ON THE COMPATIBILITY OF BÉZOUT COEFFICIENTS BETWEEN PYTHAGOREAN PAIRS UNDER UNIMODULAR TRANSFORMATIONS

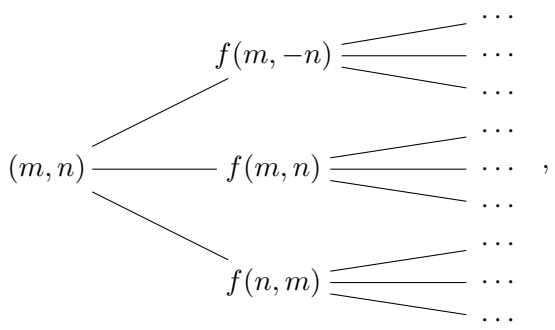
ABSTRACT. In a recent preprint, Gullerud and Walker [2] proved a theorem and made a conjecture about the correctness of efficiently generating Bézout trees for Pythagorean pairs. In this note, we give a simple proof of their theorem, confirm that their conjecture is true, and furthermore we give a generalization.

1. Introduction

The integers triple (x, y, z) is called a Pythagorean triple if $x^2 + y^2 = z^2$. It is called primitive if they are relatively prime. It is well known that all positive primitive Pythagorean triples (x, y, z) with y even can be written as

$$x = m^2 - n^2, y = 2mn, z = m^2 + n^2,$$

for some relative prime integers m and n such that $m > n > 0$ [4]. Following the authors of [2], we call such (m, n) a *Pythagorean pair*. Given (m, n), it is clear that (n, m) and $(m, -n)$ also generate Pythagorean triples; such pairs are called *associated pairs* of (m, n). Note that if (m, n) is a Pythagorean pair, then $f(m, n) := (2m + n, m)$ (where f is defined on $\mathbb{Z} \times \mathbb{Z}$) is another Pythagorean pair. Similarly, $f(n, m) = (2n + m, n)$ and $f(m, -n) = (2m - n, m)$ are also Pythagorean pairs. Define now a trinary tree generated by (m, n) as follows:



where recursively, each node on a given level produces three nodes on a next level by applying the functions $f_1(m, n) = f(m, -n), f_2(m, n) = f(m, n)$

and $f_3(m,n) = f(n,m)$ and so on. Randall and Saunders [5] proved that the trinary tree produced from $(3,1)$ contains all pairs of relatively prime odd integers. Similarly, the trinary tree produced from $(2,1)$ contains all pairs of relatively prime integers of opposite parity. Thus these together generate all relatively prime Pythagorean pairs (m,n) with $m > n > 0$.

We call (r,s) the Bézout coefficients associated with (m,n) if (r,s) is obtained from the standard division algorithm so that $rm + sn = \gcd(m,n)$. For comparison, for an input of (m,n) in the Matlab `gcd` function

$$[G,U,V] = \texttt{gcd}(m,n),$$

the output will be $G = \gcd(m,n)$ in the usual notation, and $U = r, V = s$ are the Bézout coefficients. In an attempt to efficiently generate the Bézout coefficients for Pythagorean pairs, Gullerud and Walker introduced the notion of **Bézout tree of** (m,n) **generated by** (u,v), which is defined by

$$g(u,v) = (v, u-2v),$$

$$g(v,u) = (u, v-2u) \text{ and}$$

$$g(u,-v) = (-v, u+2v),$$

and the tree is arranged in the analogous format as in the tree starting with (m,n). Gullerud and Walker proved the following result, for which we offer a simple argument.

Theorem 1.1. (cf. Theorem 1.2 of [2]) Let (m,n) be a Pythagorean pair with $m > n$ with associated pairs (n,m) and $(m,-n)$. Let f and g be as defined above, and let $mu + nv = 1$ for some $u, v \in \mathbb{Z}$. Then $g(u,v), g(v,u)$ and $g(u,-v)$ respectively yield the necessary coefficients u', v' such that

$$(2m+n)u' + mv' = 1,$$

$$(2n+m)u' + nv' = 1, \text{ and}$$

$$(2m-n)u' + mv' = 1$$

respectively.

Proof. In terms of matrices, we have

$$f(m,n) = \begin{bmatrix} 2m+n \\ m \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} m \\ n \end{bmatrix}.$$

If we let $A = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}$, then it is clear that $g(u,v)$ is given by

$$g(u,v) = \begin{bmatrix} v \\ u-2v \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = (A^{-1})^T \begin{bmatrix} u \\ v \end{bmatrix}.$$

It follows that $\begin{bmatrix} u' \\ v' \end{bmatrix} := g(u,v)$ satisfies

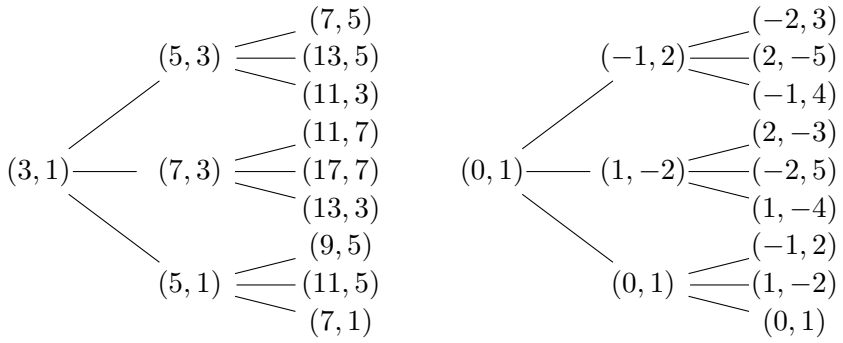
$$(2m+n)u' + mv' = [u', v'] \begin{bmatrix} 2m+n \\ m \end{bmatrix} = g(u,v)^T f(m,n)$$

$$= [u,v]A^{-1}A \begin{bmatrix} m \\ n \end{bmatrix} = um + vn = 1,$$

as required. The other two cases are handled in exactly the same way. □

Before stating the conjecture (and we call it Theorem 1.4 now), let's look at the following example (Example 1.3 of [2]), where on the left it is the trinary tree generated by $(3,1)$ up to a depth of 2, and on the right, it is the Bézout tree of $(3,1)$ generated by $(0,1)$ up to the same depth. Note that the defining rule for the second tree is analogous to the first: one proceeds from one node at a given level to three nodes at the next level by applying the functions $g_1(u,v) = g(u,-v), g_2(u,v) = g(u,v)$ and $g_3(u,v) = g(v,u)$.

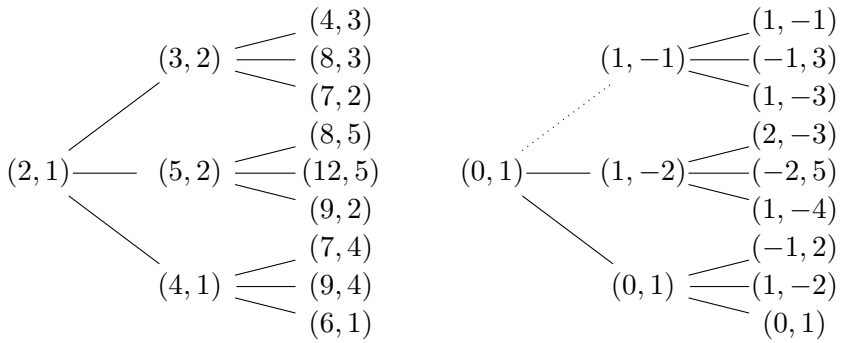
Example 1.2.



Comparing the above two trees shows that the second tree yields the Bézout coefficients for entries in the first tree. This is not completely true for the Bézout tree of $(2,1)$ generated by $(0,1)$ (which is the same as the second tree in the above example). To fix the situation, simply change the top entry in

the second level (i.e. at depth 1) from $(-1, 2)$ into $(1, -1)$, then propagate accordingly using the functions defined by $g_i(u, v), i = 1, 2, 3$. We make this precise by introducing the following trees up to a depth of 2 (cf. Figure 2.1 of [2], where there are typos regarding two entries in the upper right subtree).

Example 1.3.



The dotted line in the second tree above means that the entry $(1, -1)$ does not come from $(0, 1)$ by applying the function $g_1(u, v)$, instead one defines the entry $(1, -1)$ using the Bézout coefficients of the corresponding Pythagorean pair $(3, 2)$. After this modification, it appears that the new Bézout tree yields all the Bézout coefficients of the first tree. Hence the merit of this construction is that (if it is proven to be true) it gives an efficient way to construct the Bézout coefficients for all Pythagorean pairs.

We can state Conjecture 2.1 of [2] (and now a theorem) as follows.

Theorem 1.4. Consider the trinary trees generated by $(2, 1)$ and $(3, 1)$. Let (u, v) be the pair in the Bézout tree corresponding to the relatively prime pair (m, n) and (U, V) be the pair given by the `gcd` function for the same pair (m, n). Then the following hold:

(1) For all (u, v) in the Bézout tree of $(3, 1)$ generated by $(0, 1)$, $(u, v) = (U, V)$.

(2) One third of the (u, v) in the Bézout tree of $(2, 1)$ generated by $(0, 1)$ are not equal to (U, V). Changing the value of $g(0, -1)$ in the second level of the Bézout tree from $(-1, 2)$ to $(1, -1)$ results in a tree in which $(u, v) = (U, V)$ for all (u, v).

The above theorem is clearly implied by the following theorem.

Theorem 1.5. For a relatively prime Pythagorean pair (m, n) with $m > n > 0$, except for $(m, n) = (2, 1)$ and for $f_1(m, n) := f(m, -n)$, the following diagram is commutative, i.e. $\beta(f_i(m, n)) = g_i(\beta(m, n)), i = 1, 2, 3$:

$$\begin{array}{ccc} (m,n) & \xrightarrow{f_i} & (m',n') \\ \downarrow\beta & & \downarrow\beta \\ (r,s) & \xrightarrow{g_i} & (r',s') \end{array}$$

where $\beta(m, n)$ gives the Bézout coefficients (r, s) for (m, n).

We will prove Theorem 1.5 in Section 2 using the standard Euclidean algorithm. In Section 3, we give a generalization (see Theorem 3.2).

2. Euclidean Algorithm and the Proof of Theorem 1.5

For simplicity we will assume that all ordered pairs (m, n) consist of relatively prime integers, even though the result can be generalized to the case when $\gcd(m, n) = d > 1$.

2.1. Euclidean Algorithm

We recall that for relatively prime integers $m > n > 0$, the Division Algorithm is given by

$$m = q_1 n + r_1$$

$$n = q_2 r_1 + r_2$$

$$\cdots$$

$$r_{k-2} = q_k \cdot r_{k-1} + r_k,$$

where $0 < r_1 < n, 0 < r_i < r_{i-1}$ for $i = 2, 3, \cdots, k-1$, $r_{k-1} = \gcd(m, n) = 1$ and $r_k = 0$ if $n > 1$, and $q_1 = m, r_1 = 0$ if $n = 1 = \gcd(m, n)$. We can record the process using matrices as follows:

$$\begin{bmatrix} m \\ n \end{bmatrix} = \begin{bmatrix} q_1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} n \\ r_1 \end{bmatrix} = \begin{bmatrix} q_1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} q_2 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix}$$

$$= \begin{bmatrix} q_1 & 1 \\ 1 & 0 \end{bmatrix} \cdots \begin{bmatrix} q_k & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Similarly for $n > m > 0$ and relatively prime, we have

$$\begin{bmatrix} m \\ n \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} q_1 & 1 \\ 1 & 0 \end{bmatrix} \cdots \begin{bmatrix} q_k & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Note that these intermediate matrices with left upper corner entry q_i or 0 are uniquely determined. Note that the Bézout coefficients (r, s) for the relatively prime pair (m, n) with $m, n > 0$ can be found by backward substitutions from the steps of the division, or what amounts to be the same, from the first row vector of A^{-1}, where A is the matrix

$$A = \begin{bmatrix} q_1 & 1 \\ 1 & 0 \end{bmatrix} \cdots \begin{bmatrix} q_k & 1 \\ 1 & 0 \end{bmatrix}$$

or

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} q_1 & 1 \\ 1 & 0 \end{bmatrix} \cdots \begin{bmatrix} q_k & 1 \\ 1 & 0 \end{bmatrix}.$$

To prove Theorem 1.5, we need a few lemmas.

Lemma 2.2. Assume that $m > n > 0$ and $\gcd(m, n) = 1$. Then (r, s) gives the Bézout coefficients for (m, n), i.e. $(r, s) = \beta(m, n)$ if and only if $(s, r - 2s) = \beta(2m + n, m)$.

Proof. By division algorithm, we can write

$$\begin{bmatrix} m \\ n \end{bmatrix} = A \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

where A is an invertible integer matrix of the form

$$A = \begin{bmatrix} q_1 & 1 \\ 1 & 0 \end{bmatrix} \cdots \begin{bmatrix} q_k & 1 \\ 1 & 0 \end{bmatrix}, \ \det(A) = \pm 1.$$

By the given assumption, A^{-1} is of the form

$$A^{-1} = \begin{bmatrix} r & s \\ * & * \end{bmatrix}.$$

Now we perform the division algorithm for $2m + n$ and m, where the first step is the following:

$$2m + n = 2 \cdot m + n,$$

which is followed by the division of m by n. Hence we can write

$$\begin{bmatrix} 2m+n \\ m \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix} A \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

where the first row of the matrix $\left(\begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix} A \right)^{-1}$ gives the Bézout coefficient (r', s') of the division of $2m+n$ by m. But

$$\left(\begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix} A \right)^{-1} = A^{-1} \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} r & s \\ * & * \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & -2 \end{bmatrix} = \begin{bmatrix} s & r-2s \\ * & * \end{bmatrix},$$

from which we have $(r', s') = (s, r-2s)$ as required. □

Lemma 2.3. Assume that $m > n > 0$ and $\gcd(m, n) = 1$. Then $(r, s) = \beta(m, m-n)$ if and only if $(s, r-s) = \beta(2m-n, m)$.

Proof. Clearly we have $\gcd(m, m-n) = \gcd(2m-n, m) = 1$. Similar to the proof of Lemma 2.2, we may write

$$\begin{bmatrix} m \\ m-n \end{bmatrix} = A \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

where the first row of A^{-1} gives the Bézout coefficients for the division of m by $m-n$, i.e.

$$A^{-1} = \begin{bmatrix} r & s \\ * & * \end{bmatrix}.$$

Performing the first step of the division of $2m-n$ by m, we have

$$2m - n = 1 \cdot m + (m-n),$$

i.e.

$$\begin{bmatrix} 2m-n \\ m \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} m \\ m-n \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} A \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

whence the Bézout coefficients for the division of $2m-n$ by m are given by the first row of the matrix

$$\left(\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} A \right)^{-1} = A^{-1} \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} r & s \\ * & * \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} s & r-s \\ * & * \end{bmatrix},$$

as required. □

Lemma 2.4. Assume that $m > n > 0$, $n < \frac{m}{2}$ and $\gcd(m, n) = 1$. Then

$$\begin{bmatrix} m \\ n \end{bmatrix} = \begin{bmatrix} q_1 & 1 \\ 1 & 0 \end{bmatrix} A \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

and

$$\begin{bmatrix} m \\ m-n \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} q_2' & 1 \\ 1 & 0 \end{bmatrix} A \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

where $q_1 > 1, 1 + q_2' = q_1$ and A is void if $n = 1$.

Proof. By assumption, we can write

$$m = q_1 n + r$$

with $q_1 = \lfloor \frac{m}{n} \rfloor \geq 2$ and $r < n$.

Letting $q_2' = q_1 - 1$, we have

$$m = 1 \cdot (m - n) + n,$$

where $n < m - n$ by assumption, and

$$m - n = (q_1 - 1)n + r = q_2' n + r.$$

Writing

$$\begin{bmatrix} n \\ r \end{bmatrix} = A \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

and expressing the above divisions in terms of matrices, the result is clear. □

Lemma 2.5. Assume that $m > n > 0$, $\gcd(m, n) = 1$ and $(m, n) \neq (2, 1)$. Then $(r, s) = \beta(m, n)$ if and only if $(r + s, -s) = \beta(m, m - n)$.

Proof. Since $(m, n) \neq (2, 1)$, there are only the following two cases to consider.
Case 1: $n < \frac{m}{2}$. Using Lemma 2.4, we see that the division of m by n is described by the procedure

$$\begin{bmatrix} q_1 & 1 \\ 1 & 0 \end{bmatrix} A, \tag{1}$$

if and only if the division of m by $m-n$ is described by the following procedure:

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} q_2' & 1 \\ 1 & 0 \end{bmatrix} A, \tag{2}$$

where

$$q_1 = q_2' + 1.$$

The fact that (r, s) is the Bézout coefficient for division of m by n means precisely that

$$A^{-1} \begin{bmatrix} q_1 & 1 \\ 1 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} r & s \\ * & * \end{bmatrix},$$

while the Bézout coefficient for the division of m by $m-n$ is given by the first row vector of the matrix

$$A^{-1} \begin{bmatrix} q_2' & 1 \\ 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^{-1}.$$

But

$$\begin{aligned} & A^{-1} \begin{bmatrix} q_2' & 1 \\ 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^{-1} \\ &= A^{-1} \begin{bmatrix} q_1 - 1 & 1 \\ 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^{-1} \\ &= A^{-1} \left(\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} q_1 & 1 \\ 1 & 0 \end{bmatrix} \right)^{-1} \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} \\ &= A^{-1} \begin{bmatrix} q_1 & 1 \\ 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} \\ &= \begin{bmatrix} r & s \\ * & * \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} r+s & -s \\ * & * \end{bmatrix}, \end{aligned}$$

which shows the result.

Case 2: $n > \frac{m}{2}$. Assume that $(r, s) = \beta(m, n)$ and $(r', s') = \beta(m, m-n)$. Since $m - n < \frac{m}{2}$, by the result of Case 1, we have

$$r = r' + s' \text{ and } s = -s',$$

which shows that $r' = r + s$ and $s' = -s$, as required. □

Remark 2.6. The above lemma does not hold when $(m,n) = (2,1)$, since $(0,1) = \beta(m,n) = \beta(m, m-n)$ here, but $(r,s) = (0,1) \neq (r+s, -s)$.

Lemma 2.7. Assume that $m > n > 0, \gcd(m,n) = 1$ and $(m,n) \neq (2,1)$. Then $(r,s) = \beta(m,n)$ if and only if $(-s, r+2s) = \beta(2m-n, m)$.

Proof. The mapping $(m,n) \mapsto (2m-n, m)$ can be factored as $(m,n) \mapsto (m, m-n) \mapsto (2m-n, m)$, so by Lemma 2.5 and Lemma 2.3, the corresponding Bézout coefficients are given by $(r,s) \mapsto (r+s, -s) \mapsto (-s, (r+s) - (-s)) = (-s, r+2s)$, as required. □

Proof of Theorem 1.5

The proof for the pair f_1 and g_1 follows from Lemma 2.7. Note that the condition $(m,n) \neq (2,1)$ is precisely used here. The proof for the pair f_2 and g_2 follows from Lemma 2.2. The proof for the pair f_3 and g_3 is essentially the same as that of the previous case. Here are the details. Let

$$\begin{bmatrix} m \\ n \end{bmatrix} = \begin{bmatrix} q_1 & 1 \\ 1 & 0 \end{bmatrix} A \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

where the first step of the division process is written out. Now for the division of $2n+m$ by n, one has

$$\begin{bmatrix} 2n+m \\ n \end{bmatrix} = \begin{bmatrix} q_1+2 & 1 \\ 1 & 0 \end{bmatrix} A \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Let (r,s) be the Bézout coefficients for the division of m by n, i.e.

$$\begin{bmatrix} r & s \\ * & * \end{bmatrix} = A^{-1} \begin{bmatrix} q_1 & 1 \\ 1 & 0 \end{bmatrix}^{-1}.$$

If (r', s') is the Bézout coefficients for the division of $2n+m$ by n, then

$$\begin{aligned} \begin{bmatrix} r' & s' \\ * & * \end{bmatrix} &= A^{-1} \begin{bmatrix} q_1+2 & 1 \\ 1 & 0 \end{bmatrix}^{-1} \\ &= A^{-1} \left(\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} q_1 & 1 \\ 1 & 0 \end{bmatrix} \right)^{-1} \\ &= A^{-1} \begin{bmatrix} q_1 & 1 \\ 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}^{-1} \end{aligned}$$

$$= \begin{bmatrix} r & s \\ * & * \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} r & s-2r \\ * & * \end{bmatrix},$$

so

$$r' = r, s' = s - 2r,$$

as required. □

3. A Generalization

We first extend the definition of Bézout coefficients to general ordered pairs $(m, n) \in \mathbb{Z} \times \mathbb{Z}$. The following definition seems to yield the same output as the MATLAB's function $[G, U, V] = \texttt{gcd}(m, n)$ [3] or SAGE's `xgcd` function [6]. We have tested this by writing a Sage script using the following definitions for relatively prime (m, n) up to a reasonable size. In any case, our proof will be based on the following definitions.

Definition 3.1. The Bézout coefficients $\beta(a, b)$ for an ordered pair $(a, b) \in \mathbb{Z} \times \mathbb{Z}$ are defined as follows:

3.1.1. For $a > b > 0$, $\beta(a, b) = (r, s)$ with $ra + sb = \gcd(a, b)$ is given by the Euclidean algorithm which is uniquely determined. One writes this as $(a, b) \mapsto (\gcd(a, b), r, s)$.

This is extended to all ordered pairs by the following rules:

3.1.2. $(0, a) \mapsto (|a|, 0, \text{sign}(a))$, where $\text{sign}(a)$ is the sign of a, and by convention $\text{sign}(0) = 0$.

3.1.3. $(\pm a, a) \mapsto (|a|, 0, \text{sign}(a))$.

3.1.4. If $|a| \neq |b|$ and $\beta(|a|, |b|) = (r, s)$, then $(a, b) \mapsto (\gcd(a, b), \text{sign}(a)r, \text{sign}(b)s)$.

3.1.5. If $|a| \neq |b|$ and $(a, b) \mapsto (\gcd(a, b), r, s)$, then $(b, a) \mapsto (\gcd(a, b), s, r)$.

We leave the readers to check that these formulas are consistent.

Theorem 3.2. Let A be a unimodular 2×2 matrix, i.e. an integer matrix such that $\det(A) = \pm 1$. Consider the mappings

$$f\left(\begin{bmatrix} m \\ n \end{bmatrix}\right) := A \begin{bmatrix} m \\ n \end{bmatrix}$$

and

$$g\left(\begin{bmatrix} m \\ n \end{bmatrix}\right) := (A^{-1})^T \begin{bmatrix} m \\ n \end{bmatrix}$$

for $(m, n) \in \mathbb{Z} \times \mathbb{Z}$ such that $\gcd(m, n) = 1$. Then with only finitely many exceptions of relatively prime ordered pairs (m, n), one has $g(\beta(m, n)) = \beta(f(m, n))$, where β maps an ordered pair $(m, n) \in \mathbb{Z} \times \mathbb{Z}$ to its Bézout coefficients.

Proof. The idea of the proof is that if the group of unimodular 2×2 matrices, denoted $\mathrm{GL}_2(\mathbb{Z})$, is finitely generated, then we can decompose each element A in the group as a product of its generators and their inverses (let $\mathcal{S} = \{U_1, U_1^{-1}, \cdots, U_r, U_r^{-1}\}$ be a set of generators of $\mathrm{GL}_2(\mathbb{Z})$ and their inverses), say

$$A = V_k V_{k-1} \cdots V_1,$$

where each $V_i \in \mathcal{S}, i = 1, \cdots, k$. Since

$$(A^{-1})^T = (V_k^{-1})^T (V_{k-1}^{-1})^T \cdots (V_1^{-1})^T,$$

the proof of compatibility of Bézout coefficients of relatively prime ordered pairs under the transformation A is reduced to the simple case when $A = V_i$, where V_i is in the set of generators, and we check the relation $\beta(f(m, n)) = g(\beta(m, n))$ for f defined by V_i and for g defined by $(V_i^{-1})^T$. This is because for a factorization A into the generators (so A is a series of compositions of the generator functions), if compatibility holds at each step of the successive composition with a finite number of exceptions, then it is easy to see that there will be only finitely many exceptions for the final composite function, which is A (we illustrate this in Example 3.3). Now we start to prove the result for its generators.

It is well known [1] that $\mathrm{GL}_2(\mathbb{Z})$ is generated by

$$\mathrm{U} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \mathrm{S} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \text{ and } \mathrm{T} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

It is easy to check that the sets of exceptional pairs for U, S and S^{-1} are all given by

$$\{(-1, -1), (-1, 1), (1, -1), (1, 1)\}.$$

To determine the exceptional set E_{T} for T, we check first the following special cases (m, n) such that

$$(m, n) \in \{(0, \pm 1), (\pm 1, 0), (\pm 1, \pm 1), (\pm 1, \mp 1)\}$$

or

$$(m+n,n) \in \{(0,\pm 1),(\pm 1,0),(\pm 1,\pm 1),(\pm 1,\mp 1)\}.$$

This gives exceptional ordered pairs $(\pm 1, 0)$, for which $g(\beta(m,n)) \neq \beta(f(m,n))$.

For the remaining cases, we may assume that

$$(|m|,|n|),(|m+n|,|n|) \notin \{(1,0),(0,1),(1,1)\}.$$

After excluding the special cases above, we use 3.14 and 3.15 to reduce the checking to the following cases, noting that $\beta(-m,-n) = -\beta(m,n)$ and $\beta(-m-n,-n) = -\beta(m+n,n)$.

Case: $n > m > 0$. Let

$$\begin{bmatrix} m \\ n \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} A \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} m+n \\ n \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} A \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Then

$$\begin{bmatrix} r' & s' \\ * & * \end{bmatrix} = A^{-1} \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^{-1} = A^{-1} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} r & s \\ * & * \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} r & s-r \\ * & * \end{bmatrix},$$

where $(r,s) = \beta(m,n)$ and $(r',s') = \beta(m+n,n)$.

Case: $m > n > 0$.

Let

$$\begin{bmatrix} m \\ n \end{bmatrix} = \begin{bmatrix} q_1 & 1 \\ 1 & 0 \end{bmatrix} A \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Then

$$\begin{bmatrix} r & s \\ * & * \end{bmatrix} = A^{-1} \begin{bmatrix} q_1 & 1 \\ 1 & 0 \end{bmatrix}^{-1},$$

$$\begin{bmatrix} m+n \\ n \end{bmatrix} = \begin{bmatrix} q_1+1 & 1 \\ 1 & 0 \end{bmatrix} A \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

hence

$$\begin{bmatrix} r' & s' \\ * & * \end{bmatrix} = A^{-1} \begin{bmatrix} q_1+1 & 1 \\ 1 & 0 \end{bmatrix}^{-1}$$

$$= A^{-1} \begin{bmatrix} q_1 & 1 \\ 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} r & s \\ * & * \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} r & s-r \\ * & * \end{bmatrix}.$$

It follows that for both of the above cases,

$$\begin{bmatrix} r' \\ s' \end{bmatrix} = \left(\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^{-1} \right)^T \begin{bmatrix} r \\ s \end{bmatrix},$$

as required.

Case: $m > 0, n < 0$. Let $n = -n'$.

Subcase 1: $m < n'$.

Tracing the relations (here and in what follows, the notation $\leftrightarrow$ means a one-to-one correspondence that can be determined from Definition 3.1)

$$\begin{bmatrix} m \\ n \end{bmatrix} \leftrightarrow \begin{bmatrix} m \\ n' \end{bmatrix} \leftrightarrow \begin{bmatrix} n' \\ m \end{bmatrix}$$

and

$$\begin{bmatrix} m+n \\ n \end{bmatrix} \leftrightarrow \begin{bmatrix} n'-m \\ n' \end{bmatrix} \leftrightarrow \begin{bmatrix} n' \\ n'-m \end{bmatrix},$$

it suffices to find the relation

$$\begin{bmatrix} n' \\ m \end{bmatrix} \leftrightarrow \begin{bmatrix} n' \\ n'-m \end{bmatrix}.$$

When $(n', m) \neq (2, 1)$ (i.e. when $(m, n) \neq (1, -2)$), this can be determined by Lemma 2.5. Using 3.1.4, 3.1.5 and Lemma 2.5, we find the same relation between (r', s') and (r, s) as above. The special cases $(m, n) = (\pm 1, \mp 2)$ are checked directly to be exceptional.

Subcase 2: $m > n'$.

Tracing the relations

$$\begin{bmatrix} m \\ n \end{bmatrix} \leftrightarrow \begin{bmatrix} m \\ n' \end{bmatrix}$$

and

$$\begin{bmatrix} m+n \\ n \end{bmatrix} \leftrightarrow \begin{bmatrix} m-n' \\ n' \end{bmatrix},$$

it suffices to find the relation

$$\begin{bmatrix} m \\ n' \end{bmatrix} \leftrightarrow \begin{bmatrix} m-n' \\ n' \end{bmatrix}.$$

We have

$$\begin{bmatrix} m \\ n' \end{bmatrix} = \begin{bmatrix} q_1'+1 & 1 \\ 1 & 0 \end{bmatrix} A \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

and

$$\begin{bmatrix} m-n' \\ n' \end{bmatrix} = \begin{bmatrix} q_1' & 1 \\ 1 & 0 \end{bmatrix} A \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

where $q_1' = 0$ if $m < 2n'$ and $q_1' \geq 1$ if $m \geq 2n'$. As a result, we find the same relation as above. In summary, for the transformation T, the set E_{T} of exceptional cases is given by

$$E_{\mathrm{T}} = \{(-1,0),(1,0),(1,-2),(-1,2)\}.$$

Similarly the exceptional set $E_{\mathrm{T}^{-1}}$ is given by

$$E_{\mathrm{T}^{-1}} = \{(-1,-2),(-1,0),(1,0),(1,2)\}.$$

This concludes the proof. □

Example 3.3. The factorization of

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} = \mathrm{T}^2$$

$$\begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix} = \mathrm{T}^2\mathrm{U}$$

and

$$\begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix} = \mathrm{T}^2\mathrm{S}$$

allows us to determine the exceptional set of these transformations. For example, using the proof of the above theorem, let's determine the exceptional set $E_{\mathrm{T}^2\mathrm{S}}$ for the transformation $\begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}$, which is described by the following process:

$$\begin{array}{ccccccc} (m,n) & \to & \mathrm{S}(m,n) & \to & \mathrm{T}(\mathrm{S}(m,n)) & \to & \mathrm{T}(\mathrm{TS}(m,n)) \\ \text{(a)} & \to & \text{(b)} & \to & \text{(c)} & \to & \text{(d)} \end{array},$$

where compatibility of Bézout coefficients can fail at (a) for ordered pairs in the exceptional set E_{S} of S, or at (b) for ordered pairs in the exceptional set E_{T} of T, or at (c) for ordered pairs in the exceptional set E_{T} of T. Taking preimage of these exceptional sets to the beginning step (a), we see that the compatibility of Bézout coefficients for $\mathrm{T}^2\mathrm{S}$ can only possibly fail for ordered pairs in the set

$$E_{\mathrm{S}} \cup \mathrm{S}^{-1}E_{\mathrm{T}} \cup (\mathrm{TS})^{-1}E_{\mathrm{T}}.$$

By direct checking, this turns out to be all the exceptional cases, i.e.

$$E_{\mathrm{T}^2\mathrm{S}} = \{(-2,-3),(-2,-1),(-1,-1),(-1,1),(0,-1),(0,1),(1,-1),(1,1),(2,1),(2,3)\}.$$

Similarly, we get that

$$E_{\mathrm{T}^2} = \{(-3,2),(-1,0),(-1,2),(1,-2),(1,0),(3,-2)\}$$

and

$$E_{\mathrm{T}^2\mathrm{U}} = \{(-2,1),(-2,3),(-1,-1),(-1,1),(0,-1),(0,1),(1,-1),(1,1),(2,-3),(2,-1)\}.$$

References

[1] T. Apostol, Modular functions and Dirichlet series in number theory, Springer-Verlag, 1976.

[2] E. Gullerud and J.S. Walker, Generalized Bézout trees for Pythagorean pairs, arXiv:1803.04875v1 [math.NT] 13 Mar 2018.

[3] Matlab. A language and environment for technical computing. Product of MathWorks.

[4] I. Niven, H.S. Zuckerman, and H.L. Montgomery, An introduction to the theory of numbers, 5th edition, John Wiley & Sons, Inc.

[5] T. Randall and R. Saunders, "The family tree of the Pythagorean triplets revisited." *Math. Gaz.* (**78**) (482), pp. 190-193, 1994.

[6] William A. Stein et al. Sage Mathematics Software (Version 6.10), http://www.sagemath.org.