

8079.pdf

International Mathematical Forum, Vol. 8, 2013, no. 38, 1873 - 1879

HIKARI Ltd, www.m-hikari.com

<http://dx.doi.org/10.12988/imf.2013.39180>

Primitive Pythagorean Codes

Massoud Malek

California State University, East Bay

Hayward, CA 94542, USA

massoud.malek@csueastbay.edu

Copyright © 2013 Massoud Malek. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

In this paper, we shall define a simple generating matrix which produces all the primitive Pythagorean triples. Then we define classes of codes with primitive Pythagorean triples. These classes are very simple to encode and decode messages. They also have a very good data compression quality.

Keywords: Primitive Pythagorean triple; PPT; Pythagorean Triangle; Pythagoras, Plato, Euclid; generating matrix; Pythagorean alphabet; Coding theory; Coset Leader decoding; and Data Compression

Introduction

A *Pythagorean triple*, is a triple of positive integers a , b , and c such that a right angle triangle exists with legs a , b and hypotenuse c . By the Pythagorean theorem, this is equivalent to finding positive integers a , b , and c , satisfying

$$a^2 + b^2 = c^2.$$

If a and b are relatively prime, then the Pythagorean triple is called *primitive*. The smallest and best-known Pythagorean triple is $(3, 4, 5)$. The triangle generated by a Pythagorean triple is called a *Pythagorean triangle*. In all that follows we denote the primitive Pythagorean triple or triangle by *PPT*.

Plato (380 B. C.) is attributed with the formula:

$$\langle 1, n \rangle = (n^2 - 1, 2n, n^2 + 1), \quad \text{for } n > 1.$$

A more general formula for obtaining all triples was given by Euclid in c. 300 B.C.

Primitive Pythagorean Codes

Massoud Malek

California State University, East Bay

Hayward, CA 94542, USA

massoud.malek@csueastbay.edu

Copyright © 2013 Massoud Malek. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

In this paper, we shall define a simple generating matrix which produces all the primitive Pythagorean triples. Then we define classes of codes with primitive Pythagorean triples. These classes are very simple to encode and decode messages. They also have a very good data compression quality.

Keywords: Primitive Pythagorean triple; PPT; Pythagorean Triangle; Pythagoras, Plato, Euclid; generating matrix; Pythagorean alphabet; Coding theory; Coset Leader decoding; and Data Compression

Introduction

A *Pythagorean triple*, is a triple of positive integers a , b , and c such that a right angle triangle exists with legs a , b and hypotenuse c . By the Pythagorean theorem, this is equivalent to finding positive integers a , b , and c , satisfying

$$a^2 + b^2 = c^2.$$

If a and b are relatively prime, then the Pythagorean triple is called *primitive*. The smallest and best-known Pythagorean triple is $(3, 4, 5)$. The triangle generated by a Pythagorean triple is called a *Pythagorean triangle*. In all that follows we denote the primitive Pythagorean triple or triangle by *PPT*.

Plato (380 B. C.) is attributed with the formula:

$$\langle 1, n \rangle = (n^2 - 1, 2n, n^2 + 1), \quad \text{for } n > 1.$$

A more general formula for obtaining all triples was given by Euclid in c. 300 B.C.

Lemma 1. For positive integers m and n , where $1 \leq m < n$, the triple

$$\langle m, n \rangle = (n^2 - m^2, 2mn, n^2 + m^2)$$

is Pythagorean. Moreover, if m and n are relatively prime of opposite parity, then they generate a primitive triple.

Basic Properties of Primitive Pythagorean Triangles

In a primitive Pythagorean triangle $T = (a, b, c)$, the following conditions hold:

- The hypotenuse and one of the legs are always odd and the other leg is divisible by 4.
- The hypotenuse of every PPT exceeds the odd leg by twice the square of an integer r and the even leg by the square of an odd integer s ,

$$c = a + 2r^2 = b + s^2$$

Generating Matrix

Consider the matrix $\mathcal{G} = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 2 & 2 \\ 0 & 2 & 2 \end{pmatrix}$ and the vector $u[m, n] = (m^2, mn, n^2)$, where m is odd and $\gcd(m, n) = 1$. Then

$$u[m, n] \mathcal{G} = (m^2, mn, n^2) \begin{pmatrix} 1 & 0 & 1 \\ 2 & 2 & 2 \\ 0 & 2 & 2 \end{pmatrix} = (m^2 + 2mn, 2mn + 2n^2, m^2 + 2mn + 2n^2) = \langle n, m + n \rangle.$$

Proposition 1. Let $\mathcal{P}$ be the set of all PPTs. For fixed positive integers m_0 and n_0 , where m_0 is odd, define the following sets of PPTs:

$$\begin{aligned} \mathcal{R}_{m_0} &= \{T_n = (a_n, b_n, c_n) = u[m_0, n] \mathcal{G} : n = 1, 2, 3, \dots, \gcd(m_0, n) = 1\}, \\ \mathcal{S}_{n_0} &= \{T_m = (a_m, b_m, c_m) = u[m, n_0] \mathcal{G} : m = 1, 3, 5, \dots, \gcd(m, n_0) = 1\}. \end{aligned} \quad \text{and}$$

Then $\mathcal{R}_{m_0} \cap \mathcal{S}_{n_0} = u[m_0, n_0] \mathcal{G}$, whenever $\gcd(m_0, n_0) = 1$; also for all $T_n \in \mathcal{R}_{m_0}$ and $T_m \in \mathcal{S}_{n_0}$, $c_n - b_n = m_0^2$ and $c_m - a_m = 2n_0^2$; and finally for $i_1 \neq i_2$ and $j_1 \neq j_2$,

$$\mathcal{P} = \bigcup_{i=0}^{\infty} \mathcal{R}_{2i+1} = \bigcup_{j=1}^{\infty} \mathcal{S}_j, \quad \mathcal{R}_{2i_1+1} \cap \mathcal{R}_{2i_2+1} = \emptyset, \quad \mathcal{S}_{2j_1} \cap \mathcal{S}_{2j_2} = \emptyset.$$

Moreover, $\mathcal{R}_1$ contains all the PPTs, where the hypotenuse exceeds its even leg by one; and $\mathcal{S}_1$ contains all the PPTs, where the hypotenuse exceeds its odd leg by two.

Proof. The identity $u[m, n] \mathcal{G} = \langle n, m + n \rangle$ and Lemma 1 imply that any PPT can be obtained from the product of the vector $u[m, n]$ and the generating matrix $\mathcal{G}$. Thus

$$\mathcal{P} = \bigcup_{i=0}^{\infty} \mathcal{R}_{2i+1} = \bigcup_{j=1}^{\infty} \mathcal{S}_j.$$

We also conclude that if $\gcd(m_0, n_0) = 1$, then $\mathcal{R}_{m_0} \cap \mathcal{S}_{n_0}$ contains only one PPT, which is $u[m_0, n_0] \mathcal{G}$.

From the definitions of $u[r, s] \mathcal{G}$, we obtain

$$T = (a, b, c) = u[r, s] \mathcal{G} = (r^2 + 2rs, 2rs + 2s^2, r^2 + 2rs + 2s^2)$$

Notice that $c - b = r^2$ and $c - a = 2s^2$. Thus for the same a, b , and c ; r and s must be unique. Hence for $i_1 \neq i_2$ and $j_1 \neq j_2$,

$$\mathcal{R}_{2i_1+1} \cap \mathcal{R}_{2i_2+1} = \emptyset \quad \text{and} \quad \mathcal{S}_{2j_1} \cap \mathcal{S}_{2j_2} = \emptyset.$$

Finally, from $c - b = r^2$ and $c - a = 2s^2$, we conclude that $c_n - b_n = (2i + 1)^2$ and $c_m - a_m = 2j^2$. Hence only $\mathcal{R}_1$ contains all the PPTs, where the hypotenuse exceeds its even leg by one; and only $\mathcal{S}_1$ contains all the PPTs, where the hypotenuse exceeds its odd leg by two. ■

Proposition 2. Let $T = (a, b, c)$ be a PPT.

$$(i) \text{ If } T \in \mathcal{R}_{m_0}; \text{ then } T = u \left[m_0, \frac{a - m_0^2}{2m_0} \right] \mathcal{G} = \left(a, \frac{a^2 - m_0^4}{2m_0^2}, \frac{a^2 + m_0^4}{2m_0^2} \right).$$

$$(ii) \text{ If } T \in \mathcal{S}_{n_0}; \text{ then } T = u \left[-n_0 + \sqrt{n_0^2 + a}, n_0 \right] \mathcal{G} = \left(a, 2n_0 \sqrt{n_0^2 + a}, a + 2n_0^2 \right).$$

Proof. (i) If $T \in \mathcal{R}_{m_0}$; then there exists a positive integer n with $\gcd(m_0, n) = 1$ such that

$$T = (a, b, c) = u[m_0, n] \mathcal{G} = (m_0^2 + 2m_0n, 2n(m_0 + n), m_0^2 + 2n(m_0 + n)).$$

From $a = m_0^2 + 2m_0n$, we obtain $n = \left(\frac{a - m_0^2}{2m_0} \right)$. Hence

$$T = u \left[m_0, \frac{a - m_0^2}{2m_0} \right] \mathcal{G} = \left(a, \frac{a^2 - m_0^4}{2m_0^2}, \frac{a^2 + m_0^4}{2m_0^2} \right).$$

(ii) If $T \in \mathcal{S}_{n_0}$; then there exists an odd number m with $\gcd(m, n_0) = 1$ such that

$$T = (a, b, c) = u[m, n_0] \mathcal{G} = (m^2 + 2mn_0, 2n_0(m + n_0), m^2 + 2n_0(m + n_0)).$$

By solving the quadratic equation $a = m^2 + 2mn_0$ for a positive m , we obtain

$$m = \left(-n_0 + \sqrt{n_0^2 + a} \right). \text{ Hence}$$

$$T = u \left[-n_0 + \sqrt{n_0^2 + a}, n_0 \right] \mathcal{G} = \left(a, 2n_0 \sqrt{n_0^2 + a}, a + 2n_0^2 \right). \quad \blacksquare$$

From the fact that the last digit of any integer, could not be different from 0, 1, 2, 3, 4, 5, 6, 7, 8 and 9; we conclude that any $\mathcal{R}_{m_0}$ and $\mathcal{S}_{n_0}$ may be partitioned into a finite number of disjoint classes, based on the last digits of a , b , and c , of $T = (a, b, c)$.

The first 30 PPTs $T_n = (a_n, b_n, c_n)$, generated by $[1, n, n^2] \mathcal{G}$:

(3, 4, 5)	(5, 2, 3)	(7, 4, 5)	(9, 0, 1)	(1, 0, 1)
3 4 5	5 12 13	7 24 25	9 40 41	11 60 61
13 84 85	15 112 113	17 144 145	19 180 181	21 220 221
23 264 265	25 312 313	27 364 365	29 420 421	31 480 481
33 544 545	35 612 613	37 684 685	39 760 761	41 840 841
43 924 925	45 1012 1013	47 1104 1105	49 1200 1201	51 1300 1301
53 1404 1405	55 1512 1513	57 1624 1625	59 1740 1741	61 1860 1861

The above table shows that the set of T_n , is partitioned into five disjoint classes, with coset leaders:

$$(3, 4, 5), (5, 2, 3), (7, 4, 5), (9, 0, 1), \text{ and } (1, 0, 1).$$

According to Proposition 1, $c_n - b_n = 1$. From $a_n = 1 + 2n$, $b_n = 2n(1 + n)$, and $c_n = 1 + 2n(1 + n)$, one may readily show the following identities which appear in the above table:

$$a_{n+1} = a_n + 2, a_{5k+n} = a_n + 10k, a_{n+6} = a_n + 12, a_n^2 = b_n + c_n, \text{ and } a_{6k+1} \text{ is divisible by } 3.$$

The first 30 PPTs $T_m = (a_m, b_m, c_m)$, generated by $[m^2, m, 1] \mathcal{G}$ (Plato PPT), where m is odd:

(3, 4, 5)	(5, 8, 7)	(5, 2, 7)	(3, 6, 5)	(9, 0, 1)
3 4 5	15 8 17	35 12 37	63 16 65	99 20 101
143 24 145	195 28 197	255 32 257	323 36 325	399 40 401
483 44 485	575 48 577	675 52 677	783 56 785	899 60 901
1023 64 1025	1155 68 1157	1295 72 1297	1443 76 1445	1599 80 1601
1763 84 1765	1935 88 1937	2115 92 2117	2303 96 2305	2499 100 2501
2703 104 2705	2915 108 2917	3135 112 3137	3363 116 3365	3599 120 3601

According to the above table, the set of T_m is partitioned into five disjoint classes, with coset leaders:

$$(3, 4, 5), (5, 8, 7), (5, 2, 7), (3, 6, 5), \text{ and } (9, 0, 1).$$

According to Proposition 1, $c_m - a_m = 2$. From $a_m = m^2 + 2m$, $b_m = 2(m + 1)$, and $c_m = m^2 + 2(m + 1)$, one may readily show the following identities which appear in the above table:

$$b_{m+2} = b_m + 4, b_{m+10k} = b_m + 20k, b_m^2 = 2(b_m + c_m), \text{ and } a_{12k+1} \text{ is divisible by } 3.$$

Application to Coding Theory

Primitive Pythagorean triple can lend itself to Coding Theory. There are infinitely many ways to encode letters of alphabet with PPTs.

To define a class of codewords, first we select either $\mathcal{R}_{m_0}$ or $\mathcal{S}_{n_0}$. Suppose, we choose $\mathcal{R}_{m_0}$, where m_0 is a prime odd number; then we assign a positive integer k with $\gcd(m_0, k) = 1$ to each letter of the alphabet. Finally, we create a Pythagorean alphabet using the generating matrix $\mathcal{G}$.

Here is a Pythagorean alphabet using $\mathcal{R}_7$. For the sake of simplicity, the alphabet was generated by $u[7, k]\mathcal{G}$, where $k = 1, 2, 3, \dots, 30$, excluding $k = 7, 14, 21$, and 28 which are multiples of 7.

The Pythagorean alphabet is as follows:

$A \rightarrow 63\ 16\ 65$	$B \rightarrow 77\ 36\ 85$	$C \rightarrow 91\ 60\ 109$	$D \rightarrow 105\ 88\ 137$
$E \rightarrow 119\ 120\ 169$	$F \rightarrow 133\ 156\ 205$	$G \rightarrow 161\ 240\ 289$	$H \rightarrow 175\ 288\ 337$
$I \rightarrow 189\ 340\ 389$	$J \rightarrow 203\ 396\ 445$	$K \rightarrow 217\ 456\ 505$	$L \rightarrow 231\ 520\ 569$
$M \rightarrow 259\ 660\ 709$	$N \rightarrow 273\ 736\ 785$	$O \rightarrow 287\ 816\ 865$	$P \rightarrow 301\ 900\ 949$
$Q \rightarrow 315\ 988\ 1037$	$R \rightarrow 329\ 1080\ 1129$	$S \rightarrow 357\ 1276\ 1325$	$T \rightarrow 371\ 1380\ 1429$
$U \rightarrow 385\ 1488\ 1537$	$V \rightarrow 399\ 1600\ 1649$	$W \rightarrow 413\ 1716\ 1765$	$X \rightarrow 427\ 1836\ 1885$
$Y \rightarrow 455\ 2088\ 2137$	$Z \rightarrow 469\ 2220\ 2269$		

This set contains five different disjoint classes:

$(3, 6, 5)$	$(7, 6, 5)$	$(1, 0, 9)$	$(5, 8, 7)$	$(9, 0, 9)$
$A\ F\ J\ N\ W$	$B\ K\ O\ S\ X$	$C\ L\ T\ G\ P$	$D\ H\ Q\ U\ Y$	$E\ I\ M\ R\ V\ Z$

Error Detection

The fact that $\mathcal{R}_7$ is partitioned into five disjoint classes; the received codeword must be in one of those five classes. If the last digits of a , b , or c do not match the digits of any of the coset leaders, the received codeword may be corrupted. Once the error is detected, it could eventually be corrected either directly or by a retransmission. It is customary that all codewords have the same length. We could achieve this task by adding some digits, such as zeros or ones at the beginning or the end of the shorter codewords.

Data Compression

Sometimes, it is necessary to shorten the length of the codewords without compromising the integrity of the message. According to Proposition 1, $\mathcal{R}_{m_0}$ contains all the PPTs, where the hypotenuse exceeds its even leg by m_0^2 ; and $\mathcal{S}_{n_0}$ contains all the PPTs, where the hypotenuse exceeds its odd leg by $2n_0^2$. Removing the hypotenuse of the encrypted code in these sets, would not alter the message and its integrity.

Notice that in $\mathcal{R}_7$, $c - b = 49 = 7^2$, for all c 's and b 's; so the alphabet could be shorten, by removing all the hypotenuses. According to Proposition 2, from the odd leg a of the codeword, we obtain $k = \frac{a - 49}{14}$, the number assigned to the letter which was encoded

with the generating matrix $\mathcal{G}$. Therefore we could also remove the even leg b .

In order to be able to detect some transmission errors, we should add the last two digits of the coset leader of the class which contains the codeword. Here is the short form of the alphabet with parity-check digits:

$A \rightarrow 63\ 65$	$B \rightarrow 77\ 65$	$C \rightarrow 91\ 09$	$D \rightarrow 10\ 87$	$E \rightarrow 119\ 65$	$F \rightarrow 133\ 65$
$G \rightarrow 161\ 09$	$H \rightarrow 175\ 87$	$I \rightarrow 189\ 09$	$J \rightarrow 203\ 65$	$K \rightarrow 217\ 65$	$L \rightarrow 231\ 09$
$M \rightarrow 259\ 09$	$N \rightarrow 273\ 65$	$O \rightarrow 287\ 65$	$P \rightarrow 301\ 09$	$Q \rightarrow 315\ 87$	$R \rightarrow 329\ 09$
$S \rightarrow 357\ 65$	$T \rightarrow 371\ 09$	$U \rightarrow 385\ 87$	$V \rightarrow 399\ 09$	$W \rightarrow 413\ 65$	$X \rightarrow 427\ 65$
$Y \rightarrow 455\ 87$	$Z \rightarrow 469\ 09$				

Error Correction and Decoding

Suppose the codewords $w_1 = 133\ 67$ and $w_2 = 359\ 75$ were received. Since $(3, 6, 7)$ and $(9, 7, 5)$ are not coset leaders; but $133\ 65$ is in our table, generated by $u[7, 6]\mathcal{G}$, representing the letter F ; and $359\ 65$ is very close to $357\ 75$ generated by $u[7, 22]\mathcal{G}$, representing the letter S . We therefore conclude that it is most likely that $v_1 = 133\ 65$ and $v_2 = 357\ 65$ are the transmitted codewords.

Although the generating matrix $\mathcal{G}$ is needed to encode a message; the received codeword by itself may decode the message without the need of the inverse of the generating matrix $\mathcal{G}$. We could recreate the message, using 7 , $k = \frac{a-49}{14}$, and $\mathcal{G}$.

One could obviously generate a Pythagorean alphabet, using $\mathcal{S}_{n_0}$. To decode the message, $k = \left(-n_0 + \sqrt{n_0^2 + a}\right)$ must be used.

In order to increase the probability of correcting errors, one could assign for example, three PPTs to each letter in the following way: "Choose the odd leg of the first PPT, the even leg of the second PPT, and the hypotenuse of the third PPT." This way, if one of the legs is error-free, then the codeword may be identified.

Finally, another way of increasing the probability of correcting errors and decreasing the number of retransmission; is to choose k 's with large gaps between them. This way, it is most likely that the closest codeword to the received codeword is the codeword sent.

Clearly, all codewords must be converted to binary codewords, before transmission.

References

- [1] Carmichael, R. D., 1914, "Diophantine analysis," in second half of R. D. Carmichael, *The Theory of Numbers and Diophantine Analysis*, Dover Publ., 1959.
- [2] Edenfield, Kelly, "Pythagorean Triples." <http://jwilson.coe.uga.edu/EMT668/EMT668.Folders.F97/Edenfield/Pythtriples/Pythtriples.html>
- [3] Posamentier, Alfred; Lehmann, Ingmar (2007). *The (Fabulous) FIBONACCI Num-*

bers. Prometheus Books. p. 305. ISBN 978-1-59102-475-0.

Received: September 27, 2013