POLYNOMIAL SOLUTIONS OF BINOMIAL CONGRUENCES

H. LINDGREN

(received 3 April 1959, revised 10 March 1960)

Polynomial solutions of a few binomial congruences have been known for a long time. For instance Legendre showed that the congruence

$$(1) \qquad x^2 \equiv a \pmod{2^m;\ a = 8q+1}$$

has a solution

$$x \equiv \pm\{1 + \tfrac{1}{2}(a-1) - \tfrac{1}{8}(a-1)^2 + \cdots\} \pmod{2^{m-1}},$$

this being the expansion of $\sqrt{(1 + a - 1)}$ as far as the term of degree $m - 3$. [1] It seems that only restricted types, e.g. (1), have been investigated.

1. Summary

This paper relates to solutions of the congruence

$$(2) \qquad x^n \equiv a \pmod{M}$$

that are polynomials in a; a has any admissible value that is prime to the modulus, but otherwise there is no restriction on any of the integers in (2). Fundamental is

THEOREM 1. *All congruences* (2) *have polynomial solutions.*

The other theorems relate to the minimum degree of a solution, and to the number of terms. To facilitate their enunciation, some notations used throughout are given here.

(i) The congruence that gets most attention is

$$(3) \qquad x^n \equiv a \pmod{p^m},$$

a being any n-ic residue (not divisible by p).

(ii) d, n_1, w, and r are defined as follows:

$$(4) \qquad (n, p-1) = d, \quad n = dn_1p^w \ (p \nmid n_1), \quad p - 1 = dr.$$

Although n in (3) is unrestricted, we need only consider values such that

$$1 < n < \phi(p^m) = p^{m-1}(p-1),$$

https://doi.org/10.1017/S1446788700025957 Published online by Cambridge University Press

in view of the Fermat-Euler theorem. With this limitation

$$w \leqq m-2 \ (p=2), \qquad w \leqq m-1 \ (p \geqq 3).$$

(iii) E denotes the p-adic valuation:

$$E(p^iM/N)=i, \quad E(M/N)=0, \quad E(M/p^iN)=-i \ (p \nmid MN).$$

Extensive use will be made of the obvious properties

$$E(MN)=E(M)+E(N),$$
$$\text{If } E(M)<E(N), \text{ then } E(M+N)=E(M).$$

(iv) Most other symbols (e.g. k, μ, M) are used with various meanings, being redefined each time the meaning is changed.

THEOREM 2. *Let $D_1, D_2, \cdots$ be the minimum degrees of polynomial solutions of* (3) *when their moduli are the prime powers $M_1, M_2, \cdots$ respectively. Then the minimum degree of a polynomial solution of* (2), *when its modulus is $M_1M_2\cdots$, is*

$$D=\max(D_1, D_2, \cdots).$$

THEOREM 3. *When $p \nmid n$, the minimum degree of a polynomial solution of* (3) *satisfies*

$$D \leqq rs-1,$$

where s is the least integer such that

$$s+E(s!) \geqq m.$$

THEOREM 4. *When $p=2$ and $2|n$, the minimum degree of a polynomial solution of* (3) *is*

$$D=[\tfrac{1}{2}(m-w-1)].$$

As a particular case of Theorem 4, take (1), i.e. $n=2$. Then

$$D=[\tfrac{1}{2}(m-2)].$$

The degree need not be $m-3$, but only about half of this.

THEOREM 5. *When $p \geqq 3$ and $p|n$, the minimum degree of a polynomial solution of* (3) *satisfies*

$$r(m-w-1) \leqq D \leqq r(m-w)-1.$$

THEOREM 8. *If every prime in the modulus of* (2) *satisfies* $(d, r)=1$, *and if* $(n, M)=1$, *then* (2) *has a single-term solution.*

This generalizes the well-known solution $x \equiv a^q$ of

$$x^2 \equiv a \pmod{p, \ = 4q-1}.$$

Another generalization is

THEOREM 9. *When $p \nmid n$,* (3) *has a solution consisting of δ terms, where δ is*

https://doi.org/10.1017/S1446788700025957 Published online by Cambridge University Press

the least divisor of r such that

$$(d, r/\delta) = 1.$$

2. Preliminary

When p divides n the modulus of a solution need only be p^{m-w}. This is because

$$(x + p^{m-w}q)^n = x^n + \cdots + C(n, s)x^{n-s}p^{s(m-w)}q^s + \cdots,$$

where the E of the general term with $s \geqq 1$ is at least

$$\begin{aligned} E\{C(n, s)\} + s(m-w) &= E\{C(n-1, s-1)n/s\} + s(m-w) \\ &\geqq 0 + w - E(s) + s(m-w) \\ &= m + (s-1)(m-w) - E(s) \\ &\geqq m + s - 1 - E(s) \geqq m \quad (p \geqq 2). \end{aligned}$$

Thus

$$(x + p^{m-w}q)^n \equiv x^n \pmod{p^m},$$

justifying the assertion. It can also be seen that when $w = 0$ and p does not divide q, the E of the term with $s = 1$ is m and the E of every term with $s > 1$ is greater than m. Therefore, writing x_h for $x + p^m q$ and x_k for x, we have

$$(5) \qquad E(x_h^n - x_k^n) = E(x_h - x_k),$$

provided the latter is greater than 0.

When $p \geqq 3$, the number of n-ic residues mod p^m in a complete set is

$$(6) \qquad R = \phi(p^m)/dp^w = p^{m-w-1}r,$$

and each residue has dp^w roots mod p^m. [2] Since the modulus of a solution need only be p^{m-w}, each residue has d roots mod p^{m-w}. In particular, making $m = w + 1$ shows that there are d roots mod p to each of r n-ic residues mod p^{w+1}.

Divided differences will play a prominent part. The notations

$$\delta^i(x/a)_j, \quad \delta^i_j(x/a)$$

will be used according to convenience for the ith forward divided difference of x_j with respect to a_j in the sequences

$$x_0, x_1, \cdots, x_j, \cdots,$$
$$a_0, a_1, \cdots, a_j, \cdots,$$

and the variables may be omitted where no confusion arises. These differences will be used in the well-known interpolation formula similar to the congruence (19) below.

As is usual, a fraction M/N in a congruence denotes the solution of

https://doi.org/10.1017/S1446788700025957 Published online by Cambridge University Press

$Nx \equiv M$. Normally the denominator is prime to the modulus, but in this paper it is convenient to use fractions such as M/Np^{μ} (mod p^m; $p \nmid MN$). The reason is that they can occur as coefficients in a polynomial solution; for instance they occur in Legendre's solution of (1). Such a fraction can always be brought to a standard form M'/p^{μ}, for M' is the solution of

$$Nx \equiv M \pmod{p^{m+\mu}}.$$

It is also convenient to use a fractional modulus. The interpretation of a congruence in these cases is quite orthodox, namely that the difference between the members (any fractions being in standard form) is an integral multiple of the modulus.

Theorem 1 is proved in three stages:

1A. (3) has polynomial solutions when $p \nmid n$, and
1B. when $p | n$.
1C. (2) has polynomial solutions.

3. Proof of Theorem 1A

The formula mentioned in § 2 gives x in terms of x^n for any R values of x:

$$(7) \qquad x = x_0 + \delta(x/x^n)_0(x^n - x_0^n) + \delta^2(x/x^n)_0(x^n - x_0^n)(x^n - x_1^n) + \cdots \\ + \delta^{R-1}(x/x^n)_0(x^n - x_0^n) \cdots (x^n - x_{R-2}^n).$$

If $x_0^n, \cdots, x_{R-1}^n$ is a complete set of n-ic residues mod p^m and the formula is turned into a congruence mod p^m, it will become a polynomial solution of (3), provided that the divided differences are such that x^n can be replaced by any a congruent to it. This will certainly be so if

$$(8) \qquad E\{\delta^i(x/x^n)_0\} \geqq 0.$$

It will be proved that, when p does not divide n, the x_j can be chosen so that (8) is true.

The differences can be written in the form

$$\delta^i = \frac{P}{\Pi(x_h^n - x_k^n)} \qquad (h, k = 0, \cdots, i;\ k < h),$$

where P is a polynomial in $x_0, \cdots, x_i$ with integral coefficients, because every algebraic expression we divide by in forming δ^i is in the denominator. Since δ^i is a symmetric function of $x_0, \cdots, x_i$, and its denominator is the product of a symmetric polynomial and the simple alternant $\Pi(x_h - x_k)$, its numerator too has this form. Therefore

$$(9) \qquad \delta^i = Q\Pi\left\{\frac{x_h - x_k}{x_h^n - x_k^n}\right\} \qquad (h, k \text{ as before}),$$

where Q is a polynomial with integral coefficients.

https://doi.org/10.1017/S1446788700025957 Published online by Cambridge University Press

Choose the roots, if there are more than one per residue, so that

(10) If and only if $p|a_h - a_k$, then $p|x_h - x_k$.

This can be done because the congruences

$$x^n \equiv a_h, \qquad x^n \equiv a_k \pmod{p},$$

being the same if p divides $a_h - a_k$, have the same set of roots, and both x_h and x_k in (10) can be made congruent mod p to the same root in the set; while if p does not divide $a_h - a_k$, the two sets of roots mod p have no common member. From (5) and (9), when the roots satisfy (10), and p does not divide n,

$$E(\delta^i) = E(Q) \geqq 0,$$

proving (8). Hence one can make any set of substitutions

$$x^n - x_h^n \equiv a - a_h \pmod{p^m}$$

in (7). This proves Theorem 1A.

4. Proof of Theorem 1B

When p divides n, the simple argument based on divided differences is not applicable, for their E's may be negative.

Case 1: $p \geqq 3$. Write a in the form

$$a = \frac{a^{kr+1}}{(1 + a^r - 1)^k}, \tag{11}$$

where k is the least positive integer such that

$$kr + 1 \equiv 0 \ (\mathrm{mod}\, n/d, \ = n_1 p^w), \qquad kr + 1 = \nu n/d, \text{ say;}$$

such a k exists because, by (4),

$$(r, n_1 p^w) = 1. \tag{12}$$

Then (3) has a formal solution

$$x \equiv a^{\nu/d}(1 + a^r - 1)^{-k/n} \pmod{p^{m-w}}.$$

Let A denote a polynomial solution of

$$x^d \equiv a \pmod{p^{m-w}};$$

it can be found by virtue of Theorem 1A because p does not divide d. Then $a^{1/d}$ is a formal expression for A,

$$a^{\nu/d} \equiv A^\nu \pmod{p^{m-w}},$$

and the formal solution becomes

$$\begin{aligned} (13) \quad x &\equiv A^\nu(1 + a^r - 1)^{-k/n} \\ &\equiv A^\nu \left\{1 - \frac{k}{n}(a^r - 1) + \frac{k}{n} \cdot \frac{k+n}{2n}(a^r - 1)^2 - \cdots\right\} \pmod{p^{m-w}}, \end{aligned}$$

the expansion terminating at the last term not divisible by p^{m-w}. This will be a polynomial solution, provided that we justify this truncation of a divergent infinite series, and show that all terms from some point on are divisible by p^{m-w}.

Let

$$(1+u)^{-k/n} = 1 + c_1 u + c_2 u^2 + \cdots,$$

where k and n are any integers such that $k \leqq n$ (as in (13), where $k < n/d \leqq n$). Then, when $0 < u \leqq \frac{1}{2}$,

$$|(1+u)^{-k/n} - (1 + c_1 u + \cdots + c_{s-1}u^{s-1})| < |c_s|u^s + |c_{s+1}|u^{s+1} + \cdots$$
$$\leqq u^s + u^{s+1} + \cdots = O(u^s).$$

Therefore

$$1 + c_1 u + \cdots + c_{s-1}u^{s-1} = (1+u)^{-k/n} + O(u^s),$$
$$(1 + c_1 u + \cdots + c_{s-1}u^{s-1})^n = (1+u)^{-k} + O(u^s),$$
$$(1 + c_1 u + \cdots + c_{s-1}u^{s-1})^n (1+u)^k = 1 + O(u^s),$$
$$(1 + c_1 u + \cdots + c_{s-1}u^{s-1})^n (1+u)^k = 1 + \text{terms in } u^s, \cdots, u^{n(s-1)+k}.$$

The last equation holds for an infinite number of values of u, so it is an identity that holds for all. Hence if p^m divides each term on the right containing u, then

$$(1 + c_1 u + \cdots + c_{s-1}u^{s-1})^n (1+u)^k \equiv 1 \pmod{p^m};$$

if p does not divide $1+u$, then

$$(1 + c_1 u + \cdots + c_{s-1}u^{s-1})^n \equiv (1+u)^{-k} \pmod{p^m};$$

and one root is given by

$$(1+u)^{-k/n} \equiv 1 + c_1 u + \cdots + c_{s-1}u^{s-1} \pmod{p^{m-w}}.$$

The use of the divergent series is thus justified.

It remains to be shown that all terms in (13) from some point on are divisible by p^{m-w}. By (4) and the Fermat-Euler theorem,

(14) $$a^r - 1 \equiv x^{nr} - 1 \equiv 0 \pmod{p^{w+1}},$$

since

$$nr = dn_1 p^w r = n_1 p^w (p-1) = n_1 \phi(p^{w+1}).$$

Hence

$$E\{(a^r - 1)/n\} \geqq 1,$$

and the E of the term in (13) containing $(a^r - 1)^s$ is at least $s - E(s!)$.

We now apply the formula

(15) $$E(s!) = (s - \sigma)/(p-1),$$

where σ is the sum of the digits in s, when expressed in the number scale with radix p. [3]

By (15), when $p \geqq 3$,

$$s - E(s!) = s - (s-\sigma)/(p-1) \geqq s - (s-\sigma)/2 = \tfrac{1}{2}(s+\sigma) \geqq \tfrac{1}{2}(s+1).$$

If therefore s is chosen so that

$$\tfrac{1}{2}(s+1) = m - w,$$

then the term in (13) containing $(a^r - 1)^s$ and all terms thereafter are divisible by p^{m-w}, and the earlier terms constitute a polynomial solution of (3). This proves Case 1.

Case 2: $p = 2$. Here (and whenever $r = 1$) it is simpler to write

$$(16)\ x \equiv (1+a-1)^{1/n} \equiv 1 + \frac{1}{n}(a-1) + \frac{1}{n} \cdot \frac{1-n}{2n}(a-1)^2 + \cdots \pmod{2^{m-w}}.$$

The $O(u^s)$ argument to justify the expansion is similar to Case 1, but shorter. Instead of (14) we have

$$(17) \qquad a - 1 \equiv x^n - 1 \equiv 0 \pmod{2^{w+2}},$$

as is seen from the expansion

$$(4q \pm 1)^n = 1 \pm n \cdot 4q + \tfrac{1}{2}n(n-1) \cdot 16q^2 \pm \cdots \equiv 1 \pmod{2^{w+2}}.$$

Hence

$$E\{(a-1)/n\} \geqq 2,$$

and the E of the term in (16) containing $(a-1)^s$ is by (15) at least

$$2s - E(s!) = 2s - (s-\sigma) \geqq s + 1.$$

The conclusion is as before. This proves Case 2, and completes the proof of Theorem 1B.

5. Proof of Theorems 1C and 2

Denote a solution of minimum degree, when its modulus is the prime power M_h $(h = 1, 2, \cdots)$, by

$$x \equiv c_{h0} + c_{h1}a + c_{h2}a^2 + \cdots \pmod{M_h}.$$

There are values of c_k that satisfy the simultaneous congruences

$$c_k \equiv c_{hk} \quad (\text{mod}\ M_h;\ h = 1, 2, \cdots;\ k = 0, 1, \cdots, \max D_h),$$

for the moduli are coprime in pairs. These give a polynomial

$$c_0 + c_1 a + c_2 a^2 + \cdots$$

that is a solution for every prime power in the modulus, and so is a solution of (2). Moreover its degree is $\max D_h$, for k is not greater than this. This proves Theorem 1C, completing the proof of Theorem 1, and Theorem 2.

A congruence for c_k containing a fraction such as b/p^μ (mod p^m) does not cause any difficulty, for one can solve the set of congruences for Lc_k, where

https://doi.org/10.1017/S1446788700025957 Published online by Cambridge University Press

L is the LCM of the denominators of all such fractions, and the work is then all in integers. Thus if

$$c \equiv \tfrac{1}{2} \pmod{64}, \quad \tfrac{4}{5} \pmod{25},$$

then

$$\begin{aligned} 10c &\equiv 5 \pmod{128}, \quad 8 \pmod{125}, \\ 10c &\equiv 133 \pmod{16000}, \\ c &\equiv 133/10 \pmod{1600}. \end{aligned}$$

In proving Theorem 1A, the roots were chosen so as to satisfy (10). The roots given by a solution such as (13) in Theorem 1B, Case 1, also satisfy (10). For if p divides $a_h - a_k$, then p divides $A_h - A_k$ since its coefficients satisfy (8), and by (13)

$$x_h - x_k \equiv A_h^\nu(1 - \cdots) - A_k^\nu(1 - \cdots) \equiv A_h^\nu - A_k^\nu \equiv 0 \pmod{p},$$

that is, p divides $x_h - x_k$. A similar restriction in Theorem 1B, Case 2, evident from (16) and (17), is

$$4|x_h - x_k \quad (p = 2,\ 2|n). \tag{18}$$

The restrictions (10) and (18), on which the proof of Theorem 1 depends, will be complied with henceforth. They are not always necessary, but in complying with them one gets a solution of lower degree. A proof of this is included in the proofs of Theorems 3 and 4.

6. Serial order of residues

Theorem 1 shows that on turning (7) into a congruence with the appropriate modulus and replacing $x^n - x_h^n$ by $a - a_h$, it becomes a solution of (2). It is further desirable to replace the divided differences with respect to x^n by those with respect to a. The solution of (3) will then be the expected

$$x \equiv x_0 + \delta(x/a)_0(a - a_0) + \delta^2(x/a)_0(a - a_0)(a - a_1) + \cdots \pmod{p^{m-w}}. \tag{19}$$

The condition for (19) to be a solution of (3) is seen to be

$$\delta^i(x/a)_0 \equiv \delta^i(x/x^n)_0 \quad (\operatorname{mod} p^\mu;\ \mu = m - w - e_i), \tag{20}$$

where e_i denotes the minimum E (for $a = a_0, a_1, \cdots$, ad inf.) of the cofactor

$$(a - a_0)(a - a_1) \cdots (a - a_{i-1}) \tag{21}$$

of δ^i in (19). On the other hand divisions by terms such as $a_h - a_k$, performed in calculating the divided differences, reduce the modulus to p^ν, say. If $\nu \geqq \mu$, (20) is satisfied. If $\nu < \mu$, then the values of $\delta^i(x/a)_0$, differing by multiples of p^ν, that satisfy the congruence mod p^ν include those that satisfy it mod p^μ. Restricting ourselves to the latter, we can *make* (20) true, whatever the values of $a_0, a_1, \cdots$ in (19). But this fact is useless in practice, for we would still need to calculate the unwieldy divided differences with respect to x^n.

It will be shown that (20) is true (i.e. $\nu = \mu$) if the residues are in any order such that

$$(22)\quad E(a_h - a_k) = \begin{cases} 0 & (r \nmid h - k), \\ E(h-k) + w + 1 & (p=2, 2\nmid n; r|h-k, p \geqq 3, \text{any } n), \\ E(h-k) + w + 2 & (p = 2,\ 2|n). \end{cases}$$

Simply defined orders that satisfy (22), when the modulus is p^m, are numerical order, ascending or descending, and orders with $a_0, a_1, a_2, \cdots$ congruent to $1, g^n, g^{2n}, \cdots$, where g is a primitive root.

When (22) is satisfied, every rth factor in (21) from some point on (not later than the factor $a - a_{r-1}$) is divisible by p^{w+1} (or 2^{w+2}), the $[i/r]$ or $[i/r] + 1$ quotients are such that every pth quotient from some point on (not later than the pth) is divisible by p, every p^2-th quotient by p^2, and so on, and no other factors are divisible by p. Thus we find e_i, which, when p does not divide n, is the right-hand side of

$$(23)\quad E\{(a-a_0)(a-a_1)\cdots(a-a_{i-1})\} \\ \geqq [i/r] + [i/pr] + [i/p^2r] + \cdots = [i/r] + E([i/r]!).$$

The minimum value is attained when $a = a_i$, for then the rth factor from the end is divisible by p^{w+1} (or 2^{w+2}) but not by p^{w+2} (or 2^{w+3}), the pth quotient from the end is divisible by p but not by p^2, and so on. Therefore

$$e_i = E\{(a_i - a_0)(a_i - a_1)\cdots(a_i - a_{i-1})\} \\ = E\{(a_{jr+k} - a_k)(a_{jr+k} - a_{r+k})\cdots(a_{jr+k} - a_{jr-r+k})\},$$

where $jr + k$ is written for i, and so e_i is equal to

$$(24)\quad E\{(a_{jr} - a_0)(a_{jr-r} - a_0)\cdots(a_r - a_0)\} = E\{(x^n_{jr} - x^n_0)\cdots(x^n_r - x^n_0)\}.$$

Now calculating the respective sides of (20) involves division by quantities identical with or equal to the factors on the respective sides of (24), among others; but those in (24) are the only ones that reduce the modulus, and their total E is e_i. Thus the condition (20) for (19) to be true is satisfied automatically, if the residues are ordered as in (22). And if (22) is complied with in respect of each prime in the modulus of (2), then (19), with the appropriate modulus, is a solution of (2).

The restrictions (22) on the order of the residues will be complied with henceforth. They can be combined with (10) and (18) to give

$$(25)\quad E(x_h - x_k) = \begin{cases} 0 & (r \nmid h - k), \\ E(h-k) + 1 & (p = 2, 2 \nmid n; r|h-k, p \geqq 3, \text{any } n), \\ E(h-k) + 2 & (p = 2,\ 2|n). \end{cases}$$

https://doi.org/10.1017/S1446788700025957 Published online by Cambridge University Press

Similarly to (22), (25) is satisfied, when the modulus is p^{m-w}, by roots in the numerical and primitive-root orders.

7. Proof of Theorem 3

In view of (8) the terms in (19) containing δ^{rs} and all terms thereafter can be omitted if

$$E\{(a-a_0)(a-a_1)\cdots(a-a_{rs-1})\} \geqq m.$$

By (23) with $i = rs$, this condition becomes

$$s + E(s!) \geqq m,$$

so if s is the least integer satisfying this inequality, the degree is at most $rs-1$. This proves Theorem 3, the roots chosen satisfying (10); it will now be shown that the result is only weakened by disregarding (10).

It will be proved that if compliance with (10) gives a solution whose degree is lower than $r(m-w)$, then the degree of any solution constructed from roots not satisfying (10) is at least $r(m-w)$. The w is included in preparation for Theorem 5. We shall use the formula [4]

$$(26) \qquad \delta^i(y/a)_j = \sum \frac{y_h}{\Pi(a_h-a_k)} \qquad (h, k = j, \cdots, j+i;\ k \neq h).$$

The condition

$$(27) \qquad E\{\delta^{\mu r}(x/x^n)_0(a-a_0)\cdots(a-a_{\mu r-1})\} \geqq \mu \qquad (\mu = m-w)$$

is satisfied irrespective of the values of $a, a_0, \cdots$, if and only if the degree of the solution is less than μr. It is known (when $w=0$) and assumed (when $w>0$) that (27) is satisfied if (10) is complied with.

The n-ic residues mod p^m, of which by (6) there are $p^{\mu-1}r$, are distributed among r residue classes mod p^{w+1}, each containing $p^{\mu-1}$ members (and only $\mu \geqq 2$ need be considered). The corresponding roots for each class may have any of d residues mod p, but however one chooses them, there are for each class at least

$$[p^{\mu-1}/d] + 1 \geqq p^{\mu-2} + 1 \geqq \mu$$

roots that have the same residue mod p; so there are always at least μr roots that satisfy (10). They will be denoted by $x_0, \cdots, x_{\mu r-1}$, while $x_{\mu r}$ and ξ will denote roots of a further residue $a_{\mu r}$ such that

$$\text{if } p|a_{\mu r}-a_k, \text{ then } p|x_{\mu r}-x_k,\ p \nmid \xi - x_k,$$

where k is any integer from 0 to $\mu r-1$. This means that

$$\xi = x_{\mu r} + \rho,$$

where ρ is not divisible by p.

https://doi.org/10.1017/S1446788700025957 Published online by Cambridge University Press

The μrth divided difference of x_0 with respect to x^n_0 in the sequences

$$\begin{gathered} x_0, \cdots, x_{\mu r-1}, \xi, \\ x^n_0, \cdots, x^n_{\mu r-1}, x^n_{\mu r}, \end{gathered}$$

which will be distinguished by a bar, is

$$\bar{\delta}^{\mu r}(x/x^n)_0 = \delta^{\mu r}(x/x^n)_0 + \rho/\Pi(x^n_{\mu r} - x^n_k) \quad (k = 0, \cdots, \mu r - 1).$$

The first term on the right is the same as in (27), and the second term is obtained from (26) with $y_{\mu r} = \rho$ and $y_h = 0$ otherwise. Multiplying this equation by the denominator under ρ and replacing $x^n_{\mu r} - x^n_k$ by $a_{\mu r} - a_k$, we get a term whose E is by (27) at least μ, plus ρ whose E is 0. Therefore

$$E\{\bar{\delta}^{\mu r}(x/x^n)_0 (a_{\mu r} - a_0) \cdots (a_{\mu r} - a_{\mu r-1})\} = 0 < \mu,$$

which shows that if (10) is not complied with, the degree of the solution is at least μr, and so is greater than $rs - 1$.

That (10) is not necessary will be shown by finding a solution of

$$x^2 \equiv a = 3q + 1 \pmod{27},$$

based on the relaxed condition

$$\text{If } p^{w+2} | a_h - a_k, \text{ then } p | x_h - x_k.$$

(A similar relaxation of (18) is

$$\text{If } 2^{w+3} | a_h - a_k, \text{ then } 4 | x_h - x_k.)$$

The condition is complied with by making

$$\begin{aligned} x &\equiv 1 \pmod 3 \text{ if } a \equiv 1 \text{ or } 4 \pmod 9, \\ x &\equiv 2 \pmod 3 \text{ if } a \equiv 7 \pmod 9. \end{aligned}$$

A complete set of such roots ($q = 0, \cdots, 8$) is

$$x \equiv 1, 25, 14, 19, 16, 23, 10, 7, 5 \ (\text{mod } 27 \text{ pro tem.}).$$

Instead of (19), Newton's formula

$$(28) \quad x \equiv x_0 + q\Delta x_0 + \tfrac{1}{2}q(q-1)\Delta^2 x_0 + \cdots + \{q(q-1) \cdots (q-D+1)/D!\}\Delta^D x_0$$

can be used, because here $r = 1$. The first six differences of x_0 are 24, 19, 24, 6, 9, 9, and all 7th differences are zero. Substituting the differences in (28) and simplifying, we get

$$x \equiv 1 - 3q + 6q^2 + 4q^3 + 9q^4 + 9q^5 - q^6.$$

In contrast, a solution whose roots 1, 25, 13, 19, $\cdots$ comply with (10) should by Theorem 3 be of degree 2 at most, and in fact

$$\Delta x_0 \equiv 24, \qquad \Delta^2 x_0 \equiv 18, \qquad \Delta^3 x \equiv 0,$$

whence

$$x \equiv 1 + 24q + 9q(q - 1) \equiv 1 - 12q + 9q^2.$$

8. Theorem 4. Preliminary

When $p = 2$ and n is even, we can, in view of (17), make the substitution

$$a - 1 = 2^{w+2}q \qquad (q = 0, 1, \cdots)$$

in (16), getting

$$(29) \qquad x \equiv 1 + \frac{1}{n} \cdot 2^{w+2}q + \frac{1}{n} \cdot \frac{1-n}{2n} \cdot 2^{2w+4}q^2 + \cdots \quad (\text{mod } 2^{m-w}).$$

The degree in q of (29) may be higher than necessary. If there is a solution of lower degree, it will be given by (28) mod 2^{m-w}. We shall find D from the condition

$$(30) \qquad E(\varDelta^i x_0) \geqq m - w \qquad (i > D),$$

by virtue of which further terms in (28) are omitted.

The left-hand side of (30) in terms of i is found from (29), which expresses x as a power series in q. We use

Lemma 1. *$\varDelta^i q^{i+j}$ is an integral multiple of $(i+j)!/(2j)!$*

Proof. The lemma is nugatory and obvious for any j when $i \leqq j$, and will be proved by induction from $i - 1$ to i.

Reducing by 1 the order of the differences gives

$$\varDelta^i q^{i+j} = \varDelta^{i-1}(q+1)^{i+j} - \varDelta^{i-1}q^{i+j} = \sum C(i+j, s)\varDelta^{i-1}q^{i+j-s} \quad (s = 1, \cdots, j+1).$$

The maximum s is $j + 1$ because $\varDelta^u q^v$ vanishes when $v - u < 0$.

The general term shown, on the assumption that the lemma is true for $\varDelta^{i-1}$ and any exponent of q, is a multiple of

$$\frac{(i+j)!}{s!(i+j-s)!} \cdot \frac{(i+j-s)!}{(2j-2s+2)!} = \frac{(i+j)!}{(2j)!} \cdot \frac{(2j)!}{(2j-2s+2)!(2s-2)!} \cdot \frac{(2s-2)!}{s!}.$$

Since $1 \leqq s \leqq j + 1$, the second and third factors on the right are non-zero integers. Therefore each term in the $\sum$ is a multiple of $(i+j)!/(2j)!$, so is their sum, and the lemma is proved.

9. Proof of Theorem 4

Write (29) in the form

$$(31) \qquad x \equiv 1 + c_1 q + c_2 q^2 + \cdots \quad (\text{mod } 2^{m-w}),$$

where

$$c_k = (1-n)(1-2n) \cdots 2^{k(w+2)}/(k!n^k),$$

$$(32) \qquad E(c_k) = k(w+2) - E(k!) - kw = 2k - E(k!).$$

Applying $\varDelta^i$ to (31) with $x = x_0$, $q = 0$, we get

$$(33) \qquad \Delta^i x_0 \equiv c_i \Delta^i 0^i + c_{i+1} \Delta^i 0^{i+1} + \cdots \quad (\text{mod } 2^{m-w}),$$

where, as is well known,

$$(34) \qquad \Delta^i 0^i = i!,$$

and by Lemma 1

$$(35) \qquad E(\Delta^i 0^{i+j}) \geqq E\{(i+j)!\} - E\{(2j)!\}.$$

Therefore, by (32) with $k = i$ and (34),

$$E(c_i \Delta^i 0^i) = 2i,$$

and, by (32) with $k = i + j$ $(j \geqq 1)$, (35), and (15),

$$E(c_{i+j} \Delta^i 0^{i+j}) \geqq 2(i+j) - E\{(2j)!\} = 2(i+j) - 2j + \sigma(2j) \geqq 2i + 1.$$

These two relations show that in (28)

$$E(\Delta^i x_0) = E(c_i \Delta^i 0^i) = 2i.$$

From this and (30), D is given by

$$2(D+1) \geqq m - w,$$

which is satisfied by

$$D = [\tfrac{1}{2}(m - w - 1)]$$

and by no lower value of D, when (18) is complied with.

A similar argument can be used whenever $r = 1$, that is, when by (6) there is only one n-ic residue mod p^{w+1}.

It will now be proved that any solution constructed from roots not satisfying (18) is of degree higher than $[\frac{1}{2}(m - w - 1)]$.

Congruences with $m - w \leqq 2$ are trivial. When $m - w = 3$, so that $n = 2^{m-3} n_1$, there are only the two residues 1 and $2^{m-1} + 1$ (mod 2^m) with respective roots $\pm\, 1$ and $\pm\, 5$ (mod 8). Any linear polynomial constructed from roots not satisfying (18), e.g. from roots $8h + 1$, $8k + 3$ and residues $2^m i + 1$, $2^m j + 2^{m-1} + 1$, where h, i, j, k are any integers, will be found to be congruent mod 8 to $\pm\, \{1 + 2^{-m+2}(a - 1)\}$, which shows that it fails when $a = 2^m + 1$. Thus the degree of a true polynomial solution is at least 2. But a polynomial solution constructed from roots satisfying (18) is linear.

When $m - w \geqq 4$, at least half of the 2^{m-w-2} roots, however one chooses them, have the same residue mod 4, and

$$2^{m-w-3} \geqq [\tfrac{1}{2}(m - w + 1)] = D + 1.$$

Therefore there are always $D + 1$ roots satisfying (18). They will be denoted by $x_0, \cdots, x_D$ and the corresponding residues by $a_0, \cdots, a_D$, while x_{D+1} and ξ will denote roots of a further residue a_{D+1} such that

$$4 | x_{D+1} - x_0, \qquad 4 \nmid \xi - x_0.$$

This means that

$$\xi = x_{D+1} + \rho,$$

where ρ is singly even.

The argument applied in § 7 to $\bar{\delta}^{\mu r}$, when applied here to $\bar{\delta}^{D+1}$, leads to

$$E\{\bar{\delta}^{D+1}(x/x^n)_0 (a_{D+1} - a_0) \cdots (a_{D+1} - a_D)\} = 1 < m - w.$$

If then (18) is not complied with, the degree is higher than $[\frac{1}{2}(m - w - 1)]$. This completes the proof of Theorem 4.

10. Theorem 5. Preliminary

When $r > 1$, the preceding argument based on Lemma 1 cannot be applied direct. The plan adopted is to find a relation between divided differences and ordinary ones, and use it in conjunction with a solution valid only for $a_0, a_r, \cdots$, to which ordinary differences and so Lemma 1 can be applied.

The relation is expressed between Δy, which for the time being will denote the increment in y corresponding to an increment r in the suffix or suffixes in y, and an operator Θ, whose definition is

$$\Theta y_0 = (a_r - a_0)(a_r - a_1) \cdots (a_r - a_{r-1})\delta^r (y/a)_0,$$

and more generally

(36) $$\Theta^i y_j = (a_{ir+j} - a_j)(a_{ir+j} - a_{j+1}) \cdots (a_{ir+j} - a_{ir+j-1})\delta^{ir}(y/a)_j.$$

Lemma 2. *If y is any function of the n-ic residue a such that*

$$E(\Delta^i y_j) = i + E(y),$$

then also

$$E(\Theta^i y_j) = i + E(y).$$

For this lemma n is such that $w \geqq 1$, all values of y have the same E, and the residues are in an order, complying with (22), such that

(37) $$a_{qr+j} = a_j + p^{w+1}q.$$

In proving it we make $j = 0$ without loss in generality, for any y and a in the sequence can be labelled with the suffix 0.

Proof. Reducing by 1 the order of the differences in (36) gives

(38) $$\Theta^i y_0 = \Pi(a_{ir} - a_s)(\delta_1^{ir-1} - \delta_0^{ir-1}) \quad (s = 1, \cdots, ir - 1;\ \delta = \delta(y/a)).$$

Now in calculating δ_j^{ir-1} from the δ^{ir-r}'s we divide only by

$$a_{ir-r+h} - a_k \quad (h, k = j, \cdots, j + r - 1;\ k \neq h),$$

$$\equiv a_h - a_k \pmod{p^{w+1}}$$

(39) $$\not\equiv 0 \pmod{p}$$

by (22). But $a_h - a_k$ (h, k as above) are the quantities we divide by in calculating a δ_j^{r-1}. Therefore

$$\delta_j^{ir-1} \equiv \delta_j^{r-1}\delta^{ir-r} \quad (\text{mod}\, p^{\mu} \text{ pro tem.}; \; \mu = w + 1 + \min E(\delta^{ir-r})),$$

and, by (26) and (22), (38) becomes

$$\Theta^i y_0/\Pi(a_{ir} - a_s) \equiv \delta_1^{r-1}\delta^{ir-r} - \delta_0^{r-1}\delta^{ir-r}$$

$$= \sum \frac{\delta_{h+1}^{ir-r}}{\Pi(a_{h+1} - a_{k+1})} - \sum \frac{\delta_h^{ir-r}}{\Pi(a_h - a_k)} \quad (h, k = 0, \cdots, r-1; \; k \neq h)$$

$$= \frac{\delta_r^{ir-r}}{\Pi(a_r - a_k)} + \sum \left\{ \frac{\delta_h^{ir-r}}{\Pi(a_h - a_r)} - \frac{\delta_h^{ir-r}}{\Pi(a_h - a_0)} \right\}$$

$$- \frac{\delta_0^{ir-r}}{\Pi(a_0 - a_k)} \quad (h, k = 1, \cdots, r-1)$$

$$\equiv \frac{\delta_r^{ir-r} - \delta_0^{ir-r}}{\Pi(a_{ir} - a_k)}.$$

Therefore

$$\Theta^i y_0 \equiv \Pi(a_{ir} - a_s)(\delta_r^{ir-r} - \delta_0^{ir-r}) \quad (\text{mod}\, p^M; \; s = r, \cdots, ir-1),$$

where

$$M = w + 1 + \min E\{\Pi(a_{ir} - a_s)\delta^{ir-r}\},$$

and this, by (37) and (26), is the same as

$$\Theta^i y_0 \equiv \Pi(a_{ir-r} - a_s) \sum \left\{ \frac{y_{r+h}}{\Pi(a_{r+h} - a_{r+k})} - \frac{y_h}{\Pi(a_h - a_k)} \right\} \quad (\text{mod}\, p^M)$$

$$= \Pi(a_{ir-r} - a_s) \sum \frac{\Delta y_h}{\Pi(a_h - a_k)} = \Theta^{i-1}\Delta y_0,$$

where

$$s = 0, \cdots, ir - r - 1; \; h, k = 0, \cdots, ir - r, \; k \neq h;$$

$$M = w + 1 + \min E\{\Pi(a_{ir-r} - a_s)\delta^{ir-r}\} = w + 1 + \min E(\Theta^{i-1}y).$$

The same argument, with Θ^{i-k} applied to $\Delta^k y_j$ instead of Θ^i applied to y_0, establishes

(40) $$\Theta^{i-k}\Delta^k y_j \equiv \Theta^{i-k-1}\Delta^{k+1}y_j \quad (\text{mod}\, p^M),$$

where

$$k = 0, \cdots, i-1; \; M = w + 1 + \min E(\Theta^{i-k-1}\Delta^k y).$$

By (40), with $i = 1$, $k = 0$,

$$E(\Theta y_j) = E(\Delta y_j) = 1 + E(y).$$

Assume that, for $k = 1, \cdots, i-2$,

$$E(\Theta^{i-1}y_j) = E(\Theta^{i-k-1}\Delta^k y_j) = E(\Delta^{i-1}y_j) = i - 1 + E(y).$$

Then the M in the modulus of (40) is $w + i + E(y)$, whatever the value of

k, and we have the chain of congruences

$$\Theta^i y_j \equiv \Theta^{i-k} \Delta^k y_j \equiv \Delta^i y_j \quad (\operatorname{mod} p^{w+i+E(y)};\ k = 1, \cdots, i-1).$$

The E of the last member is $i + E(y)$, hence so is that of every other member. Thus the assumption, if true for $i - 1$, is also true for i, and so is always true. Lemma 2 follows.

11. Proof of Theorem 5

When (37) applies with $a_0 = 1$, a solution of (3), valid only for $a_0, a_r, \cdots$, is given by (16) mod p^{m-w}. Writing it in the form

$$x \equiv 1 + c_1 q + c_2 q^2 + \cdots \quad (\operatorname{mod} p^{m-w}),$$

where

$$c_k = (1-n)(1-2n) \cdots p^{k(w+1)}/(k!n^k), \quad q = (a-1)/p^{w+1},$$

we have, corresponding to (32) and (33) in § 9,

$$E(c_k) = k(w+1) - E(k!) - kw = k - E(k!)$$

and

$$\Delta^i x_0 \equiv c_i \Delta^i 0^i + c_{i+1} \Delta^i 0^{i+1} + \cdots \quad (\operatorname{mod} p^{m-w}),$$

Δ being the difference for unit increment of q. Using (34) and (35), this time we get

$$E(c_i \Delta^i 0^i) = i,$$

and, with (15) also,

$$\begin{aligned} E(c_{i+j} \Delta^i 0^{i+j}) &\geqq i + j - E\{(2j)!\} = i + j - \{2j - \sigma(2j)\}/(p-1) \\ &\geqq i + j - (2j-2)/2 = i + 1 \quad (j \geqq 1). \end{aligned}$$

These two relations show that

$$E(\Delta^i x_0) = E(c_i \Delta^i 0^i) = i,$$

where Δ, the difference for unit increment of q, is by (37) the same as the difference for an increment r of the suffix. Hence Lemma 2 can be applied to the present Δ operating on x, all values of which have the same E, and it relates Δx to Θx.

By the argument following (23),

$$E\{(a-a_0)(a-a_1) \cdots (a - a_{rs-1})\}$$

has its minimum value when $a = a_{rs}$. Therefore by (36) and Lemma 2

$$\begin{aligned} &E\{(a-a_0) \cdots (a-a_{rs-1}) \delta^{rs} (x/a)_0\} \\ &\geqq E\{(a_{rs}-a_0) \cdots (a_{rs} - a_{rs-1}) \delta^{rs} (x/a)_0\} = E(\Theta^s x_0) = s, \end{aligned}$$

there being equality when $a = a_{rs}$. This shows that the term in (19) containing δ^{rs} can be omitted if and only if $s \geqq m - w$. Intermediate terms containing δ^{rs+j} ($j < r$) can then be omitted also, for in calculating δ^{rs+j} from δ^{rs} we do not divide by p (compare (39)), so the E of an intermediate term is not less than s. Therefore the minimum degree is at least $r(m-w-1)$ and at most $r(m-w)-1$, if when $d>1$ the roots chosen satisfy (10). It has been proved in § 7 that (10) will be necessary for minimum degree, so the proof of Theorem 5 is complete.

When $r=1$, the limits are equal and we have a precise value of the minimum degree, as in Theorem 4.

12. Worked example

To illustrate the divided-difference method, we solve

$$x^3 \equiv a \pmod{27}.$$

Here $r=2$, $m=3$, $w=1$, so by Theorem 5 the degree of the solution is at most 3. Also, since $w=1$, its modulus need only be 9. To satisfy (25), the roots are put in ascending numerical order.

$$\begin{array}{llllllllllll}
x & \equiv 1 & & 2 & & 4 & & 5 & & 7 & & \pmod 9 \\
a & = 1 & & 8 & & 10 & & 17 & & 19 & & \\
\delta & \equiv & 4 & & 1 & & 4 & & 1 & & & \pmod 9 \\
\delta^2 & \equiv & & -\frac{1}{3} & & \frac{1}{3} & & -\frac{1}{3} & & & & \pmod{3^0} \\
\delta^3 & \equiv & & & \dfrac{1}{8\cdot 3} & & \dfrac{-2}{11\cdot 3} & & & & & \pmod{3^0} \\
 & \equiv & & & -\frac{1}{3} & & -\frac{1}{3} & & & & & \text{since } 1/8 \equiv -2/11 \equiv -1 \pmod{3\cdot 3^0} \\
\delta^4 & \equiv & & & & 0 & & & & & & \pmod{1/9}
\end{array}$$

The modulus of δ^2 and δ^3 is written 3^0 instead of 1, to indicate that powers of 3 dividing a denominator are not to be treated like other integers.

From the table we have the solution

$$\begin{aligned}
x &\equiv 1+4(a-1)-\tfrac{1}{3}(a-1)(a-8)-\tfrac{1}{3}(a-1)(a-8)(a-10) \quad (\text{mod } 9 \text{ pro tem.})\\
&\equiv 3+\tfrac{4}{3}a-3a^2-\tfrac{1}{3}a^3 \equiv \tfrac{1}{3}a(4-a^2) \quad (\text{since } a^2 \equiv x^6 \equiv 1)\\
&= a\{1-\tfrac{1}{3}(a^2-1)\}.
\end{aligned}$$

This is a particular case of Theorem 6 below.

Such a solution can always be checked with the help of the Fermat-Euler theorem. In the present case

$$x^3 \equiv a^3\{1-(a^2-1)\} = 2a^3-a^5 = a-a(1-a^2)^2 \equiv a \pmod{27}.$$

13. Indicial difference greater than 1

The next three theorems make use of ν_i, the least positive integer satisfying

https://doi.org/10.1017/S1446788700025957 Published online by Cambridge University Press

$$\nu_i n \equiv 1 \pmod{\phi(p^i)/d, \ = p^{i-1}r}.$$

THEOREM 6. *When* $(d, r) = 1$, (3) *has polynomial solutions of degree*

$$D \leqq \nu_1 + r(s-1) \quad (p \nmid n;\ s + E(s!) \geqq m),$$
$$D \leqq \nu_1 + r(m-w-1) \quad (p|n),$$

in which the indicial difference is r. That is, the polynomial is of the form

$$c_0 a^\nu + c_1 a^{\nu+r} + c_2 a^{\nu+2r} + \cdots$$

PROOF. There is nothing to prove when $p = 2$, since then $r = 1$; so we consider only $p \geqq 3$. Write a in the form (11), but making k the least positive integer such that

$$kr + 1 \equiv 0 \pmod{n}, \qquad kr + 1 = \nu_1 n.$$

Such a k exists because by (12)

$$(n, r) = (dn_1 p^w, r) = (d, r) = 1.$$

The formal solution is now

(41) $$x \equiv a^\nu(1+a^r-1)^{-k/n} \equiv a^\nu\{1-(k/n)(a^r-1)+\cdots\} \pmod{p^{m-w};\ \nu = \nu_1},$$

where instead of the polynomial A in (13) we have the single term a.

Make in (41) the substitutions

(42) $$x/a^\nu \equiv y \pmod{p^{m-w}}, \qquad a^r - 1 \equiv p^{w+1}q \pmod{p^m},$$

the second of which preserves the congruences

$$\frac{1}{i!}\left\{\frac{a^r-1}{n}\right\}^i \equiv \frac{1}{i!}\left\{\frac{p^{w+1}q}{n}\right\}^i \pmod{p^{m-w}},$$

however grpeat i may be. For if

$$B = C + jp^{m-w} \equiv 0 \pmod{p},$$

then

$$B^i = C^i + ijp^{m-w}C^{i-1} + \cdots \equiv C^i \pmod{p^{m-w+i-1}},$$

and

$$B^i/i! \equiv C^i/i! \pmod{p^M},$$

where

$$M = m - w + i - 1 - E(i!) \quad (i \geqq 1),$$

and so by (15)

$$M \geqq m - w + i - 1 - (i-1)/(p-1) \geqq m - w.$$

Moreover q runs through the consecutive integers $0, 1, \cdots$, because by (6) the number of rth powers of n-ic residues is p^{m-w-1}, so differences for unit increment of q exist. The substitutions give

(43) $$y \equiv 1 - \frac{k}{n}p^{w+1}q + \frac{k}{n}\cdot\frac{k+n}{2n}p^{2w+2}q^2 - \cdots \pmod{p^{m-w}}$$
$$= 1 - c_1 q + c_2 q^2 - \cdots, \quad \text{say},$$

https://doi.org/10.1017/S1446788700025957 Published online by Cambridge University Press

where

$$c_i = k(k+n)\cdots\{k+(i-1)n\}p^{i(w+1)}/(i!n^i).$$

This is equivalent to a solution of (3), and corresponds to (31) in § 9.

The rest of the proof, whether p divides n or not, is similar to that of Theorem 4. A formula for y corresponding to (28) indicates that D_q, the degree in q of the formula, must satisfy

$$E(\Delta^i y_0) \geqq m - w \quad (i > D_q,\ w \geqq 0),$$

whose left-hand side is found from (43), by an argument using Lemma 1, to be the left-hand side of

$$(44) \qquad E(c_i \Delta^i 0^i) \geqq m - w \quad (i > D_q,\ w \geqq 0).$$

The case in which p does not divide n requires

Lemma 3. *If $p \nmid n$, then*

$$E[k(k+n)\cdots\{k+(i-1)n\}] \geqq E(i!).$$

Proof. If κ is an integer such that

$$k \equiv \kappa n \quad (\text{mod}\, p^M;\ M > E[k(k+n)\cdots\{k+(i-1)n\}]),$$

then

$$E[k(k+n)\cdots\{k+(i-1)n\}] = E\{n^i\kappa(\kappa+1)\cdots(\kappa+i-1)\} \geqq E(i!).$$

Lemma 3 shows that

$$E(c_i) \geqq i \qquad (p \nmid n),$$

its companion being

$$E(c_i) = i - E(i!) \qquad (p|n).$$

From these and (34) condition (44) becomes

$$i + E(i!) \geqq m \quad (p \nmid n), \qquad i \geqq m - w \quad (p|n), \qquad (i > D_q).$$

Sufficient values of the degree in a are now seen from (41) and (42) to be those given in the enunciation of Theorem 6. This proves the theorem.

14. Fewer terms of higher degree

Instead of (11) write

$$(45) \qquad a = \frac{a^{hkr+1}}{(1+a^{hr}-1)^k}.$$

Let $h = p$, and let k be the least positive integer such that

$$kpr + 1 \equiv 0 \quad (\text{mod}\, n), \qquad kpr + 1 = \nu_2 n;$$

this is possible when p does not divide n and $(d, r) = 1$. Then (3) has a solution

$$x \equiv a^\nu(1 + a^{pr} - 1)^{-k/n} \equiv a^\nu\{1 \dot{-} (k/n)(a^{pr}-1) + \cdots\} \quad (\text{mod}\, p^m;\ \nu = \nu_2),$$

corresponding to (41), to which the argument of Theorem 6 can be applied. Instead of

$$a^r - 1 \equiv 0 \pmod{p}$$

we have

$$a^{pr} - 1 \equiv 0 \pmod{p^2},$$

so there are fewer terms, of higher degree.

On replacing h in (45) by p^2, p^3, $\cdots$, solutions are obtained in which the number of terms is progressively smaller, and the degree progressively higher. This process culminates in

THEOREM 7. *When $p \nmid n$ and $(d, r) = 1$, (3) has a solution*

$$x \equiv a^\nu \pmod{p^m},$$

where

$$\nu = \nu_{m-1} \quad (p = 2), \qquad \nu = \nu_m \quad (p \geqq 3).$$

PROOF.

$$(a^\nu)^n = a \cdot a^{\nu n-1} \equiv a x^{n(\nu n-1)} \pmod{p^m}.$$

When $p = 2$ the exponent of x is

$$n(\nu_{m-1} n - 1) = nk \cdot 2^{m-2},$$

so [5]

$$(a^\nu)^n \equiv a \quad (\text{mod } 2^m;\ \nu = \nu_{m-1}).$$

When $p \geqq 3$ the exponent of x is

$$n(\nu_m n - 1) = dn_1 k p^{m-1} r = n_1 k p^{m-1}(p - 1) = n_1 k \phi(p^m),$$

so

$$(a^\nu)^n \equiv a \quad (\text{mod } p^m;\ \nu = \nu_m).$$

As ν_i exists for all i when p does not divide n, this proves Theorem 7.

When p divides n, the only ν that exists is ν_1, so the only possible substitution for a is (11), leading to (41). Since (41) consists of a single term only if the modulus is p, the modulus of the congruence to be solved cannot be greater than p^{w+1} $(p \geqq 3)$ or 2^{w+2}.

15. Proof of Theorem 8

Corresponding to each prime power p^m in the modulus, there is a ν_m (or ν_{m-1}) such that

$$\nu_m n \equiv 1 \quad (\text{mod } p^{m-1} r;\ p \geqq 3) \quad \text{or} \quad \nu_{m-1} n \equiv 1 \pmod{2^{m-2}},$$

because n and $p^{m-1}r$ (or 2^{m-2}) are coprime. The latter is also true of n and L, the LCM of all the expressions such as $p^{m-1}r$. Therefore there is a ν such that

$$\nu n \equiv 1 \pmod{L},$$

and, by the argument of Theorem 7, a^{ν} is a solution for each prime power in the modulus, so it is a solution of (2). This proves Theorem 8.

The condition $(n, M) = 1$ is sufficient, but not necessary. For, as pointed out in § 14, p may divide n provided $m \leqq w + 1$ $(p \geqq 3)$, $m \leqq w + 2$ $(p = 2)$.

16. Proof of Theorem 9

Only $p \geqq 3$ need be considered, since $p = 2$ is covered by Theorem 7. It will be proved that there is a solution consisting of δ terms, giving roots that are congruent mod p^{m-w} to

(46) $$g^{i+jd\delta} \quad (i = 0, \cdots, \delta - 1;\ j = 0, \cdots, \phi/d\delta - 1),$$

where g is a primitive root and ϕ denotes $\phi(p^{m-w})$. This will be a complete set of roots provided that their number is ϕ/d, which is clearly true, and that their nth powers are all different mod p^m. But if

$$g^{n(i+jd\delta)} \equiv g^{n(h+kd\delta)} \quad (\text{mod}\, p^m;\ h, i < \delta;\ j, k < \phi/d\delta),$$

then

$$\phi(p^m) | n(i-h) + n(j-k)d\delta,$$

and it is easy to show, using (12) and the fact that $\delta | r$, that this divisibility requires $h = i$ and $j = k$. Hence (46) is a complete set of roots.

Let powers of some quantity y be connected with ω, a primitive δth root mod p^{m-w} of 1, by the relations

(47) $$y^i \equiv c_0 + c_1\omega^i + \cdots + c_{\delta-1}\omega^{(\delta-1)i} \quad (\text{mod}\, p^{m-w} \text{ pro tem.};\ i = 0, \cdots, \delta-1),$$

which in matrix notation are

$$\begin{bmatrix} 1 & 1 & 1 & \cdots \\ 1 & \omega & \omega^2 & \cdots \\ \cdot & \cdot & \cdot & \cdot \\ 1 & \omega^{\delta-1} & \cdots & \cdots \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \cdots \\ c_{\delta-1} \end{bmatrix} \equiv \begin{bmatrix} 1 \\ y \\ \cdots \\ y^{\delta-1} \end{bmatrix}.$$

Multiplying the matrix above by that below yields a unit matrix.

$$\begin{bmatrix} c_0 \\ c_1 \\ \cdots \\ c_{\delta-1} \end{bmatrix} \equiv \frac{1}{\delta} \begin{bmatrix} 1 & 1 & 1 & \cdots \\ 1 & \omega^{-1} & \omega^{-2} & \cdots \\ \cdot & \cdot & \cdot & \cdot \\ 1 & \omega^{-\delta+1} & \cdots & \cdots \end{bmatrix} \begin{bmatrix} 1 \\ y \\ \cdots \\ y^{\delta-1} \end{bmatrix}.$$

Therefore

(48) $$\begin{aligned} c_i \equiv \delta^{-1}\{1 + y\omega^{-i} + \cdots + (y\omega^{-i})^{\delta-1}\} &= \delta^{-1}\{(y\omega^{-i})^{\delta} - 1\}/(y\omega^{-i} - 1) \\ &\equiv \delta^{-1}(y^{\delta} - 1)/(y\omega^{-i} - 1). \end{aligned}$$

Now let

(49) $$y \equiv g^{-k\phi/d\delta}, \qquad \omega \equiv g^{n\phi/d\delta}.$$

The expression for ω is a primitive δth root of 1 because δ divides r which by (12) is prime to n/d, whence so is δ. The exponent k is derived from the least positive integer ν satisfying

$$\nu n \equiv 1 \pmod{\phi/d\delta}$$

by making

$$\nu n - 1 = k\phi/d\delta;$$

k and ν exist when p does not divide n (as in Theorem 9) with any power of p as modulus, and when p divides n with modulus p^{w+1}.

On using (48) and (49), and for brevity writing

$$g^{\phi/d\delta} = \gamma$$

where convenient, (47) becomes

$$g^{-ik\phi/d\delta} \equiv \frac{1}{\delta}(\gamma^{-k\delta} - 1)\left\{\frac{1}{\gamma^{-k} - 1} + \frac{g^{in\phi/d\delta}}{\gamma^{-k-n} - 1} + \cdots + \frac{g^{(\delta-1)in\phi/d\delta}}{\gamma^{-k-(\delta-1)n} - 1}\right\} \pmod{p^{m-w}}.$$

Increasing i to $i + jd\delta$, where j is any integer, does not affect this congruence, for i always has $\phi/d\delta$ as a cofactor, and

$$(i + jd\delta)\phi/d\delta \equiv i\phi/d\delta \pmod{\phi}.$$

So we make the substitutions

$$g^{-ik\phi/d\delta} \equiv g^{-(i+jd\delta)k\phi/d\delta} = g^{(i+jd\delta)(1-\nu n)} \equiv xa^{-\nu} \quad (\mathrm{mod}\, p^{m-w} \text{ pro tem.}),$$

$$g^{in\phi/d\delta} \equiv g^{(i+jd\delta)n\phi/d\delta} \equiv a^{\phi/d\delta},$$

and get the following solution of (3):

$$(50) \qquad x \equiv \frac{1}{\delta}(\gamma^{-k\delta} - 1)a^{\nu}\left\{\frac{1}{\gamma^{-k} - 1} + \frac{a^{\phi/d\delta}}{\gamma^{-k-n} - 1} + \cdots + \frac{a^{(\delta-1)\phi/d\delta}}{\gamma^{-k-(\delta-1)n} - 1}\right\},$$

which gives the roots in (46). This proves Theorem 9. Theorem 7 is the particular case in which $\delta = 1$, and gives roots that are powers of g^d.

Though substitution in (50) provides a solution of a given congruence, it is far from convenient. But (50) shows, for instance, that the congruence

$$x^2 \equiv a \quad (\mathrm{mod}\, p^m \text{ pro tem.};\ p = 8q + 5),$$

for which

$$d = 2, \quad \delta = 2, \quad \nu = (p + 4)/8,$$

has a solution of the form

$$x \equiv \pm\, a^{(p+4)/8}(c_0 + c_1 a^{\phi/4}).$$

Squaring shows that

$$c_0^2 + c_1^2 \equiv 0, \qquad 2c_0c_1 \equiv 1,$$

https://doi.org/10.1017/S1446788700025957 Published online by Cambridge University Press

and these are satisfied by

$$c_0 = \tfrac{1}{2}(i+1), \quad c_1 = -\tfrac{1}{2}(i-1) \quad (i^2 \equiv -1).$$

17. A question of minimum degree

The ν_1 of Theorem 6 is often appreciably less than $r-1$, so it was thought worth while to see if solutions as in Theorem 6 were always of minimum degree. For if this were so, there would be a precise value of the minimum degree, as in Theorem 4. It is disproved by counter-example as follows.

Let $x_0, x_1, \cdots, x_{r-1}$ be arbitrary nth roots of 1, $g^n, \cdots, g^{(r-1)n}$ respectively. The coefficients in

$$x_i \equiv c_0 + c_1 a_i + \cdots + c_{r-1} a_i^{r-1} \equiv c_0 + c_1 g^{in} + \cdots + c_{r-1} g^{(r-1)in}$$
$$(\text{mod } p;\ i = 0, \cdots, r-1),$$

corresponding to (47), are found by the procedure of § 16 to be

$$c_i \equiv r^{-1}(x_0 + x_1 g^{-in} + \cdots + x_{r-1} g^{-(r-1)in}) \equiv r^{-1} \sum x^{1-in} \quad (\text{mod } p),$$

whence the congruence

$$x^n \equiv a \ (\text{mod } p^{w+1})$$

has a solution [6]

$$(51) \qquad x \equiv r^{-1}(\sum x + a \sum x^{1-n} + \ldots + a^{r-1} \sum x^{1-(r-1)n}) \quad (\text{mod } p).$$

In the congruence $x^9 \equiv a$ (mod 31 pro tem.), $(d, r) = (3, 10) = 1$ and $\nu_1 = 9$ $(= r-1)$, so by Theorem 6 or 7 there is a solution a^9, giving roots that are powers of g^3. With such roots the last sum in (51) is

$$\sum x^{10} \equiv 1 + \cdots + 1 = 10.$$

Now multiply any one root by 5, and any two further roots by 25; these are the "complex" 9th roots mod 31 of 1. The 10th powers of the roots affected are also multiplied by 5 or 25, since

$$5^{10} = 5^9 \cdot 5 \equiv 5,$$

and the last sum in (51) is now

$$\sum x^{10} \equiv 7 \cdot 1 + 1 \cdot 5 + 2 \cdot 25 \equiv 0.$$

Hence in this case there are solutions of degree lower than ν_1.

18. Residues not prime to modulus

When a includes values divisible by p, polynomial solutions of (3) are still possible, at least if m is fairly small. Thus the congruence

$$x^2 \equiv a \quad (\text{mod } 64)$$

https://doi.org/10.1017/S1446788700025957 Published online by Cambridge University Press

has a solution

$$x \equiv \pm\, a(a^4 - 69a^3 - 77a^2 + 189a - 28)/16 \pmod{32}$$

that is valid for all admissible values of a, whether odd or even.

Acknowledgements

I am grateful to Professor E. S. Barnes and Dr B. C. Rennie for the interest they showed in this paper, and for criticism and advice in its preparation. I am also indebted to the referee for pointing out an error in Theorem 9.

References

[1] Proved in Uspensky, J. V. and Heaslet, M. A., Elementary Number Theory, pp. 311—4. For further polynomial and other explicit solutions see Dickson, L. E., History of the Theory of Numbers, vol. 1, ch. VII.

[2] Nagell, T., Introduction to Number Theory, p. 118, Theorem 71.

[3] Uspensky and Heaslet, Elementary Number Theory, p. 103, ex. 1. Proved in Griffin, H., Elementary Theory of Numbers, pp. 42—3.

[4] Fort, T., Finite Differences, p. 10, Theorem IV.

[5] Nagell, Introduction to Number Theory, p. 104, Theorem 63.

[6] M. Cipolla obtained (51) for a quadratic congruence in the following form. The solution of

$$x^2 \equiv q \pmod{p}$$

is

$$x \equiv \pm\, 2(qs_1 + q^2 s_3 + q^3 s_5 + \cdots + y^{(p-3)/2} s_{p-4} + s_{p-2}),$$

where

$$s_r = 1^r + 2^r + \cdots + \{(p-1)/2\}^r.$$

(In (51) let $x \equiv 1, \frac{1}{2}, \cdots, 2/(p-1)$ on the right.) He later extended the exponent of x to any divisor of $p - 1$ (i.e. $n = d$ and $m = 1$ in the notation used here), and found an equivalent of (50) for this case. (Dickson's History, vol. 1, pp. 219, 220; Rendiconto Accad. Sc. Fis. e Mat. Napoli (3), 11 (1905), 13—19; Math. Annalen, 63 (1907), 54—61.)

Patent Office, Canberra.

https://doi.org/10.1017/S1446788700025957 Published online by Cambridge University Press