Unique Factorization

What Not Everyone Knows

Paul Pollack

University of Georgia

September 26, 2023

We start with the layperson's definition of **prime**: a number > 1 not a product of two smaller positive integers.

Fundamental Theorem of Arithmetic

Every positive integer can be written as a product of primes. This expression is unique up to the order of the factors.

The standard narrative: what "everyone" knows

When it's proved in courses, the Fundamental Theorem is usually established by something isomorphic to the following chain of reasoning.

First, one shows existence of prime factorizations, by descent. The smallest number with no prime factorization couldn't be prime, so must factor. But then those factors are smaller and so they themselves have prime factorizations. . .

Uniqueness is trickier.

Suppose we have two different factorizations of the same positive integer n, say

$$n = p_1 \cdots p_k = q_1 \ldots q_\ell.$$

To proceed, we call upon the wisdom of the ancients.

Euclid's Lemma

If p is prime and $p \mid ab$, then $p \mid a$ or $p \mid b$.

Continuing from above, we can assume $n > 1$, so $k, \ell > 0$. Then $p_1 \mid q_1 \ldots q_\ell$, so $p_1 \mid q_j$ for some j. Relabeling, we can assume $j = 1$. But $p_1 \mid q_1$ implies $p_1 = q_1$. Canceling,

$$n/p_1 = p_2 \cdots p_k = q_2 \cdots q_\ell.$$

Then n/p_1 is a smaller counterexample, etc...

But where does Euclid's lemma come from?
Chain of reasoning:

1. $\mathbb{Z}$ has a division algorithm, so is a Euclidean domain.
2. Every ideal in $\mathbb{Z}$ is principal; in fact, it consists precisely of all $\mathbb{Z}$-multiples of its smallest nonnegative element.
3. For any positive integers a, b, the set $\{ax + by : a, b \in \mathbb{Z}\}$ is an ideal. Then its least nonnegative element d is the greatest common divisor of a and b. In particular, the greatest common divisor of a, b is a linear combination of a, b.

We can now prove Euclid's lemma.
Suppose $p \mid ab$, where p is prime and a and b are positive integers. If $p \mid a$, we are done. Otherwise, the only common divisors of p and a are ± 1. It follows that

$$1 = ax + py \quad \text{for some } x, y \in \mathbb{Z}.$$

Multiplying through by b gives

$$b = abx + pby.$$

Since $p \mid ab$ and $p \mid pb$, we deduce that p divides the RHS, and so p also divides the LHS, so p divides b. Done!

Think deeply of simple things.

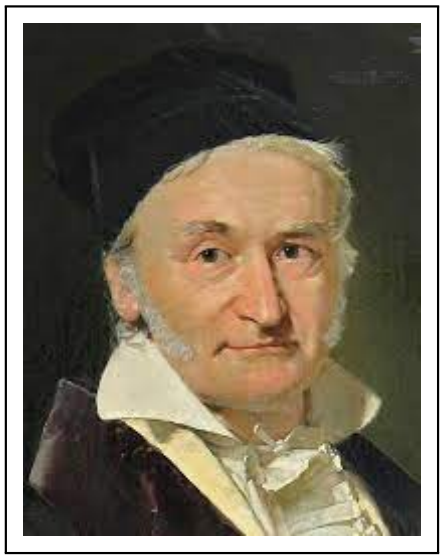
Arnold Ross

Answer the questions, then question the answers.

Glenn Stevens

A few of my favorite proofy things

Over the summer, I tracked down about 30 papers either offering new proofs of the Fundamental Theorem or collecting existing arguments.



Perhaps surprisingly, the first explicit statement of the unique factorization theorem is due to Gauss in the *Disquisitiones*. Gauss deduces the theorem from Euclid's lemma, for which he gives a beautiful, self-contained proof. Dressed up a bit using the language of groups, Gauss's proof goes as follows.

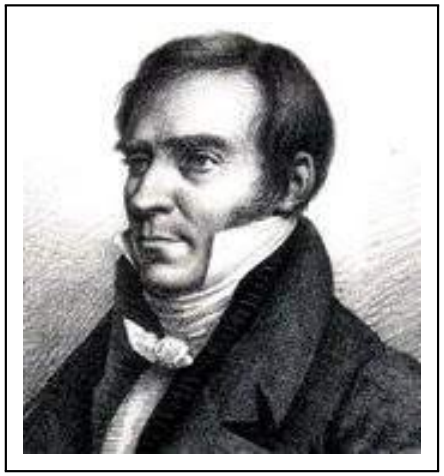
Euclid's Lemma

If p is prime and $p \mid ab$, then $p \mid a$ or $p \mid b$.

Let G be a group, written additively. If $x \in G$ and n is a nonzero integer with $nx = 0$, then the order of x is a divisor of n.

Gauss's proof, revisited.

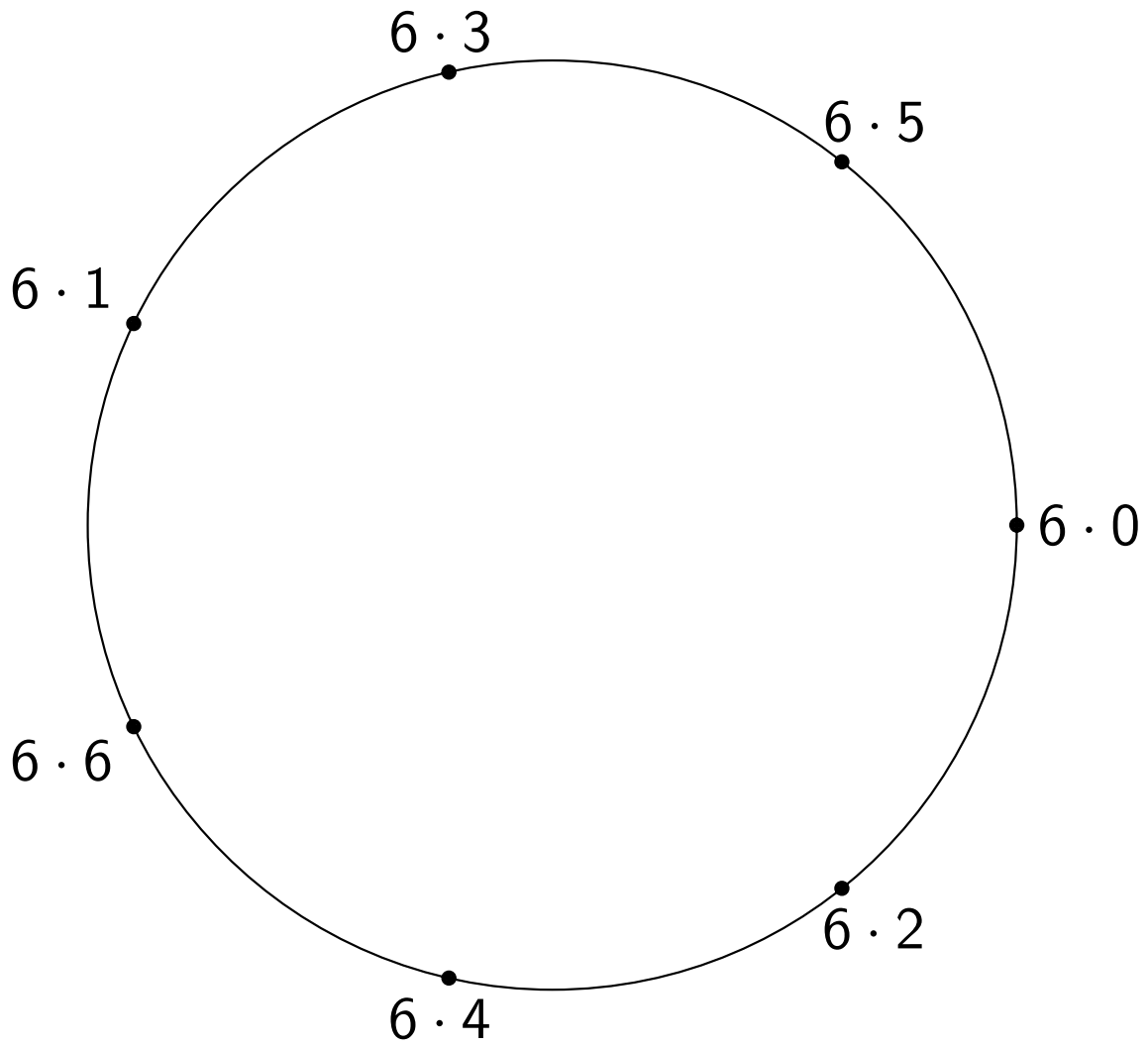
Look at the order of a mod p, as an element of the group $(\mathbb{Z}/p\mathbb{Z}, +)$. Clearly, $pa = 0$ in this group, and so the order of a mod p is a divisor of p. Hence, this order is 1 or p.
If the order of a mod p is 1, then a mod p is the identity of $\mathbb{Z}/p\mathbb{Z}$, so a mod $p = 0$ mod p and $p \mid a$.
Now suppose the order is p. Since $p \mid ab$, we have in $\mathbb{Z}/p\mathbb{Z}$ that $ba = 0$. Thus, the order of a divides b. That is, $p \mid b$.



There is a beautiful geometric approach of Louis Poinsot that I learned about from an old text of Bachmann.

Let n be a positive integer. We envision $\mathbb{Z}/n\mathbb{Z}$ as $\mathbb{Z}$ wrapped around a circle of circumference n.

Let a be a positive integer with $a < n$ (inessential restriction). Mark the points ma, with $m \in \mathbb{Z}$, on the circumference of the circle. It's enough to mark $m = 0, 1, 2, \ldots, n-1$, and sometimes even fewer.



Example:
$n = 14$,
$a = 6$

Observe:

- Diagram is homogeneous: you have the same view from any vertex.
- As a consequence, the distance along the circle between neighboring vertices is constant; let's call this d.
- The distance along the circle from $a \cdot 0$ to $a \cdot 1$ has to be a multiple of d. That is, d divides a.
- The entire circumference is a multiple of d also, so d divides n.
- Look at the nearest point to $a \cdot 0$, counterclockwise (say). This is d units away and has the form $a \cdot x$ for some $x \in \mathbb{Z}$. Hence,

$$ax \equiv d \pmod{n}.$$

  Therefore: $ax + ny = d$ for some $x, y \in \mathbb{Z}$. It follows that d is divisible by every common divisor and so is a gcd of a, b.

From here, our proof sketch of UFT can be completed as before.

Here's looking at Euclid

Euclid never stated unique factorization, but he got close.

In Book IX, Proposition 14, Euclid writes

> *If a number be the least that is measured [divisible] by prime numbers, it will not be measured by any other prime number except those originally measuring it.*

The Proposition is deduced from Euclid's lemma.

Euclid's lemma, in turn, is deduced from the following proposition.

Proposition (Common reduction lemma)

Suppose $a/b = c/d$ is an equation of rational numbers, where a, b, c, d are positive integers. Then a/b and c/d have a common reduction s/u.

Euclid's lemma, in turn, is deduced from the following proposition.

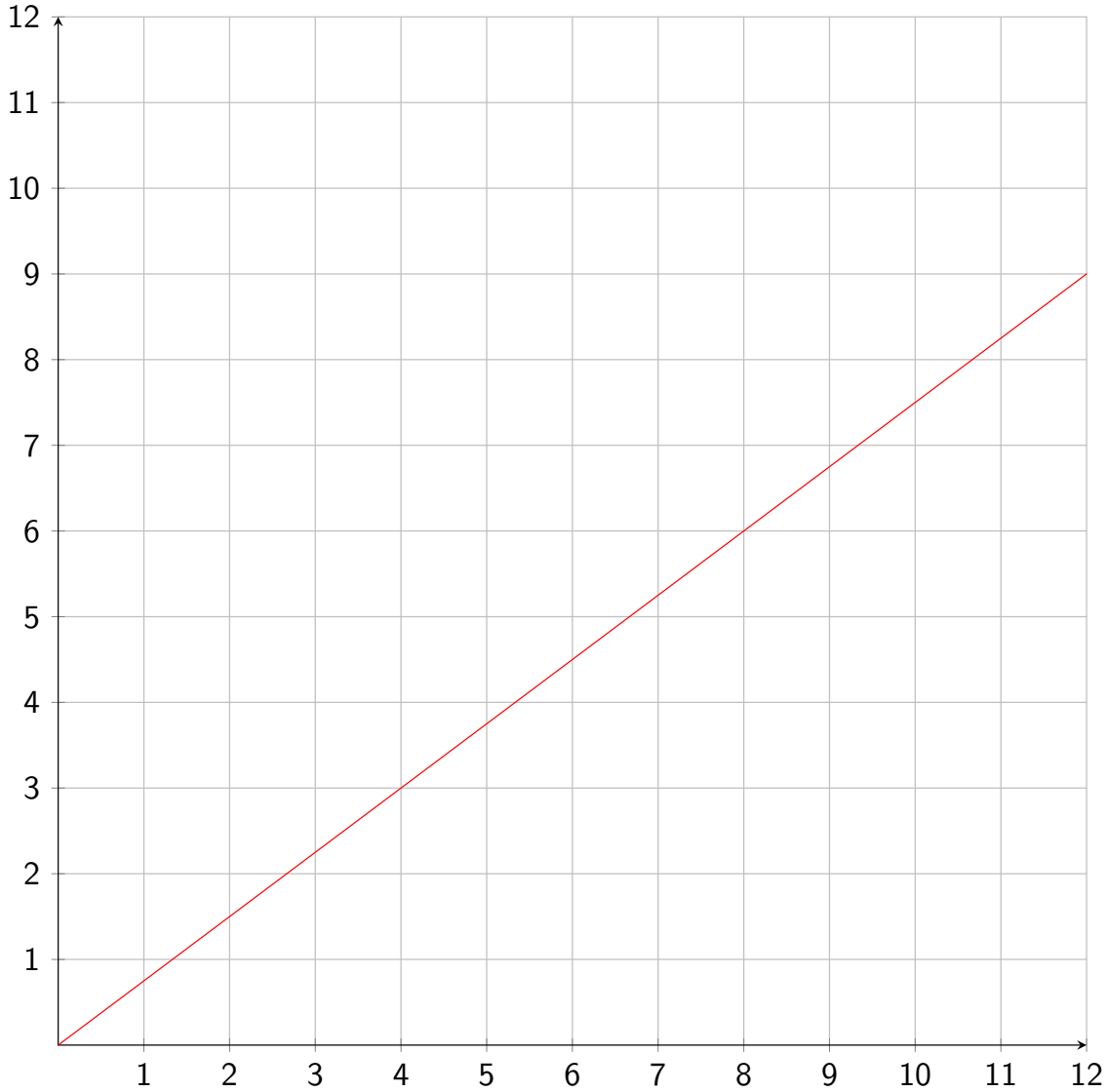
Proposition (Common reduction lemma)

Suppose $a/b = c/d$ is an equation of rational numbers, where a, b, c, d are positive integers. Then a/b and c/d have a common reduction s/u.

Somewhat remarkably, there is widespread (but not universal) consensus that Euclid's proof of the common reduction lemma doesn't quite work. Barry Mazur writes (diplomatically):

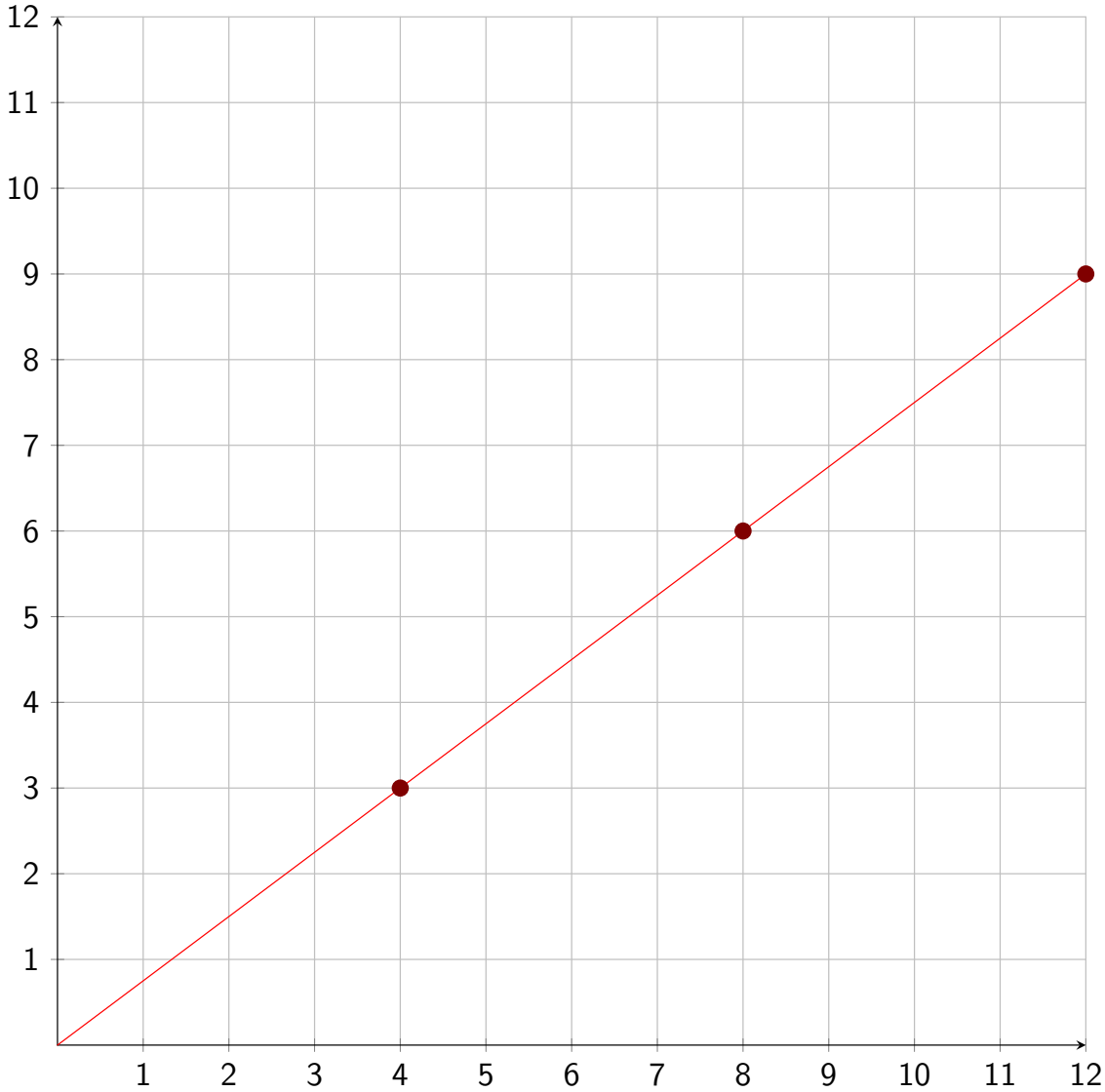
> *Now I don't quite follow Euclid's proof of this pivotal proposition, and I worry that there may be a tinge of circularity in the brief argument given in the text.*

For a detailed analysis, see the 2006 *Monthly* article 'Did Euclid Need the Euclidean Algorithm to Prove Unique Factorization?' by David Pengelley and Fred Richman.



Example:

The line $x = \alpha y$ for $\alpha = \frac{12}{9} = \frac{8}{6}$.



Example:

The line $x = \alpha y$ for $\alpha = \frac{12}{9} = \frac{8}{6}$.

A common reduction is 4/3.

[In mathematics,] a picture is worth a thousand words, but it takes another thousand words to justify the picture.
Harold Stark

Unique factorization's greatest hits and misses

Stating unique factorization explicitly, as Gauss did, makes it natural to consider when unique factorization fails.

Let D be a domain and let π be a nonzero, nonunit element of D. We say π is **irreducible** if π cannot be written as a product of nonunits.

A domain D is a **unique factorization domain (UFD)** if every nonzero nonunit is a product of irreducibles and this expression is unique up to order and up to unit factors.

In a first course on algebraic number theory, one studies "rings of integers of number fields". A **number field** is a finite extension $K/\mathbb{Q}$. For each number field K, we set

$$\mathbb{Z}_K := \{\alpha \in K : f(\alpha) = 0 \text{ for some monic polynomial } f(x) \in \mathbb{Z}[x]\}.$$

While not obvious at first glance, $\mathbb{Z}_K$ is a ring, and Dedekind realized that it is somehow the "correct" analogue of the integers inside the number field K.

Ex: $\mathbb{Z}_{\mathbb{Q}(\sqrt[3]{2})} = \mathbb{Z}[\sqrt[3]{2}]$, $\mathbb{Z}_{\mathbb{Q}(i)} = \mathbb{Z}[i]$, $\mathbb{Z}_{\mathbb{Q}(\sqrt{5})} = \mathbb{Z}[(1+\sqrt{5})/2]$.

The rings $\mathbb{Z}_K$ are not always UFDs. Dedekind showed though that they always enjoy unique factorization into ideals.

How could we quantify the failure of elementwise unique factorization?

Well, we don't have unique factorization into elements of $\mathbb{Z}_K$, but we do have unique factorization into ideals. So we could introduce a gadget that measures how far away ideals are from being elements. This motivates something called the **class group**: this is defined as a kind of quotient* of the monoid of nonzero ideals of $\mathbb{Z}_K$ by the monoid of nonzero principal ideals. This is always be a finite abelian group (a big deal!), and it is trivial precisely when $\mathbb{Z}_K$ is a UFD.

If you ask most number theorists how to measure the failure of non-uniqueness, they would tell you to read about the class group.

There's a lowbrow way we might try to quantify nonunique factorization. Our starting point a well-known example of nonuniqueness: In $\mathbb{Z}[\sqrt{-5}]$,

$$(1+\sqrt{-5})(1-\sqrt{-5}) = 2 \cdot 3.$$

These are genuinely different factorizations. . . but they have the same length.

It's tempting to not make much of this one example, but it turns out that in $\mathbb{Z}[\sqrt{-5}]$, any two irreducible factorizations of the same element share the same length.

Let's say factorization is **half-unique** in a domain D (or D is an **HFD**) if every nonzero nonunit element of D factors as a product of irreducibles, and any two factorizations of the same element share the same number of irreducibles.

Theorem (Carlitz, 1960)

Let K be a number field. If $\#\mathrm{Cl}(\mathbb{Z}_K) = 1$ or 2, then factorization is half-unique in $\mathbb{Z}_K$. The converse also holds.

Since $\mathbb{Q}(\sqrt{-5})$ has class number 2, its ring of integers $\mathbb{Z}[\sqrt{-5}]$ is an HFD, as claimed.

On the other hand, $\mathbb{Q}(\sqrt{-23})$ has class number 3. Here $3 \cdot 3 \cdot 3 = (2 + \sqrt{-23})(2 - \sqrt{-23})$, showing half-uniqueness fails.

Stretching the truth about unique factorization

Let D be a domain where every nonzero nonunit factors into irreducibles. (This is true for all the $\mathbb{Z}_K$.) For each nonzero nonunit $\alpha \in \mathbb{Z}_K$, we define the **length spectrum** of α by

$$\mathcal{L}(\alpha) = \{\text{all lengths } k \text{ of irreducible factorizations } \alpha = \pi_1 \cdots \pi_k\}.$$

We define the **elasticity** of α by

$$\rho(\alpha) = \frac{\sup \mathcal{L}(\alpha)}{\inf \mathcal{L}(\alpha)}.$$

Finally, we define the elasticity $\rho(D)$ of D by

$$\rho(D) = \sup_{\alpha} \rho(\alpha).$$

So $\rho(D) = 1$ if and only if D is an HFD.

Let K be a number field.

Theorem (Valenza, Narkiewicz, Steffan)

Assume $\mathbb{Z}_K$ is not a UFD. Then

$$\rho(\mathbb{Z}_K) = \frac{1}{2} \cdot Davenport\ constant\ of\ \mathrm{Cl}(\mathbb{Z}_K).$$

OK, but what is the Davenport constant? For a finite abelian group G, the Davenport constant $D(G)$ is the smallest D for which every length D sequence

$$g_1, g_2, \ldots, g_D$$

of elements of G contains a nonempty subsequence whose product is the identity. Not hard: $D(G) \leq \#G$ and $D(G) \to \infty$ as $\#G \to \infty$.

The proof that $\rho(\mathbb{Z}_K) \leq \frac{1}{2} D(\mathrm{Cl}(\mathbb{Z}_K))$ is not so bad, once one knows to try to prove it! The key idea is that $D(\mathrm{Cl}(\mathbb{Z}_K))$ is the largest number of prime ideals that can contain (equivalently, divide) an irreducible element of $\mathbb{Z}_K$.

The inequality $\rho(\mathbb{Z}_K) \geq \frac{1}{2} D(\mathrm{Cl}(\mathbb{Z}_K))$ lies deeper; this depends on analytic results of class field theory.

One blemish of the above result is that we don't understand how to compute Davenport constants. The answer is known for finite abelian groups of rank ≤ 2, but the general rank 3 case is still open.

Let's make all of this even more concrete.

We'll zero in on quadratic fields: extensions $K/\mathbb{Q}$ with $[K : \mathbb{Q}] = 2$. The family of quadratic fields has the virtue of being easily describable by a single integer parameter: each such $K = \mathbb{Q}(\sqrt{d})$ for a unique squarefree integer d.

In this family of K (equivalently, family of d), how good (or how bad) does factorization look in $\mathbb{Z}_K$?

It is an open problem, going back essentially to Gauss, to decide whether $\mathbb{Z}_K$ has unique factorization (equivalently, whether $\mathrm{Cl}(\mathbb{Z}_K)$ is trivial) for infinitely many quadratic fields K.

For imaginary quadratic fields, meaning $K = \mathbb{Q}(\sqrt{d})$ with $d < 0$, we know that unique factorization holds only finitely often. The largest (in absolute value) is $d = -163$ (Baker–Heegner–Stark). Moreover, from work of Heilbronn, $\#\mathrm{Cl}(\mathbb{Q}(\sqrt{d})) \to \infty$ as $d \to -\infty$. So factorization gets "worse and worse".

For $d > 0$, the situation is expected to be rather different. We expect that $\mathrm{Cl}(\mathbb{Q}(\sqrt{d})) = 1$ infinitely often. In fact, heuristics of Cohen–Lenstra predict that the class of $\mathbb{Q}(\sqrt{p})$ should be 1 for about 75.4% of all primes p.

There's been remarkable progress towards the Cohen–Lenstra heuristics in recent years. But existing methods do not establish even that

$$\#\mathrm{Cl}(\mathbb{Q}(\sqrt{d})) < 10^{10^{10}}$$

for infinitely many squarefree d !

So it seems that if we want to find infinitely many UFDs, we have to look beyond quadratic fields.

Question (Coykendall): What about HFDs? Can we find infinitely many half-factorial domains by wandering in the land of quadratic fields?

It's tempting to answer no. And if we restrict attention to the rings $\mathbb{Z}_K$, we run into the same problems as before. For $\mathbb{Z}_K$ to be half-factorial, one needs (Carlitz) that $h_K \leq 2$. This inequality happens for only finitely many imaginary quadratic fields K. And for all we can prove, it happens for only finitely many real K too.

But ... $\mathbb{Z}_K$ is not the only game in town. We can look at subrings of $\mathbb{Z}_K$.

Let K be a quadratic field. An **order** in K is a subring of $\mathbb{Z}_K$ properly containing $\mathbb{Z}$. The ring $\mathbb{Z}_K$ itself is referred to as the **maximal order**.

For example, $\mathbb{Z}[2023i]$ is an order in $\mathbb{Q}(i)$, the maximal order being $\mathbb{Z}[i]$. And $\mathbb{Z}[\sqrt{5}]$ is an order in $\mathbb{Q}(\sqrt{5})$ (there the maximal order is $\mathbb{Z}[(1+\sqrt{5})/2]$).

It's elementary to show that nonmaximal orders can never be UFDs. (UFDs are integrally closed.) But they can still be HFDs!

Conjecture (Coykendall)

(a) *There are infinitely many HFDs as you vary over all quadratic fields and all orders contained in those fields.*

(b) *There are infinitely many HFDs as you vary among the orders in the quadratic field* $\mathbb{Q}(\sqrt{2})$.

Theorem (P., 2023)

(a) *is true, and* (b) *is true assuming GRH.*

Perhaps surprisingly, the method of proof builds on techniques used to study a seemingly unrelated question.

Observe that $1/7 = 0.\overline{142857}$ has period length $7 - 1$, and $1/19 = 0.\overline{052631578947368421}$ has period length $19 - 1 = 18$. On the other hand, $1/13 = 0.\overline{076923}$, so the period length is $\frac{13-1}{2}$.

It's not so hard to prove that for a prime $p \neq 2, 5$, the period length is always a divisor of $p - 1$ — in fact, it's nothing other than the order of 10 in the multiplicative group $(\mathbb{Z}/p\mathbb{Z})^\times$. Is it equal to $p - 1$ infinitely often?

Artin conjectured YES. He gave a heuristic argument that this should happen $\approx 37.4\%$ of the time.

Artin's conjecture is still open. However, in 1967 Hooley proved that Artin's conjecture follows from the Generalized Riemann Hypothesis.

In 1984, Gupta and Murty found a clever proof that there is some base $b > 2$ for which infinitely many primes p have repeating period $p - 1$. In fact, their method produces many such bases; Heath-Brown has shown that one of the bases $2, 3$, or 5 works (though the method does not allow one to decide which).

The theorem about HFDs is proved by similar methods but in the quadratic field setting. Here the connection is provided by the relative class number formula, and the role of the base is played by the "fundamental unit" of the quadratic field.

To make this all explicit. It turns out that proving conjecture (b) is equivalent to showing there are infinitely many primes p inert in $\mathbb{Z}[\sqrt{2}]$ for which $\ell = p + 1$ is the least positive integer ℓ with

$$(1 + \sqrt{2})^{\ell} \equiv (\text{some rational integer}) \pmod{p}.$$

Thank You!