch quadratische Formen im Körper der rationalen Zahlen.* J. Reine Angew. Math. 152 (1923) 129–148 34

[H:1924] —— *Darstellbarkeit von Zahlen durch quadratische Formen in einem beliebigen algebraischen Zahlkörper.* J. Reine Angew. Math. 153 (1924) 113–130 36

[H:1926] —— *Bericht über neuere Untersuchungen und Probleme aus der Theorie der algebraischen Zahlkörper. I. Klassenkörpertheorie*. Jber. Deutsch. Math. Verein. 35 (1926) 1–55 38, 44

[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 44

[H:1927b] —— *Existenz gewisse algebraischer Zahlkörper.* Sitzungsber. Preuss. Akad. Wiss. Berlin (1927) 229–234 36

[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 44

[H:1932] —— *Vorlesungen über Klassenkörpertheorie.* Mimeographed Notes, Marburg 1932/32. Reprinted as a book: Physica Verlag, Würzburg 1967 41

[H:1932a] —— *Zu Hilbert's algebraisch-zahlentheoretischen Arbeiten.* In: David Hilbert, Gesammelte Abhandlungen, Band 1 (Zahlentheorie). Springer 1932

[H:1933] —— *Beweis des Analogons der Riemannschen Vermutung für die Artinschen und F.K.Schmidtschen Kongruenzzetafunktionen in gewissen zyklischen Fällen. Vorläufige Mitteilung.* Nachr. Ges. Wiss. Göttingen I. Math.-Phys. Kl. Fachgr. I Math. Nr.42 (1933) 253–262 3, 59

[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 40, 53, 56, 59

[H:1934a] —— *Abstrakte Begründung der komplexen Multiplikation und Riemannsche Vermutung in Funktionenkörpern.* Abh. Math. Sem. Univ. Hamburg 10 (1934) 250–263 4, 59, 60

[H:1934b] —— *Riemannsche Vermutung bei den F.K. Schmidtschen Kongruenzzetafunktionen.* Jber. Deutsch. Math. Verein. 44 (1934) 44 (kursiv) 7

[H:1934c] —— *Theorie der relativ-zyklischen algebraischen Funktionenkörper, insbesondere bei endlichem Konstantenkörper.* J. Reine Angew. Math. 172 (1934) 37–54 5, 43

[H:1935a] —— *Theorie der Differentiale in algebraischen Funktionenkörpern mit vollkommenem Konstantenkörper.* J. Reine Angew. Math. 172 (1934) 55–64 51

[H:1942b] —— *Überblick über die neuere Entwicklung der arithmetischen Theorie der algebraischen Funktionen.* In: Atti del Convegno Matematico, Roma (8.–12. Nov. 1942) Roma 1942, 25-33 53

[H:1975] —— *Mathematische Abhandlungen.* (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) 31, 41, 53

[H-FK:1933] H. Hasse, F.K. Schmidt, *Die Struktur diskret bewerteter Körper.* Journ. f.d. reine u. angewandte Math. 170 (1933) 4–63 41

[Hau:1929] O. Haupt, *Einführung in die Algebra.* 2 vols. (Berlin 1929) 48

[He:1917] E. Hecke, *Über die Zetafunktion beliebiger algebraischer Zahlkörper.* Nachr. d. K. Gesellschaft der Wissenschaften Göttingen, Math.-Phys. Kl. (1917) 77–89

[He:1923] —— *Vorlesungen über die Theorie der algebraischen Zahlen.* Leipzig (1923) 30

[Hen:1904] K. Hensel, *Neue Grundlagen der Arithmetik.* J. Reine Angew. Math. 127 (1904) 51–84

[Hen:1908] K. Hensel, *Theorie der algebraischen Zahlen I.* Teubner Leipzig (1908) 32, 33

[Hen-La:1902] K. Hensel, E.Landsberg, *Theorie der algebraischen Funktionen einer Variablen*, (Leipzig 1902) 51, 52, 53

[Her:1921] G. Herglotz, *Zur letzten Eintragung im Gaußschen Tagebuch.* Ber. Verhandl. Sächs. Akad. Wiss. Math.-Phys. Kl. 73 (1921) 271–276 22, 24

[Her:1979] —— *Gesammelte Schriften.* Vandenhoek Göttingen & Ruprecht, Göttingen (1979) 25

[Hil:1898] D. Hilbert, *Über die Theorie des relativquadratischen Zahlkörpers.* Math. Annalen 51 (1898) 1–127

[Hil:1900] —— *Mathematische Probleme. Vortrag, gehalten auf dem internationalen Mathematiker-Congress zu Paris 1900.* Nachr. Ges. Wiss. Göttingen 1900, 253-297. 49

[Ja:1906] E. Jacobsthal, *Anwendungen einer Formel aus der Theorie der quadratischen Reste.* Dissertation Berlin 1906. 20

[Ja:1910] —— *Über die Darstellung der Primzahlen der Form* $4n + 1$ *als Summe zweier Quadrate.* J. Reine Angew. Math. 137 (1910) 167–309 20

[Kor:1919] H. Kornblum, *Über die Primfunktionen in einer arithmetischen Progression.* Aus dem Nachlaß heausgegeben von E. Landau. Math. Zeitschr. 5 (1919) 100-111 6, 12, 15, 32

[Kru:1932] W. Krull, *Allgemeine Bewertungstheorie.* Journ. f. d. reine u. angewandte Math. 167 (1932) 160–196 40

[Kru:1999] —— *Gesammelte Abhandlungen – Collected Papers.* Ed. P. Ribenboim. de Gruyter Berlin 1999 38

[Kue:1902] H. Kühne, *Eine Wechselbeziehung zwischen Funktionen mehrerer Unbestimmten, die zu Reciprocitätsgesetzen führt.* J. Reine Angew. Math. 124 (1902) 121–133 9, 43

[Kue:1903] ——, *Angenäherte Auflösung von Kongruenzen nach Primmodulsystemen in Zusammenhang mit den Einheiten gewisser Körper.* J. Reine Angew. Math. 126 (1903) 102–115 33, 43

[Ku:1913] J. Kürschák, *Über Limesbildung und allgemeine Körpertheorie.* Journ.f. d. reine u. angewandte Math. 142 (1913) 211–253 32, 33

[Kum:1861] E. Kummer, *Zwei neue Beweise der allgemeinen Reciprocitätsgesetze unter den Resten und Nichtresten der Potenzen, deren Grad eine Primzahl ist.* J. Reine Angew. Math. 100 (1887) 10–50 = Abhandl. Königl. Akademie d. Wissenschaften zu Berlin (1861) 81–122 9

[Lan:1909] E. Landau, *Handbuch der Lehre von der Verteilung der Primzahlen.* Leipzig 1909 16

[Lan:1918] —— *Einführung in die elementare und analytische Theorie der algebraischen Zahlen und der Ideale.* Teubner, Leipzig (1918) 7, 13, 30, 58

[Lem:2000] F. Lemmermeyer, *Reciprocity Laws. From Euler to Eisenstein.* Springer Berlin (2000) 23, 42

[MacR:1971] R. E. MacRae, *On unique factorization in certain rings of algebraic functions.* Journal of Algebra 17 (1971) 243–261 14, 15

[Mad:1972] M. L. Madan, Clifford Queen, *Algebraic function fields of class number one.* Acta Arithmetica 20 (1972) 423–432

[Mad:1975] J.R.C. Leitzel, M.L. Madan, C.S. Queen, *Algebraic function fields with small class number.* Journ. Number Theory 7 (1975) 11–27 49

[Noe:1924] E. Noether. *Abstrakter Aufbau der Idealtheorie im algebraischen Zahlkörper.* Jber. Deutsch. Math. Verein. 33 (1924) 102 29, 30, 37

[Noe:1926] —— *Abstrakter Aufbau der Idealtheorie in algebraischen Zahl- und Funktionenkörpern.* Mathematische Annalen 96 (1926) 26–61 29, 37

[Noe:1927] —— *Der Diskriminantensatz für die Ordnungen eines algebraischen Zahl- und Funktionenkörpers.* J. Reine Angew. Math. 157 (1927) 82–104 30

[Noe:1929] —— *Hyperkomplexe Größen und Darstellungstheorie.* Math. Zeitschr. 30 (1929) 641–692 44

[OMe:1963] O. T. O'Meara, *Introduction to quadratic forms.* Springer 1963 36

[Ore:1934] O. Ore, *Contributions to the theory of finite fields.* Trans. of the American Math. Society 36 (1934) 243-274 43, 48

[Rau:1928] H. Rauter, *Studien zur Theorie des Galoisschen Körpers über dem Körper der rationalen Funktionen einer Unbestimmten t mit Koeffizienten aus einem beliebigen endlichen Körper von* p^m *Elementen.* J. Reine Angew. Math. 159 (1928) 117–132 36, 38

[Rau:1928a] —— *Bemerkung zu der Arbeit: "Studien zur Theorie des Galoisschen Körpers. . . "*. J. Reine Angew. Math. 159 (1928) 228 36

[Rei:1976] C. Reid, *Courant in Göttingen and New York.* Springer Verlag New York (1976) 6, 17

[Rq:1953] P. Roquette, *Arithmetischer Beweis der Riemannschen Vermutung in Kongruenzfunktionenkörpern beliebigen Geschlechts.* J. f. die reine und angewandte Mathematik 191 (1953), 199–252. 7

[Rq:1958] —–, *Riemann-Rochscher Satz in Funktionenkörpern vom Transzendenzgrad 1.* Mathematische Nachrichten 19 (1958) 375–404 53

[Rq:1997] —–, *Zur Geschichte der Zahlentheorie in den dreißiger Jahren*. Jahrbuch 1996 der Braunschweigischen Wissenschaftlichen Gesellschaft (Göttingen 1997) 155–191. Reprinted in: Math. Semesterberichte 45 (1998) 1–38 18

[Rq:2001] —–, *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 30, 43, 44, 55

[Rq:2002] —– *History of Valuation Theory, Part I.* In: Fields Institute communication series vol. 32 (2002) 66pp. 33, 41

[Ros:2002] M. Rosen, *Number Theory in Function fields.* Springer Graduate Texts New York (2002) 16

[Ry:1919] K. Rychlík, *Beitrag zur Körpertheorie* (Czech). Časopis 48 (1919) 145–165 33

[Ry:1923] —– *Zur Bewertungstheorie der algebraischen Körper.* Journ. f. d. reine u. angewandte Math. 153 (1924) 94–107 33

[Sch:1997] N. Schappacher, *Some Milestones of Lemniscatomy.* In: Sertöz, Sinan (ed.), *Algebraic geometry.* Proceedings of the Bilkent summer school, Ankara, Turkey, August 7–19 1995. New York, NY: Marcel Dekker. Lect. Notes Pure Appl. Math. 193, 257–290 (1997) 25

[FK:1925] F.K. Schmidt, *Allgemeine Körper im Gebiet der höheren Kongruenzen.* Dissertation Freiburg 1925 10, 49, 50

[FK:1926] —–, *Zur Zahlentheorie in Körpern der Charakteristik p. (Vorläufige Mitteilung.)* Sitz.-Ber. phys. med. Soz. Erlangen 58/59 (1926/27) 159–172 3, 42, 43, 48, 50, 55

[FK:1931] —–, *Analytische Zahlentheorie in Körpern der Charakteristik p.* Math. Zeitschr. 33 (1931) 1–32 43, 44, 45, 47, 50, 51, 53, 54, 55, 56

[FK:1931a] —–, *Die Theorie der Klassenkörper über einem Körper algebraischer Funktionen in einer Unbestimmten und mit endlichem Koeffizientenbereich.* Sitz.-Ber. phys. med. Soz. 62 (1931) 267–284 11, 50

[FK:1936] —–, *Zur arithmetischen Theorie der algebraischen Funktionen I. Beweis des Riemann-Rochschen Satz für algebraische Funktionenkörper mit beliebigem Konstantenkörper.* Math. Zeitschr. 41 (1936) 415–438 53

[FK:1936a] —–, *Über die Erhaltung der Kettensätze der Idealtheorie bei beliebigen endlichen Körpererweiterungen.* Math. Zeitschr. 46 (1936) 443–450 29

[Se:1925] P. Sengenhorst, *Über Körper der Charakteristik p.* Math. Zeitschr. 24 (1925) 1–39 32, 36

[Se:1927] —– *Bemerkungen zu meiner Arbeit "Über Körper der Charakteristik p" in Band 24 (1925) dieser Zeitschrift S. 1-39.* Math. Zeitschr. 26 (1927) 495 33

[Ste:1910] E. Steinitz, *Algebraische Theorie der Körper*. J. Reine Angew. Math. 137 (1910) 167–309 = *Algebraische Theorie der Körper*, new edition by H. Hasse und R. Baer, Verlag de Gruyter (Berlin 1930) 7, 32

[Ta:1952] J. Tate, *Genus change in inseparable extensions of function fields.* Proc. Amer. Math. Soc. 3 (1952) 400-406 37

[Ul:1999] P. Ullrich, *Die Entdeckung der Analogie zwischen Zahl- und Funktionenkörpern: Der Ursprung der "Dedekind-Ringe". (On the discovery of the analogy between number and function fields: The origin of Dedekind rings).* Jahresber. Dtsch. Math.-Ver. 101 (1999) 116–134 8, 27

[Ul:2000a] —— *Der wissenschaftliche Nachlass Emil Artins.* Mitt. Math. Ges. Hamburg 19 (2000) 113–134 16

[Ul:2000] —— *Emil Artins unveröffentlichte Verallgemeinerung seiner Dissertation.* Mitt. Math. Ges. Hamburg 19 (2000) 173–194 5, 16, 21, 22

[vdW:1975] B. L. van der Waerden, *On the sources of my book "Moderne Algebra".* Historia Mathematica 2 (1975) 31–40 45

[Web:1895] H. Weber, *Lehrbuch der Algebra.* In zwei Bänden. Vieweg & Sohn. Braunschweig (1895) 6

[W:1938] A. Weil, *Zur algebraischen Theorie der algebraischen Funktionen. Aus einem Brief an H. Hasse.* J. Reine Angew. Math. 179 (1938) 129–133 53

[W:1948a] A. Weil, *Sur les courbes algébriques et les variétés qui s'en déduisent*. Act. Sc. et Ind. no.1041 (Paris 1948) 4

[We-Zi:1991] B. Weis, H.G. Zimmer, *Artins Theorie der quadratischen Kongruenzfunktionenkörper und ihre Anwendung auf die Berechnung der Einheiten- und Klassengruppen.* Mitt. Math. Ges. Hamburg 19 (2000) 173-194 7

[Wi:1934b] E. Witt, *Über ein Gegenbeispiel zum Normensatz.* Math. Zeitschr. 39 (1934) 462-467 55

[Wi:1935] E. Witt, *Über die Invarianz des Geschlechts eines algebraischen Funktionenkörpers.* J. Reine Angew. Math. 172 (1935) 75–76 51