ACTA ARITHMETICA
107.2 (2003)

The difficulty of the local solubility problem for additive equations

by

Trevor D. Wooley (Ann Arbor, MI, and Cambridge, MA)

Dedicated to Donald J. Lewis, in celebration of his 75th birthday

1. Introduction. The investigation of additive equations dates from at least the time of Gauss, who investigated quadratics, cubics and quartics. More recent progress concerning the solubility of systems of additive equations stems from a fundamental series of papers by Davenport and Lewis spanning the 1960's (see in particular [11–14]). When $k \in \mathbb{N}$ and p is a prime number, denote by $\Gamma^*(k;p)$ the least number s_0 with the property that whenever $s \geq s_0$, and $a_1, \ldots, a_s \in \mathbb{Q}_p$, then the additive equation

$$a_1 x_1^k + \ldots + a_s x_s^k = 0 \tag{1.1}$$

possesses a solution $\mathbf{x} \in \mathbb{Q}_p^s \setminus \{\mathbf{0}\}$. Also, define $\Gamma^*(k)$ to be the supremum, over prime numbers p, of $\Gamma^*(k;p)$. Then Davenport and Lewis [11] established that for each natural number k, one has $\Gamma^*(k) \leq k^2+1$. Indeed, equality holds here whenever $k = p-1$ for some prime number p. However, there are comparatively few primes p for which so many variables are required to guarantee the solubility of the associated equations. Thus, although a simple argument demonstrates that for each natural number k exceeding 1, one has $\Gamma^*(k;p) > 2k$ for infinitely many primes p, it is essentially a classical result that whenever $p > k^4$, one has $\Gamma^*(k;p) \leq 2k+1$ (see, for example, the introduction of Atkinson and Cook [4] for a discussion of this matter). A number of authors, moreover, have provided analogous conclusions for systems of additive equations (see Cook [7], Atkinson, Brüdern and Cook [1–3], Meir

2000 *Mathematics Subject Classification*: Primary 11D79.
Key words and phrases: diagonal equations, p-adic solubility.
The work presented in this paper originated in a project supported by a Rackham Faculty Fellowship at the University of Michigan in 1992, and was completed with support from NSF grant DMS-9970440 and a Fellowship from the David and Lucile Packard Foundation.

[20]), and indeed non-trivial computational effort has also been expended on the problem of calculating $\Gamma^*(k;p)$ explicitly for smaller exponents k (see, in particular, Cook [8, 9] and Atkinson and Cook [4]).

Motivated by such work, and especially the latter computations, the object of this paper is to provide estimates which, as k varies, provide some indication of the number of "exceptional" primes p for which $\Gamma^*(k;p) > 2k+1$, and also to measure the complexity of the task of cataloguing the corresponding insoluble equations of the shape (1.1). One could loosely describe our results as measuring the "difficulty" of the local solubility problem for additive equations. Our methods and results extend naturally to corresponding problems concerning the solubility of systems of additive equations. Although such extensions are not inherently difficult, they lead to technical and expository complications which obscure the themes and ideas that we wish to highlight herein, and thus we defer any consideration of such issues to a possible future occasion.

We begin by discussing the existence of *exceptional primes* p with

$$\Gamma^*(k;p) > 2k+1.$$

In order to set this discussion in context, we first observe that whenever p is a prime number with $p \equiv 1 \pmod{k}$, and q is not a kth power residue modulo p, then the congruence $x^k \equiv qy^k \pmod{p}$ has only the trivial solution with $x \equiv y \equiv 0 \pmod{p}$. It follows that the equation

$$\sum_{i=0}^{k-1} p^i(x_i^k - qy_i^k) = 0$$

has only the trivial solution $\mathbf{x} = \mathbf{y} = \mathbf{0}$ over $\mathbb{Q}_p$, whence for infinitely many primes p one has $\Gamma^*(k;p) > 2k$. Our first theorem, which we establish in Section 2, provides a simple criterion for the existence of primes p with $\Gamma^*(k;p) > tk$.

THEOREM 1.1. *Let t, δ and k be natural numbers with $t \geq 3$ and $\delta \mid k$. Suppose that p is a prime number with $(p-1,k) = \delta$. Then the following conclusions hold*:

(i) *Whenever $\delta > 4^{t-2}$ and*

$$p \leq \delta^{(t-1)/(t-2)} - 3\delta + 1,$$

one has $\Gamma^*(k;p) \geq tk+1$.

(ii) *Define $\delta_0(t)$ and $p_0(\delta,t)$ for each natural number t with $t \geq 3$ by*

$$\delta_0(t) = \begin{cases} (\sqrt{2})^{-1}18^{t/2-1} & \text{when } t \text{ is odd}, \\ 18^{t/2-1} & \text{when } t \text{ is even}, \end{cases}$$

and

$$p_0(\delta,t)=\begin{cases}(\sqrt{2}\,\delta)^{(t-1)/(t-2)}-4\delta+1 & \text{when } t \text{ is odd},\\ \sqrt{2}\,\delta^{(t-1)/(t-2)}-4\delta+1 & \text{when } t \text{ is even}.\end{cases}$$

Then whenever $\delta\ge\delta_0(t)$, *and the prime* p *satisfies* $p\equiv\delta+1 \pmod{2\delta}$ *and* $p\le p_0(\delta,t)$, *one has* $\Gamma^*(k;p)\ge tk+1$.

A less explicit version of Theorem 1.1(i) appears in Theorem 3.25 of Norton [23], with explicit bounds for smaller values of t appearing in Corollary 3.28 of the latter technical report.

At present, our knowledge concerning the distribution of prime numbers in arithmetic progressions is unfortunately insufficient to guarantee that for a given large number δ, there exists a prime number p, in a given arithmetic progression modulo δ, with $p<2\delta^2$. Thus, although the existence of such primes is widely anticipated, the conclusion of Theorem 1.1 for individual exponents k remains a conditional result. However, by employing relatively recent versions of the Bombieri–Vinogradov theorem due to Fouvry [16] and Bombieri, Friedlander and Iwaniec [5], one may successfully establish the desired conclusion in an average sense. In order to be precise, we require some notation. When K is a large real number and r is a natural number, denote by $E_r(K)$ the number of exponents k, with $1\le k\le K$, for which $\Gamma^*(k)\le(r+1)k$. Let $\mathfrak{P}_r(k)$ denote the set of exceptional primes p for which $\Gamma^*(k;p)>rk+1$, and write $P_r(k)=\operatorname{card}(\mathfrak{P}_r(k))$. Define also

$$p_r^{\max}(k)=\max_{p\in\mathfrak{P}_r(k)}p,$$

when $\mathfrak{P}_r(k)\ne\emptyset$, and otherwise take $p_r^{\max}(k)=-\infty$. Then in Section 3 we extract the following consequences from Theorem 1.1.

Corollary 1. *For each large number* K, *one has*

$$E_2(K)\ll K\biggl(\frac{\log\log K}{\log K}\biggr)^2.$$

It follows from Corollary 1, in particular, that for almost all exponents k, there is some prime number p with $\Gamma^*(k;p)>3k$. Thus "exceptional" primes are essentially ubiquitous.

Corollary 2. *For every positive number* ε, *and for almost all exponents* k, *one has*

$$p_2^{\max}(k)>k^2\,(1-1/(\log k)^{2-\varepsilon}).$$

Thus we find that "large" exceptional primes exist for almost all exponents.

Corollary 3. *For every positive number ε, and for almost all exponents k, one has*

$$P_2(k) > \frac{k}{\log k}\,(1 - 1/(\log k)^{2-\varepsilon}).$$

We may conclude from Corollary 3 that the set of "exceptional" primes is almost always somewhat large.

So far as upper bounds for $p_r^{\max}(k)$ are concerned, the most direct argument involves applying Weil's bound for the number of points on the hypersurface defined by the equation

$$a_0x_0^k + a_1x_1^k + \ldots + a_rx_r^k = 0, \tag{1.2}$$

for fixed $a_i \in \mathbb{F}_p^\times$ $(0 \le i \le r)$, over the finite field $\mathbb{F}_p$ (see Weil [25], though earlier work of Davenport and Hasse [10] establishes the same conclusion when $r = 2$). Such an approach shows that $\Gamma^*(k;p) \le rk + 1$ whenever

$$p^r > (1 - 1/k)((k-1)^r - (-1)^r)p^{(r-1)/2}(p-1), \tag{1.3}$$

and from this one immediately obtains the upper bound

$$p_r^{\max}(k) < k^{2r/(r-1)} \tag{1.4}$$

(see, for example, Theorem 1 of Meir [20]). Indeed, the lower bound (1.3) permits one to establish an upper bound somewhat sharper than

$$p_2^{\max}(k) < (k-1)^2(k-2)^2. \tag{1.5}$$

We remark here that all of these questions are well-understood for $k = 2$, and so there is no loss in supposing throughout that $k \ge 3$. In Section 4, as a consequence of an argument designed to bound the number of insoluble additive equations of degree k, we obtain a modest refinement of the upper bound (1.4).

Theorem 1.2. *For every natural number k, and for each integer r with $r \ge 2$, one has $p_r^{\max}(k) < \frac{1}{2}k^{2r/(r-1)}$.*

It transpires that our methods are rather more effective when the underlying prime p and exponent k satisfy the condition that -1 is a kth power residue modulo p.

Theorem 1.3. *Suppose that k is a natural number, and that r is an integer with $r \ge 2$. Suppose also that p is a prime number, and write $\delta = (p-1,k)$. Then whenever $p \equiv 1 \pmod{2\delta}$ and*

$$p \ge (r!)^{-1/(r-1)}\delta^{2r/(r-1)},$$

one has $\Gamma^*(k;p) \le rk + 1$.

The ideas underlying the proofs of Theorems 1.2 and 1.3 may be applied to provide, under most circumstances, estimates for the number of

$\mathbb{F}_p$-rational points on hypersurfaces of the shape (1.2) superior to those arising directly from Weil's methods. We discuss such issues in Section 4. When it comes to bounding $p_r^{\max}(k)$ for larger values of r, the methods of Weil are superseded by arguments stemming from Stepanov's methods. In Section 4, we exploit recent results of Heath-Brown and Konyagin [19] so as to improve on the bound provided by Theorem 1.2 for $r \geq 4$.

THEOREM 1.4. *For every natural number k, and for each integer r with $r \geq 2$, one has*

$$p_r^{\max}(k) \ll \min\{k^{(5r+3)/(3r-3)}, k^{(3r+5)/(2r-2)}\}.$$

The bound provided by Theorem 1.4 is superior to that of Theorem 1.2 for $r \geq 4$, so long as k is sufficiently large. The second estimate implicit in Theorem 1.4 supersedes the first only for $r \geq 10$.

We refer to the primes p with $p \mid k$ as *singular primes* (for the exponent k), and those with $p \nmid k$ as *regular primes* (for the exponent k). An application of the prime number theorem reveals that there are at most $O(\log k/\log\log k)$ singular primes for each large exponent k, and thus singular primes are comparatively few in number. By combining this observation with the conclusions of Theorems 1.2 and 1.4, one easily obtains the upper bounds for $P_r(k)$ recorded below.

THEOREM 1.5. *For every natural number k, and for each integer r with $2 \leq r \leq 8$, one has*

$$P_r(k) \ll \begin{cases} k^{(r+1)/(r-1)} \log\log k/\log k & \text{when } 2 \leq r \leq 3, \\ k^{(2r+6)/(3r-3)} \log\log k/\log k & \text{when } 4 \leq r \leq 8. \end{cases}$$

When $r \geq 9$, meanwhile, one has the upper bound

$$P_r(k) \ll k^{(r+7)/(2r-2)+\varepsilon},$$

valid for each positive number ε.

We turn our attention now to the anticipated discussion concerning the complexity of the task of cataloguing the insoluble equations of the shape (1.1). Since this project is at present too ill-defined to permit the announcement of our conclusions, we first introduce some notation and conventions with which to make sense of the problem. Unfortunately, such issues consume a fair amount of space, but this seems unavoidable.

Loosely speaking, our idea is to describe a reduction procedure that associates to an arbitrary diagonal form

$$\Phi(\mathbf{x}) = a_1 x_1^k + \ldots + a_s x_s^k, \tag{1.6}$$

a related form

$$\Omega(\mathbf{y}) = \alpha_1 y_1^\delta + \ldots + \alpha_t y_t^\delta,$$

for suitable t, δ and $\boldsymbol{\alpha}$, with the property that the form $\Phi(\mathbf{x})$ possesses a non-trivial p-adic zero if and only if $\Omega(\mathbf{y})$ does not belong to a certain finite list of exceptional additive forms. Our objective is to determine the number of exceptional (insoluble) additive forms $\Omega(\mathbf{y})$ that exist with a given number, t, of variables.

Before proceeding further, we briefly discuss the reduction procedure alluded to above. Given additive forms $\Phi(\mathbf{x})$, as in (1.6), and

$$\Psi(\mathbf{y}) = b_1 y_1^k + \ldots + b_s y_s^k, \tag{1.7}$$

we say that $\Phi(\mathbf{x})$ and $\Psi(\mathbf{y})$ are *p-equivalent*, and we write $\Phi \sim \Psi$, when there exist integers $v, u_1, \ldots, u_s$ with the property that

$$\Psi(\mathbf{x}) = p^v \Phi(p^{u_1} x_1, \ldots, p^{u_s} x_s). \tag{1.8}$$

The notion of p-equivalence of additive forms yields an equivalence relation, and moreover an additive form $\Phi(\mathbf{x})$ possesses a non-trivial p-adic solution if, and only if, every member of the equivalence class containing $\Phi(\mathbf{x})$ does so, as is evident from (1.8). In the first step of our reduction, given the form $\Phi(\mathbf{x})$ which is of central interest to us, we determine a p-equivalent form $\Psi(\mathbf{x})$ of the shape (1.7) that is representative of the equivalence class containing $\Phi(\mathbf{x})$, this representative being selected by means of the Davenport–Lewis p-normalisation procedure (see Davenport and Lewis [11]). The most succinct description of this process for the problem at hand employs the function $\partial(\Psi)$ defined on the coefficients $\mathbf{b}$ of Ψ by $\partial(\Psi) = b_1 \ldots b_s$. We begin by noting that if any coefficient a_i of $\Phi(\mathbf{x})$ is zero, then plainly $\Phi(\mathbf{x})$ possesses a non-trivial p-adic zero, and we declare that $\Phi(\mathbf{x})$ is not associated with any form on the list of exceptional forms. Otherwise, when all of the coefficients of $\Phi(\mathbf{x})$ are non-zero, we consider any form $\Phi^*(\mathbf{x}) = a_1^* x_1^k + \ldots + a_s^* x_s^k$, with $a_i^* \in \mathbb{Z}_p$ $(1 \le i \le s)$, satisfying the property that

$$|\partial(\Phi^*)|_p = \max_{\Psi \sim \Phi} |\partial(\Psi)|_p,$$

where the maximum is taken over $\Psi(\mathbf{x}) \in \mathbb{Z}_p[\mathbf{x}]$, and $|\cdot|_p$ denotes the usual p-adic valuation, normalised with $|p|_p = p^{-1}$. Here we note that the discreteness of the p-adic valuation implies that the maximum does indeed exist. Moreover, the theory developed by Davenport and Lewis (see [11]) ensures that when $\Psi(\mathbf{x}) \in \mathbb{Z}_p[\mathbf{x}]$, and $\Psi \sim \Phi$ satisfies $|\partial(\Psi)|_p = |\partial(\Phi^*)|_p$, then necessarily $\Psi(\mathbf{y})$ takes the shape (1.7) with no coefficient b_i divisible by p^k, and in which the number of coefficients not divisible by p^j is at least $\lceil js/k \rceil$ for $1 \le j \le k$. With only modest contemplation, one may engineer simple algorithms to accomplish the goal of reducing Φ to the described shape Φ^*.

Next, for each prime number p, we define $\tau = \tau(p,k)$ by means of the relation $p^\tau \parallel k$, by which we mean that $p^\tau \mid k$ but $p^{\tau+1} \nmid k$. We then define

$\gamma = \gamma(p,k)$ by

$$\gamma(p,k) = \begin{cases} \tau + 2 & \text{when } p = 2 \text{ and } \tau > 0, \\ \tau + 1 & \text{otherwise.} \end{cases}$$

By making use of a suitable version of Hensel's lemma, the Davenport–Lewis theory of additive equations ensures that the additive form $\Phi^*(\mathbf{y})$ possesses a non-trivial p-adic solution whenever the congruence $\Phi^*(\mathbf{z}) \equiv 0 \pmod{p^\gamma}$ possesses a solution in which, for some i with $1 \le i \le s$, one has $(a_i^* z_i, p) = 1$. Moreover, Davenport and Lewis [11] show that such is guaranteed whenever $s > k^2$.

In the second step of our reduction, we relabel variables so that the coefficients of $\Phi^*(\mathbf{x})$ satisfy the condition

$$|a_1^*|_p \ge \ldots \ge |a_s^*|_p. \tag{1.9}$$

We consider first the situation in which p is an odd prime, and fix a primitive root g modulo p^2. Recall that g is then necessarily a primitive root for all powers of p. We write $\phi(\cdot)$ for Euler's totient, and put $\delta = (\phi(p^\gamma), k)$. Note that for each index i there exist integers m_i, u_i, v_i with

$$a_i^* \equiv p^{m_i} g^{u_i \delta + v_i} \pmod{p^\gamma},$$

and satisfying

$$0 \le m_i < k, \quad 0 \le v_i < \delta \quad \text{and} \quad 0 \le u_i < \phi(p^\gamma)/\delta.$$

But $(k/\delta, \phi(p^\gamma)/\delta) = 1$, and thus there exists an integer l with

$$l(k/\delta) \equiv 1 \pmod{\phi(p^\gamma)/\delta},$$

whence also $lk \equiv \delta \pmod{\phi(p^\gamma)}$. It follows that g^δ is congruent to a kth power modulo p^γ. In light of the latter observation, we put $b_i = p^{m_i} g^{v_i}$ $(1 \le i \le s)$, and write

$$\Upsilon(\mathbf{z}) = b_1 z_1^k + \ldots + b_s z_s^k.$$

Then we see that $\Phi^*(\mathbf{x})$ has a non-trivial p-adic zero if and only if the congruence $\Upsilon(\mathbf{z}) \equiv 0 \pmod{p^\gamma}$ possesses a solution $\mathbf{z}$ with $(b_i z_i, p) = 1$ for some i with $1 \le i \le s$. By multiplying all of the coefficients of $\Upsilon(\mathbf{z})$ through by $g^{\delta - v_i}$ for a suitable index i, and relabelling variables, if necessary, it is apparent that there is no loss of generality in supposing that

$$m_1 \le \ldots \le m_s \le \gamma \tag{1.10}$$

and

$$1 = (\delta+1)^{m_1}(v_1+1) \le (\delta+1)^{m_2}(v_2+1) \le \ldots \le (\delta+1)^{m_s}(v_s+1). \tag{1.11}$$

Let us write $\Upsilon^*(\mathbf{z})$ for the additive form thus obtained, which we now declare to be the *reduced form* corresponding to $\Phi(\mathbf{x})$ central to our discussion. It has the property that $\Phi(\mathbf{x})$ possesses a non-trivial p-adic zero if and only if

$\Upsilon^*(\mathbf{z})$ likewise possesses a non-trivial p-adic zero. When $p = 2$ and $\tau > 0$ we proceed similarly, though now making use of the fact that $(\mathbb{Z}/2^\gamma\mathbb{Z})^\times$ is generated by the residues $(-1)^l 5^m$.

We discuss singular primes only briefly, noting that in principle one may construct a list of all coefficients $(\alpha_1, \dots, \alpha_s)$ satisfying $\alpha_i = p^{m_i} g^{v_i}$, with $m_i \ge 0$, $0 \le v_i < \delta$, and subject to the conditions (1.10) and (1.11), and with at least $\lceil js/k \rceil$ of the α_i not divisible by p^j for $1 \le i \le s$ and $j \ge 1$, such that the congruence

$$\alpha_1 z_1^k + \dots + \alpha_s z_s^k \equiv 0 \ (\mathrm{mod}\, p^\gamma)$$

has no solution with $(\alpha_i z_i, p) = 1$ for any i with $1 \le i \le s$. In this way, one may make a list of the exceptional forms for which one is unable to find a non-trivial p-adic solution. We discuss singular primes no further, pausing only to note that this discussion demonstrates via a trivial estimate that there are at most $O((k \log k)^{k^2})$ exceptional forms for each singular prime p for the exponent k.

So far as regular primes p are concerned, the above procedure simplifies significantly on account of the fact that $\tau = 0$, and hence $\gamma = 1$. In this situation we find that $\Phi(\mathbf{x})$ possesses a non-trivial p-adic zero whenever the congruence $\Upsilon(\mathbf{z}) \equiv 0 \ (\mathrm{mod}\, p)$ possesses a solution with $(b_i z_i, p) = 1$ for some i with $1 \le i \le s$. Thus we are reduced to considering the solubility of congruences of the shape

$$b_1 x_1^k + \dots + b_t x_t^k \equiv 0 \ (\mathrm{mod}\, p), \tag{1.12}$$

where $b_i = g^{v_i}$ for some v_i with $0 \le v_i < \delta$ $(1 \le i \le t)$, where $t \ge \lceil s/k \rceil$, and in which

$$0 = v_1 \le \dots \le v_t. \tag{1.13}$$

A modicum of additional thought reveals that a list of insoluble congruences of the shape (1.12) enables us to completely determine the solubility of $\Phi(\mathbf{x})$ over $\mathbb{Q}_p$ (not merely providing a sufficient condition for solubility), by examining the sets of coefficients b_i in $\Upsilon(\mathbf{z})$ divisible by the same power of p. Moreover, the existence of an insoluble congruence of the shape (1.12) ensures the existence of an additive form

$$\Lambda(\mathbf{x}) = \sum_{i=1}^{k} p^{i-1}(b_1 x_{i1}^k + \dots + b_t x_{it}^k),$$

with only the trivial solution over $\mathbb{Q}_p$, whence $\Gamma^*(k; p) > tk$. In view of these observations, we henceforth restrict attention to the solubility of the congruence (1.12).

Consider then a regular prime p, and recall that $\delta = (p-1, k)$ and that g is a primitive root modulo p. For a natural number t, we consider the set $\mathcal{B}$ of t-tuples $(b_1, \dots, b_t)$ with $b_i = g^{v_i}$ for some v_i with $0 \le v_i < \delta$ $(1 \le i \le t)$,

and with $v_1 = 0$. We write $\mathfrak{G}_t(k;p)$ for the subset of the latter t-tuples for which the congruence

$$b_1x_1^k + \ldots + b_tx_t^k \equiv 0 \pmod{p}$$

has only the trivial solution $\mathbf{x} \equiv \mathbf{0} \pmod{p}$, and we write $G_t(k;p)$ for the cardinality of $\mathfrak{G}_t(k;p)$. Notice here that we do not order the v_i as in (1.13), this extra freedom making a comparison of our conclusions with the trivial estimate more transparent. However, by noting that whenever $(g^{v_1}, \ldots, g^{v_t}) \in \mathfrak{G}_t(k;p)$, then necessarily $(g^{w_1}, \ldots, g^{w_t}) \in \mathfrak{G}_t(k;p)$ for any t-tuple of integers $(w_1, \ldots, w_t)$, with $0 \le w_i < \delta$ and $w_i \equiv v_{\sigma i} - v_{\sigma 1} \pmod{\delta}$ $(1 \le i \le t)$ for some permutation σ of $\{1, \ldots, t\}$, we find that a more precise classification may be achieved in which the number of exceptional t-tuples is reduced by a factor roughly of $1/t!$. By exploiting our work from Sections 2 and 4, we establish the following bounds for $G_t(k;p)$ in Section 6.

THEOREM 1.6. *Let δ and k be natural numbers with $\delta \mid k$, and suppose that p is a prime number with $(p-1, k) = \delta$.*

(i) *One has*

$$G_t(k;p) \le \delta^{2t-2}p^{2-t}.$$

(ii) *Whenever $\delta(p^{1/(t-1)} - 1) > p - 1$, one has*

$$G_t(k;p) \ge \delta^{t-1}(p-1)^{-t}\prod_{r=1}^{t}(p - ((p-1)/\delta + 1)^{r-1}).$$

Observe that the total number of t-tuples $(b_1, \ldots, b_t)$, subject to $v_1 = 0$ and $b_i = g^{v_i}$ for some v_i with $0 \le v_i < \delta$ $(2 \le i \le t)$, is δ^{t-1}. Thus, given any positive number τ, and prime number p with $(p-1, k) = \delta$ and $p > \delta^{(t-1)/(t-2)+\tau}$, almost all "reduced congruences" in t variables are soluble non-trivially, since in such circumstances one has $G_t(k;p) \ll \delta^{t-1-\tau}$.

We next consider the size of the catalogue of all exceptional congruences in t variables for regular primes p, defining $G_t(k)$ by

$$G_t(k) = \sum_{p \nmid k} G_t(k;p).$$

As an immediate consequence of Theorem 1.6, we are able to show that this catalogue of exceptional congruences is not particularly long.

COROLLARY 1. *Suppose that k is a natural number with $k \ge 3$.*

(i) *For each exponent k one has*

$$G_3(k) \ll k^3(\log\log k)^2,$$

and for almost all exponents k, one has

$$G_3(k) \gg k^3/\log k.$$

(ii) *When $t \geq 4$, one has*

$$G_t(k) \ll k^t \log\log k.$$

We note that the conclusion of Theorem 1.6(ii), combined with standard conjectures concerning the distribution of prime numbers in arithmetic progressions, shows that when t is fixed and k is large, one has the conditional lower bound $G_t(k) \gg k^{t-1+1/(t-2)}/\log k$. Meanwhile, when $t \geq 3$ one finds that a trivial bound yields the estimate $G_t(k) \ll k^{t+2/(t-2)}$. When $t > k$, of course, the classical theory readily demonstrates that $G_t(k) = 0$. In such circumstances, the conclusions of Theorem 1.6 and its corollary become inconsequential (see also Dodson [15] for refinements of [11] relevant in this context). We should note also that when k is odd, the upper bound

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

due to Tietäväinen [24] (see Chowla and Shimura [6] for earlier work) demonstrates that $G_t(k) = 0$ for $t > (1 + o(1))(\log k)/(\log 2)$.

Throughout this paper, the letter k denotes an integer exceeding 2, and ε denotes a sufficiently small positive number. The implicit constants in Vinogradov's well-known notation, $\ll$ and $\gg$, will depend at most on quantities occurring as subscripts to the latter notation. When $\alpha \in \mathbb{R}$, we write $\lceil \alpha \rceil$ for the smallest integer greater than or equal to α, and $[\alpha]$ for the largest integer not exceeding α. On occasion we will abuse notation by speaking interchangeably of the finite field $\mathbb{F}_p$, and the set of congruence classes modulo p.

The author is grateful to Professor P. T. Bateman for supplying a copy of Norton's technical report (see [23]). I am also very grateful to Don Lewis for encouragement, support, advice, and much else besides. The subject area of this paper would, of course, be much the poorer were it not for his work and influence.

2. Criteria for the existence of exceptional primes. At the core of our proof of Theorem 1.1 is a simple counting argument. In order to describe the conclusion of this elementary argument, we introduce some notation. When p is a prime number and t and k are natural numbers, we define $E_t(k;p)$ to be the number of t-tuples $(g_1, \dots, g_t)$ with $g_i \in \mathbb{F}_p^\times$ $(1 \leq i \leq t)$, and satisfying the property that the equation

$$\sum_{i=1}^{t} g_i x_i^k = 0 \tag{2.1}$$

has only the trivial solution $\mathbf{x} = \mathbf{0}$ over $\mathbb{F}_p$.

LEMMA 2.1. *Let t, δ and k be natural numbers with $t \geq 2$ and $\delta \geq 2$, and let p be a prime number with $(p-1,k) = \delta$ and*

$$\delta(p^{1/(t-1)} - 1) > p - 1. \tag{2.2}$$

Then

$$E_t(k,p) \geq \prod_{r=1}^{t} (p - ((p-1)/\delta + 1)^{r-1}). \tag{2.3}$$

Proof. We establish the lemma by induction, noting that when $t = 1$ the lower bound (2.3) follows immediately from the observation that whenever $a \in \mathbb{F}_p^\times$, then the equation $ax^k = 0$ has only the trivial solution $x = 0$. Suppose then that $T \geq 2$, and that the conclusion of the lemma holds for $t < T$. We may suppose that p is a prime number with $(p-1,k) = \delta$, satisfying the condition (2.2) with $t = T$. For the sake of concision, we write $q = (p-1)/\delta + 1$. By the definition of $E_t(k;p)$, there exist $E_{T-1}(k;p)$ distinct $(T-1)$-tuples $(a_1, \dots, a_{T-1})$ with $a_i \in \mathbb{F}_p^\times$ $(1 \leq i \leq T-1)$, and such that the equation

$$a_1 x_1^k + \dots + a_{T-1} x_{T-1}^k = 0 \tag{2.4}$$

has only the trivial solution over $\mathbb{F}_p$. Fix any one such $(T-1)$-tuple $\mathbf{a}$. Then on noting that the monomial x^k takes precisely q distinct values as x varies over $\mathbb{F}_p$, it follows that as the x_i vary over $\mathbb{F}_p$ for $1 \leq i \leq T-1$, the polynomial $a_1 x_1^k + \dots + a_{T-1} x_{T-1}^k$ takes at most q^{T-1} distinct values, including 0. Consequently, there are at least $p - q^{T-1}$ distinct choices for a_T in $\mathbb{F}_p$ for which the equation

$$a_1 x_1^k + \dots + a_{T-1} x_{T-1}^k = -a_T$$

is insoluble, none of which is zero. But by hypothesis, the equation (2.4) has only the trivial solution over $\mathbb{F}_p$, so that by homogeneity, each of the associated equations

$$a_1 x_1^k + \dots + a_{T-1} x_{T-1}^k = -a_T x_T^k$$

has only the trivial solution. We thus conclude that

$$E_T(k;p) \geq (p - q^{T-1}) E_{T-1}(k;p),$$

and so the inductive hypothesis follows with T replaced by $T+1$. This completes the proof of the lemma.

When -1 is not a kth power residue modulo p, one may exploit additive equations with symmetric features in order to extract further dividends from the argument of the proof of Lemma 2.1. In principle, this argument applies also for equations in which more than two coefficients are equal, but in such circumstances the hypothesis that -1 be a kth power non-residue must be replaced by a more complicated condition with little utility.

LEMMA 2.2. *Let t, δ and k be natural numbers with $t \ge 2$ and $\delta \ge 2$, and let p be a prime number with $(p-1,k) = \delta$ and $p \equiv \delta + 1 \pmod{2\delta}$. Write $q = (p-1)/\delta + 1$. Then one has the following conclusions.*

(i) *Suppose that t is an even integer with $t \ge 2$. Then whenever*

$$p > q\big(\tfrac{1}{2}q(q+1)\big)^{(t-2)/2}, \tag{2.5}$$

one has

$$E_t(k;p) \ge (p-1)\prod_{r=1}^{(t-2)/2}\big(p - q\big(\tfrac{1}{2}q(q+1)\big)^r\big). \tag{2.6}$$

(ii) *Suppose that t is an odd integer with $t \ge 3$. Then whenever*

$$p > \big(\tfrac{1}{2}q(q+1)\big)^{(t-1)/2}, \tag{2.7}$$

one has

$$E_t(k;p) \ge (p-1)\big(p - \big(\tfrac{1}{2}q(q+1)\big)^{(t-1)/2}\big)\prod_{r=1}^{(t-3)/2}\big(p - q\big(\tfrac{1}{2}q(q+1)\big)^r\big). \tag{2.8}$$

Proof. We begin by observing that when $p \equiv \delta + 1 \pmod{2\delta}$, then

$$(-1)^{(p-1)/\delta} = -1,$$

so that -1 is a kth power non-residue modulo p. It therefore follows that whenever $a \in \mathbb{F}_p^\times$, then the equation $a(x^k + y^k) = 0$ has only the trivial solution $x = y = 0$. The lower bound (2.6) is therefore immediate when $t = 2$. When u is an integer with $u \ge 1$, write

$$F_u(k;p) = (p-1)\prod_{r=1}^{u-1}\big(p - q\big(\tfrac{1}{2}q(q+1)\big)^r\big).$$

We claim that when u is an integer with $u \ge 1$, and p is a prime number with $(p-1,k) = \delta$ and $p \equiv \delta + 1 \pmod{2\delta}$, satisfying the lower bound

$$p > q\big(\tfrac{1}{2}q(q+1)\big)^{u-1}, \tag{2.9}$$

then there exist at least $F_u(k;p)$ distinct u-tuples $(a_1, \dots, a_u)$, with $a_i \in \mathbb{F}_p^\times$ $(1 \le i \le u)$, such that the equation

$$\sum_{i=1}^{u} a_i(x_{2i-1}^k + x_{2i}^k) = 0$$

has only the trivial solution over $\mathbb{F}_p$.

Suppose that the above claim holds for $u < U$. We may suppose that p is a prime number with $(p-1,k) = \delta$ and $p \equiv \delta + 1 \pmod{2\delta}$, satisfying the

condition (2.9) with $u = U$. Then there exist $F_{U-1}(k;p)$ distinct $(U-1)$-tuples $(a_1, \dots, a_{U-1})$, with $a_i \in \mathbb{F}_p^\times$ $(1 \le i \le U-1)$, such that the equation

$$\sum_{i=1}^{U-1} a_i(x_{2i-1}^k + x_{2i}^k) = 0 \tag{2.10}$$

has only the trivial solution over $\mathbb{F}_p$. Fix any one such $(U-1)$-tuple $\mathbf{a}$. Note that by symmetry, the polynomial $x^k + y^k$ takes at most $\frac{1}{2}q(q-1)$ distinct values as x and y vary over $\mathbb{F}_p$ with $x^k \ne y^k$, and that this polynomial takes an additional q values with $x^k = y^k$. It follows that as the x_i vary over $\mathbb{F}_p$ for $1 \le i \le 2U-2$, the polynomial

$$\sum_{i=1}^{U-1} a_i(x_{2i-1}^k + x_{2i}^k)$$

takes at most $\big(\frac{1}{2}q(q+1)\big)^{U-1}$ distinct values, including zero. Moreover, the polynomial $1 + y^k$ takes precisely q distinct values as y varies over $\mathbb{F}_p$, and since -1 is a kth power non-residue modulo p, none of the aforementioned values are zero. We therefore conclude that the expression

$$(1+y^k)^{-1} \sum_{i=1}^{U-1} a_i(x_{2i-1}^k + x_{2i}^k)$$

takes at most $q\big(\frac{1}{2}q(q+1)\big)^{U-1}$ distinct values as y and x_i $(1 \le i \le 2U-2)$ vary over $\mathbb{F}_p$. Consequently, there are at least $p - q\big(\frac{1}{2}q(q+1)\big)^{U-1}$ distinct choices for a_U in $\mathbb{F}_p$ for which the equation

$$(1+y^k)^{-1} \sum_{i=1}^{U-1} a_i(x_{2i-1}^k + x_{2i}^k) = -a_U$$

is insoluble, none of which is zero. But by hypothesis, the equation (2.10) has only the trivial solution over $\mathbb{F}_p$, so that by homogeneity, each of the associated equations

$$\sum_{i=1}^{U-1} a_i(x_{2i-1}^k + x_{2i}^k) = -a_U(y_1^k + y_2^k)$$

has only the trivial solution with $\mathbf{x} = \mathbf{0}$ and $\mathbf{y} = \mathbf{0}$. We therefore conclude that our initial claim holds for each $U \ge 1$, and part (i) of the lemma is an immediate consequence of the trivial lower bound $E_t(k;p) \ge F_{t/2}(k;p)$, valid for every even integer t with $t \ge 2$.

Suppose next that U is an integer with $U \ge 2$, and consider a $(U-1)$-tuple of integers $(a_1, \dots, a_{U-1})$ satisfying the condition that the equation (2.10) has no non-trivial solution over $\mathbb{F}_p$. We may suppose on this occasion that p is a prime number with $(p-1,k) = \delta$ and $p \equiv \delta+1 \pmod{2\delta}$, satisfying

the condition $p > \big(\frac{1}{2}q(q+1)\big)^{U-1}$. In view of our earlier discussion, there are at least $F_{U-1}(k;p)$ such $(U-1)$-tuples. As we observed earlier, the polynomial

$$\sum_{i=1}^{U-1} a_i(x_{2i-1}^k + x_{2i}^k)$$

takes at most $\big(\frac{1}{2}q(q+1)\big)^{U-1}$ distinct values as the x_i vary over $\mathbb{F}_p$, including zero. Thus, there are at least $p - \big(\frac{1}{2}q(q+1)\big)^{U-1}$ distinct choices for a_U in $\mathbb{F}_p$ for which the equation

$$\sum_{i=1}^{U-1} a_i(x_{2i-1}^k + x_{2i}^k) = -a_U$$

is insoluble, none of which is zero. But the equation (2.10) has only the trivial solution over $\mathbb{F}_p$, so that by homogeneity, each of the associated equations

$$\sum_{i=1}^{U-1} a_i(x_{2i-1}^k + x_{2i}^k) = -a_U x_{2U-1}^k$$

has only the trivial solution $\mathbf{x} = \mathbf{0}$. We thus conclude that for each integer U with $U \ge 2$, one has

$$E_{2U-1}(k;p) \ge \big(p - \big(\tfrac{1}{2}q(q+1)\big)^{U-1}\big)E_{2U-2}(k;p),$$

and part (ii) of the lemma is now an immediate consequence of the conclusion of part (i).

We are now equipped to establish Theorem 1.1. Suppose that t, δ and k are natural numbers with $t \ge 3$ and $\delta \mid k$. Suppose also that p is a prime number with $(p-1,k) = \delta$.

In order to establish Theorem 1.1(i), we suppose that p and δ satisfy the inequalities $\delta > 4^{t-2}$ and

$$p \le \delta^{(t-1)/(t-2)} - 3\delta + 1.$$

Under the latter condition, one has

$$\begin{aligned} \frac{\delta(p^{1/(t-1)} - 1)}{p-1} &\ge \frac{\delta((\delta^{(t-1)/(t-2)}(1 - 3\delta^{-1/(t-2)}))^{1/(t-1)} - 1)}{\delta(\delta^{1/(t-2)} - 3)} \\ &= \frac{\delta^{1/(t-2)}(1 - 3\delta^{-1/(t-2)})^{1/(t-1)} - 1}{\delta^{1/(t-2)} - 3}. \end{aligned} \tag{2.11}$$

But whenever $t \ge 3$ and $0 < \xi < 3/4$, one has

$$(1-\xi)^{1/(t-1)} \ge (1-\xi)^{1/2} > 1 - 2\xi/3.$$

We therefore deduce that for $\delta > 4^{t-2}$, one has the lower bound

$$(1 - 3\delta^{-1/(t-2)})^{1/(t-1)} > 1 - 2\delta^{-1/(t-2)}.$$

On recalling (2.11), we conclude that

$$\frac{\delta(p^{1/(t-1)}-1)}{p-1} > \frac{\delta^{1/(t-2)}(1-2\delta^{-1/(t-2)})-1}{\delta^{1/(t-2)}-3} = 1.$$

Thus the condition (2.2) holds, and it follows from the lower bound (2.3) of Lemma 2.1 that $E_t(k;p) \geq 1$. The remarks in the introduction following (1.13) in this case provide the conclusion that $\Gamma^*(k;p) > tk$, and this completes the proof of Theorem 1.1(i).

We now turn our attention to the proof of Theorem 1.1(ii). Suppose first that t is an odd integer with $t \geq 3$. We may then suppose in addition that δ and p satisfy the inequalities

$$\delta \geq (\sqrt{2})^{-1}18^{t/2-1} \quad \text{and} \quad p \leq (\sqrt{2}\,\delta)^{(t-1)/(t-2)} - 4\delta + 1.$$

Write $u = (\sqrt{2})^{-1}(\sqrt{2}\,\delta)^{-1/(t-2)}$ and $q = (p-1)/\delta + 1$. Observe that when $\delta \geq (\sqrt{2})^{-1}18^{t/2-1}$, one has $0 < u \leq 1/6$. Then under the latter conditions, one finds that

$$\begin{aligned}\frac{\big(\frac{1}{2}q(q+1)\big)^{(t-1)/2}}{p-1} &\leq \frac{((\sqrt{2}\,\delta)^{2/(t-2)}(1-3u)(1-2u))^{(t-1)/2}}{(\sqrt{2}\,\delta)^{(t-1)/(t-2)}(1-4u)}\\ &\leq \frac{(1-3u)(1-2u)}{1-4u}.\end{aligned}$$

But when $0 < u \leq 1/6$, one has

$$\frac{(1-3u)(1-2u)}{1-4u} = 1 - \frac{u(1-6u)}{1-4u} \leq 1,$$

and thus we deduce that, under the aforementioned conditions, one has

$$\big(\tfrac{1}{2}q(q+1)\big)^{(t-1)/2} \leq p-1 < p.$$

The condition (2.7) therefore holds, and consequently it follows from the lower bound (2.8) of Lemma 2.2(ii) that $E_t(k;p) \geq 1$. The remarks in the introduction following (1.13) thus provide the desired conclusion $\Gamma^*(k;p) > tk$. This completes the proof of the first case of Theorem 1.1(ii), in which t is odd.

Finally, we suppose that t is an even integer with $t \geq 4$. We may now suppose that δ and p satisfy the conditions

$$\delta \geq 18^{t/2-1} \quad \text{and} \quad p \leq \sqrt{2}\,\delta^{(t-1)/(t-2)} - 4\delta + 1.$$

We now write $u = (\sqrt{2})^{-1}\delta^{-1/(t-2)}$ and $q = (p-1)/\delta + 1$. Observe that when $\delta \geq 18^{t/2-1}$, one has $0 < u \leq 1/6$. Then under these circumstances, one has

$$\frac{q\big(\frac{1}{2}q(q+1)\big)^{t/2-1}}{p-1}$$
$$\le \frac{(\sqrt{2}\,\delta^{1/(t-2)}-3)((\delta^{1/(t-2)}-3/\sqrt{2})(\delta^{1/(t-2)}-2/\sqrt{2}))^{t/2-1}}{\sqrt{2}\,\delta^{(t-1)/(t-2)}-4\delta}$$
$$\le \frac{(1-3u)((1-3u)(1-2u))^{t/2-1}}{1-4u} \le \frac{(1-3u)(1-2u)}{1-4u}.$$

But as in the case for odd values of t, the latter quantity is at most 1 for $0 < u \le 1/6$, and thus the aforementioned conditions ensure that

$$q\big(\tfrac{1}{2}q(q+1)\big)^{t/2-1} \le p-1 < p.$$

The condition (2.5) therefore holds, and hence it follows from the lower bound (2.6) of Lemma 2.2(i) that $E_t(k;p) \ge 1$. We consequently conclude as in previous cases that $\Gamma^*(k;p) > tk$. This completes the proof of the remaining case of Theorem 1.1(ii).

3. The abundance of exceptional primes. In order to exploit the criteria recorded in Theorem 1.1, we require useful information concerning the distribution of small primes in arithmetic progressions. This is a subject of considerable difficulty, and indeed the conclusions available for a single arithmetic progression fall far short of what would be required to obtain unconditional conclusions from Theorem 1.1. Thus the sharpest available estimates for Linnik's constant due to Heath-Brown [18] show only that whenever $(a,q) = 1$, then the least prime p with $p \equiv a \pmod q$ satisfies $p < cq^{11/2}$, in which c is an effectively computable constant. It is generally believed that there should be an abundance of primes p with $p \equiv 1 \pmod k$ and $p \ll k(\log k)^2$, for example, and such would provide positive conclusions stemming from Theorem 1.1 for every natural number t. However, so far as unconditional conclusions are concerned, we are forced to resort to average results available only relatively recently from work of Bombieri, Friedlander and Iwaniec [5]. As is usual, we define

$$\pi(x) = \sum_{\substack{p\le x\\ p\text{ prime}}} 1 \quad\text{and}\quad \pi(x;q,a) = \sum_{\substack{p\le x\\ p\equiv a\,(\mathrm{mod}\,q)\\ p\text{ prime}}} 1.$$

THEOREM 3.1. *Let a be a non-zero integer, and let A, x and Q be positive real numbers with $2 \le Q \le x^{3/4}$. Let Q' be a positive real number with $Q' < Q$, and denote by $\mathcal{Q}$ the set of integers q with $Q' < q \le Q$ satisfying $(a,q) = 1$. Write $\theta = \log Q/\log x$ and $L = \log x$. Then there is an absolute constant B such that*

$$\sum_{q\in\mathcal{Q}} |\pi(x;q,a)-\pi(x)/\phi(q)|$$
$$\le (B(\theta-1/2)^2xL^{-1}+O_A(xL^{-3}(\log L)^2))\sum_{q\in\mathcal{Q}}\phi(q)^{-1}+O_{a,A}(xL^{-A}).$$

Proof. This is the main theorem of Bombieri, Friedlander and Iwaniec [5].

Before tackling the proof of Corollary 1 to Theorem 1.1, we provide an elementary analytic estimate familiar to multiplicative number theorists.

LEMMA 3.2. *Let X and Y be real numbers with $X > Y > 2$. Then one has*

$$\sum_{Y<n\le X}\phi(n)^{-1}\ll \log(X/Y)+(\log X)/Y.$$

Proof. One has the relation

$$\phi(n)^{-1}=n^{-1}\prod_{p|n}(1-1/p)^{-1}\ll n^{-1}\prod_{p|n}(1+1/p)\le n^{-1}\sum_{d|n}d^{-1}.$$

Thus one obtains

$$\sum_{Y<n\le X}\phi(n)^{-1}\ll \sum_{Y<n\le X}n^{-1}\sum_{d|n}d^{-1}=\sum_{1\le d\le X}d^{-1}\sum_{Y/d<m\le X/d}(md)^{-1}$$
$$\ll \sum_{1\le d\le X}d^{-2}(\log(X/Y)+O(d/Y))$$

and the conclusion of the lemma follows immediately.

The proof of Corollary 1 to Theorem 1.1. Let K be a large real number, and recall that $E_2(K)$ denotes the number of exponents k, with $1\le k\le K$, for which $\Gamma^*(k)\le 3k$. By Theorem 1.1(i) with $t=3$, whenever $k\ge 5$ and there exists a prime p with $p\equiv 1 \pmod k$ and $p\le k^2-3k+1$, then one has $\Gamma^*(k)\ge 3k+1$. Thus we find that $E_2(K)-E_2(K/2)$ is bounded above by the number of exponents k, with $K/2<k\le K$, for which there exists no prime p with $p\equiv 1 \pmod k$ and $p\le K^2/5$. Writing $x=K^2/5$, we therefore deduce that

$$E_2(K)-E_2(K/2)\le \sum_{K/2<k\le K}\frac{\phi(k)}{\pi(x)}|\pi(x;k,1)-\pi(x)/\phi(k)|$$
$$\le \frac{K}{\pi(x)}\sum_{K/2<k\le K}|\pi(x;k,1)-\pi(x)/\phi(k)|.$$

But in view of Theorem 3.1, it follows that

$$E_2(K)-E_2(K/2)\ll \frac{K}{\pi(x)}\Big(\Xi\sum_{K/2<k\le K}\phi(k)^{-1}+x(\log x)^{-3}\Big), \tag{3.1}$$

where we write

$$\Xi = \left(\frac{\log K}{\log x} - \frac{1}{2}\right)^2 \frac{x}{\log x} + \frac{x(\log\log x)^2}{(\log x)^3}.$$

We next observe that Lemma 3.2 delivers the bound

$$\sum_{K/2<k\le K} \phi(k)^{-1} \ll \log 2 + (\log K)/K \ll 1,$$

and moreover that

$$\frac{\log K}{\log x} = \frac{\log K}{2\log K - \log 5} = \frac{1}{2} + O\bigg(\frac{1}{\log K}\bigg).$$

On substituting the latter estimates into (3.1), we therefore conclude from the prime number theorem that

$$\begin{aligned} E_2(K) - E_2(K/2) &\ll \frac{\log K}{K}\bigg(\frac{K^2}{(\log K)^3} + \frac{K^2(\log\log K)^2}{(\log K)^3}\bigg) \\ &\ll K\bigg(\frac{\log\log K}{\log K}\bigg)^2. \end{aligned}$$

The conclusion of Corollary 1 now follows on summing over dyadic intervals.

The proof of Corollary 2 to Theorem 1.1. Let ε be a fixed positive number. We wish to establish that for almost all exponents k, there exists a prime p with $p \equiv 1 \pmod k$ that satisfies the inequalities

$$k^2(1 - 1/(\log k)^{2-\varepsilon}) < p \le k^2 - 3k + 1. \tag{3.2}$$

Given such a prime p, it follows from Theorem 1.1(i) that one has $p \in \mathfrak{P}_2(k)$, whence

$$p_2^{\max}(k) > k^2(1 - 1/(\log k)^{2-\varepsilon}).$$

The conclusion of Corollary 2 follows immediately.

Let K be a large real number, and set

$$K_1 = K\big(1 - \tfrac{1}{4}(\log K)^{\varepsilon-2}\big), \quad \xi_1 = K_1^2 - 3K_1 + 1 \tag{3.3}$$

and

$$\xi_2 = \xi_1\big(1 - \tfrac{1}{4}(\log K)^{\varepsilon-2}\big).$$

Notice that whenever n lies between ξ_2 and ξ_1, and $K_1 < k \le K$, then necessarily one has

$$n > K^2(1 - (\log K)^{\varepsilon-2}) \ge k^2(1 - (\log k)^{\varepsilon-2}).$$

Thus, in order to establish this corollary, it suffices to show that for almost all exponents k with $K_1 < k \le K$, there is a prime number p with $p \equiv 1 \pmod k$ and $\xi_2 < p \le \xi_1$.

Next write

$$\varpi(k) = \pi(\xi_1; k, 1) - \pi(\xi_2; k, 1)$$

and

$$\nu(k) = (\pi(\xi_1) - \pi(\xi_2))/\phi(k).$$

By the prime number theorem with error term, one has

$$\nu(k) \gg \phi(k)^{-1} K^2 (\log K)^{\varepsilon-3}.$$

Let $\mathcal{E}(X)$ denote the number of exponents k, with $1 \le k \le X$, for which there exists no prime p with $p \equiv 1 \pmod k$ satisfying (3.2). Then one has

$$\begin{aligned}\mathcal{E}(K) - \mathcal{E}(K_1) &\le \sum_{K_1 < k \le K} \nu(k)^{-1} |\varpi(k) - \nu(k)| \\ &\ll K^{-1} (\log K)^{3-\varepsilon} \sum_{K_1 < k \le K} |\varpi(k) - \nu(k)|.\end{aligned}$$

But by the triangle inequality,

$$|\varpi(k) - \nu(k)| \le |\pi(\xi_1; k, 1) - \pi(\xi_1)/\phi(k)| + |\pi(\xi_2; k, 1) - \pi(\xi_2)/\phi(k)|.$$

Thus we deduce from Theorem 3.1 that

$$\mathcal{E}(K) - \mathcal{E}(K_1) \ll K^{-1} (\log K)^{3-\varepsilon} \Big(\Xi \sum_{K_1 < k \le K} \phi(k)^{-1} + K^2 (\log K)^{-9} \Big), \tag{3.4}$$

where we write

$$\Xi = \left(\frac{\log K}{\log \xi_2} - \frac{1}{2} \right)^2 \frac{K^2}{\log K} + \frac{K^2 (\log\log K)^2}{(\log K)^3}.$$

From Lemma 3.2, we have the upper bound

$$\sum_{K_1 < k \le K} \phi(k)^{-1} \ll -\log\big(1 - \tfrac{1}{4}(\log K)^{\varepsilon-2}\big) + (\log K)/K \ll (\log K)^{\varepsilon-2}.$$

Further, on recalling the definition of ξ_2, one finds that

$$\frac{\log K}{\log \xi_2} = \frac{\log K}{2 \log K + O(|\log(1 - (\log K)^{\varepsilon-2})|)} = \frac{1}{2} + O((\log K)^{\varepsilon-3}).$$

Consequently, on substituting these estimates into (3.4), we arrive at the estimate

$$\begin{aligned}\mathcal{E}(K) - \mathcal{E}(K_1) &\ll \frac{K}{(\log K)^{6-2\varepsilon}} + \frac{K (\log\log K)^2}{(\log K)^2} \\ &\ll K (\log K)^{\varepsilon/2-2}.\end{aligned} \tag{3.5}$$

Since the interval $(K_1, K]$ contains $\gg K(\log K)^{\varepsilon-2}$ integers, we may conclude that almost all integers k with $K_1 < k \le K$ satisfy the property that there exists a prime p, with $p \equiv 1 \pmod k$, and satisfying (3.2). The conclusion of the corollary follows.

The proof of Corollary 3 to Theorem 1.1. Let ε be a fixed positive number, let K be a large real number, and define K_1 and ξ_1 as in (3.3). We aim

to show that for all but $K/(\log K)^{\varepsilon/3}$ values of k with $K_1 < k \le K$, one has

$$|\pi(\xi_1;k,1) - \pi(\xi_1)/\phi(k)| < \tfrac{1}{4}\pi(\xi_1)\phi(k)^{-1}(\log K)^{\varepsilon-2}. \tag{3.6}$$

Whenever $K_1 < k \le K$ and p is a prime number with $p \equiv 1 \pmod k$ and $p \le \xi_1$, it is a consequence of Theorem 1.1(i) that $p \in \mathfrak{P}_2(k)$. Thus we find that $P_2(k) \ge \pi(\xi_1;k,1)$, and hence (3.6) will provide the desired lower bound on $P_2(k)$. Let $\mathfrak{E}(X)$ denote the number of exponents k, with $1 \le k \le X$, for which the inequality (3.6) fails. Then an application of the prime number theorem reveals that

$$\begin{aligned}\mathfrak{E}(K) - \mathfrak{E}(K_1) &\le \sum_{K_1<k\le K} \frac{4\phi(k)}{\pi(\xi_1)}(\log K)^{2-\varepsilon}|\pi(\xi_1;k,1) - \pi(\xi_1)/\phi(k)| \\ &\ll K^{-1}(\log K)^{3-\varepsilon} \sum_{K_1<k\le K} |\pi(\xi_1;k,1) - \pi(\xi_1)/\phi(k)|.\end{aligned}$$

But the argument leading to (3.4) and (3.5) above now leads to the conclusion that

$$\mathfrak{E}(K) - \mathfrak{E}(K_1) \ll K(\log K)^{\varepsilon/2-2}.$$

Furthermore, the interval $(K_1, K]$ contains $\gg K(\log K)^{\varepsilon-2}$ integers, and thus we deduce that (3.6) holds for almost all integers k with $K_1 < k \le K$. In particular, as a consequence of the prime number theorem, we deduce that for almost all integers k with $K_1 < k \le K$, one has

$$\begin{aligned}P_2(k) &\ge \pi(\xi_1;k,1) > \pi(\xi_1)\phi(k)^{-1}\big(1 - \tfrac{1}{4}(\log K)^{\varepsilon-2}\big) \\ &\ge \frac{K^2}{\log K}\,\phi(k)^{-1}\big(1 - \tfrac{5}{6}(\log K)^{\varepsilon-2}\big) \ge \frac{k}{\log k}\,(1 - (\log k)^{\varepsilon-2}).\end{aligned}$$

The conclusion of the corollary follows on covering the interval $(x/\log x, x]$ by a union of such intervals.

4. Upper bounds for exceptional primes. The ideas required in our proof of Theorem 1.2 lay the foundations also for our proof of Theorem 1.6. The definition of $\mathfrak{G}_t(k;p)$ restricts coefficients to the coset representatives of $\mathfrak{F}_k = \mathbb{F}_p^\times/(\mathbb{F}_p^\times)^k$, wherein we write $(\mathbb{F}_p^\times)^k$ for the set of kth powers of elements of $\mathbb{F}_p^\times$. In order to establish Theorem 1.6(i), we instead consider all possible choices for the coefficients. By averaging over the set of all such coefficients, one may estimate in mean square the discrepancy between the number of solutions of the corresponding equation over $\mathbb{F}_p$, and the expected number of solutions. From such an estimate one may infer an upper bound for the number of congruences that possess only the trivial solution. The ideas underlying this argument are motivated by the use of Bessel's inequality in the estimation of exceptional sets in Waring's problem.

Let k be a natural number with $k \geq 2$, and let p be a prime number. We denote by $N_s(\mathbf{h}; k, p)$ the number of solutions $(x_1, \dots, x_s)$ of the congruence

$$h_1 x_1^k + \dots + h_s x_s^k \equiv 0 \pmod{p}, \tag{4.1}$$

with $0 \leq x_i < p$ $(1 \leq i \leq s)$. We first record a lemma that estimates the difference, in mean square, between $N_s(\mathbf{h}; k, p)$ and its expected value p^{s-1}.

LEMMA 4.1. *Let δ and k be natural numbers with $\delta \mid k$, and suppose that p is a prime number with $(p-1, k) = \delta$. Suppose also that s is a natural number with $s \geq 3$. Then*

$$\sum_{h_1=1}^{p-1} \dots \sum_{h_s=1}^{p-1} |N_s(\mathbf{h}; k, p) - p^{s-1}|^2 \leq \frac{\delta - 1}{\delta}((\delta-1)^{s-1} - (-1)^{s-1})p^{s-2}(p-1)^{s+2}.$$

Proof. Define the exponential sum $f(u)$ by

$$f(u) = \sum_{x=1}^{p} e_p(ux^k),$$

where, as usual, we write $e_p(z)$ for $\exp(2\pi i z/p)$. Then it follows from orthogonality that

$$N_s(\mathbf{h}; k, p) = p^{-1} \sum_{u=1}^{p} f(h_1 u) \dots f(h_s u).$$

Thus we see that

$$\begin{aligned} \sum_{h_1=1}^{p-1} \dots \sum_{h_s=1}^{p-1} |N_s(\mathbf{h}; k, p) - p^{s-1}|^2 &= p^{-2} \sum_{h_1=1}^{p-1} \dots \sum_{h_s=1}^{p-1} \Big| \sum_{u=1}^{p-1} f(h_1 u) \dots f(h_s u) \Big|^2 \\ &= p^{-2} \sum_{u=1}^{p-1} \sum_{v=1}^{p-1} \Delta(u, v)^s, \end{aligned} \tag{4.2}$$

where

$$\Delta(u, v) = \sum_{h=1}^{p-1} f(hu) f(-hv).$$

But, again by orthogonality, one finds that $\Delta(u, v) = p\Theta(u, v) - p^2$, where $\Theta(u, v)$ denotes the number of solutions of the congruence $ux^k \equiv vy^k \pmod{p}$, with $0 \leq x, y < p$. When $u/v \in (\mathbb{F}_p^\times)^k$, it is apparent that $\Theta(u, v) = \delta(p-1) + 1$. On the other hand, when $1 \leq u, v < p$ and $u/v \notin (\mathbb{F}_p^\times)^k$, the

congruence $ux^k \equiv vy^k \pmod{p}$ has only the trivial solution $x = y = 0$, so that $\Theta(u,v) = 1$. We therefore conclude that for $1 \le u, v < p$, one has

$$\Delta(u,v) = \begin{cases} (\delta-1)p(p-1) & \text{when } u/v \in (\mathbb{F}_p^\times)^k, \\ -p(p-1) & \text{when } u/v \notin (\mathbb{F}_p^\times)^k, \end{cases}$$

whence by (4.2),

$$\begin{aligned} \sum_{h_1=1}^{p-1} \dots \sum_{h_s=1}^{p-1} |N_s(\mathbf{h};k,p) - p^{s-1}|^2 & \\ &= p^{-2}(\delta^{-1}(p-1)^2((\delta-1)p(p-1))^s \\ &\quad + ((p-1)^2 - \delta^{-1}(p-1)^2)(-p(p-1))^s) \\ &= p^{s-2}(p-1)^{s+2}(\delta^{-1}(\delta-1)^s + \delta^{-1}(\delta-1)(-1)^s). \end{aligned}$$

The conclusion of the lemma follows immediately.

It is convenient to establish Theorem 1.3 before turning our attention to the proof of Theorem 1.2. We first introduce some additional notation. We say that the s-tuples $(h_1, \dots, h_s)$ and $(h_1', \dots, h_s')$ of non-zero elements of $\mathbb{F}_p$ are equivalent, and we write $\mathbf{h} \sim \mathbf{h}'$, when there exists an element $\varrho \in \mathbb{F}_p^\times$ with the property that $\varrho h_i' h_i^{-1} \in (\mathbb{F}_p^\times)^k$ for $1 \le i \le s$. With a moment's thought, it is apparent that this notion of equivalence does indeed constitute an equivalence relation on the set of s-tuples with coordinates in $\mathbb{F}_p^\times$, and it is also evident that there are δ^{s-1} equivalence classes each containing the same number, $(p-1)^s/\delta^{s-1}$, of elements. Whenever $\mathbf{h} \sim \mathbf{h}'$, moreover, we see that $N_s(\mathbf{h};k,p) = N_s(\mathbf{h}';k,p)$.

Observe next that whenever σ is a permutation on the set $\{1, \dots, s\}$, then

$$N_s((h_1, \dots, h_s);k,p) = N_s((h_{\sigma 1}, \dots, h_{\sigma s});k,p).$$

Motivated by this observation, we define an equivalence relation $\approx_{\mathbf{h}}$ on the set Σ_s of permutations of $\{1, \dots, s\}$ by defining two permutations σ and τ to be equivalent, whereupon we write $\sigma \approx_{\mathbf{h}} \tau$, when

$$(h_{\sigma 1}, \dots, h_{\sigma s}) \sim (h_{\tau 1}, \dots, h_{\tau s}).$$

Once again it is apparent that equivalence classes contain the same number of elements. Write $\mathfrak{S}(\mathbf{h})$ for the number of distinct equivalence classes in Σ_s with respect to $\approx_{\mathbf{h}}$.

LEMMA 4.2. *Let δ and k be natural numbers with $\delta \mid k$, and suppose that p is a prime number with $(p-1,k) = \delta$. Suppose also that s is a natural number with $s \ge 3$. Then whenever $h_i \in \mathbb{F}_p^\times$ $(1 \le i \le s)$, one has*

$$|N_s(\mathbf{h};k,p) - p^{s-1}| \le \mathfrak{S}(\mathbf{h})^{-1/2}(\delta(\delta-1))^{(s-1)/2}p^{(s-2)/2}(p-1). \tag{4.3}$$

Proof. Let σ_i $(1 \le i \le \mathfrak{S}(\mathbf{h}))$ be a list of representatives of the equivalence classes of Σ_s with respect to $\approx_{\mathbf{h}}$. Then for $1 \le i \le \mathfrak{S}(\mathbf{h})$, one has

$$N_s((h_{\sigma_i 1}, \dots, h_{\sigma_i s}); k, p) = N_s((h_1, \dots, h_s); k, p).$$

But given any one permutation σ_i of this type, there are $(p-1)^s/\delta^{s-1}$ distinct s-tuples $\mathbf{h}'$ with $\mathbf{h}' \sim (h_{\sigma_i 1}, \dots, h_{\sigma_i s})$, and moreover, each such s-tuple $\mathbf{h}'$ is equivalent to $(h_{\sigma_j 1}, \dots, h_{\sigma_j s})$ for no j with $j \ne i$. Thus we see that

$$\sum_{g_1=1}^{p-1} \dots \sum_{g_s=1}^{p-1} |N_s(\mathbf{g}; k, p) - p^{s-1}|^2 \ge (\mathfrak{S}(\mathbf{h})(p-1)^s/\delta^{s-1})|N_s(\mathbf{h}; k, p) - p^{s-1}|^2,$$

whence, by Lemma 4.1,

$$\mathfrak{S}(\mathbf{h})(p-1)^s \delta^{1-s} |N_s(\mathbf{h}; k, p) - p^{s-1}|^2 \le (\delta - 1)^{s-1} p^{s-2} (p-1)^{s+2}.$$

The conclusion of the lemma follows immediately.

It is evident from Lemma 4.2 that whenever $\mathbf{h}$ is an s-tuple for which $\mathfrak{S}(\mathbf{h}) > 1$, then the estimate for $N_s(\mathbf{h}; k, p)$ provided by (4.3) will be superior to Weil's bound (at least, for large enough values of δ). Roughly speaking, one gains a factor of $\mathfrak{S}(\mathbf{h})^{-1/2}$, and this can be as large as $(s!)^{-1/2}$. By working more closely with Weil's argument, one may obtain slightly more precise conclusions that go beyond the latter factor (see long-forthcoming work of Granville and Wooley [17]). We are now equipped to establish Theorem 1.3.

The proof of Theorem 1.3. Recall the hypotheses of the statement of Theorem 1.3, write $s = r + 1$, and consider an s-tuple $(h_1, \dots, h_s)$ with $h_i \in \mathbb{F}_p^\times$ $(1 \le i \le s)$. We seek to establish that the congruence (4.1) possesses a solution $\mathbf{x} \not\equiv \mathbf{0} \pmod p$, whence we may infer from the argument leading to (1.13) above that $\Gamma^*(k; p) \le rk + 1$.

Suppose first that for some i and j with $1 \le i < j \le s$, one has $h_i h_j^{-1} \in (\mathbb{F}_p^\times)^k$. Then by making a change of variables, we see that there is no loss of generality in supposing that $h_1 = h_2 = 1$. But the hypothesis that $p \equiv 1 \pmod{2\delta}$ ensures that $(-1)^{(p-1)/\delta} = 1$, whence -1 is necessarily a kth power modulo p. Then the congruence $x_1^k + x_2^k \equiv 0 \pmod p$ has a non-trivial solution, and so (4.1) likewise possesses a non-trivial solution. We may suppose henceforth that $h_i h_j^{-1} \in (\mathbb{F}_p^\times)^k$ for no indices i and j with $1 \le i < j \le s$.

We consider next the situation in which $h_i h_j^{-1} \in (\mathbb{F}_p^\times)^k$ for no indices i and j with $1 \le i < j \le s$. Suppose first that $\mathfrak{S}(\mathbf{h}) < s!$. In such circumstances, there exists a permutation $\sigma \in \Sigma_s$, different from the identity permutation, with

$$(h_1, \dots, h_s) \sim (h_{\sigma 1}, \dots, h_{\sigma s}). \tag{4.4}$$

From the definition of equivalence, we see that there exist residues $\nu \in \mathbb{F}_p^\times \setminus (\mathbb{F}_p^\times)^k$ and $\xi_1, \dots, \xi_s \in \mathbb{F}_p^\times$ for which one has $h_{\sigma i} = \nu h_i \xi_i^k$ $(1 \le i \le s)$. Let g be a primitive root modulo p, and let $e(\sigma)$ denote the smallest positive integer e with the property that $g^e = \nu$ for some element ν associated with the permutation σ in the above manner. Let $\widehat{e}$ be the least value of $e(\sigma)$ as σ varies over all permutations σ, different from the identity permutation, for which (4.4) holds. Then in fact every permutation σ for which (4.4) holds is generated by any permutation τ for which $e(\tau) = \widehat{e}$. For suppose otherwise. Then there is a permutation σ for which $\widehat{e} \nmid e(\sigma)$. We put $f = (\widehat{e}, e(\sigma))$, and observe that $f = u\widehat{e} + ve(\sigma)$ for some integers u and v. Consequently, if we write μ for g^f, we find that there is a permutation $\omega = \tau^u \sigma^v$ for which there exist residues $\zeta_1, \dots, \zeta_s \in \mathbb{F}_p^\times$ with $h_{\omega i} = \mu h_i \zeta_i^k$ $(1 \le i \le s)$. Since $f < \widehat{e}$, we have contradicted the minimality of $\widehat{e}$. We are therefore forced to conclude that all permutations σ satisfying (4.4) are generated by τ. But the order of τ is at most s, and so there are at most s permutations within each equivalence class defined by $\approx_{\mathbf{h}}$. We thus deduce that $\mathfrak{S}(\mathbf{h}) \ge s!/s = r!$.

In view of the argument of the previous paragraph, we may suppose in what follows that $\mathfrak{S}(\mathbf{h}) \ge r!$. If one were to have no non-trivial solution of the congruence (4.1), then $N_s(\mathbf{h}; k, p)$ would count only the trivial solution $\mathbf{x} \equiv \mathbf{0} \pmod p$. Under such circumstances, we infer from Lemma 4.2 that

$$p^{s-1} - 1 \le \mathfrak{S}(\mathbf{h})^{-1/2} (\delta(\delta-1))^{(s-1)/2} p^{(s-2)/2} (p-1),$$

whence $p^{s-2} < \mathfrak{S}(\mathbf{h})^{-1} \delta^{2s-2}$. On recalling that in present circumstances, we may assume that $\mathfrak{S}(\mathbf{h}) \ge r!$, we conclude that necessarily

$$p < (r!)^{-1/(s-2)} \delta^{(2s-2)/(s-2)}.$$

It therefore follows that whenever $p \ge (r!)^{-1/(r-1)} \delta^{2r/(r-1)}$, then the congruence (4.1) has a non-trivial solution. The conclusion of the theorem now follows in this case, and the proof of Theorem 1.3 is complete.

Before embarking on the proof of Theorem 1.2, we recall the estimate for Gauss sums provided by Theorem 1 of Montgomery, Vaughan and Wooley [22].

LEMMA 4.3. *Let k be an even positive integer. Suppose that p is an odd prime number with $p \equiv k+1 \pmod{2k}$ and $p \nmid a$. Then one has*

$$\Big| \sum_{x=1}^{p} e_p(ax^k) \Big| \le 2^{-1/2} (k^2 - 2k + 2)^{1/2} p^{1/2}.$$

The proof of Theorem 1.2. Let k and r be natural numbers with $k \ge 3$ and $r \ge 2$, and write $p = p_r^{\max}(k)$. We may assume without loss of generality that $p > 0$. Put $\delta = (p-1, k)$, and observe that whenever $p \equiv 1 \pmod{2\delta}$,

it follows from Theorem 1.3 that

$$p < (r!)^{-1/(r-1)}\delta^{2r/(r-1)}.$$

But plainly $r! \geq 2^{r-1}$, and thus we obtain

$$p < \tfrac{1}{2}\delta^{2r/(r-1)} \leq \tfrac{1}{2}k^{2r/(r-1)}.$$

Suppose next that $p \not\equiv 1 \pmod{2\delta}$, whence $p \equiv \delta + 1 \pmod{2\delta}$ and δ is necessarily even. Write $s = r + 1$, consider integers $h_1, \dots, h_s$ with $1 \leq h_i \leq p - 1$ $(1 \leq i \leq s)$, and recall the definition of the exponential sum $f(u)$ from the proof of Lemma 4.1. We have

$$N_s(\mathbf{h}; k, p) - p^{s-1} = p^{-1} \sum_{u=1}^{p-1} f(h_1 u) \dots f(h_s u).$$

But whenever $(h_i u, p) = 1$, it follows from Lemma 4.3 that

$$|f(h_i u)| \leq 2^{-1/2}(\delta^2 - 2\delta + 2)^{1/2} p^{1/2}.$$

On considering the number of solutions of the underlying congruence $x_i^k \equiv y_i^k \pmod{p}$, moreover, it follows from orthogonality that, under the same conditions, one has

$$\sum_{u=1}^{p-1} |f(h_i u)|^2 = (\delta - 1)p(p - 1).$$

In this way, an application of Hölder's inequality yields the upper bound

$$\begin{aligned} |N_s(\mathbf{h}; k, p) - p^{s-1}| &\leq \prod_{i=1}^{s} \Big(p^{-1} \big(\max_{1 \leq a \leq p-1} |f(h_i a)| \big)^{s-2} \sum_{u=1}^{p-1} |f(h_i u)|^2 \Big)^{1/s} \\ &\leq 2^{1-s/2}(\delta^2 - 2\delta + 2)^{s/2-1}(\delta - 1)p^{s/2-1}(p - 1). \end{aligned}$$

Since $p = p_r^{\max}(k)$, the congruence (4.1) can have only the trivial solution $\mathbf{x} \equiv \mathbf{0} \pmod{p}$. Then one has $N_s(\mathbf{h}; k, p) = 1$, whence

$$p^{s-1} - 1 \leq 2^{1-s/2}(\delta^2 - 2\delta + 2)^{s/2-1}(\delta - 1)p^{s/2-1}(p - 1).$$

But this inequality implies that

$$p^{s/2-1} < 2^{1-s/2}\delta^{s-1},$$

and thus we conclude that

$$p_r^{\max}(k) < 2^{-1}\delta^{(2s-2)/(s-2)} \leq 2^{-1}k^{2r/(r-1)}.$$

This completes the proof of Theorem 1.2.

Our proof of Theorem 1.4 makes use of recent work of Heath-Brown and Konyagin [19] concerning upper bounds for the Gauss sum $f(u)$ defined in the course of the proof of Lemma 4.1.

LEMMA 4.4. *Suppose that* $(a,p)=1$. *Then one has*

$$\sum_{x=1}^{p} e_p(ax^k) \ll \begin{cases} k^{5/8}p^{5/8} & \text{when } p \ge k^2, \\ k^{3/8}p^{3/4} & \text{when } k^{3/2} \le p < k^2. \end{cases}$$

Proof. This is immediate from Theorem 1 of Heath-Brown and Konyagin [19].

The proof of Theorem 1.4. Suppose that k and r are natural numbers with $k \ge 3$ and $r \ge 2$, and write $p = p_r^{\max}(k)$. We may assume without loss of generality that $p > 0$. Put $\delta = (p-1,k)$, and recall the definition of the exponential sum $f(u)$ from the proof of Lemma 4.1. Also, write $s = r+1$, and consider integers $h_1, \dots, h_s$ with $1 \le h_i \le p-1$ $(1 \le i \le s)$. As in the proof of Theorem 1.2, we obtain the upper bound

$$\begin{aligned} |N_s(\mathbf{h};k,p) - p^{s-1}| &\le \prod_{i=1}^{s} \Big(p^{-1} \big(\max_{1 \le a \le p-1} |f(h_i a)| \big)^{s-2} \sum_{u=1}^{p-1} |f(h_i u)|^2 \Big)^{1/s} \\ &\le \big(\max_{1 \le a \le p-1} |f(a)| \big)^{s-2} (\delta - 1)(p-1). \end{aligned}$$

Since $p = p_r^{\max}(k)$, we find that $N_s(\mathbf{h};k,p) = 1$. When $p \ge k^2$, it therefore follows from Lemma 4.4 that

$$p^{s-1} - 1 \ll (k^{5/8}p^{5/8})^{s-2} k(p-1),$$

whence

$$p^{s-2} \ll k(kp)^{5(s-2)/8}.$$

We conclude in this case that

$$p \ll k^{(5s-2)/(3s-6)} = k^{(5r+3)/(3r-3)}. \tag{4.5}$$

When $k^{3/2} \le p < k^2$, meanwhile, one obtains in a similar fashion the relation

$$p^{s-1} - 1 \ll (k^{3/8}p^{3/4})^{s-2} k(p-1),$$

whence

$$p^{s-2} \ll k(kp^2)^{3(s-2)/8}.$$

Thus we conclude that in this second case, one has

$$p \ll k^{(3s+2)/(2s-4)} = k^{(3r+5)/(2r-2)}. \tag{4.6}$$

The conclusion of Theorem 1.4 is an immediate consequence of (4.5) and (4.6).

5. The number of exceptional primes. Upper bounds for $P_r(k)$ may be obtained directly from Theorems 1.2 and 1.4 via estimates of Brun–Titchmarsh type. Let A and α be real numbers with $A > 1$ and $\alpha > 3/2$.

Then Theorem 2 of Montgomery and Vaughan [21] establishes that

$$\sum_{\substack{p\equiv 1 \pmod{\delta}\\ \delta<p\le A\delta^{\alpha}}} 1 \le \frac{2A\delta^{\alpha}}{\phi(\delta)\log(A\delta^{\alpha-1})}, \tag{5.1}$$

uniformly in A, α and δ. When $\beta > 2$, it is a simple exercise in multiplicative number theory to establish that

$$\sum_{\delta\mid k} \delta^{\beta}\phi(\delta)^{-1} \ll k^{\beta}\phi(k)^{-1} \ll k^{\beta-1}\log\log k.$$

When $1 < \beta \le 2$, meanwhile, one obtains in similar fashion an estimate somewhat sharper than

$$\sum_{\delta\mid k} \delta^{\beta}\phi(\delta)^{-1} \ll k^{\beta-1}\exp(c\log k/\log\log k),$$

for a suitable positive number c. We therefore find from (5.1) that when $A > 1$ and $\alpha > 2$, one has

$$\begin{aligned}\sum_{\delta\mid k}\sum_{\substack{p\equiv 1 \pmod{\delta}\\ \delta<p\le A\delta^{\alpha}}} 1 &\ll k^{(\alpha-2)/2}(\log k)^{-1}\sum_{\delta\mid k}\delta^{(\alpha+2)/2}\phi(\delta)^{-1}\\ &\ll k^{\alpha-1}\,\frac{\log\log k}{\log k}.\end{aligned} \tag{5.2}$$

When $A > 1$ and $1 < \alpha \le 2$, meanwhile, it nonetheless follows that for each positive number ε, one has

$$\sum_{\delta\mid k}\sum_{\substack{p\equiv 1 \pmod{\delta}\\ \delta<p\le A\delta^{\alpha}}} 1 \ll \sum_{\delta\mid k}\delta^{\alpha}\phi(\delta)^{-1} \ll k^{\alpha-1+\varepsilon}. \tag{5.3}$$

Define the exponent α_r by

$$\alpha_r = \begin{cases} 2r/(r-1) & \text{when } 2 \le r \le 3,\\ (5r+3)/(3r-3) & \text{when } 4 \le r \le 8,\\ (3r+5)/(2r-2) & \text{when } r \ge 9.\end{cases}$$

Then the upper bounds presented in Theorem 1.5 follow by applying (5.2) and (5.3) with $\alpha = \alpha_r$, making use of the upper bounds for $p_r^{\max}(k)$ recorded in Theorem 1.2 for $2 \le r \le 3$, and those made available via Theorem 1.4 for $r > 3$.

6. Cataloguing exceptional equations. We now return to the topic of exceptional equations discussed, en passant, in Section 4. Let δ and k be natural numbers with $\delta \mid k$, and suppose that p is a prime number with $(p-1,k) = \delta$. Suppose also that s is a natural number with $s \ge 3$, and let

$(h_1, \dots, h_s)$ be an s-tuple of integers with

$$1 \le h_i \le p-1 \quad (1 \le i \le s). \tag{6.1}$$

If the congruence (4.1) has only the trivial solution $\mathbf{x} \equiv \mathbf{0} \pmod p$, then one has $N_s(\mathbf{h}; k, p) = 1$. Then on recalling the conclusion of Lemma 4.1, one finds that the number of such s-tuples $\mathbf{h}$ satisfying (6.1) is bounded above by

$$\sum_{h_1=1}^{p-1} \dots \sum_{h_s=1}^{p-1} \left| \frac{N_s(\mathbf{h}; k, p) - p^{s-1}}{p^{s-1} - 1} \right|^2 \le \frac{(\delta-1)^{s-1} p^{s-2} (p-1)^{s+2}}{(p^{s-1}-1)^2}. \tag{6.2}$$

By the definition of $\mathfrak{G}_s(k; p)$ and the discussion of Section 4, moreover, we see that each element of $\mathfrak{G}_s(k; p)$ generates $(p-1)^s/\delta^{s-1}$ s-tuples $\mathbf{h}$ satisfying (6.1) for which (4.1) has only the trivial solution $\mathbf{x} \equiv \mathbf{0} \pmod p$. We may therefore conclude from (6.2) that

$$G_s(k; p) \le \frac{\delta^{2s-2} p^{s-2} (p-1)^2}{(p^{s-1}-1)^2} < \delta^{2s-2} p^{2-s}.$$

This completes the proof of Theorem 1.6(i).

In order to establish Theorem 1.6(ii), we recall from Lemma 2.1 that the number of s-tuples $\mathbf{h}$ satisfying (6.1), for which the congruence (4.1) has only the solution $\mathbf{x} \equiv \mathbf{0} \pmod p$, is at least as large as

$$\prod_{r=1}^{s} (p - ((p-1)/\delta + 1)^{r-1}),$$

provided only that $\delta(p^{1/(s-1)} - 1) > p - 1$. But as in the previous paragraph, there are $(p-1)^s/\delta^{s-1}$ such s-tuples $\mathbf{h}$ corresponding to each element of $\mathfrak{G}_s(k; p)$. Thus we infer that

$$G_s(k; p) \ge \delta^{s-1} (p-1)^{-s} \prod_{r=1}^{s} (p - ((p-1)/\delta + 1)^{r-1}),$$

and the proof of Theorem 1.6 is now complete.

The upper bounds recorded in the corollary to Theorem 1.6 follow directly from Theorem 1.6(i) via the Brun–Titchmarsh theorem. Thus one obtains

$$G_t(k) \le \sum_{\delta \mid k} \sum_{\substack{\delta < p \le \delta^{(2t-2)/(t-2)} \\ p \equiv 1 \,(\mathrm{mod}\, \delta)}} \delta^{2t-2} p^{2-t},$$

and by dividing the summation over p into dyadic intervals, it follows from the Brun–Titchmarsh theorem that

$$G_t(k) \ll \sum_{\delta|k} \sum_{\substack{i=0 \\ 2^i\delta \le \delta^{(2t-2)/(t-2)}}}^{\infty} \delta^{2t-2}(2^i\delta)^{2-t} \sum_{\substack{2^i\delta < p \le 2^{i+1}\delta \\ p \equiv 1 \pmod{\delta}}} 1$$

$$\ll \sum_{\delta|k} \delta^t \sum_{\substack{i=0 \\ 2^i\delta \le \delta^{(2t-2)/(t-2)}}}^{\infty} (2^i)^{2-t} \frac{2^i\delta}{\phi(\delta)\log(2^{i+1})}.$$

But

$$\sum_{\substack{i=0 \\ 2^i\delta \le \delta^{(2t-2)/(t-2)}}}^{\infty} (i+1)^{-1}(2^i)^{3-t} \ll (\log\log(3\delta))^{\nu},$$

where ν is 1 or 0 according to whether $t = 3$ or $t > 3$. Thus, on recalling that $\phi(\delta) \gg \delta/\log\log\delta$, we find that

$$G_t(k) \ll \sum_{\delta|k} \delta^t(\log\log(3\delta))^{1+\nu} \ll k^t(\log\log k)^{1+\nu}.$$

This establishes the upper bounds of the corollary.

The lower bound for $G_3(k)$ recorded in part (i) of the corollary is an easy consequence of the argument yielding the proof of Corollary 3 to Theorem 1.1, combined with the conclusion of Theorem 1.6(ii). For the former shows that for almost all exponents k, there are $\gg k/\log k$ prime numbers p with $p \equiv 1 \pmod{k}$ and

$$\tfrac{1}{3}k^2 < p \le \tfrac{2}{3}k^2,$$

and the latter shows that for each such prime number p, one has $G_3(k;p) \gg k^2$. Thus we conclude that for almost all exponents k, one has

$$G_3(k) \gg k^2(k/\log k) = k^3/\log k.$$

This completes the proof of the corollary to Theorem 1.6.

References

[1] O. D. Atkinson, J. Brüdern and R. J. Cook, *Three additive cubic equations*, Acta Arith. 60 (1991), 29–83.

[2] —, —, —, *Simultaneous additive congruences to a large prime modulus*, Mathematika 39 (1992), 1–9.

[3] —, —, —, *Three additive congruences to a large prime modulus*, J. Austral. Math. Soc. Ser. A 55 (1993), 355–368.

[4] O. D. Atkinson and R. J. Cook, *Pairs of additive congruences to a large prime modulus*, ibid. 46 (1989), 438–455.

[5] E. Bombieri, J. B. Friedlander and H. Iwaniec, *Primes in arithmetic progressions to large moduli, III*, J. Amer. Math. Soc. 2 (1989), 215–224.

[6] S. Chowla and G. Shimura, *On the representation of zero by a linear combination of kth powers*, Norske Vid. Selsk. Forh. (Trondheim) 36 (1963), 169–176.

[7] R. J. Cook, *Pairs of additive congruences*: *cubic congruences*, Mathematika 32 (1985), 286–300.
[8] —, *Pairs of additive congruences*: *quintic congruences*, Indian J. Pure Appl. Math. 17 (1986), 786–799.
[9] —, *Computations for additive Diophantine equations*: *pairs of quintic congruences. II*, in: Computers in Mathematical Research, N. M. Stephens and M. P. Thorne (eds.), Clarendon Press, Oxford, 1988, 93–117.
[10] H. Davenport und H. Hasse, *Die Nullstellen der Kongruenzzetafunktionen in gewissen zyklischen Fällen*, J. Reine Angew. Math. 172 (1934), 151–182.
[11] H. Davenport and D. J. Lewis, *Homogeneous additive equations*, Proc. Roy. Soc. London Ser. A 274 (1963), 443–460.
[12] —, —, *Cubic equations of additive type*, Philos. Trans. Roy. Soc. London Ser. A 261 (1966), 97–136.
[13] —, —, *Two additive equations*, in: Proc. Sympos. Pure Math. 12, Amer. Math. Soc., Providence, 1967, 74–98.
[14] —, —, *Simultaneous equations of additive type*, Philos. Trans. Roy. Soc. London Ser. A 264 (1969), 557–595.
[15] M. M. Dodson, *Homogeneous additive congruences*, ibid. 261 (1966), 163–210.
[16] É. Fouvry, *Autour du théorème de Bombieri–Vinogradov*, Acta Math. 152 (1984), 219–244.
[17] A. Granville and T. D. Wooley, *Twist-induced cancellation in certain estimates of Weil*, in preparation.
[18] D. R. Heath-Brown, *Zero-free regions for Dirichlet L-functions, and the least prime in an arithmetic progression*, Proc. London Math. Soc. (3) 64 (1992), 265–338.
[19] D. R. Heath-Brown and S. Konyagin, *New bounds for Gauss sums derived from kth powers, and for Heilbronn's exponential sum*, Quart. J. Math. Oxford (2) 51 (2000), 221–235.
[20] I. D. Meir, *Simultaneous diagonal p-adic equations*, Mathematika 45 (1998), 337–349.
[21] H. L. Montgomery and R. C. Vaughan, *The large sieve*, ibid. 20 (1973), 119–134.
[22] H. L. Montgomery, R. C. Vaughan and T. D. Wooley, *Some remarks on Gauss sums associated with kth powers*, Math. Proc. Cambridge Philos. Soc. 118 (1995), 21–33.
[23] K. K. Norton, *On homogeneous diagonal congruences of odd degree*, Technical Report No. 16, Mathematics Branch, Office of Naval Research, August, 1966.
[24] A. Tietäväinen, *On a problem of Chowla and Shimura*, J. Number Theory 3 (1971), 247–252.
[25] A. Weil, *Numbers of solutions of equations in finite fields*, Bull. Amer. Math. Soc. 55 (1949), 497–508.

Department of Mathematics
University of Michigan
East Hall, 525 East University Ave.
Ann Arbor, MI 48109-1109, U.S.A.
E-mail: wooley@math.lsa.umich.edu

Current address:
Department of Mathematics
Harvard University
Science Center, One Oxford Street
Cambridge, MA 02138, U.S.A.
E-mail: wooley@math.harvard.edu

Received on 8.10.2001 (4123)