

SQUARE PATTERNS AND INFINITELY MANY PRIMES

KEITH CONRAD

1. INTRODUCTION

Numerical data suggest the following patterns for prime numbers p :

$$\begin{aligned} -1 \equiv \square \pmod{p} &\iff p = 2 \text{ or } p \equiv 1 \pmod{4}, \\ 2 \equiv \square \pmod{p} &\iff p = 2 \text{ or } p \equiv 1, 7 \pmod{8}, \\ -2 \equiv \square \pmod{p} &\iff p = 2 \text{ or } p \equiv 1, 3 \pmod{8}, \\ 3 \equiv \square \pmod{p} &\iff p = 2, 3 \text{ or } p \equiv 1, 11 \pmod{12}, \\ -3 \equiv \square \pmod{p} &\iff p = 2, 3 \text{ or } p \equiv 1 \pmod{3}, \\ 5 \equiv \square \pmod{p} &\iff p = 2, 5 \text{ or } p \equiv 1, 4 \pmod{5}. \end{aligned}$$

As an application of such equivalences, we will use them to prove there are infinitely many primes in certain arithmetic progressions by adapting a proof going back to Euclid that there are infinitely many primes.

2. EUCLID'S PROOF OF THE INFINITUDE OF THE PRIMES

Euclid's *Elements*, which is famous mostly for its rigorous development of the theorems of plane geometry from five axioms, contains a fair bit of number theory: the Euclidean algorithm gets its name from its appearance in this work, and the existence of prime factorization is proved here as well. Proposition 20 of Book IX of the *Elements* proves the infinitude of the set of prime numbers. Here is that argument, in modern language.

Theorem 2.1 (Euclid). *There are infinitely many prime numbers.*

Proof. We know some primes already, such as 2. (We could list some more, but we just need one of them.) Suppose p_1, \dots, p_r are all prime. We want to show there is another prime off this list. The key idea is to consider the number

$$N = p_1 \cdots p_r + 1.$$

That is the product of all the primes in the list, plus one. The number N is not divisible by any of p_1, \dots, p_r since N has remainder 1 when divided by each p_i . Since $N > 1$, N has a prime factor, say p . This prime is different from p_1, \dots, p_r since N is divisible by p but not by any p_i .

If there were finitely many primes, then running through the above argument with p_1, \dots, p_r being the complete list of primes shows there is another prime, which is a contradiction. Therefore there are infinitely many primes. \square

A common misunderstanding of this proof is that it is saying if p_1, \dots, p_r are all prime then $p_1 \cdots p_r + 1$ is prime. *This need not be true.* For example, $2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 + 1 = 30031 = 59 \cdot 509$. What the proof says is that if p_1, \dots, p_r are all prime then any prime factor of $p_1 \cdots p_r + 1$ will be a prime other than one of the p_i 's, but not that $p_1 \cdots p_r + 1$ is itself prime.

Remark 2.2. Here is a recursive way to find new primes, motivated by Euclid’s proof: set $p_1 = 2$, and if we have primes p_1, \dots, p_r then let p_{r+1} be the smallest prime factor of $p_1 p_2 \cdots p_r + 1$. For instance, $p_1 + 1 = 3$ is prime, so $p_2 = 3$, and $p_1 p_2 + 1 = 7$ is prime, so $p_3 = 7$. This list of primes falls out in the following order:

2, 3, 7, 43, 13, 53, 5, 6221671, 38709183810571, 139, 2801, 11, 17, ...

This process of building prime numbers appeared in a paper of Mullin [1] and the resulting list is called the Euclid–Mullin sequence. See <https://oeis.org/A000945> for the first 43 terms. Does this list eventually include all primes? Nobody knows.

3. EXTENDING EUCLID’S PROOF TO PRIMES IN ARITHMETIC PROGRESSION

An arithmetic progression is a sequence with a common difference between successive terms. It has the form $a, a + m, a + 2m, a + 3m, a + 4m, \dots$. For example, the (positive) odd numbers are an arithmetic progression with $a = 1$ and $m = 2$. We will focus on arithmetic progressions where $0 < a < m$. In the language of congruences, an arithmetic progression is the set of (positive) integers n satisfying a congruence condition $n \equiv a \pmod{m}$.

If $(a, m) > 1$ then the arithmetic progression $a, a + m, a + 2m, a + 3m, a + 4m, \dots$ contains at most one prime number since every term in this arithmetic progression is a multiple of (a, m) . For example, there is only one prime $p \equiv 2 \pmod{4}$ and there are no primes $p \equiv 6 \pmod{8}$. If $(a, m) = 1$, on the other hand, there is no obvious reason there couldn’t be infinitely many primes $p \equiv a \pmod{m}$, and Dirichlet proved there really are infinitely many such primes.

Theorem 3.1 (Dirichlet, 1837). *If $(a, m) = 1$ then there are infinitely many prime numbers $p \equiv a \pmod{m}$.*

The proof of Dirichlet’s theorem in general is hard, but special cases are accessible to the strategy of Euclid’s proof that there are infinitely many primes. We will show for the a and m in the table below that there are infinitely many primes $p \equiv a \pmod{m}$. Most of the proofs in Section 3 will use the square patterns in the introduction.

| $a \pmod{m}$ | Theorem |
|--------------|----------------------|
| 1 mod 3 | 3.2 |
| 2 mod 3 | 3.3 |
| 1 mod 4 | 3.4 |
| 3 mod 4 | 3.5 |
| 4 mod 5 | 3.6 |
| 3 mod 8 | 3.7 |
| 5 mod 8 | 3.8 |
| 7 mod 8 | 3.9 |
| 5 mod 12 | 3.10 |
| 7 mod 12 | 3.11 |
| 11 mod 12 | 3.12 |

Theorem 3.2. *There are infinitely many primes $p \equiv 1 \pmod{3}$.*

Proof. One such prime is 7. If p_1, \dots, p_r are primes $\equiv 1 \pmod{3}$, let

$$N = (2p_1 p_2 \cdots p_r)^2 + 3.$$

Then N is not divisible by 2, 3, or by any of p_1, \dots, p_r (why?). Since $N > 1$, N has a prime factor, say p . Writing the condition $N \equiv 0 \pmod{p}$ as $(2p_1 \cdots p_r)^2 + 3 \equiv 0 \pmod{p}$, we

have $-3 \equiv (2p_1 \cdots p_r)^2 \pmod{p}$, so $-3 \equiv \square \pmod{p}$. Therefore, since $p \neq 2$ or 3 , the pattern for $-3 \equiv \square \pmod{p}$ tells us $p \equiv 1 \pmod{3}$. This prime is different from p_1, \dots, p_r , since $N \equiv 3 \not\equiv 0 \pmod{p_i}$ while $N \equiv 0 \pmod{p}$, so there are infinitely many primes $\equiv 1 \pmod{3}$. \square

Theorem 3.3. *There are infinitely many primes $p \equiv 2 \pmod{3}$.*

Proof. One such prime is 2. If p_1, \dots, p_r are primes $\equiv 2 \pmod{3}$, let

$$N = 3p_1p_2 \cdots p_r - 1.$$

Then N is not divisible by 3 or by any of p_1, \dots, p_r . Since $N > 1$, N has a prime factor. Since $N \equiv -1 \equiv 2 \pmod{3}$, the prime factors of N are not all $1 \pmod{3}$; otherwise $N \equiv 1 \pmod{3}$, because an integer greater than 1 is the product of its primes factors to some powers. Therefore N has a prime factor p that is $\equiv 2 \pmod{3}$. This prime is different from p_1, \dots, p_r , since $N \equiv -1 \not\equiv 0 \pmod{p_i}$ while $N \equiv 0 \pmod{p}$, so there are infinitely many primes $\equiv 2 \pmod{3}$. \square

Theorem 3.4. *There are infinitely many primes $p \equiv 1 \pmod{4}$.*

Proof. One such prime is 5. If p_1, \dots, p_r are primes $\equiv 1 \pmod{4}$, let

$$N = (2p_1p_2 \cdots p_r)^2 + 1.$$

Then N is not divisible by 2 or by any of p_1, \dots, p_r . Since $N > 1$, N has a prime factor, say p . Then the condition $N \equiv 0 \pmod{p}$ implies $-1 \equiv \square \pmod{p}$ (why?). Since $p \neq 2$, the pattern for $-1 \equiv \square \pmod{p}$ tells us $p \equiv 1 \pmod{4}$. This prime is different from p_1, \dots, p_r , since $N \equiv 1 \not\equiv 0 \pmod{p_i}$ while $N \equiv 0 \pmod{p}$, so there are infinitely many primes $\equiv 1 \pmod{4}$. \square

Theorem 3.5. *There are infinitely many primes $p \equiv 3 \pmod{4}$.*

Proof. One such prime is 3. If p_1, \dots, p_r are primes $\equiv 3 \pmod{4}$, let

$$N = 4p_1p_2 \cdots p_r - 1 > 1.$$

Then N is not divisible by 2 or by any of p_1, \dots, p_r . Since $N \equiv -1 \equiv 3 \pmod{4}$, the prime factors of N are not all $1 \pmod{4}$ (otherwise $N \equiv 1 \pmod{4}$). Therefore N has a prime factor p that is $3 \pmod{4}$. This prime is different from p_1, \dots, p_r , since $N \equiv -1 \not\equiv 0 \pmod{p_i}$ while $N \equiv 0 \pmod{p}$, so there are infinitely many primes $\equiv 3 \pmod{4}$. \square

The proofs of Theorems 3.3 and 3.5 did not use square patterns, but they relied on there being just two possible remainders for primes modulo 3 other than 3 and primes modulo 4 other than 4: 1 and another choice. If we try to extend the proofs of those cases to other moduli we quickly run into problems.¹ For example, if we want to show there are infinitely many primes $p \equiv 4 \pmod{5}$ then we could observe there are such primes, like 19, and if p_1, \dots, p_r are all $\equiv 4 \pmod{5}$ then the product $N = 5p_1 \cdots p_r - 1$ satisfies $N > 1$ and $N \equiv -1 \equiv 4 \not\equiv 1 \pmod{5}$, so N has a prime factor p that is not $\equiv 1 \pmod{5}$, but this doesn't mean $p \equiv 4 \pmod{5}$. For example, $5 \cdot 19 - 1 = 94 = 2 \cdot 47$ has both prime factors $\equiv 2 \pmod{5}$.

To extend Euclid's proof of the infinitude of primes in arithmetic progressions to moduli besides 3 and 4 we will use quadratic expressions to define N in the proof (by comparison, the formula for N in Theorems 3.3 and 3.5 is linear in the product $p_1 \cdots p_r$). This was already seen in Theorems 3.2 and 3.4.

¹For modulus 6 there is not a problem: the same ideas show there are infinitely many primes $p \equiv 5 \pmod{6}$. But this is not interesting since for odd p the condition $p \equiv 5 \pmod{6}$ is the same as the condition $p \equiv 2 \pmod{3}$, and we already handled this in Theorem 3.3.

Theorem 3.6. *There are infinitely many primes $p \equiv 4 \pmod{5}$.*

Proof. One such prime is 19. If p_1, \dots, p_r are primes $\equiv 4 \pmod{5}$, let

$$N = (2p_1p_2 \cdots p_r)^2 - 5 > 1.$$

Then N is not divisible by 2, 5, or p_1, \dots, p_r . Let p be any prime factor of N , so $5 \equiv \square \pmod{p}$ (why?). Therefore, since $p \neq 2$ or 5, the pattern for $5 \equiv \square \pmod{p}$ tells us $p \equiv 1$ or $4 \pmod{5}$: all prime factors of N are $1 \pmod{5}$ or $4 \pmod{5}$. To show N has a prime factor that is $4 \pmod{5}$ we argue by contradiction. If every prime factor of N is $1 \pmod{5}$, then $N \equiv 1 \pmod{5}$, but in fact $N \equiv 4 \pmod{5}$ since $p_i^2 \equiv 1 \pmod{5}$ for all i . (Here we use $p_i \equiv 4 \pmod{5}$.) Therefore some prime factor of N is not $1 \pmod{5}$. The only option left is that this prime factor is $4 \pmod{5}$. This prime is different from p_1, \dots, p_r , since $N \equiv -5 \not\equiv 0 \pmod{p_i}$ while $N \equiv 0 \pmod{p}$, so there are infinitely many primes $\equiv 4 \pmod{5}$. \square

Theorem 3.7. *There are infinitely many primes $p \equiv 3 \pmod{8}$.*

Proof. One such prime is 3. If p_1, \dots, p_r are primes $\equiv 3 \pmod{8}$, let

$$N = (p_1p_2 \cdots p_r)^2 + 2 > 1.$$

Then N is not divisible by 2 or by any of p_1, \dots, p_r . Let p be any prime factor of N , so $-2 \equiv \square \pmod{p}$. Therefore, since $p \neq 2$, the pattern for $-2 \equiv \square \pmod{p}$ says $p \equiv 1$ or $3 \pmod{8}$. We want to show N has a prime factor that is $3 \pmod{8}$, and will show this by contradiction. If every prime factor of N is $\equiv 1 \pmod{8}$, then $N \equiv 1 \pmod{8}$, but in fact $N \equiv 3 \pmod{8}$ since $p_i^2 \equiv 1 \pmod{8}$ for all i . Therefore some prime factor p of N is not $1 \pmod{8}$, so $p \equiv 3 \pmod{8}$. This prime is different from p_1, \dots, p_r , since $N \equiv 2 \not\equiv 0 \pmod{p_i}$ while $N \equiv 0 \pmod{p}$, so there are infinitely many primes $\equiv 3 \pmod{8}$. \square

Theorem 3.8. *There are infinitely many primes $p \equiv 5 \pmod{8}$.*

Proof. One such prime is 5. If p_1, \dots, p_r are primes $\equiv 5 \pmod{8}$, let

$$N = (2p_1p_2 \cdots p_r)^2 + 1 > 1.$$

Then N is not divisible by 2 or by any of p_1, \dots, p_r . Let p be any prime factor of N , so $-1 \equiv \square \pmod{p}$. Therefore, since $p \neq 2$, we have $p \equiv 1 \pmod{4}$, which is the same as $p \equiv 1$ or $5 \pmod{8}$. If every prime factor of N is $1 \pmod{8}$, then $N \equiv 1 \pmod{8}$, but in fact $N \equiv 5 \pmod{8}$ since $p_i^2 \equiv 1 \pmod{8}$ for all i . Therefore some prime factor p of N is not $1 \pmod{8}$, so $p \equiv 5 \pmod{8}$. This prime is different from p_1, \dots, p_r , since $N \equiv 1 \not\equiv 0 \pmod{p_i}$ while $N \equiv 0 \pmod{p}$, so there are infinitely many primes $\equiv 5 \pmod{8}$. \square

Theorem 3.9. *There are infinitely many primes $p \equiv 7 \pmod{8}$.*

Proof. One such prime is 7. If p_1, \dots, p_r are primes $\equiv 7 \pmod{8}$, let

$$N = (p_1p_2 \cdots p_r)^2 - 2 > 1.$$

Then N is not divisible by 2 or by any of p_1, \dots, p_r . Let p be a prime factor of N , so $2 \equiv \square \pmod{p}$. Therefore, since $p \neq 2$, the pattern for $2 \equiv \square \pmod{p}$ implies $p \equiv 1$ or $7 \pmod{8}$. If every prime factor of N is $1 \pmod{8}$, then $N \equiv 1 \pmod{8}$, but in fact $N \equiv -1 \pmod{8}$ since $p_i^2 \equiv 1 \pmod{8}$. Therefore some prime factor p of N is not $1 \pmod{8}$, so $p \equiv 7 \pmod{8}$. This prime is different from p_1, \dots, p_r , since $N \equiv -2 \not\equiv 0 \pmod{p_i}$ while $N \equiv 0 \pmod{p}$, so there are infinitely many primes $\equiv 7 \pmod{8}$. \square

Theorem 3.10. *There are infinitely many primes $p \equiv 5 \pmod{12}$.*

Proof. One such prime is 5. If p_1, \dots, p_r are primes $\equiv 5 \pmod{12}$, let

$$N = (2p_1p_2 \cdots p_r)^2 + 1 > 1.$$

Then N is not divisible by 2 or by any of p_1, \dots, p_r . Let p be any prime factor of N , so $-1 \equiv \square \pmod{p}$. Therefore, since $p \neq 2$, we have $p \equiv 1 \pmod{4}$, which is the same as $p \equiv 1$ or $5 \pmod{12}$. (The choice $p \equiv 9 \pmod{12}$ is satisfied by no prime.) If every prime factor of N is $1 \pmod{12}$, then $N \equiv 1 \pmod{12}$, but in fact $N \equiv 5 \pmod{12}$ since $p_i^2 \equiv 1 \pmod{12}$ for all i . Therefore some prime factor p of N is not $1 \pmod{12}$, $p \equiv 5 \pmod{12}$. This prime is different from p_1, \dots, p_r , since $N \equiv 1 \not\equiv 0 \pmod{p_i}$ while $N \equiv 0 \pmod{p}$, so there are infinitely many primes $\equiv 5 \pmod{12}$. \square

Theorem 3.11. *There are infinitely many primes $p \equiv 7 \pmod{12}$.*

Proof. One such prime is 7. If p_1, \dots, p_r are primes $\equiv 7 \pmod{12}$, let

$$N = (2p_1 \cdots p_r)^2 + 3.$$

Then N is not divisible by 2 or by any of p_1, \dots, p_r . Let p be any prime factor of N , so $-3 \equiv \square \pmod{p}$. Therefore, since p is not 2 or 3, the pattern for $-3 \equiv \square \pmod{p}$ implies $p \equiv 1 \pmod{3}$. Lifting this mod 3 congruence to modulus 12 tells us $p \equiv 1, 4, 7$ or $10 \pmod{12}$. No primes are $4 \pmod{12}$ or $10 \pmod{12}$, so $p \equiv 1$ or $7 \pmod{12}$. If every prime factor of N is $\equiv 1 \pmod{12}$, then $N \equiv 1 \pmod{12}$, but in fact $N \equiv 7 \pmod{12}$ since $p_i^2 \equiv 1 \pmod{12}$ for all i (so $N \equiv 4 + 3 \pmod{12}$). Therefore some prime factor p of N is not $1 \pmod{12}$, so $p \equiv 7 \pmod{12}$. This prime is different from p_1, \dots, p_r , since $N \equiv 3 \not\equiv 0 \pmod{p_i}$ while $N \equiv 0 \pmod{p}$, so there are infinitely many primes $\equiv 7 \pmod{12}$. \square

Theorem 3.12. *There are infinitely many primes $p \equiv 11 \pmod{12}$.*

Proof. One such prime is 11. If p_1, \dots, p_r are primes $\equiv 11 \pmod{12}$, let

$$N = 3(p_1p_2 \cdots p_r)^2 - 4 > 1.$$

Then N is not divisible by 2, 3, or any of p_1, \dots, p_r . Let p be a prime factor of N , so $3 \equiv \square \pmod{p}$ (why?). Therefore, since $p \neq 2$ or 3 , the pattern for $3 \equiv \square \pmod{p}$ implies $p \equiv 1$ or $11 \pmod{12}$. If every prime factor of N is $1 \pmod{12}$, then $N \equiv 1 \pmod{12}$, but in fact $N \equiv -1 \pmod{12}$ since $p_i^2 \equiv 1 \pmod{12}$ for all i . Therefore some prime factor p of N is not $1 \pmod{12}$, so $p \equiv 11 \pmod{12}$. This prime is different from p_1, \dots, p_r , since $N \equiv -4 \not\equiv 0 \pmod{p_i}$ while $N \equiv 0 \pmod{p}$, so there are infinitely many primes $\equiv 11 \pmod{12}$. \square

In all these proofs, we used a polynomial whose values on integers have special congruence conditions on their prime factors, *e.g.*, to show $p \equiv 4 \pmod{5}$ infinitely often we relied on the fact that any integer of the form $n^2 - 5$ with n even and $n \not\equiv 0 \pmod{5}$ is only divisible by primes $p \equiv 1, 4 \pmod{5}$: if $p \mid (n^2 - 5)$ then $5 \pmod{p}$ is a square, so $p \equiv 1, 4 \pmod{5}$ if $p \neq 2, 5$. Thus the proof of Theorem 3.6 relies on a feature of the polynomial $T^2 - 5$. The table below is a summary of the polynomial and the square condition used for each congruence condition above. Euclid's proof of the infinitude of the primes is associated to the linear polynomial $T + 1$. (Recall the role of $p_1 \cdots p_r + 1$ in that proof.) The proofs using square patterns all involve a quadratic polynomial.

| Congruence | Polynomial | Square condition |
|------------|------------|------------------------------|
| 1 mod 3 | $T^2 + 3$ | $-3 \equiv \square \pmod{p}$ |
| 2 mod 3 | $T - 1$ | None |
| 1 mod 4 | $T^2 + 1$ | $-1 \equiv \square \pmod{p}$ |
| 3 mod 4 | $T - 1$ | None |
| 4 mod 5 | $T^2 - 5$ | $5 \equiv \square \pmod{p}$ |
| 3 mod 8 | $T^2 + 2$ | $-2 \equiv \square \pmod{p}$ |
| 5 mod 8 | $T^2 + 1$ | $-1 \equiv \square \pmod{p}$ |
| 7 mod 8 | $T^2 - 2$ | $2 \equiv \square \pmod{p}$ |
| 5 mod 12 | $T^2 + 1$ | $-1 \equiv \square \pmod{p}$ |
| 7 mod 12 | $T^2 + 3$ | $-3 \equiv \square \pmod{p}$ |
| 11 mod 12 | $3T^2 - 4$ | $3 \equiv \square \pmod{p}$ |

We have proved Dirichlet's theorem (Theorem 3.1) for all cases where $m = 3, 4, 5, 8,$ and 12 except for $1, 2, 3 \pmod{5}$, $1 \pmod{8}$, and $1 \pmod{12}$. These remaining cases for $m = 5, 8,$ and 12 can be handled by elementary techniques in the style we used above, but we have to replace quadratic polynomials with higher degree polynomials. We do not discuss this further here.

REFERENCES

- [1] A. A. Mullin, "Recursive function theory (A modern look at a Euclidean idea)," *Bull. Amer. Math. Soc.* **69** (1963), 737.