ARTICLES

The Group of Rational Points on the Unit Circle

LIN TAN
West Chester University
West Chester, PA 19383

Introduction

Recently, Fermat's Last Theorem was proved. A long chain of arguments, based on many mathematicians' deep work, culminated in Wiles' last and decisive step ([**28**, **25**]).

FERMAT'S LAST THEOREM. *Let* $n, a, b, c \in \mathbb{Z}$ *with* $n > 2$. *If* $a^n + b^n = c^n$, *then* $abc = 0$.

There have been several attempts to present the basic idea of this marvelous proof to a wider mathematical audience ([**2**, **7**, **8**]). Another recent paper [**3**] provides more details.

Wiles' proof is based on the theory of elliptic curves, i.e., curves defined by cubic equations. A big part of this theory is devoted to understanding the "rational points" (points whose coordinates are rational numbers) on these curves. The set of rational points on an elliptic curve has a natural group structure, which will be described very briefly later. It is often very difficult, however, to find all the rational points on an elliptic curve.

In this paper we take a much easier and more familiar example—the unit circle—and show how to compute the group structure of its rational points. Next, some applications are given. Finally, for comparison, we give a brief summary of known results on the group structure for rational points on elliptic curves.

Rational Points on the Unit Circle

Let C be the unit circle in the real plane, defined by $x^2 + y^2 = 1$. The *rational points* on C are those for which both coordinates are in $\mathbb{Q}$. For example, $\left(\frac{3}{5}, \frac{4}{5}\right)$, $\left(-\frac{5}{13}, \frac{12}{13}\right)$ and $(0, 1)$ are rational points, while $\left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$ is not. We denote the set of rational points on C by $C(\mathbb{Q})$. A rational point $\left(\frac{a}{c}, \frac{b}{c}\right)$ on C corresponds to an *integer* solution to $X^2 + Y^2 = Z^2$, with $X = a$, $Y = b$, and $Z = c$. (More generally, a rational point on the curve $x^n + y^n = 1$ corresponds to an integer solution to $X^n + Y^n = Z^n$.)

C is an abelian group under the "angle addition" $\oplus$, defined by

$$(x_1, y_1) \oplus (x_2, y_2) = (x_1 x_2 - y_1 y_2, x_1 y_2 + x_2 y_1) \qquad (1)$$

for $(x_1, y_1), (x_2, y_2) \in C$. The identity element is $(1, 0)$, and the inverse of (x, y) is $(x, -y)$. Note that (1) is merely the familiar "addition formula" in trigonometry using the correspondence $\theta \mapsto (x, y) = (\cos\theta, \sin\theta)$; (1) is also the usual formula for the multiplication in the field of complex numbers.

It is clear from (1) that $C(\mathbb{Q})$ is a subgroup of C. This raises a natural question:

What is the group structure of $C(\mathbb{Q})$?

We begin our search for an answer with some elementary facts from number theory and algebra (see, e.g., **[12]**).

Every *Pythagorean triple* (a, b, c) (i.e., a triple of integers a, b, c with $c \neq 0$, satisfying $a^2 + b^2 = c^2$) corresponds to the rational point $\left(\frac{a}{c}, \frac{b}{c}\right)$ on $C(\mathbb{Q})$. Two Pythagorean triples (a, b, c) and (a', b', c') correspond to the same point on $C(\mathbb{Q})$ if and only if $(a, b, c) = r(a', b', c')$ for some $r \in \mathbb{Q} \setminus \{0\}$. Therefore, if (a, b, c) is *primitive* (i.e., if $c > 0$ and the greatest common divisor of a, b, and c is 1), then every Pythagorean triple corresponding to $\left(\frac{a}{c}, \frac{b}{c}\right)$ must have the form (na, nb, nc) for some nonzero integer n.

From elementary number theory we know that the parametrization $(m^2 - n^2, 2mn, m^2 + n^2)$, with $m, n \in \mathbb{Z}$, not both 0, gives all Pythagorean triples (a, b, c) with $c > 0$. Those m and n that satisfy $(m, n) = 1$ and $m - n \equiv 1 (\mathrm{mod}\, 2)$ produce all the primitive triples. (See, e.g., **[23**, p. 13].) It is perhaps illuminating to have a geometric interpretation of this parametrization. The expressions $m^2 - n^2$ and $2mn$ remind us of the double-angle formulas for cosine and sine. They come from the well-known "rational parametrization" $\rho\colon \mathbb{R} \to C$ of the unit circle, defined by

$$\rho(t) = \left(\frac{1 - t^2}{1 + t^2}, \frac{2t}{1 + t^2}\right). \tag{2}$$

(See FIGURE 1; for more details, see **[23**, p. 11].) In particular,

$$\rho\left(\frac{n}{m}\right) = \left(\frac{m^2 - n^2}{m^2 + n^2}, \frac{2mn}{m^2 + n^2}\right).$$

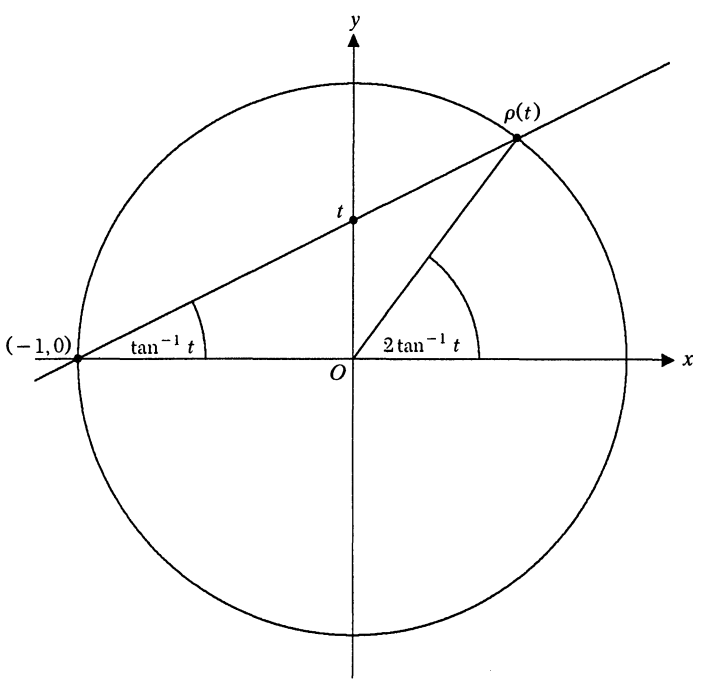

FIGURE 1

As the lines through $(-1,0)$ sweep through the rational slopes, their intersections with C, other than $(-1,0)$, sweep through $C(\mathbb{Q})$. Notice too that $\frac{1-t^2}{1+t^2} = \cos(2\tan^{-1}(t))$, and that $\rho|_{\mathbb{Q}}\colon \mathbb{Q} \to C(\mathbb{Q})\setminus\{(-1,0)\}$ is onto, since $\rho(\frac{v}{1+u}) = (u,v)$.

Next we consider the ring $\mathbb{Z}[i] = \{m+ni \mid m,n \in \mathbb{Z}\}$ of Gaussian integers. For $m+ni \in \mathbb{Z}[i]$ with $m \neq 0$, the line connecting $m+ni$ and O intersects the vertical line $x=1$ at $(1, \frac{n}{m})$ (see FIGURE 2). Combining the above observations (by superimposing FIGURE 1 and FIGURE 2, with the line $x=1$ in FIGURE 2 aligned with the y-axis in FIGURE 1), we get a map $f\colon \mathbb{Z}[i]\setminus\{0\} \to C(\mathbb{Q})$ defined by

$$f(m+ni) = \left(\frac{m^2-n^2}{m^2+n^2}, \frac{2mn}{m^2+n^2}\right); \tag{3}$$

$f(ni)$ is defined to be $(-1,0)$. It is clear from (3) that any Gaussian integer on the line of slope n/m through O yields the same value for $f(m+ni)$.

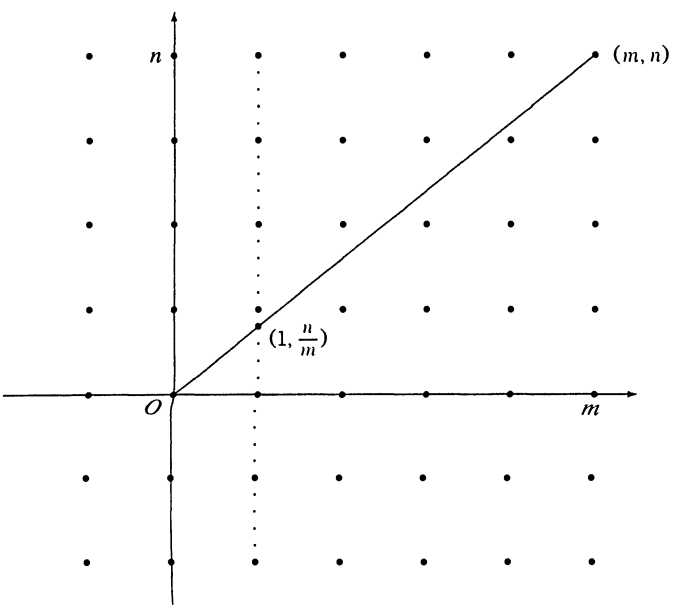

FIGURE 2

The keys to our approach are the simple facts that f is an onto map from $\mathbb{Z}[i]\setminus\{0\}$ (which is almost a group under multiplication but for the lack of inverses, i.e., a semigroup) to the group $C(\mathbb{Q})$, and that f preserves multiplication. This, again, is the double-angle formulas (see FIGURE 1): If $m+ni$ corresponds to angle $\theta = \tan^{-1}(n/m)$, then $f(m+ni) \in C(\mathbb{Q})$ corresponds to 2θ. The following properties are also clear:

$$f(m-ni) \oplus f(m+ni) = (1,0), \tag{4}$$

$$f(m+ni) = (1,0) \text{ (resp. } (-1,0)) \quad \text{if and only if } n = 0 \text{ (resp. } m = 0). \tag{5}$$

The homomorphism $f\colon \mathbb{Z}[i]\setminus\{0\} \to C(\mathbb{Q})$ will enable us to explore $C(\mathbb{Q})$, using the well-known properties of $\mathbb{Z}[i]$. The Lemma below shows the advantage of doing so.

The Group Structure of $C(\mathbb{Q})$

Now for $k = 1, 2, 3$, let $(c_k, s_k) \in C(\mathbb{Q}), (c_k, s_k) = f(m_k + n_k i)$. Then $(c_1, s_1) = (c_2, s_2) \oplus (c_3, s_3)$ if and only if $(m_1 + n_1 i)l' = (m_2 + n_2 i)(m_3 + n_3 i)l$ for some $l, l' \in \mathbb{Z} \setminus \{0\}$. Thus, the elements $f(m + ni) \in C(\mathbb{Q})$ with $m + ni$ irreducible in $\mathbb{Z}[i]$ suffice to generate $C(\mathbb{Q})$.

Let us recall some basic properties of $\mathbb{Z}[i]$. (For more details, see **[12]**.) First, $\mathbb{Z}[i]$ is a unique factorization domain (in fact, a Euclidean domain) with units ± 1 and $\pm i$. Second, the norm N in $\mathbb{Z}[i]$, defined by

$$N(m + ni) = (m + ni)(m - ni) = m^2 + n^2,$$

is a homomorphism from the multiplicative semigroup $\mathbb{Z}[i] \setminus \{0\}$ to the multiplicative semigroup $\mathbb{N}$ of natural numbers. Third, $\mathbb{Z}[i]$ has three types of irreducible elements:

(i) For a rational prime (i.e., a prime number in $\mathbb{Z}$) $p \equiv 1 (\mathrm{mod}\, 4)$, p can be written, uniquely up to sign and order, as the sum $m^2 + n^2$ of two squares. For such m and n, $m + ni$ is irreducible in $\mathbb{Z}[i]$ of norm p.
(ii) For a rational prime $p \equiv 3 (\mathrm{mod}\, 4)$, p remains irreducible in $\mathbb{Z}[i]$ of norm p^2.
(iii) $1 + i$ and $1 - i$ are irreducible, each with norm $(1 + i)(1 - i) = 2$.

For irreducible elements p of type (ii), $f(p) = (1, 0)$, the identity in $C(\mathbb{Q})$. For irreducible elements $1 \pm i$ of type (iii), $f(1 \pm i) = (0, \pm 1)$, which have order 4 in $C(\mathbb{Q})$. For each $p \equiv 1 (\mathrm{mod}\, 4)$, we pick, for definiteness, integers m_p and n_p so that $p = m_p^2 + n_p^2$, $m_p > n_p > 0$. Then

$$\left\{ \left(\frac{m_p^2 - n_p^2}{m_p^2 + n_p^2}, \frac{2m_p n_p}{m_p^2 + n_p^2} \right) \right\}_{p \equiv (\mathrm{mod}\, 4)} \cup \{(0, 1)\}$$

is a set of generators for $C(\mathbb{Q})$. Now we show that no relation exists among these generators, other than that $(0, 1)$ has order 4. First we handle generators of type (i):

LEMMA. *There is no non-trivial relation in*

$$\left\{ \left(\frac{m_p^2 - n_p^2}{m_p^2 + n_p^2}, \frac{2m_p n_p}{m_p^2 + n_p^2} \right) \right\}_{p \equiv 1 (\mathrm{mod}\, 4)}$$

Proof. Let $p_1, \ldots, p_k$ be rational primes, all congruent to 1 modulo 4. Let $m_j = m_{p_j}$, $n_j = n_{p_j}$, and suppose that

$$a_1 f(m_1 + n_1 i) \oplus \cdots \oplus a_k f(m_k + n_k i) = (1, 0) \tag{6}$$

for integers $a_1, \ldots, a_k$. Then

$$(m_1 + n_1 i)^{a_1} \ldots (m_k + n_i i)^{a_k} l = l' \tag{7}$$

for some $l, l' \in \mathbb{Z} \setminus \{0\}$. Let $l = q_1 q_2 \ldots$ and $l' = q_1' q_2' \ldots$ with q_j and q_j' rational primes. We may assume, by unique factorization, that $l = 1$. So

$$(m_1 + n_1 i)^{a_1} \ldots (m_k + n_k i)^{a_k} = q_1' q_2' \ldots . \tag{8}$$

By unique factorization again, each q_j' is associated with a product of two factors, each the complex conjugate of the other, of the left side of (8). The images of such a pair of

factors under f are inverses of each other by (4) and (5). Unless all $a_j = 0$ this leads to a contradiction, since all m_j and n_j are positive. Thus (6) is a trivial relation. □

If, on the other hand, a relation involves $(0,1)$ $(=f(1+i))$, then the "quadruple" of this relation has the form (6), and is therefore trivial by the Lemma. It follows that the original relation is a consequence of the relation $4\cdot(0,1) = (1,0)$.

To finish the analysis of $C(\mathbb{Q})$ we invoke a special case of Dirichlet's Theorem on primes in arithmetic progression. This case was proved by Euler **[5]** in 1775:

There are infinitely many primes congruent to 1 *modulo* 4.

Combining the results above, we get the structure theorem for $C(\mathbb{Q})$.

THEOREM 1. *The abelian group $C(\mathbb{Q})$ is the direct sum of infinitely many cyclic subgroups:*

$$C(\mathbb{Q}) \cong C_2 \oplus \left(\bigoplus_{p \equiv 1 \pmod 4} C_p \right),$$

where C_2 is generated by $(0,1)$ (an element of order 4), and C_p is the infinite cyclic group generated by

$$\left(\frac{m_p^2 - n_p^2}{m_p^2 + n_p^2}, \frac{2m_p n_p}{m_p^2 + n_p^2} \right),$$

with m_p, n_p being the unique solution to $m_p^2 + n_p^2 = p$, $m_p > n_p > 0$.

Examples and Corollaries

We illustrate the significance of Theorem 1 with some examples and corollaries.

Example A. C_5, C_{13}, C_{17}, C_{29}, C_{37}, and C_{41} are generated, respectively, by $\left(\frac{3}{5}, \frac{4}{5}\right) = f(2+i)$, $\left(\frac{5}{13}, \frac{12}{13}\right) = f(3+2i)$, $\left(\frac{15}{17}, \frac{8}{18}\right) = f(4+i)$, $\left(\frac{21}{29}, \frac{20}{29}\right) = f(5+2i)$, $\left(\frac{35}{37}, \frac{12}{37}\right) = f(6+i)$, and $\left(\frac{9}{41}, \frac{40}{41}\right) = f(5+4i)$.

Example B. $\left(\frac{-76}{1445}, \frac{1443}{1445}\right) = (0,1) \oplus (-1)\cdot\left(\frac{3}{5}, \frac{4}{5}\right) \oplus 2\cdot\left(\frac{15}{17}, \frac{8}{17}\right)$. This decomposition can be computed as follows. From the prime factorization $1445 = 5\cdot 17^2$, we have $\left(\frac{-76}{1445}, \frac{1443}{1445}\right) = n\cdot(0,1) \oplus (\pm 1)\cdot\left(\frac{3}{5}, \frac{4}{5}\right) \oplus (\pm 2)\cdot\left(\frac{15}{17}, \frac{8}{17}\right)$. Now we choose the coefficient of $\left(\frac{3}{5}, \frac{4}{5}\right)$ to be 1 (resp. -1) if the denominators are 17^2 in $\left(\frac{-76}{1445}, \frac{1443}{1445}\right) \oplus (-1)\cdot\left(\frac{3}{5}, \frac{4}{5}\right)$ (resp. in $\left(\frac{-76}{1445}, \frac{1443}{1445}\right) \oplus (1)\cdot\left(\frac{3}{5}, \frac{4}{5}\right)$), when reduced. The coefficient ± 2 is chosen similarly. Finally, the coefficient of $(0,1)$ is chosen to get the right signs and the right order of the two coordinates.

The following corollaries are straightforward consequences of Theorem 1.

COROLLARY 1. *Let α and β be real numbers, and suppose $P_\alpha = (\cos\alpha, \sin\alpha)$ and $P_\beta = (\cos\beta, \sin\beta)$ are in $C(\mathbb{Q})$, and that $\frac{\alpha}{\beta} \in \mathbb{Q}$. Then there exist $r, s \in \mathbb{Z}$, $P_\gamma \in C(\mathbb{Q})$, and $c_\alpha, c_\beta \in C_2$, such that $P_\alpha = rP_\gamma \oplus c_\alpha$ and $P_\beta = sP_\gamma \oplus c_\beta$. In particular, if $P_\alpha \in C(\mathbb{Q})$ and α is a rational multiple of π, then $P_\alpha \in C_2$.* **([26])**

Proof. Let $\alpha/\beta = r/s \in \mathbb{Q}$, with r and s relatively prime. Then $sP_\alpha = rP_\beta$, because in $C(\mathbb{Q})$ we have $n\cdot(\cos\theta, \sin\theta) = (\cos(n\theta), \sin(n\theta))$. So the first assertion follows from Theorem 1 by comparing the C_p-components of P_α and P_β: Each C_p-component of P_α (resp. P_β) is "divisible" by r (resp. by s), and we can construct P_γ by defining its C_p-component as $\frac{1}{r}\cdot\{$the C_p-component of $P_\alpha\}$, which is equal to $\frac{1}{s}\cdot\{$the

C_p-component cf P_β}. (Because C_2 is finite, we cannot compare the C_2-components here, and so must include the "c_α" and "c_β" terms.)

The second assertion follows from the first by taking $\beta = \pi$. □

COROLLARY 2. *Let* $P_\alpha = (\cos\alpha, \sin\alpha)$ *and* $P_\beta = (\cos\beta, \sin\beta)$ *be in* $C(\mathbb{Q})$ *and suppose that* $\alpha - \beta$ *is a rational multiple of* π. *Then* $\alpha = \beta + \frac{k\pi}{2}$ *for some* $k \in \mathbb{Z}$.

Proof. This follows directly by applying Corollary 1 to $\alpha - \beta$. □

COROLLARY 3. *On the "square geoboard" (the lattice* $\mathbb{Z} \times \mathbb{Z}$ *in* $\mathbb{R}^2$*), the only angles of rational measure (in degrees) that can be formed by three lattice points are integer multiples of* 45°. (*This was conjectured in* 1921 **[26]** *and first proved in* 1945 **[20]**.)

Proof. The claim follows directly from Corollary 2 and the double-angle correspondence (3) between the geoboard $\mathbb{Z}[i]$ and $C(\mathbb{Q})$. □

On a geoboard, therefore, one cannot "construct" angles measuring 30°, 60°, 22.5°, 36° and 20°. Thus, Corollary 3 solves a "teaser" in the premier issue of the *Math Horizons* **[16]**.

We mention in passing that although for $r \in \mathbb{Q}$, $\cos(r\pi)$ and $\sin(r\pi)$ are both irrational (the only possible exception is $r = \frac{n}{6}$ for $n \in \mathbb{Z}$), these values are known to be algebraic **[24, 14, 9]**. By contrast, the values $\cos r$ and $\sin r$, for $r \in \mathbb{Q} \setminus \{0\}$, are all transcendental **[15, 27]**. (The first result follows easily from De Moivre's Theorem; the second is much harder.)

Remarks. (a) The unique factorization property of $\mathbb{Z}[i]$ gives an algebraic interpretation of the double-angle correspondence f in (3). In fact, from $a^2 + b^2 = c^2$, we get

$$(a + bi)(a - bi) = c^2. \tag{9}$$

Writing $a + bi$, $a - bi$ and c as products of irreducible elements in $\mathbb{Z}[i]$ shows that $a + bi$ is a square in $\mathbb{Z}[i]$, i.e., $a + bi = (m + ni)^2 = (m^2 - n^2) + (2mn)i$ for some $m, n \in \mathbb{Z}$. Hence $\tan^{-1}(\frac{b}{a}) = 2\tan^{-1}(\frac{n}{m})$. For instance, $3 + 4i = (2 + i)^2$. Thus the simple factorization (9) links the additive and multiplicative properties of numbers, and leads to the starting point of Kummer's work on Fermat's Last Theorem. (Cf. **[1**, Ch. 3] and **[4**, Ch. 5].)

(b) $C(\mathbb{Q})$ can be identified with the (multiplicative) subgroup

$$\mathbb{Q}(i)_1 = \left\{ \frac{a}{c} + \frac{b}{c}i \in \mathbb{Q}(i) \,\middle|\, \left(\frac{a}{c}\right)^2 + \left(\frac{b}{c}\right)^2 = 1 \right\}$$

of $\mathbb{Q}(i)$ consisting of the norm 1 elements, where $\mathbb{Q}(i)$ is realized as the field of fractions of $\mathbb{Z}[i]$. Then for $\frac{a}{c} + \frac{b}{c}i \in \mathbb{Q}(i), (a, b, c) = 1$, and the prime decomposition in $\mathbb{Z}$ of $c = p_1 p_2 \ldots q_1 q_2 \ldots$ where $p_i \equiv 1 (\mathrm{mod}\, 4)$ and $q_j \equiv 3 (\mathrm{mod}\, 4)$, we have, by unique factorization, that

$$\frac{a + bi}{c} = \frac{(r_1 + s_1 i)(r_2 + s_2 i)\ldots}{p_1 p_2 \ldots}.$$

As remarked above,

$$\frac{a + bi}{c} = \frac{(m_1 + n_1 i)^2 (m_2 + n_2 i)^2 \ldots}{p_1 p_2 \ldots} = \frac{m_1 + n_1 i}{m_1 - n_1 i} \frac{m_2 + n_2 i}{m_2 - n_2 i} \ldots.$$

By multiplying both $m_j + n_j i$ and $m_j - n_j i$ by an element in C_2, we can assume that each $m_j > |n_j| > 0$. Thus, we have the generation part of Theorem 1. The argument

using $\mathbb{Q}(i)_1$ on the relation part is similar. The map $\mathbb{Q}(i)^* \to \mathbb{Q}(i)_1$ defined by (3) is the map in Hilbert's Satz 90 ([**10**]), applied to the current special case.

Rational Points on the Hyperbola

Our analysis of the rational points on the unit circle $x^2 + y^2 = 1$ can be modified to study the rational points on the hyperbola H defined by $x^2 - y^2 = 1$. In this case, we have a rational parametrization ρ: $\mathbb{R} \setminus \{\pm 1\} \to H$ defined by

$$\rho(t) = \left(\frac{1+t^2}{1-t^2}, \frac{2t}{1-t^2} \right), \tag{10}$$

obtained by intersecting the line of slope t, through $(-1, 0)$, with H. We define an addition $\oplus$ on H by

$$(x_1, y_1) \oplus (x_2, y_2) = (x_1 x_2 + y_1 y_2, x_1 y_2 + x_2 y_1) \tag{11}$$

for $(x_1, y_1), (x_2, y_2) \in H$; the identity element is $(1, 0)$; and the inverse of (x, y) is $(x, -y)$. It is clear from (11) that the rational points $H(\mathbb{Q})$ form a subgroup of H. In the current case, $\mathbb{Z}[i]$ is replaced by $\mathbb{Z}[\varepsilon] \cong \mathbb{Z}[X]/(X^2 - 1)$. Note that $\mathbb{Z}[\varepsilon]$ is not an integral domain, since $(\varepsilon - 1)(\varepsilon + 1) = 0$. To remedy this, we take the subset $R = \{m + n\varepsilon \mid m > n\}$ of $\mathbb{Z}[\varepsilon]$. Then R is closed under multiplication and has the unique factorization property. Since H consists of two branches $H_1 = \{(x, y) \in H \mid x > 0\}$ and $H_2 = \{(x, y) \in H \mid x < 0\}$, the subgroup $H(\mathbb{Q})$ has a corresponding decomposition

$$H(\mathbb{Q}) = H_1(\mathbb{Q}) \cup H_2(\mathbb{Q}). \tag{12}$$

It can be readily verified that H_1 is a subgroup of H and $H_1(\mathbb{Q})$ is a subgroup of $H(\mathbb{Q})$. Thus, (12) is the coset decomposition of $H(\mathbb{Q})$ with respect to $H_1(\mathbb{Q})$, with $H_2(\mathbb{Q}) = (-1, 0) \oplus H_1(\mathbb{Q})$ and $H(\mathbb{Q}) = H' \oplus H_1(\mathbb{Q})$, where $H' = \{(\pm 1, 0)\}$. The map f: $R \to H_1(\mathbb{Q})$ defined by

$$f(m + n\varepsilon) = \left(\frac{m^2 + n^2}{m^2 - n^2}, \frac{2mn}{m^2 - n^2} \right) \tag{13}$$

is an onto *homomorphism*. The following theorem, whose proof is left to the reader, gives a complete description of the group structure of $H(\mathbb{Q})$.

THEOREM 2. *The abelian group $H(\mathbb{Q})$ is the direct sum of infinitely many cyclic subgroups*

$$H(\mathbb{Q}) = H' \oplus \left(\bigoplus_{p \text{ prime}} H_p \right),$$

where, for each prime p, H_p is the infinite cyclic group generated by $\left(\frac{p^2+1}{2p}, \frac{p^2-1}{2p}\right)$.

Note that for $p > 2$, $\left(\frac{p^2+1}{2p}, \frac{p^2-1}{2p}\right) = f\left(\frac{p+1}{2} + \frac{p-1}{2}\varepsilon\right)$, and $p = 2$, $\left(\frac{5}{4}, \frac{3}{4}\right) = f(3 + \varepsilon)$.

Example C. $\left(-\frac{409}{120}, -\frac{391}{120}\right) = (-1, 0) \oplus 2 \cdot \left(\frac{5}{4}, \frac{3}{4}\right) \oplus (-1) \cdot \left(\frac{5}{3}, \frac{4}{3}\right) \oplus \left(\frac{13}{5}, \frac{12}{5}\right)$.

General Conic Sections

The curves C and H studied above are but two special cases of the conic sections (defined by integral polynomials of degree 2). The structure of the rational points on a general conic section may be more complicated. Indeed, some conic sections have no rational points at all. For example, the curve $ax^2 + by^2 = c$, where $a, b, c \in \mathbb{Z}$, not all of the same sign, and abc is square-free, has rational points if and only if $-bc$, $-ca$, $-ab$ are quadratic residues modulo $|a|, |b|, |c|$ respectively. This is Legendre's Criterion **[13]**. Analyzing the general case requires the arithmetic theory of quadratic forms. Once we know that a single rational point P exists, we can get all the rational points by sweeping the secant lines through P of rational slope; this gives a rational parametrization of the curve. (This is a special case of the Hilbert-Hurwitz Theorem **[11]** on the rational parametrization of curves of "genus 0".) In the cases of the unit circle and the hyperbola, we took P to be $(-1, 0)$. These cases are exceptionally interesting because of their natural group structure that can be described in a simple way.

Conclusion

We conclude with a brief discussion of the situation for cubic curves.

Let E be a nonsingular cubic curve defined over $\mathbb{Q}$, and let $E(\mathbb{Q})$ be the set of rational points on E. We may assume that E is defined by an equation in Weierstrass form: $y^2 = x^3 + ax^2 + bx + c$. Now call the unique point at infinity $\mathscr{O}$; it will be the identity element of our group. The "negative" of a point (x, y) is $(x, -y)$. In general, a line intersects E in three points, since a general cubic equation has three roots. We define the addition on E and $E(\mathbb{Q})$ by decreeing that collinear points on E "add up" to $\mathscr{O}$. The reader is referred to **[23**, Ch. I] for detailed discussion. The picture on the cover of **[23]** illustrates the definition of the addition on E. On such a curve E, we have these group-theoretic properties:

(i) *E is an abelian group.* (This has been known since Euler.)
(ii) *$E(\mathbb{Q})$ is a finitely generated abelian group.* (This is the celebrated Mordell's Theorem **[10]**, first conjectured by Poincaré **[21]** in 1901.)
(iii) *The torsion subgroup of $E(\mathbb{Q})$ can have only the following forms, each of which is realizable: A cyclic group of order n with $1 \le n \le 10$ or $n = 12$, or the product of a cyclic group of order 2 and a cyclic group of order $2n$ with $1 \le n \le 4$.* (This is Mazur's Theorem **[17, 18]**.)
(iv) *E has only finitely many points with integer coordinates.* (This is Siegel's Theorem **[22]**.)

Note that in the unit circle case, the torsion subgroup coincides with the set of integer points on the curve; this is not the case for cubic curves. The difference as regards finite generation between degree 2 and degree 3 (or higher) should not be surprising: Fermat's equation $X^n + Y^n = Z^n$ has infinitely many solutions if $n = 2$ and no nontrivial solution if $n \ge 3$. The corresponding curves have different topology as well—they have "genus" 0 and 1 (or $n - 2$) respectively.

We hope that our elementary discussion of $C(\mathbb{Q})$ will stimulate the reader's interest in the deeper theory of quadratic forms and cubic curves, and in the discoveries on these subjects by Fermat, Euler, Lagrange, Legendre, Minkowski, Mordell, Hasse, Weil, Siegel, Mazur, Wiles, and others.

Acknowledgement. We thank Sebastian Koh and Joe Silverman for sharing insights with us, the referees and the editor for many valuable comments and suggestions, and Shiv Gupta and Cliff Johnston for technical help.

REFERENCES

1. Z. Borevich and I. Shafarevich, *Number Theory*, Academic Press, New York, 1966.
2. D. Cox, Introduction to Fermat's Last Theorem, *Amer. Math. Monthly* 101(1994), 3–14.
3. H. Darmon, F. Diamond and R. Taylor, Fermat's Last Theorem, in: *Current Developments in Mathematics, 1995*, International Press, Cambridge, MA, 1995, pp. 1–107.
4. H. Edwards, *Fermat's Last Theorem*, Springer-Verlag, New York, 1977.
5. L. Euler, De summa seriei ex numeris primis formitatae ubi numeri primi formae $4n - 1$ habent signum positivum formae autem, $4n + 1$ signum negativum, *Opera Omnia*, Teubner, Leipzig, I, 4, pp. 146–162.
6. G. Faltings, Endlichkeitssätze für abelsche Varietäten über Zahlköopern, *Invent. Math.* 73 (1983), 349–366. (English translation: *Arithmetic Geometry*, G. Cornell and J. Silverman (*ed.*), pp. 9–27. Springer-Verlag, New York, 1986.)
7. G. Faltings, The proof of Fermat's Last Theorem by R. Taylor and A. Wiles, *Notices of the Amer. Math. Soc.* 42(1995), 743–746.
8. F. Gouvêa, A marvelous proof, *Amer. Math. Monthly* 101(1994), 203–222.
9. R. Hamming, The transcendental character of cox x, *Amer. Math. Monthly* 52(1945), 336–337.
10. D. Hilbert, Die Theorie der algebraischer Zahlköper, *Jahesber. Deutsch. Math. Verein.* 4(1897), 175–546.
11. D. Hilbert and A. Hurwitz, Über die diophantischen Gleichungen vom Geschlecht Null, *Acta Math.* 14(1890), 217–224.
12. K. Ireland and M. Rosen, *A Classical Introduction to Modern Number Theory*, 2nd edition, Springer-Verlag, New York, 1990.
13. A. Legendre, *Recherches d'Analyse Indéterminée*. Histoire de l'Académie Royale des Sciences, 1785. (Also in: A. Legendre, *Thèorie des Nombres*, 3^{me} ed., t.I, §§III, IV, Mazer, Leipzig, 1886.)
14. D. Lehmer, A note on trigonometric numbers, *Amer. Math. Monthly* 40(1933), 165–166.
15. F. Lindermann, Über die Zahl π, *Math. Ann.* 20(1882), 213–225.
16. *Math Horizons* 1(1993), 21.
17. B. Mazur, Modular curves and the Eisenstein ideals, *IHES Publ. Math.* 47(1977), 33–186.
18. B. Mazur, Rational isogenies of prime degree, *Invent. Math.* 44(1978), 129–162.
19. L. J. Mordell, On the rational solutions of the indeterminate equations of the third and fourth degrees, *Proc. Camb. Philos. Soc.* 21(1922), 179–192.
20. J. Olmsted, Rational values of trigonometric functions, *Amer. Math. Monthly* 52(1945), 507–508.
21. H. Poincaré, Sur la propriétés arithmétiques des courbes algébriques, *Journal de Mathematiques Pures et Appliquees* 7(1901), 161–233.
22. C. L. Siegel, Über einige Anwendungen diophantischer Approximationen, *Abh. Preuss. Akad. Wiss. Phys.-Mat. Kl.* (1929) (*Collected Works*, Springer-Verlag, New York, 1966, 209–266.)
23. J. Silverman and J. Tate, *Rational Points on Elliptic Curves*, Springer-Verlag, New York, 1992.
24. E. Swift, Note on trigonometric functions, *Amer. Math. Monthly* 29(1922), 404–405.
25. R. Taylor and A. Wiles, Ring theoretic properties of certain Hecke algebras, *Annals of Math.* 141(1995), 553–572.
26. R. Underwood, On the irrationality of certain trigonometric functions, *Amer. Math. Monthly* 28(1921), 374–376.
27. K. Weierstrass, Zu Lindermanns Abhandlung "Über die Ludolphsche Zahl," *Sitzgsber. Preuss. Akad. Wiss.* (1885), 1067–1085. (*Math. Werke*, II, 341–362.)
28. A. Wiles, Modular elliptic curves and Fermat's Last Theorem, *Annals of Math.* 141(1995), 443–551.