# DIVISIBILITY AND GREATEST COMMON DIVISORS

KEITH CONRAD

## 1. Introduction

We will begin with a review of divisibility among integers, mostly to set some notation and to indicate its properties. Then we will look at two important theorems involving greatest common divisors: Euclid's algorithm and Bezout's identity.

The set of integers is denoted $\mathbf{Z}$ (from the German word Zahl = number).

## 2. The Divisibility Relation

**Definition 2.1.** When $a$ and $b$ are integers, we say $a$ *divides* $b$ if $b = ak$ for some $k \in \mathbf{Z}$. We then write $a|b$ (read as "$a$ divides $b$").

**Example 2.2.** We have $2|6$ (because $6 = 2 \cdot 3$), $4|(-12)$, and $5|0$. We have $\pm 1|b$ for every $b \in \mathbf{Z}$. However, 6 does not divide 2 and 0 does not divide 5.

Divisibility is a relation, much like inequalities. In particular, the relation $2|6$ is *not* the number 3, even though $6 = 2 \cdot 3$. Such an error would be similar to the mistake of confusing the relation $5 < 9$ with the number $9 - 5$.

Notice divisibility is not symmetric: if $a|b$, it is usually not true that $b|a$, so you should not confuse the roles of $a$ and $b$ in this relation: $4|20$ but $20 \nmid 4$.

**Remark 2.3.** Learn the definition of $a|b$ as given in Definition 2.1, and not in the form "$\frac{b}{a}$ is an integer." It essentially amounts to the same thing (exception: $0|0$ but $\frac{0}{0}$ is not defined), however thinking about divisibility in terms of ratios will screw up your understanding of divisibility in other settings in algebra. That is why it is best to regard Definition 2.1, which makes no reference to fractions, as the correct definition of divisibility.

The following three theorems about divisibility are simple applications of the definition. They should all make intuitive sense.

**Theorem 2.4.** *Let $a, b \in \mathbf{Z}$ with $a|b$. Then $a|bc$ for any $c \in \mathbf{Z}$.*

*Proof.* We have $b = ak$ for some $k \in \mathbf{Z}$. Therefore $bc = (ak)c = a(kc)$ and $kc \in \mathbf{Z}$, so $a|bc$. □

This just says a factor of a number is a factor of any multiple of it. (Or, equivalently, a multiple of a multiple is a multiple.)

**Theorem 2.5.** *If $a|b$ and $b|c$ then $a|c$.*

*Proof.* We have $b = ak$ and $c = b\ell$ for some $k$ and $\ell$ in $\mathbf{Z}$. Then

$$c = b\ell = a(k\ell)$$

and $k\ell \in \mathbf{Z}$, so $a|c$. □

As a mantra, "a factor of a factor is a factor."

**Theorem 2.6.** *If $a|b$ and $a|c$ then $a|(br+cs)$ for every $r$ and $s$ in $\mathbf{Z}$. In particular, if $a|b$ and $a|c$ then $a|(b+c)$ and $a|(b-c)$.*

*Proof.* We have $b = ak$ and $c = a\ell$ for some $k$ and $\ell$ in $\mathbf{Z}$. Then

$$br + cs = akr + a\ell s = a(kr + \ell s)$$

and $kr + \ell s \in \mathbf{Z}$, so $a|(br+cs)$. □

Using the language of linear algebra, Theorem 2.6 says any factor of two integers is also a factor of any $\mathbf{Z}$-linear combination of the two integers.

To avoid silly errors, keep in mind the following false implications: generally

$$a|bc \not\Rightarrow a|b \text{ or } a|c$$

(for instance, $6|(4 \cdot 9)$ but 6 divides neither 4 nor 9) and

$$a|b^n \not\Rightarrow a|b$$

(for instance, $12|6^2$ but 12 doesn't divide 6).

## 3. Greatest Common Divisors

For two nonzero integers $a$ and $b$, their *greatest common divisor* is the largest integer which is a factor of both of them. It is denoted $(a, b)$. For instance, $(12, 18) = 6$ and $(-9, 15) = 3$. Do not confuse our usage of parentheses in $(a, b)$ with the notation for open intervals in calculus. The number 1 is always a common divisor, and it is the greatest common divisor exactly when $a$ and $b$ are relatively prime.

The naive method of finding the greatest common divisor of two integers is to factor each into primes and extract the greatest common divisor from the prime power factors that appear.

**Example 3.1.** Consider $a = 19088597$ and $b = 39083$. Since

$$19088597 = 11^2 \cdot 19^3 \cdot 23, \quad 39083 = 11^2 \cdot 17 \cdot 19,$$

we have $(19088597, 39083) = 11^2 \cdot 19 = 2299$.

Factoring is hard (even on a computer when the integer has several hundred digits), so this method of computing $(a, b)$ is not good when $a$ and $b$ are large. There is a method of computing greatest common divisors, going back to Euclid, which avoids the need to factor at all. Instead of factoring, we will do successive divisions with remainder in such a way that the remainder keeps dropping. The *last nonzero remainder* will turn out to be the greatest common divisor.

**Theorem 3.2** (Euclid)**.** *Let $a$ and $b$ be nonzero integers. Divide $b$ into $a$ and carry out further divisions according to the following method, where the old remainder becomes the new divisor:*

$$\begin{aligned} a &= bq_1 + r_1, \quad 0 \leq r_1 < |b|, \\ b &= r_1 q_2 + r_2, \quad 0 \leq r_2 < r_1, \\ r_1 &= r_2 q_3 + r_3, \quad 0 \leq r_3 < r_2, \\ &\vdots \quad \vdots \quad \vdots \end{aligned}$$

*The non-negative remainders $r_1, r_2, \ldots$ are strictly decreasing, and thus must eventually become 0. The last nonzero remainder is the greatest common divisor.*

This algorithm of Euclid for finding $(a, b)$ can be carried out very rapidly on a computer, even for very large integers which are not easy to factor into primes.

**Example 3.3.** Before we prove Euclid's algorithm works, let's see how it looks for the pair in Example 3.1:

$$\begin{aligned} 19088597 &= 39083 \cdot 488 + 16093 \\ 39083 &= 16093 \cdot 2 + 6897 \\ 16093 &= 6897 \cdot 2 + 2299 \\ 6897 &= 2299 \cdot 3 + 0. \end{aligned}$$

The last nonzero remainder is 2299, and we said $(19088597, 39083) = 2299$ in Example 3.1. Notice we did not need to factor the two numbers to find their greatest common divisor.

Let's prove Theorem 3.2.

*Proof.* The key idea that makes Euclid's algorithm work is this: if $a = b + mk$ for some $k$ in $\mathbf{Z}$, then $(a, m) = (b, m)$. That is, two numbers whose difference is a multiple of $m$ have the same gcd with $m$. Indeed, any common divisor of $a$ and $m$ is a divisor of $b = a - mk$ (Theorem 2.6), and therefore is a common divisor of $b$ and $m$. This tells us $(a, m) \leq (b, m)$. Similarly, any common divisor of $b$ and $m$ is a divisor of $a = b + mk$, and therefore is a common divisor of $a$ and $m$. Thus $(b, m) \leq (a, m)$ too, so $(a, m) = (b, m)$.

Another way of putting this is:

$$m|(a - b) \Longrightarrow (a, m) = (b, m). \tag{3.1}$$

Now we look at the successive equations in Euclid's algorithm:

$$\begin{aligned} a &= bq_1 + r_1, \quad 0 \leq r_1 < |b|, \\ b &= r_1 q_2 + r_2, \quad 0 \leq r_2 < r_1, \\ r_1 &= r_2 q_3 + r_3, \quad 0 \leq r_3 < r_2, \\ &\vdots \quad \vdots \quad \vdots \end{aligned}$$

The first equation says $b|(a - r_1)$, so by (3.1) we have $(a, b) = (r_1, b)$. The second equation says $r_1|(b - r_2)$, so again by (3.1) we have $(b, r_1) = (r_2, r_1)$. The third equation says $r_2|(r_1 - r_3)$, so again by (3.1) we have $(r_1, r_2) = (r_3, r_2)$. Comparing these results,

$$(a, b) = (b, r_1) = (r_1, r_2) = (r_2, r_3),$$

so the later greatest common divisors continue to be equal to $(a, b)$. The last equations in Euclid's algorithm look like this:

$$\begin{aligned} r_n &= r_{n+1} q_{n+2} + r_{n+2}, \quad 0 < r_{n+2} < r_{n+1}, \\ r_{n+1} &= r_{n+2} q_{n+3} + 0. \end{aligned}$$

Thus

$$(a, b) = (b, r_1) = \cdots = (r_n, r_{n+1}) = (r_{n+1}, r_{n+2}).$$

The final equation in Euclid's algorithm tells us $(r_{n+1}, r_{n+2}) = r_{n+2}$, so $(a, b)$ equals $r_{n+2}$, which is the last nonzero remainder. □

**Example 3.4.** We compute $(322345, 21419)$:

$$\begin{aligned} 322345 &= 21419 \cdot 15 + 1060, \\ 21419 &= 1060 \cdot 20 + 219, \\ 1060 &= 219 \cdot 4 + 184, \\ 219 &= 184 \cdot 1 + 35, \\ 184 &= 35 \cdot 5 + 9, \\ 35 &= 9 \cdot 3 + 8, \\ 9 &= 8 \cdot 1 + 1, \\ 8 &= 1 \cdot 8 + 0. \end{aligned}$$

Therefore $(322345, 21419) = 1$. The last equation was superfluous: if we ever reach a remainder of 1, then the next remainder is $\geq 0$ and less than 1 and therefore must be 0, so 1 is the last nonzero remainder.

Not only is the last equation superfluous, but we could have stopped already in the fourth equation: here we meet a remainder of 35, which is small enough that we can factor it in our heads as $5 \cdot 7$. Therefore

$$(322345, 21419) = (184, 35),$$

and we can easily check 5 and 7 are not factors of 184, so this greatest common divisor must be 1. However, this early cutoff in the algorithm misses something important: as we will soon see, *all* the steps of Euclid's algorithm are needed to carry out one of the algorithm's most crucial consequences.

The significance of Euclid's algorithm goes beyond its computation of the greatest common divisor. By reversing the steps of Euclid's algorithm starting with the equation having the last nonzero remainder, we are able to write $(a, b)$ in an especially useful form, as follows.

**Theorem 3.5** (Bezout)**.** *For nonzero $a$ and $b$ in $\mathbf{Z}$, there are $x$ and $y$ in $\mathbf{Z}$ such that*

$$(a, b) = ax + by. \tag{3.2}$$

*In particular, when $a$ and $b$ are relatively prime, there are $x$ and $y$ in $\mathbf{Z}$ such that $ax+by = 1$.*

Adopting terminology from linear algebra, expressions of the form $ax + by$ with $x, y \in \mathbf{Z}$ are called $\mathbf{Z}$-linear combinations of $a$ and $b$. Equation (3.2) is called *Bezout's identity*.

Before we prove Theorem 3.5 we illustrate the idea of the proof in some examples.

**Example 3.6.** In Example 3.3 we used Euclid's algorithm to show $(19088597, 39083) = 2299$. Reversing the steps of that algorithm,

$$\begin{aligned} 2299 &= 16093 - 6897 \cdot 2 \\ &= 16093 - (\mathbf{39083} - \mathbf{16093} \cdot 2) \cdot 2 \\ &= 16093 \cdot (1 + 2 \cdot 2) - 39083 \cdot 2 \\ &= 16093 \cdot 5 - 39083 \cdot 2 \\ &= (\mathbf{19088597} - \mathbf{39083} \cdot 488) \cdot 5 - 39083 \cdot 2 \\ &= 19088597 \cdot 5 + 39083 \cdot (-488 \cdot 5 - 2) \\ &= 19088597 \cdot 5 - 39083 \cdot 2442. \end{aligned}$$

Therefore Bezout's identity is satisfied with the integers $x = 5$ and $y = -2442$ (*not* $y = 2442$).

**Example 3.7.** Let $a = 121$ and $b = 38$. Then by Euclid's algorithm,

$$\begin{aligned} 121 &= 38 \cdot 3 + 7, \\ 38 &= 7 \cdot 5 + 3, \\ 7 &= 3 \cdot 2 + 1, \end{aligned}$$

and we can stop here since we have reached a remainder of 1. Unwinding,

$$\begin{aligned} 1 &= 7 - 3 \cdot 2 \\ &= 7 - (\mathbf{38} - \mathbf{7} \cdot 5) \cdot 2 \\ &= 7 \cdot 11 - 38 \cdot 2 \\ &= (\mathbf{121} - \mathbf{38} \cdot 3) \cdot 11 - 38 \cdot 2 \\ &= 121 \cdot 11 - 38 \cdot 35, \end{aligned}$$

so $121x + 38y = 1$ for $x = 11$ and $y = -35$.

Of course it would be completely trivial (but also useless!) to solve $121x + 38y = 1$ in real numbers $x$ and $y$: use $x = 0$ and $y = 1/38$, for instance. Being able to solve such an equation with *integers* is the key point.

**Example 3.8.** The reader should reverse the steps of Euclid's algorithm in Example 3.4 to find

$$1 = 322345 \cdot 2445 + 21419 \cdot (-36796).$$

Now we'll prove Theorem 3.5.

*Proof.* Write out the equations in Euclid's algorithm in their natural order:

$$\begin{aligned} a &= bq_1 + r_1, \quad 0 < r_1 < |b|, \\ b &= r_1q_2 + r_2, \quad 0 < r_2 < r_1, \\ r_1 &= r_2q_3 + r_3, \quad 0 < r_3 < r_2, \\ &\vdots \quad \vdots \quad \vdots \\ r_{n-1} &= r_nq_{n+1} + r_{n+1}, \quad 0 < r_{n+1} < r_n, \\ r_n &= r_{n+1}q_{n+2} + r_{n+2}, \quad 0 < r_{n+2} < r_{n+1}, \\ r_{n+1} &= r_{n+2}q_{n+3}. \end{aligned}$$

In the equation with the last nonzero remainder, solve for it:

$$r_{n+2} = r_n - r_{n+1}q_{n+2}. \tag{3.3}$$

This expresses $r_{n+2}$ as a $\mathbf{Z}$-linear combination of $r_n$ and $r_{n+1}$. Now feed in the expression for the remainder from the preceding equation:

$$\begin{aligned} r_{n+2} &= r_n - (\underline{r_{n-1}} - \underline{r_n}q_{n+1})q_{n+2} \\ &= r_n(1 + q_{n+1}q_{n+2}) - r_{n-1}q_{n+2}. \end{aligned}$$

Now we have the last nonzero remainder $r_{n+2}$ as a $\mathbf{Z}$-linear combination of $r_{n-1}$ and $r_n$. Proceeding up the equations in Euclid's algorithm, eventually we reach

$$r_{n+2} = bu + r_1v$$

for some $u, v \in \mathbf{Z}$. Finally, writing $r_1$ as $a - bq_1$, we get

$$r_{n+2} = av + b(u - vq_1)$$

and we have obtained Bezout's identity. □

## 4. Consequences of Bezout's identity

In this section we collect several corollaries of Bezout's identity (3.2). The main thing to learn from each proof is how Bezout's identity is used: whenever you have relatively prime integers $a$ and $b$, you should immediately think "Oh, so 1 can be written in the form $ax+by$ for some integers $x$ and $y$." This is the *main technique* to handle proofs involving relatively prime integers. Since Bezout's identity is not intuitive, you will find most of the proofs are not intuitive even if the statements of the theorems feel like common sense. On the other hand, the proofs are quite short. After reading the proofs, write out the corollaries on a separate sheet of paper and check that you can reproduce the proofs on your own.

**Corollary 4.1.** *If $a|bc$ and $(a,b)=1$, then $a|c$.*

*Proof.* Since $a|bc$, $bc=ak$ for some integer $k$. Because $(a,b)=1$,

$$1 = ax + by$$

for some integers $x$ and $y$. Multiplying through by $c$ (why? because we want to show $c$ is $a$ times something and $c\cdot 1 = c$), we have

$$c = acx + (bc)y = acx + (ak)y = a(cx+ky).$$

Since $cx+ky \in \mathbf{Z}$, $a|c$. □

**Corollary 4.2.** *If $a|c$, $b|c$, and $(a,b)=1$, then $ab|c$.*

*Proof.* We have $c=ak$ and $c=b\ell$ where $k$ and $\ell$ are integers. Also, since $(a,b)=1$ we have

$$ax + by = 1$$

for some integers $x$ and $y$. Therefore

$$c = c\cdot 1 = cax + cby = (b\ell)ax + (ak)by = ab(\ell x + ky).$$

Since $\ell x + ky \in \mathbf{Z}$, $ab|c$. □

**Corollary 4.3.** *If $(a,c)=1$ and $(b,c)=1$, then $(ab,c)=1$.*

*Proof.* Write

$$ax + cy = 1, \quad bx' + cy' = 1$$

for some integers $x, y, x', y'$. Multiplying the above equations,

$$1 = (ax+cy)(bx'+cy') = ab(xx') + c(axy' + bx'y + cyy').$$

This expresses 1 as a $\mathbf{Z}$-linear combination of $ab$ and $c$, so $ab$ and $c$ are relatively prime (any common factor would divide 1, and thus is $\pm 1$). □

All these corollaries generalize to more than two pairs of relatively prime integers, by using induction on the number of pairs (and associativity to write a product of several integers as a product of two integers, *e.g.*, $abcd = (abc)d$). We will state these generalizations, but leave the proofs to the reader:

(1) if $a|b_1b_2\dots b_rc$ and $(a,b_i)=1$ for all $i$, then $a|c$,
(2) if $a_1,\dots,a_r$ are all factors of $m$ and $(a_i,a_j)=1$ for all $i\neq j$ then $a_1a_2\cdots a_r$ is a factor of $m$,
(3) if $a_1,\dots,a_r$ are relatively prime to $m$ then $a_1a_2\cdots a_r$ is relatively prime to $m$.