1. Introduction.

Let E be an elliptic curve over a finite field $\mathbf{F}_q$. Then E is a smooth cubic in $\mathbf{P}^2$. It can be given by a Weierstrass equation, an affine version of which is

$$Y^2 + a_1XY + a_3Y = X^3 + a_2X^2 + a_4X + a_6, \qquad \text{with } a_1, a_2, a_3, a_4, a_6 \in \mathbf{F}_q.$$

The unique point at infinity is the neutral element of the group law. It is denoted by ∞. The zeta-function $Z_E(T)$ of E is the power series defined by

$$Z_E(T) \;=\; \sum_{D \geq 0} T^{\deg D} \qquad \text{in } \mathbf{Z}[[T]].$$

Here D runs over the effective divisors of E that are defined over $\mathbf{F}_q$. In this note we prove two theorems concerning $Z_E(T)$.

Theorem 1.1. *Let E be an elliptic curve over $\mathbf{F}_q$. Then the power series $Z_E(T)$ is equal to the rational function*

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

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

And we prove Hasse's Theorem:

Theorem 1.2. *Let E be an elliptic curve over $\mathbf{F}_q$. Then the complex zeroes of the numerator $1 - \tau T + qT^2$ of its zeta function have absolute value $1/\sqrt{q}$.*

Theorem 1.2 is the analogue of the Riemann Hypothesis for the curve E. It implies the inequalities

$$q + 1 - 2\sqrt{q} \;\leq\; \#E(\mathbf{F}_q) \;\leq\; q + 1 + 2\sqrt{q}.$$

It was proved by H. Hasse in 1933. Our approach is elementary and follows a method invented by S.A. Stepanov around 1969. We only make use of the Weierstrass equation and the group law.

2. Rationality of the zeta function.

In this section we prove Theorem 1.1. First we review some properties of elliptic curves. Let E be an elliptic curve given by a Weierstrass equation as in the introduction. The ring R of functions on E without poles outside ∞ is the $\mathbf{F}_q$-algebra generated by the functions X and Y. 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 order of the pole of f at ∞. Then $\deg X = 2$ 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 functions of degree 1. We call $f \in R$ *monic* if the coefficient of its highest degree term is equal to 1. Any $f \in R$ has, counting multiplicities, precisely $\deg f$ zeroes on E^0. 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 $\deg f$.

For two divisors D, D' of E we write $D \sim D'$ if $D - D'$ is principal, i.e. if $D - D'$ is the divisor of a function on E.

Lemma 2.1. *Let E be an elliptic curve and let P, Q be two points on E. Then*

$$P + Q \ \sim\ (P + Q) + \infty.$$

Here the leftmost and rightmost plus signs indicate addition of divisors, while the one in the middle refers to the group law on E.

Proof. The quotient of the equations of the chords or tangents used to add the points P and Q is a function g on E whose divisor is precisely $P + Q - (P + Q) - \infty$. Note that if P and Q are defined over $\mathbf{F}_{q^d}$, then so is g.

Proposition 2.2. *Let E be an elliptic curve over $\mathbf{F}_q$ and let D be a divisor of E of degree d. Then we have $D \sim (d-1)\infty + P$ for a unique point $P \in E(\mathbf{F}_q)$. Moreover, if D is defined over $\mathbf{F}_q$, then $D - (d-1)\infty - P$ is the divisor of a function in $\mathbf{F}_q(E)$.*

Proof. Let $D = \sum_Q n_Q Q$ for certain integers n_Q. We prove the first statement by induction with respect to $w(D) = \sum |n_Q|$. If $w(D) = 0$, then $D = 0$ and we can take $P = \infty$. If $w(D) = 1$, then $D = \pm Q$ for some point $Q \in E(\mathbf{F}_q)$. If $D = Q$ we take $P = Q$. If $D = -Q$, then we take $P = -Q$ where $-Q$ denotes the inverse of Q with respect to the group law on E. Indeed, Lemma 2.1 implies then that $-Q \sim -2\infty + P$.

If $w(D) > 1$, we can write $D = D' \pm Q$ with $w(D')$ is strictly smaller than $w(D)$. When $D = D' + Q$, we have by induction $D \sim (d-2)\infty + P + Q$ and hence $D' \sim (d-1)\infty + (P+Q)$ by Lemma 2.1. If $D = D' - Q$, we have $D \sim d\infty + P + (-Q) \sim (d-1)\infty + (P - Q)$.

To see that the point P is unique, let $D = (d-1)\infty + P = (d-1)\infty + P'$. Then $P \sim P'$ and Lemma 2.1 implies that $P - P' \sim \infty$. If $P \neq P'$, this implies that there is a function on E of degree 1, which is impossible. So $P = P'$. This proves the first statement.

For the second statement, suppose that $D = \sum_Q n_Q Q$ is defined over $\mathbf{F}_q$ and that $D - (d-1)\infty - P$ is the divisor of some function g on E. By the remark made in the proof of Lemma 2.1, the function g is contained in $\mathbf{F}_{q^d}(E)$, where d is so large that all points Q for which $n_Q \neq 0$, are defined over $\mathbf{F}_{q^d}$. Let σ be the Frobenius automorphism given by $\sigma(t) = t^q$ for all $t \in \mathbf{F}_{q^d}$. Since $D - (d-1)\infty - P$ is defined over $\mathbf{F}_q$, the divisors of g and

$\sigma(g)$ are equal. Therefore $\lambda = \sigma(g)/g$ is in $\mathbf{F}^*_{q^d}$. Since the norm of λ from $\mathbf{F}_{q^d}$ to $\mathbf{F}_q$ is 1, there exists an element $\mu \in \mathbf{F}^*_{q^d}$ for which $\lambda = \sigma(\mu)/\mu$. This is "Hilbert 90", which in this case follows easily from the fact that $\mathbf{F}^*_{q^d}$ is a cyclic group.

It follows that $f = g/\mu$ is invariant under σ. Therefore we have $f \in \mathbf{F}_q(E)$, as required.

Example 2.3. We first compute the zeta function of the projective line $\mathbf{P}_1$ over $\mathbf{F}_q$ and then deal in a similar way with zeta functions of elliptic curves E. The zeta function of $\mathbf{P}^1$ over $\mathbf{F}_q$ is defined by

$$Z_{\mathbf{P}_1}(T) = \sum_{D \geq 0} T^{\deg D} \qquad \text{in } \mathbf{Z}[[T]],$$

where D runs over the effective divisors of $\mathbf{P}^1$ that are defined over $\mathbf{F}_q$. Since every divisor is a sum of points, we have

$$Z_{\mathbf{P}_1}(T) = \prod_P \frac{1}{1 - T^{\deg P}}.$$

Here P runs over the conjugacy classes of points of $\mathbf{P}_1$. The zeta function of the affine line $\mathbf{A}^1$ is obtained by omitting the factor $1/(1-T)$ corresponding to the point at infinity. So we have

$$Z_{\mathbf{A}^1}(T) = \sum_{D \geq 0} T^{\deg D} \qquad \text{in } \mathbf{Z}[[T]],$$

where D runs over the effective divisors of $\mathbf{A}^1$. Since the ring $\mathbf{F}_q[X]$ is a principal ideal domain, every divisor $D \geq 0$ of $\mathbf{A}^1$ that is defined over $\mathbf{F}_q$ is the divisor of a unique monic polynomial f in $\mathbf{F}_q[X]$. Moreover, the degree of D is equal to the degree of f. We can therefore compute the zeta function of $\mathbf{A}^1$ by counting polynomials. We find

$$Z_{\mathbf{A}^1}(T) = \sum_{d \geq 0} c_d T^d = \sum_{d \geq 0} q^d T^d = \frac{1}{1 - qT}.$$

Here c_d denotes the number of effective divisors of $\mathbf{A}^1$ of degree d. Since the number of monic degree d polynomials in $\mathbf{F}_q[X]$ is q^d, we have $c_d = q^d$. Going back to the projective line $\mathbf{P}_1$, we obtain the following fomula for the zeta function of $\mathbf{P}^1$ over $\mathbf{F}_q$.

$$Z_{\mathbf{P}_1}(T) = \frac{1}{(1-T)(1-qT)}.$$

This completes the computation of the zeta function of $\mathbf{P}_1$.

Proof of Theorem 1.1. We determine the zeta function of an elliptic curve E over $\mathbf{F}_q$ in a similar way. Let E^0 be the affine curve that is obtained by removing the point ∞ from E. We first determine the zeta-function of E^0. This means that we must count effective divisors on E^0 that are defined over $\mathbf{F}_q$. These are simply divisors on E of the form $\sum_P n_P P$ with $n_P \geq 0$ for all P in E and with $n_\infty = 0$. The only effective divisor of

E^0 of degree 0 is the divisor 0. The effective divisors over $\mathbf{F}_q$ of degree 1 are precisely the points in $E(\mathbf{F}_q) - \{\infty\}$. Denoting $\#E(\mathbf{F}_q)$ by h, there are $h-1$ of them.

Let D be an effective divisor on E^0 of degree $d > 1$. By Proposition 2.2 there exists a unique point $P \in E(\mathbf{F}_q)$ for which we have $-D \sim (-d-1)\infty + P$. More precisely, there exists a function $f \in E(\mathbf{F}_q)$ whose divisor is $D + P - (d+1)\infty$. The function f is unique up to a non-zero constant. Since D is effective, f is contained in the ring R. It has degree d or $d+1$ depending on whether $P = \infty$ or not. If $P \neq \infty$, the function f vanishes in P. Conversely, the divisor on E^0 of any $f \in R$ satisfying these properties is effective and has degree d. Therefore it suffices to count the functions f up to multiplication by non-zero constants. There are q^{d-1} monic $f \in R$ of degree d and there are q^{d-1} monic $f \in R$ of degree $d+1$ that vanish in a given point $P \in E(\mathbf{F}_q) - \{\infty\}$. Therefore there are $q^{d-1} + (h-1)q^{d-1} = hq^{d-1}$ effective divisors on E^0 of degree d.

This computation shows that

$$Z_{E^0}(T) \;=\; 1 + (h-1) + \sum_{d\geq 2} hq^{d-1}T^d \;=\; \frac{1 + (h-q-1)T + qT^2}{1-qT}.$$

The zeta function of E is obtained from the one of E^0 in the same way the zeta function of $\mathbf{P}_1$ is obtained from the one of $\mathbf{A}^1$. In order to take into account the point at infinity, we multiply $Z_{E^0}(T)$ by the factor $1/(1-T)$. This gives

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

where $\tau = q + 1 - h$. This proves Theorem 1.1.

3. An upper bound.

In this section we obtain an upper bound for the number of points of an elliptic curve over a finite field. This is the key ingredient in the proof of Theorem 1.2. Our method is due to S.A. Stepanov.

We introduce some notation. Recall that E is given by a Weierstrass equation and that R is the $\mathbf{F}_q$-algebra generated by the functions X and Y. For $a \geq 0$ let H_a denote the $\mathbf{F}_q$-vector space

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

Since R does not contain any functions $f \in R$ with $\deg f = 1$, the space H_a consists only of constant functions when $a = 0$ or 1 and therefore has dimension 1. In general we have the folllowing result.

Lemma 3.1. *For $a \geq 1$, the monomials X^i and YX^j with $2i \leq a$ and $2j+3 \leq a$ are an $\mathbf{F}_q$-basis for H_a. In particular, H_a has $\mathbf{F}_q$-dimension a.*

Proof. The monomials certainly generate H_a. On the other hand, the orders of their poles at ∞ are all distinct. Therefore the monomials are linearly independent and hence form a basis of H_a. One checks that there are precisely a distinct monomials of degree $\leq a$. This proves the lemma.

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

Lemma 3.2. *Let $a, b \geq 1$ and let $H_a^q H_b$ denote the $\mathbf{F}_q$-vector space generated by the functions fg where $f \in H_a^q$ and $g \in H_b$. If $b < q$, Then $H_a^q H_b$ has $\mathbf{F}_q$-dimension ab.*

Dimostrazione. There exists a basis $e_1, \ldots, e_a$ of H_a and there exists a basis $f_1, \ldots, f_b$ di H_b of monomials as in Lemma 2. Clearly the functions $e_i^q f_j$ with $1 \leq i \leq a$ and $1 \leq j \leq b$ generate $H_a^q H_b$. We have

$$\deg e_i^q f_j = q \deg e_i + \deg f_i.$$

Since $\deg f_i \leq b < q$ the degrees $\deg e_i^q f_j$ are all distinct. If an $\mathbf{F}_q$-linear combination $\sum_{i,j} \lambda_{ij} e_i^q f_j$ is zero, then necessarily $\lambda_{ij} = 0$ for every i, j. This proves that the functions $e_i^q f_j$ are independent and form an $\mathbf{F}_q$-basis. Therefore the dimension of $H_a^q H_b$ is ab. This proves the lemma.

From now on we assume that $a, b \geq 1$ with $b < q$. Lemma 3.1 implies that the $\mathbf{F}_q$-linear map

$$\vartheta : H_a^q H_b \longrightarrow H_a H_b^q$$

given by

$$e_i^q f_j \mapsto e_i f_j^q, \qquad \text{per } 1 \leq i \leq a \text{ e } 1 \leq j \leq b,$$

is well defined.

The following proposition is the key ingredient in the proof of Theorem 3.4.

Proposition 3.3. *Let $a, b \geq 1$ with $b < q$. If the map $\vartheta$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 $F = \sum \lambda_{ij} e_i^q f_j$ for certain $\lambda_{ij} \in \mathbf{F}_q$ and let $P \in E(\mathbf{F}_{q^2}) - \{\infty\}$. Then

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

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.

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

The rightmost inequality follows from the fact that $F \in H_a^q H_b \subset H_{aq+b}$. This proves the proposition.

Theorem 3.4. *Let E be an elliptic curve defined over a finite field $\mathbf{F}_q$. 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

$$\dim H_a^q H_b > \dim H_a H_b^q.$$

Since $b < q$, Lemma 3.2 implies that $H_a^q H_b$ has dimension ab. Lemma 3.2 cannot be applied to $H_a H_b^q$. In some sense this is the point of the proof. But we still know that $H_a H_b^q$ is a subspace of H_{a+bq} and hence 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 3.3, we choose a as small as possible. Since the inequality $ab > a + bq$ must be satisfied, the minimal choice for a 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 3.3 becomes $(q + 2)q + q - 1 + 1 = q^2 + 3q$, as required.

4. The Riemann Hypothesis.

Let E be an elliptic curve over $\mathbf{F}_q$. In this section we prove that the complex zeroes of the numerator of its zeta function have absolute value $1/\sqrt{q}$. The key ingredient is the inequality af Theorem 3.4. First we use the proof of Theorem 3.4 to obtain a lower bound for $\#E(\mathbf{F}_{q^2})$.

Proposition 4.1. *Let E be an elliptic curve over $\mathbf{F}_q$ and suppose that $q \geq 5$. Then we have*

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

Proof. The set Ω of points (x, y) of E^0 for which $x \in \mathbf{F}_q$ admits two commuting involutions. The elliptic involution switches (x, y) and $(x, \overline{y})$, where $\overline{y}$ denotes $-y - a_1 x - a_3$. The automorphism σ of $\mathbf{F}_{q^4}$ given by $\sigma(t) = t^{q^2}$ also acts on Ω. It maps a point $(x, y) \in \Omega$ to $(\sigma(x), \sigma(y)) = (x, y^{q^2})$. Since for a given point (x, y) there ia at most one other point on E with X-coordinate equal to x, namely $(x, \overline{y})$, we either have $\sigma(y) = y$ or $\sigma(y) = \overline{y}$.

The subset $\{(x, y) \in \Omega : \sigma(y) = y\}$ is the set $E(\mathbf{F}_{q^2}) - \{\infty\}$. Theorem 3.4 provides an estimate for its size. In this section we estimate the size of the set

$$W = \{(x, y) \in \Omega : \sigma(y) = \overline{y}\}$$

with the method of section 3. Let a, b be as in the proof of Theorem 3.4. Note that the spaces H_a and H_b are preserved by the automorphism of R given by $f(X, Y) \mapsto (X, -Y - a_1 X - a_3)$. Consider the $\mathbf{F}_q$-linear map

$$\vartheta' : H_a^q H_b \; \longrightarrow H_a H_b^q$$

defined by

$$e_i^q f_j \; \mapsto \; \overline{e_i} f_j^q.$$

Every function $F \in \ker \vartheta$ vanishes on the set W. Indeed, let $F = \sum \lambda_{ij} e_i^q f_j$ for certain $\lambda_{ij} \in \mathbf{F}_q$ and let $P \in W$.

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

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

$$\#W \;\leq\; q^2 + 3q.$$

Since the Weierstrass equation is cubic, there are at most three points $(x, y) \in \Omega$ with $y = \overline{y}$. Therefore we have $\#\Omega \geq 2q^2 - 3$. Since $\Omega = W \cup E(\mathbf{F}_{q^2}) - \{\infty\}$, we find

$$\#E(\mathbf{F}_{q^2}) \;\geq\; \#\Omega - \#W \;\leq\; 2q^2 - 3 - (q^2 + 3q) = q^2 - 3q - 3$$

as required.

Let $1 - \tau T + qT^2$ be the numerator of the zeta function of E and let π and π' be the complex zeroes of the reciprocal polynomial $T^2 - \tau T + q$.

Lemma 4.2. *For every $d \geq 1$, we have*

$$\#E(\mathbf{F}_{q^d}) \;=\; q^d + 1 - \pi^d - \pi'^d.$$

Proof. By Theorem 1.1 we have

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

Combining this with the identity

$$Z_E(T) = \sum_{D \geq 0} T^{\deg D} \;=\; \prod_P \frac{1}{1 - T^{\deg P}},$$

we obtain

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

For $d \geq 1$ we write here a_d for the number of points on E of degree d up to conjugacy. For every $e \geq 1$ we have $\#E(\mathbf{F}_{q_e}) = \sum_{d|e} da_d$. Taking the logarithmic derivative of this identity, expanding the geometric series and comparing coefficients shows that we have $q^e + 1 - \pi^e - \pi'^e = \sum_{d|e} da_d$ for every $d \geq 1$. This proves the Lemma.

Theorem 4.3. *The complex zeroes π and π' of the polynomial $T^2 - \tau T + q$ have absolute value $\sqrt{q}$. In particular $\pi' = \overline{\pi}$.*

Proof. Lemma 4.2, Theorem 3.4 and Proposition 4.1 provide us with the inequalities

$$q^d - 3q^{d/2} - 3 \;\leq\; q^d + 1 - \pi^d - \pi'^d \;\leq\; q^d + 3q^{d/2}, \qquad \text{for even } d \geq 0.$$

Therefore we have

$$|\pi^d + \pi'^d| \;\leq\; 3q^{d/2} + 3, \qquad \text{for even } d \geq 0.$$

This implies that the power series $\sum_{d \geq 0,\, \text{even}} (\pi^d + \pi'^d) t^d$ converges for $t \in \mathbf{C}$ of absolute value $< 1/\sqrt{q}$. On the other hand, summing the geometric series we obtain an expression with denominators equal to $1 - (\pi t)^2$ and $1 - (\pi' t)^2$. Therefore $|\pi^2|, |\pi'^2| \leq q$. Since the product of π and π' is q, we find that $|\pi| = |\pi'| = \sqrt{q}$ as required.

The inequalities of Theorem 3.4 and Proposition 4.1 have only been proved for $q \geq 5$. However, when $q < 5$, we have $q^d > 5$ for $d \geq 3$. This implies that we still have the inequality for even degrees $d \geq 6$. Therefore the convergence radius of the series is not affected and the conclusion is the same for $q < 5$. This proves the theorem.