ON THE DISTRIBUTION OF PYTHAGOREAN TRIPLES

Edward K. Hinson
University of New Hampshire, Durham, NH 03824
(Submitted January 1991)

1. Introduction

A triple (a, b, c) of natural numbers is a *pythagorean triple* if $a^2 + b^2 = c^2$, that is, if there exists a right triangle whose sides are lengths a, b, and c. If $\gcd(a, b) = 1$, then the triple is *primitive*. The family of such triples was among the earliest mathematical objects to be completely characterized.

Theorem 1: Every primitive pythagorean triple (x, y, z) with x even and x, y, $z > 0$ is given by

$$x = 2st, \quad y = s^2 - t^2, \quad z = s^2 + t^2$$

for positive integers s, t such that $\gcd(s, t) = 1$ and $s \not\equiv t \pmod 2$. Conversely, each such pair s, t gives a primitive pythagorean triple by the formula.

In this paper we pursue alternate descriptions of the family of pythagorean triples. We approach this by way of functions which map the set of triples into subsets of $\mathbf{R}$ in which their distribution can be represented topologically and algebraically.

2. The Counting Function ν

We wish to characterize pythagorean triples in terms of two parameters: the positive differences between the lengths of the hypotenuse and the respective legs. In order that this be unambiguous, we must verify that any pair (a, b) in $\mathbf{N} \times \mathbf{N}$, $a \le b$, corresponds to at most one triple. But this amounts to showing that the quadratic equation

$$x^2 + (x + a)^2 = (x + b)^2 \tag{1}$$

has at most one natural number solution—an easy exercise using the quadratic formula. Thus, we have a function

$$\nu_0 : (\mathbf{N} \cup \{0\}) \times \mathbf{N} \to \{0, 1\},$$

where $\nu_0(a, b) = 1$ if and only if there exists a natural number solution for the equation (1).

One can formulate this more concisely. Let $S = \mathbf{Q} \cap [0, 1)$, the set of all rational points in the unit interval except the right endpoint 1. Define

$$\nu : S \to \{0, 1\}$$

by

$$\nu(a/b) = \nu_0(a, b).$$

For ν to be well defined, it suffices that, for all a, b, d in $\mathbf{N}$, we have

$$\nu_0(a, b) = \nu_0(ad, bd).$$

But this holds since

$$(b - a) + \sqrt{2b(b - a)} \in \mathbf{N}$$

if and only if

$$d(b - a) + d\sqrt{2b(b - a)} = (db - da) + \sqrt{2(db)(db - da)} \in \mathbf{N}.$$

Note that any common divisor of x, $x + a$, and $x + b$ must divide both a and b. Since every fraction can be represented in lowest terms, it follows that a one-to-one correspondence exists between the elements of $\nu^{-1}(1)$ and the primitive pythagorean triples. Considering S to have the topology induced by the usual one on $\mathbf{R}$ we may use ν to represent the primitive triples in S and study them from a topological viewpoint.

For example, consider the infinite family of triples

$$(2) \qquad (2n + 1,\ 2n^2 + 2n,\ 2n^2 + 2n + 1),\ n \in \mathbf{N}.$$

Under ν these correspond to the rational numbers

$$q_n = \frac{2n^2 - 1}{2n^2},\ n \geq 1.$$

Thus, in the real unit interval $I = [0, 1]$, the accumulation point 1 of the set $\nu^{-1}(1)$ reflects the asymptotic equality of the longer leg and the hypotenuse in the family (2).

We shall use the following basic property of ν in the next section.

Proposition 2: Let a, b be natural numbers. If a is even and b is odd, then $\nu(a/b) = 0$.

Proof: It suffices to show that $\sqrt{2b(b - a)}$ cannot be an integer. Under the hypotheses, both b and $b - a$ are odd; thus, there is not the second factor of 2 necessary in $2b(b - a)$ for it to be a square.

3. A Density Theorem for ν

Most of the easily represented families of triples yield sequences in I converging to 1; e.g.,

$$(2n,\ n^2 - 1,\ n^2 + 1),$$
$$(4n^2,\ n^4 - 4,\ n^4 + 4),$$
$$(2n + 1,\ 2n^2 + 2n,\ 2n^2 + 2n + 1).$$

But there may be many other accumulation points of $\nu^{-1}(1)$. We can use Theorem 1 to determine the inverse images of the counting function ν.

Theorem 3: The sets $\nu^{-1}(0)$ and $\nu^{-1}(1)$ are both dense in the real unit interval I with respect to the usual metric.

Proof: We shall use Proposition 2 to show the density of $\nu^{-1}(0)$. Since $\nu(0) = \nu(1) = 0$, choose r in $(0, 1)$ and $\varepsilon > 0$. Choose b to be an even natural number satisfying $1/(b^2 + 1) < \varepsilon/2$. Now for some nonnegative integer a the interval $(r - \varepsilon,\ r + \varepsilon)$ contains both $a/(b^2 + 1)$ and $(a + 1)/(b^2 + 1)$. Exactly one of a and $a + 1$ is even (say it's a), and now $\nu(a/(b^2 + 1)) = 0$ by Proposition 2. Since ε is arbitrary we have r in the closure of $\nu^{-1}(0)$.

To show the density of $\nu^{-1}(1)$ in I it suffices to show that every neighborhood in I contains some a/b with $\nu(a/b) = 1$. Choose r and ε from $(0, 1)$ such that $0 < \varepsilon < \min\{r, 1 - r\}$. We can restrict ourselves (thus slightly strengthening the result) to those triples whose longer leg has even length, i.e., for which $2st > s^2 - t^2$ in the characterization of Theorem 1. Solving the quadratic inequality resulting from the substitution $\gamma = s/t$ gives $s < (1 + \sqrt{2})t$ as a necessary and sufficient condition for this restriction. Thus, by Theorem 1, we wish to find relatively prime s and t, exactly one of which is even, so that

$$(3) \qquad r - \varepsilon < \frac{2st - (s^2 - t^2)}{(s^2 + t^2) - (s^2 - t^2)} < r + \varepsilon.$$

Again using $\gamma = s/t$ and the quadratic formula, and setting $R = 1 - r - \varepsilon$, we have (3) if and only if

$$(4) \qquad \sqrt{R} < \frac{1}{\sqrt{2}}\left(\frac{s}{t} - 1\right) < \sqrt{R + 2\varepsilon}.$$

The density of $\mathbf{Q}$ in $\mathbf{R}$ insures that relatively prime s_0 and t_0 exist which satisfy (4). Furthermore, $\sqrt{R + 2\varepsilon} < 1$ implies that $s_0 < (1 + \sqrt{2})t_0$. If exactly one of s_0 and t_0 is even, we may take $s = s_0$, $t = t_0$ and be done. If s_0 and t_0 are both odd, choose $N > 0$ odd and large enough so that

$$\sqrt{R} < \frac{1}{\sqrt{2}}\left(\frac{Ns_0 + 1}{Nt_0} - 1\right) < \sqrt{R + 2\varepsilon}.$$

Let s and t be the numerator and denominator, respectively, of the lowest terms representation of $(Ns_0 + 1)/Nt_0$; it follows from the choice of N that s is even and t is odd. In this way we can construct a rational a/b with $\nu(a/b) = 1$ and $|(a/b) - r| < \varepsilon$, and the theorem is proved.

4. A Representation in the Multiplicative Positive Rationals

There is another formulation of the counting function which is of interest. Define a function

$$\eta : \mathbf{Q}^+ \to \{0, 1\}$$

by

$$\eta(a/b) = \nu(a/(a + b))$$

and note that it, too, is well defined. There is again a one-to-one correspondence between primitive triples and the elements of $\eta^{-1}(1)$. Realizing η as $\nu \circ f$, where $f : \mathbf{Q}^+ \to [0, 1)$ is given by $f(x) = x/(1 + x)$, allows one to deduce from the continuity of f that $\eta^{-1}(0)$ and $\eta^{-1}(1)$ are both dense in $\mathbf{Q}^+$.

The natural multiplicative closure in $\mathbf{Q}^+$ suggests the possibility of an induced closure in $\eta^{-1}(0)$, $\eta^{-1}(1)$, or related subsets. But direct calculations yield

$$\eta(7) = \eta\left(\frac{1}{8}\right) = 1, \quad \eta\left(\frac{7}{8}\right) = \eta\left(\frac{1}{2}\right) = \eta\left(\frac{1}{4}\right) = 0,$$

which taken together show the failure of closure in $\eta^{-1}(0)$ and $\eta^{-1}(1)$. One may observe some slight structure, however, from the following point of view. Let

$$I = \left\{\frac{p}{q} \in \mathbf{Q}^+ : \eta\left(\frac{p}{q}\right) = \eta\left(\frac{q}{p}\right) = 1\right\}$$

and

$$I' = \left\{\frac{p}{q} \in \mathbf{Q}^+ : \eta\left(\frac{p}{q}\right) = \eta\left(\frac{q}{p}\right)\right\}.$$

Clearly, I contains 1, and thus one has a chain $I \subseteq I' \subseteq \mathbf{Q}^+$ of nonempty sets. In fact, we can further characterize the elements of I.

Proposition 4: Let p and q be in $\mathbf{Z}^+$ with $\gcd(p, q) = 1$. Then p/q is in I if and only if $\eta(p/q) \cdot \eta(q/p) = 1$ if and only if p and q are each squares and $p + q$ is twice a square.

Proof: The first equivalence is immediate. Note that

$$f(p/q) = p/(p + q) \quad \text{and} \quad f(q/p) = q/(p + q)$$

and so $\eta(p/q) \cdot \eta(q/p) = 1$ if and only if both

$$\sqrt{2(p + q)q} \quad \text{and} \quad \sqrt{2(p + q)p}$$

are integers. Suppose that p, q, and $(p + q)/2$ are each squares. Then the above radicals are clearly integers. Conversely, if

$$\sqrt{2(p + q)q} \quad \text{and} \quad \sqrt{2(p + q)q}$$

are both integers, then so is

$$\sqrt{2(p + q)q} \cdot \sqrt{2(p + q)p} = 2(p + q)\sqrt{pq}$$

and thus pq is a square. Moreover, since they are relatively prime, each of p and q must be a square. Letting p be a square, it follows from the integrality of $\sqrt{2(p + q)p}$ that $(p + q)/2$ is also a square, as required.

One sees as a corollary that a given p/q from $\eta^{-1}(1)$ is in I if and only if p and q are squares. This observation is useful in proving the following result.

Proposition 5: Let p, p_i, q, and q_i be positive integers.

(i) If p_i/q_i is in I, $i = 1, 2$, then p_1p_2/q_1q_2 is in I';

(ii) for any positive rational p/q, I' contains $(p/q)^2$;

(iii) if p/q is in I, then $(p/q)^n$ is in I' for all $n \geq 1$.

Proof: If, under the hypothesis of (i), $p_1p_2 + q_1q_2$ is twice a square, then p_1p_2/q_1q_2 is in I by Proposition 4. If $p_1p_2 + q_1q_2$ is not twice a square then

$$\eta(p_1p_2/q_1q_2) \cdot \eta(q_1q_2/p_1p_2) = 0;$$

but each factor must be 0 since, otherwise, the above remark would force their product to be 1. A similar argument proves (ii) immediately, and (iii) follows from (ii) using Proposition 4.

As in the previous section, one may wish to know the accumulation points of I and I' in the nonnegative half-line $\mathbf{R}^+ \cup \{0\}$.

Theorem 6: The sets I and I' are dense in $\mathbf{R}^+$.

Proof: The density of I' will follow from that of I by the inclusion $I \subseteq I'$. We know from Proposition 4 that p/q is in I if and only if p and q are squares and $p + q$ is twice a square. Note that such p/q, in lowest terms, correspond to the primitive solutions of the diophantine equation $u^2 + v^2 = 2w^2$ when $p = u^2$ and $q = v^2$. One may calculate that

$$(b - a)^2 + (b + a)^2 = 2c^2$$

if and only if (a, b, c) is a pythagorean triple. Thus, it will suffice to show that as a and b vary among primitive pythagorean triples (a, b, c) the fractions $(b - a)/(b + a)$ are dense in the interval $(0, 1)$. We argue as in Theorem 3. Characterizing the primitive triples as in Theorem 1, restricting our attention to those triples in which $2st > s^2 - t^2$ and setting $\gamma = s/t$ gives

$$\frac{b - a}{b + a} = \frac{2\gamma - \gamma^2 + 1}{2\gamma + \gamma^2 - 1}.$$

But now, differentiating this expression with respect to the real variable γ shows that its range on the restricted domain $(\sqrt{2} - 1, \sqrt{2} + 1)$ is all of $\mathbf{R}^+$; as in Theorem 3, the restriction above on s and t holds in this interval. We complete the proof by using the technique of Theorem 3 to produce s/t corresponding to primitive pythagorean triples arbitrarily close to any rational in $(0, 1)$.

Acknowledgment

The author gratefully acknowledges the referee's observation of the result in Theorem 6.

AMS Classification number: 10A99

*****