en remarquant que c'est le coefficient de $z_1^{r_1} z_2^{r_2}$ dans le polynôme

$$\sum_{n \leqslant x} \chi(n) z_1^{f_1(n)} z_2^{f_2(n)}.$$

Par exemple, en prenant $f_1 = \omega, f_2 = \Omega - \omega$ et $\chi(n) = 1$ pour tout n, on obtient ainsi le résultat suivant.

THÉORÈME 4. *Étant donné les entiers $m \geqslant 2$ et $h \geqslant 1$, il existe une suite de polynômes $P_0, P_1, \ldots, P_j, \ldots$ de degré $m-2$ telle que, $\nu_{m,h}(x)$ étant le nombre des $n \leqslant x$ pour lesquels $\omega(n) = m$ et $\Omega(n) = m+h$, quel que soit q entier $\geqslant 0$, on a quand x tend vers $+\infty$*

$$\nu_{m,h}(x) = \sum_{j=0}^{q} \frac{x P_j(\log\log x)}{(\log x)^{j+1}} + O\left(\frac{x(\log\log x)^{m-2}}{(\log x)^{q+2}}\right).$$

Le coefficient de X^{m-2} dans $P_j(X)$ est

$$\frac{(-1)^j}{(m-2)!} H_h^{(j)}(1), \quad \text{où} \quad H_h(s) = \frac{1}{s} \sum \frac{1}{p^{(h+1)s}}.$$

Ajoutons que le complément au théorème C mentionné au paragraphe 6.4.2 permet d'obtenir un développement asymptotique du nombre des n au plus égaux à x tels que

$$\omega(n) = m, \quad \Omega(n) = m+h \quad \text{et} \quad n \equiv l \pmod{k}.$$

Reçu le 15. 2. 1970 (38)

ACTA ARITHMETICA
XIX (1971)

On Waring's Problem in GF[p]

by

M. M. DODSON (Heslington)

In this paper we consider the solubility of Waring's Problem and of two related equations in GF[p], the Galois field of p elements. It is convenient to consider the equivalent question of the solubility of the congruences (mod p). Throughout, p will denote a prime, k a positive integer and $d = (k, p-1)$ will denote the highest common factor of k and $p-1$. We shall always write $t = (p-1)/d$.

The first congruence to be considered then is

$$x_1^k + \ldots + x_s^k \equiv N \pmod{p}, \tag{1}$$

where N is an arbitrary integer. Let $\Gamma(k, p)$ be the least positive integer s such that this congruence has a non-trivial solution (i.e. a solution with not all of the integral variables $x_1, \ldots, x_s$ divisible by p) for all integers N. Hardy and Littlewood ([12]) showed that if $d < \frac{1}{2}(p-1)$, then $\Gamma(k,p) \leqslant k$ and Chowla, Mann and Straus ([6]) showed that if $d < \frac{1}{2}(p-1)$, then $\Gamma(k, p) \leqslant [\frac{1}{2}(k+4)]$. I. Chowla ([4]) showed that if k is sufficiently large, then for all primes p with $d = (k, p-1) < \frac{1}{2}(p-1)$, we have

$$\Gamma(k, p) < k^{1-c+\varepsilon}$$

where, as always, ε is a positive number, and $c = (103 - 3\sqrt{641})/220$ lies strictly between 1/8 and 1/9, a result which for large k is much stronger than Hardy and Littlewood's and that of Chowla, Mann and Straus. Here this estimate is improved slightly to the simpler result that if $d < \frac{1}{2}(p-1)$, then

$$\Gamma(k, p) < k^{7/8},$$

provided k is sufficiently large. This is probably a long way from the truth and indeed Heilbronn ([14], p. 5) has conjectured that given $\varepsilon > 0$, and k large enough,

$$\Gamma(k, p) < k^{\varepsilon},$$

for t sufficiently large or at least, that if $t > 2$ (corresponding to $d < \frac{1}{2}(p-1)$), then $\Gamma(k, p) = O(k^{1/2})$. We note that the cases $d = p-1$ and

$d = \frac{1}{2}(p-1)$ are exceptional in the sense that in these cases $\Gamma(k,p)$ can be determined exactly and is in general not small. For instance, Hardy and Littlewood showed ([12], p. 524, Lemma 7) that if $p = k+1 > 2$, then

$$\Gamma(p-1,p) = p = k+1$$

and provided $p > 3$, if $p = 2k+1$, then

$$\Gamma\left(\tfrac{1}{2}(p-1),p\right) = \tfrac{1}{2}(p-1) = k.$$

Secondly we let $\theta(k,p)$ be the least s such that the congruence

$$x_1^k + \ldots + x_s^k \equiv 0 \pmod p \tag{2}$$

has a non-trivial solution. Plainly if k is odd or more generally if -1 is a kth power residue (mod p), $\theta(k,p) = 2$. Clearly for all $k \geqslant 1$

$$\theta(k,p) \leqslant \Gamma(k,p)$$

and S. Chowla ([5], p. 62) has conjectured that if $d < p-1$, then for k sufficiently large,

$$\theta(k,p) < k^{1/2+\varepsilon}.$$

The simplification of having $N = 0$ enables us to combine an estimate developed for $\Gamma(k,p)$ which is effective when d is near $p^{1/2}$, with the inequality

$$\theta(k,p) \leqslant t = (p-1)/d,$$

which is effective when d is near p, and we are able to show that if $d < p-1$, then

$$\theta(k,p) < k^{2/3+\varepsilon}$$

for k sufficiently large. This goes some way towards proving Chowla's conjecture.

Thirdly we consider the function $\gamma^*(k,p)$, defined to be the least positive integer s such that the congruence

$$a_1 x_1^k + \ldots + a_s x_s^k \equiv 0 \pmod p, \tag{3}$$

where $a_1, \ldots, a_s$ are arbitrary integers not divisible by p, has a non-trivial solution. When k is odd and sufficiently large, A. Tietäväinen ([15]) has proved using beautifully arguments that

$$\gamma^*(k) = \operatorname*{Max}_p \gamma^*(k,p) < (1+\varepsilon)(\log k)/(\log 2),$$

a result which is also best possible. In general, such arguments are not available and the problem of estimating $\gamma^*(k,p)$ is much more difficult.

However the author ([9], p. 167, Lemma 2.6.7) has shown that if $1 \leqslant d < p-1$, then for $k > 1$,

$$\gamma^*(k,p) < 12(\log k)^2 k^{7/8}$$

so that for k sufficiently large,

$$\gamma^*(k,p) < k^{7/8+\varepsilon}.$$

By using estimates obtained for $\Gamma(k,p)$ and $\theta(k,p)$, this is improved to

$$\gamma^*(k,p) < k^{2/3+\varepsilon}$$

for $d < p-1$ and k sufficiently large. Of course this result is of interest only when k is even.

We conclude by observing that the improvement gained for $\gamma^*(k,p)$ leads to a corresponding improvement in the estimate for the related function $\Gamma^*(k,p)$.

1. The congruence $x_1^k + \ldots + x_s^k \equiv N \pmod p$. The arguments and results given in this section are modifications of those due to I. Chowla ([4]) in his investigation of Waring's Problem (mod p). Unfortunately, his paper which is not easily obtainable contains a number of misprints and obscurities and for these reasons and in order to keep this paper reasonably self-contained, we give his arguments in some detail.

It is well-known that the non-zero residue classes (mod p) form a cyclic group of order $p-1$, and it follows from this that the values assumed by x^k, for a given exponent k and arbitrary x are the same as the values assumed by x^d, where as always, $d = (k, p-1)$. Hence

$$\Gamma(k,p) = \Gamma(d,p)$$

and therefore it suffices to investigate $\Gamma(d,p)$ where d divides $p-1$. The case $d = 1$ is trivial and plainly

$$\Gamma(1,p) = 2. \tag{4}$$

From now on we shall assume that p is odd, since when $p = 2$ the only possibility is $d = 1$, covered by (4). Also for the rest of this section, we assume that $d < \frac{1}{2}(p-1)$, since the cases $d = p-1$ and $d = \frac{1}{2}(p-1)$ are somewhat special and do not concern us here.

It is convenient to introduce the function $\gamma(d,p)$, which is defined to be the least positive integer s such that every residue class (mod p) is representable as a sum of s dth powers (mod p). Plainly

$$\gamma(d,p) \leqslant \Gamma(d,p) \leqslant \gamma(d,p)+1,$$

and since $\Gamma(d,p) \leqslant d$ when $d < \frac{1}{2}(p-1)$ ([12], p. 533, Theorem 4) we have

$$\gamma(d,p) \leqslant \Gamma(d,p) \leqslant d.$$

This has been improved by Chowla, Mann and Straus ([6]) to

$$I(d,p) \leqslant [\tfrac{1}{2}(d+4)],$$

where $[x]$ denotes the integral part of x, but nothing would be gained by using their result here.

As is well known, the use of exponential sums provides a powerful method of attack in additive problems. We make some definitions: we define

$$e_p(y) = e^{(2\pi i y)/p}, \qquad S(a) = \sum_{x=0}^{p-1} e_p(ax^d), \qquad T(a) = \sum_{u} e_p(au),$$

where the last sum is over the $t = (p-1)/d$ distinct non-zero dth power residues (mod p). We shall always let u denote a non-zero dth power residue (mod p), so that for instance the t distinct non-zero dth power residue classes (mod p) will be written $u_1, \ldots, u_t$. Since the congruence $x^d \equiv u \pmod p$ has just d solutions, we have

$$S(a) = 1 + dT(a).$$

We recall that the $p-1$ non-zero residue classes (mod p) fall into d disjoint equivalence classes, one such class consisting of the dth power residues, and the others being the various classes of dth power non-residues. We denote by $\sum_b^*$ a sum in which b runs through a set of d distinct representatives of these classes. It can be proved readily that

$$\sum_b^* |S(b)|^2 = d(d-1)p \tag{5}$$

([9], p. 170, Lemma 2.5.1).

Also, we shall write $L = (\log p)/r$, where r is a large positive number so that $0 < L < 1$. Then we have (loc. cit. p. 170, Lemma 2.5.2)

LEMMA 1. *Suppose $\gamma(d,p) > r$. Then there exists an integer a, prime to p, such that*

$$|S(a)| > p(1-L)$$

and this implies that

$$|S(ma)| > p(1-m^2L)$$

for every non-zero integer m.

As an immediate consequence we have the

COROLLARY. *If for all integers a prime to p*

$$|S(a)| \leqslant p(1-L)$$

then

$$\gamma(d,p) \leqslant r.$$

Moreover, if $r < p$ and if for all integers a prime to p,

$$|T(a)| \leqslant p(1-2L),$$

then

$$\gamma(d,p) \leqslant r.$$

For

$$|S(a)| = |1+dT(a)| \leqslant 1+d|t(1-2L) \leqslant p(1-L).$$

Using this lemma in combination with the equation (5) it can be shown ([9], p. 171, Lemma 2.5.4) that if $p > d^2$, then

$$\gamma(d,p) \leqslant [8\log p]+1 \leqslant \mathrm{Max}(3, [32\log d]+1) \tag{6}$$

since exponential sum arguments (loc. cit. § 2.4) give $\gamma(d,p) \leqslant 3$ if $p > d^4$. Thus unless otherwise stated we shall suppose that from now on $p < d^2$ and we shall write

$$p = d^{1+\varrho}, \qquad 0 < \varrho < 1.$$

A natural approach to additive problems is by means of theorems on the addition of residue classes (mod p), and here following I. Chowla we make repeated use of the Cauchy–Davenport Theorem (Cauchy [1] and Davenport [7], [8]) in order to obtain an estimate for $\gamma(d,p)$ which is effective for small ϱ, i.e. when p is not much greater than d.

LEMMA 2. *Let $p = d^{1+\varrho}$ $(0 < \varrho < 1)$ be a prime > 243 such that $p-1$ is a multiple of d with $d < \frac{1}{2}(p-1)$. Then*

$$\gamma(d,p) < 54 \cdot d^{(3/5)(1+\varrho)}.$$

Proof. Since by hypothesis $t = (p-1)/d > 2$, we can find a non-zero dth power residue (mod p), R say, which is not congruent to ± 1 (mod p). Consider the least positive residues (mod p) of the numbers

$$0, R, 2R, \ldots, ([p^{1/2}]+1)R.$$

These are all distinct and define $[p^{1/2}]+2$ different points distributed amongst the $[p^{1/2}]+1$ intervals defined by

$$rp^{1/2} \leqslant \xi < (r+1)p^{1/2}; \qquad r = 0, 1, \ldots, [p^{1/2}].$$

By the Box Principle at least one of these intervals must contain at least two points and it follows that there exist integers x and y satisfying

$$1 \leqslant |x| < p^{1/2}, \qquad 1 \leqslant y < p^{1/2}$$

such that

$$R \equiv x \cdot y^{-1} \pmod p.$$

Moreover we can assume without loss of generality that x and y are coprime and that $|x| > y$, since otherwise we can replace R by R^{-1} in the preceding argument. Thus we have

$$R \equiv x \cdot y^{-1} \pmod p,$$

where $1 \leqslant y < |x| < p^{1/2}$, $(x,y) = 1$.

We consider three separate cases:

Case 1. $p^{2/5} < |x| < p^{1/2}$. The $([\frac{1}{2}p^{2/5}]+1)^2$ numbers of the form

$$m+nR, \quad 0 \leqslant m, n < \tfrac{1}{2}p^{2/5}, \tag{7}$$

are mutually incongruent (mod p). For suppose

$$m_1+n_1R \equiv m_2+n_2R \pmod p.$$

Then

$$(n_1-n_2)R \equiv m_2-m_1 \pmod p,$$

i.e.

$$(n_1-n_2)x \equiv (m_2-m_1)y \pmod p.$$

But

$$|x(n_1-n_2)-y(m_2-m_1)| < p^{1/2}\cdot p^{2/5} < p,$$

whence

$$x(n_1-n_2) = y(m_2-m_1).$$

Moreover, since x and y are coprime, it follows that x divides m_2-m_1, and in view of the inequality $|x| > p^{2/5} > |m_2-m_1|$, we deduce that $m_1 = m_2$, which plainly implies that $n_1 = n_2$. Thus the numbers (7) form $([\frac{1}{2}p^{2/5}]+1)^2$ distinct residues (mod p), each number being the sum of at most $p^{2/5}$ dth power residues (mod p).

Now by the Cauchy–Davenport Theorem, the expression

$$m_1+n_1R+\ldots+m_r+n_rR, \quad 0 \leqslant m_i, n_i < \tfrac{1}{2}p^{2/5}, \ 1 \leqslant i \leqslant r,$$

represents at least $\min(r\nu-r+1, p)$ distinct residue classes (mod p), where $\nu = ([\frac{1}{2}p^{2/5}]+1)^2$. It follows by definition that provided $p > 5$,

$$\gamma(d,p) < p^{2/5}(p-1)/(\nu-1),$$

whence

$$\gamma(d,p) < 8\cdot p^{3/5}.$$

Case 2. $p^{1/5} < |x| < p^{2/5}$. The argument in this case is similar to the previous one, except that here we consider numbers of the form

$$l+mR+nR^2, \quad 0 \leqslant l, m, n \leqslant \tfrac{1}{3}p^{1/5}. \tag{8}$$

We note that $R^2 \not\equiv \pm 1 \pmod p$, since otherwise we would have $x^2 \pm y^2 \equiv 0 \pmod p$, where $y < |x| < p^{2/5}$, which is impossible. Suppose two of the numbers of the form (8) are congruent (mod p), i.e. suppose that

$$l_1+m_1R+n_1R^2 \equiv l_2+m_2R+n_2R^2 \pmod p.$$

Then

$$(l_1-l_2)+(m_1-m_2)R+(n_1-n_2)R^2 \equiv 0 \pmod p,$$

whence

$$(l_1-l_2)y^2+(m_1-m_2)xy+(n_1-n_2)x^2 \equiv 0 \pmod p.$$

But the inequality

$$|(l_1-l_2)y^2+(m_1-m_2)xy+(n_1-n_2)x^2| < 3\cdot\tfrac{1}{3}\cdot p^{1/5}\cdot p^{4/5} = p$$

implies that

$$(l_1-l_2)y^2+(m_1-m_2)xy+(n_1-n_2)x^2 = 0.$$

Since $(x, y) = 1$, we must have that x divides l_1-l_2, and by the inequality $|x| > p^{1/5} > |l_1-l_2|$, we have further that $l_1 = l_2$. Thus we have that

$$(m_1-m_2)y+(n_1-n_2)x = 0.$$

Again x must divide m_1-m_2 and it follows from the inequalities $|x| > p^{1/5} > |m_1-m_2|$ that $m_1 = m_2$, whence plainly $n_1 = n_2$. Therefore the integers (8) form $([\frac{1}{3}p^{1/5}]+1)^3$ distinct residue classes (mod p), each number being the sum of at most $p^{1/5}$ dth power residues (mod p).

Now by the Cauchy–Davenport Theorem the expression

$$l_1+m_1R+n_1R^2+\ldots+l_r+m_rR+n_rR^2, \quad 0 \leqslant l_i, m_i, n_i < \tfrac{1}{3}p^{1/5}, \ 1 \leqslant i \leqslant r,$$

represents at least $\min(r\nu-r+1, p)$ distinct residue classes (mod p), where $\nu = ([\frac{1}{3}p^{1/5}]+1)^3$. Again, provided $p > 243$, it follows from the definition of $\gamma(d,p)$ that

$$\gamma(d,p) < p^{1/5}\frac{p-1}{\nu-1} < 2\cdot p^{1/5}\frac{p}{\nu} < 54p^{3/5}.$$

Case 3. $1 < |x| < p^{1/5}$. Since $|x| \geqslant 2$, we can find a positive integer f such that

$$p^{2/5} < |x|^f < p^{3/5}.$$

Thus we have

$$R^f \equiv x^f y^{-f} \pmod p,$$

where $(x^f, y^f) = 1$, $1 \leqslant y^f < |x|^f < p^{3/5}$, $|x|^f > p^{2/5}$ and it follows that R^f, which is plainly congruent to a dth power, is not congruent to ± 1 (mod p).

The argument now proceeds on the same lines as before: numbers of the form

$$m+nR^f, \quad 0 \leqslant m, n < \tfrac{1}{2}p^{2/5}, \tag{9}$$

are all mutually incongruent (mod p). For suppose

$$m_1+n_1R^f \equiv m_2+n_2R^f \pmod p.$$

Then

$$R^f \equiv x^f y^{-f} \equiv (m_1-m_2)(n_1-n_2)^{-1} \pmod p,$$

whence

$$x^f(n_1-n_2) \equiv y^f(m_1-m_2) \pmod p.$$

But we have that $|x^f(n_1-n_2)-y^f(m_1-m_2)| < p^{3/5}\cdot p^{2/5} = p$, which implies

$$x^f(n_2-n_1) = y^f(m_1-m_2)$$

and since $(x^l, y^l) = 1$, it follows that x^l divides $m_1 - m_2$. However since $|x|^l > p^{2/5} > |m_1 - m_2|$, we have $m_1 = m_2$, whence $n_1 = n_2$.

Thus the numbers (9) form $([\frac{1}{2}p^{2/5}]+1)^2$ distinct residues (mod p), each number being the sum of at most $p^{2/5}$ dth power residues. The Cauchy–Davenport Theorem implies that the expression

$$m_1 + n_1 R^l + \ldots + m_r + n_r R^l, \qquad 0 \leqslant m_i, n_i < \tfrac{1}{2}p^{2/5},\ 1 \leqslant i \leqslant r$$

represents at least $\min(rv - r + 1, p)$ distinct residue classes (mod p), where $v = ([\frac{1}{2}p^{2/5}]+1)^2$.

It follows that provided $p > 5$,

$$\gamma(d, p) < p^{2/5}\frac{p-1}{v-1} < 2p^{2/5}\frac{p}{v} < 8p^{3/5},$$

which combined with the other two estimates for $\gamma(d, p)$ gives us that

$$\gamma(d, p) < 54p^{3/5} = 54d^{(3/5)(1+\varrho)},$$

since we are given that $p = d^{1+\varrho}$, $0 < \varrho < 1$.

We see that this result, which was obtained by elementary considerations, gives us an effective estimate for $\gamma(d, p)$ when p is not much greater than d, i.e. when ϱ is small. The principal difficulty arises when p is not much less than d^2, and we now proceed to discuss this case.

The underlying idea here is that $\gamma(d, p)$ is small providing the exponential sums $T(a)$ are all sufficiently small in modulus for every a prime to p, while if for some a, $|T(a)|$ is large, then there is some regularity in the distribution of the residues au (mod p). This regularity is exploited in developing the arguments which lead to the estimate for $\gamma(d, p)$. We begin with a lemma whose proof has some features in common with that of Lemma 1.

Lemma 3. *Suppose a is a non-zero residue class* (mod p) *such that*

$$|T(a)| > t(1-L).$$

Then the t residue classes au (mod p) *have representatives which lie in an interval of length $p\sqrt{L/2\lambda}$ with less than λt exceptions, where $0 < \lambda < 1$.*

Proof. For suitable real θ $(0 < \theta < p)$ we have

$$t(1-L) < |T(a)| = \sum_u e_p(au-\theta) = \sum_u \cos\left(2\pi(au-\theta)/p\right) < t,$$

whence

$$\sum_u \sin^2 \pi(au-\theta)/p < \tfrac{1}{2}tL.$$

Now suppose that m of the residue classes au (mod p) have representatives b say, which satisfy the inequality

$$\tfrac{1}{2}x \leqslant |b-\theta| \leqslant p - \tfrac{1}{2}x.$$

where $0 < x < p$. Then since $\sin y \geqslant 2y/\pi$ for $0 \leqslant y \leqslant \pi/2$, we have

$$\sin^2(\pi(au-\theta)/p) \geqslant \sin^2(\pi x/2p) \geqslant (x/p)^2$$

for m values of u. Hence

$$m(x/p)^2 < \tfrac{1}{2}tL$$

which implies

$$m < p^2 tL/2x^2$$

which on putting $\lambda = p^2 L/2x^2$ gives the desired result.

The following lemma is based on a dissection argument and the preceding result with $\lambda = 1/\log p$ is used repeatedly. For convenience we shall write

$$\delta = 3/(\log\log p).$$

Lemma 4. *Let p be sufficiently large and let a_0 be a non-zero residue* (mod p). *Suppose that*

$$|T(a_0)| > t(1-L).$$

Then there exists a positive number $Y \leqslant p\sqrt{\frac{1}{2}\log p \cdot L}$ and a positive integer $a < Yp^{\delta}/t$, such that at least $t/(2\log^2 p)$ of the residue classes au have representatives, b say, which satisfy

$$Y \leqslant |b| \leqslant 2Y.$$

Proof. Let $n = [\frac{1}{2}\log\log p]+1$ and take p large enough to ensure that $2^{2n+1} < \log p$ and note that $n\delta > 1$. Let $X = p\sqrt{\frac{1}{2}\log p \cdot L}$. Then by Lemma 7, at least $t\left(1 - \dfrac{1}{\log p}\right)$ of the numbers $a_0 u$ have representatives which lie in an interval of length X. It follows by a box argument that there exist two such residues, B_0 and B_0' say, such that

$$0 < B_0 - B_0' < X \Big/ \left[t\left(1 - \frac{1}{\log p}\right)\right] < 2X/t$$

and we note that there exist two dth power residues, v and v' say, such that

$$B_0 - B_0' \equiv a_0(v-v') \pmod p.$$

We write

$$a_1 = B_0 - B_0' < 2X/t.$$

Now since $T(a_0) = T(a_0 v) = T(a_0 v')$, Lemma 3 and the above remarks also apply to numbers of the form $a_0 uv$ and $a_0 uv'$ where u runs through t distinct dth power residues (mod p). Thus there are at least $t\left(1 - \dfrac{1}{\log p}\right)$ representatives of the residue classes $a_0 uv$ (mod p) and at least $t\left(1 - \dfrac{1}{\log p}\right)$

representatives of the residue classes $a_0uv' \pmod p$ which lie in the same interval of length X. Hence there are at least $t\left(1-\frac{2}{\log p}\right)$ residue classes $\pmod p$, which we will denote by b_1, in the interval $(-2X, 2X)$ which satisfy

$$b_1 \equiv a_0uv - a_0uv' \equiv a_1u \pmod p.$$

Write $Y_0 = 2X$. Dissect the interval $(0, Y_0)$ into $[2\log p]+1$ subintervals of the type $(2^{-m-1}Y_0, 2^{-m}Y_0)$ where $0 \leqslant m < 2\log p$, and the interval $(-Y_0, 0)$ into subintervals of the type $(-2^{-m}Y_0, -2^{-m-1}Y_0)$, where $0 \leqslant m < 2\log p$. None of the integers b_1 lies in the interval

$$(-2^{-[2\log p]-1}Y_0, 2^{-[2\log p]-1}Y_0)$$

since the length of this interval is

$$2^{-[2\log p]}Y_0 < 2^{-(2\log p)+1}2p\cdot\sqrt{\tfrac{1}{2}\log p\cdot L} < 1$$

for p sufficiently large, since $0 < L < 1$. Thus the numbers $|b_1|$ lie in less than $2\log p$ subintervals with less than $2t/\log p$ exceptions and so at least one of these subintervals contains at least $t\left(1-\frac{2}{\log p}\right)\Big/(\log p)^2$ of the numbers $|b_1|$. Let $(2^{-m_0-1}Y_0, 2^{-m_0}Y_0)$ be the longest such subinterval $(0 \leqslant m_0 < 2\log p)$, and write

$$Y_1 = 2^{-m_0-1}Y_0 \leqslant X,$$

so that at least $t\left(1-\frac{2}{\log p}\right)\Big/(\log p)^2$ of the b_1 satisfy the inequality

$$Y_1 \leqslant |b_1| \leqslant 2Y_1. \tag{10}$$

Now suppose that

$$2^{m_0+1} \leqslant p^\delta$$

so that

$$Y_1 = 2^{-m_0-1}Y_0 \geqslant p^{-\delta}Y_0.$$

Then

$$0 < a_1 < 2X/t = Y_0/t \leqslant Y_1p^\delta/t,$$

and in view of (10), the lemma follows on putting $Y = Y_1$ and $a = a_1$.

On the other hand, suppose

$$2^{m_0+1} > p^\delta$$

so that

$$Y_1 < p^{-\delta}Y_0.$$

Then it is evident from the definition of Y_1 that less than $2t\left(1-\frac{2}{\log p}\right)\Big/\log p$ of the b_1 satisfy

$$2Y_1 < |b_1| \leqslant Y_0.$$

Since there are at least $t\left(1-\frac{2}{\log p}\right)$ of the b_1 in the interval $(-Y_0, Y_0)$ there are at least

$$t\left(1-\frac{2}{\log p}\right) - 2t\left(1-\frac{2}{\log p}\right)\Big/\log p = t\left(1-\frac{2}{\log p}\right)^2 \geqslant t\left(1-\frac{4}{\log p}\right)$$

in $(-2Y_1, 2Y_1)$ and by a box argument, this implies that there exist two such residues, B_1 and B_1' say, such that

$$0 < B_1 - B_1' < 4Y_1\Big/\left[t\left(1-\frac{4}{\log p}\right)\right] < 8Y_1/t.$$

Define $a_2 = B_1 - B_1'$ and note that there exist two dth power residues $\pmod p$, v_1 and v_1' say, such that $B_1 \equiv a_1v_1 \pmod p$ and $B_1' \equiv a_1v_1' \pmod p$, whence $a_2 \equiv a_1(v_1 - v_1') \pmod p$.

Moreover at least $t\left(1-\frac{4}{\log p}\right)$ of the numbers a_1v_1u and at least $t\left(1-\frac{4}{\log p}\right)$ of the numbers $a_1v_1'u$ have residues $\pmod p$ which lie in the interval $(-2Y_1, 2Y_1)$ since our arguments still apply when we replace a_0 by a_1v_1 or by a_1v_1'. Hence there are at least $t\left(1-\frac{8}{\log p}\right)$ residues $\pmod p$, which we will denote by b_2, in the interval $(-4Y_1, 4Y_1)$ which satisfy

$$b_2 \equiv a_1(v_1 - v_1')u \equiv a_2u \pmod p.$$

As before let the largest interval of the form $(2^{-m-1}Y_1, 2^{-m}Y_1)$ where $-2 \leqslant m < 2\log p$ containing at least $t\left(1-\frac{4}{\log p}\right)\Big/(\log p)^2$ of the numbers $|b_2|$ be $(Y_2, 2Y_2)$, where $0 < Y_2 \leqslant 2Y_1$. Then at least, since p is sufficiently large

$$t\left(1-\frac{4}{\log p}\right) - 2(\log p+1)t\left(1-\frac{4}{\log p}\right)\Big/(\log p)^2 \geqslant t\left(1-\frac{8}{\log p}\right)$$

of the numbers b_2 lie in $(-2Y_2, 2Y_2)$. If $Y_2 \geqslant 8Y_1p^{-\delta}$ then

$$0 < a_2 < Y_2p^\delta/t \quad \text{and} \quad Y_2 \leqslant 2Y_1 \leqslant X$$

and the lemma is proved with $a = a_2$ and $Y = Y_2$.

If $Y_2 < 8Y_1 p^{-\delta}$ we repeat this process. Suppose that for each r $2 \leqslant r \leqslant n = [\frac{1}{2}\log\log p]+1$, we have at the rth stage,

$$Y_r \geqslant 8Y_{r-1}p^{-\delta} \tag{11}$$

where at least $t\left(1-\frac{2^{2r}}{\log p}\right)\Big/(\log p)^2 \geqslant t/2(\log p)^2$ of the residue classes $a_r u \pmod p$ have representatives $b_r \equiv a_r u \pmod p$ which satisfy

$$Y_r \leqslant |b_r| \leqslant 2Y_r$$

where

$$0 < ar < 4Y_{r-1}\Big/\left[t\left(1-\frac{2^{2r-1}}{\log p}\right)\right] < 8Y_{r-1}/t$$

and

$$Y_r \leqslant 2Y_{r-1} < 2\cdot 8Y_{r-2}p^{-\delta} < \ldots < 2\cdot 8^{r-1}p^{-\delta(r-1)}Y_0 \leqslant X$$

since $\delta = 3/\log\log p$ and $r \leqslant [\frac{1}{2}\log\log p]+1$. The lemma follows on putting $Y = Y_r$, $a = a_r$.

Thus it remains to show that the inequality (11) holds for some $r \leqslant n$. Assume the contrary, i.e. assume

$$Y_r < 8Y_{r-1}p^{-\delta}$$

for each $r = 2, \ldots, n$. Then in particular

$$Y_n < 8Y_{n-1}p^{-\delta}$$

and at least $t\left(1-\frac{2^{2n}}{\log p}\right) \geqslant \frac{1}{2}t$ of the residue classes $a_n u \pmod p$ have representatives which lie in the interval $(-2Y_n, 2Y_n)$. But

$$4Y_n < 4\cdot 8Y_{n-1}p^{-\delta} < 2^{3n-1}Xp^{-\delta n} < 1 \leqslant t/2,$$

by the choice of n and δ, which is a contradiction, and the lemma is proved.

We now give a series of lemmas which provide estimates for $\gamma(d, p)$ in special cases.

LEMMA 5. *Let $c > 1$ be congruent to the difference of two distinct non-zero d-th powers* (mod p). *Then*

$$\gamma(d, p) \leqslant 2c2^{\log p/\log c}.$$

Proof. We have to show that every integer N is congruent to a sum of at most $2c^{\log p/\log c}$ dth powers (mod p). Without loss of generality we

can suppose $0 \leqslant N < p$ and then we can express N as

$$N = a_0 + a_1 c + \ldots + a_h c^h,$$

where $0 \leqslant a_i < c$ for each $i = 0, 1, \ldots, h-1$, and $0 < a_h < c$. Plainly $c^h < p$.

Now we are given that $c \equiv u - v \pmod p$ where u, v are dth power residues (mod p). Hence we have

$$N \equiv \sum_{i=0}^{h} a_i c^i \equiv \sum_{i=0}^{h} a_i \sum_{j=0}^{i} (-1)^j \binom{i}{j} u^{i-j}v^j \pmod p,$$

i.e.

$$N \equiv -\sum_{i=0}^{h}(c-1)\sum_{\substack{j=1\\ j\,\text{odd}}}^{i}\binom{i}{j}u^{i-j}v^j + \sum_{i=0}^{h}(c-1-a_i)\sum_{\substack{j=1\\ j\,\text{odd}}}^{i}\binom{i}{j}u^{i-j}v^j + \sum_{i=0}^{h}a_i\sum_{\substack{j=0\\ j\,\text{even}}}^{h}\binom{i}{j}u^{i-j}v^j \pmod p.$$

We see that the first sum is a fixed residue class (mod p), while the other two sums consist of positive terms only which give rise to at most $(c-1)(2^{h+1}-1) < 2c2^h$ dth powers. Since $h < \log p/\log c$, the proof of the lemma is complete.

LEMMA 6. *Let c and r be integers >1 and suppose c or $-c$ is congruent to the sum of r d-th powers* (mod p). *Then*

$$\gamma(d, p) \leqslant 2cr^{\log p/\log c}.$$

Proof. We have to show that every integer N is congruent to a sum of at most $2cr^{\log p/\log c}$ dth powers (mod p). Without loss of generality we can assume $0 \leqslant N < p$ so that we can express N as

$$N = a_0 + a_1 c + \ldots + a_h c^h,$$

where $0 \leqslant a_i < c$ for each $i = 0, 1, \ldots, h-1$, and $0 < a_h < c$. Plainly $c^h < p$.

First suppose c is congruent to a sum of r dth powers (mod p). Then N is congruent to a sum of at most

$$(c-1)\sum_{i=0}^{h} r^i = (c-1)\frac{r^{h+1}-1}{r-1} < 2cr^h < 2cr^{\log p/\log c}$$

dth powers (mod p).

Next suppose $-c$ is congruent to a sum of r dth powers (mod p), i.e. suppose that

$$c \equiv -(u_1 + \ldots + u_r) \pmod p.$$

Then

$$N \equiv \sum_{\substack{i=0\\ i\,\text{even}}}^{h} a_i(u_1+\ldots+u_r)^i - \sum_{\substack{i=1\\ i\,\text{odd}}}^{h} a_i(u_1+\ldots+u_r)^i \pmod p$$

$$\equiv -(c-1)\sum_{\substack{i=1\\ i\,\text{odd}}}^{h}(u_1+\ldots+u_r)^i + \sum_{\substack{i=0\\ i\,\text{even}}}^{h} a_i(u_1+\ldots+u_r)^i +$$

$$+\sum_{\substack{i=1\\ i\,\text{odd}}}^{h}(c-1-a_i)(u_1+\ldots+u_r)^i \pmod p.$$

The first term on the right hand side of this congruence is a fixed residue class (mod p) and the other two terms consist of at most $(c-1)\sum_{i=0}^{h} r^i < 2cr^h$ dth power residues (mod p), whence the lemma.

LEMMA 7. *Let n be a positive integer and let $a_1, \ldots, a_l$ be l distinct residue classes* (mod n). *Let $b_1, \ldots, b_m$ be m distinct residue classes* (mod n), *one of which is* 0 *and the remainder prime to n. Then the number of distinct residue classes representable as*

$$a_i+b_j, \quad 1 \leqslant i \leqslant l,\ 1 \leqslant j \leqslant m,$$

is at least

$$\min(l+m-1, n).$$

This modification of the Cauchy–Davenport Theorem is due to I. Chowla ([2]) but a more convenient reference is Halberstam and Roth ([11], p. 49, Theorem 15).

LEMMA 8. *Let n be a positive integer and let $a_1, \ldots, a_l$ be l distinct residues* (mod p), *where $l < n$. Then the congruence*

$$x_1+\ldots+x_s \equiv 0 \pmod n,$$

where $x_1, \ldots, x_s$ are chosen from among the residues $a_1, \ldots, a_l$, repetitions being allowed, is soluble with

$$s \ll (n\log\log n)/l.$$

Proof. Suppose $a_1, \ldots, a_l$ are all prime to n. Then by Lemma 7 we can solve the congruence

$$x_1+\ldots+x_r \equiv 0 \pmod n,$$

providing $r \geqslant (n-1)/l+1$. Thus the least integer s for which this congruence is soluble satisfies

$$s < \left[\frac{n-1}{l}\right]+2 < \frac{3n}{l}.$$

Let $\nu(n)$ be the number of divisors of n and let d_i, $i = 1, \ldots, \nu(n)$, be the divisors of n. Let $b_{i_1}, \ldots, b_{i_{f_i}}$ be those f_i residues from $a_1, \ldots, a_l$ which satisfy $(b, n) = d_i$. Then by considering the congruence

$$x_1+\ldots+x_r \equiv 0 \left(\mathrm{mod}\,\frac{n}{d_i}\right)$$

where $x_1, \ldots, x_r$ are chosen from $b_{i_1}|d_i, \ldots, b_{i_{f_i}}|d_i$ we deduce that for each $i = 1, \ldots, \nu(n)$,

$$s < \frac{3n}{d_i f_i}. \tag{12}$$

Also it is plain that

$$\sum_{i=1}^{\nu(n)} f_i = l.$$

It is well known that $\sigma(n) < cn\log\log n$ where $\sigma(n)$ is the sum of the divisors of n and where c is an absolute constant ([13], Theorem 323). Hence we have

$$\sum_{i=1}^{\nu(n)} \frac{1}{d_i} = \frac{1}{n}\sigma(n) < c\log\log n.$$

Now for some $i, 1 \leqslant i \leqslant \nu(n)$, we must have

$$d_i f_i \geqslant l/(c\log\log n), \tag{13}$$

since otherwise

$$l = \sum_{i=1}^{\nu(n)} f_i < l\Big(\sum_{i=1}^{\nu(n)} 1/d_i\Big)\Big/(c\log\log n) < l,$$

which is impossible. The lemma follows on substituting the inequality (13) in the estimate (12).

LEMMA 9. *Let $0 < \beta < 1$ and let $p = d^{1+\varrho}$ where p is sufficiently large and where*

$$\tfrac{1}{2}(1+\beta)+2\delta+(3\log\log p)/\log d+(8\log 2)/\log d \leqslant \varrho \leqslant 1. \tag{14}$$

Suppose that at least $t/(2\log^2 p)$ of the residue classes au (mod p) *where u runs through the t non-zero d-th power residues* (mod p) *and where*

$$0 < a < Yp^{\delta}/t \tag{15}$$

have representatives, b say, such that

$$Y \leqslant |b| \leqslant 2Y,$$

where $Y \leqslant (p\log p)/d^{(1/2)(1-\beta)}$. Then

$$\gamma(d, p) \leqslant 2^{17}\cdot\log^3 d\cdot d^{(1+\beta)/2}.$$

Proof. At least one of the intervals $(Y, 2Y)$ and $(-2Y, -Y)$ must contain at least $t/(4\log^2 p)$ of the numbers b. Let us suppose that w of the numbers b lie in $(Y, 2Y)$ and that $w \geqslant t/(4\log^2 p)$; i.e. we suppose that there are $w \geqslant t/(4\log^2 p)$ numbers $b \equiv au \pmod p$ such that

$$Y \leqslant b \leqslant 2Y.$$

The inequalities (14) and (15) together with $Y \leqslant (p\log p)/d^{(1/2)(1-\beta)}$ imply that $a < t/(8\log p)^2 \leqslant w/16$. We deduce that there exists a residue class $\pmod a$ which contains at least $w/a \geqslant 16$ of the numbers b, say $b_1, \ldots, b_{l+1}$ where $l \geqslant [w/a]$ and where

$$Y \leqslant b_1 < b_2 < \ldots < b_{l+1} \leqslant 2Y.$$

Consider the $[l/16]$ positive integers

$$(b_{17} - b_1)/a, \ \ldots, \ (b_{16[l/16]+1} - b_{16[l/16-1]+1})/a.$$

Their sum is at most Y/a and each integer is at least 16, since a divides each $b_{i+1} - b_i$ $(1 \leqslant i \leqslant l)$. Hence the least such integer, h say, satisfies

$$16 \leqslant h < Y/a\,[l/16] < 32Y/al \leqslant 64Y/w < 256Y\,(\log p)^2/t$$
$$\leqslant 256\,(\log p)^3 p/(d^{(1/2)(1-\beta)}t) < 512\,(\log p)^3 d^{(1/2)(1+\beta)},$$

since $p = dt+1 < 2dt$. Also each of these integers is congruent $\pmod p$ to a difference of two dth powers, since we have $b_r = au_r + \lambda_r p$, $b_s = au_s + \lambda_s p$ for some integers λ_r, λ_s, whence $(b_r - b_s)/a \equiv u_r - u_s \pmod p$, where u_r, u_s, are dth power residues $\pmod p$. In particular h is congruent $\pmod p$ to two distinct non-zero dth powers, so that we can apply Lemma 5 which asserts that

$$\gamma(d, p) \leqslant 2hp^{\log 2/\log h}.$$

The right hand side of this estimate is for a fixed p, a decreasing function of h for $h \leqslant e^{\sqrt{\log 2 \cdot \log p}}$, and is an increasing function of h for $h \geqslant e^{\sqrt{\log 2 \cdot \log p}}$. Hence

$$\gamma(d, p) \leqslant \max\{32p^{1/4}, 2^{10}(\log p)^3 d^{(1/2)(1+\beta)} p^{\log 4/\log d}\}$$
$$\leqslant \max\{32d^{1/2}, 2^{17}(\log d)^3 d^{(1/2)(1+\beta)}\}$$

since $p < d^2$. It follows that

$$\gamma(d, p) \leqslant 2^{17}\,(\log d)^3 d^{(1/2)(1+\beta)}$$

which is the desired result.

Finally, if at least $t/(4\log^2 p)$ of the numbers b lie in the interval $(-2Y, -Y)$, i.e. if

$$-2Y \leqslant b_{l+1} < b_l < \ldots < b_1 \leqslant -Y$$

where $l \geqslant w/a$, then the result still holds. For we consider instead the integers

$$(b_1 - b_{17})/a, \ \ldots, \ (b_{16[l/16]-15} - b_{16[l/16]+1})/a,$$

which are positive with a sum of at most Y/a and where each integer is at least 16 and is congruent $\pmod p$ to a difference of two distinct non-zero dth powers.

Lemma 10. *Let $0 < \beta < 1$ and let $p = d^{1+\varrho}$ where $3/7 \leqslant \varrho \leqslant (1+\beta)/2$. Suppose that at least $t/(2\log^2 p)$ of the residue classes $au \pmod p$, where u runs through the t non-zero d-th power residues $\pmod p$ and where*

$$0 < a < Yp^{\delta}/t,$$

have representatives, b say, such that

$$Y \leqslant |b| \leqslant 2Y$$

where $Y \leqslant p\log p/d^{(1/2)(1-\beta)}$. Then for d sufficiently large we have

$$\gamma(d, p) \ll \log^4 d \cdot d^{(1/2)(1+\beta)+2\lambda_0(1+\varrho)^2/l+8\delta},$$

where $\lambda_0 = \left(\sqrt{2\varrho(1+\beta) - 3\varrho^2}\right) - \varrho\big)/2(1+\varrho)$.

Proof. We can choose at least $t/(4\log^2 p)$ of the numbers b, say $b_1, \ldots, b_w$ where $w \geqslant t/(4\log^2 p)$, to be of the same sign. Also we can suppose without loss of generality that each residue class $\pmod a$ contains less than 16 of the numbers $b_1, \ldots, b_w$, since otherwise arguing as in the preceding lemma we deduce that

$$\gamma(d, p) \leqslant 2^{17}\,(\log d)^3 d^{(1+\beta)/2}.$$

It follows that the numbers $b_1, \ldots, b_w$ must be distributed amongst at least $[w/16]$ different residue classes $\pmod a$. By applying Lemma 8 with $n = a$, $l = [w/16]$, we can solve the congruence

$$b_{i_1} + \ldots + b_{i_s} \equiv 0 \pmod a$$

where $1 \leqslant i_1 \leqslant \ldots \leqslant i_s \leqslant w$, with $s \ll a(\log p)^3/t$. We recall that $b_i \equiv au_i \pmod p$ whence there exists an integer c say such that

$$c = (b_{i_1} + \ldots + b_{i_s})/a \equiv u_{i_1} + \ldots + u_{i_s} \pmod p,$$

where $sY/a \leqslant |c| \leqslant 2sY/a$. Therefore by Lemma 6 we have

$$\gamma(d, p) < 2\,|c|\,s^{(\log p)/(\log|c|)} = 2\,|c|\,p^{(\log s)/(\log|c|)}. \tag{16}$$

Using the estimates for s, Y and a, we obtain

$$Ys/a \ll (\log d)^4 d^{(1/2)(1+\beta)},$$

whence if $s = 1$,

$$\gamma(d, p) < 4Ys/a \ll (\log d)^4 d^{(1/2)(1+\beta)}$$

and if $s \geqslant 2$

$$d^{\varrho-2\delta} < 2tp^{-\delta} < 2Y/a \leqslant Ys/a \ll (\log d)^4 d^{(1/2)(1+\beta)}. \tag{17}$$

Now let $s = p^\lambda \geqslant 2$ without loss of generality and suppose first that $0 \leqslant \lambda \leqslant \lambda_0$. Then we have

$$\gamma(d,p) \leqslant 4Ys \cdot p^{(\log s)/\log(Ys/a)} a^{-1} \ll (\log d)^4 d^{(1/2)(1+\beta)} p^{\lambda(1+\varrho)/(\varrho-2\delta)},$$

i.e.

$$\gamma(d,p) \ll (\log d)^4 d^{(1/2)(1+\beta)+\{\lambda_0(1+\varrho)^2/\varrho\}+8\delta} \tag{18}$$

for d sufficiently large.

Next suppose $\lambda_0 \leqslant \lambda \leqslant 1$, so that s satisfies

$$p^{\lambda_0} \leqslant p^\lambda = s \ll a(\log p)^3/t \ll (\log d)^4 d^{-(1/2)(1-\beta)} p^{1+\delta} t^{-2},$$

i.e. since $p = d^{1+\varrho} = dt+1$,

$$s \ll (\log d)^4 d^{(1/2)(1+\beta)-\varrho+2\delta}. \tag{19}$$

Also $Ys/a > tsp^{-\delta}$, whence

$$\log(Ys/a) \geqslant \{\varrho+(1+\varrho)\lambda-2\delta\}\log d$$

and providing d is large enough we have

$$(\log s)/(\log(Ys/a)) \leqslant \{(1+\beta-2\varrho)/2(\lambda+\varrho+\lambda\varrho)\}+4\delta.$$

Thus from the estimate (16) and since $\lambda \leqslant \lambda_0$, we get

$$\gamma(d,p) \ll (\log d)^4 d^{(1/2)(1+\beta)+\{(1+\varrho)(1+\beta-2\varrho)/2(\varrho+\lambda_0+\varrho\lambda_0)\}+8\delta}. \tag{20}$$

Now λ_0 was choosen so that

$$(1+\beta-2\varrho)/2(\varrho+\lambda_0+\varrho\lambda_0) = \lambda_0(1+\varrho)/\varrho,$$

and the lemma follows on comparing the estimates (18) and (20).

We see that Lemmas 9 and 10 do not provide estimates for $\gamma(d,p)$ when

$$\tfrac{1}{2}(1+\beta) \leqslant \varrho \leqslant \tfrac{1}{2}(1+\beta)+2\delta+3\log\log p/\log d+8\log 2/\log d. \tag{21}$$

However we deal with this case in the following lemma:

LEMMA 11. *Let $0 < \beta < 1$ and let $p = d^{1+\varrho}$ where ϱ satisfies (21). Suppose that at least $t/(2\log^2 p)$ of the residue classes $au \pmod p$ where u runs through the t non-zero d-th power residues $\pmod p$ and where*

$$0 < a < Yp^\delta/t,$$

have representatives, b say, such that

$$Y \leqslant |b| \leqslant 2Y,$$

where $Y \leqslant p\log p/d^{(1/2)(1-\beta)}$. Then for d sufficiently large

$$\gamma(d,p) \ll \log^4 d \cdot d^{(1+\beta)/2+12\delta}.$$

Proof. The estimates (16), (17) and (19) of the preceding lemma still hold, and we have

$$d^{\varrho-2\delta} \leqslant Ys/a \ll (\log d)^4 d^{(1/2)(1+\beta)},$$

and using (21),

$$s \ll (\log d)^4 d^{2\delta}.$$

It follows similarly that

$$\gamma(d,p) \ll (\log d)^4 d^{(1/2)(1+\beta)+12\delta},$$

since d is sufficiently large.

Now we are able to establish

THEOREM 1. *Suppose d is a positive integer strictly less than $\frac{1}{2}(p-1)$ and which divides $p-1$. Then for d sufficiently large*

$$\gamma(d,p) < \tfrac{1}{2} \cdot d^{7/8}.$$

Proof. First we recall that if $p > d^2$ we have

$$\gamma(d,p) \leqslant \max(3, 9\log d),$$

which plainly implies the theorem, and so from henceforth we suppose that $p < d^2$ and write

$$p = d^{1+\varrho}, \quad 0 < \varrho < 1.$$

Now choose ϱ_0 so that

$$\tfrac{3}{5}(1+\varrho_0) = \tfrac{1}{2}(1+\beta)+\lambda_0(1+\varrho_0)^2/\varrho_0 \tag{22}$$

and in view of the choice of λ_0 in Lemma 10, we also have that

$$\tfrac{3}{5}(1+\varrho_0) = \tfrac{1}{2}(1+\beta)+(1+\varrho_0)(1+\beta-2\varrho_0)/2(\varrho_0+\lambda_0+\varrho_0\lambda_0). \tag{23}$$

Next choose β_0 so that

$$1-\beta_0 = \tfrac{3}{5}(1+\varrho_0). \tag{24}$$

We remark that the exponent of d in the estimate for $\gamma(d,p)$ will be $1-\beta_0+\varepsilon$. Dropping the suffix $_0$ from λ_0 and β_0, we have

$$\varrho_0 = \tfrac{1}{3}(2-5\beta)$$

and

$$\lambda = \varrho_0(1-3\beta)/2(1+\varrho_0)^2.$$

Using (22) and (23) to eliminate λ and ϱ_0, we obtain the following cubic for β:

$$3(1-3\beta)(2-5\beta)(13-19\beta)+50(1-\beta)^2(1-13\beta) = 0. \tag{25}$$

This has one real root, B say, and B satisfies

$$1/8 < B < 1/7,$$

which implies that $3/7 < \varrho_0 < 11/24$ and $33/490 < \lambda < 3/25$.

Next choose K to be a constant greater than 2^{17} and greater than the constants implied in the estimates for $\gamma(d,p)$ in Lemmas 10 and 11. Then providing $0 \leqslant \varrho \leqslant \varrho_0 = (1/3)(2-5B)$, Lemma 2 gives

$$\gamma(d,p) < 54 \cdot d^{(3/5)(1+\varrho)} < K \cdot d^{1-B}.$$

So we take $\varrho_0 \leqslant \varrho \leqslant 1$ and assume that

$$\gamma(d,p) > K(\log d)^4 d^{1-B+12\delta}$$

where $\delta = 3/(\log\log p)$, and obtain a contradiction.

This assumption implies that there exists a non-zero residue class (mod p), a_0 say, such that

$$|T(a_0)| = \Big|\sum_u e_p(a_0 u)\Big| > t\{1-(2\log p)/d^{1-B}\},$$

since otherwise the corollary to Lemma 1 with $r = d^{1-B}$ tells us that

$$\gamma(d,p) \leqslant d^{1-B}.$$

We see that the conditions for the application of Lemma 4 are fulfilled with $L = (2\log p)/d^{1-B}$ and therefore we deduce that there exists a positive number $Y \leqslant (p\log p)/d^{(1/2)(1-B)}$ and a positive integer $a < Y \cdot p^{\delta}/t$ such that at least $t/(2\log^2 p)$ of the t residue classes au (mod p) have representatives, b say, which satisfy

$$Y \leqslant |b| \leqslant 2Y.$$

By comparing the estimates for $\gamma(d,p)$ in Lemmas 9, 10 and 11, we have that for $\varrho_0 \leqslant \varrho \leqslant 1$,

$$\gamma(d,p) \leqslant K(\log d)^4 d^{(1/2)(1+B)+\{\lambda(1+B)^2/\varrho_0\}+12\delta},$$

whence by the choice of λ, ϱ_0 and B,

$$\gamma(d,p) \leqslant K(\log d)^4 d^{1-B+12\delta}$$

which gives us the required contradiction. Since $B > 1/8$, the theorem follows on taking d sufficiently large.

The estimate for $\Gamma(k,p)$ now follows easily:

THEOREM 2. *Let k be a sufficiently large positive integer. Then for all primes p with $(k, p-1) < \frac{1}{2}(p-1)$, we have*

$$\Gamma(k,p) < k^{7/8}.$$

Proof. We recall that $(k, p-1) = d$ divides k and that $d < \frac{1}{2}(p-1)$ whence

$$\Gamma(k,p) = \Gamma(d,p) \leqslant d$$

and

$$\Gamma(d,p) \leqslant \gamma(d,p)+1.$$

Also we can suppose that $d > k^{7/8}$, since otherwise

$$\Gamma(k,p) = \Gamma(d,p) \leqslant d \leqslant k^{7/8}.$$

But by choosing k sufficiently large we can ensure that d is large enough for Theorem 1 to hold, so that we have

$$\Gamma(k,p) = \Gamma(d,p) \leqslant \gamma(d,p)+1 < d^{7/8} \leqslant k^{7/8},$$

since d divides k, and the theorem is proved.

Plainly this estimate could be made a little sharper and the exponent $7/8$ could be reduced slightly by determining the real root B of the cubic (25) more precisely. But since we are very likely a long way from the final answer, we leave the exponent in its present simple form.

2. The congruence $x_1^k + \ldots + x_s^k \equiv 0 \pmod p$. As in the previous section, $\theta(k,p) = \theta(d,p)$ and so again it suffices to consider $\theta(d,p)$, the least positive integer s such that the congruence

$$x_1^d + \ldots + x_s^d \equiv 0 \pmod p, \tag{26}$$

where $p-1$ is a multiple of d, has a nontrivial solution. Since the congruence (26) is a specialization of the congruence (1) it is immediate that

$$\theta(d,p) \leqslant \Gamma(d,p). \tag{27}$$

As before the cases $d = 1$ and $d = p-1$ are somewhat special: plainly

$$\theta(1,p) = 2$$

and since when $d = p-1$, the only values of $x^d \pmod p$ are 0 and 1, and since $\theta(d,p) \leqslant d+1$ by the Cauchy–Davenport Theorem, we have

$$\theta(p-1,p) = p.$$

Also when d divides $\frac{1}{2}(p-1)$ (as is always the case when k is odd), -1 is a dth power residue (mod p), whence clearly

$$\theta(d,p) = 2.$$

Moreover since $\Gamma(d,p) \leqslant \gamma(d,p)+1$, it follows from (6) and (27) that when $d^2 < p$,

$$\theta(d,p) < 8\log p + 2 < d^{1/2}$$

for d sufficiently large. Accordingly we can assume without loss of generality that $d^2 > p$ and that $1 < d < \frac{1}{2}(p-1)$ which is equivalent to $2 < t < p-1$, and as before we write

$$p = d^{1+\varrho}, \quad 0 < \varrho < 1.$$

If we can find a set of values of x^d whose sum is congruent to 0 (mod p) then we have an upper bound for $\theta(d, p)$. In fact the set of distinct non-zero values $y_1, \ldots, y_t$, where $t = (p-1)/d$, of $x^d \pmod p$ provide such a set, for by Fermat's Theorem, they are given by the roots of the congruence

$$y^t - 1 \equiv 0 \pmod p$$

and the sum of the t roots of this congruence is congruent to 0 (mod p), i.e.

$$y_1 + \ldots + y_t \equiv 0 \pmod p.$$

Thus by its definition,

$$\theta(d, p) \leqslant t = (p-1)/d < d^{\varrho}, \tag{28}$$

an estimate which is effective when d is near p or when t is composite. For if $t = t_1 t_2$, $t_1 > 1$, $t_2 > 1$, then the dth power residues (mod p) satisfy

$$y^{t_1 t_2} - 1 \equiv 0 \pmod p,$$

so that there exist t_1 dth power residues $z_1, \ldots, z_{t_1}$ say, which satisfy

$$z^{t_1} - 1 \equiv 0 \pmod p$$

and hence which satisfy

$$z_1 + \ldots + z_{t_1} \equiv 0 \pmod p.$$

Clearly we can choose $t_1 \leqslant t^{1/2}$ and so if t is composite,

$$\theta(d, p) \leqslant t^{1/2} < d^{1/2},$$

since $d > p^{1/2}$. Therefore if t is not an odd prime greater than $d^{1/2}$ then S. Chowla's conjecture holds.

Here we show that when $d < p-1$ and d is sufficiently large

$$\theta(d, p) < d^{2/3+\varepsilon}$$

by combining (28) with the estimates for $\gamma(d, p)$ used in Theorem 1. Instead of the equations (22) and (23), we seek to solve the equations

$$\begin{aligned} \varrho_0 = \tfrac{1}{2}(1+\beta) &= \lambda(1+\varrho_0)^2/\varrho_0 \\ &= \tfrac{1}{2}(1+\beta) + (1+\varrho_0)(1+\beta-2\varrho_0)/2(\varrho_0+\lambda+\varrho_0\lambda) \end{aligned} \tag{29}$$

with β chosen so that

$$\varrho_0 = 1-\beta. \tag{30}$$

We note that the term ϱ_0 on the left hand side of equations (29) and (30) corresponds to the exponent of d in the estimate (28) for $\theta(d, p)$. On eliminating λ and ϱ_0, we obtain the following cubic for β:

$$(1-3\beta)[5(1-\beta)^2 + 2(2-\beta)^2] = 0.$$

Evidently this has one real root, namely 1/3, whence $\varrho_0 = 2/3$ and $\lambda = 0$. The significance of the vanishing of λ is that the improved version obtained above of I. Chowla's estimates is no more effective in this case than his original estimates. We can suppose without loss of generality that $2/3 \leqslant \varrho \leqslant 1$ since otherwise by (28) $\theta(d, p) \leqslant d^{\varrho} \leqslant d^{2/3} < d^{2/3+\varepsilon}$. Let K be the constant introduced in Theorem 1 and assume that

$$\theta(d, p) > K(\log d)^4 d^{2/3+12\delta} + 1$$

where $\delta = 3/(\log\log p)$. Then it follows from (27) that

$$\gamma(d, p) > K(\log d)^4 d^{2/3+12\delta},$$

which implies exactly as in Theorem 1, that the conditions for Lemma 4 are fulfilled with $L = (2\log p)/d^{2/3}$, which in turn implies that the estimates for $\gamma(d, p)$ in Lemmas 9 and 11 hold. Thus for $\frac{2}{3} \leqslant \varrho \leqslant 1$,

$$\gamma(d, p) \leqslant \max\left(2^{17}(\log d)^3 d^{2/3}, K(\log d)^4 d^{2/3+12\delta}\right)$$

which is a contradiction. Thus we must have for $1 < d \leqslant \frac{1}{2}(p-1)$

$$\theta(d, p) \leqslant K(\log d)^4 d^{2/3+12\delta} + 1,$$

so that if d is sufficiently large

$$\theta(d, p) < d^{2/3+\varepsilon},$$

and we have

THEOREM 3. *Let $\varepsilon > 0$ be given and let p be prime such that $(k, p-1) < p-1$. Then for k sufficiently large, we have*

$$\theta(k, p) < k^{2/3+\varepsilon}.$$

Proof. We recall that as $d < p-1$ by hypothesis,

$$\theta(k, p) = \theta(d, p) \leqslant d,$$

and take $d \geqslant k^{2/3}$ since otherwise $\theta(k, p) < k^{2/3}$. Thus by choosing k sufficiently large, we can ensure that

$$\theta(k, p) < d^{2/3+\varepsilon} \leqslant k^{2/3+\varepsilon},$$

which gives us the theorem.

3. The congruence $a_1x_1^k+\ldots+a_sx_s^k\equiv 0 \pmod p$. We recall that $a_1, \ldots, a_s$ are arbitrary integers not divisible by p. Arguing as in the previous section, we can confine ourselves to studying $\gamma^*(d, p)\,(=\gamma^*(k, p))$, where d divides $p-1$. The case $d=1$ is trivial and plainly

$$\gamma^*(1, p) = 2$$

and more generally it is easily proved using addition of residue classes ([9], p. 167, Lemma 2.3.1) that

$$\gamma^*(d, p) \leqslant d+1,$$

while in the case $d=p-1$,

$$\gamma^*(p-1, p) = p.$$

Actually the above estimate can be improved by using a theorem due to Chowla, Mann and Straus ([6]) but the estimate just given is sufficient for our purposes. Thus for the rest section we assume unless otherwise stated that $p>2$ and $1<d<p-1$. Further, when $p>d^2$ the use of exponential sums gives ([9], p. 172, Lemma 2.5.5 and p. 168, Lemma 2.4.1) that

$$\gamma^*(d, p) \leqslant \max(3, 48\log d+1)+1 < d^{2/3}$$

for d sufficiently large.

When d divides $\frac{1}{2}(p-1)$ (as is always the case when k is odd), -1 is a dth power residue (mod p) and a simple box argument gives ([9], p. 166, Lemma 2.2.1)

$$\gamma^*(d, p) < [(\log p)/(\log 2)]+1 < d^{2/3}$$

for d large enough. In general such arguments are not available but there is an argument (the possibility of which was suggested to me by Dr. Erdös through Professor Davenport) which applies without the hypothesis that d divides $\frac{1}{2}(p-1)$ and only requires that $t>1$, which has been assumed already.

Suppose we can find $\theta(d, p)$ disjoint sets of coefficients

$$a_1, \ldots, a_{r_1};\ a_{r_1+1}, \ldots, a_{r_2};\ \ldots;\ a_{r_{\theta-1}+1}, \ldots, a_{r_\theta},$$

where $\theta = \theta(d, p)$, such that their sums are all mutually congruent (mod p). Then we can solve the congruence

$$a_1x_1^d+\ldots+a_sx_s^d \equiv 0 \pmod p$$

non-trivially by taking

$$x_i \equiv x_j' \pmod p \quad \text{for} \quad r_{j-1} < i \leqslant r_j \ (j=1, \ldots, \theta)$$

where $(x_1', \ldots, x_\theta')$ is a non-trivial solution of (26) and where $r_0=0$ and by taking

$$x_i = 0 \quad \text{for} \quad i > r_\theta.$$

The possibility of finding $\theta(d, p)$ such sets of coefficients is guaranteed by a purely combinatorial theorem of Erdös and Rado ([10], Theorem III). It is easily seen that the results of § 3.3 ([9]) still hold when t is replaced by $\theta=\theta(d, p)$ and when considered (mod p). In particular we get Lemma 3.3.3 there in the following modified form:

LEMMA 12. *Suppose* $1<d<p-1$ *and let*

$$r = [(\log pt)/(\log 4)]+2.$$

Then

$$\gamma^*(d, p) \leqslant r^2\theta(d, p)+r.$$

Hence we have

THEOREM 4. *Let* $\varepsilon>0$ *be given and let* k *be a sufficiently large positive integer. Let* p *be a prime with* $d=(k, p-1)<p-1$. *Then*

$$\gamma^*(k, p) < k^{2/3+\varepsilon}.$$

The proof is similar to Theorem 3. We can suppose that $d>\frac{1}{2}k^{2/3}$ since otherwise

$$\gamma^*(k, p) = \gamma^*(d, p) \leqslant d+1 \leqslant \tfrac{1}{2}k^{2/3}+1 < k^{2/3}$$

and then if k is large enough it follows readily from Theorem 3 and Lemma 12 that we can ensure that

$$\gamma^*(k, p) = \gamma^*(d, p) < r^2d^{2/3+\varepsilon/2}+r \leqslant k^{2/3+\varepsilon},$$

since we can assume $p<d^2$, so that $r<3\log k$.

In conclusion we note that this improved estimate for $\gamma^*(k, p)$ implies a sharper estimate for the related function $\Gamma^*(k, p)$, which is defined to be the least value of s for which the congruence

$$a_1x_1^k+\ldots+a_sx_s^k \equiv 0 \pmod{p^n}, \tag{31}$$

where $a_1, \ldots, a_s$ are arbitrary non-zero integers, has a non-trivial solution for every positive integer n, for the particular prime p. In the author's paper ([9]), it is shown that (Lemma 4.4.2) if $d<p-1$ and $k\geqslant 7$, then

$$\Gamma^*(k, p) < 12(\log k)^2k^{15/8}$$

so that if k is sufficiently large we have

$$\Gamma^*(k, p) < k^{15/8+\varepsilon}.$$

In fact it is shown that except for the case when p does not divide k,

$$\Gamma^*(k,p) < 12(\log k)^2 k^{3/2}.$$

But if p does not divide k, then (loc. cit., equation (4.2.2))

$$\Gamma^*(k,p) \leqslant k\{\gamma^*(k,p)-1\}+1,$$

so that by Theorem 4, if k is large enough

$$\Gamma^*(k,p) < k^{5/3+\varepsilon}.$$

Collecting these results we have

THEOREM 5. *Let $\varepsilon > 0$ be given and let p be a prime such that $p-1$ does not divide k. Then for k sufficiently large*

$$\Gamma^*(k,p) < k^{5/3+\varepsilon}.$$

The function $\Gamma^*(k)$ is defined to be the least s such that the congruence (31) has non-trivial solutions for every prime power p^n. Thus

$$\Gamma^*(k) = \operatorname*{Max}_p \Gamma^*(k,p).$$

Recently A. Tietäväinen [16] has shown very elegantly that for all sufficiently large odd k,

$$\Gamma^*(k) < (1+\varepsilon)k\log k/(\log 2).$$

Such an estimate cannot hold for $\Gamma^*(k)$ in general ([9], p. 200, § 5.2), but by adapting the proof of Theorem 5.4.2 ([9], p. 205) it can be readily verified that Theorem 5 implies that there are infinitely many even k such that

$$\Gamma^*(k) < k^{5/3+\varepsilon}.$$

Finally let $\Gamma(k,p^n)$ be the least s such that the congruence

$$x_1^k + \ldots + x_s^k \equiv N \pmod{p^n},$$

where N is any integer, always has a non-trivial solution; i.e. $\Gamma(k,p^n)$ is the least s such that every integer N has a non-trivial representation as a sum of s kth powers in GF$[p^n]$. It is likely that I. Chowla's results ([4], § 7, 8) can be adapted to establish that if $p>3$ and $(k,p-1) < \frac{1}{2}(p-1)$, then for all $n \geqslant 1$,

$$\Gamma(k,p^n) < k^{7/8}$$

for k sufficiently large. In view of Theorem 2, it is well-known that only the case when p divides k needs to be considered. If this result is true, then it would follow that for large enough k, any integer N can be represented non-trivially as a sum of less than $k^{7/8}$ kth powers in every p-adic field with $(k,p-1) < \frac{1}{2}(p-1)$ and $p>3$.

I am very grateful to the late Professor H. Davenport for his advice and encouragement. Also I am indebted to the referee for pointing out a number of mistakes and suggesting some improvements.

References

[1] A. L. Cauchy, *Recherches sur les nombres*, J. Ec. Polytech. 9 (1813), pp. 99–116.

[2] I. Chowla, *A theorem on the addition of residue classes*, Proc. Indian Acad. Sci. 2 (1935), pp. 242–243.

[3] — *On the number of solutions of some congruences in two variables*, Proc. Indian Nat. Acad. Sci. A, 5 (1937), pp. 40–44.

[4] — *On Waring's Problem* (mod p), Proc. Indian Nat. Acad. Sci. A, 13 (1943), pp. 195–200.

[5] S. Chowla, *Proceedings of the 1963 Number Theory Conference*, University of Colorado, Boulder, Colorado, 1963.

[6] — H. B. Mann and E. G. Straus, *Some applications of the Cauchy-Davenport theorem*, K. Norske Vidensk Selsk. Forh. 32 (13) (1959), pp. 74–80.

[7] H. Davenport, *On the addition of residue classes*, Journ. London Math. Soc. 10 (1935), pp. 30–32.

[8] — *A historical note*, Journ. London Math. Soc. 22 (1947), pp. 100–107.

[9] M. Dodson, *Homogeneous additive congruences*, Philos. Trans. Roy. Soc. London, Ser. A, 261 (1967), pp. 163–210.

[10] P. Erdös and R. Rado, *Intersection theorems for systems of sets*, Journ. London Math. Soc. 35 (1960), pp. 85–90.

[11] H. Halberstam and K. Roth, *Sequences*, vol. I, Oxford 1966.

[12] G. H. Hardy and J. E. Littlewood, *Some problems of 'Partitio Numerorum': VIII. The number $\Gamma(k)$ in Waring's Problem*, Proc. London Math. Soc. 28 (1927), pp. 518–542.

[13] — and E. M. Wright, *An Introduction to the Theory of Numbers*, Oxford 1960.

[14] H. Heilbronn, *Lecture notes on additive number theory* mod p, California Institute of Technology, 1964.

[15] A. Tietäväinen, *On a homogeneous congruence of odd degree*, Ann. Univ. Turk., A, 121 (1969), pp. 3–6.

[16] — *On a problem of Chowla and Shimura*, to appear in the Journal of Number Theory.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF YORK

Received on 19. 3. 1970 (63)