On the Computation of Representations
of Primes as Sums of Four Squares

Michele Elia
Politecnico di Torino, Italy
elia@polito.it

Abstract
Lagrange proved that every positive integer is the sum of four squares of natural numbers. Although Lagrange’s proof is constructive, it is not known whether the relative algorithm produces a four-square representation of any prime or non-prime integer with deterministic polynomial complexity. Limited to prime numbers, it is proved that their representation as the sums of four squares can be obtained with deterministic polynomial complexity. The key to this computational accomplishment is the evaluation of square roots modulo a prime, through Schoof’s method for counting the number of points on elliptic curves over prime fields.

Mathematics Subject Classification: 11T24

Keywords: number theory, computational complexity, quadratic forms.

1 Introduction
Fermat formulated the celebrated four-square theorem in a letter to Carcavi in 1650, [19]. Lagrange proved the theorem in a paper that appeared in Berlin Memoirs in 1772 [22, p.316]. In view of the Euler multiplication formula of 4-tuples, Lagrange had only to show that any prime is the sum of four squares of natural numbers, zero addends permitted. Although the characterization of all integers, in particular primes, as the sum of two, three, or four squares
was completed by Gauss, some computational aspects have only been accomplished recently. In particular, the representation of primes $p = 1 \mod 4$ is computed with deterministic polynomial complexity by a direct application of an algorithm for counting the number of points on elliptic curves over finite fields proposed by Schoof in 1985 [20]. Lagrange’s constructive proof holds for every prime, but apparently it is not known whether the relative algorithm produces a four-square representation with deterministic polynomial complexity. This paper describes a possible interpretation of Lagrange’s procedure that works with deterministic polynomial complexity. The key to the computational accomplishment is Schoof’s algorithm for counting the number of points on elliptic curves over prime fields $\mathbb{F}_p$ of order $p$, and subordinately square roots modulo $p$ [20].

2 Preliminaries

Let $E[\mathbb{F}_p]$ be an elliptic curve of equation $y^2 = x^3 - ax - b$ over a prime field $\mathbb{F}_p$. The number $N_p(a, b)$ of $\mathbb{F}_p$-points on $E[\mathbb{F}_p]$, counting the point at infinity, is

$$N_p(a, b) = 1 + \sum_{x \in \mathbb{F}_p} \left[ 1 + \left( x^3 - ax - b \mid p \right) \right],$$

where $(\cdot \mid p)$ denotes the Legendre symbol. Thus defining the Jacobsthal sum as

$$J_p(a, b) = \sum_{x \in \mathbb{F}_p} \left( x^3 - ax - b \mid p \right),$$

we have $N_p(a, b) = p + 1 + J_p(a, b)$. It is remarkable that the representation $p = u^2 + v^2$ of primes $p = 1 \mod 4$ is obtained counting the number of points on elliptic curves of equation $y^2 = x^3 - ax$ with $a \neq 0$, over $\mathbb{F}_p$. Precisely, the representation $u = \frac{J_p(1, 0)}{2}$, and $v = \frac{J_p(\zeta, 0)}{2}$ with $\zeta$ any quadratic nonresidue modulo $p$, can be proved using elementary methods [1], and both $J_p(1, 0)$ and $J_p(\zeta, 0)$ are evaluated with deterministic polynomial complexity using Schoof’s algorithm. In principle, a similar method could be applied for computing a representation of every prime $p = 3 \mod 4$ as a sum of three or four squares, however the realization of this plan requires different methods for different residues of $p$ modulo 24 [10]. Remarkably, this approach was outlined by Fermat, in a letter to Pascal in 1654 [7, p.187], for proving the four square
Computation of Representations of Primes

507

theorem, although that proof was never finished conclusively. Since Schoof’s counting algorithm plays a central role in the proof of the main theorem, the following well-known facts concerning elliptic curves with complex multiplication are summarized for easy reference, and can be found with more detail in [21, 16, 6].

2.1 Elliptic Curves with Complex Multiplication

The additive group [21] of an elliptic curve $E[\mathbb{K}]$ over a field $\mathbb{K}$ implies that point multiplication by an integer $n$ induces an endomorphism of the curve. An endomorphism is a mapping of the plane into itself that leaves invariant the elliptic curve equation $y^2 = x^3 - ax - b$. However, curve equations may be invariant for other transformations. For example the equation $y^2 = x^3 - x$ is left invariant by the transformation $x = -X$, $y = iY$, whereas the equation $y^2 = x^3 - 1$ is left invariant by the transformation $x = \zeta_3 X$, $y = Y$, $\zeta_3$ being a primitive cubic root of unity. In these cases we say that the curve has complex multiplication. Moreover, the expression $j = \frac{1728}{4a^3 - 27b^2} \in \mathbb{K}$ is called the $j$-invariant of the elliptic curve $E(\mathbb{K})$. In the two examples, we have $j = 1728$, and $j = 0$, respectively. The fact that both $j$-invariant are integers is not a coincidence, for it can be proved that any $j$ of a curve with complex multiplication is an algebraic integer [16, Theorem 6.10, p.402]. Conversely, any $j \in \mathbb{K}$, different from 0 and 1728, is a $j$-invariant of an elliptic curve $E[\mathbb{K}]$ with equation

$$y^2 = x^3 - \frac{3j}{1728 - j} x - \frac{2j}{1728 - j}.$$  (1)

This defining equation is useful to identify an elliptic curve with complex multiplication as specified by Theorem 1. The number of points on $E[\mathbb{F}_p]$, including the point at infinity, is $N_p = p + 1 - a_p$, where the defeat number $a_p$ of $E[\mathbb{F}_p]$ is a Frobenius trace, and satisfies the Hasse bound $|a_p| \leq 2\sqrt{q}$ [5, 1, 21, p.119]. A direct computation of $a_p$ requires the execution of $p$ checks on the elements of $\mathbb{F}_p$, a task which is certainly not feasible in deterministic polynomial-time. However, Schoof’s algorithm computes $a_p$ with deterministic polynomial complexity. A description of the algorithm can be found in the original paper [20], or in [4, 21], and improved versions have been developed by Atkin, and by Elkies [4, 8].

Let $\mathbb{K} = \mathbb{Q}(\sqrt{-D})$ be an imaginary quadratic field, where $D$ is a positive
integer. A Theorem of Lagrange’s tells us that if $-D$ is a quadratic residue modulo $p$, then $p$ is represented by some binary quadratic form of discriminant $-D$ [6]. In particular, Theorem 1 below, borrowed without proof from [6] jointly with a Theorem of Deuring’s [6, Theorem 14.16, p.317], allows us to compute the representation $p = u^2 + Dv^2$ from the $a_p$s of certain elliptic curves with complex multiplication.

**Theorem 1 (Cox: Theorem 9.2, p.180,[6]).** A prime $p$ is represented by the quadratic form $u^2 + Dv^2$ if and only if $(-D \mid p) = 1$ and the minimal polynomial $H_{-4D}(x)$ (the class equation) of a $j$-invariant of an elliptic curve over $K = \mathbb{Q}(\sqrt{-D})$ has a linear factor in $\mathbb{F}_p$.

The degree of $H_{-4D}(x)$ is the class number $h(K)$ of $K$. When $p$ satisfies the conditions of Theorem 1, and $j$ is a root of $H_{-4D}(x)$, there is an integral prime $\pi = u + \sqrt{-D}v$ in $K$, such that $p = \pi \bar{\pi} = u^2 + Dv^2$. Moreover, we have $N_p = |E(\mathbb{F}_p)| = p + 1 - (\pi + \bar{\pi}) = (\pi - 1)(\bar{\pi} - 1)$, thus, the identity $a_p = -(\pi + \bar{\pi})$ holds. In particular, if $K$ has class number $h(K) = 1$, any prime $p$ such that $(-D \mid p) = 1$ is represented by the form $u^2 + Dv^2$, where $u = \frac{a_p}{2}$, $v = \sqrt{\frac{p-u^2}{D}}$. Schoof’s algorithm computes $a_p$ with deterministic complexity not greater than $O((\log p)^5)$ [20].

2.2 Equations $x^2 + y^2 \pm 1 = 0 \text{ mod } p$

In the next section, for complexity evaluation purposes, we need to know that any solution of each of the two equations

$$x^2 + y^2 + 1 = 0 \text{ mod } p \quad \text{and} \quad x^2 + y^2 - 1 = 0 \text{ mod } p,$$

with $p \equiv 3 \text{ mod } 4$, is computed with deterministic polynomial complexity. It is proved by direct counting that both equations have a solution [14, p.135]. Let $(x_0, y_0)$ denote any solution of $x^2 + y^2 \pm 1 = 0 \text{ mod } p$, therefore, in view of a method of Diophantus’ [2] that yields every solution

$$\begin{align*}
x &= \frac{x_0(-1+m^2)-2my_0}{1+m^2} \\
y &= \frac{-2mx_0-m(-1+m^2)}{1+m^2}
\end{align*} \quad (2)$$

by drawing a straight line of slope $m$ through $(x_0, y_0)$, for each sign choice we have exactly $p + 1$ solutions. Remarkably, the whole set of solutions of
\[ x^2 + y^2 \pm 1 = 0 \mod p \] form a multiplicative group \( \mathcal{G} \) of order \( 2(p+1) \), the set of solutions of \( x^2 + y^2 - 1 = 0 \mod p \) forms a subgroup \( \mathcal{H} \) of index 2, and the set of solutions of \( x^2 + y^2 + 1 = 0 \mod p \) is a coset \( \mathcal{L} \) of \( \mathcal{H} \). This property is a consequence of the composition of quadratic forms \( x^2 + y^2 \), and can also be seen using a matrix representation

\[
(x, y) \mapsto A = \begin{bmatrix} x & -y \\ y & x \end{bmatrix} \quad \text{with} \quad \det(A) = x^2 + y^2 = \pm 1.
\]

Moreover, the square of every element of \( \mathcal{L} \) is in \( \mathcal{H} \), conversely the square root of an element of \( \mathcal{H} \) may be either in \( \mathcal{L} \) or again in \( \mathcal{H} \). Thus, considering the factorization \( 2(p+1) = 2^\alpha(2k+1) \), the computation of at most \( 2^\alpha \) iterated square roots of elements of \( \mathcal{H} \) yields an element of \( \mathcal{L} \).

Let \((X_0, Y_0) = ((-1 + m^2)/(1 + m^2), -2m/(1 + m^2))\) be a solution of \( x^2 + y^2 - 1 = 0 \mod p \), with \( m \neq 0 \) to avoid degeneracy. A solution \((u, v)\) of either \( x^2 + y^2 - 1 = 0 \mod p \), or \( x^2 + y^2 + 1 = 0 \mod p \), whose composition (square) yields \((X_0, Y_0)\), is obtained by solving one of two systems

\[
u^2 - v^2 = \frac{(-1 + m^2)}{1 + m^2}, \quad u^2 + v^2 = 1 \mod p, \quad 2uv = \frac{-2m}{1 + m^2} \mod p
\]
\[
u^2 - v^2 = \frac{(-1 + m^2)}{1 + m^2}, \quad u^2 + v^2 = -1 \mod p, \quad 2uv = \frac{-2m}{1 + m^2} \mod p
\]

depending on whether \( 1 + m^2 \) is a QR or a QNR modulo \( p \), hence we have either

\[
(u, v) = ((-\frac{m^2}{1 + m^2})^{\frac{p+1}{4}}, (\frac{1}{1 + m^2})^{\frac{p+1}{4}}) \mod p,
\]
or

\[
(u, v) = ((\frac{-1}{1 + m^2})^{\frac{p+1}{4}}, (\frac{-m^2}{1 + m^2})^{\frac{p+1}{4}}) \mod p.
\]

The process is iterated until a solution of \( x^2 + y^2 + 1 = 0 \mod p \) is obtained. The number of iterations is \( \alpha < 1 + \log_2(p+1) \), thus the complexity is deterministic polynomial.

### 3 Main Theorem

Let every prime \( p \) be written in the form \( 24\ell + \delta \), where \( \delta \) belongs to the set \( \{1, 5, 7, 11, 13, 17, 19, 23\} \); the four partition of the whole set of primes identified by the four sets \( \mathcal{P}_1 = \{1, 5, 13, 17\} \), \( \mathcal{P}_2 = \{7, 19\} \), \( \mathcal{P}_3 = \{11\} \), and \( \mathcal{P}_4 = \{23\} \) will be considered separately in the proof of the main theorem.
Theorem 2 (Lagrange four-square theorem). Any prime can be represented as the sum of at most four squares with deterministic polynomial complexity.

Proof. Case 1: $\delta \in \{1, 5, 13, 17\}$. Every prime in $\mathcal{P}_1$ admits of the representation $p = X^2 + Y^2$, which is computed using the elliptic curve $y^2 = x^3 - x$ with complex multiplier $I$. We have

$$X = \frac{a_p}{2} \quad Y = \sqrt{p - \frac{a_p^2}{4}},$$

where the square root is evaluated with complexity $O(\log^2 p)$, see [12, Problem 33, p.247].

Case 2: $\delta \in \{7, 19\}$. Every prime in $\mathcal{P}_2$ admits of the representation $p = X^2 + 3Y^2$, which is computed using the elliptic curve $y^2 = x^3 - 1$ with complex multiplier $\zeta_3 = \frac{-1 + \sqrt{3}}{2}$. We have

$$X = \frac{a_p}{2} \quad Y = \sqrt{p - \frac{X^2}{3}},$$

thus, a representation $p = x^2 + 3y^2 = x^2 + y^2 + y^2 + y^2$ of $p$ as a sum of four squares is obtained.

Case 3: $\delta \in \{11\}$. Every prime in $\mathcal{P}_3$ admits of the representation $p = X^2 + 2Y^2$, which is computed using the elliptic curve $y^2 = x^3 - 30x - 56$ with complex multiplier $\omega_2 = I\sqrt{2}$. We have

$$X = \frac{a_p}{2} \quad Y = \sqrt{p - \frac{X^2}{2}},$$

thus, a representation $p = x^2 + y^2 + y^2$ of $p$ as a sum of three squares is obtained.

Case 4: $\delta \in \{23\}$. Every prime in $\mathcal{P}_4$ admits of a representation $p = X^2 + Y^2 + W^2 + Z^2$, where necessarily none of the four squares is zero. This representation is obtained with deterministic polynomial complexity through the joint use of the descent method, as in Lagrange’s proof, and a quaternion division, as in the Euclidean algorithm [10]. Specifically, the procedure consists of a sequence of steps of two kinds that we will call "Lagrangian" and "Euclidean" steps. The use of Euclidean steps is motivated by a limit of the
Lagrange descent method, which ultimately gives a quaternion of \( p \) norm, but does not guarantee that computations are completed with deterministic polynomial complexity. Since the number of Lagrangian steps will be shown not to be greater than the number of Euclidean steps, the complexity is dominated by the Euclidean algorithm, which is known to have deterministic polynomial complexity [12].

A quaternion \( q \) is written as a vector \((q_1, q_2, q_3, q_4)\), and \( \bar{q} = (q_1, -q_2, -q_3, -q_4) \) is the \( q \) conjugate. The norm of \( q \) is \( N(q) = qq = q\bar{q} = q_1^2 + q_2^2 + q_3^2 + q_4^2 \), and is a multiplicative function. Integer quaternions have all four entries either integral or semi-integral. An integer quaternion of norm 1 is called unity. The unities form a finite multiplicative group of order 24 consisting of 8 quaternions with three 0-entries and one entry \( \pm 1 \), and 16 quaternions \((\pm\frac{1}{2}, \pm\frac{1}{2}, \pm\frac{1}{2}, \pm\frac{1}{2})\). Two quaternions differing for a unity factor are said to be associated. If we write \( \rho = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}) \), any integral quaternion can be written as \( q = k_0\rho + n \) with integer \( k_0 \), and \( n \) a quaternion of integer entries.

The ”Euclidean step” is an instance of the Euclidean algorithm for integer quaternions [10, p.306]. It is motivated and described in Theorem 372 [10, p.306], which is the key to a practical algorithm performing the Euclidean step. Given two integral quaternions \( q_a \) and \( q_b \) with a common left divisor \( q_d \) and \( N(q_a) < N(q_b) \), two quaternions \( \mathfrak{D} \) and \( q_r \) can effectively be computed such that

\[
q_a = q_b\mathfrak{D} + q_r \tag{3}
\]

with \( N(q_r) < N(q_b) \) and \( q_d \) left divisor of \( q_r \). Hardy-Wright’s clever description of the division process is summarized for the sake of easy reference. Let \( l = (0, l_2, l_3, l_4) \) denote a purely imaginary integral quaternion; if \( q \) is an integral quaternion, and \( m > 1 \) is a positive integer, then an integral quaternion \( \mathfrak{d} = l_0\rho + l \) such that \( N(q - m\mathfrak{d}) < m^2 \) is found by imposing that \(|q_1 - ml_0| < \frac{1}{2}m \), \(|q_1 + 2q_2 - m(l_0 + 2l_2)| < m \), \(|q_1 + 2q_3 - m(l_0 + 2l_3)| < m \), \(|q_1 + 2q_4 - m(l_0 + 2l_4)| < m \). To find \( \mathfrak{D} \) and \( q_r \) in (3), take \( q = q_bq_a \) and \( m = q_b, q_b \), then

\[
q - m\mathfrak{d} = q_bq_a - q_bq_b\mathfrak{d} = q_b[q_a - q_b\mathfrak{d}]
\]

\[
N(q_bq_b) = m^2 > N(q - m\mathfrak{d}) = N(q_b)N(q_a - q_b\mathfrak{d})
\]

Thus, setting \( \mathfrak{D} = \mathfrak{d} \), and \( q_r = q_a - q_b\mathfrak{d} \), equation (3) is satisfied, with \( q_d \) a left divisor of \( q_r \), as required.
The "Lagrangian step" is an instance of the method that Lagrange used to prove the four-square theorem [10]. A Lagrangian step is based on the observation that, given a quaternion \( z = (Z_1, Z_2, Z_3, Z_4) \) with norm \( N(z) = mp \) divisible by \( m \), taking the integer \( \mu_i \) such that \( y_i = Z_i + \mu_i m \) satisfies \( |y_i| < m/2 \), the norm of the quaternion \( \eta = (y_1, y_2, y_3, y_4) \) is again divisible by \( m \), that is \( mm_1 = N(\eta) \), with \( m_1 < m \). This is easily seen considering the quaternion \( \mu = (\mu_1, \mu_2, \mu_3, \mu_4) \), whereby we have

\[
N(\eta) = (z + m\mu)(\overline{z + m\mu}) = N(z) + m(3\overline{\mu} + \mu\overline{z}) + m^2N(\mu) = mm_1,
\]

because \( N(z) \) is divisible by \( m \). For the same reason, each component of the quaternions

\[
w_L = \overline{z\eta} = z(\overline{3 + m\mu}) = N(z) + m3\overline{\mu} = m(p + 3\overline{\mu})
\]

\[
w_R = \overline{\eta z} = (\overline{3 + m\mu})z = N(z) + m\overline{\mu}3 = m(p + \overline{\mu}3)
\]

is divisible by \( m \). For the \( w \) norm, we have \( N(w) = N(z)N(\eta) = m_1m^2p \), thus \( \frac{w}{m} \) is an integral quaternion with norm \( m_1p \). Since \( z \) is either left or right divisible by a quaternion of norm \( p \), then either \( w_l = w_L/m \) or \( w_r = w_R/m \) is left or right divisible by a quaternion of norm \( p \). Which is the case is established by performing a Euclidean step on both products: of \( w_l \) and \( w_r \), the quaternion sharing divisibility with \( z \) yields a remainder with norm divisible by \( p \).

The algorithm for computing the quaternion \( p \) starts with a solution \((a_1, b_1)\) of \( x^2 + y^2 + 1 = 0 \mod p \), computed with deterministic polynomial complexity as shown in Section 2. Therefore, quaternion \( z_0 = (1, 0, a_1, b_1) \) has norm \( N(z_0) = k_5p = 1 + a_1^2 + b_1^2 \), a multiple of \( p \), with \( k_5 \leq 2p \).

First perform a Lagrangian step to produce \( w_0 \) from \( z_0 \), where both \( z_0 \) and \( w_0 \) are either left or right multiples of \( p \). Then perform a Euclidean step on the pair \( z_0 \) and \( w_0 \) to produce \( w_1 \). Iterate the Euclidean step on the sequence of pairs \( w_{n-1} \) and \( w_n \), unless the remainder \( w_{n+1} \) is zero and \( N(w_n) \neq p \). In this event, perform a Lagrangian step on \( w_n \), then continue with Euclidean steps. At the end, a quaternion with norm \( p \) is found. It is evident that the complexity is dominated by the complexity of the Euclidean algorithm, which is known to have deterministic polynomial complexity. This completes the proof.
References


Received: June 25, 2006