Preliminary version
February 24, 2015

II Università degli
Studi di Roma

TOR VERGATA

NZMRI summer school, Nelson, January 2015

René Schoof

Dipartimento di Matematica
2^{a} Università di Roma "Tor Vergata"
I-00133 Roma ITALY
Email: schoof@mat.uniroma2.it

Abstract. These are lectures notes of three talks given at the NZMRI summer school, Nelson, New Zealand, January 2015

1. Zeta functions.

The Riemann zeta function $\zeta(s)$ of a complex variable s satisfying $\operatorname{Re} s > 1$ is defined by

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}.$$

It is also given by the convergent 'Euler product'

$$\zeta(s) = \prod_{p \text{ prime}} \frac{1}{1-p^{-s}}.$$

This follows by writing the Euler factors as geometric series

$$\frac{1}{1-p^{-s}} = 1 + p^{-s} + p^{-2s} + \ldots$$

and using the fact that every natural number is a product of primes in a unique way. It follows that the zeta function has no zeroes s with $\mathrm{Re}\, s > 1$. In his 1859 paper Riemann proved that $\zeta(s)$ admits a meromorphic extension to $\mathbf{C}$ and satisfies a functional equation [17] . More precisely, the function

$$Z(s) \;=\; \pi^{-\frac{s}{2}}\Gamma(\frac{s}{2})\zeta(s)$$

admits an alternative expression that visibly converges for all $s \in \mathbf{C}$ except for $s = 0$ and $s = 1$, where the function has simple poles. Moreover, $Z(s)$ satisfies

$$Z(s) \;=\; Z(1-s).$$

This implies that any zero ρ of $Z(s)$ must satisfy $0 \leq \mathrm{Re}\,\rho \leq 1$. Since the Γ-function has simple poles in integers $n \leq 0$, the function $\zeta(s)$ has 'trivial' zeroes in the negative even integers. All its other zeroes are in the so-called critical strip $\{s \in \mathbf{C} : 0 \leq \mathrm{Re}\, s \leq 1\}$.

Riemann conjectured in his 1859 paper that, in fact, all non-trivial zeroes of the zeta function have real part equal to $\frac{1}{2}$. This is the celebrated Riemann Hypothesis. In 1900 it was included by Hilbert in his list of problems and, about one century later, it appears in the list of millenium problems of the Clay Institute [15, 4]. The zeroes form a discrete set and the 10^{13} zeroes with imaginary part at most $\approx 2,4 \cdot 10^{12}$ have been shown to have real part $\frac{1}{2}$. See [12].

In Appendix XI of the Dirichlet-Dedekind monograph on algebraic number theory [7], Dedekind proposed a generalization of the zeta function to rings of integers of number fields. He observed that the summation over the natural numbers n in the definition of the zeta function, can be viewed as a summation over the non-zero ideals $n\mathbf{Z}$ of the ring $\mathbf{Z}$. Moreover, the number n in the summand $\frac{1}{n^s}$ is precisely the index of $n\mathbf{Z}$ in $\mathbf{Z}$. This leads to the following definition for the zeta function of a ring of integers R of a number field:

$$\zeta_R(s) \;=\; \sum_{0 \neq I \subset R} \frac{1}{N(I)^s}.$$

Here $N(I)$ denotes the cardinality of the finite ring R/I. The rings R are in general not unique factorization domains, but their non-zero ideals are products of maximal ideals in a unique way. Therefore we have an Euler product

$$\zeta_R(s) \;=\; \prod_{\mathfrak{m}\ \max} \frac{1}{1 - N(\mathfrak{m})^{-s}}.$$

Both the sum and the product converge for $s \in \mathbf{C}$ satisfying $\mathrm{Re}\, s > 1$. Hecke [14] showed in 1934 that $\zeta_R(s)$ admits a meromorphic continuation to $\mathbf{C}$, and that it satisfies a functional equation relating $\zeta_R(s)$ to $\zeta_R(1-s)$. As in the case of the Riemann zeta function, this implies that all non-trivial zeroes of $\zeta_R(s)$ are in the critical strip. The Generalized Riemann Hypothesis states that they all have real part equal to $\frac{1}{2}$.

A different kind of rings with the property that their non-zero ideals are of finite index and are products of maximal ideals, are the polynomial rings $R = \mathbf{F}_q[X]$, where $\mathbf{F}_q$ denotes a finite field with q elements. In this case, it is easy to obtain a closed form for $\zeta_R(s)$. We have $N(I) = q^{\dim R/I}$ and hence

$$\zeta_R(s) = \sum_{0\neq I\subset R} \frac{1}{N(I)^s} = \sum_{d=0}^{\infty} \#\{I \subset R \text{ of codimension } d\} q^{-ds}.$$

Since R is a principal ideal domain and the unit group R^* is equal to $\mathbf{F}_q^*$, every ideal I is generated by a unique monic polynomial f. There are precisely q^d monic polynomials of degree d in R. It follows that

$$\zeta_R(s) = \sum_{d=0}^{\infty} q^d q^{-ds} = \frac{1}{1-q^{1-s}}.$$

This formula provides us with a meromorphic continuation of $\zeta_R(s)$ to $\mathbf{C}$. There are infinitely many poles. They are of the form $s = 1 + \frac{2\pi i k}{\log q}$ for $k \in \mathbf{Z}$. The zeta function $\zeta_R(s)$ satisfies the analogue of the Riemann Hypothesis, since it has no zeroes at all!

In his thesis Artin studied in 1921 quadratic extensions of the ring $\mathbf{F}_q[X]$. See [2]. When the characteristic of $\mathbf{F}_q$ is not 2, these are of form $R = \mathbf{F}_q[X][\sqrt{f(X)}]$ where $f(X)$ is a squarefree polynomial in $\mathbf{F}_q[X]$. Artin viewed the rings R as analogues of the rings of integers of quadratic number fields and he defined and studied the analogues of Dedekind's zeta function for these rings. The ideals of these rings are of finite index and are products of maximal ideals. Artin computed the zeta function explicitly for a handful of rings R. In each case the analog of the Riemann Hypothesis turned out to be true.

In order to describe the successive developments, it is useful to adopt a more geometric language. Writing $Y = \sqrt{f(X)}$, we see that R is equal to the ring of regular functions $\mathbf{F}_q[X,Y]/(Y^2 - f(X))$ on the algebraic curve given by $Y^2 = f(X)$. For any point $P = (x,y)$ of this curve with $x, y \in \overline{\mathbf{F}}_q$, the kernel $\mathfrak{m}$ of the evaluation map $R \longrightarrow \overline{\mathbf{F}}_q$ given by $g \mapsto g(P)$ is a maximal ideal of R. This is in fact a bijective correspondence between maximal ideals $\mathfrak{m}$ of R and $\overline{\mathbf{F}}_q$-points P of the curve $Y^2 = f(X)$. Here the points are to be taken up to conjugacy by the Galois group of $\overline{\mathbf{F}}_q$ over $\mathbf{F}_q$. In this correspondence, $R/\mathfrak{m}$ is the subfield of $\overline{\mathbf{F}}_q$ that is generated by the coordinates of P. Therefore, we can write

$$\zeta_R(s) = \prod_P \frac{1}{1 - N(P)^{-s}},$$

where P runs over the $\overline{\mathbf{F}}_q$-points P of the curve given by $Y^2 - f(X)$ and $N(P)$ is the number of elements of the field of definition of P.

It is well known that there is a projective, absolutely irreducible smooth curve C over $\mathbf{F}_q$ whose function field is the field of fractions of $R = \mathbf{F}_q[X,Y]/(Y^2 - f(X))$. The curve C is unique up to isomorphism and its zeta function is defined by

$$\zeta_C(s) = \prod_{P\in C(\overline{\mathbf{F}}_q)} \frac{1}{1 - N(P)^{-s}}.$$

Here $C(\overline{\mathbf{F}}_q)$ denotes the set of $\overline{\mathbf{F}}_q$-points of C, counted up to Galois conjugacy. The product converges for $s \in \mathbf{C}$ for which $\operatorname{Re} s > 1$. The function $\zeta_C(s)$ is equal to $\zeta_R(s)$ up to a finite number of Euler factors. In particular, the analog of the Riemann Hypothesis holds for $\zeta_R(s)$ if and only if it holds for $\zeta_C(s)$. It is more convenient to describe the properties of the zeta functions $\zeta_C(s)$.

In 1927 the German mathematician F.K. Schmidt proved the Riemann-Roch Theorem for curves over finite fields [18]. This enabled him to compute closed forms for the zeta functions $\zeta_C(s)$. It is convenient to introduce the power series

$$Z_C(T) \;=\; \prod_{P\in C(\overline{\mathbf{F}}_q)} \frac{1}{1-T^{\deg P}},$$

where $\deg P$ denotes the degree of the field of definition of P over $\mathbf{F}_q$. It is related to $N(P)$ by the formula $N(P) = q^{\deg P}$ and we have $\zeta_C(s) = Z_C(q^{-s})$. Schmidt showed that

$$Z_C(T) \;=\; \frac{P(T)}{(1-T)(1-qT)},$$

where $P(T)$ is a polynomial having the following palindrome shape

$$P(T) = 1 + b_1T + \ldots + b_{g-1}T^{g-1} + b_gT^g + b_{g-1}qT^{g+1} + \ldots + b_1q^{g-1}T^{2g-1} + q^gT^{2g},$$

for certain coefficients $b_i \in \mathbf{Z}$. The integer g is the genus of C. The palindrome property means

$$q^gT^{2g}P(\frac{1}{qT}) \;=\; P(T).$$

This formula provides the meromorphic continuation of $\zeta_C(s)$ to all of $\mathbf{C}$. The palindrome property translates into a functional equation relating $\zeta_C(s)$ to $\zeta_C(1-s)$.

The analogue of the Riemann Hypothesis is trivially true when the genus $g = 0$. In 1934 Hasse [13] proved it for curves with $g = 1$. For curves of genus $g > 1$ it was proved by Weil in the period 1940–1948. See Weil's publications [22, 23, 24] and [1] for historical context.

We have $\zeta_C(s) = 0$ if and only if $P(q^{-s}) = 0$. If we factor the polynomial $P(T)$ in $\mathbf{C}[T]$ and write

$$P(T) \;=\; \prod_{\pi}(1-\pi T),$$

where π runs over the reciprocal roots of $P(T)$, then $\zeta_C(s) = 0$ if and only if $1 - \pi q^{-s} = 0$ for one of the reciprocal zeroes π. Therefore, the analogue of the Riemann Hypothesis says precisely that the absolute values of the complex numbers π are all equal to $\sqrt{q}$.

2. Stepanov's method.

In this section we estimate the number of rational points of an elliptic curve over a finite field. This is the key ingredient for the proof of the analogue of the Riemann Hypothesis. Our method is due to Stepanov [19]. His proof easily generalizes to curves of higher genus. See Bombieri's Bourbaki lecture [3] or Hindry's note [16].

An elliptic curve E over a finite field $\mathbf{F}_q$ is a smooth cubic curve given by a Weierstrass equation

$$Y^2 + a_1XY + a_3Y \;=\; X^3 + a_2X^2 + a_4X + a_6,$$

with $a_i \in \mathbf{F}_q$. We write $E(\mathbf{F}_q^n)$ for the set of points with coordinates in the subfield $\mathbf{F}_{q^n}$ of $\overline{\mathbf{F}}_q$. By $E(\overline{\mathbf{F}}_q)$ we denote the set of points with coordinates in $\overline{\mathbf{F}}_q$.

We consider the $\mathbf{F}_q$-algebra

$$R = \mathbf{F}_q[X,Y]/(Y^2 + a_1XY + a_3Y - X^3 - a_2X^2 - a_4X - a_6).$$

The elements of R can be viewed as functions on E. Every element $f \in R$ has the form $g(X) + Yh(X)$ for unique polynomials $g, h \in \mathbf{F}_q[X]$. For every non-zero $f \in R$, let $\deg f$ denote the dimension of the $\mathbf{F}_q$-vector space R/fR. We have $\deg X = 2$ and $\deg Y = 3$. In general, for $f = g(X) + Yh(X)$ with $g, h \in \mathbf{F}_q[X]$ polynomials of degrees d, e respectively, one has $\deg f = \max(2d, 3+2e)$. In particular, R contains no function f with $\deg f = 1$. Any non-zero $f \in R$ has at most $\deg f$ zeroes on $E(\overline{F}_q) - \{\infty\}$. Indeed, if $f = g(X) + Yh(X)$ as above, then the equation obtained by substituting $Y = -g(X)/h(X)$ in the Weierstrass equation has degree $\deg f$ in X.

For $a \geq 0$, let L_a denote the $\mathbf{F}_q$-vector space

$$L_a = \{f \in R : \deg f \leq a\}.$$

For $a = 0$ or 1 the space L_a consists only of constant functions and has dimension 1. This follows from the fact that R does not contain any functions f of degree 1. The following Lemma describes what happens for $a \geq 2$. We put $e_1 = 1$ and

$$e_{2i} = X^i \quad \text{and} \quad e_{2i+1} = X^{i-1}Y \quad \text{for } i \geq 1.$$

Then e_i has degree i for $i \geq 1$. Every function $f \in R$ is of the form $\sum_{k=1}^{d} \lambda_k e_k$ with unique coefficients $\lambda_k \in \mathbf{F}_q$. The degree of f is equal to the largest index k for which $\lambda_k \neq 0$. We call f *monic* if $\lambda_k = 1$.

Lemma 2.1. *For $a \geq 1$, the monomials e_i with $i \leq a$ are an $\mathbf{F}_q$-basis for L_a. In particular, L_a has $\mathbf{F}_q$-dimension a.*

Proof. The monomials e_i certainly generate L_a. On the other hand, the orders of their poles at ∞ are all different. Therefore, they are linearly independent and hence form a basis of L_a. This proves the lemma.

For $a \geq 1$, the set $L_a^q = \{f^q : f \in L_a\}$ is an $\mathbf{F}_q$-vector space of dimension $a = \dim L_a$. Indeed, the map $f \mapsto f^q$ is an $\mathbf{F}_q$-linear bijection $L_a \leftrightarrow L_a^q$.

Lemma 2.2. *Let a, b be positive integers and let $L_a^q L_b$ denote the $\mathbf{F}_q$-vector space generated by the functions $f^q g$ where $f \in L_a$ and $g \in L_b$. Then we have*
(a) $\dim L_a^q L_b \leq aq + b$;
(b) $\dim L_a^q L_b \leq ab$;
(c) if $b < q$, the elements $e_i^q e_j$ for $1 \leq i \leq a$ and $1 \leq j \leq b$, form an $\mathbf{F}_q$-basis of $L_a^q L_b$ and we have equality in (b).

Proof. Part (a) follows from the fact that $L_a^q L_b \subset L_{aq+b}$. The inequality of part (b) follows from the fact that the functions $e_i^q e_j$ with $1 \leq i \leq a$ and $1 \leq j \leq b$ generate $L_a^q L_b$. For (c) we observe that

$$\deg e_i^q e_j = q \deg e_i + \deg e_j$$

Thus, if $b < q$, we have $\deg e_j < q$ for all j. It follows that the degrees $\deg e_i^q e_j$ are all distinct. So any $\mathbf{F}_q$-linear combination $\sum_{i,j} \lambda_{ij} e_i^q e_j$ that is zero, necessarily has $\lambda_{ij} = 0$ for every i, j. This proves that the functions $e_i^q e_j$ are independent. It follows that they are a basis for $L_a^q L_b$. This proves the lemma.

From now on we assume that $a, b \geq 1$ with $b < q$. Lemma 2.2 (c) implies that the $\mathbf{F}_q$-linear map

$$\vartheta : L_a^q L_b \longrightarrow L_a L_b^q$$

given by

$$e_i^q e_j \mapsto e_i e_j^q, \qquad \text{for } 1 \leq i \leq a \text{ and } 1 \leq j \leq b,$$

is well defined.

The following proposition is the key ingredient in the proof of Theorem 2.4.

Proposition 2.3. *Let $a, b \geq 1$ with $b < q$. If the map ϑ is not injective, then*

$$\#E(\mathbf{F}_{q^2}) \leq aq + b + 1.$$

Proof. Every function $F \in \ker \vartheta$ vanishes on $E(\mathbf{F}_{q^2}) - \{\infty\}$. Indeed, let $F = \sum \lambda_{ij} e_i^q e_j$ for certain $\lambda_{ij} \in \mathbf{F}_q$, and let $P \in E(\mathbf{F}_{q^2}) - \{\infty\}$. Then

$$F(P)^q = \sum \lambda_{ij} e_i^{q^2}(P) e_j^q(P) = \sum \lambda_{ij} e_i(P) e_j^q(P) = (\sum \lambda_{ij} e_i e_j^q)(P) = \vartheta(F)(P) = 0,$$

which is zero when $F \in \ker \vartheta$. The second equality follows from the fact that $P \in E(\mathbf{F}_{q^2})$ so that $f^{q^2}(P) = f(P)$ for every function $f \in R$.

Since ϑ is not injective, there exists a non-zero F in $\ker \vartheta$. Therefore, we obtain the following estimate.

$$\#E(\mathbf{F}_{q^2}) - 1 \leq \#\{\text{zeroes of } F\} \leq \deg(F) \leq aq + b.$$

The rightmost inequality follows from Lemma 2.2 (a). This proves the proposition.

Theorem 2.4. *Let E be an elliptic curve defined over $\mathbf{F}_q$ and suppose that $q \geq 5$. Then we have*

$$\#E(\mathbf{F}_{q^2}) \leq q^2 + 3q.$$

Proof. The map ϑ defined above cannot be injective if $a, b \geq 1$ have the property that

$$\dim L_a^q L_b \; > \; \dim L_a L_b^q.$$

Since $b < q$, Lemma 2.2 (b) implies that $L_a^q L_b$ has dimension ab. Lemma 2.2 (b) cannot be applied to $L_a L_b^q$. In some sense this is the point of the proof. However, Lemma 2.2 (a) implies that $L_a L_b^q$ has dimension $\leq a + bq$. Therefore the map ϑ is *not* injective when

$$ab \; > \; a + bq.$$

In order to deduce a sharp estimate from Proposition 2.3, we choose a as small as possible. Since the inequality $ab > a + bq$ must be satisfied, the minimal choice is $a = q + 2$. Once a is chosen, we can take $b = q - 1$, at least for $q \geq 5$. With these choices the quantity $aq + b + 1$ in Proposition 2.3 becomes $(q + 2)q + q - 1 + 1 = q^2 + 3q$, as required.

3. The Riemann Hypothesis.

Let E be an elliptic curve over $\mathbf{F}_q$.

Proposition 3.1. *The zeta function of the elliptic curve E is given by*

$$Z_E(T) \; = \; \frac{1 - \tau T + qT^2}{(1 - T)(1 - qT)},$$

where τ is an integer given by $\#E(\mathbf{F}_q) = q + 1 - \tau$.

Before proving Proposition 3.1, we prove the analogue of the Riemann Hypothesis. In other words, we prove that the complex zeroes of the numerator of $Z_E(T)$ have absolute value $1/\sqrt{q}$. The key ingredient is the upper bound for $\#E(\mathbf{F}_{q^2})$ of Theorem 2.4. We first use the method of the proof of Theorem 2.4 to obtain a lower bound for $\#E(\mathbf{F}_{q^2})$.

Proposition 3.2. *Let E be an elliptic curve over $\mathbf{F}_q$ and suppose that $q \geq 5$. Then we have*

$$\#E(\mathbf{F}_{q^2}) \; > \; q^2 - 3q.$$

Proof. Let Ω denote the set of points (x, y) of $E(\overline{\mathbf{F}}_q) - \{\infty\}$ for which $x \in \mathbf{F}_{q^2}$. For every $x \in \mathbf{F}_{q^2}$ there are at most two points $(x, y) \in \Omega$. If (x, y) is one such point, then $(x, \overline{y})$ where $\overline{y} = -y - a_1 x - a_3$, is the other. We have

$$\#\Omega \; = \; 2q^2 - r.$$

where r is the number of values of x for which $y = \overline{y}$. We have $r \leq 3$.

The automorphism σ of $\overline{\mathbf{F}}_q$ given by $\sigma(t) = t^{q^2}$ acts on Ω. It maps a point $(x, y) \in \Omega$ to $(\sigma(x), \sigma(y)) = (x^{q^2}, y^{q^2}) = (x, y^{q^2})$. It follows that either $\sigma(y) = y$, or $\sigma(y) = \overline{y}$. Therefore, we have

$$\Omega = \Omega^+ \cup \Omega^-,$$

where $\Omega^+ = \{(x,y) \in \Omega : \sigma(y) = y\}$ and $\Omega^- = \{(x,y) \in \Omega : \sigma(y) = \overline{y}\}$. The intersection $\Omega^+ \cap \Omega^-$ consists of the r points (x,y) for which $y = \overline{y}$. Clearly, Ω^+ is the set $E(\mathbf{F}_{q^2}) - \{\infty\}$. Theorem 2.4 provides an estimate for its size.

We estimate the size of the set Ω^-. Let a, b be as in the proof of Theorem 2.4. Note that the spaces L_a and L_b are preserved by the automorphism $f \mapsto \overline{f}$ of R given by $\overline{f}(X,Y) = f(X, -Y - a_1X - a_3)$. Consider the $\mathbf{F}_q$-linear map

$$\vartheta' : L_a^q L_b \longrightarrow L_a L_b^q$$

defined by

$$e_i^q e_j \mapsto \overline{e_i} e_j^q.$$

Every function $F \in \ker \vartheta'$ vanishes on the set Ω^-. Indeed, let $F = \sum \lambda_{ij} e_i^q e_j$ for certain $\lambda_{ij} \in \mathbf{F}_q$ and let $P \in \Omega^-$. Then

$$F(P)^q = \sum \lambda_{ij} e_i^{q^2}(P) e_j^q(P) = \sum \lambda_{ij} \overline{e}_i(P) e_j^q(P) = (\sum \lambda_{ij} \overline{e}_i f_j^q)(P) = \vartheta'(F)(P) = 0,$$

and hence $F(P) = 0$. Therefore, we can draw the same conclusion as in the previous section. We have

$$\#\Omega^- \leq q^2 + 3q.$$

and hence

$$\begin{aligned} \#E(\mathbf{F}_{q^2}) - 1 &= \#\Omega^+, \\ &= \#\Omega - \#\Omega^- + \#(\Omega^+ \cap \Omega^-), \\ &\geq (2q^2 - r) - (q^2 + 3q) + r, \\ &\geq q^2 - 3q. \end{aligned}$$

as required.

Theorem 3.3. *Let E be an elliptic curve over $\mathbf{F}_q$. The inverse zeroes π and π' of the numerator $1 - \tau T + qT^2$ of the zeta function of E have absolute value $\sqrt{q}$. In particular, we have $\pi' = \overline{\pi}$.*

Proof. By Proposition 3.1 we have

$$\frac{1 - \tau T + qT^2}{(1-T)(1-qT)} = Z_E(T) = \prod_P \frac{1}{1 - T^{\deg P}},$$

where P runs over the points in $E(\overline{\mathbf{F}}_q)$ up to Galois conjugacy. This gives

$$\frac{(1 - \pi T)(1 - \pi' T)}{(1-T)(1-qT)} = \prod_{d \geq 1} (1 - T^d)^{-a_d}.$$

Here a_d denotes the number of points on E of degree d up to Galois conjugacy. Taking the logarithmic derivative of this identity, expanding the geometric series and comparing coefficients, shows that for $e \geq 1$ we have $q^e + 1 - \pi^e - \pi'^e = \sum_{d|e} d a_d$. Therefore

$$\#E(\mathbf{F}_{q^e}) = \sum_{d|e} d a_d = q^e + 1 - \pi^e - {\pi'}^e, \qquad \text{for every } e \geq 1.$$

Theorem 2.4 and Proposition 3.2 imply that

$$|\pi^e + {\pi'}^e| \leq 3q^{e/2} + 1, \qquad \text{for all even exponents } e.$$

It follows that $|(\frac{\pi}{\sqrt{q}})^e + (\frac{\pi'}{\sqrt{q}})^e|$ remains bounded as $e \to \infty$. Since $\pi\pi' = q$, this implies that $|\pi| = |\pi'| = \sqrt{q}$, as required.

The inequalities of Theorem 2.4 and Proposition 3.2 have only been proved for $q \geq 5$. However, when $q < 5$, we have $q^k > 5$ for $k \geq 3$. This implies that we still have the inequality for even degrees $e \geq 6$. So, the argument involving $e \to \infty$ is not affected and the conclusion is the same for $q < 5$. This proves the theorem.

Corollary 3.4. *Let E be an elliptic curve over $\mathbf{F}_q$. Then*

$$|q^e + 1 - \#E(\mathbf{F}_{q^e})| \leq 2q^{e/2}, \qquad \textit{for every } e \geq 1.$$

We briefly explain how Proposition 3.1 can be proved. Since the elliptic curve has a unique point at infinity and since this point is defined over $\mathbf{F}_q$, it suffices to show that

$$Z_R(T) = \frac{1 - \tau T + qT}{1 - qT},$$

where R is the $\mathbf{F}_q$-algebra given by

$$R = \mathbf{F}_q[X,Y]/(Y^2 + a_1XY + a_3Y - X^3 - a_2X^2 - a_4X - a_6).$$

Since $Z_R(T)$ is equal to $\sum_{d\geq 1} c_d T^d$ where c_d denotes the number of ideals of R of codimension d, this boils down to counting ideals of R of fixed codimension. We already did this in section 1 for the ring $R = \mathbf{F}_q[X]$. Here we proceed in a similar way. To show that

$$Z_R(T) = \sum_{d\geq 0} c_d T^d = (1 - \tau T + qT^2)(1 + qT + (qT)^2 + \ldots),$$

we must show that $c_1 = q - \tau$ and $c_d = q^{d-1}(1 - \tau + q)$ for $d \geq 2$. Since the ideals of codimension 1 are maximal with residue field $\mathbf{F}_q$, they correspond bijectively to the set of points $E(\mathbf{F}_q) - \{\infty\}$. Therefore $c_1 = \#E(\mathbf{F}_q) - 1 = q - \tau$ as required.

For $d \geq 2$, Lemma 2.1 implies that there are q^{d-1} monic functions in R of degree d. Since the unit group of R is equal to $\mathbf{F}_q^*$, this implies that there are q^{d-1} *principal* ideals of R of codimension d.

Let I be a *non-principal* ideal of R of codimension d. Then the monomials e_i with $1 \leq i \leq d+1$ of section 2 are linearly dependent in the $\mathbf{F}_q$-vector space R/I. Therefore I contains a monic function f of degree $d+1$. Since R contains no functions of degree 1, the function f is unique. In order to count the functions f we use the fact that I is an invertible ideal of R and write $(f) = I\mathfrak{m}$ for a unique codimension 1 ideal $\mathfrak{m}$. Let $P \in E(\mathbf{F}_q)$ be the rational point that corresponds to $\mathfrak{m}$. The number of ideals of codimension d is then seen to be equal to the number of points in $E(\mathbf{F}_q) - \{\infty\}$ times the number of monic degree $d+1$ functions that vanish in P. This gives $q^{d-1}(q-\tau)$ ideals. Adding the principal ideals we find that $c_d = q^{d-1}(1 - \tau + q)$ as required.

4. Counting points on elliptic curves over finite fields.

Let E be an elliptic curve over a finite field $\mathbf{F}_q$. For convenience sake we assume in this section that the characteristic of $\mathbf{F}_q$ is not 2 or 3. Then E is given by a Weierstrass equation

$$Y^2 = X^3 + AX + B$$

for some A, $B \in \mathbf{F}_q$ satisfying $4A^3 + 27B^2 \neq 0$ in $\mathbf{F}_q$. We let $E(\mathbf{F}_q)$ denote the set of points on E with coordinates in $\mathbf{F}_q$. In this section we describe two methods to determine the cardinality of $E(\mathbf{F}_q)$.

First we describe the straightforward naive method. Given $x \in \mathbf{F}_q$, it is easily seen that the number of points (x, y) in $E(\mathbf{F}_q)$ whose X-coordinate is equal to x, is equal to $1 + \chi(x^3 + Ax + B)$. Here $\chi : \mathbf{F}_q \longrightarrow \{-1, 0, +1\}$ is the function given by

$$\chi(t) \;=\; \begin{cases} -1, & \text{if } t \text{ is not a square in } \mathbf{F}_q; \\ 0, & \text{if } t = 0; \\ 1, & \text{if } t \text{ is a non-zero square in } \mathbf{F}_q. \end{cases}$$

Including the point at infinity, the set $E(\mathbf{F}_q)$ has therefore cardinality

$$\#E(\mathbf{F}_q) \;=\; 1 + \sum_{x \in \mathbf{F}_q} (1 + \chi(x^3 + Ax + B)) = 1 + q + \sum_{x \in \mathbf{F}_q} \chi(x^3 + Ax + B).$$

This implies that evaluating the sum $\sum_{x \in \mathbf{F}_p} \chi(x^3 + Ax + B)$ is the same problem as computing $\#E(\mathbf{F}_q)$. For very small values of q, a straightforward evaluation of this sum is an efficient way to compute $\#E(\mathbf{F}_q)$. The running time of this algorithm is proportional to q. It is an *exponential* algorithm.

Next we describe an deterministic polynomial time algorithm that is based on calculations with torsion points. Since the running time is $O(\log^8 q)$, the algorithm is asymptotically fast. However, in the form we present it, it is not very efficient in practice. Successive improvements by Atkin and Elkies [9] have made the algorithm much faster at the cost of not being deterministic anymore. These enabled Sutherland in 2010 to compute the number of points on the curve

$$y^2 = x^3 + 2718281828X + 3141592653,$$

modulo the 5011 digit prime $q = 16219299585 \cdot 2^{16612} - 1$. It is equal to $q + 1 - t$, where t is the integer listed in the appendix of these notes [20].

When the characteristic p of $\mathbf{F}_q$ is very small, there are better algorithms [21]. In contrast to the present algorithm, which may be said to be 'l-adic', those algorithms are p-adic in nature.

In order to explain the algorithm, it is useful to first explain how to compute $\#E(\mathbf{F}_q)$ modulo 2. The cardinality of the group $E(\mathbf{F}_q)$ is even if and only if it contains a point of order 2. Since the points of order 2 have the form $(x, 0)$, this means precisely that the polynomial $X^3 + AX + B$ has a zero in $\mathbf{F}_q$. This, in turn, is equivalent to

$$\gcd(X^q - X, X^3 + AX + B) \neq 1 \qquad \text{in the ring } \mathbf{F}_q[X].$$

This can be tested efficiently; the bulk of the computation is the calculation of X^q in the ring $\mathbf{F}_q[X]/(X^3+AX+B)$, which can be done by repeated squarings and multiplications, using the binary presentation of the exponent q. The amount of work involved is $O(\log^3 q)$.

We generalize this calculation to other primes l. We compute $\#E(\mathbf{F}_q)$ modulo the first few small odd primes $l = 3, 5, 7, \ldots$ Since, by the analog of the Riemann Hypothesis, we have

$$q + 1 - 2\sqrt{q} < \#E(\mathbf{F}_p) < q + 1 + 2\sqrt{q},$$

it suffices that

$$\prod_l l > 4\sqrt{q}$$

in order to determine the cardinality uniquely by means of the Chinese Remainder Theorem. A weak form of the prime number theorem shows that this can be achieved with at most $O(\log q)$ primes l, each of size at most $O(\log q)$. Since q is large, the primes l are very small with respect to q. We avoid $l = \operatorname{char} \mathbf{F}_q$.

As in the case where $l = 2$, we use the subgroup $E[l]$ of l-torsion points of $E(\overline{\mathbf{F}}_q)$:

$$E[l] = \{P \in E(\overline{\mathbf{F}}_q) : [l]P = 0\}.$$

The group $E[l]$ is isomorphic to $\mathbf{Z}/l\mathbf{Z} \times \mathbf{Z}/l\mathbf{Z}$. It is the kernel of the multiplication by l morphism $[l] : E \longrightarrow E$. There exist polynomials, the socalled *division polynomials*

$$\Psi_l(X) \in \mathbf{F}_q[X],$$

that vanish precisely in the l-torsion points. For example

$$\begin{aligned}
\Psi_3(X) &= 3X^4 + 6AX^2 + 12BX - A^2,\\
\Psi_5(X) &= 5X^{12} + 62AX^{10} + 380BX^9 - 105A^2X^8 + 240BAX^7 + (-300A^3 - 240B^2)X^6\\
&\quad - 696BA^2X^5 + (-125A^4 - 1920B^2A)X^4 + (-80BA^3 - 1600B^3)X^3\\
&\quad + (-50A^5 - 240B^2A^2)X^2 + (-100BA^4 - 640B^3A)X + (A^6 - 32B^2A^3 - 256B^4).
\end{aligned}$$

The division polynomials can be calculated recursively by means of the addition formulas [8]. The degree of $\Psi_l(X)$ is $(l^2-1)/2$. The amount of work involved in calculating them is dominated by the rest of the computation, so we don't bother estimating it.

The Frobenius endomorphism $\varphi : E \longrightarrow E$ satisfies the quadratic relation

$$\varphi^2 - [\tau]\varphi + [q] = 0,$$

where τ is the integer for which $\#E(\mathbf{F}_q) = q + 1 - \tau$. In the algorithm we check which of the relations

$$\varphi^2 - [t]\varphi + [q] = 0, \qquad t = 0, 1, 2, \ldots, l-1,$$

holds *on the group $E[l]$ of l-torsion points.* It is easily seen that the relation can only hold for $t \equiv \tau \pmod{l}$. In this way we obtain the value of τ modulo l.

The key point is that the relations can be expressed by means of polynomials and that they can be checked efficiently: we have that

$$\varphi^2(x,y) + [q](x,y) = [t]\varphi(x,y) \qquad \text{for all } (x,y) \in E[l]$$

if and only if

$$(X^{q^2}, Y^{q^2}) + [q'](X,Y) \equiv [t](X^q, Y^q)$$

modulo the polynomials $\Psi_l(X)$ and $Y^2 - X^3 - AX - B$. Here q' denotes the integer congruent to q (mod l) that satisfies $0 \le q' < l$. Note that the "+" that occurs in the formula is the addition on the elliptic curve, and that the multiplications are repeated additions.

The bulk of the computation is, first, the computation of the powers X^q, X^{q^2}, etc. in the ring

$$\mathbf{F}_q[X,Y]/(\Psi_l(X), Y^2 - X^3 - AX - B),$$

and then, l times, the addition of the point (X^q, Y^q), which boils down to a few additions and multiplications in the same ring. Since the elements of the ring have size $l^2 \log q$, the amount of work involved is $O(\log q (l^2 \log q)^2)$ and $O(l(l^2 \log q)^2)$ respectively. Here we assume that the usual multiplication algorithms are being used, so that multiplying two elements of size n takes time proportional to n^2.

Keeping in mind that $l = O(\log q)$ and that we do this calculation for each l, we conclude that the amount of work involved for the entire calculation is

$$O(\log^8 q).$$

Therefore, this is a deterministic polynomial time algorithm.

We mention three applications of this algorithm.

Application 4.1. *Cryptography.*

Let p be a prime and let g be a primitive root modulo p. Then every $x \in \mathbf{F}_p^*$ can be written as $x = g^l$ for some $l \in \mathbf{Z}$ that is unique modulo $p-1$. The number l is the *discrete logarithm* of x with respect to the primitive root g. When p is large, it is difficult to compute l given g and x. The difficulty of this problem has been used to design cryptosystems such as the Diffie-Hellman key exchange. The best methods to compute discrete logarithms are based on index calculus and use variations of the number field sieve.

The elliptic discrete logarithm is analogous [10]. Let E be an elliptic curve over a finite field $\mathbf{F}_q$ and suppose that the group of points $E(\mathbf{F}_q)$ is generated by a point P. Then every $Q \in E(\mathbf{F}_p)$ can be written as $Q = [l]P$ for some $l \in \mathbf{Z}$ that is unique modulo the order of the group $E(\mathbf{F}_q)$. The number l is the *elliptic discrete logarithm* of Q with respect to the point P. There are no good methods to compute elliptic discrete logarithms. In particular, there is no analogue of the methods that are based on index calculus. Therefore, elliptic curve cryptosystems are even more secure than cryptosystems based on the usual discrete logarithm or on the difficulty of factoring large numbers [11]. As a consequence the key size can be smaller and the encryption and decryption algorithms are faster.

In order to create secure elliptic curve cryptosystems, it is necessary to count the number of elements of the groups of points $E(\mathbf{F}_q)$ of elliptic curves E over finite fields $\mathbf{F}_q$.

Application 4.2. *An algorithm to compute square roots modulo primes.*

Consider the elliptic curve $Y^2 = X^3 - X$. It has good reduction modulo any prime $p > 2$. For primes $p \equiv 1 \pmod 4$ the ring of endomorphisms of E over $\mathbf{F}_p$ is isomorphic to the ring of Gaussian integers $\mathbf{Z}[i]$. In terms of this isomorphism the Frobenius element of E over $\mathbf{F}_p$ is equal to an element $a + bi \in \mathbf{Z}[i]$ where $a, b \in \mathbf{Z}$ satisfy $a^2 + b^2 = p$. The trace τ of Frobenius is equal to $2a$. Since $\#E(\mathbf{F}_p) = p + 1 - \tau$, we can compute a by counting the points on E over $\mathbf{F}_p$. Since $(a/b)^2 \equiv -1 \pmod p$, this computation also yields a square root of -1 modulo p.

In a similar way, one obtains for each $d \in \mathbf{Z}$ a deterministic polynomial time algorithm to compute the square root of d modulo the primes for which d is a square modulo p. The dependence of the running time on d is exponential however.

Application 4.3. *An algorithm to compute coefficients a_p of modular forms of weight 2.*

Let $\tau \in \mathbf{C}$ be a variable satisfying $\operatorname{Im} \tau > 0$ and put $q = e^{2\pi i \tau}$. The Fourier series

$$\sum_{n=1}^{\infty} a_n q^n \;=\; q \prod_{n=1}^{\infty} (1 - q^n)^2 (1 - q^{11n})^2,$$

is the unique normalized weight 2 cusp form for the group $\Gamma_0(11)$. The Fourier coefficients a_n are a multiplicative function of n satisfying $a_{p^{m+1}} = a_p a_{p^m} - p a_{p^{m-1}}$ for prime p and $m \geq 1$. Therefore, the series is determined by the coefficients a_p for prime p.

The elliptic curve E given by

$$Y^2 - Y = X^3 - X^2$$

has good reduction modulo primes $p \neq 11$. The cardinality of the set $E(\mathbf{F}_p)$ is equal to $p + 1 - a_p$. Therefore, the coefficients a_p of the modular form can be computed in deterministic polynomial time by counting points on E over $\mathbf{F}_p$. This algorithm generalizes easily to normalized weight 2 cusp form for the groups $\Gamma_0(N)$ for any $N \geq 1$, as long as the Fourier coefficients are in $\mathbf{Z}$.

Since Pila generalized the elliptic curve algorithm to abelian varieties, it is also possible to determine the coefficients a_p of arbitrary modular forms of weight 2 in deterministic polynomial time. Recently Couveignes and Edixhoven have proposed an algorithm to determine the Fourier coefficients of modular forms f of weight $k > 2$. See [5]. Their algorithm computes the coefficients modulo small primes l by exploiting the 2-dimensional Galois representations associated to f. An important example is the unique cusp form of weight 12 for the group $\mathrm{SL}_2(\mathbf{Z})$. It is given by

$$\Delta(q) = \sum_{n=1}^{\infty} \tau(n) q^n \;=\; q \prod_{n=1}^{\infty} (1 - q^n)^{24},$$

where $\tau(n)$ is Ramanujan's τ-function. At present it is possible to compute $\tau(p)$ modulo primes $l \leq 31$ for p a 1000 digit prime. See [6] and the references there.

It would be very interesting to have a deterministic polynomial time algorithm to compute the coefficients of modular forms of half integral weight.

Bibliography

[1] Audin, M.: La guerre de recensions autour d'une note de André Weil en 1940. http://arxiv.org/pdf/1109.5230.pdf

[2] Artin, E.: Quadratische Körper im Gebiete der höheren Kongruenzen I, II. Math. Zeitschr. 19 (1924) 153–246

[3] Bombieri, E., Counting points on curves over finite fields (daprs S.A.. Stepanov), Séminaire Bourbaki, Exp. 430, p. 234–241. Lecture Notes in Math., 383, Springer, Berlin, 1974.

[4] http://www.claymath.org/millennium-problems

[5] Couveignes, J.-M. and Edixhoven, B.: Computational aspects of Modular Forms and Galois Representations. Annals of Mathematics Studies 176, Princeton University Press 2011.

[6] Derickx, M., Van Hoeij, M. and Zeng, J.: Computing Galois representations and equations for modular curves $X_H(l)$. http://arxiv.org/pdf/1307.5719.pdf

[7] Dirichlet P.G.L and Dedekind R.: *Vorlesungen ber Zahlentheorie* F. Vieweg und Sohn, 1879 Supplement XI: Über die Theorie der ganzen algebraischen Zahlen.

[8] http://en.wikipedia.org/wiki/Division_polynomials

[9] Elkies, N.: Elliptic and modular curves over finite fields and related computational issues, pages 21–76 in Computational Perspectives on Number Theory: Proceedings of a Conference in Honor of A.O.L. Atkin (D.A. Buell and J.T. Teitelbaum, eds.; AMS 1998).

[10] http://en.wikipedia.org/wiki/Discrete_logarithm_records

[11] http://en.wikipedia.org/wiki/Elliptic_curve_cryptography

[12] Gourdon, X.: Computation of zeros of the Zeta function. http://numbers.computation.free.fr/Constants/Miscellaneous/zetazeroscompute.html

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

[14] Hecke, E.: Über die Zetafunktion beliebiger algebraischer Zahlkörper. Mathematische Werke 7, 159–171. Vandenhoeck-Ruprecht, Göttingen 1970.

[15] http://en.wikipedia.org/wiki/Hilbert's_problems

[16] Hindry, M.: La preuve d'André Weil de l'hypothèse de Riemann pour une courbe sur un corps fini, http://www.math.polytechnique.fr/xups/xups12-02.pdf

[17] Riemann, B.: Über die Anzahl der Primzahlen unter einer gegebenen Grösse (19. Oktober 1859). In: Monatsberichte der Königlichen Preussischen Akademie der Wissenschaften zu Berlin, 1860, S. 671–680

[18] Schmidt, F.K.: Analytische Zahlentheorie in Körpern der Charakteristik p. Math. Zeitschr. 33 (1931)

[19] Stepanov, S.A., On the number of points of a hyperelliptic curve over a finite prime field, Izv. Akad. Nauk SSSR, Ser. Math. 33 (1969), p. 1103–1114.]

[20] Sutherland, D.: Genus 1 point counting records over prime fields https://math.mit.edu/ drew/SEArecords.html

[21] Vercauteren, F.: Counting points on elliptic curves; p-adic algorithms. https://www.cosic.esat.kuleuven.be/publications/talk-63.pdf

[22] Weil, A.: Sur les fonctions algébriques à corps de constantes fini. C. R. Acad. Sci. Paris 210 (1940), p. 592–594.

[23] Weil, A.: On the Riemann hypothesis in function fields. Proc. Nat. Acad. Sci. USA 27 (1941), p. 345–347.

[24] Weil, A.: Sur les courbes algébriques et les variétés qui s'en déduisent. Paris, Hermann, 1948.

Appendix.

The integer t computed by Sutherland is

-20567600940562956287072354495171197849376963352269100129034298282439457997902049595922754072525033751309735814750343243350915783036596028980929619519035647569198933574237708567449417804544527593253489171272835036914287573195186259060507937805196714223995712481127173386249668152305287210010467470890781901246147332517212125625521605551726558627169208412849002751285600599984303993072321301552354770318037680806930569295989240725111217683349520179828098497727732803344520638064502733460261681405906832066362169985659821976097669264357627875294743014049165954756296678742436340184885860523455985052065691113193899099153218693471679493155868676978108334106790272982037127662933425771749796702645059239637511409201528922052535078505128176132033358943834841606855854254746506091627832013963074429073195635233683726249485915962067021637894353186997757910481489556784358787421897666214061826149830661747871603324711207139815142803476791860168691157478407200331589018911087289759611642200722891147432020886233450416233641386483673621053597154449098053198386076704931756218684058977187585742581068679101771458435380874391942946098290760561460645783184379042763328777066300543506565157925043754361183782424472580992276211254913722850401007606958020934687660434823913623842581183947150182078360194646874876064029364183240939177664172102165418956775711526763272193305850435146386442265129449541112284818987192814321299484336893724691274133400966448278123085892349178818111918905465031088986990085862997465940891344284839251695647465377829644880597529413153209333875284267156198318175154856244734335275132494621001952446852583583199264754496938732527393372326375034790053408189958309972564692129575244756666003790317338492900301592175777229587037567689253179901253009116461150887177176391053827898214582098967853943629259218055618911632548625725986031204760147307727593699728775128379092168271121767274773752596424664623997644837861427024527213468730990728496193860866923718742302063574668996192928427531185600483118522050092167617602019573180382041196237505080071936406702382528439958955208274496842320127279061845291668832974850434817931979782757051137260734360596411347250460510493721624777276758648494174926903279111476140958964001568141098863226394847386613566304160265758982926219719676262300861812490242979836754266268188351661135061847887688824809576465291476689413020982267770616019765627112172371744755702275872542089372775277692666711536832348498650264895069527394293908683422549839672589103379227123369049 12