Polynomial Congruences and Density

GERRY MYERSON
Macquarie University
NSW 2109 Australia
gerry@math.mq.edu.au

What do the solutions of a congruence look like, as the modulus varies? Let $f(t)$ be a polynomial with integer coefficients, let the solutions of $f(t) \equiv 0 \pmod{m}$, if there are any, be $r_1, r_2, \ldots, r_k$, with $0 \le r_j \le m-1$; how are the numbers $r_1, r_2, \ldots, r_k$ distributed within the set $\{0, 1, \ldots, m-1\}$? For a fixed f, how does the answer change as m increases?

In order to compare answers for different values of the modulus, it is convenient to divide through by the modulus. Thus, we'll define

$$S_f(m) = \{r/m : 0 \le r \le m-1, \gcd(r, m) = 1, \text{ and } m \text{ divides } f(r)\}$$

This makes $S_f(m)$ a (finite, possibly empty) subset of $[0, 1)$ for each m, so the sets for various m are directly comparable.

We'll also let S_f be the union of the sets $S_f(m)$ over all positive integers m. It is this set that will concern us.

Polynomials of degree zero are supremely boring in this context, so let's start by looking at polynomials of degree one. Let $f(t) = at + b$. If $at + b \equiv 0 \pmod{m}$, then $at + b = mv$ for some integer v, so

$$\frac{t}{m} = \frac{v}{a} - \frac{b}{am} = \frac{v}{a} + O(m^{-1}).$$

Thus, most of the points in S_f are very close to one or another of the points $0, 1/a, 2/a, \ldots, 1$. Moreover, if v is relatively prime to a then there will be infinitely many m such that $mv \equiv b \pmod{a}$, and for such m there will be a point in $S_f(m)$ close to v/a. Thus, we have a pretty good idea of what S_f looks like; very crowded near the points v/a with v relatively prime to a (and perhaps also near the other points of the form v/a—we encourage the reader to look at this more closely), very sparse everywhere else.

Matters get considerably more complicated when we go to quadratics and polynomials of yet higher degree. Hooley [**1**] proved that if f is an irreducible polynomial of degree at least 2, with integer coefficients, then the sequence formed by ordering S_f by increasing denominator is uniformly distributed in $[0, 1)$. What does that mean?

Consider, for example, the polynomial $f(t) = t^2 + 1$. The relevant sequence is then 0/1, 1/2, 2/5, 3/5, 3/10, 7/10, 5/13, 8/13, 4/17, 13/17, 7/25, 18/25, 5/26, 21/26, etc. In fact, it makes no difference for our purposes how fractions with the same denominator are ordered, so long as the denominators are in nondecreasing order.

A sequence $u_1, u_2, \ldots$ of numbers in $[0, 1)$ is said to be uniformly distributed in $[0, 1)$ if in the limit each subinterval I of $[0, 1)$ contains a proportion of terms of the sequence equal to the length of I. That is to say, $u_1, u_2, \ldots$ is uniformly distributed in $[0, 1)$ means $\lim_{n\to\infty} \#\{i \le n : a \le u_i < b\}/n = b - a$ for all a and b with $0 \le a < b \le 1$. Chapter 21 of Roberts [**2**] provides a gentle introduction to the theory of uniformly distributed sequences.

Hooley's proof uses tools at a level beyond that suitable for this MAGAZINE. The main purpose of this paper is to prove a weaker, but still interesting, result, using only readily accessible methods.

THEOREM. *Let $f(t) = t^e g(t)$ where e is a nonnegative integer, g is a polynomial of degree at least 2 with integer coefficients, and $g(0) \neq 0$. Define T_f by*

$$T_f = \{r/m : \gcd(r, m) = 1, \text{ and } m \text{ divides } f(r)\}.$$

Then T_f is dense in the reals.

It is easy to see that if a sequence is uniformly distributed in $[0, 1)$ then the set underlying the sequence is dense in $[0, 1)$. Our theorem is thus an immediate consequence of Hooley's, provided only that f have an irreducible factor of degree at least two. We will prove it without this requirement, and without reference to Hooley's result.

We note that the theorem is best possible, in the sense that if f is a polynomial with integer coefficients but doesn't satisfy the hypothesis of the theorem then it is easy to see from the introductory remarks on first degree polynomials that T_f is not dense in the reals.

Our proof proceeds along the following lines. Given f satisfying the hypotheses, and given a real number x, we choose rational numbers h_1/k_1 and h_2/k_2 close to x. Then $(h_1a + h_2b)/(k_1a + k_2b)$ is also close to x for all positive integers a and b. We show how to choose a and b in such a way that if $r = h_1a + h_2b$ and $m = k_1a + k_2b$ then m divides $f(r)$.

First we show that a "weighted mediant" of two fractions lies between the fractions and is in lowest terms:

LEMMA 1. *If h_1, h_2, k_1, k_2, a and b are positive integers and $h_1/k_1 < h_2/k_2$ then*

$$\frac{h_1}{k_1} < \frac{h_1a + h_2b}{k_1a + k_2b} < \frac{h_2}{k_2}.$$

Moreover, if $h_1k_2 - h_2k_1 = -1$ and a and b are relatively prime then $h_1a + h_2b$ and $k_1a + k_2b$ are relatively prime.

Proof. The inequalities are evident on viewing the weighted mediant as a weighted average of the two fractions h_1/k_1 and h_2/k_2 with positive weights k_1a and k_2b:

$$\frac{h_1a + h_2b}{k_1a + k_2b} = \frac{(k_1a)(h_1/k_1) + (k_2b)(h_2/k_2)}{k_1a + k_2b}$$

If $h_1k_2 - h_2k_1 = -1$, let $r = h_1a + h_2b$ and $m = k_1a + k_2b$ and solve for a and b; $a = h_2m - k_2r$, $b = k_1r - h_1m$. Now any common divisor of r and m divides both a and b, and the last assertion of the lemma follows. ∎

Next we show that good approximations to x can be chosen to satisfy certain divisibility and coprimality conditions.

LEMMA 2. *Given positive reals x and ϵ, a positive integer c, a nonzero integer c', and an integer $n \geq 2$, there are positive integers h_1, h_2, k_1, k_2 and d such that d is relatively prime to c', $k_2 = cd^{n-1}$, $h_1k_2 - h_2k_1 = -1$, and $|x - h_i/k_i| < \epsilon$ for $i = 1, 2$.*

Proof. Given any positive integer D, there is a positive integer H such that $|x - H/D^{n-1}| \leq 1/(2D^{n-1})$. Let $d = D^2$, let $h_2 = cD^{n-1}H + 1$, and let $k_2 = cd^{n-1} = cD^{2n-2}$; then

$$|x - h_2/k_2| \leq |x - H/D^{n-1}| + 1/(cD^{2n-2}) \leq 1/(2D^{n-1}) + 1/(cD^{2n-2}).$$

Moreover, $\gcd(h_2, k_2) = 1$, so there are positive integers h_1 and k_1 such that $h_1k_2 - h_2k_1 = -1$, which entails that

$$|x - h_1/k_1| \le |x - h_2/k_2| + h_2/k_2 - h_1/k_1 = |x - h_2/k_2| + 1/(k_1k_2)$$
$$\le 1/(2D^{n-1}) + 2/(cD^{2n-2}).$$

Now it suffices to choose D relatively prime to c' and large enough to ensure $1/(2D^{n-1}) + 2/(cD^{2n-2}) < \epsilon$. ∎

Finally, we need a version of the Remainder Theorem, which we present without proof.

LEMMA 3. *Given a polynomial $f(t)$ of degree n and real numbers a and b, $a \neq 0$, there is a polynomial $q(t)$ such that*

$$a^n f(t) = (at + b)q(t) + a^n f(-b/a).$$

Moreover, if a, b, and the coefficients of f are integers, then so are the remainder and the coefficients of q.

Proof of the Theorem. Let $f(t) = t^e g(t) = c_0t^n + c_1t^{n-1} + \cdots + c_n$ satisfy the hypotheses. Note that if T_g is dense then so is T_f, so we may assume without loss of generality that $e = 0$. Thus, $c_n \neq 0$. We may assume that c_0 is positive, since we may replace f with $-f$, if need be. If m divides $f(r)$ then also m divides $f(r + mQ)$ for any integer Q, so it is enough to prove that T_f is dense in the positive reals (indeed, in $[0, 1)$).

Let x and ϵ be positive. By Lemma 2, there are positive integers h_1, h_2, k_1, k_2, and d such that $\gcd(d, c_n) = 1$, $k_2 = c_0d^{n-1}$, $h_1k_2 - h_2k_1 = -1$, and $|x - h_i/k_i| < \epsilon$ for $i = 1, 2$. Note that $\gcd(k_1, k_2) = 1$, whence $\gcd(d, k_1) = 1$.

Recall that we want to choose positive integers a and b in such a way that if $r = h_1a + h_2b$ and $m = k_1a + k_2b$ then m divides $f(r)$. We claim this can be achieved by letting $b = d + k_1s$, where s is any multiple of c_n large enough to guarantee $k_1^n f(b/k_1) > k_2b$, and then defining a by $k_1^n f(b/k_1) = k_1a + k_2b$.

Clearly, a is positive. Moreover, a is an integer, because

$$a = c_1b^{n-1} + c_2b^{n-2}k_1 + \cdots + c_nk_1^{n-1} + b(c_0b^{n-1} - k_2)/k_1 \tag{1}$$

and

$$(c_0b^{n-1} - k_2)/k_1 = c_0(b^{n-1} - d^{n-1})/k_1 = c_0s(b^{n-1} - d^{n-1})/(b - d).$$

From $b = d + k_1s$ we deduce

$$\gcd(b, k_1) = \gcd(d, k_1) = 1. \tag{2}$$

We claim that a and b are relatively prime. For let p be a prime dividing both a and b. Then by (1) p divides c_n or k_1. By (2), p divides c_n. We chose s to be a multiple of c_n, so p divides s. We defined b by $b = d + k_1s$, so p divides d. But now we have reached a contradiction, as d was chosen relatively prime to c_n.

Let $r(t) = h_1t + h_2b$ and $m(t) = k_1t + k_2b$. By Lemma 1, $r(a)$ and $m(a)$ are relatively prime, and $|x - r(a)/m(a)| < \epsilon$. All that remains is to prove $m(a)$ divides $f(r(a))$.

By Lemma 3, there is a polynomial $q(t)$ with integer coefficients such that $k_1^n f\big(r(t)\big) = m(t)q(t) + k_1^n f\big(r(-k_2b/k_1)\big)$. Thus,

$$k_1^n f\big(r(t)\big) - m(t)q(t) = k_1^n f\big(r(-k_2b/k_1)\big) = k_1^n f\big(h_1(-k_2b/k_1) + h_2b\big)$$
$$= k_1^n f\big((-h_1k_2 + h_2k_1)(b/k_1)\big) = k_1^n f(b/k_1)$$
$$= k_1a + k_2b = m(a).$$

Evaluating at $t = a$, we see that $m(a)$ divides $k_1^n f\big(r(a)\big)$. Any common prime divisor of $m(a)$ and k_1^n divides $k_2 b$, but k_1 is relatively prime to b by (2) and also to k_2. Thus $m(a)$ is relatively prime to k_1, so it divides $f\big(r(a)\big)$, and we are done. ■

REFERENCES

1. Christopher Hooley, On the distribution of the roots of polynomial congruences, *Mathematika* **11** (1964) 39–49, MR 29 #1173.
2. Joe Roberts, *Elementary Number Theory*, MIT Press, 1977, MR 58 #16472.

A Curious Way to Test for Primes

DENNIS P. WALSH
Middle Tennessee State University
Murfreesboro, TN, 37132
dwalsh@mtsu.edu

Upon hearing that you are a student of mathematics, a cab driver says to you, "Check this out. The second derivative of e^x is e^x, right? And e^x evaluated at 0 is equal to 1, right? Therefore 2 has got to be a prime number." Your first reaction is a condescending chuckle for the cabby who seems to dabble in mathematics and appears to make a jumble of it. "Well, at least she's logically correct—2 is a prime number," you mumble to yourself with some smugness. But the cab driver hears you and takes a long route to your destination. You end up paying a hefty fare. The cab driver smirks as she drives off.

In fact, the smirking cabby stated a specific case of the theorem we provide below, a theorem which offers an unusual characterization of prime numbers based on differentiation. The cab driver could give another specific case by stating, "The third derivative of $e^x + e^{x^2/2}$ evaluated at $x = 0$ is 1, and thus 3 is a prime number." Offering another example, she could state accurately that, "The fourth derivative of $e^x + e^{x^2/2} + e^{x^3/3}$ evaluated at 0 does not equal 1, and hence 4 is not a prime number." You can probably guess the pattern. We give it below in a theorem with a surprisingly simple proof that uses the series expansion of $e^{x^k/k}$ and the power rule for differentiation.

THEOREM. *For each positive integer $n > 1$, define the function g_n by $g_n(x) = \sum_{k=1}^{n-1} e^{x^k/k}$. A positive integer n is prime if and only if $\frac{d^n}{dx^n} g_n(0) = 1$.*

Proof. Let n be a positive integer greater than 1. Note that $e^{x^k/k}$ has a series expansion given by

$$e^{x^k/k} = \sum_{j=0}^{\infty} \frac{(x^k/k)^j}{j!} = \sum_{j=0}^{\infty} \frac{x^{kj}}{k^j j!}$$

for all real x. Hence, we have

$$g_n(x) = \sum_{k=1}^{n-1} \sum_{j=0}^{\infty} \frac{x^{kj}}{k^j j!},$$