Euclid's algorithm - since Euclid *

P. M. Cohn
University College London

It has been said that every good idea has been thought of before†. This may not seem true in Mathematics, where so much spectacular progress has been made in our day; nevertheless there is some truth in it, in the sense that every new idea has its ancester somewhere among the great old ideas. Here I want to trace one idea down the ages : Euclid's algorithm.

1. Consider the natural numbers:

$$\mathbf{N} : 1, 2, 3, \ldots .$$

The two operations that can be performed on them, addition and multiplication, show very different behaviour. Using + we can get all positive integers by starting from $1 : 2 = 1 + 1, 3 = 1 + 1 + 1, \ldots$; we say that **N** is generated by 1, using +. For multiplication we need an infinite generating set, consisting of all the prime numbers: $p_1 = 2$, $p_2 = 3$, $p_3 = 5, \ldots$. Every number a is uniquely expressible in the form

$$a = p_1^{\alpha_1} p_2^{\alpha_2} \ldots, \tag{1}$$

where the α_i are non-negative integers and all but a finite number of them are 0. This is often called the Fundamental Theorem of Arithmetic. Nowadays one usually states it for the ring **Z** of all integers (got from **N** by throwing in zero and the negative integers). Thus we can say that in **Z** every non-zero element is either a unit (i.e. invertible, namely 1 or -1), or a product of unfactorable elements, which are unique except for the order in

---

* Lecture given to the Singapore Mathematical Society on April 10, 1991. The article first appeared in EUREKA 44(1984), 39-45 and thanks are due to the editors for permission to reprint it here (in slightly revised form).

† Alles gescheite ist schon gedacht worden; man muss nur versuchen es noch einmal zu denken (Goethe: Sprüche in Prosa).

which they occur and for unit factors (e.g. $6 = 2 \cdot 3 = (-3) \cdot (-2)$ etc.). This is expressed more briefly by saying: $\mathbf{Z}$ is a unique factorization domain, or more briefly still, a UFD.

In a UFD it is easy to describe the highest common factor (HCF) and least common multiple (LCM) of two members. Instead of (1) we can briefly write $a = \Pi p_i^{\alpha_i}$. If $b = \Pi p_i^{\beta_i}$ is another element, then we have

$$\text{(2)} \qquad \text{HCF} : (a,b) = \Pi\, p_i^{\delta_i}, \quad \text{where } \delta_i = \min\{\alpha_i, \beta_i\},$$

$$\text{(3)} \qquad \text{LCM} : [a,b] = \Pi\, p_i^{\mu_i}, \quad \text{where } \mu_i = \max\{\alpha_i, \beta_i\}.$$

For example, take $a = 36 = 2^2 \cdot 3^2$, $b = 50 = 2.5^2$; then $(a,b) = 2$, $[a,b] = 900$. If we know one of (2), (3), we can find the other by the formula : $(a,b)[a,b] = ab$, but this is of no help in finding the HCF and LCM themeselves, unless all factorizations (1) are known.

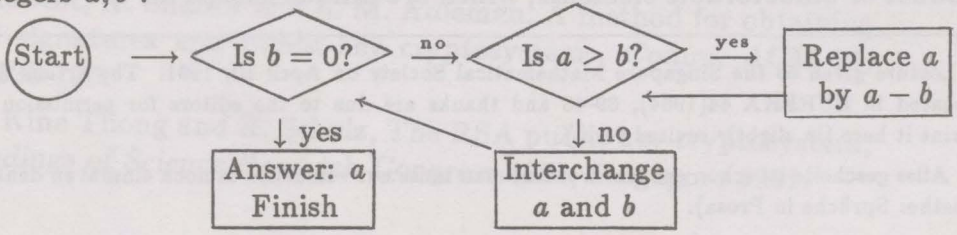
To find the HCF of two numbers a, b without factorizing them, Euclid [2] uses the division with remainder : if $b > 0$, there exist numbers q, r such that

$$\text{(4)} \qquad a = bq + r, \qquad 0 \leq r < b.$$

We now repeat the process with a, b replaced by b, r and continue in this way, getting a chain of equations

$$\begin{aligned} a &= bq_1 + r_1, \\ b &= r_1 q_2 + r_2, \\ r_1 &= r_2 q_3 + r_3, \quad b > r_1 > r_2 > \dots \\ &\dots \quad \dots \\ r_{n-1} &= r_n q_{n+1}. \end{aligned} \qquad \text{(5)}$$

Since the remainders are decreasing positive integers, the chain must break off, when $r_{n+1} = 0$. It is easy to check that the last non-zero remainder r_n is the HCF of a and b, as we shall see in a moment, and we have an algorithm, because the answer is always reached in a finite number of steps. Here is a flow-chart to find the HCF of two non-negative integers a, b :



The chain (5) of equations can be rewritten as follows in terms of matrices. If we put $P(x) = \begin{pmatrix} x & 1 \\ 1 & 0 \end{pmatrix}$, then $P(x)^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & -x \end{pmatrix}$ and (5) takes the form (where (a, b) denotes a 2-component vector, not the HCF)

$$
\begin{aligned}
(a, b) &= (b, r_1)P(q_1), \\
(b, r_1) &= (r_1, r_2)P(q_2), \\
&\cdots \quad \cdots \\
(r_{n-1}, r_n) &= (r_n, 0)P(q_{n+1}).
\end{aligned} \tag{6}
$$

Let us write $C = P(q_{n+1})P(q_n)\ldots P(q_1)$; then the matrix C has an inverse, because each P has, and we obtain from (6),

$$(a, b) = (r_n, 0)C, \quad (r_n, 0) = (a, b)C^{-1}. \tag{7}$$

We shall write $d|a$ to indicate that $a = md$ for some m, i.e. that a is divisible by d. The first equation (7) shows that $r_n|a$, $r_n|b$, so r_n is a common factor of a, b. The second equation has the form $r_n = au + bv$, (where u, v are the entries of the first column of C^{-1}, again integers). Hence any common factor of a, b also divides r_n and this shows r_n to be the required HCF.

2. It is not difficult to obtain explicit formulae for a, b in terms of the quotients and remainders occuring in the Euclidean algorithm. We define polynomials p_n in n variables recursively by $p_o = 1$, $p_1(t_1) = t_1$, and for $n \geq 2$,

$$p_n(t_1, \ldots, t_n) = p_{n-1}(t_1, \ldots, t_{n-1})t_n + p_{n-2}(t_1, \ldots, t_{n-2}).$$

This definition shows incidentally that $p_n(1, 1, \ldots, 1)$ is the nth Fibonacci number, cf. [3]. The first few p's are $1, t_1, t_1t_2+1, t_1t_2t_3+t_3+t_1, t_1t_2t_3t_4+t_3t_4+t_1t_4+t_1t_2+1$. In general p_n is formed by the "'leapfrog rule": write down $t_1t_2\ldots t_n$ and add to it all products obtained by omitting one or more pairs t_it_{i+1}. Alternatively p_n may be described as the polynomial part of the rational function

$$(t_1 + t_2^{-1})(t_2 + t_3^{-1})\ldots(t_{n-1} + t_n^{-1})t_n.$$

The p_n occur as numerators and denominators of continued fractions [3], and so are called continuant polynomials, and they can also be described

as determinants (continuants), but our interest in them here stems from the fact that (by an easy induction),

$$P(t_1)...P(t_n) = \begin{pmatrix} p(t_1,...,t_n) & p(t_1,...,t_{n-1}) \\ p(t_2,...,t_n) & p(t_2,...,t_{n-1}) \end{pmatrix}, \tag{8}$$

(where the suffix on p_i has been omitted for simplicity). Since $P(x)$ has determinant -1, the inverse of (8) follows from the formula

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = (-1)^n \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}. \tag{9}$$

If we now apply (8) to the Euclidean algorithm we find that for any positive integers a, b with HCF d, we have

$$a = dp(q_{n+1},...,q_1), \quad b = dp(q_{n+1},...,q_2), \tag{10}$$

and

$$d = bu - av, \tag{11}$$

where $u = (-1)^n p(q_1,..,q_n)$, $v = (-1)^n p(q_2,...,q_n)$.

3. One can ask similar questions about the ring of polynomials $a_0 x^n + a_1 x^{n-1} + ... + a_n$ with rational (or real) coefficients. The answer is now part of most second year courses, but it was not always so easy. Pedro Nuñez [5] (inventor of the Vernier scale), writing in 1567 tries to find the HCF of two polynomials, but without success; he does not get beyond some generalities. Yet only 18 years later Simon Stevin [6] sets it as a problem and says: the answer is obtained by applying the Euclidean algorithm, as for integers. Of course this is not quite true, we need to use the degree of the polynomial in place of the absolute value of a. But why did Nuñez find it so hard? My guess is that he took integer coefficients; then the Euclidean algorithm does not apply and in fact the problem is quite difficult. Stevin (who among other things introduced decimal notation) would be more likely to use rational coefficients and so make the problem more tractable.

4. We now have two examples and true to modern habits, we try to find a description to fit both. Our starting point is any integral domain, i.e. a ring (not necessarily commutative) without zero-divisors and with $1 \neq 0$.

An integral domain R is called a Euclidean domain if for each $a \in R$ there is a non-negative integer $\phi(a)$ with the properties:

E.1 $\phi(0) = 0$,

E.2 $\phi(ab) \geq \phi(a)$ *for all* $a, b \in R$, $b \neq 0$,

E.3 *For any* $a, b \in R$ *if* $b \neq 0$ *and* $\phi(a) \geq \phi(b)$, *there exists* $c \in R$ such that

$$\phi(a - bc) < \phi(a). \tag{12}$$

From E.3 we can easily derive the more usual form of the division algorithm:

E.3′ *For any* $a, b \in R$, *if* $b \neq 0$, *there exist* $q, r \in R$ *such that*

$$a = bq + r, \quad \phi(r) < \phi(b). \tag{13}$$

If ϕ is constant on the non-zero elements of R, then by E.3 every non-zero element of R has an inverse, so R is a field. We shall exclude this rather trivial case.

To prove that E.3 and E.3′ are equivalent, assume E.3 and for given a, b, choose q, r in (13) with $\phi(r)$ minimal. If $\phi(r) \geq \phi(b)$, then by E.3, $\phi(r - bc) < \phi(r)$ for some $c \in R$, but $r - bc = a - b(q + c)$ and this contradicts the minimality of $\phi(r)$. Conversely, if E.3′ is given and $\phi(a) \geq \phi(b)$, we take q, r as in (13); then $\phi(a - bq) = \phi(r) < \phi(b) \leq \phi(a)$.

One can now develop the Euclidean algorithm as before and use it to prove the existence of HCF and LCM, and also the following property of unfactorable elements, used to establish unique factorization:

Euclid's Lemma. Given any unfactorable element p in a (commutative) Euclidean domain, if p divides ab, then p divides a or p divides b.

The proof is in all text-books and depends on the Bezout identity:

$$HCF(a, b) = au + bv, \tag{14}$$

which expresses the HCF of a and b as a linear combination of a and b.

More generally one can show that in a Euclidean domain R any ideal can be generated by a single element, i.e. R is a principal ideal domain. We shall not give details (nor even define 'ideal'), but merely note the following examples of Euclidean domains: (i) $\mathbf{Z}$, $\phi(a) = |a|$, (ii) $k[x]$, k a field, $\phi(a) = \deg a$, (iii) $\mathbf{Z}[i]$ or more generally, the ring of integers in

$\mathbb{Q}(\sqrt{d})$, $\phi(a) = |a|$, when d = -1, -2, -3, -7, -11 and 2, 3, 5, 6, 7, 11, 13, 17, 19, 21, 29, 33, 37, 41, 55, 73 (cf. [7],p.95).

5. By a more elaborate method (using Gauss's lemma) one can show that the polynomial ring in several variables $k[x_1, ..., x_n]$ over any field k of coefficients is a UFD, but for $n > 1$ the Euclidean algorithm seems to have got lost. This may well be connected with the fact that whereas in $k[x]$ every ideal is principal, this is no longer the case in the ring $k[x_1, ..., x_n]$ for $n > 1$. Some efforts to find a Euclidean algorithm or a substitute were made (cf. e.g. [4]), but without much success.

Nevertheless there is an analogue applying to polynomial rings which enables us to prove almost everything the usual algorithm does (as far as it is true). It applies to polynomials in any number of variables over any field k, but with the proviso that the variables do not commute. This polynomial ring in non-commuting variables is called the free associative algebra on $x_1, ..., x_n$ over k, written $k\langle x_1, ..., x_n\rangle$.

To find this general algorithm, let us take two variables x, y for simplicity. Given two elements of $k\langle x, y\rangle$, say $f = x^2y + yx + 1$ and $g = xyx + yxy$, in general neither can be divided by the other (even with remainder), for a very good reason: f and g have no common right multiple at all, apart from zero. This is a new feature which did not appear in the commutative case, where any two non-zero elements f, g have the common multiple $fg = gf$. If we now restrict attention to pairs with a non-zero common right multiple, we find that the Euclidean algorithm is restored. E.g., if $f = xyz + z + x$, $g = xy + 1$, then $f = gz + x$, $g = x \cdot y + 1$, $x = 1 \cdot x$. In this example, $f = p(x, y, z)$, $g = p(x, y)$ in the notation of Section 2; this is no accident, for as we saw, the Euclidean algorithm can always be expressed in terms of the p's; of course in general the arguments of the p's will be much more complicated than this example.

What we actually need is a kind of n-term algorithm which applies whenever an appropriate right multiple condition is satisfied. This is the

Weak algorithm. Given $a_1, ..., a_m$, if there exist $b_1, ..., b_m$ such that

$$\phi(\sum a_i b_i) < \max\{\phi(a_1 b_1), ..., \phi(a_m b_m)\},$$

and if the a's are numbered in such a way that $\phi(a_1) \leq \phi(a_2) \leq ... \leq \phi(a_m)$, then there exists j in the range $2 \leq j \leq m$ and $c_1, ..., c_{j-1} \in R$ such that

$$\phi(a_j - \sum_1^{j-1} a_i c_i) < \phi(a_j), \quad \phi(a_i c_i) \leq \phi(a_j) \quad (i = 1, ..., j-1).$$

This is satisfied by the free algebra $k\langle X\rangle$ in any set X of non-commuting variables over a field k, taking ϕ to be the usual degree. In fact free algebras may be characterized in this way, and more generally, all rings with a week algorithm can be determined (cf. [1], Ch.2).

The weak algorithm enables one to prove a unique factorization property: any two complete factorizations of a given element have the same number of factors and the factors on the two sides can be paired off so that corresponding factors are 'similar'; we shall not define this term here but merely remark that in the commutative case two elements are similar precisely if they differ by a unit factor. As a simple example we have the factorizations in $k\langle x, y\rangle$:

$$xyx + x = (xy + 1)x = x(yx + 1); \tag{15}$$

here $xy+1$ is similar to $yx+1$. To give another example, over $\mathbf{R}\langle x, y\rangle$ the element $a = xy^2x + xy + yx + x^2 + 1$ is unfactorable, but when the ground field is extended to $\mathbf{C}$, a can be factorized as

$$a = (xy + ix + 1)(yx - ix + 1).$$

Writing $b = xy^2 + x + y$, we have $b^2 + 1 = a(y^2 + 1)$ and this shows that a and b have a non-zero commont right multiple. Hence one can carry out the Euclidean algorithm for a, b (more precisely, the right Euclidean algorithm); this is an amusing and not too difficult exercise.

For any two elements a, b which have a common non-zero right multiple, there is a highest common left factor d, which can again be written as a right linear combination of a and b, as in (11). In fact, exactly the same formulae (10), (11) apply, where the q's are the quotients obtained from the weak algorithm (cf.[1], Ch.2). Here the p_n are defined as before and (8) still holds. The formula (9) for the inverse matrix cannot now be used, but the inverse exists and is easily written down, bearing in mind that $P(x)$ has an inverse.

Of course, the ideals in $k\langle X\rangle$ are not principal, but they are free, as modules over the ring, of well-defined rank, thus $k\langle X\rangle$ is a free ideal ring (fir for short). One can work out a theory of firs which in many respects parallels the theory of principal ideal domains, and one finds that many important rings are firs. Better still, some have a weak algorithm; this enables one to take over most of the formulae of Section 2, but in spite of their very explicit form, many questions can be asked about these rings which are still unanswered. The first Edition of [1] appeared in 1971 and

contained 86 open problems, of which 15 were solved in the next 14 years. The Second Edition (1985) contains 103 open problems, of which 8 have been solved to date, leaving 95 still unsolved.

References

[1] P. M. Cohn, *Free rings and their relations*, 2nd Ed. LMS Monographs No. 19 (Academic Press, Londong 1985).

[2] Euclid, *Elements*, -300.

[3] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers* (Clarendon Press, Oxford 1938).

[4] H. Levi, A characterization of polynomial rings by means of order relations, *Amer. J. Math.* 65 (1943), 221-234.

[5] P. Nuñez, *Libro de Algebra* (Antwerp 1567).

[6] S. Stevin, *Arithmétique 1585* (Vol. II of collected works, 1958).

[7] I. N. Stewart and D. O. Tall, *Algebraic Number Theory* (Chapman and Hall, London 1979).