ELEMENTARY PROPERTIES OF THE SUBTRACTIVE EUCLIDEAN ALGORITHM

Arnold Knopfmacher
University of the Witwatersrand, P O Wits 2050, Johannesburg, South Africa
(Submitted May 1990)

1. Introduction

In a recent article in this journal, T. Moore [4] used a microcomputer to make a study of the length of the Euclidean algorithm in determining the greatest common divisor of two nonzero integers m, n. Our intention is to make a similar study of the lengths of the *subtractive Euclidean algorithm*. Recall that to determine for example gcd(11, 3) by the subtractive algorithm we perform the operations:

$$11 - 3 = 8,\ 8 - 3 = 5,\ 5 - 3 = 2,\ 3 - 2 = 1,\ 2 - 1 = 1,\ 1 - 1 = 0;$$

a total of six steps. By contrast, the ordinary Euclidean algorithm yields

$$\begin{aligned} 11 &= (3)(3) + 2 \\ 3 &= (1)(2) + 1 \\ 2 &= (2)(1), \end{aligned}$$

in only three steps. However, if we use the Euclidean algorithm to express 11/3 as a regular continued fraction,

$$\frac{11}{3} = 3 + \frac{1}{1 + \frac{1}{2}} \equiv [3,\ 1,\ 2],$$

we notice that the partial quotient 3 corresponds to the number of subtractions of 3 above, etc. In general, it is easy to see that the length of the subtractive Euclidean algorithm for (m, n) is equal to the sum of the partial quotients in the regular continued fraction expansion of m/n.

2. Analysis

Following the approach of T. Moore [4], we begin our investigation by representing the pair of integers m, n as a lattice point (m, n) in the plane and plotting this point only if it has a subtractive length equal to the fixed value in which we are interested. In view of the equivalence of the subtractive length to the sum of the continued fraction partial quotients for (m, n), we can implement the subtractive algorithm computations merely by changing line 280 in the basic program given by Moore [4, fig. 1] for the Euclidean algorithm to read

```
280   DC = DC + INT(N1/M1).
```

It is also necessary to swap m and n in the case $m > n$. The graphic results from four different choices of subtractive lengths are shown in Figure 1 below for all pairs (m, n) belonging to the range $-320 \le m \le 320$, $-175 \le n \le 175$. The range of coordinates here is a consequence of the EGA resolution of an IBM compatible computer.

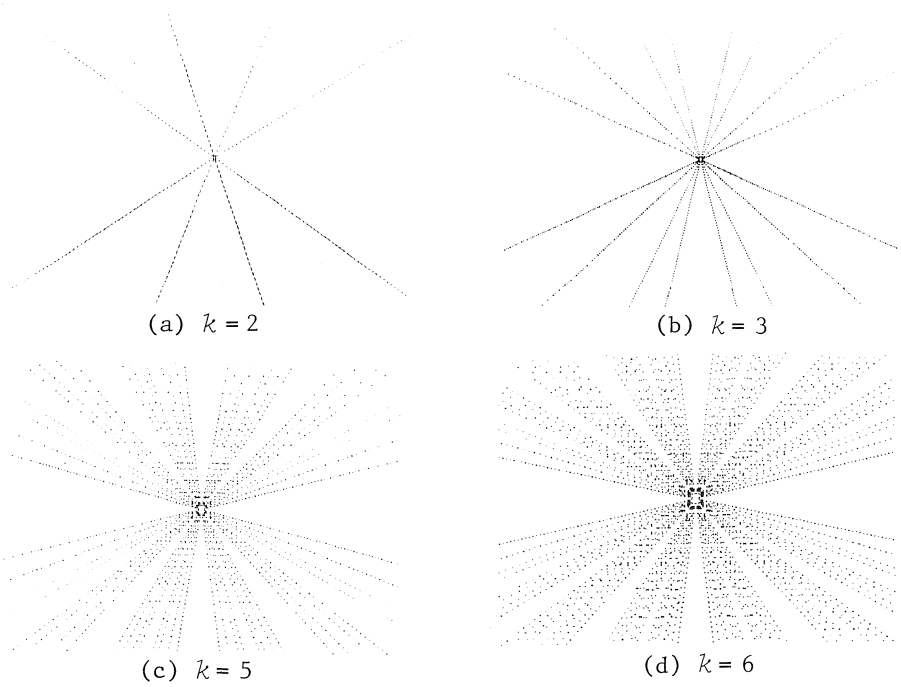

Figure 1

Screendumps showing all integer pairs (m, n) in the range $-320 \le m \le 320$, $-175 \le n \le 175$, whose gcds are obtained in exactly k steps of the subtractive Euclidean algorithm

In contrast to the output from the ordinary Euclidean algorithm [4, fig. 2], the patterns shown here are surprisingly regular. In each case, the pairs (m, n) having subtractive length k are seen to lie on 2^k straight lines. Since

$$\gcd(\pm m, \pm n) = \gcd(m, n) = \gcd(n, m),$$

we can restrict our attention to the pairs $m \ge 1$, $n \ge 1$, and $m \ge n$. A mathematical description of these pictures is then given by the following theorem.

Theorem 1: For any fixed integer $k \ge 1$, the pair of coprime positive integers (m, n) with $m \ge n$ has subtractive length k iff $m/n = [a_1, a_2, \ldots, a_r]$, where

$$a_1 + a_2 + \cdots + a_r = k,\ a_i \ge 1,\ 1 \le i \le r - 1 \text{ and } a_r \ge 2.$$

Furthermore, there are 2^{k-2} such coprime pairs which, together with their multiples and symmetry, make up the lattice points lying on the 2^k lines in the corresponding diagram.

Proof: If (m, n) has subtractive length k, then, by previous remarks,

$$\frac{m}{n} = [a_1, a_2, \ldots, a_r] \text{ where } a_1 + a_2 + \cdots + a_r = k,\ a_i \ge 1,\ 1 \le i \le r.$$

In addition, $a_r \ge 2$ since a value $a_r = 1$ corresponds to the final step

$$r_{k-1} = (1) \times r_k + 0$$

in the Euclidean algorithm, which contradicts $0 < r_k < r_{k-1}$. Since

$$\frac{jm}{jn} = \frac{m}{n},\ j \ge 1,$$

we can restrict our attention to m, n coprime; the multiples jm, jn thus give the other integer points lying on the line determined by (m, n). The number of lines is therefore determined by the number of solutions in positive integers, $a_1, \ldots, a_r$ of the equation $a_1 + \cdots + a_r = k$ where $a_r \geq 2$. Since any solution with $a_r = 1$ can be paired uniquely to a solution

$$a_1 + \cdots + (a_{r-1} + 1) = k, \text{ with } a_{r-1} + 1 \geq 2,$$

the quantity we require is half of the total number of solutions to $a_1 + \cdots + a_r = k$ in positive integers. Now, for fixed r, the number of such solutions is $\binom{k-1}{r-1}$ (see, e.g., Brualdi [1, p. 38]). Since r can take on any value from 1 to k, the total number of solutions is

$$\sum_{r=1}^{k} \binom{k-1}{r-1} = \sum_{r=0}^{k-1} \binom{k-1}{r} = 2^{k-1},$$

and the result follows.

It is interesting to note that for each k various integer pairs consisting of Fibonacci and Lucas numbers occur among the coprime pairs with subtractive length k. In particular, the pairs

$$(F_{k+1}, F_k), (F_{k+1}, F_{k-1}), (L_{k-1}, L_{k-2}), (L_{k-1}, L_{k-3}), (L_{k-1}, F_k)$$

are included in the set. To see this, we can use the recurrence relationships for the Fibonacci and Lucas numbers to derive the following continued fraction expansions, the partial quotients of which sum in each case to k:

$$\frac{F_{k+1}}{F_k} = [1, \ldots, 1, 2] \quad (k - 2 \text{ ones}), \quad k \geq 2,$$

$$\frac{F_{k+1}}{F_{k-1}} = [2, 1, \ldots, 1, 2] \quad (k - 4 \text{ ones}), \quad k \geq 4,$$

$$\frac{L_{k-1}}{L_{k-2}} = [1, \ldots, 1, 3] \quad (k - 3 \text{ ones}), \quad k \geq 3,$$

$$\frac{L_{k-1}}{L_{k-3}} = [2, 1, \ldots, 1, 3] \quad (k - 5 \text{ ones}), \quad k \geq 5,$$

$$\frac{L_{k-1}}{F_k} = [1, 2, 1, \ldots, 1, 2] \quad (k - 5 \text{ consecutive ones}), \quad k \geq 5.$$

In addition, it is well known that among the pairs (m, n), with $m \geq n$, that require k steps of the ordinary Euclidean algorithm, the Fibonacci pair (F_{k+1}, F_k) is the smallest. By contrast, we show (F_{k+1}, F_k) is the largest coprime pair that has subtractive length k. [The smallest such pair is, of course, $(k, 1)$.]

To see this, suppose inductively that F_k/F_{k-1} is the largest pair requiring $k - 1$ subtractive steps. Now to each of the 2^{k-3} positive integer pairs (c_{k-1}, d_{k-1}), with $c_{k-1} \geq d_{k-1}$ with subtractive length $k - 1$, we can associate two of the 2^{k-2} pairs of subtractive length k, namely, (A_k, B_k) where

$$\frac{A_k}{B_k} = 1 + \frac{c_{k-1}}{d_{k-1}} = \frac{c_{k-1} + d_{k-1}}{d_{k-1}}$$

and (A_k', B_k') where

$$\frac{A_k'}{B_k'} = 1 + \frac{1}{c_{k-1}/d_{k-1}} = \frac{c_{k-1} + d_{k-1}}{c_{k-1}}.$$

By our inductive hypothesis, the largest pairs of the forms (A_k, B_k) and (A_k', B_k') will be F_{k+1}/F_{k-1} and F_{k+1}/F_k, respectively. The latter pair gives the result.

3. Estimates for Almost All Pairs

We can use the above results to derive some elementary bounds for lengths of the subtractive Euclidean algorithm valid for almost all pairs (m, n) with $1 \le n \le x$, $1 \le m \le x$, as $x \to \infty$. For convenience, we denote the subtractive length for the pair (m, n) by $L(m, n)$ and the set of all x^2 pairs (m, n) with $1 \le m \le x$, $1 \le n \le x$ by $S(x)$. We first show that the proportion of pairs in $S(x)$ for which $c \log_2 x < L(m, n) \le x$, $\to 1$ as $x \to \infty$ for any $0 < c < 1$.

For any fixed positive integer k, the pairs in $S(x)$ with subtractive length k lie on at most 2^{k-1} straight lines. Each such line contains at most x such pairs. It follows that for any $m \in \mathbb{N}$, the number of pairs in $S(x)$ with subtractive length $\le m$ is not greater than $\sum_{k=1}^{m} 2^{k-1}x = x(2^m - 1)$. Thus, the proportion of pairs with subtractive length $\le m$ is bounded above by $2^m/x$. This tends to zero as $x \to \infty$ provided $m < c \log_2 x$, for any $0 < c < 1$.

If we consider only coprime pairs in $S(x)$, then the corresponding result is as follows: The proportion of coprime pairs in $S(x)$ for which $c \log_2 x < L(m, n) < x$, $\to 1$ as $x \to \infty$ for any $0 < c < 2$. In this case, the number of coprime pairs with subtractive length $\le m$ is at most $\sum_{k=1}^{m} 2^{k-1} = 2^m - 1$. Now, by Theorem 330 of Hardy & Wright [2], the number of coprime pairs in $S(x)$ is asymptotically

$$\frac{6x^2}{\pi^2} + O(x \log x) \text{ as } x \to \infty.$$

Hence, the proportion of coprime pairs in $S(x)$ with subtractive length $\le m$ is bounded above by

$$\frac{\pi^2}{6}\,\frac{2^m}{x^2} + O\left(\frac{\log x}{x^3}\right) \text{ as } x \to \infty,$$

which tends to zero, provided $m < c \log_2 x$ for any $0 < c < 2$.

4. Final Remarks

A graphical representation led us to various observations as well as estimates for the length of the subtractive Euglidean algorithm by elementary means. By a much deeper analytical approach, Knuth & Yao [3] have shown that for fixed m the average length of the subtractive algorithm over all pairs (m, n) with $1 \le n \le m$ is

$$6\pi^{-2}(\ln m)^2 + O(\ln m(\ln\ln m)^2).$$

References

1. R. A. Brualdi. *Introductory Combinatorics*. New York: North Holland, 1977.
2. G. H. Hardy & E. M. Wright. *An Introduction to the Theory of Numbers*. 5th ed. Oxford: Oxford University Press, 1979.
3. D. E. Knuth & A. C. C. Yao. "Analysis of the Subtractive Algorithm for Greatest Common Divisors." *Proc. Nat. Acad. Sci. U.S.A.* *72.12* (1975):4720-22.
4. T. E. Moore. "Euclid's Algorithm and Lamé's Theorem on a Microcomputer." *Fibonacci Quarterly* *27.4* (1989):290-95.

*****