ARTIN'S CONJECTURE FOR PRIMITIVE ROOTS

M. Ram Murty

§0. Introduction

In his preface to *Diophantische Approximationen*, Hermann Minkowski expressed the conviction that the "deepest interrelationships in analysis are of an arithmetical nature." Gauss described one such remarkable interrelationship in articles 315 - 317 of his *Disquisitiones Arithmeticae*. There, he asked why the decimal fraction of 1/7 has period length 6:

$$\frac{1}{7} = 0.142857\ 142857\ 142857\ ...$$

whereas 1/11 has period length of only 2:

$$\frac{1}{11} = 0.09\ 09\ 09\ ...$$

Why does 1/99007599, when written as a binary fraction (that is, expanded in base 2), have a period of nearly 50 million 0's and 1's ? To answer these questions, Gauss introduced the concept of a primitive root.

To motivate our discussion, let p be a prime ($\neq 2, 5$), and let

$$\frac{1}{p} = .a_1 a_2 ... a_k ...$$

be its decimal expansion with period k. Then, it is easily seen that

$$\frac{1}{p} = \left(\frac{a_1}{10} + \frac{a_2}{10^2} + ... + \frac{a_k}{10^k}\right)\left(1 + \frac{1}{10^k} + \frac{1}{10^{2k}} + ...\right) = \frac{M}{10^k - 1},$$

where M is some integer. Therefore, $10^k - 1 = Mp$. That is,

$$(\ddagger) \qquad 10^k \equiv 1 (\text{mod } p).$$

The period k must satisfy the above congruence, and k is characterized as the smallest exponent for which (‡) is satisfied.

If k is the smallest integer satisfying (‡), we say that 10 has *order* k (mod p). By Fermat's little theorem,

$$10^{p-1} \equiv 1 (\text{mod } p)$$

and therefore

$$0 < k \leq p - 1.$$

Since by a theorem of Lagrange, $k|p-1$, the largest period of $1/p$ can occur if and only if 10 has order $p-1 \pmod p$. In such a case, 10 is called a *primitive root* $\pmod p$. More generally, if $p \not| a$, and the smallest k such that

$$a^k \equiv 1 (\text{mod } p)$$

is $p-1$, then a is a *primitive root* $\pmod p$. If n is the product of distinct primes p_i, the period of $1/n$ in base a is the least common multiple of the orders of $a (\text{mod } p_i)$, provided a and n are relatively prime. In the case $n = 99007599 = (9851)(9949)$, 9851 and 9949 are both prime, 2 is a primitive root for both primes and the period of $1/99007599$ in base 2 is

$$\text{lcm } [9850,\ 9948] = 48,993,900.$$

Therefore, $1/99007599$ has a binary expansion of period 48,993,900. Such facts are used by the computer scientists to generate pseudo- random binary sequences.

This remarkable interrelationship does not end here. Gauss raised the question of how often 10 is a primitive root mod p, as p varies over the primes, but made no specific conjecture. A precise conjecture was formulated by E. Artin [1, p.viii-x] in 1927 during a conversation with H. Hasse. He hypothesised that for any given non-zero integer a other than 1, -1 or a perfect square, there exist infinitely many primes p for which a is a primitive root $\pmod p$. Moreover, if $N_a(x)$ denotes the number of such primes up to x, he conjectured an asymptotic formula of the form

$$N_a(x) \sim A(a) \frac{x}{\log x}$$

as $x \to \infty$, where $A(a)$ is a certain constant depending on a. This is now known as Artin's conjecture.

> "We all believe," Artin wrote [1, p. 534], "that mathematics is an art. The author of a book, the lecturer in a classroom tries to convey the structural beauty of mathematics to his readers, to his listeners. In this attempt, he must always fail. Mathematics is logical to be sure, each conclusion is drawn from previously derived statements. Yet the whole of it, the real piece of art, is not linear; worse than that, its perception should be instantaneous. We all have experienced

on some rare occasions the feeling of elation in realizing that we have enabled our listeners to see at a moment's glance the whole architecture and all its ramifications."

Artin had a profound effect in the shaping of the mathematics of our time. His deep insights into class field theory, reciprocity laws, non-abelian L series and the arithmetic of function fields have guided the development of modern number theory. Artin's conjecture is one celebrated instance of his mathematical intuition and creativity. It is the focal point of diverse areas of mathematics such as group theory, algebraic and analytic number theory, and algebraic geometry. In fact, Artin's motivation originated in algebraic number theory. His intuition was as follows.

If a is a primitive root (mod p), then it is necessary and sufficient that

$$a^{(p-1)/q} \not\equiv 1 \pmod p$$

for every prime divisor q of $p-1$. For if k is the order of a mod p, then $k|p-1$ and if $k \neq p-1$, then $k|(p-1)/q$ for some prime divisor q of $p-1$. Heuristically, a is a primitive root (mod p) if the "events"

$$(*) \qquad \begin{aligned} p &\equiv 1 \pmod q \\ a^{(p-1)/q} &\equiv 1 \pmod p \end{aligned}$$

do not occur. To invert the problem, fix q and find the probability that a prime p satisfies the above conditions. By Dirichlet's theorem, $p \equiv 1 \pmod q$ is true for primes p with frequency

$$\frac{1}{q-1}.$$

One would expect that

$$a^{(p-1)/q} \equiv 1 \pmod p$$

occurs with probability $1/q$. The probability that both events occur is $1/q(q-1)$, since these events can be assumed to be independent. To ensure that a is a primitive root (mod p), the above events must not occur for every q. This suggests a probability of

$$\prod_q \left(1 - \frac{1}{q(q-1)}\right)$$

for such primes.

In 1967, Hooley [10] proved both Artin's conjecture and an asymptotic formula for $N_a(x)$ subject to the assumption of the generalised Riemann hypothesis. This hypothesis, which is still unproved, is the natural extension of the classical Riemann hypothesis to the Dedekind zeta function of a number field (see next section). The implication of Hooley's theorem is that if Artin's conjecture is false, then the generalised Riemann hypothesis is false.

In 1983, Rajiv Gupta and the author [6] proved *without any hypothesis*, that there is a set of 13 numbers such that, for at least one of these 13 numbers, Artin's conjecture is true. This established, for the first time, the existence of some a for which Artin's conjecture is true. Moreover, this proof demonstrated that the conjecture was also true for almost all a. Gupta, Kumar Murty and the author [8] subsequently reduced the size of this set to 7. In 1985, Heath-Brown [9] refined this result to obtain a set of three numbers, by an application of the "Chen - Iwaniec switching". (Switching first occurs in Lemma 4.4 of Iwaniec [13] in the study of primes of the form $\phi(x,y)+A$, where ϕ is a quadratic form. It was also discovered independently by Chen in his quasi - resolution of the twin prime problem and the Goldbach conjecture.) More precise results will be stated below. One consequence of the Heath-Brown refinement is

Theorem 1. One of 2, 3, 5 is a primitive root (mod p) for infinitely many primes p.

In order that a be a primitive root for a prime p not dividing a, it is clearly necessary and sufficient that for each prime q,

$$(\dagger) \qquad p \equiv 1 (\text{mod } q) \Rightarrow a^{(p-1)/q} \not\equiv 1 (\text{mod } q).$$

Using this criterion, several nineteenth century mathematicians observed that 2 is a primitive root (mod p) whenever p is of the form $4q+1$, where q is prime. In such a case, $p-1$ has only two prime divisors, namely 2 and q. Since q is odd,

$$p = 4q + 1 \equiv 5 (\text{mod } 8)$$

and

$$2^{(p-1)/2} \equiv -1 (\text{mod } p)$$

by a special case of quadratic reciprocity. Also,

$$2^{(p-1)/q} = 2^4 \equiv 1 (\text{mod } p)$$

implies that $p = 3$ or 5 and neither of these primes is of the form $4q + 1$. Therefore, (†) is satisfied and 2 is a primitive root (mod p). It is a classical unsolved problem to determine whether there are infinitely many primes of the form $(p-1)/4$. It is known by sieve methods that $(p-1)/4$ is infinitely often a product of at most two primes and both these prime factors are greater than p^θ, with $\theta > \frac{1}{4}$. Thus, there are not many conditions specified by (†) to ensure that a is a primitive root (mod p) for such primes. This is the essential fact that enables us to prove Theorem 1.

§1. Intuition of Artin and Hooley's theorem

An *algebraic number* α is a complex number which satisfies an equation of the form

$$c_n\alpha^n + c_{n-1}\alpha^{n-1} + ... + c_1\alpha + c_0 = 0 \tag{0}$$

where $c_i,\ 0 \le i \le n$, are rational numbers. If n is the smallest natural number for which an equation of the form (0) holds, α is said to be of *degree* n. The set of all numbers of the form

$$b_0 + b_1\alpha + ... + b_{n-1}\alpha^{n-1}$$

with $b_i \in \mathbf{Q}$, forms a field K, which we denote by $\mathbf{Q}(\alpha)$, and we say that K has degree n. The set of all $\beta \in \mathbf{Q}(\alpha)$ which satisfy a relation of the form

$$\beta^n + a_{n-1}\beta^{n-1} + ... + a_1\beta + a_0 = 0$$

with $a_i \in \mathbf{Z}$ forms a *ring* called the *ring of integers* of K, which we denote by O_K.

O_K enjoys some remarkable properties which were first discovered and systematically used by Dedekind. For instance, every ideal of O_K can be factored uniquely into a product of prime ideals. This is the number field analogue of the ancient theorem that every natural number is a unique product of prime numbers.

If p is a prime number, the ideal generated by p in O_K, namely pO_K, factorises as a product of distinct prime ideals $\mathcal{P}_i$:

$$pO_K = \mathcal{P}_1^{e_1} ... \mathcal{P}_g^{e_g}.$$

Dedekind proved the important relation

$$n = \sum_{i=1}^{g} e_i f_i$$

where f_i, defined by the cardinality of the quotient ring $O_K/\mathcal{P}_i = p^{f_i}$, is called the *degree* of the prime ideal $\mathcal{P}_i$. A prime p is said to *split completely* in K if

$$pO_K = \mathcal{P}_1 ... \mathcal{P}_n$$

with distinct prime ideals $\mathcal{P}_i$ of degree one. Let $\pi_K(x)$ denote the number of primes $p \le x$ which split completely in K. If K is a normal extension of $\mathbf{Q}$, then a theorem of Chebotarev states that the density of primes p which split completely in K is $1/n$. That is,

$$\lim_{x\to\infty} \frac{\pi_K(x)}{\pi_{\mathbf{Q}}(x)} = \frac{1}{n}.$$

Artin realized that the two conditions of $(*)$ are satisfied if and only if p splits completely in the (normal) extension $L_q = \mathbf{Q}(\zeta_q, a^{1/q})$, where ζ_q denotes a primitive q-th root of unity. If a is squarefree, then the degree of $L_q/\mathbf{Q}$ is $q(q-1)$ and by the theorem of Chebotarev, the density of primes which split completely in L_q is

$$\frac{1}{q(q-1)}.$$

Now, a is a primitive root (mod p) if and only if the above two conditions $(*)$ do not hold for all q. That is, a is a primitive root (mod p) if and only if p does not split completely in any L_q. We would therefore expect the density of such primes to be

$$\prod_q \left(1 - \frac{1}{q(q-1)}\right).$$

This was the heuristic reasoning which led Artin to formulate his conjecture. Computations by Lehmer revealed that some adjustment is needed in the above conjectured density for a general a, in order to take into account the possible dependence in the fields L_q. It is clear that if we let (for each squarefree integer k),

$$L_k = \prod_{q|k} L_q$$

be the compositum of the fields L_q for primes q dividing k and let n_k be the degree of L_k over $\mathbf{Q}$, then the set of primes which do not split completely in any L_k has density

$$A(a) = \sum_{k=1}^{\infty} \frac{\mu(k)}{n_k}$$

by the inclusion-exclusion principle. The Möbius function μ is defined by

$$\mu(k) = \begin{cases} (-1)^r, & \text{if } k \text{ is the product of } r \text{ distinct primes} \\ 0, & \text{otherwise.} \end{cases}$$

It can be shown that if k is odd, then the fields L_q for $q|k$ are completely linearly disjoint. Therefore, for odd k,

$$n_k = \prod_{q|k} n_q,$$

and for even subscripts, n_{2k} is equal to n_k or $2n_k$ according to whether $\sqrt{a}$ is or is not contained in the field of k-th roots of unity. If $a = bc^2$ with b squarefree, then the criterion for $\sqrt{a}$ to be contained in the field of k-th roots of unity (for odd squarefree k) is that $b|k$ and $b \equiv 1 \pmod 4$. The formula

$$A(a) = \delta \prod_q (1 - \frac{1}{n_q})$$

where

$$\delta = \begin{cases} 1 & \text{if } b \not\equiv 1 (\text{mod } 4), \\ 1 - \mu(b)\prod_{q|b} \frac{1}{n_q - 1} & \text{if } b \equiv 1 (\text{mod } 4). \end{cases}$$

is the "obvious" modification for the formulation of the precise conjecture. Subject to the generalised Riemann hypothesis, Hooley proved that this modified density is the correct density of primes for which a is a primitive root.

To make this precise, we need an effective version of the Chebotarev density theorem. In retrospect, it seems rather surprising that such a theorem was not proved until half a century later. One reason for this might be inadequate communication between analytic and algebraic number theorists.

The *Dedekind zeta function* of a number field K is defined by

$$\zeta_K(s) = \sum_{\mathcal{A}} N(\mathcal{A})^{-s},$$

where $N(\mathcal{A})$ denotes the cardinality of $O_K/\mathcal{A}$, and the sum is over all ideals $\mathcal{A}$ of O_K. $\zeta_K(s)$ converges absolutely for $Re(s) > 1$. Hecke first showed that $\zeta_K(s)$ admits an analytic continuation to the entire complex plane, except for a simple pole at $s = 1$, and satisfies a functional equation analogous to that of the classical Riemann zeta function. In the same spirit, there is the *generalised Riemann hypothesis*: for $Re(s) > 0$,

$$\zeta_K(s) = 0 \Rightarrow Re(s) = \frac{1}{2}.$$

Under this hypothesis, Hooley proceeded as follows. If π_q denotes the set of primes p which split completely in L_q and $\pi_k(x)$ denotes the number of primes p up to x contained in $\cap_{q|k}\pi_q$, then by the inclusion-exclusion principle

$$N_a(x) = \sum_{k=1}^{\infty} \mu(k)\pi_k(x).$$

Since $\pi_k(x) = 0$ for $k > x$, this is a finite sum. If the analogue of the Riemann hypothesis for the Dedekind zeta function corresponding to the field L_d is assumed, then one can prove along classical lines, the following prime number theorem:

$$(C) \qquad \pi_d(x) = \frac{li\ x}{n_d} + O(x^{1/2} \log dx)$$

where

$$li\ x = \int_2^x \frac{dt}{\log t}$$

and the constant implied by the O symbol is absolute. (Here and elsewhere, the O notation means the following: we write that $f(x) = O(g(x))$ if $|f(x)| \leq Cg(x)$ for some constant C.) It is clear that if we insert this in the above formula for $N_a(x)$, then the contribution from the error term is too large. For this reason, Hooley decomposed the sum in the following way, for $k = \prod_{p<z} p$:

$$N_a(x) \leq \sum_{d|k} \mu(d)\pi_d(x),$$

because the right hand side enumerates primes that satisfy a proper subset of the conditions specified by (†). Moreover, if $M(x; z, w)$ denotes the number of primes $p \leq x$ that satisfy $(*)$ for some $z < q < w$, then

$$N_a(x) \geq \sum_{d|k} \mu(d)\pi_d(x) - M(x; z, x).$$

If $z = \frac{1}{6}\log x$, then from (C),

$$N_a(x) = A(a)li\ x + O(\frac{x}{\log^2 x}) + O(M(x; z, x)).$$

To treat the term $M(x; z, x)$, write

$$M(x; z, x) \leq M(x; z, \frac{x^{1/2}}{\log^2 x}) + M(x; \frac{x^{1/2}}{\log^2 x}, x^{1/2} \log x) + M(x; x^{1/2} \log x, x).$$

The first two terms on the right hand side are shown to be

$$O(\frac{x \log\log x}{\log^2 x})$$

by using (C) and estimates derived from sieve methods. The generalised Riemann hypothesis is unable to estimate the third term in a satisfactory way. The treatment of $M(x; x^{1/2}\log x, x)$ is ingenious and simple. A variation of this method appears in our quasi-resolution of Artin's conjecture [6] and the details are given here. Firstly, the term in question enumerates primes p such that

$$a^{(p-1)/q} \equiv 1 \pmod p$$

for some prime $q > x^{1/2}\log x$. Thus, such a prime p divides

$$\prod_{m < \frac{x^{1/2}}{\log x}} (a^m - 1)$$

since

$$\frac{p-1}{q} < \frac{x^{1/2}}{\log x}.$$

But the number of prime divisors of $a^m - 1$ is at most $m \log a$ (using the fact that a natural number n has at most $\log n$ prime factors). Therefore, the total number of prime factors in question cannot exceed

$$\sum_{m < \frac{x^{1/2}}{\log x}} m \log a = O(\frac{x}{\log^2 x}).$$

Putting these three estimates together, we conclude that

$$M(x; z, x) = O(\frac{x \log\log x}{\log^2 x}).$$

This proves that, subject to the generalised Riemann hypothesis,

$$N_a(x) = A(a)\frac{x}{\log x} + O(\frac{x \log\log x}{\log^2 x}),$$

which completes the account of Hooley's theorem.

§2. Elliptic analogues

In the fall of 1983, at the Institute for Advanced Study in Princeton, the author and Rajiv Gupta were considering some unresolved conjectures of Lang and Trotter concerning elliptic curves. In 1976, Lang and Trotter [14] formulated the elliptic analogue of the Artin conjecture. More precisely, let E be an elliptic curve over $\mathbf{Q}$. That is, E can be viewed as the solutions of the equation

$$y^2 = x^3 + g_2 x + g_3, \quad g_2, g_3 \in \mathbf{Q}.$$

If K is a field, then we define

$$E(K) = \{(x, y) \mid x, y \in K, y^2 = x^3 + g_2 x + g_3\}.$$

Poincaré turned this set into an additive abelian group by defining the addition of two points $P = (x_1, y_1)$, $Q = (x_2, y_2)$ as $R = (x_3, y_3)$, where for $x_1 \neq x_2$,

$$x_3 = -x_1 - x_2 + \frac{(y_2 - y_1)^2}{(x_2 - x_1)^2}$$

and

$$y_3 = -\{x_3\left(\frac{y_2 - y_1}{x_2 - x_1}\right) + \frac{x_2 y_1 - x_1 y_2}{x_2 - x_1}\},$$

and if $x_1 = x_2$, then

$$x_3 = -2x_1 + \left(\frac{3x_1^2 + g_2}{2y_1}\right)^2$$

$$y_3 = (x_1 - x_3)\left(\frac{3x_1^2 + g_2}{2y_1}\right) + y_1.$$

(These formulas are valid for any field of characteristic $\neq 2, 3$.) It is a classical theorem of Mordell and Weil that $E(K)$ thus defined is a finitely generated abelian group. The number of independent generators is called the *rank* of $E(K)$. Geometrically, the addition of two points of the cubic is the reflection in the x-axis of the third point determined by the secant (or tangent) joining the two given points (see Figure 1).

The Mordell-Weil theorem says therefore that all of the points with rational coordinates can be obtained by tangent-secant process from a finite number of points. The map ϕ_n defined by $\phi_n(P) = nP$, for $P \in E(K)$ defines an endomorphism of $E(K)$, for every integer n. Therefore, the endomorphism ring of $E(K)$ contains a natural copy of $\mathbf{Z}$. Deuring proved the remarkable fact that if the endomorphism ring is strictly larger than the copy of $\mathbf{Z}$, then it can be naturally identified to a subring of the ring of integers in an imaginary quadratic field. These are called elliptic curves *with complex multiplication* whenever these extra endomorphisms exist. The corresponding imaginary quadratic field is called the CM field. If a is a rational point of infinite order, the elliptic analogue of Artin's conjecture is to determine the density of primes p for which $E(\mathbf{F_p})$ (the rational points on the curve E viewed over the finite field $\mathbf{F_p}$) is generated by $\overline{a}$, the reduction of $a(\text{mod } p)$. Such a point is called a primitive point. It was conjectured by Lang and Trotter that the density of primes p for which a is a primitive point always exists. Attempts to prove this conjecture by the method of Hooley, outlined in section 1, proved futile. The error terms in the Chebotarev density theorem were too large. Nevertheless, Gupta and the author were successful in carrying out a variation of the method of Hooley for curves with complex multiplication and primes p which split completely in the corresponding CM field. The problem still remains open in the non-CM case.

Lang and Trotter formulated higher rank analogues of their conjecture. A natural question arises: does the problem get simpler if the rank of the curve goes up ? The elliptic analogue of lemma 3 in the next section was the key. Thanks to lemma 3, the problem does become simpler when the rank increases and the full strength of the Riemann hypothesis need not be invoked. In fact, Gupta and the author proved:

Theorem 2. Let E be an elliptic curve with CM by an order in an imaginary quadratic field. If the rank of $E(\mathbf{Q}) \geq 6$, then there is a set S of 2^{18} rational points such that at least one of these points is a primitive point $(\text{mod } p)$ for infinitely many primes p.

But then, what about the classical case ? It seems that no one had previously formulated the higher rank analogue of the classical conjecture of Artin. Going back and formulating the higher rank analogue, the authors of [7] were led to the beginnings of the results described in the previous section. Together with Heath-Brown's refinement, the theorem stated at the outset is the most that we can say on Artin's conjecture at present.

§3. Quasi-resolution of Artin's conjecture

Suppose now that r integers $a_1, a_2, ..., a_r$ are given which are multiplicatively independent. That is, if there are integers $n_1, n_2, ..., n_r$ such that

$$a_1^{n_1} \cdots a_r^{n_r} = 1$$

then $n_1 = n_2 = \cdots = n_r = 0$. Let Γ be the subgroup of $\mathbf{Q}^\times$ generated by $a_1, ..., a_r$. Denote by Γ_p the image of Γ (mod p). Thus, if $r = 1$, and $a = a_1$, then a is a primitive root (mod p) if and only if Γ_p is the full group of co-prime residue classes (mod p). In the general case of arbitrary r, consider the size of Γ_p as p varies. The key lemma is:

Lemma 3. The number of primes p such that $|\Gamma_p| \le y$ is $O(y^{1+\frac{1}{r}})$.

Proof. Consider the set $S = \{a_1^{n_1} \cdots a_r^{n_r} : 0 \le n_i \le y^{\frac{1}{r}}, 1 \le i \le r\}$. As $a_1, a_2, ..., a_r$ are multiplicatively independent, the number of elements of S exceeds

$$([y^{\frac{1}{r}}] + 1)^r > y.$$

If p is a prime such that $|\Gamma_p| \le y$, then two distinct elements of S are congruent (mod p). Hence, p divides the numerator N of

$$a_1^{m_1} \cdots a_r^{m_r} - 1$$

for some $m_1, m_2, ..., m_r$ satisfying

$$|m_i| \le y^{1/r}, \quad 1 \le i \le r.$$

For a fixed choice of $m_1, ..., m_r$, the number of such primes is bounded by

$$\log N \le y^{1/r} \sum_{i=1}^{r} \log a_i = O(y^{1/r}).$$

Taking into account the number of possibilities for $m_1, ..., m_r$, the total number of primes p cannot exceed

$$O(y^{1+\frac{1}{r}}).$$

This completes the proof of the lemma.

Remark. This is the higher dimensional version of the argument used to treat

$$M(x; x^{1/2} \log x, x)$$

of the previous section. In a different context, this lemma was first proved by Matthews [15].

A higher rank analogue of Artin's conjecture can be formulated. Let Γ be the subgroup of $\mathbf{Q}^\times$ generated by r multiplicatively independent natural numbers $a_1, ..., a_r$.

Question. When is Γ_p the set of coprime residue classes (mod p) for infinitely many primes p ?

This question is considered for $r \geq 3$ in a leisurely fashion, ignoring technicalities. Suppose that $(p-1)/2$ is a product of two primes each greater than p^θ, with $\theta > 1/4$. Let $\mathcal{B}$ be the set of such primes and denote by $\mathcal{B}(x)$ the number of such primes up to x. By sieve methods, it can be shown that

$$\mathcal{B}(x) \geq \frac{cx}{\log^2 x}$$

for some positive constant c. The interested reader should consult the excellent exposition of Bombieri [2, pp. 65-75] for the technical details. Then, since Γ_p is a subgroup of the coprime residue classes (mod p), the size of Γ_p does not have many possibilities. This is because $|\Gamma_p|$ divides $p-1$. Consequently, if either of the two prime factors divides the index of Γ_p in $\mathbf{F}_\mathbf{p}^\times$, then

$$|\Gamma_p| \leq p^{1-\theta} \;\leq\; x^{1-\theta},$$

for $p \leq x$. Since $r \geq 3$, we find by the lemma that the number of such primes does not exceed

$$O(x^{4(1-\theta)/3})$$

and since $\theta > 1/4$, this bound is certainly

$$o(\frac{x}{\log^2 x}).$$

Therefore, if the primes for which $|\Gamma_p| \leq p^{1-\theta}$ are thrown away, then for almost all primes contained in $\mathcal{B}$,

$$|\Gamma_p| > p^{1-\theta}.$$

For $p \in \mathcal{B}$, the only divisors of $p-1$ which satisfy the above inequality are $(p-1)/2$ and $p-1$. Therefore

$$|\Gamma_p| = \frac{p-1}{2} \quad or \quad p-1.$$

If the first possibility can be eliminated, then Γ_p is the group of coprime residue classes (mod p). If $|\Gamma_p| = \frac{p-1}{2}$, then $\Gamma_p = (\mathbf{F}_\mathbf{p}^\times)^\mathbf{2}$. Impose the restriction that at least one of $a_1, ..., a_r$ is not a perfect square. The sieve methods alluded to earlier produce a set of primes $\mathcal{B}'$ such that $|\Gamma_p| \neq \frac{p-1}{2}$ and $\frac{p-1}{2}$ is either prime or a product of two primes, each greater than p^θ, $\theta > 1/4$. The method indicated above forces $\Gamma_p = \mathbf{F}_\mathbf{p}^\times$.

Thus, if $r \geq 3$, and if at least one of $a_1, a_2, ..., a_r$ is not a perfect square, then there are infinitely many primes p such that Γ_p is the set of coprime residue classes (mod p).

To take a specific case, we deduce that 2, 3, and 5 together generate the coprime residue classes (mod p) for at least

$$\frac{cx}{\log^2 x}$$

primes $p \leq x$. To produce an a which generates $\mathbf{F}_\mathbf{p}^\times$, consider a variation of Hooley's argument in section 2. Suppose that none of 2, 3, 5 is a primitive root (mod p) for $p \in \mathcal{B}'$. If $\frac{p-1}{2}$ is a prime, then one of 2, 3, 5 is a primitive root (mod p) for otherwise, 2, 3, 5 would generate a subgroup strictly smaller than $\mathbf{F}_\mathbf{p}^\times$, which is a contradiction. Therefore $\frac{p-1}{2}$ is not prime. In this case,

$$\frac{p-1}{2} = q_1 q_2, \quad q_1 < q_2,$$

with $q_1 > p^\theta$, $\theta > 1/4$, the order (mod p) of each of 2, 3, 5 must be one of $q_1, q_2, 2q_1$, or $2q_2$. Clearly $q_1 < p^{1/2}$ for otherwise, $\frac{p-1}{2} > p$, a contradiction. Now let $\eta > 0$, be a parameter to be chosen. For a fixed q_1, the number of solutions of

$$\frac{p-1}{2} = q_1 q_2$$

where p and q_2 are primes, can be shown by elementary sieve methods to satisfy the bound

$$\frac{Cx}{q_1 (\log \frac{x}{q_1})^2}$$

for some positive constant C and for $q_1 < x$. From this, it can be deduced that the number of $p \in \mathcal{B}'$ such that $q_1 > p^{\frac{1}{2}-\eta}$ cannot exceed

$$\sum_{x^{\frac{1}{2}-\eta} < q_1 < x^{\frac{1}{2}}} \frac{Cx}{4q_1 \log^2 x} \leq C_1 \eta \frac{x}{\log^2 x}$$

for some absolute constant C_1. If η is chosen sufficiently small, then it may be assumed without loss that for a certain positive constant c_1, at least

$$\frac{c_1 x}{\log^2 x}$$

primes $p \in \mathcal{B}',\ p \le x$ are such that $\frac{p-1}{2}$ is either prime or

$$\frac{p-1}{2} = q_1 q_2, \quad p^{\theta} < q_1 < p^{\frac{1}{2}-\eta} < q_2$$

with $\theta > \frac{1}{4}$. This facilitates matters considerably because if a is any natural number such that the order of $a \pmod p$ is $< p^{\frac{1}{2}-\eta}$, then applying the lemma with $r = 1$, and $\Gamma = \langle a \rangle$, the subgroup of $\mathbf{Q}^{\times}$ generated by a, shows that the number of such primes cannot exceed $O(x^{1-2\eta})$. Thus the number of primes p such that 2, 3, or 5 has order (mod p) equal to q_1 or $2q_1$ is at most $O(x^{1-2\eta})$. Therefore, if these primes are eliminated from the set $\mathcal{B}'$, then for x sufficiently large, there remain at least $c_1 x/\log^2 x$ primes such that the order (mod p) of 2, 3, and 5 is q_2 or $2q_2$. But then, 2, 3, 5 together generate a subgroup strictly smaller than $\mathbf{F}_{\mathbf{p}}^{\times}$, which contradicts the fact that they generate $\mathbf{F}_{\mathbf{p}}^{\times}$. Therefore, one of 2, 3, 5 is a primitive root mod p for infinitely many primes p.

By a variation of the method observed earlier, if $\frac{p-1}{2}$ is prime, then Artin's conjecture can be proved for some specific a. Sieve methods are unable to separate those primes p such that $\frac{p-1}{2}$ has *exactly* two prime factors and this constitutes the famous parity problem of sieve theory.

§4. Refinements and concluding remarks

In 1977, Iwaniec [11, 12] discovered an improved form of the remainder term in the linear sieve which leads to the result that there are at least $cx/\log^2 x$ primes $p \le x$ such that $c > 0$ and $\frac{p-1}{2}$ has all prime factors $> p^{\theta}$ with $\theta = \frac{1}{4} - \epsilon$. In 1982, Fouvry and Iwaniec [4] proved a theorem of the form

$$(\S) \qquad \sum_{m < x^{\frac{9}{17}-\epsilon}} \lambda(m)\left(\pi(x,m,1) - \frac{li\ x}{\phi(m)}\right) = O\left(\frac{x}{\log^A x}\right)$$

where $\lambda(m)$ is a certain convolution of arithmetical functions, $\pi(x, m, 1)$ denotes the number of primes $p \le x,\ p \equiv 1 \pmod m$, and ϕ denotes Euler's function. With the improved form of the error term in the linear sieve referred to above, this result produces the desired proportion of primes with $\theta = \frac{9}{34} - \epsilon > \frac{1}{4}$. Such a result gives an affirmative answer to the question of the previous section for the rank $r \geq 3$. Indeed, utilising (§), Rajiv Gupta and the author [5] proved:

Theorem 4. For any distinct primes q, r, s, at least one element in the set

$$\{qs^2, q^3r^2, q^2r, r^3s^2, r^2s, q^2s^3, qr^3, q^3rs^2, rs^3, q^2r^3s, q^3s, qr^2s^3, qrs\}$$

is a primitive root (mod p) for infinitely many primes p.

In his doctoral thesis, Fouvry [5] discussed various results and techniques to extend the range in (§). In particular, he proved that the exponent 9/17 can be improved to 17/32. Finally, in 1983, Bombieri, Friedlander and Iwaniec [3] proved that the exponent can be improved to 4/7. This result enables us to obtain $cx/\log^2 x$ primes $p \leq x$ and $c > 0$, such that $\frac{p-1}{2}$ is a product of three prime factors, each greater than $p^{\frac{2}{7}-\epsilon}$. Heath-Brown observed that those primes p such that $\frac{p-1}{2}$ is a product of precisely three prime factors can be removed from this set, so that many primes p for which $\frac{p-1}{2}$ is a product of <u>*two*</u> large prime factors are obtained. (As was mentioned earlier, such an idea was previously used before by Iwaniec and Chen, in different contexts.) This yields at least $cx/\log^2 x$ primes $p \leq x$ such that $\frac{p-1}{2}$ is either prime or

$$\frac{p-1}{2} = q_1q_2, \quad p^\theta < q_1 < p^{\frac{1}{2}-\eta} < q_2, \quad \theta > \frac{1}{4}.$$

By the method described in section 3, it follows that one of 2,3,5 is a primitive root mod p for infinitely many primes p.

It is conjectured by Halberstam and Elliott that

$$(**) \qquad \sum_{m<x^{1-\epsilon}} |\pi(x,m,1) - \frac{li\ x}{\phi(m)}| = O(\frac{x}{\log^A x}).$$

In fact, if we had (§) with an exponent of $\frac{2}{3} + \epsilon$, instead of $9/17 - \epsilon$, then this would give the result that one of 2 or 3 is a primitive root (mod p) for infinitely many primes p. Any further improvement in (§) does not seem to give any better result. It may be that within a decade (§) will be proved with the exponent $\frac{2}{3} + \epsilon$. (Recently, it was announced [3] that (**) is true with $x^{1-\epsilon}$ replaced by $x^{1/2}(\log x)^{1987}$ and $A < 3$. This is certainly a significant development.) But this still would not resolve Artin's conjecture. Perhaps, some new simple idea is still lurking in the background that would settle the whole conjecture.

The methods under discussion actually give better results than stated. A special case version of the results was adopted for the sake of clarity. It is clear that one can prove along the same lines that if there are three distinct prime numbers q, r, s, then at least one of them is a primitive root (mod p) for infinitely many primes p. It follows that there can be at most two exceptional primes for which Artin's conjecture is false.

Let E be the set of integers, which are not perfect squares, for which Artin's conjecture is false. Let $E(x)$ denote the number of elements of E which are $\leq x$. Srinivasan and the author [16] proved that

$$E(x) = O(\log^6 x)$$

utilising (§). Heath-Brown [9] independently obtained the slightly finer result

$$E(x) = O(\log^2 x)$$

by incorporating the results of [3]. In a similar vein, he proved that there are at most three squarefree integers for which Artin's conjecture is false.

The result (§) with an exponent of $\frac{2}{3} + \epsilon$ would prove $E(x) = O(\log x)$ and one exceptional a for which Artin's conjecture is false. Of course, if there is an exceptional a, then the generalised Riemann hypothesis would be false in view of Hooley's result.

In the winter of 1984, the author had the privelege of visiting the Tata Institute of Fundamental Research in Bombay. At that time, Srinivasan and the author formulated certain plausible conjectures which will lead to the resolution of Artin's conjecture. These conjectures are interesting in their own right. One such conjecture is

Conjecture. Let $f_a(p)$ denote the order of $a \bmod p$. Is it true that

$$\sum_{p \leq x} \frac{1}{f_a(p)} = O(x^{\frac{1}{4}}) \ ?$$

(Here the summation is over prime numbers.)

To see that this conjecture implies Artin's conjecture, let us first note that if $p | a^m - 1$, then $f_a(p) | m$. Hence, for a fixed p, the number of $m < w$ such that $p | a^m - 1$ cannot exceed $w / f_a(p)$ so that the total number of primes $p \leq x$ which divide the product

$$\prod_{m<w} (a^m - 1)$$

is

$$\sum_{p\le x} \frac{w}{f_a(p)} = O(wx^{\frac{1}{4}}),$$

on the basis of the above conjecture. In the notation of the previous section, there is a set $\mathcal{B}'$ of primes $p \le x$ such that

$$p-1 = 2q_1q_2, \quad p^\theta < q_1 < p^{\frac{1}{2}-\eta} < q_2, \quad \theta > \frac{1}{4}$$

and the cardinality of $\mathcal{B}'$ is at least $cx/\log^2 x$, for some constant $c > 0$. If $f_a(p) = q_1, q_2, 2q_1$ or $2q_2$ then p divides

$$\prod_{m<w} (a^m - 1)$$

with $w = x^{1-\theta}$. The conjecture implies that the number of such primes is $O(x^{\frac{5}{4}-\theta})$ which is $o(x/\log^2 x)$ since $\theta > 1/4$. It follows that for almost all $p \in \mathcal{B}'$, a is a primitive root (mod p).

Perhaps it is also true that

$$\sum_{p\le x} \frac{1}{f_a(p)} = O(x^\epsilon),$$

but it is not even known how to prove the simpler statement that

$$\sum_{p\le x} \frac{1}{f_a(p)} = O(x^{\frac{1}{2}-\epsilon}).$$

The bound of $O(x^{\frac{1}{2}})$ is easily deduced from standard sieve arguments.

The lower bound sieve techniques do not give a positive proportion of primes for which a given a is a primitive root (mod p). The full strength of the generalised Riemann hypothesis is not necessary to derive a positive proportion of such primes. Hooley [17] makes this comment at the end of chapter 3 of his beautiful monograph. He states that it suffices to assume that the Dedekind zeta function has no zeroes in the half-plane

$$Re(s) > 1 - \frac{1}{2e}.$$

In October of 1982, while Hooley was visiting McGill University in Montréal, the author asked him about this remark. Hooley explained that if one assumes for a moment that the Artin constant $A(a)$ is 1 (which it isn't), then by choosing z sufficiently small, one can ensure that p does not split completely in any L_q for $q < z$ by using unconditional versions of the Chebotarev density theorem. To treat $M(x; z, x^{1/2}/\log^A x)$, the Bombieri - Vinogradov result that

$$\sum_{m<x^{1/2}/\log^A x} |\pi(x,m,1) - \frac{li\ x}{\phi(m)}| \ll \frac{x}{\log^B x}.$$

is used to show that

$$M(x; x^\theta, x^{1/2}/\log^A x) \leq (1-\epsilon)\frac{x}{\log x}$$

provided $\theta > 1/2e$. The term $M(x; z, x^\theta)$ is then treated assuming the quasi - Riemann hypothesis stated above. But of course, the Artin constant is not 1. When a is squarefree, it is approximately .375. To take this into account, we must make use of the fact that we need only sieve out those primes which split completely in some L_q, $q > z$, but not in any previous field, in particular L_2, and L_3. Then, incorporating a "twisted" version of the Bombieri - Vinogradov Theorem leads to the result stated above.

Problems in mathematics and number theory in particular, largely serve as motivating forces for the discovery and understanding of new concepts. They provide the background for the play of ideas. It may be that we shall not see a resolution of Artin's conjecture in the near future. Nevertheless, it has provided us with a rich interplay of algebraic and analytic number theory and demonstrated a profound relationship between analysis and arithmetic.

REFERENCES

[1] E. Artin, Collected Papers, Addison-Wesley, 1965.

[2] E. Bombieri, Le grand crible dans la théorie analytique des nombres, *Asterisque,* **18** (1974).

[3] E. Bombieri, J.B. Friedlander and H. Iwaniec, Primes in arithmetic progressions to large moduli, *Acta Math.* **156**(1986)203-251.

[4] E. Fouvry and H. Iwaniec, Primes in arithmetic progressions, *Acta Arith.,* **42**(1983)197-218.

[5] E. Fouvry, Autour du théorème de Bombieri - Vinogradov, *Acta Math.,* **152**(1984)219-244.

[6] R. Gupta and M. Ram Murty, A remark on Artin's conjecture, *Inventiones Math.,* **78**(1984)127-130.

[7] R. Gupta and M. Ram Murty, Primitive points on elliptic curves, *Compositio Math.,* **58**(1986)13-44.

[8] R. Gupta, V. Kumar Murty and M. Ram Murty, The Euclidean algorithm for S integers, *CMS Conference Proceedings*, Volume 7, (1985)189-202.

[9] D.R. Heath-Brown, Artin's conjecture for primitive roots, *Quart. J. Math. Oxford,* **(2) 37**(1986)27-38.

[10] C. Hooley, On Artin's conjecture, *J. reine angew. Math.,* **226** (1967) 209-220.

[11] H. Iwaniec, Rosser's sieve, *Acta Arith.,* **36**(1980) 171-202.

[12] H. Iwaniec, A new form of the error term in the linear sieve, *Acta Arith.,* **37**(1980)307-320.

[13] H. Iwaniec, Primes of the type $\phi(x, y) + A$, where ϕ is a quadratic form, *Acta Arith.* **21**(1972)203-224.

[14] S. Lang and H. Trotter, Primitive points on elliptic curves, *Bulletin Amer. Math. Soc.,* **83**(1977)289-292.

[15] C.R. Matthews, Counting points modulo p for some finitely generated subgroups of algebraic groups, *Bulletin London Math. Soc.,* **14**(1982)149-154.

[16] M. Ram Murty and S. Srinivasan, Some remarks on Artin's conjecture, *Canadian Math. Bulletin*, **30** (1987)80-85.

[17] C. Hooley, Applications of sieve methods to the theory of numbers, *Cambridge Tracts in Mathematics*, **70**, Cambridge University Press, (1976).

M. Ram Murty,
Department of Mathematics,
McGill University,
Montréal, Canada