EUCLID'S ALGORITHM AND LAMÉ'S THEOREM ON A MICROCOMPUTER

Thomas E. Moore
Bridgewater State College, Bridgewater, MA 02324
(Submitted July 1987)

To the memory of my friend and colleague Hugo D'Alarcao.

1. Introduction

We denote the greatest common divisor of two nonzero integers m and n by $\gcd(m, n)$. Since it is true that $\gcd(m, n) = \gcd(\pm m, \pm n)$, and since $\gcd(m, n) = \gcd(n, m)$, we may assume that both m and n are positive and $m \leq n$.

The Euclidean algorithm for computing $\gcd(m, n)$ is a familiar process of iterated long division which can be written as follows:

$$\begin{aligned}
n &= mq_1 + r_1;\ 0 < r_1 < m,\\
m &= r_1q_2 + r_2;\ 0 < r_2 < r_1,\\
r_1 &= r_2q_3 + r_3;\ 0 < r_3 < r_2,\\
&\vdots\\
r_{k-2} &= r_{k-1}q_k + r_k;\ 0 < r_k < r_{k-1},\\
r_{k-1} &= r_kq_{k+1} + r_{k+1};\ r_{k+1} = 0.
\end{aligned}$$

The process halts when a remainder 0 is obtained and then $\gcd(m, n)$ is the divisor r_k in the last step of division.

A theorem of Gabriel Lamé (1795-1870) asserts that the number of divisions required to find $\gcd(m, n)$ by Euclid's algorithm is no more than five times the number of digits (base 10) in the smaller of m and n. For proofs see [1] and [2].

Our idea is to keep a count of the number of divisions required to produce $\gcd(m, n)$ by Euclid's algorithm, for a range of values of m and n, and to study the distribution of these numbers.

2. Implementation

The actual computations were accomplished using a BASIC program (written for the APPLE II computers but easily modified for other equipment). The program is listed in Figure 1.

In this program, the variable DC represents a division count, that is, the number of steps of division in using Euclid's algorithm to obtain $\gcd(m, n)$.

The program actually calculates both $\gcd(m, n)$ and $\gcd(n, m)$ within the nested loops of lines 140-230 and, while the second computation is redundant, we have chosen to allow it because it gives us a program that should be easier to follow than otherwise. The program is also fairly slow to execute, and a compiled version of it is preferred.

```
100   REM  DYNAMIC VIEW OF LAME'S THM
110   PRINT "PLOT WHAT DIV. COUNT "
    : INPUT CH
    : REM  USER CHOICE
120   HGR2
130   H = 140
    : V = 95
140   FOR M = 1 TO H
150   FOR N = 1 TO V
160   DC = 0
170   IF M > N THEN DC = - 1
180   GOSUB 240
190   IF DC < > CH THEN 220
200   HCOLOR= 3
210   HPLOT M + H,V - N
    : HPLOT M + H,V + N
    : HPLOT H - M,V + N
    : HPLOT H - M,V - N
220   NEXT N
230   NEXT M
240   REM  SUBROUTINE FOR GCD VIA EUCLID
250   M1 = M
260   N1 = N
270   R = N1 - M1 * INT (N1 / M1)
280   DC = DC + 1
290   N1 = M1
300   M1 = R
310   IF R > 0 THEN 270
320   RETURN
330   END
```

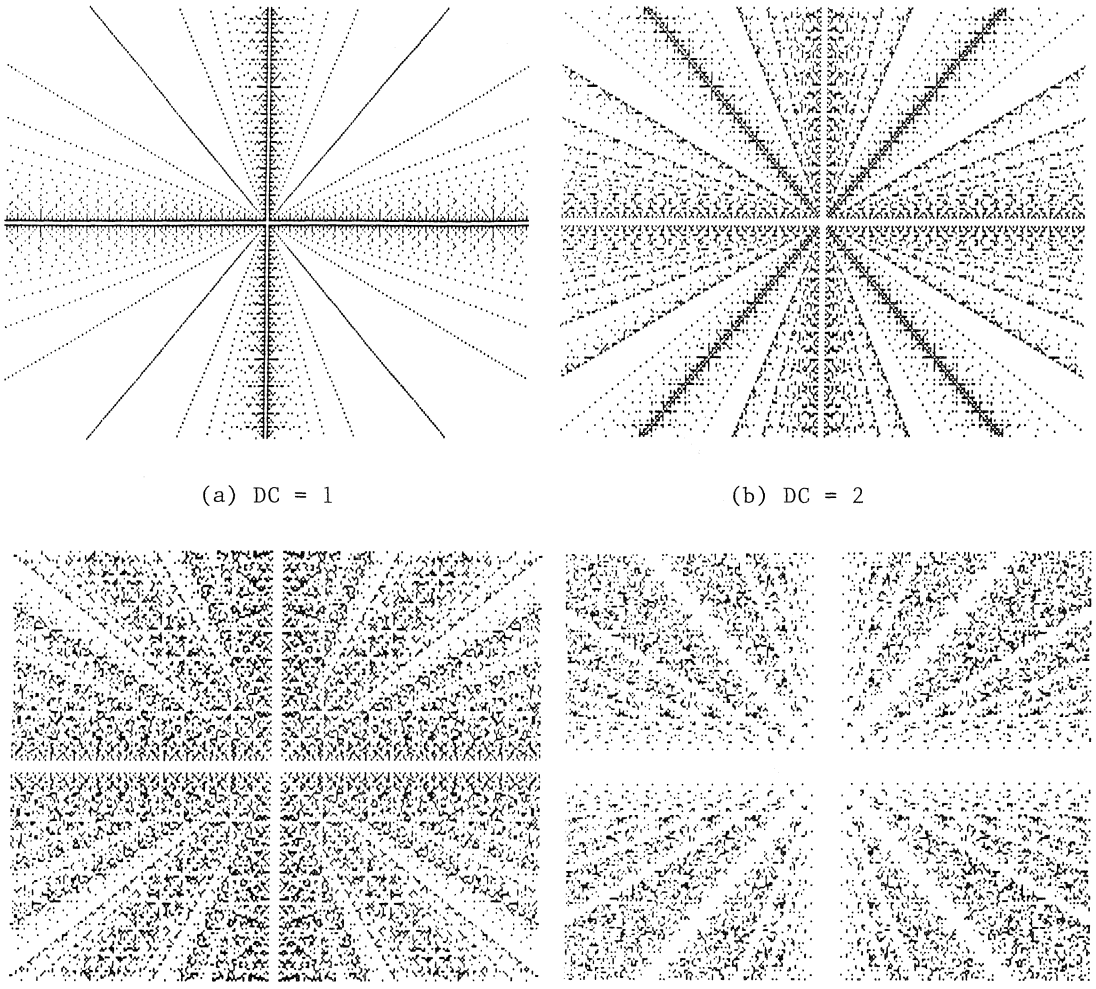
FIGURE 1

The graphics display capability of the computer with a monitor suggested that we interpret each pair of integers m and n as a lattice point (m, n) in the plane and that we plot or do not plot this point on the monitor screen according to the value DC obtained in finding gcd(m, n). Thus, the program asks the user to declare the value of DC in which he is interested.

From the observation that the values of gcd$(\pm m, \pm n)$ are all equal, we note that a fourfold symmetry can be achieved if the display includes all four quadrants. Hence, the origin (0, 0) is translated to screen coordinates (140, 95) and all subsequently lit points are, similarly, translates of the actual $(\pm m, \pm n)$.

The screen images resulting from four different choices of division counts are shown in Figure 2. In each case, the range of positive integers for which gcd(m, n) is computed and a division count kept is $1 \le m \le 140$, $1 \le n \le 95$. (These bounds were determined by the graphics page of memory HGR2 in the APPLE II and by the decision to display four quadrants.)

Figure 2 not only illustrates the expected symmetry but also shows patterns of distribution that invite further investigation.



(a) DC = 1 (b) DC = 2

(c) DC = 3 (d) DC = 5

FIGURE 2

Screen dumps showing integer pairs (m, n) in the range $-140 \le m \le 140$, $-95 \le n \le 95$ whose gcd has been obtained by Euclid's algorithm in the same number of steps

3. Analysis

Consider the displays in Figure 2 and the striking fact that the plotted points arrange themselves along various lines. For example, in Figure 2(a) these are the lines in the x-y plane with the equations $y = kx$ and $y = (1/k)x$, for integers k: $k \neq 0$.

Indeed, if gcd(m, n) is found in one step, then m divides n (recall $m \le n$) and, if $n = mq$, then the point (m, n) is on $y = qx$. Since $x = (1/q)y$, then (n, m) is on $y = (1/q)x$.

Again in Figure 2(a), scanning it in the direction of increasing x, we can observe vertical segments at x-values that are multiples of 6 and still longer segments at multiples of 12. For example, at $x = 60$ (see Fig. 3), we note that

the plotted points are $(60, \pm k)$ for $k = 1, 2, \ldots, 6$, and clearly $\gcd(60, \pm k)$ is accomplished in one step by Euclid's algorithm. We point out here that $60 = \text{lcm}(1, 2, \ldots, 6)$ and that similar vertical segments will occur at all x such that $x = \text{lcm}(1, 2, \ldots, m)$.

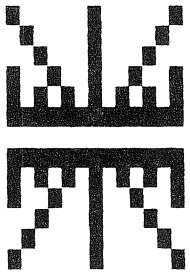


FIGURE 3

Still in Figure 2(a), there is also an X-shape of plotted points at the locations where x is a multiple of 6. The arms of the X-shape at $x = 60$ (see Fig. 3), for example, are just the pairs $(60 \pm k, \pm k)$, for $k = 1, 2, \ldots, 6$, whose greatest common divisor is obtained in one step.

In Figure 2(b), if we scan along the line $y = x$ in the first quadrant, then we can observe +-shapes centered on this line at x-values that are once more multiples of 6. For example, locating (60, 60) as the unplotted (white) center of one such shape (see Fig. 4), we find that this shape is the collection of points $(60, 60 \pm k)$ and $(60 \pm k, 60)$, for $k = 1, 2, \ldots, 6$. Each pair has corresponding $\gcd(m, n)$ obtained in two steps by Euclid's algorithm, as follows:

$$60 + k = (60)(1) + k;\ 0 < k < 60,$$
$$60 = (k)(60/k) + 0$$

or

$$60 = (60 - k)(1) + k;\ 0 < k < 60 - k,$$
$$60 - k = (k)((60 - k)/k) + 0.$$

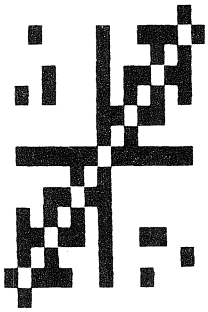


FIGURE 4

Another strongly recurring visual element in Figure 2(b) are blocks of four consecutively plotted horizontal or vertical points. In quadrant one, these occur at the points $(12a + k, 12)$ and $(12, 12a + k)$ for $a \geq 1$, $k = 1, 2, 3, 4$. For these integer pairs, Euclid's algorithm is done in two steps as follows:

$$12a + k = (12)(a) + k;\ 0 \le k < 12,$$

$$12 = (k)(12/k) + 0.$$

There are other discernible patterns in these figures, such as in Figure 2(c) where a pattern of mostly white lines parallel to the axes defines an irregular grid. What is behind it? What is the rule for spacing between successive lines? The interested reader may pursue this line of questioning.

4. Cyclic Behavior

In another direction, we study the distribution of the values DC for fixed $m \ge 1$ and $n \ge m$.

Example 1: $m = 4$.

n	4	5	6	7	8	9	10	11	12	13	14	15 ...
DC	1	2	2	3	1	2	2	3	1	2	2	3 ...

That is, the values of DC for consecutive $n \ge 4$ form the cycle (1223) of length 4.

Example 2: $m = 5, 6, 7$.

	n	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20 ...
($m = 5$)	DC	1	2	3	4	3	1	2	3	4	3	1	2	3	4	3	1 ...
($m = 6$)	DC		1	2	2	2	3	3	1	2	2	2	3	3	1	2	2 ...
($m = 7$)	DC			1	2	3	3	4	4	3	1	2	3	3	4	4	3 ...

In each case, the values DC form a cycle of length m. In fact, we find this is easy to prove generally.

Theorem: For fixed $m \ge 1$ and integers $n \ge m$, let DC be the number of steps required to find $\gcd(m, n)$ by Euclid's algorithm. Then the successive values of DC form a cycle of length m.

Proof: Let r be fixed, $0 \le r < m$. It is sufficient to prove that the values DC are the same for the computations of $\gcd(m, m + r)$ and $\gcd(m, km + r)$ for all integers $k \ge 1$. This follows at once from the initial division in each computation. The former begins

$$m + r = (m)(1) + r$$

and the latter begins

$$km + r = (m)(k) + r.$$

Thus, in each case, the second step of division and all succeeding steps are correspondingly equal.

Corollary: If $\gcd(m, n)$ is accomplished in s steps by Euclid's algorithm, then $\gcd(m, n + km)$ is accomplished in s steps for all integers $k \ge 1$.

Example 3: It is well known that gcd(89, 144) takes 10 steps of division, the maximum predicted by Lamé's theorem. It follows that infinitely many integers can be paired to 89 in this way, namely, the integers $144 + 89k$, and each gcd computation takes 10 steps.

5. Queries and Conclusion

The cycles for all $m \geq 3$ necessarily have the form (12...3), with the remaining DC values of the cycle showing considerable variety. We ask for a rule in terms of m and the position of a value within the cycle that will deliver this value. We have also observed that DC values can be consecutively repeated within a cycle. Is there a rule governing this? Specifically, for a given value DC and any positive integer k, is there a cycle such that DC is repeated consecutively k times?

The microcomputer has been used to gain insight into both the Euclidean algorithm and Lamé's theorem. More can be gained, and some directions to pursue have been given.

References

1. J. V. Uspensky & M. A. Heaslet. *Elementary Number Theory*. New York and London: McGraw-Hill, 1939.
2. D. E. Thoro. "The Euclidean Algorithm II." *Fibonacci Quarterly* 2 (1964): 135-137.

*****