$\mathbb{Z}[\sqrt{-5}]$: halfway to unique factorization

Paul Pollack

Abstract. It is well known that factorization is not unique in $\mathbb{Z}[\sqrt{-5}]$. We give a short, self-contained proof that $\mathbb{Z}[\sqrt{-5}]$ is "halfway" towards being a unique factorization domain: For every nonzero, nonunit $\alpha \in \mathbb{Z}[\sqrt{-5}]$, any two factorizations of α into irreducibles involve the same number of factors.

1. INTRODUCTION. Our jumping-off point is the familiar definition of a **unique factorization domain** (UFD): An integral domain D is a UFD if every nonzero nonunit element of D can be expressed as a product of irreducible elements of D in a unique way, where uniqueness is up to order and unit multiplication. This last clause ("up to order and…") is a bit slippery, and the precise conditions for uniqueness are most clearly expressed in two parts. Whenever $\pi_1, \dots, \pi_k$ and $\rho_1, \dots, \rho_\ell$ are irreducibles having $\pi_1 \cdots \pi_k = \rho_1 \cdots \rho_\ell$, uniqueness requires that

(i) $k = \ell$, *and*

(ii) for some permutation σ of $\{1, 2, \dots, k\}$, and some units $\epsilon_1, \dots, \epsilon_k$ of D,

$$\rho_{\sigma(i)} = \epsilon_i \pi_i \quad \text{for all } i = 1, 2, \dots, k.$$

Unique factorization domains are strewn throughout the landscape of a first ring theory course, customary examples being the ring $\mathbb{Z}$ of ordinary integers, the ring $F[x]$ of one-variable polynomials over a field F, and the ring $\mathbb{Z}[i]$ of Gaussian integers.

Lest one form the impression that all reasonable domains are UFDs, it is common for instructors in these courses to trot out $\mathbb{Z}[\sqrt{-5}] = \{a + b\sqrt{-5} : a, b \in \mathbb{Z}\}$ as a cautionary tale. Making use of the norm map (whose definition is recalled below), it is simple to prove that every nonzero nonunit in $\mathbb{Z}[\sqrt{-5}]$ has some expression as a product of irreducibles. But this expression is not always unique! Arguing again with norms, one can show that ± 1 are the only units in $\mathbb{Z}[\sqrt{-5}]$ (see §2, below) and that all of $2, 3, 1 + \sqrt{-5}$, and $1 - \sqrt{-5}$ are irreducible. It follows that the innocent-seeming and easily-noticed identity

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

exhibits two distinct factorizations of the element 6.

Both sides of (1) involve the same number of irreducibles ($k = \ell = 2$), meaning that condition (i) for uniqueness holds. Thus (1) is a counterexample to uniqueness only on account of condition (ii) and so it might be objected, somewhat cheekily, that (1) represents only a *half*-failure of unique factorization. Following Zaks [1, 2], we say that a domain D is **half-factorial** (or an **HFD**) if all counterexamples in D to unique-factorization satisfy condition (i) above. Precisely, D is an HFD if any two factorizations into irreducibles of the same nonzero nonunit element feature the same number of factors.

It is remarkable — and surely deserves to be better known — that half-failures of unique factorization are all one can hope for (fear for?) in $\mathbb{Z}[\sqrt{-5}]$.

Theorem 1. The domain $\mathbb{Z}[\sqrt{-5}]$ is an HFD.

Theorem 1 is a very special case of a 1960 result of Carlitz [3] which characterizes number fields whose rings of integers are half-factorial. Carlitz proves that these are precisely the number fields of class number less than or equal to 2. (Note that $\mathbb{Q}[\sqrt{-5}]$ has class number 2 and ring of integers $\mathbb{Z}[\sqrt{-5}]$, so Theorem 1 follows.) Carlitz's paper has spawned a large body of related work; some relevant surveys are [4, 5, 6, 7].

Since Theorem 1 can be appreciated by anyone who has completed a course on rings, it seems desirable to have a proof accessible to that same audience. Towards this end, Chapman, Gotti, and Gotti [8] offer a self-contained development of ideal theory in $\mathbb{Z}[\sqrt{-5}]$, sufficient to carry out Carlitz's argument for Theorem 1. Our approach here is somewhat different. We make a beeline towards Theorem 1, avoiding any reliance on unique factorization into ideals and making no mention of class numbers. Nevertheless, experts will recognize that our arguments share features with those appearing in the development of ideal theory; we trust this will be viewed as a feature and not a bug.

2. PREPARATION. Here we lay out some of the tools needed for the proof of Theorem 1.

Our argument makes heavy use of the norm map. For each $\alpha = a + b\sqrt{-5} \in \mathbb{Q}[\sqrt{-5}]$, we define the **conjugate** of α by $\tilde{\alpha} = a - b\sqrt{-5}$, and we define the **norm** of α by $N\alpha = \alpha\tilde{\alpha}$, so that $N\alpha = a^2 + 5b^2$. It is straightforward to check that conjugation is a field automorphism of $\mathbb{Q}[\sqrt{-5}]$; thus, for all $\alpha, \beta \in \mathbb{Q}[\sqrt{-5}]$,

$$N(\alpha\beta) = \alpha\beta \cdot \widetilde{\alpha\beta} = \alpha\beta \cdot (\tilde{\alpha}\tilde{\beta}) = \alpha\tilde{\alpha} \cdot \beta\tilde{\beta} = N\alpha \cdot N\beta.$$

For $\alpha \in \mathbb{Z}[\sqrt{-5}]$, the formula $N\alpha = a^2 + 5b^2$ shows that $N\alpha \in \mathbb{Z}_{\geq 0}$, with equality only when $\alpha = 0$. The norm being integer-valued on $\mathbb{Z}[\sqrt{-5}]$ allows one to transfer certain questions about the arithmetic of $\mathbb{Z}[\sqrt{-5}]$ to questions about $\mathbb{Z}$. For example, it is relatively straightforward now to determine the units of $\mathbb{Z}[\sqrt{-5}]$: If α is a unit of $\mathbb{Z}[\sqrt{-5}]$, with inverse $\beta \in \mathbb{Z}[\sqrt{-5}]$, then $1 = N(\alpha\beta) = N\alpha \cdot N\beta$. Since $N\alpha, N\beta \in \mathbb{Z}_{\geq 0}$, it must be that $N\alpha = 1$ (and $N\beta = 1$). Conversely, if $N\alpha = 1$, then α is a unit with inverse $\tilde{\alpha}$. Hence,

$$N\alpha = 1 \iff \alpha \text{ is a unit in } \mathbb{Z}[\sqrt{-5}].$$

Since the only solutions in integers of $a^2 + 5b^2 = 1$ are $a = \pm 1, b = 0$, the only units of $\mathbb{Z}[\sqrt{-5}]$ are ± 1.

Recall that a nonzero, nonunit element π of a domain D is said to be **prime** in D if, whenever $\pi \mid \alpha\beta$ with $\alpha, \beta \in D$, either $\pi \mid \alpha$ or $\pi \mid \beta$. Equivalently, π is prime when π is nonzero and π generates a prime ideal of D. We (continue to) say π is **irreducible** if, whenever $\pi = \alpha\beta$ with $\alpha, \beta \in D$, either α or β is a unit in D. It is a pleasant exercise to show that in any domain D every prime is irreducible. Irreducibles need not be prime; looking back at our earlier factorizations of 6 in the domain $\mathbb{Z}[\sqrt{-5}]$, each of the factors $2, 3, 1 + \sqrt{-5}$ and $1 - \sqrt{-5}$ is irreducible but none of these are prime.

For us it is of crucial importance that elements of prime norm in $\mathbb{Z}[\sqrt{-5}]$ are themselves prime.

Lemma 2. If $\pi \in \mathbb{Z}[\sqrt{-5}]$ has a norm that is prime in $\mathbb{Z}$, then π is prime in $\mathbb{Z}[\sqrt{-5}]$.

In the following argument, the expression $\#S$ denotes the cardinality of the set S and $\mathbb{F}_p$ denotes the finite field with p elements (p prime).

Proof (following [9, 10]). Let $p = N\pi$. Since $\pi \mid \pi\tilde{\pi} = p$, reduction mod π yields a well-defined surjection $\mathbb{Z}[\sqrt{-5}]/(p) \twoheadrightarrow \mathbb{Z}[\sqrt{-5}]/(\pi)$. The corresponding kernel contains π and so is nontrivial. Thus, $\#\mathbb{Z}[\sqrt{-5}]/(\pi)$ is a proper divisor of $\#\mathbb{Z}[\sqrt{-5}]/(p)$. Since $\mathbb{Z}[\sqrt{-5}]/(p)$ has p^2 elements (namely $a + b\sqrt{-5}$ for $0 \leq a, b < p$) and $\mathbb{Z}[\sqrt{-5}]/(\pi)$ is not the zero ring, this forces $\#\mathbb{Z}[\sqrt{-5}]/(\pi) = p$. It follows that $\mathbb{Z}[\sqrt{-5}]/(\pi) \cong \mathbb{F}_p$. Therefore (π) is a prime (and indeed, maximal) ideal of $\mathbb{Z}[\sqrt{-5}]$, and hence π is prime in $\mathbb{Z}[\sqrt{-5}]$. ∎

We also need a simple result from the theory of congruences, due essentially to Aubry, Thue, and Vinogradov (independently); see [11] for variants and a discussion of its history.

Lemma 3. Let m be a positive integer, and let A, B be positive real numbers with $AB \geq m$. For each integer μ, there are $x, y \in \mathbb{Z}$, not both 0, with

$$x \equiv y\mu \pmod{m}$$

and $|x| \leq A, |y| \leq B$.

Proof. We consider the residue classes mod m of the integers $x_0 - y_0\mu$, as (x_0, y_0) ranges over all ordered pairs of integers satisfying $0 \leq x_0 \leq A$ and $0 \leq y_0 \leq B$. The number of pairs (x, y) is $(1 + \lfloor A \rfloor)(1 + \lfloor B \rfloor) > AB$. Since $AB \geq m$, two of our pairs, say (x_1, y_1) and (x_2, y_2), must satisfy $x_1 - y_1\mu \equiv x_2 - y_2\mu \pmod{m}$. Then $x = x_1 - x_2$ and $y = y_1 - y_2$ are as in the lemma statement. ∎

3. PROOF OF THEOREM 1. We will deduce Theorem 1 from the following proposition. For a positive integer n, we write $\Omega(n)$ for the number of positive primes of $\mathbb{Z}$ dividing n, counted with multiplicity (for example, $\Omega(6) = \Omega(9) = 2$, since $6 = 2 \cdot 3$ while $9 = 3 \cdot 3$).

Proposition 4. If π is an irreducible of $\mathbb{Z}[\sqrt{-5}]$ that is not prime, then $\Omega(N\pi) = 2$.

Proof of Theorem 1, assuming Proposition 4. Assume that the statement of Theorem 1 is false. We choose a counterexample α of minimal norm. That is, α has two factorizations into irreducibles of different lengths, and $N\alpha$ is as small as possible among all such α. Write

$$\alpha = \pi_1 \cdots \pi_k = \rho_1 \cdots \rho_\ell, \qquad (\text{all } \pi_i, \rho_j \text{ irreducible, and } k \neq \ell). \tag{2}$$

If π_i is prime for some i, the primality of π_i implies that $\pi_i \mid \rho_j$ for some j. The irreducibility of ρ_j then forces ρ_j to be a unit multiple of π_i. We can now divide both our factorizations of α through by π_i to find that α/π_i is still a counterexample to Theorem 1, of smaller norm than α. But this contradicts the choice of α. Hence, no π_i is prime and similarly no ρ_j is prime.

Take norms in (2). Applying Proposition 4, we see that $\Omega(N\alpha) = 2k = 2\ell$. Thus $k = \ell$, a contradiction. ∎

Proof of Proposition 4. Assuming Proposition 4 is false, choose a counterexample π of minimal norm. Since π is not a unit, we have $\Omega(N\pi) > 0$. Also, since π is not prime, Lemma 2 gives $\Omega(N\pi) > 1$. As π is a counterexample to Proposition 4, it must be that $k := \Omega(N\pi) \geq 3$. We factor the integer $N\pi$ into positive primes in $\mathbb{Z}$:

$$\pi\tilde{\pi} = p_1 p_2 p_3 \cdots p_k, \quad \text{where} \quad p_1 \leq p_2 \leq \cdots \leq p_k.$$

Let us observe for later use that $\pi\tilde{\pi}$ cannot be divisible by any prime of $\mathbb{Z}[\sqrt{-5}]$. Indeed, if ρ is prime and $\rho \mid \pi\tilde{\pi}$, then the primality of ρ coupled with the irreducibility of π forces π to be an associate of ρ or $\tilde{\rho}$. But then π is prime, contrary to our hypothesis. (We have tacitly used here that conjugation is an automorphism of $\mathbb{Z}[\sqrt{-5}]$ and so preserves both primality and irreducibility.)

Writing $\pi = X + Y\sqrt{-5}$, the integers X and Y must be relatively prime. Otherwise, there is a rational prime p dividing π, which forces (by irreducibility) π to be an associate of p. But then $N\pi = p^2$ and so $\Omega(N\pi) = 2$ after all, a contradiction. From $\gcd(X, Y) = 1$ and

$$N\pi = X^2 + 5Y^2 \equiv 0 \pmod{p_1p_2}, \tag{3}$$

we deduce that

$$\gcd(Y, p_1p_2) = 1. \tag{4}$$

Thus, we can choose an integer μ with

$$X \equiv Y\mu \pmod{p_1p_2}. \tag{5}$$

From (3), (4), and (5),

$$\mu^2 \equiv -5 \pmod{p_1p_2}. \tag{6}$$

We now use Lemma 3 to choose integers x and y, not both 0, with

$$x \equiv y\mu \pmod{p_1p_2} \tag{7}$$

and $|x| \le 5^{1/4}\sqrt{p_1p_2}$, $|y| \le 5^{-1/4}\sqrt{p_1p_2}$. Put $\gamma = x + y\sqrt{-5}$. Then $N\gamma = x^2 + 5y^2 \equiv 0 \pmod{p_1p_2}$ while also

$$0 < x^2 + 5y^2 \le (2\sqrt{5})p_1p_2 < 5p_1p_2.$$

Thus, $N\gamma = p_1p_2, 2p_1p_2, 3p_1p_2$, or $4p_1p_2$.

Let us see what the congruence (7) buys us. Multiplying out,

$$\pi\tilde{\gamma} = (X + Y\sqrt{-5})(x - y\sqrt{-5}) = (xX + 5yY) + (xY - yX)\sqrt{-5}.$$

From (5), (6), and (7), we have that $xX \equiv y\mu \cdot Y\mu \equiv -5yY \pmod{p_1p_2}$, so that $p_1p_2 \mid xX + 5yY$. Also, $xY - yX \equiv (y\mu)Y - y(Y\mu) \equiv 0 \pmod{p_1p_2}$. Hence, $p_1p_2 \mid \pi\tilde{\gamma}$ in $\mathbb{Z}[\sqrt{-5}]$.

We now complete the proof by considering the different possibilities for $N\gamma$.

Suppose that $N\gamma = p_1p_2$. Then $\pi/\gamma = \pi\tilde{\gamma}/N\gamma = \pi\tilde{\gamma}/p_1p_2$. We have just seen that $p_1p_2 \mid \pi\tilde{\gamma}$, and so $\gamma \mid \pi$. Since π is irreducible, π is a unit multiple of γ. But then $N\pi = N\gamma = p_1p_2$, contradicting that $\Omega(N\pi) \ge 3$.

Next, suppose that $N\gamma = 2p_1p_2$. Then γ is irreducible. Otherwise, we can factor $\gamma = \alpha\beta$ for nonunits α, β. Taking norms, $2p_1p_2 = N\alpha \cdot N\beta$. Since 2 is not a norm from $\mathbb{Z}[\sqrt{-5}]$, this forces $N\alpha$ or $N\beta$ to be one of the primes p_1 or p_2. But then (by Lemma 2) α or β is a prime of $\mathbb{Z}[\sqrt{-5}]$ dividing $p_1 \cdots p_k = \pi\tilde{\pi}$; however, we ruled out the existence of primes dividing $\pi\tilde{\pi}$ at the start of this proof.

Furthermore, γ is not prime: Otherwise, as $\gamma \mid \gamma\tilde{\gamma} = 2p_1p_2$, we have that γ divides either $2, p_1$, or p_2. But then $N\gamma$ divides $2^2, p_1^2$, or p_2^2 in $\mathbb{Z}$, contrary to $\Omega(N\gamma) = 3$.

Since γ is irreducible, non-prime, and $\Omega(N\gamma) = 3$, the element γ is itself a counterexample to Proposition 4. The minimality of $N\pi$ therefore implies that $N\pi = p_1 \cdots p_k \leq 2p_1p_2 = N\gamma$. This forces $k = 3$ and $p_1 = p_2 = p_3 = 2$. But then $N\pi = 2 \cdot 2 \cdot 2 = 8$, which is absurd as there are no integer solutions to $u^2 + 5v^2 = 8$.

The case when $N\gamma = 3p_1p_2$ is similar. We find that γ is irreducible, non-prime, and that $k = 3$, with $p_3 \leq 3$. The case $p_3 = 2$ is ruled out as above. If $p_3 = 3$, then $N\pi$ is one of $2 \cdot 2 \cdot 3$, $2 \cdot 3 \cdot 3$, or $3 \cdot 3 \cdot 3$. But none of these are of the form $u^2 + 5v^2$: They are all 2 or 3 mod 5, whereas $u^2 + 5v^2 \equiv u^2 \equiv 0, 1$ or 4 mod 5.

The only remaining possibility is $N\gamma = 4p_1p_2$. In this case,

$$4\pi/\gamma = 4\pi\tilde{\gamma}/N\gamma = \pi\tilde{\gamma}/p_1p_2 \in \mathbb{Z}[\sqrt{-5}].$$

From here, we can conclude with a bit of trickery. Suppose $\eta = a + b\sqrt{-5}$ is any element of $\mathbb{Z}[\sqrt{-5}]$ whose norm $a^2 + 5b^2$ is a multiple of 4. Working mod 4, we see that a and b are both even, so that $\eta/2 \in \mathbb{Z}[\sqrt{-5}]$. Taking $\eta = \gamma$ we get from this argument that $\gamma/2 \in \mathbb{Z}[\sqrt{-5}]$. But now we notice that

$$N(4\pi/\gamma) = 16 \cdot N(\pi)/N(\gamma) = 4p_3 \cdots p_k,$$

so we can take $\eta = 4\pi/\gamma$ and deduce that $\frac{1}{2}(4\pi/\gamma) = \frac{\pi}{\gamma/2} \in \mathbb{Z}[\sqrt{-5}]$. Thus, $\gamma/2$ is a divisor of the irreducible π. Hence, π is a unit multiple of $\gamma/2$, and $N\pi = N(\gamma/2) = p_1p_2$, contradicting that $\Omega(N\pi) \geq 3$. ∎

Remark. Nothing in our argument requires $\mathbb{Z}[\sqrt{-5}]$ to be the full ring of algebraic integers inside $\mathbb{Q}[\sqrt{-5}]$. In fact, the method of this note can be applied equally well to establish 'half-unique'-factorization for certain nonmaximal quadratic orders, such as $\mathbb{Z}[\sqrt{-3}]$ and $\mathbb{Z}[\sqrt{5}]$. Nonmaximal quadratic orders satisfying the conclusion of Theorem 1 are investigated extensively in [12] (see also [13, pp. 226–229]), [14], [15], and [16], by more sophisticated means than those employed here.

ACKNOWLEDGMENTS. The author thanks Enrique Treviño and the referees for a number of helpful comments. This work was initiated during a stint by the author as a lecturer at the 2023 Ross Mathematics Program in Terre Haute, Indiana. He would like to thank the students and staff of the Ross Program, and particularly camp director Timothy All, for a very enjoyable six weeks. Last but not least, support by the National Science Foundation (NSF) under award DMS-2001581 is gratefully acknowledged.

REFERENCES

[1] Zaks A. Half factorial domains. Bull Amer Math Soc. 1976;82:721-3. Errata in 1976;82:965.
[2] Zaks A. Half-factorial-domains. Israel J Math. 1980;37:281-302.
[3] Carlitz L. A characterization of algebraic number fields with class number two. Proc Amer Math Soc. 1960;11:391-2.
[4] Anderson D. Elasticity of factorizations in integral domains: a survey. In: Factorization in integral domains (Iowa City, IA, 1996). vol. 189 of Lecture Notes in Pure and Appl. Math. Dekker, New York; 1997. p. 1-29.
[5] Chapman S, Coykendall J. Half-factorial domains, a survey. In: Non-Noetherian commutative ring theory. vol. 520 of Math. Appl. Kluwer Acad. Publ., Dordrecht; 2000. p. 97-115.
[6] Halter-Koch F. Non-unique factorizations of algebraic integers. Funct Approx Comment Math. 2008;39:49-60.
[7] Baginski P, Chapman S. Factorizations of algebraic integers, block monoids, and additive number theory. Amer Math Monthly. 2011;118:901-20.
[8] Chapman S, Gotti F, Gotti M. How do elements really factor in $\mathbb{Z}[\sqrt{-5}]$? In: Advances in commutative algebra. Trends Math.. Birkhäuser/Springer, Singapore; 2019. p. 171-95.

[9] Zaupper T. Unique factorization in quadratic number fields. Studia Sci Math Hungar. 1990;25:437-45.
[10] Pollack P, Snyder N. A quick route to unique factorization in quadratic orders. Amer Math Monthly. 2021;128:554-8.
[11] Brauer A, Reynolds R. On a theorem of Aubry-Thue. Canad J Math. 1951;3:367-74.
[12] Halter-Koch F. Factorization of Algebraic Integers. Grazer Math Berichte. 1983;191.
[13] Geroldinger A, Halter-Koch F. Non-unique factorizations. vol. 278 of Pure and Applied Mathematics (Boca Raton). Chapman & Hall/CRC, Boca Raton, FL; 2006.
[14] Coykendall J. Half-factorial domains in quadratic fields. J Algebra. 2001;235:417-30.
[15] Alan M. Half-factorial domains and quadratic orders. Int J Number Theory. 2016;12:465-72.
[16] Coykendall J, Malcolmson P, Okoh F. Inert primes and factorization in extensions of quadratic orders. Houston J Math. 2017;43:61-77.

PAUL POLLACK *Department of Mathematics, University of Georgia, Athens GA 30602*
pollack@uga.edu