Behind The Two Classic Results Of Euclid About Primes

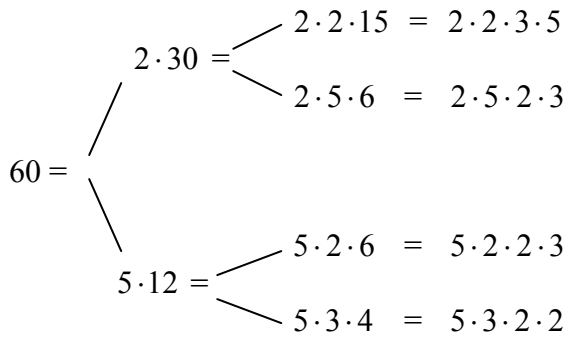
Everybody even outside mathematics heard about the fact that Euclid axiomatized Geometry. Laid down the simple facts from which geometrical theorems can be derived by Logic. The particular difficulties to find the axiom for parallelity and the general difficulties to apply Logic are less known. But amazingly a much simpler surprise should be for outsiders that Euclid also laid down the logical build up of Number Theory too. This theory deals with the natural numbers : $1 , 2 , 3 , 4 , \ldots$

D

Composite numbers are those that can be "decomposed", that is written as product of two smaller numbers. For example: $4 = 2 \cdot 2$ or $6 = 2 \cdot 3$.
There are numbers that can not be decomposed and these are called the primes except the number 1. So the primes are: $2 , 3 , 5 , 7 , 11 , 13 , 17 , \ldots$

R

The reason why we left 1 out is crucial and relates to the whole importance of primes. Namely, if we decompose the composites not merely into two numbers but those again into products as far as possible, then actually we must end up with numbers that are not decomposable any more. And here we shouldn't use the number 1 ever because it would just keep the other member the same. So, such total decomposition doesn't need the number 1. But what's much more interesting is that these decompositions are unique too. That is, the final primes are the same regardless in what order we did the decompositions:



This uniqueness claim is the Fundamental Theorem Of Arithmetic and it follows very easily from Euclid's second theorem we'll show. This claims that if a p prime divides a product $n\,b$ then p must divide n or b maybe both of them.
So what we claim is that primes divide separately.
Then for two prime products we can regard all the primes on one side as dividers of the other side and then being a divider now actually means being a member there too.
This then shows that the two sides are the same.
So the essence of the unique prime decomposition is that "primes divide separately".
The modern expression for this decomposition is by the way "factorization".

D

A factor of a number is any divider of it that is not the trivial 1.
But the also trivial divider as the number itself is allowed as factor.
For example: 1 has no factor, 2 has the only factor 2, 4 has factors $2 , 4$.
So the primes have only themselves as factors while the composites can be all factorized into primes. And The Fundamental Theorem Of Arithmetic is then U.P.F.T = Unique Prime Factorization Theorem.

R The concept of factor as non 1 divider, quite interestingly will be also very useful for the other simpler theorem that Euclid discovered.
This simply claims the: Infinity Of Primes.
It's amazing that the prime factorization does not imply in itself that there has to be infinite many primes. We could have all the infinite many numbers factorizable from finite many primes only. It would make Number Theory much simpler.
The real reason why there has to be infinite many primes is the fact that for any finite many p_1 , p_2 , . . . , p_k primes we can make a number that is definitely not dividable by any of these. First, we can simply multiply these together to get a P product which of course is dividable by each of them. And then adding 1 to this, that is forming P + 1 we got one that has 1 remainder to each p_1 , p_2 , . . . , p_k .
But we are not finished! This P + 1 doesn't have to be a prime just because it is not dividable by the listed ones. It just means that it can not have those as prime dividers.
Our relentless decomposition till primes of course means that this P + 1 has to be also decomposable into primes and all these would be different from our listed ones.
So we proved our point that the assumed finite many primes can not be all.
But the real argument for the contradiction was not really needing this decomposition process. If we knew somehow else that every number must have at least one prime factor then it would be enough to conclude that P + 1 must have a new prime factor.
Euclid was aware of this too though not as clearly as Kummer in 1878 when he made this newest version solely getting the contradiction from a prime factor existence.
The most elegant way to show that every number has a prime factor is by simply looking for the smallest factor. That is, for the smallest non 1 divider:

D min fact (n) = first non 1 that divides n.
For any n > 1 number this min fact (n) must exist because n itself is a non 1 divider. And now comes the sweet surprise:

T min fact (n) is always a prime.

P Indeed, if it were a composite then it had an s factor so that s < min fact (n).
But s would be a factor of n too, contradicting that min fact (n) was the smallest.

So we can make our arguments for the infinity of primes rock solid:

T There are infinite many primes.

P For any $p_1, p_2, . . , p_k$ primes, min fact ($p_1 \cdot p_2 \cdot . . . \cdot p_k + 1$) is a new prime.
Indeed, it's a prime by the previous theorem and its new because p_1 , p_2 , . . . , p_k do not divide $p_1 \cdot p_2 \cdot . . . \cdot p_k + 1$.

R This implies that for any n number there is a prime after n but we didn't exactly tell how far after an n we have to go to find one.
To sharpen in this direction, we should give a prime as close to n as possible.
But this search is difficult for the following reasons:

Let n^* denote the next prime after n and n_* the last prime up to n allowing n itself if it is a prime. Then $\delta(n) = n^* - n_*$ is the prime gap at n.

These gaps can be very big, in fact arbitrary big but even relative to n can vary a lot. So to guarantee a prime for sure is pretty hard. The first serious and very hard result was by Chebishev and it merely guaranteed a prime before $2\,n$.
This seems ridiculously weak but the sad fact is that the complicatedness of his proof shows how difficult must be anything better. He was also the first to get exact results about the already guessed limit behavior of these gaps.
$\delta(n)$ is in average $\log(n)$ around n.
This denotes logarithm with base $e = 2.718\ldots$ The base 10 logarithm is simply the number of 0-s in a 10 power and so to get a feel of the prime density is very easy by multiplying this 0 number by $\log(10) \approx 2.3$. So for example, around one million $\delta(1000{,}000) \approx 6 \times 2.3 \approx 14$ so every 14-th number is a prime in average.
The meaning is clear but to tend both n and some surrounding to infinity is complicated and a simpler averaging would be to regard the δ values not around, rather up to n. This average of course feels much smaller than an around average because the gaps are increasing.
Surprisingly, the difference is only 1. So the average up to n is $\log(n) - 1$.

And if $\pi(n)$ denotes the number of primes up to n then this average is $\frac{n}{\pi(n)}$.

So the limit law is $\frac{n}{\pi(n)} \approx \log(n) - 1$.

Amazingly, to prove this, it was enough to prove the consequence that the ratio of the two sides tends to 1 which is denoted by $\sim$.
This of course is true without the -1 part and so:

$$\frac{n}{\pi(n)} \sim \log(n) \quad \text{or} \quad \pi(n) \sim \frac{n}{\log(n)}.$$

This is called the “Prime Number Theorem”. A very misleading name!
Actually the $\pi(n)$ number of primes was simply abbreviated as “prime number”.
So the name actually means $\pi(n)$ - Theorem and not The Theorem about Prime Numbers in general or in most importance. In that sense it should mean the U.P.F.T.
Now back from tendencies to actually limit the gaps:
A Conjecture of Cramer claims that: $n > 7 \rightarrow \delta(n) < (\log(n))^2$.
This also means that: $p \geq 11 \rightarrow p^* < p + (\log(p))^2$.
A consequence of this would be that: $p \geq 127 \rightarrow p^* < p + \sqrt{p}$.

So now we arrived at the second Euclid theorem that primes divide separately.
Not surprisingly, it comes out of examining common multiples, so when products are not unique.

D c is a common multiple of a and b if $m\,a = n\,b = c$.
These m, n values are called the multipliers for our c.
The k-th common multiple of a and b is denoted as $[a, b]_k$.
So $[a, b]_1$ is the first or smallest common multiple of a and b.

T $[a, b]_k = [a, b]_1\, k$

P First of all, $[a, b]_1\, k$ is a common multiple because $\frac{[a,b]_1}{a}$ and $\frac{[a,b]_1}{b}$ are whole

numbers and $\frac{[a,b]_1}{a}\ a = \frac{[a,b]_1}{b}\ b = [a,b]_1$ implies:

$(\frac{[a,b]_1}{a}\ k)\ a = (\frac{[a,b]_1}{b}\ k)\ b = [a,b]_1\ k.$

There can be no common multiple under $[a,b]_1$ by its definition. So all we have to show that there can be no common multiple between $[a,b]_1\ k$ and $[a,b]_1\ (k+1)$. Any such number is $[a,b]_1\ k + u$ with $a\ u < [a,b]_1$. Such u not being a common multiple implies that it has positive remainder for a or b, maybe both.
But the remainder of $[a,b]_1\ k + u$ for a and b are the same as of u, so at least one is positive again.

D a and b are relative primes if their only common divider is 1.

T The multipliers for $[a,b]_1$ are relative primes.
But the multipliers for any other $[a,b]_k$ are not relative primes.

P Suppose d divides $\frac{[a,b]_1}{a}$ and $\frac{[a,b]_1}{b}$. Then $\frac{[a,b]_1}{a\,d}$ and $\frac{[a,b]_1}{b\,d}$ are wholes and $\frac{[a,b]_1}{a\,d}\ a = \frac{[a,b]_1}{b\,d}\ b = \frac{[a,b]_1}{d}$ so $\frac{[a,b]_1}{d}$ is a common multiple.
But $[a,b]_1$ is the smallest so $d = 1$.
For other $c = [a,b]_k = [a,b]_1\ k$ with $k > 1$ we have k as common divider of both $\frac{c}{a}$ and $\frac{c}{b}$.

D (a,b) denotes the greatest common divider of a and b.
So, a and b are relative primes if $(a,b) = 1$.

T $[a,b]_1 = \frac{a\,b}{(a,b)}$

P For any d common divider of a and b, $\frac{a\,b}{d}$ is a common multiple because:
$\frac{b}{d}\ a = \frac{a}{d}\ b = \frac{a\,b}{d}$. So, in particular $\frac{a\,b}{(a,b)}$ is a common multiple too.

But for this the multipliers are: $\frac{\frac{ab}{(a,b)}}{a} = \frac{b}{(a,b)}$ and $\frac{\frac{ab}{(a,b)}}{b} = \frac{a}{(a,b)}$.

If d is a common divider of these then $\frac{b}{(a,b)\,d} = m$ and $\frac{a}{(a,b)\,d} = n$.

So $(a,b)\,d$ is a common divider of a and b.

But (a,b) is the greatest so $d = 1$.

Thus by previous theorem $\frac{a\,b}{(a,b)}$ has to be $[a,b]_1$.

T If a , b are relative primes and $m\,a = n\,b$ then $m = k\,b$ and $n = k\,a$.

P $m\,a = n\,b = [a\,,\,b]_k = [a\,,\,b]_1\,k = \frac{a\,b}{(a\,,\,b)}\,k = \frac{a\,b}{1}\,k = a\,b\,k.$

T Primes divide separately: If p divides $n\,b$ then p divides n or b , maybe both.

P We can use p as a in the previous theorem.
Dividing $n\,b$ means $m\,a = n\,b$.
Being a prime implies that if a doesn't divide b then a , b are relative primes.
So we can use the theorem and imply that $n = k\,a$ so indeed a divides n.

R It took a pretty long detour into the common multiples to achieve what we wanted.
But actually there is a very straightforward road into this same field, giving the previous theorem about relative primes instantly.
And this goes to the heart of Number Theory.
The amazing truth is that all number theoretical proofs are "simply" inductions.
Let $R\,(x_1, x_2\,, . . , x_k\,)$ be a relation that is true for all values written into these variables. So then using the $\forall$ quantor for "every" or "universality", we must prove the $\forall\, x_1\, \forall\, x_2\, . .\, \forall\, x_k\, R\,(x_1, x_2\,, . . , x_k\,)$ statement.
Of course R may contain logical symbols like "and" , "or", $\rightarrow$ for implication, $\neg$ for "not" and also the other quantor $\exists$ for existence that is "there is".
In fact, R may contain further $\forall$ claims that are not used inductively.
The idea is very simple. We prove first R for some initial $x_1, x_2\,, . . , x_k$ values and then step by step widen these to all possible values. This step by step widening of course can not be improvised, it has to be by a single method.
So we try to find an $f\,(x_1, x_2\,, . . , x_k\,)$ function and the $x_1, x_2\,, . . , x_k$ values for which f is itself under some increasing $g\,(n)$ values will be the widening.
The start is defined by any f_o value of f. So our initial claim is :

$\{f_o\} = \forall\, x_1\, \forall\, x_2\, . .\, \forall\, x_k\, [\; f\,(x_1, x_2\,, . . , x_k\,) \leq f_o \rightarrow R\,(x_1, x_2\,, . . , x_k\,)\;]$

Then we must prove the widening, that is: $\forall\, v\; \{\; \{v\} \rightarrow \{\, g\,(v)\,\}\;\} = \forall\, v\; \{$

$\forall\, x_1\, \forall\, x_2\, . .\, \forall\, x_k\, [\; f\,(x_1, x_2\,, . . , x_k\,) \leq v \rightarrow R\,(x_1, x_2\,, . . , x_k\,)\;] \rightarrow$

$\forall\, x_1\, \forall\, x_2\, . .\, \forall\, x_k\, [\; f\,(x_1, x_2\,, . . , x_k\,) \leq g(v) \rightarrow R\,(x_1, x_2\,, . . , x_k\,)\;]\;\}$

If $g(v) > v$ that is g is an increasing function then it grows to infinity and so all f values and thus all possible $x_1, x_2\,, . . , x_k$ value combinations will be covered too.
Of course it's not obvious at all why all universal truths should hide a single provable method of widenings. And indeed as it turned out, there must be statements not provable this way. But this is the only tool we have at present inside Number Theory.

In the simplest applications $f_o = 1$ and $g\,(v) = v + 1$.
Also observe that if R is a $P \rightarrow Q$ implication then $f \leq v \rightarrow (P \rightarrow Q)$ is equivalent with: $(\,P$ and $f \leq v\,) \rightarrow Q$.
Indeed, $A \rightarrow B$ simply means $\neg A$ or B. So: $A \rightarrow (\,B \rightarrow C\,) =$
$\neg A$ or $(\neg B$ or $C\,) = (\neg A$ or $\neg B)$ or $C = \neg\,(\,A$ and $B)$ or $C = (\,A$ and $B\,) \rightarrow C$.
In our concrete case the $P \rightarrow Q$ claim is:

$$[\;(\,(a\,,\,b) = 1 \text{ and } m\,a = n\,b\,) \rightarrow \exists k\,(\,m = k\,b \text{ and } n = k\,a\,)\;]$$

One way to go is by using all the a , b, m , n universal variables as inductive variables and then the "magic" $f(a , b , m , n)$ function can be the common multiple value which already appears in P, so we just have to add that it is $\le v$:

$\{v\} = \forall a \forall b \forall m \forall n [((a , b) = 1 \text{ and } m a = n b \le v) \rightarrow \exists k (m = k b \text{ and } n = k a)]$

The f_o initial value will be 1 and the g widening will be $v + 1$.
So the initial a , b , m , n values are defined by $m a = n b \le 1$.
Thus $\{1\}$ is trivial! a , b , m , n must be all 1 and then $k = 1$ is an existence.
To be precise, this "must be 1" meant that if a , b , m , n are not all 1 then $m a = n b \le 1$ is false and so $\{1\}$ is true anyway.

Now comes the harder part, that: $\{v\} \rightarrow \{v + 1\}$.
We associate with every a , b , m , n values in $\{v + 1\}$ some a', b', m', n' values so that their f decreases.
Then for these a', b', m', n' we can apply $\{v\}$ and claim the k-s .
Finally we'll show that these k-s work as k for $\{v + 1\}$ too.
The particular trick is to keep the smaller one of a , b and of m , n and replace the bigger one with their difference. So the new pairs are:
$a , b - a$ or $a - b , b$ according to $a < b$ or $a > b$.
$m , n - m$ or $m - n , n$ according to $m < n$ or $m > n$.

Of course, from $m a = n b$ it's obvious that these accordings are actually opposite.
This assumption of a and b being different is luckily not a problem.
Indeed, if $a = b$ then $(a , b) = 1$ implies at once that $a = b = 1$ and so $m a = n b$ implies $m = n$. Then $k = m = n$ is an existence solution for our claim.

The crucial common multiple inheritance and its decrease follows by:
$m a = n b \le v + 1 \rightarrow (m - n) a = n (b - a) \le v$ or $m (a - b) = (n - m) b \le v$.
The relative primeness inheritance is trivial because all common dividers of a and b inherit to the smaller and the difference.
The final step is to see that the guaranteed k for the replacement pairs is good for the original a , b , m , n too. This is so because one of a , b and m , n remained oppositely and these define the k that will automatically be good for the other two.

A simpler way is leaving the $\forall m \forall n$ universalities as inner business, so only regard a , b as inductive variables. Then $f(a , b)$ can be chosen as simply the bigger of them, $\max(a , b)$ that means any of them if they are equal.
The $\{1\}$ start now means $\max(a , b) = 1$ so $a = b = 1$. And m , n could be anything.
But of course $m a = n b$ implies $m = n$ and so this common value is a k.
$\{v\} = \forall a \forall b$
$\max(a , b) \le v \rightarrow \forall m \forall n [((a , b) = 1 \text{ and } m a = n b) \rightarrow \exists k (m = k b \text{ and } n = k a)]$

This now implies $\{v + 1\}$ easier in spite of having infinite many m , n options.
We can not have $a = b$ because it contradicts $(a , b) = 1$ so say $b < a = v + 1$.
Then $a - b , b \le v$ and also $a - b , b$ are relative primes. Plus:
$m a = n b \rightarrow m (a - b) = (n - m) b$ so we can use $\{v\}$, getting that:
$\exists k (m = k b \text{ and } n - m = k (a - b) = k a - k b = k a - m)$.
So for this k we also have $n = k a$ automatically.

Induction is very visual as widening but is a bit over complicated.
We can strip the process to its bare essence which was to associate new $x_1', x_2', \ldots, x_k'$ values to old ones and prove that with these the f value decreases. The step of the k usability was merely a side issue.

In general this simply means that the R relation must inherit from $R(x_1', x_2', \ldots, x_k')$ back to $R(x_1, x_2, \ldots, x_k)$.
This can be reversed and then the initial values are irrelevant with the following twist:
We regard as claim not that R is universal rather that $\neg R$ is impossible.
A simple general fact is that $A \to B$ is equivalent with $\neg B \to \neg A$ so the inheritance from $R(x_1', x_2', \ldots, x_k')$ to $R(x_1, x_2, \ldots, x_k)$ is equivalent with inheritance from $\neg R(x_1, x_2, \ldots, x_k)$ to $\neg R(x_1', x_2', \ldots, x_k')$.
Then the whole argument becomes simpler:
R becomes true for all values because for any assumed $\neg R(x_1, x_2, \ldots, x_k)$ case we demonstrated a new $\neg R(x_1', x_2', \ldots, x_k')$ case with smaller f value.
But the f values are positive so they can not decrease infinitely.
The real essence is that our association from $x_1, x_2, \ldots, x_k$ to $x_1', x_2', \ldots, x_k'$ has no limit and this contradicts the f decrease.
To see this method in a better light we can avoid this imaginary infinite descent and use a more directly indirect approach. Then we say: Suppose there were cases where $\neg R$ is true! Then there has to be one with minimal f value. But this is impossible because $x_1', x_2', \ldots, x_k'$ always exists and gives a lower f value.
So "infinite descent" is the same as the single indirect use of minimality.
The point is that this indeed simplifies our arguments sometimes.
To illustrate this we revisit Euclid's result about the infinity of primes.
As I said, Kummer realized first that merely the existence of prime factors is enough.
For this itself the indirect minimality is perfect:

T Every number has a prime factor.

P Suppose there were numbers without prime factors and let the minimal such be m.
This can not be a prime because then it were a prime factor of itself.
So $m = a\,b$ with $a, b < m$.
a can not have prime factor because it were prime factor of m too.
So a is a smaller number than m having no prime factor which contradicts m.

R After this we can use the same proof we used earlier except now we just claim that $p_1 \cdot p_2 \cdot \ldots \cdot p_k + 1$ must have a prime factor and it can not be any of these.

This raises the idea that maybe the two proofs could be combined as a single indirect minimality. To get a single variable is easy because having infinite many primes simply means to have a prime after any n. The negative means to have an n that it has no prime after. The first of such is simply the first number after the biggest of the $p_1, p_2, \ldots, p_k$ primes. This gives no extra feature from which we could get an earlier one, so we are clueless and have to use the given finite many primes in a positive manner as we did before. So the indirect minimality is not a single universal method.

Infinite descent or indirect minimality became the fashionable method after Fermat discovered the claim called today as his "last" or more appropriately "lost" theorem.
It's very interesting to recognize the steps from Fermat to Euler and then to Gauss.
Fermat the judge and "amateur" mathematician rediscovered the old Chinese results in more general forms plus a lot more! But proofs were already a "must" in this western scientific atmosphere and this forced Fermat to lie. This is evident from his letters and yet I regard him as the biggest number theoretical genius ever.
The only exact proof he ever produced was the easiest $n = 4$ case of his infamous lost theorem, but the method he invented, the infinite descent remained the same for all future trials. This is amazing because stepping out of the naturals did happen already by Euler. His proof for the $n = 3$ case was faulty which he himself felt and tried to fix

later. Gauss saw clearly how the introduction of new numbers should be made precise but he did not see either what new arguments may this bring in beside the indirect minimality. And the shocking truth is that we still don't know this. So Fermat's Last Theorem is still a mystery. The outside proof by elliptic functions does not reveal whether a simple inside indirect minimality can do the job or not. So history just reinforced how crazy Gauss was when brushed off the outsiders who tried to provoke him about Fermat's Last Theorem. Him saying that there will be always unsolved problems avoids the point why this particular is so hard. The moral of the story for me is that no genius should pretend to be above common sense. Whatever can be explained can be explained clearly but this clarity is not identical with exactness alone. The new exactness of Gauss was not exact itself. He only knew instinctively what he regarded as exact proof. His distorted views dictate the present formalism of mathematics too. But most surprisingly, this false formalism is the "original sin" of all human science. Euclid's parallelity investigations and Newton's creation of mathematical physics all lack a simple honesty. To admit this is "heresy". The false idolization of these geniuses comes from the stupidity of the epigones, the mediocrity of the education system. Teachers simply do not know what they teach.

And all this happens because true teaching, elucidation, explaining has no social value.

There are no Nobel prizes for breakthroughs in clarifications. Understanding is the biggest tool for world peace but the peace prizes are also given for actions that avoid understanding the real problems.

Society simply strives for stupidity. Obedience and consumption is its only real goal.

Newton, Gauss, Einstein, Gödel were monster geniuses. Some don't even know why they were geniuses, some who know may not know the dirty sides. Finally, some can know both but excuse them as geniuses.

Of course, we should invent a new kind of meaning for this "monster" because it is very different from the over simplified everyday meaning. They were the most honest people in some respect and also they became victims of their own monstrosity.

Their honesty, the belief in scientific truth could make them even look like saints.

But to have such ability, to see the truths is an unquestioned black hole, they tip toe around for their whole lives. And this question is not about themselves, it's about the others. And not just the others who saw but all who could see. The average man.

The phony humbleness had a perfect outcome at Newton who said: "If I saw further then others it was because I stood on the shoulders of giants". How sweet! But it would have been better if he tried harder to explain to laymen what he saw.

The being victims part is the most amazing because it is true for them not just as private people but as scientists. These geniuses showed major blindness in their own fields and in quite easy details. But these details are themselves taboos, never mentioned or emphasized.

The real point is quite simple: Understanding is universal! Everybody can understand everything. Everybody can understand the Newton Laws just as well as Newton saw them. In fact much better. The same goes for Relativity and Einstein! This is the real evolution of man, not the seeing further at a given moment.

Here I am talking about so revolutionary things that they are more important than any religion, science or philosophy ever. But I must return to the sweet white liar Fermat.

His mentioned theorem claims that $x^n + y^n = z^n$ is impossible if $n > 2$.

For $n = 2$ the simplest example is $3^2 + 4^2 = 5^2$ and actually there are an infinite many such Pythagorean triplets. They were investigated long before Pythagoras by the Babylonians. Observe that 5 itself, is $1^2 + 2^2$ and indeed being square sum relates to giving Pythagorean triplets. Quite amazingly, Fermat had an other discovery about this square sumness too. He claimed that exactly those primes are square sums that are $4k + 1$ forms: $5 = 4 \cdot 1 + 1 = 1^2 + 2^2$, $13 = 4 \cdot 3 + 1 = 2^2 + 3^2$, . .

The k value of course does not reveal the square values.

These two discoveries of Fermat interrelate and I wrote a separate article about that.

Now the main concern is the method of infinite descent or indirect minimality.
As I said, Fermat invented this for proving the $n = 4$ case.
The $n = 3$ case is thousand times more difficult and the few increasing exponents all required even more tricks. Yet they all used indirect minimality.
The f size value is quite simple and common as $x\ y\ z$. So, all methods produced new x' , y' , z' cases with smaller product and still satisfying the impossible equality.
Of course, assuming $x^n + y^n = z^n$ for $n > 2$ is a dark walk in impossibilities.
Still, we can simplify our impossible equation with any common factor of x, y, z .
In other words, we can assume that these are relative primes.
This also means that only one of them can be even. But actually one must be even because two odds on the left side combines into even.
This was also the start for Euler to solve the $x^3 + y^3 = z^3$ case.

If the even one is z then let $x < y$ and let $p = \frac{x+y}{2}$ and $q = \frac{y-x}{2}$.

If one of x , y is the even, then let it be x and let $p = \frac{z-y}{2}$ and $q = \frac{y+z}{2}$

Calculating $2\,p\,(p^2 + 3\,q^2)$ we get z^3 in the first case, x^3 in the second.

Showing just the first case: $2\,p\,(p^2 + 3\,q^2) =$

$$(x+y)\left(\frac{x^2+y^2+2xy}{4} + 3\,\frac{y^2+x^2-2xy}{4}\right) = (x+y)\,(x^2+y^2-xy) =$$

$$(x+y)\,((x+y)^2 - 3xy) = ((x+y)^3 - 3x^2y - 3xy^2) = x^3 + y^3 = z^3$$

So $2\,p\,(p^2 + 3\,q^2) =$ cube no matter what. Also observe that in both cases:
1.) $p + q = y$ is odd, so p , q have opposite parity and so $p^2 + 3\,q^2$ is odd too.
2.) p and q are relative primes because their sum y, and difference x, are too.
We claim that $2\,p$ and $p^2 + 3\,q^2$ can only have c common prime factor as 3.
Obviously c can't be 2 because $p^2 + 3\,q^2$ is odd. So c must divide p.
Thus, dividing $p^2 + 3\,q^2$ it must divide $3\,q^2$ too. But it can't divide q because p and q are relative primes, so it must divide 3 that is be 3.
So, we have two cases. When $2\,p$ and $p^2 + 3\,q^2$ are relative primes and when they are both 3 multiples. This second case easy from the first so we continue with them being relative primes. This then implies that both $2\,p$ and $p^2 + 3\,q^2$ are cubes.
We might think we hit the jackpot and these will be two of the new x'^3 , y'^3 , z'^3 values. Unfortunately, not. A whole new direction is pursued by claiming that:

$p = a^3 - 9\,a\,b^2$ with some natural b and natural or negative a.

Observe that: $a^3 - 9\,a\,b^2 = (a^2 - 9\,b^2)\,a = (a + 3\,b)\,(a - 3\,b)\,a$
and also: $a^3 - 9\,a\,b^2 = (9\,b^2 - (-a)^2)\,(-a) = (3\,b - a)\,(3\,b + a)\,(-a)$

If a is natural then the first form contains all naturals and if a is negative then the second. Plus a and b are such that these are odd relative primes.
So then $2\,p = (a + 3\,b)\,(a - 3\,b)\,2\,a$ or $(3\,b - a)\,(3\,b + a)\,2\,(-a)$ is a cube with relative prime factors and thus these factors are cubes themselves.
In the first case : $x'^3 = a + 3\,b$, $y'^3 = a - 3\,b$, $z'^3 = 2\,a$.
In the second case: $x'^3 = 3\,b + a$, $y'^3 = 2\,(-a)$, $z'^3 = 3\,b - a$.

In both case we have trivially that : $x'^3 + y'^3 = z'^3$.

Plus it's easy to see that $x' y' z' < x y z$.
The big question is where this totally "ad hoc" $p = a^3 - 9 a b^2$ came from.
And why we pursued only $2 p$ not the other earlier member $p^2 + 3 q^2$.
Amazingly, this forgotten member is the real deal!
The big picture is that just like for the Pythagorean triplets we needed the square sums, here for the cube sums we need the $p^2 + 3 q^2$ square plus triple square numbers.
In short, we should just call them triple square sums!
I can at once reveal an amazing fact that corresponds to Fermat's earlier mentioned discovery about the $4 k + 1$ primes being square sums.
Now the $6 k + 1$ primes are the triple squares sums: $7 = 6 \cdot 1 + 1 = 2^2 + 3 \cdot 1^2$,
$13 = 6 \cdot 2 + 1 = 1^2 + 3 \cdot 2^2$, $19 = 6 \cdot 3 + 1 = 4^2 + 3 \cdot 1^2$, . . .
Poor Fermat would have been amazed to see this.
But most amazingly, both of these are irrelevant for our goals!
So already the $4 k + 1$ prime form was not needed for the Pythagorean triples and now again the $6 k + 1$ prime form is not needed for Euler's proof.
So the primes having these really simple characterizations are part of some bigger pictures that actually we still don't see quite clearly.
The immediately useful point is that the square sumness or triple square sumness inherits from the primes to all products and in reverse too, those products will only have such factors in general.
In fact, it's quite easy to tell what special square sums or triple square sums can be obtained in this way. They must be "simple", meaning having relative prime members.
This is a logical name because if in $a^2 + b^2$ the a , b members have a $c > 1$ common factor then c^2 could be brought out to "simplify" our form. But for the triple square sums an extra condition of simplicity is that it should be odd too.
As it turned out, the triple square sums brought out an important new feature not recognized for square sums earlier. To see this, observe that square sums or triple square sums, in fact any $a^2 + m b^2$ forms have a trivial inheritance to products:

$$(a^2 + m b^2)(c^2 + m d^2) = a^2 c^2 + a^2 m d^2 + m b^2 c^2 + m^2 b^2 d^2 =$$

$$= a^2 c^2 + m^2 b^2 d^2 + 2 m a c b d + m a^2 d^2 + m b^2 c^2 - 2 m a d b c =$$

$$= (a c + m b d)^2 + m (a d - b c)^2 .$$

Observe the rule:

These are algebraic rules that apply to any negative or zero numbers too.
In fact, they do lead to zero second members even from naturals if $a d - b c = 0$.
Which simply means $\frac{a}{c} = \frac{b}{d}$ so the two multiplied forms are a common f factored in both members. So one is f^2 times the other and so indeed their product should be a square number as it comes out from our rules if the second member is zero.
This of course doesn't mean that this square number couldn't be a proper m square sum with all natural members too.
The Pythagorean third members are all such, like: $5^2 = 3^2 + 4^2$
At triple square sums the first such is even simpler: $4^2 = 1^2 + 3 \cdot 1^2$

This of course wasn't "simple" at all by our definition of having relative prime members and being odd. It is even! The first simple example is: $49 = 1^2 + 3 \cdot 4^2$.
So these squares are always examples of numbers that have multiple forms:
$5^2 = 3^2 + 4^2 = 5^2 + 0^2$ and $7^2 = 1^2 + 3 \cdot 4^2 = 7^2 + 3 \cdot 0^2$
But multiple forms can happen without 0 members. In fact, if we check out bigger and bigger numbers many multiple forms are the typical.
Surprisingly but easily provably, the primes can only have unique empirical forms but when multiplied with other primes will already lead to multiple forms!
The first examples for the square sums and the triple square sums are:

$65 = 5 \cdot 13 = 1^2 + 8^2 = 4^2 + 7^2$ and $91 = 7 \cdot 13 = 4^2 + 3 \cdot 5^2 = 8^2 + 3 \cdot 3^2$.

We really can't see any connections between the two forms in either case.
And yet the connection is that they both come out with our established product manufacturing rules if we allow negative input numbers too. Of course, we need the numbers as products of their prime factors which have their own empirical forms:
$65 = 5 \cdot 13 = (1^2 + 2^2)(2^2 + 3^2) = 1^2 + 8^2 = 4^2 + 7^2$
$91 = 7 \cdot 13 = (2^2 + 3 \cdot 1^2)(1^2 + 3 \cdot 2^2) = 4^2 + 3 \cdot 5^2 = 8^2 + 3 \cdot 3^2$
This didn't help much! A better thing is to envision the forms of the factors under each other because our rules were formulated this way:

$1^2 + 2^2$ and $2^2 + 3 \cdot 1^2$
$2^2 + 3^2$ $1^2 + 3 \cdot 2^2$

And indeed $1 \cdot 2 + 2 \cdot 3 = 8$ and $1 \cdot 3 - 2 \cdot 2 = -1$
So the negativity entered though unimportantly because the squares are 8^2 and 1^2.
But at least we obtained the first product form by the multiplication rule.
A bit of trial and error will soon show that using negatives for the input numbers as well, can give the second form too, namely as:

$(-1) \cdot 2 + 2 \cdot 3 = 4$ and $(-1) \cdot 3 - 2 \cdot 2 = -7$.

To avoid lucky trials is easy. We should list all possible sign combinations:

1	2	−1	2	1	−2	−1	−2
2	3	2	3	2	3	2	3
1	2	−1	2	1	−2	−1	−2
−2	3	−2	3	−2	3	−2	3
1	2	−1	2	1	−2	−1	−2
2	−3	2	−3	2	−3	2	−3
1	2	−1	2	1	−2	−1	−2
−2	−3	−2	−3	−2	−3	−2	−3

Using our formulas we always get only the $1^2 + 8^2$ or $4^2 + 7^2$ resulting forms.
The same will stand for the triple square sums $4^2 + 3 \cdot 5^2 = 8^2 + 3 \cdot 3^2$.
Using more prime factors we have more and more different final forms as alternatives but they all come out by sign combinations only. This also means that to get the possible forms of an $M\,N$ product made from square sums or triple square sums, we merely have to make sign combinations of the M and N forms.
This then also means that in reverse too, knowing the product's one form and one form of one of the factors, the other will have forms that give the result this way.

For powers this means that for any form of a power, there are sign variation forms of the base, so that they give the power by our rules.
Of course this whole idea of multiplication forms applicable only if we know that the factors have forms. So this part has to be proven!
Usually we first prove that all prime factors of a simple form have forms.
The simplicity of the forms for primes is trivial because they have no factors.
Then we prove that a simple formed number multiplied by a prime with same form remains simple formed. Thus any product of such primes remains simple formed.
Then $p^2 + 3\,q^2$ being a cube r^3 of course means that r as factor of a simple triple square sum will be also simple square sum $a^2 + 3\,b^2$. Euler got to this point too.
But he didn't realize the above explained inheritance of the multiplication formula.
By that, not only is true that $r = a^2 + 3\,b^2$, but also that from any such form with mere sign variations of the a, b numbers we must get $p^2 + 3\,q^2$ by using our rules.
This then boils down to one sign choice only in the three factors and so:

$$p^2 + 3\,q^2 = (a^2 + 3\,b^2)\,(a^2 + 3\,(-b)^2)\,(a^2 + 3\,b^2)$$

must be true not just as equality of the final values but as application of our rules.

The first two multiplied by our rules gives:

$$(a^2 + 3\,b(-b))^2 + 3\,(a\,(-b) - b\,a)^2 = (a^2 - 3\,b^2)^2 + 3\,(-2\,a\,b)^2.$$

This then multiplied with the last third factor by the rules gives:

$$[\,(a^2 - 3\,b^2)\,a + 3\,(-2\,a\,b)\,b\,]^2 + 3\,[\,(a^2 - 3\,b^2)\,b - (-2\,a\,b)\,a\,]^2 =$$

$$\underbrace{a^3 - 9\,a\,b^2}_{p} \qquad\qquad \underbrace{3\,a^2 b - 3\,b^3}_{q}$$

So now we see why p must be this formed.

Euler tried to force out this form multiplication by going beyond the negatives, namely to the 3-complex numbers that is using the square root of -3.
Gauss introduced the complex name for the $a + b\sqrt{-1}$ forms.
He realized that the complex whole numbers lie behind the square sums and the above mentioned proofs for primes can be visualized and simplified with these.
This relied on the unique prime factorization of these complex wholes.
Unfortunately, this is not always true for the m-complex $a + b\sqrt{-m}$ wholes.
And as it turns out it is false for the $m = 3$ case too.
Quite amazingly, 2 the smallest prime of our natural numbers becomes a composite among the Gaussian complex wholes because:
$2 = 2 + 0\sqrt{-1} = (1 + 1\sqrt{-1})\,(1 + (-1)\sqrt{-1})$.
Even more amazingly, here among the Eulerian $a + b\sqrt{-3}$ wholes $2 = 2 + 0\sqrt{-3}$ remains a prime but $4 = 4 + 0\sqrt{-3}$ will have two different prime factorizations.
The trivial from 2 and one using new primes:

$$4 + 0\sqrt{-3} = (2 + 0\sqrt{-3})\,(2 + 0\sqrt{-3}) = (1 + 1\sqrt{-3})\,(1 + (-1)\sqrt{-3}).$$

So Euler's manipulations turned out to be unsupported by the facts.

In spite of all this, Euler's discovery of the $p = a^3 - 9 a b^2$ form is the essence of using infinite descent or indirect minimality to prove Fermat's Last Theorem for $n = 3$.
The first simple triple square sum is $a^2 + 3 b^2 = 2^2 + 3 \cdot 1^2 = 7$ and it's cube is $7^3 = p^2 + 3 q^2 = 10^2 + 3 \cdot 9^2$. So here $p = 10 = a^3 - 9 a b^2 = (-2)^3 - 9(-2) \cdot 1^2$.
Thus we need negative $a = -2$ value in the $a^2 + 3 b^2$ triple square sum form of the base, so that Euler's formula should give the cube's triple square sum form.
So we got the "bad" scenario where not the trivial $(a + 3 b) (a - 3 b) a$ factorization rather $(3 b + a) (3 b - a) (- a)$ should be used.
But this indeed gives $1 \cdot 5 \cdot 2$.
If $2 p = 20$ were a cube (which is not) then $5 , 1 , 4$ were positive cubes giving a lower product case.
Finally, I leave Fermat's Last Theorem and turn to a newer mystery, probably not as deep as that but also relating to the infinite descent.

This problem can be raised by almost starting as a parlor trick:
Think of a number! Divide it by 2 as many times its possible. When you finally get an odd, multiply it by 3 and add 1 to that. It will be obviously an even so you can divide again by 2 as many times you can. Now again multiply by 3 and add 1.
And so on, the numbers obviously decrease at halvings but increase at the steps of "times 3 plus 1". I claim that eventually you always get down to 1.
Starting from 100 we get: $100 , 50 , 25 , 76 , 38 , 19 , 58 , 29 , 88 , 44 , 22 , 11 , 34 , 17 , 52 , 26 , 13 , 40 , 20 , 10 , 5 , 16 , 8 , 4 , 2 , 1$.

By the way, if we continue from 1, we simply enter a short cycle:
$1 , 4 , 2 , 1 , 4 , 2 , 1 , \ldots .$

The claim that for any N we reach 1 is a "problem too difficult for mathematics yet" as Paul Erdős said. So he knew that we are yet "babies" but he only believed in chasing new proofs. I had a few letters exchanged with him and tried to provoke his interest in didactics without any success. He was blind as a bat to this wider truth just as to many other. To him "real mathematics" was only finding proofs for existing problems. To me proofs are not important for what they prove rather how they do that.

Here the minimality of an N where our claim is false is again almost a dead end.
The existence of a first N that doesn't return to lower value is the negative of our whole claim. This only offers two choices as return to N, that is having a cycle or never even returning to N. The next question is then whether not having cycle ever is possible or not. Keep thinking!