

UNIQUE FACTORIZATION IN \mathbf{Z} AND $F[T]$

KEITH CONRAD

1. INTRODUCTION

We call an integer $p > 1$ *prime* when its only positive factors are 1 and p . Every integer $n > 1$ has two obvious positive factors: 1 and itself. Primes are the numbers greater than 1 whose only positive factors are the obvious ones. The sequence of primes starts out as

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, \dots$$

Prime numbers are the building blocks of the positive integers under multiplication, as codified in the following theorem.

Theorem 1.1 (Unique Factorization in \mathbf{Z}). *Every integer $n > 1$ can be written as a product of primes. Moreover, the prime factorization of n is unique: if $n = p_1 \cdots p_r$ and $n = q_1 \cdots q_s$ where the p_i 's and q_j 's are prime then $r = s$ and after relabeling the factors we have $p_i = q_i$ for all i .*

Theorem 1.1 is really two statements about each $n > 1$: (i) a prime factorization of n exists and (ii) there is *only one* prime factorization for n up to the order of multiplication of the prime factors. To prove Theorem 1.1, we will prove these two statements separately.

When we talk about a product of primes in Theorem 1.1, we allow repeated factors, *e.g.*, 45 is $3 \cdot 3 \cdot 5$. Also we allow a “product” with a single term in it, so a prime number is a product of primes using only itself in the product. If we didn't allow this, then we'd have to say every $n > 1$ is a prime or a product of primes. By allowing a product with a single term, our language becomes simpler.

Polynomials can be factored, just like integers, and for historical reasons the name for the building blocks of polynomials under multiplication is irreducibles instead of primes: in $F[T]$, where F is a field, a nonconstant polynomial $p(T)$ is called *irreducible* if its only factors are nonzero constants and nonzero constant multiples of itself. Every nonconstant $f(T)$ in $F[T]$ is divisible by nonzero constants and nonzero constant multiples of $f(T)$, so the irreducible polynomials are the nonconstant polynomials whose only factors are the obvious ones. Unique factorization in $F[T]$ looks similar to unique factorization in \mathbf{Z} :

Theorem 1.2 (Unique Factorization in $F[T]$). *Let F be a field. Every nonconstant $f(T)$ in $F[T]$ can be written as a product of irreducibles. Moreover, the irreducible factorization of $f(T)$ is unique: if $f(T) = p_1(T) \cdots p_r(T)$ and $f(T) = q_1(T) \cdots q_s(T)$ where the $p_i(T)$'s and $q_j(T)$'s are irreducible then $r = s$ and after relabeling the factors we have $p_i(T) = c_i q_i(T)$ for all i , where the c_i 's are nonzero constants.*

Like Theorem 1.1, there are two statements in Theorem 1.2: for every nonconstant polynomial in $F[T]$, (i) an irreducible factorization exists and (ii) it is unique up to the order of multiplication and up to scaling by nonzero constants. Also, we adopt the convention that an irreducible polynomial is a one-term product of irreducibles.

The ambiguity in irreducible factorizations is broader than just changing the order of multiplication: we have to allow scaling of irreducible factors by nonzero constants (which doesn't change their irreducibility). To illustrate this, consider in $\mathbf{R}[T]$ the following irreducible factorizations of $T^2 - 1$:

$$T^2 - 1 = (T + 1)(T - 1) = (3T + 3) \left(\frac{1}{3}T - \frac{1}{3} \right) = \left(\frac{4}{5}T - \frac{4}{5} \right) \left(\frac{5}{4}T + \frac{5}{4} \right).$$

The second irreducible factorization scales the two factors $T + 1$ and $T - 1$ by 3 and $1/3$, while the third irreducible factorization scales these factors by $5/4$ and $4/5$ and changes the order of multiplication.

The proofs of Theorems 1.1 and 1.2 will be similar: induct on n and induct on $\deg f$.

2. PROOF OF THEOREM 1.1

Theorem 2.1. *Every $n > 1$ has a prime factorization: we can write $n = p_1 \cdots p_r$, where the p_i are prime numbers.*

Proof. We will use induction, but more precisely strong induction: assuming every integer between 1 and n has a prime factorization we will derive that n has a prime factorization.

Our base case is $n = 2$. This is a prime, so it is a product of primes by our convention that a prime is a product of primes with one term.

Now assume $n > 2$ and (here comes the strong inductive hypothesis) for all m with $1 < m < n$ that m is a product of primes. To show n is a product of primes, we take cases depending on whether m is prime or not.

Case 1: The number n is prime.

In this case, n is a product of primes with just one term. (This is the easy case.)

Case 2: The number n is not prime.

Since $n > 1$ and n is not prime, there is some nontrivial factorization $n = ab$ where $1 < a < n$ and $1 < b < n$. By our strong inductive hypothesis, both a and b are products of primes. Since n is the product of a and b , and both a and b are products of primes, n is a product of primes by stringing together the prime factorizations of a and b . More explicitly, writing $a = p_1 \cdots p_k$ and $b = p'_1 \cdots p'_\ell$ where p_i and p'_j are all prime, we have

$$n = ab = p_1 \cdots p_k p'_1 \cdots p'_\ell,$$

which is a product of primes. □

The key to proving uniqueness of prime factorization is the following property of primes.

Lemma 2.2. *If p is a prime number and $p \mid ab$ for some integers a and b , then $p \mid a$ or $p \mid b$.*

Proof. We will assume $p \mid ab$ and the conclusion is false: p does not divide a or p does not divide b . If p does not divide a then $(p, a) = 1$ because p is prime. A basic consequence of Bezout's identity tells us that from $p \mid ab$ and $(p, a) = 1$ we have $p \mid b$.

If p does not divide b , then by switching the roles of a and b (which is okay since $ab = ba$) we can conclude that $p \mid a$. □

A generalization of Lemma 2.2 is that for each finite list of integers a_1, \dots, a_k , if $p \mid a_1 \cdots a_k$ then $p \mid a_i$ for some i . This is trivial for $k = 1$, and for $k \geq 2$ it can be proved by induction on k . Lemma 2.2 is the base case $k = 2$, and if for $k \geq 3$ and the generalization is proved for a prime p dividing products of k integers, then we will prove it for p dividing

products of $k + 1$ integers: if $p \mid a_1 \cdots a_{k+1}$ then rewrite this as $p \mid (a_1 \cdots a_k)a_{k+1}$. Since $(a_1 \cdots a_k)a_{k+1}$ is a product of two integers $a_1 \cdots a_k$ and a_{k+1} , by Lemma 2.2 either $p \mid a_1 \cdots a_k$ or $p \mid a_{k+1}$. If $p \mid a_1 \cdots a_k$ then $p \mid a_i$ for some i from 1 to k by the inductive hypothesis, and including the option $p \mid a_{k+1}$ with these tells us that $p \mid a_i$ for some i from 1 to $k + 1$, which completes the inductive proof.

Now we can prove prime factorization is unique.

Theorem 2.3. *If $p_1 \cdots p_r = q_1 \cdots q_s$ where the p_i 's and q_j 's are prime, then $r = s$ and after relabeling the factors we have $p_i = q_i$ for all i .*

Proof. The key mathematical step is this: when $p_1 \cdots p_r = q_1 \cdots q_s$, p_1 must equal some q_j . This is because

$$p_1 \cdots p_r = q_1 \cdots q_s \implies p_1 \mid q_1 \cdots q_s \implies p_1 \mid q_j \text{ for some } j,$$

where the second implication is the generalization of Lemma 2.2 that we mentioned above. That uses primality of p_1 . Since q_j is prime and $p_1 \mid q_j$, we must have $p_1 = q_j$ (a prime has no factor greater than 1 other than itself).

To prove the theorem, we will induct on the total number of prime factors in the two equal prime factorizations, which is $r + s$.¹ We allow repeated primes.

The base case is $r + s = 2$ (why not $r + s = 1$?) when the equal prime factorization turns into $p_1 = q_1$. Here the conclusion of the theorem is obvious (there is no relabeling needed, since each side has one factor).

Suppose next that $r + s > 2$ and the theorem is true for all pairs of equal prime factorizations for which the total number of primes being used is less than $r + s$. If we have $p_1 \cdots p_r = q_1 \cdots q_s$ then $r > 1$ and $s > 1$: if $r = 1$ or $s = 1$ then one side is a prime number and therefore the other side has to be a prime number, so $r = s = 1$, but $r + s > 2$.

From $p_1 \cdots p_r = q_1 \cdots q_s$ we explained at the start of the proof that p_1 must be some q_j . By relabeling the factors on the right, which is okay since the order of multiplication doesn't matter, we can assume $p_1 = q_1$. Then our equal prime factorization becomes

$$p_1 p_2 \cdots p_r = p_1 q_2 \cdots q_s.$$

Canceling the common factor p_1 on both sides, we get

$$(2.1) \quad p_2 \cdots p_r = q_2 \cdots q_s.$$

In this equation of equal prime factorizations, the total number of primes appearing on both sides is $(r - 1) + (s - 1) = r + s - 2$, which is less than $r + s$. By our inductive hypothesis we conclude $r - 1 = s - 1$ (there are $r - 1$ primes on the left and $s - 1$ primes on the right), so $r = s$, and after relabeling the primes in (2.1) we have $p_i = q_i$ for all $i \geq 2$. Combining this with $p_1 = q_1$ we have $p_i = q_i$ for all i . \square

Example 2.4. For a prime p and integer $m \geq 1$, every factor of p^m is $\pm p^k$ where $0 \leq k \leq m$. Indeed, if $d \mid p^m$ then each prime factor of d is a factor of p^m , and by unique factorization the only prime factor of p^m is p . Therefore the expression of d as a product of primes up to a sign (if $d \neq \pm 1$) must be a product of p 's, so $d = \pm p^k$. Since $d \leq p^m$ we get $p^k \leq p^m$, so $0 \leq k \leq m$.

Perhaps you think the uniqueness of prime factorization is obvious, since it is consistent with all of your prior experience. Is there really a need to give a proof at all? Here are three answers to that question.

¹Another possibility is to induct on $\max(r, s)$ with very similar steps to the proof we give.

- (1) There are number systems where prime factorization exists but is *not* unique.
- (2) Even if uniqueness of prime factorization holds for small numbers, how do we know uniqueness never breaks down for big numbers without a proof? For instance, $n = 11501689$ can be written as $2747 \cdot 4187$ and $3239 \cdot 3551$. Does that violate uniqueness of prime factorization? No, because those factorizations don't use primes: $2747 = 41 \cdot 67$ and $4187 = 53 \cdot 79$, while $3239 = 41 \cdot 79$ and $3551 = 53 \cdot 67$. The prime factorization of n is really $41 \cdot 53 \cdot 67 \cdot 79$.
- (3) Historically, nobody saw a need to prove Theorem 1.1 until Gauss in 1801 (Article 16 in his book *Disquisitiones Arithmeticae*). Earlier, Legendre and Euler stated the existence of prime factorization in their work, but gave no proof of its uniqueness even if they implicitly used it. Going back further to the ancient Greeks, in Euclid's *Elements* there is proof that each $n > 1$ has a prime factor (Prop. 31 in Book VII), but no proof that n has a prime factorization or that it is unique. There is a proof of Example 2.4 when $d > 0$ (Prop. 13 in Book IX).

We turn next to unique factorization of polynomials, and we will see that almost everything we have done for integers will carry over to the polynomial setting except that we have to keep track of scalar multiples of irreducible factors.

3. PROOF OF THEOREM 1.2

To prove $F[T]$ has unique factorization, we will first prove irreducible factorizations exist for all nonconstant polynomials and then we will show they are unique (up to the order of multiplication and scaling by nonzero constants).² This is similar to the strategy in \mathbf{Z} .

Theorem 3.1. *Every nonconstant polynomial in $F[T]$ has an irreducible factorization.*

Proof. We will argue by strong induction on the degree of polynomials.

Our base case is degree 1. Every polynomial in $F[T]$ of degree 1 is irreducible, so they are each a product of irreducibles using just one term.

Now assume $d > 1$ and *every* nonconstant polynomial in $F[T]$ with degree less than d has an irreducible factorization. Pick a polynomial $f(T) \in F[T]$ of degree d . We want to show $f(T)$ has an irreducible factorization.

Case 1: The polynomial $f(T)$ is irreducible.

Here $f(T)$ is a one-term product of irreducibles.

Case 2: The polynomial $f(T)$ is not irreducible.

Since $f(T)$ is nonconstant and not irreducible, we can write $f(T) = g(T)h(T)$ where $0 < \deg g < \deg f$ and $0 < \deg h < \deg f$. By the strong inductive hypothesis, $g(T)$ and $h(T)$ have irreducible factorizations: $g(T) = p_1(T) \cdots p_k(T)$ and $h(T) = \tilde{p}_1(T) \cdots \tilde{p}_\ell(T)$ with irreducible $p_i(T)$ and $\tilde{p}_j(T)$ in $F[T]$. Putting these together gives us an irreducible factorization of $f(T)$:

$$f(T) = g(T)h(T) = p_1(T) \cdots p_k(T) \tilde{p}_1(T) \cdots \tilde{p}_\ell(T),$$

which expresses $f(T)$ as a product of irreducibles in $F[T]$. □

To prove irreducible factorization in $F[T]$ is unique, we use an analogue of Lemma 2.2.

Lemma 3.2. *If $p(T)$ is irreducible in $F[T]$ and $p(T) \mid a(T)b(T)$ in $F[T]$, then $p(T) \mid a(T)$ or $p(T) \mid b(T)$.*

²The page <https://mathoverflow.net/questions/15137> discusses the history of unique factorization in polynomials.

Proof. This is proved just like Lemma 2.2. If $p(T)$ does not divide $a(T)$ then $(p(T), a(T)) = 1$ in $F[T]$ because $p(T)$ is irreducible. From Bezout's identity in $F[T]$, the conditions $p(T) \mid a(T)b(T)$ and $(p(T), a(T)) = 1$ imply $p(T) \mid b(T)$. Similarly, if $p(T)$ does not divide $b(T)$ then $p(T) \mid a(T)$ by swapping the roles of $a(T)$ and $b(T)$. \square

Lemma 3.2 generalizes by induction on the number of terms to say that if $p(T)$ is irreducible in $F[T]$ and $p(T) \mid a_1(T) \cdots a_k(T)$ then $p(T) \mid a_i(T)$ for some i . The proof is just like its analogue in \mathbf{Z} that we discussed earlier. With this generalization we can prove the uniqueness of irreducible factorizations in $F[T]$.

Theorem 3.3. *If $p_1(T) \cdots p_r(T) = q_1(T) \cdots q_s(T)$ where the $p_i(T)$'s and $q_j(T)$'s are irreducible, then $r = s$ and after relabeling the factors we have $p_i(T) = c_i q_i(T)$ for all i where the c_i are nonzero in F .*

Proof. As in the proof Theorem 2.3, the key step is that if $p_1(T) \cdots p_r(T) = q_1(T) \cdots q_s(T)$ then $p_1(T) = c_j q_j(T)$ for some j and nonzero c_j in F . This follows from

$$p_1(T) \cdots p_r(T) = q_1(T) \cdots q_s(T) \implies p_1(T) \mid q_1(T) \cdots q_s(T) \implies p_1(T) \mid q_j(T) \text{ for some } j.$$

By relabeling, we can take $j = 1$, i.e., $p_1(T) \mid q_1(T)$. That implies $p_1(T)$ is a constant multiple of $q_1(T)$ since the only nonconstant factors of an irreducible polynomial in $F[T]$ are constant multiples of itself. (This is different from prime numbers, where the only positive factor is the number itself!) Therefore $p_1(T) = c_1 q_1(T)$ for some constant c_1 .

Our theorem will be proved by induction on the total number of irreducible factors in the equal irreducible factorizations, which is $r + s$. The base case is $r + s = 2$, when the equation is $p_1(T) = q_1(T)$, and this case is obvious.

Now suppose $r + s > 2$ and the theorem is true for all pairs of equal irreducible factorizations for which the total number of irreducibles is less than $r + s$. If $p_1(T) \cdots p_r(T) = q_1(T) \cdots q_s(T)$ then $r > 1$ and $s > 1$ by the same proof as in the integer case. By relabeling the factors we have assume $p_1(T) = c_1 q_1(T)$, so

$$p_1(T)p_2(T) \cdots p_r(T) = \frac{1}{c_1} p_1(T)q_2(T) \cdots q_s(T).$$

Canceling $p_1(T)$ from both sides,

$$(3.1) \quad p_2(T) \cdots p_r(T) = \frac{1}{c_1} q_2(T) \cdots q_s(T).$$

We need to be careful here: the factor $1/c_1$ on the right is *not* irreducible. It's just a constant. We can attach it to $q_2(T)$: the polynomial $(1/c_1)q_2(T)$ is irreducible. Therefore the left side of (3.1) has $r - 1$ irreducible factors and the right side has $s - 1$ irreducible factors (namely $(1/c_1)q_2(T), q_3(T), \dots, q_s(T)$). From $(r - 1) + (s - 1) = r + s - 2 < r + s$, by the inductive hypothesis $r - 1 = s - 1$, so $r = s$. Also by the inductive hypothesis, after relabeling the factors we have $p_2(T) = c(1/c_1)q_2(T)$ and $p_i(T) = c_i q_i(T)$ for all $i \geq 3$, where c, c_3, \dots are all nonzero constants. Set $c_2 = c/c_1$, and then along with $p_1(T) = c_1 q_1(T)$ we have shown each $p_i(T)$ is a constant multiple of $q_i(T)$ and our proof is complete. \square

Other than some extra bookkeeping to account for constant multiples, the proof of unique factorization in $F[T]$ is basically the same as the proof of unique factorization in \mathbf{Z} .

Example 3.4. If $p(T)$ is irreducible in $F[T]$ and $m \geq 1$ then every factor of $p(T)^m$ in $F[T]$ is $cp(T)^k$ where c is nonzero in F and $0 \leq k \leq m$. The proof is along the same lines as in Example 2.4. Details are left to the reader.