UNIQUE FACTORIZATION IN 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

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

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

Theorem 1.1 (Unique Factorization in **Z**)**.** *Every integer $n > 1$ can be written as a product of primes. Moreover, the prime factorization of n is unique up to the order of multiplication: 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 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.

Like integers, polynomials can be factored, but for historical reasons we call the building blocks irreducible instead of prime. 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 ones whose only factors are the obvious ones. Unique factorization in $F[T]$ has a statement very similar to unique factorization in **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 up to the order of multiplication and constant scaling factors: 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 in $F[T]$ 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 in F.*

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 scaling by nonzero constants. Also, we adopt the convention that an irreducible is considered to be a product of irreducibles using just one term.

To illustrate why the uniqueness of irreducible factorizations in $F[T]$ has to allow for scaling of irreducible factors by nonzero constants (which doesn't change their irreducibility), consider the following different irreducible factorizations of $T^2 - 1$ in $\mathbf{R}[T]$:

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

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. We want to consider all of these to be essentially the same irreducible factorization.

We will prove Theorems 1.1 and 1.2 in similar ways: induction on n and induction 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 $a = p_1 \cdots p_r$ and $b = q_1 \cdots q_s$ where p_i and q_j are all prime, we have

$$n = ab = p_1 \cdots p_r q_1 \cdots q_s,$$

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 $p|ab$ for some integers a and b, then $p|a$ or $p|b$.*

Proof. We are assuming $p|ab$ and want to show p divides a or p divides b. If p did not divide a then $(p, a) = 1$ because p is prime. A basic consequence of Bezout's identity tells us that from $p|ab$ and $(p, a) = 1$ we have $p|b$.

Similarly, if p did not divide b then by switching the roles of a and b (which is okay since $ab = ba$) we conclude that $p|a$.

We showed if either a or b were not divisible by p then the other factor would be divisible by p, so at least one of a or b has to be divisible by p. □

A generalization of Lemma 2.2 is that for any finite list of integers $a_1, \ldots, a_k$, if $p|a_1 \cdots a_k$ then $p|a_i$ for some i. This is trivial for $k = 1$, and for $k \geq 2$ it is true by induction on k with Lemma 2.2 being the base case $k = 2$. Details of the inductive step are left to the reader.

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

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

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 $p_1|q_j$, we must have $p_1 = q_j$ (a prime has no factor greater than 1 other than itself).

To prove our 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 any two 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

$$p_2 \cdots p_r = q_2 \cdots q_s. \tag{2.1}$$

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. □

Perhaps you think the uniqueness of prime factorization is obvious, since it is consistent with all of your prior experience. Is there really any need to give a proof at all? There are (at least) two answers to this question.

(1) There are number systems where prime factorization exists but is *not* unique.
(2) You have no sustained experience with very large numbers, and on strictly logical grounds why couldn't unique factorization break down for big integers if we don't have a proof that it doesn't happen? For instance, $n = 11501689$ can be written as $2747 \cdot 4187$ and $3239 \cdot 3551$. Is this a violation of uniqueness of prime factorization? No, because those two factorizations are not into 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

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

really $41 \cdot 53 \cdot 67 \cdot 79$. We had taken a number with four prime factors and paired off the prime factors in two different ways, with the resulting factors being not obviously composite. That you have never seen an actual counterexample to unique factorization in $\mathbf{Z}$ may be your reason to believe that there is no example, but it doesn't *prove* there is no example. The proof of Theorem 2.3 settles the matter unequivocally.

We turn next to unique factorization of polynomials, and happily almost everything we have done for integers will carry over to the polynomial setting.

3. Proof of Theorem 1.2

As in the integer case, we will first prove irreducible factorizations exist and then we will show they are unique (up to the order of multiplication and scaling by nonzero constants).

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 irreducible prime.

There is a factorization $f(T) = g(T)h(T)$ where $0 < \deg g < \deg f$ and $0 < \deg h < \deg f$. By the strong inductive hypothesis, $g(T)$ and $h(T)$ each have irreducible factorizations, and putting these irreducible factorizations together gives us an irreducible factorization of $f(T)$. □

To prove the irreducible factorization in $F[T]$ is unique, we need the following analogue of Lemma 2.2.

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

Proof. This is proved just like Lemma 2.2. If $p(T)$ did 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 $p(T)|a(T)b(T)$ and $(p(T), a(T)) = 1$ imply $p(T)|b(T)$. Similarly, if $p(T)$ did not divide $b(T)$ then $p(T)|a(T)$ by swapping the roles of $a(T)$ and $b(T)$. □

Lemma 3.2 generalizes by induction on the number of terms to say that if $p(T)$ is irreducible in $F[T]$ and $p(T)|a_1(T) \cdots a_k(T)$ then $p(T)|a_i(T)$ for some i. We leave the details to you to work out. 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 and for some nonzero $c_i \in F$.*

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 some $c_j \in F - \{0\}$. This follows from

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

By relabeling, we can take $j = 1$, *i.e.*, $p_1(T)|q_1(T)$. 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 it. (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 any two 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,

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

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. Since $(r - 1) + (s - 1) = r + s - 2 < r + s$, from the inductive hypothesis we get $r - 1 = s - 1$, so $r = s$. Also from the inductive hypothesis, by 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 $c, c_3, \ldots$ are all nonzero constants in F. Set $c_2 = c/c_1$, and then along with the equation $p_1(T) = c_1 q_1(T)$ we have shown each $p_i(T)$ is a constant multiple of $q_i(T)$ and our proof is complete. □

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