Factorization and Primality Testing Chapter 2 Euclid's Algorithm and Applications

Robert C. Vaughan

August 26, 2025

Euclid's Algorithm

- The question arises. We know that given integers a, b not both 0, there are integers x and y so that

$$(a, b) = ax + by.$$

How do we find x and y?

Euclid's Algorithm

- The question arises. We know that given integers a, b not both 0, there are integers x and y so that

$$(a, b) = ax + by.$$

  How do we find x and y?
- A method for solving this problem, known as Euclid's algorithm, first appeared in Euclid's *Elements* more than 2000 years ago.

Euclid's Algorithm

- The question arises. We know that given integers a, b not both 0, there are integers x and y so that

$$(a, b) = ax + by.$$

  How do we find x and y?
- A method for solving this problem, known as Euclid's algorithm, first appeared in Euclid's *Elements* more than 2000 years ago.
- Moreover this solution gives a very efficient algorithm and it is still the basis for many numerical methods in arithmetical applications.

Euclid's Algorithm

- The question arises. We know that given integers a, b not both 0, there are integers x and y so that

$$(a, b) = ax + by.$$

  How do we find x and y?
- A method for solving this problem, known as Euclid's algorithm, first appeared in Euclid's *Elements* more than 2000 years ago.
- Moreover this solution gives a very efficient algorithm and it is still the basis for many numerical methods in arithmetical applications.
- We may certainly suppose that a and $b > 0$ since multiplying either by (-1) does not change the (a, b) - we can replace x by $-x$ and y by $-y$.

- We can certainly suppose that $b \le a$. For convenience of notation put $r_0 = b$, $r_{-1} = a$.

- We can certainly suppose that $b \le a$. For convenience of notation put $r_0 = b$, $r_{-1} = a$.
- Now apply the division algorithm iteratively as follows

$$
\begin{aligned}
r_{-1} &= r_0 q_1 + r_1, \quad 0 < r_1 \le r_0,\\
r_0 &= r_1 q_2 + r_2, \quad 0 < r_2 < r_1,\\
r_1 &= r_2 q_3 + r_3, \quad 0 < r_3 < r_2,\\
&\dots\\
r_{s-3} &= r_{s-2} q_{s-1} + r_{s-1}, \quad 0 < r_{s-1} < r_{s-2},\\
r_{s-2} &= r_{s-1} q_s.
\end{aligned}
$$

- We can certainly suppose that $b \le a$. For convenience of notation put $r_0 = b$, $r_{-1} = a$.
- Now apply the division algorithm iteratively as follows

$$\begin{aligned}
r_{-1} &= r_0 q_1 + r_1, \quad 0 < r_1 \le r_0, \\
r_0 &= r_1 q_2 + r_2, \quad 0 < r_2 < r_1, \\
r_1 &= r_2 q_3 + r_3, \quad 0 < r_3 < r_2, \\
&\dots \\
r_{s-3} &= r_{s-2} q_{s-1} + r_{s-1}, \quad 0 < r_{s-1} < r_{s-2}, \\
r_{s-2} &= r_{s-1} q_s.
\end{aligned}$$

- That is, we stop the moment that there is a remainder equal to 0.

- We can certainly suppose that $b \le a$. For convenience of notation put $r_0 = b$, $r_{-1} = a$.
- Now apply the division algorithm iteratively as follows

$$\begin{aligned}
r_{-1} &= r_0 q_1 + r_1, \quad 0 < r_1 \le r_0, \\
r_0 &= r_1 q_2 + r_2, \quad 0 < r_2 < r_1, \\
r_1 &= r_2 q_3 + r_3, \quad 0 < r_3 < r_2, \\
&\dots \\
r_{s-3} &= r_{s-2} q_{s-1} + r_{s-1}, \quad 0 < r_{s-1} < r_{s-2}, \\
r_{s-2} &= r_{s-1} q_s.
\end{aligned}$$

- That is, we stop the moment that there is a remainder equal to 0.
- This could be r_1 if $b|a$, for example, although the way it is written out above it is as if s is at least 3.

- We can certainly suppose that $b \le a$. For convenience of notation put $r_0 = b$, $r_{-1} = a$.
- Now apply the division algorithm iteratively as follows

$$\begin{aligned}
r_{-1} &= r_0 q_1 + r_1, \quad 0 < r_1 \le r_0,\\
r_0 &= r_1 q_2 + r_2, \quad 0 < r_2 < r_1,\\
r_1 &= r_2 q_3 + r_3, \quad 0 < r_3 < r_2,\\
&\dots\\
r_{s-3} &= r_{s-2} q_{s-1} + r_{s-1}, \quad 0 < r_{s-1} < r_{s-2},\\
r_{s-2} &= r_{s-1} q_s.
\end{aligned}$$

- That is, we stop the moment that there is a remainder equal to 0.
- This could be r_1 if $b|a$, for example, although the way it is written out above it is as if s is at least 3.
- The important point is that because $r_j < r_{j-1}$, sooner or later we must have a zero remainder.

- Repeating

$$\begin{aligned}
r_{-1} &= r_0 q_1 + r_1, \quad 0 < r_1 \le r_0,\\
r_0 &= r_1 q_2 + r_2, \quad 0 < r_2 < r_1,\\
r_1 &= r_2 q_3 + r_3, \quad 0 < r_3 < r_2,\\
&\dots\\
r_{s-3} &= r_{s-2} q_{s-1} + r_{s-1}, \quad 0 < r_{s-1} < r_{s-2},\\
r_{s-2} &= r_{s-1} q_s.
\end{aligned}$$

- Repeating

$$\begin{aligned}
r_{-1} &= r_0 q_1 + r_1, \quad 0 < r_1 \le r_0,\\
r_0 &= r_1 q_2 + r_2, \quad 0 < r_2 < r_1,\\
r_1 &= r_2 q_3 + r_3, \quad 0 < r_3 < r_2,\\
&\dots\\
r_{s-3} &= r_{s-2} q_{s-1} + r_{s-1}, \quad 0 < r_{s-1} < r_{s-2},\\
r_{s-2} &= r_{s-1} q_s.
\end{aligned}$$

- Euclid proved that $(a, b) = r_{s-1}$.

- Repeating

$$\begin{aligned}
r_{-1} &= r_0 q_1 + r_1, \quad 0 < r_1 \le r_0,\\
r_0 &= r_1 q_2 + r_2, \quad 0 < r_2 < r_1,\\
r_1 &= r_2 q_3 + r_3, \quad 0 < r_3 < r_2,\\
&\dots\\
r_{s-3} &= r_{s-2} q_{s-1} + r_{s-1}, \quad 0 < r_{s-1} < r_{s-2},\\
r_{s-2} &= r_{s-1} q_s.
\end{aligned}$$

- Euclid proved that $(a, b) = r_{s-1}$.
- First of all $(a, b)|a$ and $(a, b)|b$, and so $(a, b)|r_1$.

- Repeating

$$\begin{aligned} r_{-1} &= r_0q_1 + r_1, \quad 0 < r_1 \le r_0, \\ r_0 &= r_1q_2 + r_2, \quad 0 < r_2 < r_1, \\ r_1 &= r_2q_3 + r_3, \quad 0 < r_3 < r_2, \\ &\dots \\ r_{s-3} &= r_{s-2}q_{s-1} + r_{s-1}, \quad 0 < r_{s-1} < r_{s-2}, \\ r_{s-2} &= r_{s-1}q_s. \end{aligned}$$

- Euclid proved that $(a, b) = r_{s-1}$.
- First of all $(a, b)|a$ and $(a, b)|b$, and so $(a, b)|r_1$.
- Repeating this we get $(a, b)|r_j$ for $j = 2, 3, \dots, s-1$.

- Repeating

$$
\begin{aligned}
r_{-1} &= r_0 q_1 + r_1, \quad 0 < r_1 \le r_0, \\
r_0 &= r_1 q_2 + r_2, \quad 0 < r_2 < r_1, \\
r_1 &= r_2 q_3 + r_3, \quad 0 < r_3 < r_2, \\
&\ \ \dots \\
r_{s-3} &= r_{s-2} q_{s-1} + r_{s-1}, \quad 0 < r_{s-1} < r_{s-2}, \\
r_{s-2} &= r_{s-1} q_s.
\end{aligned}
$$

- Euclid proved that $(a, b) = r_{s-1}$.
- First of all $(a, b)|a$ and $(a, b)|b$, and so $(a, b)|r_1$.
- Repeating this we get $(a, b)|r_j$ for $j = 2, 3, \dots, s-1$.
- On the other hand, starting at the bottom line $r_{s-1}|r_{s-2}$, $r_{s-1}|r_{s-3}$ and so on until we have $r_{s-1}|b$ and $r_{s-1}|a$. Recall that this means that $r_{s-1}|(a, b)$.

- Repeating

$$\begin{aligned} r_{-1} &= r_0 q_1 + r_1, \quad 0 < r_1 \le r_0, \\ r_0 &= r_1 q_2 + r_2, \quad 0 < r_2 < r_1, \\ r_1 &= r_2 q_3 + r_3, \quad 0 < r_3 < r_2, \\ &\dots \\ r_{s-3} &= r_{s-2} q_{s-1} + r_{s-1}, \quad 0 < r_{s-1} < r_{s-2}, \\ r_{s-2} &= r_{s-1} q_s. \end{aligned}$$

- Euclid proved that $(a,b) = r_{s-1}$.
- First of all $(a,b)|a$ and $(a,b)|b$, and so $(a,b)|r_1$.
- Repeating this we get $(a,b)|r_j$ for $j = 2, 3, \dots, s-1$.
- On the other hand, starting at the bottom line $r_{s-1}|r_{s-2}$, $r_{s-1}|r_{s-3}$ and so on until we have $r_{s-1}|b$ and $r_{s-1}|a$. Recall that this means that $r_{s-1}|(a,b)$.
- Thus we have just proved that

$$r_{s-1}|(a,b), \quad (a,b)|r_{s-1}, \quad r_{s-1} = (a,b).$$

- Consider.

Example 1

Let $a = 10678,\ b = 42$

$$\begin{aligned} 10678 &= 42 \times 254 + 10 \\ 42 &= 10 \times 4 + 2 \\ 10 &= 2 \times 5. \end{aligned}$$

Thus $(10678, 42) = 2$.

- Consider.

Example 1

Let $a = 10678,\ b = 42$

$$\begin{aligned} 10678 &= 42 \times 254 + 10 \\ 42 &= 10 \times 4 + 2 \\ 10 &= 2 \times 5. \end{aligned}$$

Thus $(10678, 42) = 2.$

- But how to compute the x and y in $(a, b) = ax + by$?

- Consider.

Example 1

Let $a = 10678,\ b = 42$

$$\begin{aligned} 10678 &= 42 \times 254 + 10 \\ 42 &= 10 \times 4 + 2 \\ 10 &= 2 \times 5. \end{aligned}$$

Thus $(10678, 42) = 2$.

- But how to compute the x and y in $(a, b) = ax + by$?
- We could just work backwards through the algorithm using back substitution,

$$\begin{aligned} 2 &= 42 - 10 \times 4 = 42 - (10678 - 42 \times 254) \times 4 \\ &= 42 \times 1017 - 10678 \times 4. \end{aligned}$$

- Consider.

Example 1

Let $a = 10678,\ b = 42$

$$\begin{aligned} 10678 &= 42 \times 254 + 10 \\ 42 &= 10 \times 4 + 2 \\ 10 &= 2 \times 5. \end{aligned}$$

Thus $(10678, 42) = 2$.

- But how to compute the x and y in $(a,b) = ax + by$?
- We could just work backwards through the algorithm using back substitution,

$$\begin{aligned} 2 = 42 - 10 \times 4 &= 42 - (10678 - 42 \times 254) \times 4 \\ &= 42 \times 1017 - 10678 \times 4. \end{aligned}$$

- In general this is tedious and computationally wasteful since it requires all our calculations to be stored.

- A simpler way is as follows.

- A simpler way is as follows.
- Define $x_{-1} = 1$, $y_{-1} = 0$, $x_0 = 0$, $y_0 = 1$ and then lay the calculations out as follows.

$$\begin{array}{lll}
r_{-1} = r_0 q_1 + r_1, & x_1 = x_{-1} - q_1 x_0, & y_1 = y_{-1} - q_1 y_0 \\
r_0 = r_1 q_2 + r_2, & x_2 = x_0 - q_2 x_1, & y_2 = y_0 - q_2 y_1 \\
r_1 = r_2 q_3 + r_3, & x_3 = x_1 - q_3 x_2, & y_3 = y_1 - q_3 y_2 \\
\vdots & \vdots & \vdots \\
r_{s-2} = r_{s-1} q_s. & &
\end{array}$$

- A simpler way is as follows.
- Define $x_{-1} = 1$, $y_{-1} = 0$, $x_0 = 0$, $y_0 = 1$ and then lay the calculations out as follows.

$$\begin{array}{lll}
r_{-1} = r_0 q_1 + r_1, & x_1 = x_{-1} - q_1 x_0, & y_1 = y_{-1} - q_1 y_0 \\
r_0 = r_1 q_2 + r_2, & x_2 = x_0 - q_2 x_1, & y_2 = y_0 - q_2 y_1 \\
r_1 = r_2 q_3 + r_3, & x_3 = x_1 - q_3 x_2, & y_3 = y_1 - q_3 y_2 \\
\vdots & \vdots & \vdots \\
r_{s-2} = r_{s-1} q_s. & &
\end{array}$$

- The claim is that $x = x_{s-1}$, $y = y_{s-1}$. More generally $r_j = ax_j + by_j$ and this can be proved by induction.

- A simpler way is as follows.
- Define $x_{-1} = 1$, $y_{-1} = 0$, $x_0 = 0$, $y_0 = 1$ and then lay the calculations out as follows.

$$\begin{array}{lll}
r_{-1} = r_0 q_1 + r_1, & x_1 = x_{-1} - q_1 x_0, & y_1 = y_{-1} - q_1 y_0 \\
r_0 = r_1 q_2 + r_2, & x_2 = x_0 - q_2 x_1, & y_2 = y_0 - q_2 y_1 \\
r_1 = r_2 q_3 + r_3, & x_3 = x_1 - q_3 x_2, & y_3 = y_1 - q_3 y_2 \\
\vdots & \vdots & \vdots \\
r_{s-2} = r_{s-1} q_s. & &
\end{array}$$

- The claim is that $x = x_{s-1}$, $y = y_{s-1}$. More generally $r_j = ax_j + by_j$ and this can be proved by induction.
- By construction we have $r_{-1} = ax_{-1} + by_{-1}$, $r_0 = ax_0 + by_0$.

- A simpler way is as follows.
- Define $x_{-1} = 1$, $y_{-1} = 0$, $x_0 = 0$, $y_0 = 1$ and then lay the calculations out as follows.

$$\begin{array}{lll}
r_{-1} = r_0 q_1 + r_1, & x_1 = x_{-1} - q_1 x_0, & y_1 = y_{-1} - q_1 y_0 \\
r_0 = r_1 q_2 + r_2, & x_2 = x_0 - q_2 x_1, & y_2 = y_0 - q_2 y_1 \\
r_1 = r_2 q_3 + r_3, & x_3 = x_1 - q_3 x_2, & y_3 = y_1 - q_3 y_2 \\
\vdots & \vdots & \vdots \\
r_{s-2} = r_{s-1} q_s. & &
\end{array}$$

- The claim is that $x = x_{s-1}$, $y = y_{s-1}$. More generally $r_j = ax_j + by_j$ and this can be proved by induction.
- By construction we have $r_{-1} = ax_{-1} + by_{-1}$, $r_0 = ax_0 + by_0$.
- Suppose $r_j = ax_j + by_j$ is established for all $j \le k$. Then

$$\begin{aligned}
r_{k+1} &= r_{k-1} - q_{k+1} r_k \\
&= (ax_{k-1} + by_{k-1}) - q_{k+1}(ax_k + by_k) \\
&= ax_{k+1} + by_{k+1}.
\end{aligned}$$

- A simpler way is as follows.
- Define $x_{-1} = 1$, $y_{-1} = 0$, $x_0 = 0$, $y_0 = 1$ and then lay the calculations out as follows.

$$\begin{array}{lll} r_{-1} = r_0 q_1 + r_1, & x_1 = x_{-1} - q_1 x_0, & y_1 = y_{-1} - q_1 y_0 \\ r_0 = r_1 q_2 + r_2, & x_2 = x_0 - q_2 x_1, & y_2 = y_0 - q_2 y_1 \\ r_1 = r_2 q_3 + r_3, & x_3 = x_1 - q_3 x_2, & y_3 = y_1 - q_3 y_2 \\ \vdots & \vdots & \vdots \\ r_{s-2} = r_{s-1} q_s. & & \end{array}$$

- The claim is that $x = x_{s-1}$, $y = y_{s-1}$. More generally $r_j = ax_j + by_j$ and this can be proved by induction.
- By construction we have $r_{-1} = ax_{-1} + by_{-1}$, $r_0 = ax_0 + by_0$.
- Suppose $r_j = ax_j + by_j$ is established for all $j \le k$. Then

$$\begin{aligned} r_{k+1} &= r_{k-1} - q_{k+1} r_k \\ &= (ax_{k-1} + by_{k-1}) - q_{k+1}(ax_k + by_k) \\ &= ax_{k+1} + by_{k+1}. \end{aligned}$$

- In particular $(a, b) = r_{s-1} = ax_{s-1} + by_{s-1}$.

- Hence laying out the example above in this expanded form we have

$$r_{-1} = 10678,\ r_0 = 42,\ x_{-1} = 1,\ x_0 = 0,\ y_{-1} = 0,\ y_0 = 1,$$

$$\begin{array}{lll} 10678 = 42 \cdot 254 + 10, & x_1 = 1, & y_1 = -254 \\ 42 = 10 \cdot 4 + 2, & x_2 = -4, & y_2 = 1017 \\ 10 = 2 \cdot 5. & & \end{array}$$

$$(10678, 42) = 2 = 10678 \cdot (-4) + 42 \cdot (1017).$$

- Hence laying out the example above in this expanded form we have

$$r_{-1} = 10678,\ r_0 = 42,\ x_{-1} = 1,\ x_0 = 0,\ y_{-1} = 0,\ y_0 = 1,$$

$$\begin{aligned} &10678 = 42 \cdot 254 + 10, && x_1 = 1, && y_1 = -254 \\ &42 = 10 \cdot 4 + 2, && x_2 = -4, && y_2 = 1017 \\ &10 = 2 \cdot 5. \end{aligned}$$

$$(10678, 42) = 2 = 10678 \cdot (-4) + 42 \cdot (1017).$$

- It is also possible to set this up using matrices.

- Lay out the sequences in rows

$$\begin{array}{lll} r_{-1}, & x_{-1}, & y_{-1} \\ r_0, & x_0, & y_0 \\ \vdots & \vdots & \vdots \end{array}$$

- Lay out the sequences in rows

$$\begin{array}{lll} r_{-1}, & x_{-1}, & y_{-1} \\ r_0, & x_0, & y_0 \\ \vdots & \vdots & \vdots \end{array}$$

- Now proceed to compute each successive row as follows.

- Lay out the sequences in rows

$$\begin{matrix} r_{-1}, & x_{-1}, & y_{-1} \\ r_0, & x_0, & y_0 \\ \vdots & \vdots & \vdots \end{matrix}$$

- Now proceed to compute each successive row as follows.
- If the s-th row is the last one to be computed, calculate $q_s = \lfloor r_{s-1}/r_s \rfloor$.

- Lay out the sequences in rows

$$\begin{matrix} r_{-1}, & x_{-1}, & y_{-1} \\ r_0, & x_0, & y_0 \\ \vdots & \vdots & \vdots \end{matrix}$$

- Now proceed to compute each successive row as follows.
- If the s-th row is the last one to be computed, calculate $q_s = \lfloor r_{s-1}/r_s \rfloor$.
- Then take the last two rows computed and pre multiply by $(1, -q_s)$

$$(1, -q_s)\begin{pmatrix} r_{s-1}, & x_{s-1}, & y_{s-1} \\ r_s, & x_s, & y_s \end{pmatrix} = (r_{s+1}, x_{s+1}, y_{s+1})$$

to obtain the $s+1$-st row.

- Here is a simple example.

Example 2

Let $a = 4343$, $b = 973$. We can lay this out as follows

	4343	1	0
4	973	0	1
2	451	1	-4
6	71	-2	9
2	25	13	-58
1	21	-28	125
5	4	41	-183
	1	-233	1040

Thus $(4343, 973) = 1 = (-233)4343 + (1040)973$.

- We can use Euclid's algorithm to find the complete solution in integers to linear diophantine equations of the kind

$$ax + by = c.$$

- We can use Euclid's algorithm to find the complete solution in integers to linear diophantine equations of the kind

$$ax + by = c.$$

- Here a, b, c are integers and we wish to find all integers x and y which satisfy this.

- We can use Euclid's algorithm to find the complete solution in integers to linear diophantine equations of the kind
$$ax + by = c.$$
- Here a, b, c are integers and we wish to find all integers x and y which satisfy this.
- There are some obvious necessary conditions.

- We can use Euclid's algorithm to find the complete solution in integers to linear diophantine equations of the kind
$$ax + by = c.$$
- Here a, b, c are integers and we wish to find all integers x and y which satisfy this.
- There are some obvious necessary conditions.
- First of all if $a = b = 0$, then it is not soluble unless $c = 0$ and then it is soluble by any x and y, which is not very interesting.

- We can use Euclid's algorithm to find the complete solution in integers to linear diophantine equations of the kind
$$ax + by = c.$$
- Here a, b, c are integers and we wish to find all integers x and y which satisfy this.
- There are some obvious necessary conditions.
- First of all if $a = b = 0$, then it is not soluble unless $c = 0$ and then it is soluble by any x and y, which is not very interesting.
- Thus it makes sense to suppose that one of a or b is non-zero.

- We can use Euclid's algorithm to find the complete solution in integers to linear diophantine equations of the kind
$$ax + by = c.$$
- Here a, b, c are integers and we wish to find all integers x and y which satisfy this.
- There are some obvious necessary conditions.
- First of all if $a = b = 0$, then it is not soluble unless $c = 0$ and then it is soluble by any x and y, which is not very interesting.
- Thus it makes sense to suppose that one of a or b is non-zero.
- Then since (a, b) divides the left hand side, we can only have solutions if $(a, b)|c$.

- We are considering $ax + by = c$ and we are assuming that a and b are not both 0 and $(a, b)|c$.

- We are considering $ax + by = c$ and we are assuming that a and b are not both 0 and $(a, b)|c$.
- If we choose x and y so that $ax + by = (a, b)$, then we have

$$a(xc/(a, b)) + b(yc/(a, b)) = (ax + by)c/(a, b) = c$$

so we certainly have a solution of our equation.

- We are considering $ax + by = c$ and we are assuming that a and b are not both 0 and $(a, b)|c$.
- If we choose x and y so that $ax + by = (a, b)$, then we have

$$a(xc/(a, b)) + b(yc/(a, b)) = (ax + by)c/(a, b) = c$$

so we certainly have a solution of our equation.
- Call it x_0, y_0.

- We are considering $ax+by=c$ and we are assuming that a and b are not both 0 and $(a,b)|c$.
- If we choose x and y so that $ax+by=(a,b)$, then we have

$$a(xc/(a,b))+b(yc/(a,b))=(ax+by)c/(a,b)=c$$

so we certainly have a solution of our equation.
- Call it x_0, y_0.
- Now consider any other solution. Then

$$ax+by-ax_0-by_0=c-c=0,\ a(x-x_0)=b(y_0-y).$$

- We are considering $ax + by = c$ and we are assuming that a and b are not both 0 and $(a, b)|c$.
- If we choose x and y so that $ax + by = (a, b)$, then we have

$$a(xc/(a, b)) + b(yc/(a, b)) = (ax + by)c/(a, b) = c$$

so we certainly have a solution of our equation.
- Call it x_0, y_0.
- Now consider any other solution. Then

$$ax + by - ax_0 - by_0 = c - c = 0,\ a(x - x_0) = b(y_0 - y).$$

- Hence

$$\frac{a}{(a, b)}(x - x_0) = \frac{b}{(a, b)}(y_0 - y).$$

- We are considering $ax + by = c$ and we are assuming that a and b are not both 0 and $(a, b)|c$.
- If we choose x and y so that $ax + by = (a, b)$, then we have

$$a(xc/(a, b)) + b(yc/(a, b)) = (ax + by)c/(a, b) = c$$

so we certainly have a solution of our equation.
- Call it x_0, y_0.
- Now consider any other solution. Then

$$ax + by - ax_0 - by_0 = c - c = 0,\ a(x - x_0) = b(y_0 - y).$$

- Hence

$$\frac{a}{(a, b)}(x - x_0) = \frac{b}{(a, b)}(y_0 - y).$$

- Then as $\left(\frac{a}{(a,b)}, \frac{b}{(a,b)}\right) = 1$ we have by an earlier example that $y_0 - y = z\frac{a}{(a,b)}$ and $x - x_0 = z\frac{b}{(a,b)}$ for some z.

- We are considering $ax+by=c$ and we are assuming that a and b are not both 0 and $(a,b)|c$.
- If we choose x and y so that $ax+by=(a,b)$, then we have

$$a(xc/(a,b))+b(yc/(a,b))=(ax+by)c/(a,b)=c$$

so we certainly have a solution of our equation.
- Call it x_0, y_0.
- Now consider any other solution. Then

$$ax+by-ax_0-by_0=c-c=0,\ a(x-x_0)=b(y_0-y).$$

- Hence
$$\frac{a}{(a,b)}(x-x_0)=\frac{b}{(a,b)}(y_0-y).$$
- Then as $\left(\frac{a}{(a,b)},\frac{b}{(a,b)}\right)=1$ we have by an earlier example that $y_0-y=z\frac{a}{(a,b)}$ and $x-x_0=z\frac{b}{(a,b)}$ for some z.
- But any x and y of this form give a solution, so we have found the complete solution set.

- We have

Theorem 3

Suppose that a and b are not both 0 and $(a,b)|c$. Suppose further that $ax_0 + by_0 = c$. Then every solution of

$$ax + by = c$$

is given by

$$x = x_0 + z\frac{b}{(a,b)}, \quad y = y_0 - z\frac{a}{(a,b)}$$

where z is any integer.

- We have

Theorem 3

Suppose that a and b are not both 0 and $(a,b)|c$. Suppose further that $ax_0+by_0=c$. Then every solution of

$$ax+by=c$$

is given by

$$x=x_0+z\frac{b}{(a,b)},\quad y=y_0-z\frac{a}{(a,b)}$$

where z is any integer.

- One can see here that the solutions x all leave the same remainder on division by $\frac{b}{(a,b)}$ and likewise for y on division by $\frac{a}{(a,b)}$. This suggests that there may be a useful way of classifying integers.

- Here is an algorithm due to R. S. Lehmen based on differences of squares which is a small improvement on trial division.

- Here is an algorithm due to R. S. Lehmen based on differences of squares which is a small improvement on trial division.
- **1.** Apply trial division with $d = 2, 3, \ldots, d \le n^{1/3}$.

- Here is an algorithm due to R. S. Lehmen based on differences of squares which is a small improvement on trial division.
- **1.** Apply trial division with $d = 2, 3, \ldots, d \le n^{1/3}$.
- **2.** For $1 \le t \le n^{1/3} + 1$ consider the numbers x with

$$\sqrt{4tn} \le x \le \sqrt{4tn + n^{2/3}}.$$

Check each $x^2 - 4tn$ to see if it is a perfect square y^2 (compute $4tn - \lfloor\sqrt{4tn}\rfloor^2$).

- Here is an algorithm due to R. S. Lehmen based on differences of squares which is a small improvement on trial division.
- **1.** Apply trial division with $d = 2, 3, \ldots, d \le n^{1/3}$.
- **2.** For $1 \le t \le n^{1/3} + 1$ consider the numbers x with

$$\sqrt{4tn} \le x \le \sqrt{4tn + n^{2/3}}.$$

  Check each $x^2 - 4tn$ to see if it is a perfect square y^2 (compute $4tn - \lfloor\sqrt{4tn}\rfloor^2$).
- **3.** If there are x and y such that

$$x^2 - 4tn = y^2,$$

  then compute

$$GCD(x + y, n).$$

- Here is an algorithm due to R. S. Lehmen based on differences of squares which is a small improvement on trial division.
- **1.** Apply trial division with $d = 2, 3, \ldots, d \le n^{1/3}$.
- **2.** For $1 \le t \le n^{1/3} + 1$ consider the numbers x with

$$\sqrt{4tn} \le x \le \sqrt{4tn + n^{2/3}}.$$

  Check each $x^2 - 4tn$ to see if it is a perfect square y^2 (compute $4tn - \lfloor\sqrt{4tn}\rfloor^2$).
- **3.** If there are x and y such that

$$x^2 - 4tn = y^2,$$

  then compute

$$GCD(x + y, n).$$

- **4.** If there is no $t \le n^{1/3} + 1$ for which there are x and y, then n is prime.

- We already saw this in Example 1.23, but now it does not look like a fluke.

Example 4

Let $n = 10001$. Then $\lfloor (10001)^{1/3} \rfloor = 21$.
Trial division with $d = 2, 3, 5, 7, 11, 13, 17, 19$ finds no factors.
Let $t = 1$, so that $4tn = 40004$. Then

$$\lfloor \sqrt{4n} \rfloor = 200,\ \lfloor \sqrt{4n + n^{2/3}} \rfloor = \lfloor (40445)^{1/2} \rfloor = 201,$$

$$(201)^2 = 40401,\ 397 \neq y^2.$$

Let $t = 2$, so that $4tn = 80008$. Then

$$\lfloor \sqrt{8n} \rfloor = 282,\ \lfloor \sqrt{8n + n^{2/3}} \rfloor = \lfloor (80449)^{1/2} \rfloor = 283,$$

$$x = 283,\ (283)^2 - 8n = 80089 - 80008 = 81 = 9^2,$$

$$y = 9,\ x + y = 292,\ (292, 10001) = 73.$$

- The proof that Lehman's algorithm works depends on a subject called *diophantine approximation*.

- The proof that Lehman's algorithm works depends on a subject called *diophantine approximation*.
- The normal way in to this is *via* continued fractions, which in turn has some connections with Euclid's algorithm.

- The proof that Lehman's algorithm works depends on a subject called *diophantine approximation*.
- The normal way in to this is *via* continued fractions, which in turn has some connections with Euclid's algorithm.
- Fortunately we can take a short cut by appealing to

Theorem 5 (Dirichlet)

For any real number α and any integer $Q \geq 1$ there exist integers a and q with $1 \leq q \leq Q$ such that

$$\left|\alpha - \frac{a}{q}\right| \leq \frac{1}{q(Q+1)}.$$

- The proof that Lehman's algorithm works depends on a subject called *diophantine approximation*.
- The normal way in to this is *via* continued fractions, which in turn has some connections with Euclid's algorithm.
- Fortunately we can take a short cut by appealing to

Theorem 5 (Dirichlet)

For any real number α and any integer $Q \geq 1$ there exist integers a and q with $1 \leq q \leq Q$ such that

$$\left|\alpha - \frac{a}{q}\right| \leq \frac{1}{q(Q+1)}.$$

- As an immediate consequence of casting out all common factors of a and q in a/q we have

Corollary 6

The conclusion holds with the additional condition $(a, q) = 1$.

- **Proof of Lehman's algorithm.** We have to show that when there is a $d|n$ with $n^{1/3} < d \le n^{1/2}$, then there is a t with $1 \le t \le n^{1/3} + 1$ and x, y such that

$$4tn \le x^2 \le 4tn + n^{2/3},\ x^2 - y^2 = 4tn.$$

- **Proof of Lehman's algorithm.** We have to show that when there is a $d|n$ with $n^{1/3} < d \le n^{1/2}$, then there is a t with $1 \le t \le n^{1/3} + 1$ and x, y such that

$$4tn \le x^2 \le 4tn + n^{2/3},\ x^2 - y^2 = 4tn.$$

- We use Dirichlet's theorem with $\alpha = \frac{n}{d^2}$, $Q = \left\lfloor \frac{d}{n^{1/3}} \right\rfloor$.

- **Proof of Lehman's algorithm.** We have to show that when there is a $d|n$ with $n^{1/3} < d \le n^{1/2}$, then there is a t with $1 \le t \le n^{1/3} + 1$ and x, y such that

$$4tn \le x^2 \le 4tn + n^{2/3}, \; x^2 - y^2 = 4tn.$$

- We use Dirichlet's theorem with $\alpha = \frac{n}{d^2}$, $Q = \left\lfloor \frac{d}{n^{1/3}} \right\rfloor$.
- As $d > n^{1/3}$ we have $Q > 1$. Thus there are $a \in \mathbb{Z}$, $q \in \mathbb{N}$ such that $1 \le q \le Q$ and

$$\left| \frac{n}{d^2} - \frac{a}{q} \right| \le \frac{1}{q(Q+1)} < \frac{n^{1/3}}{qd}, \; \left| \frac{n}{d} q - ad \right| < n^{1/3}.$$

- **Proof of Lehman's algorithm.** We have to show that when there is a $d|n$ with $n^{1/3} < d \le n^{1/2}$, then there is a t with $1 \le t \le n^{1/3}+1$ and x, y such that

$$4tn \le x^2 \le 4tn + n^{2/3},\ x^2 - y^2 = 4tn.$$

- We use Dirichlet's theorem with $\alpha = \frac{n}{d^2}$, $Q = \left\lfloor \frac{d}{n^{1/3}} \right\rfloor$.
- As $d > n^{1/3}$ we have $Q > 1$. Thus there are $a \in \mathbb{Z}$, $q \in \mathbb{N}$ such that $1 \le q \le Q$ and

$$\left| \frac{n}{d^2} - \frac{a}{q} \right| \le \frac{1}{q(Q+1)} < \frac{n^{1/3}}{qd},\ \left| \frac{n}{d} q - ad \right| < n^{1/3}.$$

- Let $x = \frac{n}{d} q + ad$, $y = \left| \frac{n}{d} q - ad \right|$, $t = aq$.

- **Proof of Lehman's algorithm.** We have to show that when there is a $d|n$ with $n^{1/3} < d \le n^{1/2}$, then there is a t with $1 \le t \le n^{1/3} + 1$ and x, y such that

$$4tn \le x^2 \le 4tn + n^{2/3},\ x^2 - y^2 = 4tn.$$

- We use Dirichlet's theorem with $\alpha = \frac{n}{d^2}$, $Q = \left\lfloor \frac{d}{n^{1/3}} \right\rfloor$.
- As $d > n^{1/3}$ we have $Q > 1$. Thus there are $a \in \mathbb{Z}$, $q \in \mathbb{N}$ such that $1 \le q \le Q$ and

$$\left| \frac{n}{d^2} - \frac{a}{q} \right| \le \frac{1}{q(Q+1)} < \frac{n^{1/3}}{qd},\ \left| \frac{n}{d}q - ad \right| < n^{1/3}.$$

- Let $x = \dfrac{n}{d}q + ad$, $y = \left| \dfrac{n}{d}q - ad \right|$, $t = aq$.
- Then $x^2 = \dfrac{n^2}{d^2}q^2 + 2nqa + a^2d^2 = y^2 + 4tn$.

- **Proof of Lehman's algorithm.** We have to show that when there is a $d|n$ with $n^{1/3} < d \le n^{1/2}$, then there is a t with $1 \le t \le n^{1/3} + 1$ and x, y such that

$$4tn \le x^2 \le 4tn + n^{2/3},\ x^2 - y^2 = 4tn.$$

- We use Dirichlet's theorem with $\alpha = \frac{n}{d^2}$, $Q = \left\lfloor \frac{d}{n^{1/3}} \right\rfloor$.
- As $d > n^{1/3}$ we have $Q > 1$. Thus there are $a \in \mathbb{Z}$, $q \in \mathbb{N}$ such that $1 \le q \le Q$ and

$$\left| \frac{n}{d^2} - \frac{a}{q} \right| \le \frac{1}{q(Q+1)} < \frac{n^{1/3}}{qd},\ \left| \frac{n}{d}q - ad \right| < n^{1/3}.$$

- Let $x = \dfrac{n}{d}q + ad$, $y = \left| \dfrac{n}{d}q - ad \right|$, $t = aq$.
- Then $x^2 = \dfrac{n^2}{d^2}q^2 + 2nqa + a^2d^2 = y^2 + 4tn$.
- Moreover $y^2 < n^{2/3}$ and

$$t = aq < \frac{n}{d^2}q^2 + n^{1/3}\frac{q}{d} \le \frac{n}{d^2}Q^2 + n^{1/3}\frac{Q}{d} \le n^{1/3} + 1.$$

- I will not go into details but the runtime is bounded by $n^{1/3}$.

- I will not go into details but the runtime is bounded by $n^{1/3}$.
- A little more precisely, since $y^2 = x^2 - 4tn$ y is determined by t and x it suffices to bound the number of pairs t, x which need to be considered and this can be shown to be of order $n^{1/3}$.

- **Theorem 5 (Dirichlet).** For any real number α and any integer $Q \geq 1$ there exist integers a and q with $1 \leq q \leq Q$ such that $\left|\alpha - \frac{a}{q}\right| \leq \frac{1}{q(Q+1)}$.

- **Theorem 5 (Dirichlet).** For any real number α and any integer $Q \ge 1$ there exist integers a and q with $1 \le q \le Q$ such that $\left|\alpha - \dfrac{a}{q}\right| \le \dfrac{1}{q(Q+1)}$.
- **Proof.** Let I_n denote the interval $\left[\frac{n-1}{Q+1}, \frac{n}{Q+1}\right)$ and consider the Q numbers $\{\alpha\}$, $\{2\alpha\}$, $\dots$, $\{Q\alpha\}$ where we use $\{\beta\} = \beta - \lfloor \beta \rfloor$ to denote the "fractional" part of β.

- **Theorem 5 (Dirichlet).** For any real number α and any integer $Q \geq 1$ there exist integers a and q with $1 \leq q \leq Q$ such that $\left|\alpha - \frac{a}{q}\right| \leq \frac{1}{q(Q+1)}$.
- **Proof.** Let I_n denote the interval $\left[\frac{n-1}{Q+1}, \frac{n}{Q+1}\right)$ and consider the Q numbers $\{\alpha\}$, $\{2\alpha\}$, …, $\{Q\alpha\}$ where we use $\{\beta\} = \beta - \lfloor\beta\rfloor$ to denote the "fractional" part of β.
- If one of these, say $\{q\alpha\}$, lies in I_1, then we are done with $a = \lfloor q\alpha \rfloor$, and then $0 \leq q\alpha - a < \frac{1}{Q+1}$.

- **Theorem 5 (Dirichlet).** For any real number α and any integer $Q \geq 1$ there exist integers a and q with $1 \leq q \leq Q$ such that $\left|\alpha - \frac{a}{q}\right| \leq \frac{1}{q(Q+1)}$.
- **Proof.** Let I_n denote the interval $\left[\frac{n-1}{Q+1}, \frac{n}{Q+1}\right)$ and consider the Q numbers $\{\alpha\}, \{2\alpha\}, \dots, \{Q\alpha\}$ where we use $\{\beta\} = \beta - \lfloor \beta \rfloor$ to denote the "fractional" part of β.
- If one of these, say $\{q\alpha\}$, lies in I_1, then we are done with $a = \lfloor q\alpha \rfloor$, and then $0 \leq q\alpha - a < \frac{1}{Q+1}$.
- Similarly when one of them lies in I_{Q+1}, then $1 - \frac{1}{Q+1} \leq q\alpha - \lfloor q\alpha \rfloor < 1$, whence $-\frac{1}{Q+1} \leq q\alpha - (\lfloor q\alpha \rfloor + 1) < 0$ and we can take $a = \lfloor q\alpha \rfloor + 1$.

- **Theorem 5 (Dirichlet).** For any real number α and any integer $Q \geq 1$ there exist integers a and q with $1 \leq q \leq Q$ such that $\left|\alpha - \frac{a}{q}\right| \leq \frac{1}{q(Q+1)}$.
- **Proof.** Let I_n denote the interval $\left[\frac{n-1}{Q+1}, \frac{n}{Q+1}\right)$ and consider the Q numbers $\{\alpha\}$, $\{2\alpha\}$, ..., $\{Q\alpha\}$ where we use $\{\beta\} = \beta - \lfloor\beta\rfloor$ to denote the "fractional" part of β.
- If one of these, say $\{q\alpha\}$, lies in I_1, then we are done with $a = \lfloor q\alpha\rfloor$, and then $0 \leq q\alpha - a < \frac{1}{Q+1}$.
- Similarly when one of them lies in I_{Q+1}, then $1 - \frac{1}{Q+1} \leq q\alpha - \lfloor q\alpha\rfloor < 1$, whence $-\frac{1}{Q+1} \leq q\alpha - (\lfloor q\alpha\rfloor + 1) < 0$ and we can take $a = \lfloor q\alpha\rfloor + 1$.

- **Theorem 5 (Dirichlet).** For any real number α and any integer $Q \geq 1$ there exist integers a and q with $1 \leq q \leq Q$ such that $\left|\alpha - \frac{a}{q}\right| \leq \frac{1}{q(Q+1)}$.
- **Proof.** Let I_n denote the interval $\left[\frac{n-1}{Q+1}, \frac{n}{Q+1}\right)$ and consider the Q numbers $\{\alpha\}, \{2\alpha\}, \ldots, \{Q\alpha\}$ where we use $\{\beta\} = \beta - \lfloor\beta\rfloor$ to denote the "fractional" part of β.
- If one of these, say $\{q\alpha\}$, lies in I_1, then we are done with $a = \lfloor q\alpha\rfloor$, and then $0 \leq q\alpha - a < \frac{1}{Q+1}$.
- Similarly when one of them lies in I_{Q+1}, then $1 - \frac{1}{Q+1} \leq q\alpha - \lfloor q\alpha\rfloor < 1$, whence $-\frac{1}{Q+1} \leq q\alpha - (\lfloor q\alpha\rfloor + 1) < 0$ and we can take $a = \lfloor q\alpha\rfloor + 1$.
- When neither situation occurs the Q numbers will lie in the $Q-1$ intervals $I_2, \ldots, I_Q$, so there is at least one interval containing at least two (the *pigeon hole principle*.

- **Theorem 5 (Dirichlet).** For any real number α and any integer $Q \ge 1$ there exist integers a and q with $1 \le q \le Q$ such that $\left|\alpha - \frac{a}{q}\right| \le \frac{1}{q(Q+1)}$.
- **Proof.** Let I_n denote the interval $\left[\frac{n-1}{Q+1}, \frac{n}{Q+1}\right)$ and consider the Q numbers $\{\alpha\}, \{2\alpha\}, \ldots, \{Q\alpha\}$ where we use $\{\beta\} = \beta - \lfloor \beta \rfloor$ to denote the "fractional" part of β.
- If one of these, say $\{q\alpha\}$, lies in I_1, then we are done with $a = \lfloor q\alpha \rfloor$, and then $0 \le q\alpha - a < \frac{1}{Q+1}$.
- Similarly when one of them lies in I_{Q+1}, then $1 - \frac{1}{Q+1} \le q\alpha - \lfloor q\alpha \rfloor < 1$, whence $-\frac{1}{Q+1} \le q\alpha - (\lfloor q\alpha \rfloor + 1) < 0$ and we can take $a = \lfloor q\alpha \rfloor + 1$.
- When neither situation occurs the Q numbers will lie in the $Q-1$ intervals $I_2, \ldots, I_Q$, so there is at least one interval containing at least two (the *pigeon hole principle*.
- Thus there are q_1, q_2 with $q_1 < q_2$ such that $|(\alpha q_2 - \lfloor \alpha q_2 \rfloor) - (\alpha q_1 - \lfloor \alpha q_1 \rfloor)| < \frac{1}{Q+1}$.

- **Theorem 5 (Dirichlet).** For any real number α and any integer $Q \ge 1$ there exist integers a and q with $1 \le q \le Q$ such that $\left|\alpha - \frac{a}{q}\right| \le \frac{1}{q(Q+1)}$.
- **Proof.** Let I_n denote the interval $\left[\frac{n-1}{Q+1}, \frac{n}{Q+1}\right)$ and consider the Q numbers $\{\alpha\}, \{2\alpha\}, \ldots, \{Q\alpha\}$ where we use $\{\beta\} = \beta - \lfloor \beta \rfloor$ to denote the "fractional" part of β.
- If one of these, say $\{q\alpha\}$, lies in I_1, then we are done with $a = \lfloor q\alpha \rfloor$, and then $0 \le q\alpha - a < \frac{1}{Q+1}$.
- Similarly when one of them lies in I_{Q+1}, then $1 - \frac{1}{Q+1} \le q\alpha - \lfloor q\alpha \rfloor < 1$, whence $-\frac{1}{Q+1} \le q\alpha - (\lfloor q\alpha \rfloor + 1) < 0$ and we can take $a = \lfloor q\alpha \rfloor + 1$.
- When neither situation occurs the Q numbers will lie in the $Q - 1$ intervals $I_2, \ldots, I_Q$, so there is at least one interval containing at least two (the *pigeon hole principle*.
- Thus there are q_1, q_2 with $q_1 < q_2$ such that $|(\alpha q_2 - \lfloor \alpha q_2 \rfloor) - (\alpha q_1 - \lfloor \alpha q_1 \rfloor)| < \frac{1}{Q+1}$.
- We put $q = (q_2 - q_1)$, $a = (\lfloor \alpha q_2 \rfloor - \lfloor \alpha q_1 \rfloor)$.