FOR WHICH ARITHMETIC PROGRESSIONS IS THERE A EUCLIDEAN PROOF OF DIRICHLET'S THEOREM?

PAUL POLLACK

Abstract. Dirichlet's 1837 theorem on primes in arithmetic progressions asserts that every invertible residue class contains infinitely many prime numbers. Dirichlet's proof was analytic, but many special cases of this theorem have more elementary, purely algebraic proofs whose general strategy closely resembles Euclid's proof of the infinitude of primes. It is natural to wonder what portion of Dirichlet's theorem can be proved by such "Euclidean arguments." Murty [Mur88] showed (unconditionally) that relative to a particular formalization of "Euclidean proof," the only such progressions $a \bmod m$ are those for which $a^2 \equiv 1 \pmod{m}$. Assuming Schinzel's Hypothesis H, we establish the same result for an a priori more inclusive (and arguably more natural) notion of Euclidean proof.

1. Introduction

1.1. **Motivation.** Are there infinitely many prime numbers which end in the digit 7? This is a simple and natural question about primes which anyone learning about them for the first time might well be inclined to ask. To number theorists the answer is well known: the boldfaced sequence

$$\mathbf{7}, \mathbf{17}, 27, \mathbf{37}, \mathbf{47}, 57, \mathbf{67}, 77, 87, \mathbf{97}, \mathbf{107}, 117, \mathbf{127}, \mathbf{137}, 147, \mathbf{157}, \dots$$

does indeed go on forever. Indeed, this theorem is easily recognizable as a special case ($a = 7$, $m = 10$) of the following 1837 result of Dirichlet [Dir37], one of the crowning achievements of early analytic number theory:

Dirichlet's Theorem on Primes in Arithmetic Progressions. *Let a and m be integers with m positive, and suppose that a is relatively prime to m. Then the arithmetic progression*

$$a, a + m, a + 2m, \dots$$

contains infinitely many primes.

Unfortunately Dirichlet's argument is by no means simple, and our hypothetical questioner might well be a bit put off by all the details necessary to verify the proof – details from *analysis*, no less, an area which seems quite remote from our opening problem. Proofs which minimize analytic prerequisites exist (e.g., those of Zassenhaus [Zas49], Selberg [Sel49], and Shapiro [Sha50a], [Sha50b]), but these "elementary" proofs exhibit at least as complicated a structure as Dirichlet's original argument. The contortions necessary to establish Dirichlet's theorem stand in stark contrast to the elegant and simple proof offered by Euclid for the infinitude

2000 *Mathematics Subject Classification.* 11N32 (Primary); 11A25, 11N13 (Secondary).
The author is supported by an NSF Graduate Research Fellowship.

of the primes. Is this difficulty inherent in the problem itself or merely an artifact of our own ignorance?

For certain progressions we know how to establish the conclusion of Dirichlet's theorem by a method hardly more difficult than Euclid's. A well-known proof for the progression $-1 \bmod 4$ runs as follows:

Proof that there are infinitely many primes $p \equiv -1 \pmod 4$. Let $p_1, p_2, \dots, p_k$ be any finite (possibly empty) list of primes congruent to $-1 \bmod 4$ and let $P := 4p_1 \cdots p_k - 1$. Since $P > 1$ and P is odd, P decomposes as a product of odd primes. Moreover, since $P \equiv -1 \pmod 4$, not all of the primes dividing P can belong to the residue class 1 mod 4. So P has a prime divisor p from the residue class $-1 \bmod 4$. For the same reasons as in Euclid's proof, p is distinct from all of $p_1, \dots, p_k$, and so our original list cannot possibly be complete. □

Lebesgue [Leb56] gives a version of this proof, noting that "cette démonstration est imitée d'Euclide," and in the same paper he goes on to give a "Euclidean" proof for the progression $1 \bmod 4$. Dickson's *History* records several further attempts at giving Euclidean proofs for other special cases of Dirichlet's theorem (see the listing on [Dic66, pp. 418-420]). More recently, Bateman & Low [BL65] have given Euclidean proofs for all coprime residue classes mod 24 (in the sense of exhibiting explicit Euclidean polynomials of type I – see below).

Is there an argument for all of Dirichlet's theorem along similar Euclidean lines? If not, can we characterize the progressions for which we have such a proof?

1.2. **What is a Euclidean Proof?** To have any hope of answering this question we must formalize the notion of a Euclidean proof. Our approach to the problem begins with the following definition:

Definition 1. *A* Euclidean Polynomial of Type I for the progression $a \bmod m$ *is a polynomial f with integer coefficients possessing the following two properties:*

(i) *$f(n)$ has a prime divisor $p \equiv a \pmod m$ for all large enough integers n,*
(ii) *f has no fixed prime divisor $p \equiv a \pmod m$; in other words, if $p \equiv a \pmod m$ is prime, then there is some n for which p does not divide $f(n)$.*

Example. The reader can easily check that $f(T) = 4T - 1$ is a Type I Euclidean polynomial for the progression $-1 \pmod 4$. Note that this would not be true if the words "large enough" were omitted from condition (i) in Definition 1. Similarly, using Euler's characterization of the primes for which -1 is a square, we find that $4T^2 + 1$ is a Type I Euclidean polynomial for the progression $1 \bmod 4$.

Not surprisingly, the properties singled out in Definition 1 are exactly the properties needed to make Euclid's proof work for the progression $a \bmod m$:

Proposition 1.1. *If a Type I Euclidean polynomial exists for the progression $a \bmod m$, then there are infinitely many primes $p \equiv a \pmod m$.*

We will prove this proposition in §2.

In this paper when we speak of a Euclidean proof existing for a progression $a \bmod m$, we mean that a Euclidean polynomial of type I exists for that progression.

Progression	Type I Polynomial	Type II Polynomial
$-1 \bmod 4$	$4T-1$	T
$1 \bmod 4$	$4T^2+1$	T^2+1
$4 \bmod 5$	$100T^2+40T-1$	T^2-5
$1 \bmod 8$	$16T^4+1$	T^4+1
$3 \bmod 8$	$4T^2+4T+3$	T^2+2
$5 \bmod 8$	$4T^2+4T+5$	T^2+4
$7 \bmod 8$	$4T^2+4T-1$	T^2-2
$7 \bmod 24$	$1296T^4+864T^3+288T^2+48T+7$	T^4+2T^2+4
$1 \bmod m$	$\Phi_m(mT)$	$\Phi_m(T)$
$-1 \bmod m$	see text	$G_m(T)$

TABLE 1. Further examples of Euclidean polynomials of both types for various arithmetic progressions. See §3 for the definition of G_m.

1.3. **A Competing Notion of a Euclidean Proof.** An alternate definition of a Euclidean proof, together with a subsequent characterization of progressions for which such proofs exist, has been offered by Murty ([Mur88], and more recently [MT06]). To minimize confusion and because this notion will be helpful in the sequel, we recall some of his ideas here.

If f is a polynomial with integer coefficients, we say that a rational prime p is a *prime divisor of* f if p divides $f(n)$ for some integer n.

Definition 2. *A* Euclidean Polynomial of Type II for the progression $a \bmod m$ *is a polynomial f with integer coefficients for which both of the following hold:*

(i) *every prime divisor p of f, with at most finitely many exceptions, satisfies $p \equiv 1 \pmod m$ or $p \equiv a \pmod m$,*
(ii) *infinitely many prime divisors p of f satisfy $p \equiv a \pmod m$.*

Example. The polynomial $f(T) = T^2+2$ is a Euclidean polynomial of type II for the progression $3 \pmod 8$. Indeed, the supplements to the quadratic reciprocity law imply that every odd prime divisor of T^2+2 belongs to either the progression $1 \bmod 8$ or $3 \bmod 8$, so that (i) holds. The same supplementary laws allow us to deduce (ii): suppose that only finitely many primes from the progression $3 \pmod 8$ divide f, say $p_1, \ldots, p_k$, and then consider the prime factorization of $(p_1 \cdots p_k)^2+2$.

Example. Other examples of Euclidean polynomials of types I and II are given in Table 1.

As we will see in §3, whenever a progression possesses a Euclidean polynomial of type II, it also possesses one of type I, and so by Proposition 1.1 there are Euclidean proofs for these progressions. Euclidean Polynomials of Type II naturally arise in Schur's work on special cases of Dirichlet's theorem, which will be reviewed in §3.3. Using class field theory, Murty established the "only if" portion of the following striking theorem (the "if" portion having been earlier proved by Schur):

Murty's Impossibility Theorem. *A Euclidean Polynomial of Type II exists for the progression $a \bmod m$ if and only if $a^2 \equiv 1 \pmod m$.*

In particular, by Murty's definitions, Euclidean methods cannot provide an affirmative answer to our opening question.

1.4. **Comparison.** We have already mentioned that whenever a Euclidean polynomial of type II exists for a given progression, there is a Euclidean polynomial of type I for the same progression. Thus our notion of a Euclidean proof is a priori more inclusive than Murty's, and it is reasonable to wonder whether there are any progressions $a \bmod m$ for which Euclidean proofs exist (in our sense) other than those for which $a^2 \equiv 1 \bmod m$.

We have not been able to unconditionally characterize those progressions for which a Euclidean polynomial of type I exists. However, if we assume the following well-known conjecture of Schinzel [SS58], then we are able to make progress:

Hypothesis H. *Let $f_1, f_2, \dots, f_k$ be nonconstant polynomials which have integer coefficients, positive leading coefficients, and which are all irreducible over the rationals. Suppose $f := f_1 f_2 \dots f_k$ has no fixed prime divisor, i.e., there is no prime p dividing $f(n)$ for all integers n. Then*

$$f_1(n), f_2(n), \dots, f_k(n) \text{ are simultaneously prime}$$

for arbitrarily large positive integer values of n.

Our purpose here is to establish the following conditional theorem:

Theorem 1.2. *Assume Hypothesis H. Then whenever there is a Euclidean polynomial of type I for a progression $a \bmod m$, there is also one of type II. Thus (under Hypothesis H) the only progressions for which there exist a Euclidean proof are those for which $a^2 \equiv 1 \pmod m$.*

Example. To sense some of the difficulties involved in proving Theorem 1.2 without the assumption of Hypothesis H, consider the problem of showing that there is no Type I Euclidean polynomial for the progression 2 mod 5, or even the subproblem of showing that the specific polynomial T^2+1 is not a Type I Euclidean polynomial for this progression. For the latter problem one needs to establish that there are arbitrarily large n for which $n^2 + 1$ is free of prime factors from the progression 2 mod 5. This is straightforward if Hypothesis H is assumed but is perhaps not so obvious otherwise. In §5.2 we present clever proofs of this assertion communicated to the author by Lenstra and Coppersmith. However, the following similar question, which would need to be settled to rule out a Euclidean proof for the progression 2 mod 7, is untouched by their methods and appears to be unsolved:

Open question. *Is $n^2 + 2$ free of prime factors from the progression $2 \bmod 7$ infinitely often?*

Notation. We use the notation $d \parallel n$ to denote that d is a unitary divisor of n, i.e., that d divides n while d is coprime to n/d. We use $\mathrm{Res}(f, g)$ to denote the resultant of the two polynomials f and g. By the discriminant of a polynomial f, we always mean the so-called Sylvester discriminant defined by $\mathrm{SylDisc} f := \mathrm{Res}(f, f')$.

2. Proof of Proposition 1.1

Proof. Let $q_1, \dots, q_k$ be a finite (possibly empty) list of primes from the congruence class $a \bmod m$; we find another. Condition (ii) in the definition of a Type I Euclidean polynomial (Definition 1) allows us to choose an integer n_0 so that

$f(n_0)$ is coprime to $Q := \prod q_i$: indeed, this is a satisfiable congruence condition on $n_0 \pmod{q_i}$ for each i, so that we can obtain n_0 from the Chinese Remainder Theorem. Then for any integer n,

$$f(Qn + n_0) \equiv f(n_0) \pmod{Q}, \quad \text{hence} \quad \gcd(f(Qn + n_0), Q) = 1.$$

Now for large n condition (i) Definition 1 guarantees $f(Qn+n_0)$ has a prime factor in the progression $a \bmod m$, but since $f(Qn + n_0)$ is prime to Q, this cannot be any of the given q_i. □

3. Schur's Ideas: A Brief Review

Murty's definition of Type II Euclidean Polynomials is motivated by the work of Schur. Let m be a positive integer and let $L = \mathbf{Q}(\zeta_m)$ be the mth cyclotomic field. Then $L/\mathbf{Q}$ is Galois and we may identify the Galois group of this extension with $(\mathbf{Z}/m\mathbf{Z})^\times$. The next theorem is due to Schur (1912):

Theorem A. *Let H be a subgroup of $(\mathbf{Z}/m\mathbf{Z})^\times$, which we identify with a subgroup of $\mathrm{Gal}(L/\mathbf{Q})$ under the above correspondence. Let K be the fixed field of H, and let θ be any primitive element for the extension $K/\mathbf{Q}$, i.e., any θ for which $K = \mathbf{Q}(\theta)$. Suppose $f(T) \in \mathbf{Z}[T]$ is irreducible over $\mathbf{Q}$ and that $f(\theta) = 0$. Then the set of prime divisors of f coincides, up to finitely many exceptions, with the set of primes p for which $p \bmod m$ lies in H.*

(Here and below we say two sets *coincide with finitely many exceptions* if their symmetric difference is finite.) Proofs of Theorem A can be found in [Sch12], [Mur88], and [MT06]. Here we content ourselves with a few examples:

Example. Let m be any positive integer, and take for H the trivial subgroup. Then $K = L = \mathbf{Q}(\zeta_m)$, and we may take for f the minimal polynomial of ζ_m, which is the mth cyclotomic polynomial Φ_m. Thus it should be no surprise that the mth cyclotomic polynomial often appears in proofs that there are infinitely many primes $\equiv 1 \bmod m$ (e.g., the proof of Wendt [Wen95]).

Example. Let $\Phi_m(T_1, T_2)$ denote the mth homogenized cyclotomic polynomial, so that $\Phi_m(T_1, T_2) := T_2^{\varphi(m)}\Phi_m(T_1/T_2)$. It is not hard to check that $\Phi_m(T_1 + T_2, T_1 - T_2)$ can be written as a polynomial in T_1 and T_2^2. Thus $G_m(T) := \Phi_m(T + i, T - i)$ has integer coefficients. The characterization of the prime divisors of G_m is given by a (more general) theorem of Kronecker [Kro88]: apart from finitely many exceptions, these are exactly the primes $p \equiv \pm 1 \pmod{[4, m]}$. This result is used by Bauer [Bau06] in his proof that for each m there are infinitely many primes $p \equiv -1 \pmod m$ (see, e.g., [Nag64, §§49-50] for a modern account of this proof).

Kronecker's argument is entirely elementary but somewhat unenlightening. Theorem A removes the mystery: It is an exercise in Galois theory to show that $\cot(\pi/m)$ generates the fixed field of $\mathbf{Q}(\zeta_{[4,m]})$ corresponding to the two-element subgroup $\{\pm 1\} \subset (\mathbf{Z}/[4, m]\mathbf{Z})^\times$. One then checks that if 4 does not divide m, then the polynomial $G_m(T)$ obtained this way is exactly the minimal polynomial $p(T)$ of $\cot(\pi/m)$, and that if 4 does divide m then (with the same notation) $G_m(T) = p(T)p(-T)$.

Example. Let m be any positive integer and let $H = \{1 \bmod m, a \bmod m\}$ be any 2-element subgroup of $(\mathbf{Z}/m\mathbf{Z})^\times$. Note that this is the case precisely when $a \not\equiv 1 \pmod m$ while $a^2 \equiv 1 \pmod m$. In this case, Schur's Theorem A gives us

a polynomial whose prime divisors are exactly the primes $1 \bmod m$ or $a \bmod m$, apart from finitely many exceptions. Such a polynomial f automatically satisfies condition (i) of Definition 2, and Dirichlet's theorem guarantees that condition (ii) also holds; thus these polynomials are Euclidean Polynomials of Type II for the progression $a \bmod m$.

The reader may be a bit disconcerted to see an appeal to Dirichlet's theorem in the last example, given the theme of the paper. This is not as much a cause for alarm as it may first appear; before we can explain why, we need to march a bit further in the direction of the main result of this section, that the existence of a Euclidean polynomial of type II for a given progression implies the existence of a Euclidean polynomial of type I.

Lemma 3.1 (Schinzel)**.** *Let f be a polynomial with integer coefficients, and let D be the greatest fixed divisor of f, i.e., $D := \gcd\{f(n)\}_{n\in\mathbf{Z}}$. There exist integers A, B with $B > 0$ so that $\frac{1}{D}f(AT + B)$ has integer coefficients and no fixed prime divisor.*

We have attributed the lemma to Schinzel, since an essentially equivalent result is implicit in Schinzel's proof [Sch62] that Hypothesis H implies Bunyakovsky's conjecture on prime values of integer-valued polynomials.

Proof. We take $A = D$ and look for a suitable integer value of B. Note that with this choice of A,

$$g(T) := \frac{1}{D}f(AT + B) = \frac{1}{D}\left(f(B) + f'(B)DT + \frac{1}{2}f''(B)(DT)^2 + \dots\right)$$

has integer coefficients, regardless of our later choice of B.

If p is a fixed prime divisor of g, then p is also a fixed prime divisor of the polynomial $f(AT + B)$. If p does not divide A, then the only way p can be a fixed prime divisor of $f(AT + B)$ is if p is a fixed prime divisor of f, i.e., if p divides D. So either p divides A or p divides D. But since $A = D$, in either case we may conclude that p divides D.

Since $g(0) = \frac{1}{D}f(A\cdot 0+B) = f(B)/D$, in order that g have no fixed prime divisor it suffices to choose B in such a way that $\gcd(D, f(B)/D) = 1$. Write $D = \prod p^{e_p}$. For each prime p dividing D, choose an integer B_p so that $\gcd(f(B_p)/D, p) = 1$. We can then take for B any integer which satisfies $B \equiv B_p \pmod{p^{e_p+1}}$ for all p dividing D. □

Corollary 3.2. *Let f be an arbitrary nonzero polynomial with integer coefficients. Then one can find a polynomial g without a fixed prime divisor for which the set of prime divisors of f coincides, except for possibly finitely many exceptions, with the set of prime divisors of g.*

Proof. Let $g := \frac{1}{D}f(AT + B)$ be as in the statement of Lemma 3.1. Then apart from (possibly) the primes p dividing AD, we have that p is a prime divisor of f if and only if p is a prime divisor of g. □

Lemma 3.3. *Let $a \bmod m$ be an arbitrary arithmetic progression. Suppose that f is a polynomial with integer coefficients with the property that all but finitely many of its prime divisors belong to one of the residue classes $1 \bmod m$ or $a \bmod m$. Moreover, suppose*

(i) *f has no fixed prime divisor, and*
(ii) *there is a prime $p \equiv a \pmod m$ dividing f which does not divide* $\mathrm{SylDisc}(f)$.
Then there is a Euclidean polynomial of type I for the progression $a \bmod m$.

Before proving Lemma 3.3, we note that the main result of this section emerges as an immediate corollary:

Corollary 3.4. *If there is a Euclidean polynomial of type II for the progression $a \bmod m$, then there is a Euclidean polynomial of type I.*

Proof. Let f be a Euclidean polynomial of type II for the progression $a \bmod m$. Then some irreducible factor of f in $\mathbf{Z}[T]$ must also be a Euclidean polynomial of type II for this progression, so we may as well assume that f is irreducible. If f has any fixed prime divisors, remove them using Lemma 3.1 to obtain a new irreducible polynomial, say g, with the same prime divisors as f (up to finitely many exceptions). Then g is also a Euclidean polynomial of type II. We complete the proof by applying Lemma 3.3 to g: Hypothesis (i) of that lemma holds by Definition 2, while hypothesis (ii) holds for any large prime divisor of g from the residue class $a \bmod m$. (Note that $\mathrm{SylDisc}(g) \neq 0$ as g is irreducible.) Since in Definition 2 we are supposing that g has infinitely many prime divisors from this residue class, the result follows. □

Proof of Lemma 3.3. If the conditions of Lemma 3.3 hold for f, then they also hold for at least one of the irreducible factors of f in $\mathbf{Z}[T]$. So we can assume f is irreducible over the integers (and thus over the rationals). Moreover, multiplying through by -1 if necessary, we can assume further that the leading coefficient of f is positive.

By hypothesis (ii), we may choose a prime $p \equiv a \pmod m$ which divides f but not does divide the discriminant of f. Let n be an integer for which p divides $f(n)$. Then p does not divide $f'(n)$, otherwise p would divide the discriminant of f. So by Taylor's theorem, if we adjust n by (at most) a multiple of p we can assume $p \parallel f(n)$. By the Chinese Remainder Theorem, we may choose n' so that $n' \equiv n \pmod{p^2}$ and so that $f(n')$ is coprime to E, where E is the product of the finitely many primes that divide f but belong to neither of the residue classes $1 \bmod m$ or $a \bmod m$. Write $f(n') = pD$. Then

$$f(mDET + n') = Dg(T), \quad \text{where } g(T) \in \mathbf{Z}[T] \text{ and } g \text{ has constant term } p.$$

We claim that g is a Euclidean polynomial of type I for the progression $a \bmod m$. Indeed, suppose that n is a large integer. Since

$$Dg(n) \equiv f(n') \pmod E, \qquad \text{we have} \qquad (Dg(n), E) = 1$$

by the choice of n'. So $g(n)$ is divisible only by primes from the residue classes $1 \bmod m$ and $a \bmod m$. If $a \equiv 1 \pmod m$, this verifies condition (i) in the definition of a Euclidean polynomial of Type I (Definition 1). Otherwise, noting that

$$Dg(n) \equiv f(n') = pD \pmod{mD},$$

we see

$$g(n) \equiv p \equiv a \pmod m.$$

So if n is large enough that $g(n) > 0$, it follows that not all prime factors of $g(n)$ belong to the residue class $1 \bmod m$. Hence condition (i) of Definition 1 holds in any case.

To guarantee condition (ii), note that any fixed prime divisor of g must divide $g(0) = f(n')/D = p$. But p is not a fixed prime divisor of g, as we now check. Certainly $p \nmid mE$. But also $p \nmid D$, since

$$pD = f(n') \equiv f(n) \not\equiv 0 \pmod{p^2}.$$

If p is a fixed prime divisor of g, it is also a fixed prime divisor of $Dg(T) = f(mDET + n_2)$, but then since $p \nmid mDE$ it is fixed prime divisor of f, an absurdity. □

We are now in a position to assuage any potential fears over our earlier appeal to Dirichlet's Theorem. Suppose we are given a progression $a \bmod m$ for which $a^2 \equiv 1 \pmod{m}$. Theorem A supplies us with an irreducible polynomial f which we suspect to be a Euclidean polynomial of type II for the given progression. The method of Lemma 3.1 can be turned into an algorithm which converts f into a polynomial without a fixed prime divisor which is still a Euclidean polynomial of type II for our progression. We then compute the discriminant of the result, and look for a prime $p \equiv a \bmod m$ among the prime divisors of f which does not divide this discriminant. Having found such a prime, the proof of Lemma 3.3 describes how to construct a Euclidean polynomial of type I for $a \bmod m$. Since Dirichlet's theorem is true, this sequence of steps will always go through, but for any particular progression we need make no appeal to Dirichlet's Theorem.

By arguments of the same kind as those given here, one can prove the following clean result: if $a \bmod m$ generates a subgroup of order 2 of $(\mathbf{Z}/m\mathbf{Z})^\times$, then there is always a Euclidean polynomial of type I for the progression $a \bmod m$ provided there is a single prime p belonging to this progression obeying the inequality $p > \varphi(m)/2$. This was already worked out by Schur [Sch12]; we will not need this in the sequel and so say nothing more about it here.

4. Characterizing Euclidean Polynomials of Type I

In this section we prove Theorem 1.2. The proof requires several preliminaries beginning with an analogue of Lemma 3.2:

Lemma 4.1. *If $a \bmod m$ possesses a Euclidean polynomial of type I, then it also possesses such a polynomial with no fixed prime divisor.*

Proof. Let f be a Euclidean polynomial of type I for the given progression, and let $g = \frac{1}{D} f(AT + B)$ be a polynomial as in the conclusion of Lemma 3.1. Then g is a Euclidean polynomial of type I for $a \bmod m$. Indeed, suppose n is large. Since f satisfies condition (i) of Definition 1, there is a prime $p \equiv a \pmod{m}$ that divides $f(An + B)$; since f satisfies condition (ii), this prime p does not divide D, hence it also divides $g(n)$. This shows (i) holds for g. And since g has no fixed prime divisor at all, (ii) holds as well. □

Lemma 4.2. *Suppose that the progression $a \bmod m$ possesses a Euclidean polynomial of type I. Then we can find such a polynomial $f(T) \in \mathbf{Z}[T]$ which has a factorization in $\mathbf{Z}[T]$ of the form*

$$f(T) = f_1(T) \cdots f_k(T),$$

and where all of the following hold:

(i) *f has no fixed prime divisor,*
(ii) *f has constant term coprime to m,*

(iii) *f has positive leading coefficient,*
(iv) *the f_i have integer coefficients, positive leading coefficients and are distinct and irreducible over the rationals,*
(v) *f and the f_i are constant modulo m; i.e., the natural reductions $\overline{f}(T) \in (\mathbf{Z}/m\mathbf{Z})[T]$ and $\overline{f_i}(T) \in (\mathbf{Z}/m\mathbf{Z})[T]$ are constant polynomials.*

Proof. We begin by constructing a Euclidean polynomial $g(T)$ of type I with a factorization $g = g_1 \dots g_k$ for which all of (i) - (iv) are satisfied. We then describe the modifications necessary to obtain (v).

Choose a Euclidean polynomial $g(T)$ of type I for the progression $a \bmod m$ with no fixed prime divisor. This is possible by Lemma 4.1; then already (i) is satisfied. Apply the Chinese Remainder Theorem to find an integer n with $g(n)$ coprime to m, then replace g with $g(T+n)$; this gives (ii). Write

$$g(T) = \pm g_1(T) \cdots g_k(T) \tag{1}$$

where the g_i are irreducible over the integers and the leading coefficient of each g_i is positive. Since g has no fixed prime divisor, the g_i are nonconstant, so each g_i is irreducible also over the rationals. Moreover, because the g_i on the right hand side are all primitive with positive leading coefficient, no two are associates in $\mathbf{Q}[T]$ without being equal. Now replace g by the product of the distinct g_i. This gives (iii) and (iv); note that g remains a Euclidean polynomial of type I after this transformation.

To obtain (v) set $f(T) := g(mT)$ and $f_i(T) := g_i(mT)$ for $1 \le i \le k$. Clearly (ii)-(v) now hold. Since g was without fixed prime divisor, all fixed prime divisors of f divide m. But $f(0) = g(0)$ is coprime to m, so (i) holds as well. Finally, f is still a Euclidean polynomial of type I for the given progression, since property (i) in Definition 1 is inherited from g and property (ii) in Definition 1 is weaker than condition (i) above. □

We also require an easy but technical consequence of Hypothesis H:

Lemma 4.3. *Assume Hypothesis H. Suppose $f_1, \dots, f_k$ are distinct nonconstant polynomials with integer coefficients, positive leading coefficients, all irreducible over the rationals and such that the product $f := f_1 \dots f_k$ has no fixed prime divisor. Let $0 \le r \le k$ and suppose that for each $1 \le i \le r$ we are given a prime divisor p_i of f_i. Moreover, suppose*

$$p_1 p_2 \dots p_r \text{ is prime to } \prod_{1 \le i \le k} \mathrm{SylDisc}(f_i) \prod_{1 \le i < j \le k} \mathrm{Res}(f_i, f_j). \tag{2}$$

Then there are arbitrarily large n for which

$$f_1(n) = p_1 q_1,\ f_2(n) = p_2 q_2,\ \dots,\ f_r(n) = p_r q_r,\ f_{r+1}(n) = q_{r+1}, \dots,\ f_k(n) = q_k,$$

where all the q_i are prime.

Note that under the hypotheses of the lemma the product appearing in (2) is nonvanishing. Indeed, the discriminants are nonzero since the f_i are irreducible and the resultants are nonzero since the given conditions imply that the f_i are nonassociated over the rationals.

Proof. Let $1 \le i \le r$. Since p_i divides f_i, we can choose an integer n_i with $f_i(n_i) \equiv 0 \pmod{p_i}$. Since $p_i \nmid \mathrm{SylDisc}(f_i) = \mathrm{Res}(f_i, f_i')$, we have $f_i'(n_i) \not\equiv 0 \pmod{p_i}$, and

hence by adjusting n_i by a multiple of p we can in fact assume $p \parallel f_i(n_i)$. Choose n_0 with $n_0 \equiv n_i \pmod{(\prod p_j)^2}$ for all $1 \le i \le r$. Then $p \| f_i(n_0)$ for each i.

Let $P = (\prod p_i)^2$. Define polynomials $g_1, \dots, g_k$ by

$$f_1(PT + n_0) = p_1 g_1(T), f_2(PT + n_0) = p_2 g_2(T), \dots, f_r(PT + n_0) = p_r g_r(T),$$
$$f_{r+1}(PT + n_0) = g_{r+1}(T), f_{r+2}(PT + n_0) = g_{r+2}(T), \dots, f_k(PT + n_0) = g_k(T).$$

It suffices to check that the conditions of Hypothesis H hold for $g_1, \dots, g_k$.

The polynomials g_i all have integer coefficients. Moreover, the g_i are nonconstant and irreducible over the rationals, since the f_i are. So we need only verify that the product

$$g(T) := (g_1 \dots g_k)(T) = \frac{(f_1 \dots f_k)(PT + n_0)}{p_1 \dots p_r}$$

has no fixed prime divisor. Any such is also a fixed prime divisor of $f(PT + n_0)$ and so necessarily divides P. That is, any fixed prime divisor is one of the p_i. But

$$p_i \mid g(0) \Longrightarrow p_i^2 \mid \prod f_i(n_0) \Longrightarrow p_i \mid f_j(n_0) \text{ for some } j \neq i,$$

since $p_i \parallel f_i(n_0)$. This implies f_i and f_j have a common root mod p, so that $p_i \mid \mathrm{Res}(f_i, f_j)$, a contradiction. □

Proof of Theorem 1.2. We shall show that if f and f_i are as in Lemma 4.2, then one of the f_i is a Euclidean polynomial of type II for the progression $a \bmod m$. Renumbering if necessary, we assume that the constant term of f_i is congruent $a \pmod m$ precisely for $1 \le i \le r$.

We first show that $r \ge 1$, i.e., that some f_i has constant term congruent to $a \pmod m$. Supposing the opposite we apply Hypothesis H to obtain arbitrarily large values of n for which $f_i(n)$ is prime for each $1 \le i \le k$ (the conditions of Hypothesis H are implied by conditions (i) and (iv) of Lemma 4.2). Then, recalling (v) of the same lemma,

$$f_i(n) \equiv f_i(0) \not\equiv a \pmod m \qquad \text{for } 1 \le i \le k,$$

and so $f(n) = \prod_{i=1}^k f_i(n)$ has no prime factors from the progression $a \bmod m$ for these values of n. But this contradicts that f is a Euclidean polynomial of type I for $a \bmod m$.

Now suppose $1 \le i \le r$, and that f_i is not a Euclidean polynomial of type II for the progression $a \bmod m$. Then either property (i) in Definition 2 fails or (i) holds but (ii) fails, i.e., either

- (i′) f_i has infinitely many prime divisors outside the progressions $1 \bmod m$ and $a \bmod m$, or
- (ii′) $a \not\equiv 1 \pmod m$, and all but finitely many of f_i's prime divisors belong to the residue class $1 \bmod m$.

But (ii′) is easily seen to be impossible. For in this case let E be the product of those prime divisors of f_i that do not belong to the residue class $1 \bmod m$. Since f has no fixed prime divisor, neither does f_i, so again the Chinese Remainder Theorem allows us to find arbitrarily large n with $f_i(n)$ prime to E. Since

$$f_i(n) \equiv f_i(0) \equiv a \pmod m,$$

for large enough n of this type we get a positive integer congruent to $a \pmod m$ all of whose prime factors are from the progression $1 \bmod m$, a contradiction.

Hence if none of $f_1, \dots, f_r$ is a Euclidean polynomial of type II for $a \bmod m$, then for each $1 \le i \le r$ the polynomial f_i must have infinitely many prime divisors outside the progressions $1 \bmod m$ and $a \bmod m$. Choose such a prime divisor p_i for each $1 \le i \le r$ in such a way that Lemma 4.3 can be applied (e.g., it suffices to take choose the p_i all sufficiently large). Then we obtain arbitrarily large n for which

$$f_1(n) = p_1 q_1, \quad f_2(n) = p_2 q_2, \quad \dots, \quad f_r(n) = p_r q_r,$$
$$f_{r+1}(n) = q_{r+1}, \quad f_{r+2}(n) = q_{r+2}, \quad \dots, \quad f_k(n) = q_k,$$

where all the q_i are prime. Since $f = \prod f_i$ is a Euclidean polynomial of type I, for sufficiently large n of this type the list

$$p_1, p_2, \dots, p_r, q_1, \dots, q_r, q_{r+1}, \dots, q_k$$

must contain a prime congruent to $a \pmod m$.

But this is not possible: The p_i were chosen outside the progression $a \bmod m$, the q_j for $j > r$ satisfy

$$q_j = f_j(n) \equiv f_j(0) \not\equiv a \pmod m,$$

and for $1 \le j \le r$,

$$p_j q_j \equiv f_j(n) \equiv f_j(0) \equiv a \pmod m,$$

so that

$$q_j \equiv a p_j^{-1} \not\equiv a \pmod m$$

since $p_j \not\equiv 1 \pmod m$. This contradiction completes the proof. □

5. Concluding Remarks

5.1. **Lessons from *History*.** As we have already mentioned, the first volume of Dickson's *History of the Theory of Numbers* chronicles the early attempts to obtain special cases of Dirichlet's theorem. There are two entries that deserve special attention. While nearly all the results are for progressions $a \bmod m$ with $a^2 \equiv 1 \pmod m$, Dickson also reports, more surprisingly, that

> A.S. Bang [gave a proof] for the differences 4, 6, 8, 10, 12, 14, 18, 20, 24, 30, 42, 60,

and

> E. Lucas for $5n + 2$, $8n + 7$.

But the first progression attributed to Lucas, as well as many of the progressions here attributed to Bang, have no Euclidean proofs in our sense (assuming Hypothesis H). How then did these authors proceed?

Bang's entry is a healthy reminder that we ought not equate Euclidean proofs with elementary proofs. His arguments (which may be found in [Ban91] or [Ban37]) are based not on Euclid's approach to prime number theory but on Chebyshev's. Forty years later, similar proofs would be given by Erdős [Erd35] and Ricci ([Ric33], [Ric34]), both of whom were apparently unaware of Bang's work.

Lucas's argument [Luc78, p. 309] is intriguing but erroneous. Let $L_0 = 2$, $L_1 = 1$ and $L_{n+2} = L_{n+1} + L_n$ for $n = 0, 1, 2, \dots$; these "Lucas numbers" satisfy many well-known identities, in particular

$$L_{2n} = L_n^2 - 2(-1)^n \quad \text{and} \quad L_n^2 - 5F_n^2 = 4(-1)^n,$$

where F_n is the nth Fibonacci number, indexed so that $F_0 = 0$. The first identity implies that $L_{2^k} \equiv 2 \pmod 5$ for $k \geq 2$, while the second implies that $\left(\frac{-5}{p}\right) = 1$ for each prime divisor p of L_{2^k}, whence $p \equiv 1, 3, 7$ or $9 \pmod{20}$. From these two facts we are supposed to conclude that each L_{2^k}, with $k \geq 2$, has a prime divisor $2 \pmod 5$, but in fact

$$L_{2^7} = 5628827661246116195137236 47 = 119809 \cdot 4698167634523379875583.$$

5.2. **The Arguments of Lenstra and Coppersmith.** We now return to the example of §1.4 and the problem of establishing unconditionally that $n^2 + 1$ is free of prime factors $2 \bmod 5$ infinitely often. We might expect problems of this kind to be difficult because elementary sieve methods easily establish that for almost all n (that is, all n outside a set of density 0) the integer $n^2 + 1$ does have a prime factor congruent to $2 \pmod 5$. But when the author proposed this problem to `sci.math.research`, H. W. Lenstra soon responded with the following nearly one-line demonstration: there is certainly one n for which $n^2 + 1$ is free of prime factors from the progression $2 \bmod 5$, namely $n = 2$. But then $(n^5)^2 + 1$ is another, since the quotient

$$\frac{n^{10} + 1}{n^2 + 1} = \Phi_{10}(n^2) \qquad \text{has only prime divisors 0 or 1} \pmod 5.$$

(Here we use a more exact characterization of the prime divisors of Φ_m than that obtained in the first example of §3; see for example [Nag64, Theorem 94].) Alternatively, Don Coppersmith notes that by quadratic reciprocity, it suffices that infinitely often $n^2 + 1$ admit a primitive representation in the form $x^2 - 5y^2$ with x and y of opposite parity. He completes his proof by observing that

$$(10k^2)^2 + 1 = 100k^4 + 1 = (10k^2 + 1)^2 - 5(2k)^2.$$

Both methods are capable of generalization. For example, Coppersmith's method shows that $n^2 + 2$ is free of prime factors $2 \pmod 5$ whenever the positive integer n corresponds to a solution of the generalized Pell equation $n^2 - 3m^2 = -242$. To illustrate, taking a reasonably large solution of this Pell equation we find

$$760577608550439702331^2 + 2 =$$
$$3 \cdot 2539 \cdot 316981018521 \cdot 29579890746624425993228 9011.$$

Again, almost always $n^2 + 2$ has a prime divisor congruent to $2 \pmod 5$, so it is interesting that in the above cases we have methods for producing explicit atypical examples.

Appendix: Hypothesis H and Murty's Impossibility Theorem

In keeping with the spirit of this article, we take this opportunity to describe how Murty's Impossibility Theorem itself can be proved from Hypothesis H. Since an unconditional proof exists, our motivation here is purely methodological.

Suppose f is an irreducible polynomial with integer coefficients. Given a positive integer m, let $S(f, m)$ denote that subset of $\mathbf{Z}/m\mathbf{Z}$ consisting of those residue classes which contain infinitely many prime divisors of $f(T)$. Clearly $S(f, m)$ is a subset of $(\mathbf{Z}/m\mathbf{Z})^\times$. Actually, more is true. The next theorem appears explicitly in Conrad's discussion [Con] of Murty's ideas:

Theorem 5.1. *$S(f, m)$ is a subgroup of $(\mathbf{Z}/m\mathbf{Z})^\times$.*

In fact, if K is obtained by adjoining a root of f to $\mathbf{Q}$, then the Chebotarev density theorem implies that $S(f,m)$ is exactly the image of $\mathrm{Gal}(K(\zeta_m)/K)$ under the restriction map to $\mathrm{Gal}(\mathbf{Q}(\zeta_m)/\mathbf{Q})$.

The impossibility theorem is now an easy consequence: If the residue class a mod m has a Euclidean polynomial of type II, then it has an irreducible Euclidean polynomial of type II. Calling this polynomial f, Theorem 5.1 implies that $S(f,m)$ is a subgroup of $(\mathbf{Z}/m\mathbf{Z})^\times$ satisfying $\{a \bmod m\} \subset S(f,m) \subset \{a \bmod m, 1 \bmod m\}$. Now $a^2 \equiv 1 \pmod m$ follows easily.

Here we give a conditional proof of Theorem 5.1, relying on Hypothesis H instead of Chebotarev density. It is enough to establish the following lemma:

Lemma 5.2. *Assume Hypothesis H. Let f be a nonconstant irreducible polynomial with integer coefficients and m a positive integer. Suppose that $S(f,m)$ contains the residue classes $a_1 \bmod m, \ldots, a_k \bmod m$. Then $S(f,m)$ also contains the residue class $a_1^{-1} \cdots a_k^{-1} \bmod m$.*

Remark. As will be seen in the proof, we only need Hypothesis H for single-polynomial families.

Proof. We can assume that f is irreducible and (using Lemma 3.1) that f has no fixed prime divisor. Replacing f with $f(T+n_0)$ for a suitable n_0, we can assume additionally that the constant term of f is prime to m. Replacing f with $f(mT)$, we can also assume that f is constant modulo m, i.e., that the reduction $\overline{f}$ of f in $(\mathbf{Z}/m\mathbf{Z})[T]$ actually belongs to $\mathbf{Z}/m\mathbf{Z}$. Let $a_0 \bmod m$ be this reduction. Since f assumes infinitely many prime values by Hypothesis H, it must be that $a_0 \bmod m$ is an element of $S(f,m)$.

Now choose prime divisors $p_0, p_1, \ldots, p_k$ of f with each $p_i \equiv a_i \bmod m$ and no p_i dividing $\mathrm{SylDisc}(f)$. For each $i = 1, \ldots, k$, we can choose r_i so that $p_i \parallel f(r_i)$. Then choosing r to satisfy $r \equiv r_i \pmod{p_i^2}$ for each i, we have $D \parallel f(r)$. Define a new polynomial $g(T)$ by the equation

$$f(p_0 \ldots p_k T + r) = p_0 \cdots p_k g(T).$$

Then $g(T)$ has integer coefficients, is irreducible over $\mathbf{Q}$ and has no fixed prime divisors. To see the last of these, note that since f has no fixed prime divisors, the only possibilities for fixed prime divisors of g are p_0, ..., p_k, but none of these divide $g(0) = f(r)/(p_0 \ldots p_k)$. So by Hypothesis H, there are arbitrarily large n for which $g(n)$ is prime. Since

$$g(n) \equiv a_0 a_0^{-1} a_1^{-1} \cdots a_k^{-1} \equiv a_1^{-1} \cdots a_k^{-1} \bmod m,$$

it follows that $a_1^{-1} \cdots a_k^{-1} \bmod m$ belongs to $S(f,m)$, as we sought to show. □

Acknowledgements

I would like to thank Noah Snyder for initially suggesting the notion of a Euclidean proof examined in this paper, Keith Conrad for spot-on comments on an early draft, my advisor Carl Pomerance for helpful suggestions and constant encouragement, and H. W. Lenstra and Don Coppersmith for permission to include their ingenious arguments.

References

[Ban91] A. S. Bang, *Om Primtal af bestemte Former*, Nyt Tidsskrift for matematik, B (advanced) **2** (1891), 73–82.

[Ban37] ———, *Elementære Beviser for specielle Tilfælde af Dirichlets Sætning om Differensrækker*, H. Chr. Bakkes Boghandel, Københavns, 1937.

[Bau06] M. Bauer, *Über die arithmetische Reihe*, J. Reine Angew. Math. **131** (1906), 265–267.

[BL65] P. T. Bateman and M. E. Low, *Prime numbers in arithmetic progressions with difference* 24, Amer. Math. Monthly **72** (1965), 139–143. MR MR0173649 (30 #3859)

[Con] Keith Conrad, *Euclidean proofs of dirichlet's theorem*, Unpublished expository note.

[Dic66] L. E. Dickson, *History of the theory of numbers. Vol. I: Divisibility and primality.*, Chelsea Publishing Co., New York, 1966. MR 39 #6807a

[Dir37] P. G. L. Dirichlet, *Beweis der Satz, dass jede unbegrentze arithmetische progression, deren erstes Glied und Differenz ganze Zahlen ohne gemeinschaftliche Factor sind, unendliche viele Primzahlen enthält.*, Abh. der Königlichen Preuss. Akad. der Wiss. (1837), 45–81.

[Erd35] P. Erdős, *Über die Primzahlen gewisser arithmetischer Reihen*, Math Z. **39** (1935), 473–491.

[Kro88] Leopold Kronecker, *Über die arithmetischen Sätze, welche Lejeune Dirichlet in seiner Breslauer Habilitationsschrift entwickelt hat*, Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin **16** (1888), 417–423.

[Leb56] V. A. Lebesgue, *Remarques diverses sur les nombres premiers*, Nouv. Ann. Math. **15** (1856), 130–134, 236–239.

[Luc78] E. Lucas, *Theorie des fonctions numeriques simplement periodiques*, Amer. J. Math. **1** (1878), 289–321.

[MT06] M. R. Murty and N. Thain, *Primes in certain arithmetic progressions*, Functiones et Approximatio **35** (2006), 7–17, to appear.

[Mur88] M. R. Murty, *Primes in certain arithmetic progressions*, Journal of the Madras University (1988), 161–169.

[Nag64] Trygve Nagell, *Introduction to number theory*, Second edition, Chelsea Publishing Co., New York, 1964. MR 30 #4714

[Ric33] G. Ricci, *Sul teorema di Dirichlet relativo alla progressione aritmetica*, Boll. Un. Mat. Ital. **12** (1933), 304–309.

[Ric34] ———, *Sui teoremi di Dirichlet e di Bertrand-Tchebychef relativi alla progressione aritmetica*, Boll. Un. Mat. Ital. **13** (1934), 7–17.

[Sch12] I. Schur, *Über die Existenz unendlich vieler Primzahlen in einigen speziellen arithmetischen Progressionen*, Sitzungsber. Berl. Math. Ges. **11** (1912), 40–50.

[Sch62] A. Schinzel, *Remarks on the paper "Sur certaines hypothèses concernant les nombres premiers"*, Acta Arith. **7** (1961/1962), 1–8. MR MR0130203 (24 #A70)

[Sel49] A. Selberg, *An elementary proof of Dirichlet's theorem about primes in an arithmetic progression*, Ann. of Math. (2) **50** (1949), 297–304. MR MR0029409 (10,595a)

[Sha50a] H. N. Shapiro, *On primes in arithmetic progressions. I*, Ann. of Math. (2) **52** (1950), 217–230. MR MR0036261 (12,81a)

[Sha50b] ———, *On primes in arithmetic progression. II*, Ann. of Math. (2) **52** (1950), 231–243. MR 12,81b

[SS58] A. Schinzel and W. Sierpiński, *Sur certaines hypothèses concernant les nombres premiers*, Acta Arith. 4 (1958), 185–208; erratum **5** (1958), 259 (French). MR 21 #4936

[Wen95] E. Wendt, *Elementarer Beweis des Satzes, dass in jeder unbegrenzten arithmetischen Progression $my+1$ unendlich viele Primzahlen vorkommen*, J. Reine Angew. Math. **115** (1895), 85–88.

[Zas49] H. Zassenhaus, *Über die Existenz von Primzahlen in arithmetischen Progressionen*, Comment. Math. Helv. **22** (1949), 232–259. MR MR0029408 (10,594f)

Mathematics Department, Dartmouth College, Hanover, New Hampshire 03755
E-mail address: paul.pollack@dartmouth.edu