Waring’s Problem: A Survey

R. C. Vaughan and T. D. Wooley

1 The Classical Waring Problem

“Omnis integer numerus vel est cubus, vel e duobus, tribus, 4, 5, 6, 7, 8, vel novem cubis compositus, est etiam quadrato-quadratus vel e duobus, tribus, &c. usque ad novemdecim compositus, et sic deinceps.”

Waring [150, pp. 204-5].

“Every integer is a cube or the sum of two, three, ... nine cubes; every integer is also the square of a square, or the sum of up to nineteen such; and so forth.”

Waring [152, p. 336].

It is presumed that by this, in modern notation, Waring meant that for every $k \geq 3$ there are numbers $s$ such that every natural number is the sum of at most $s$ $k$-th powers of natural numbers and that the smallest such number $g(k)$ satisfies $g(3) = 9$, $g(4) = 19$.

By the end of the nineteenth century, the existence of $g(k)$ was known for only a finite number of values of $k$. There is an account of this work in Dickson [48], and as far as we have been able to ascertain, by 1909 its existence was known for $k = 3, 4, 5, 6, 7, 8, 9, 10$, but not for any larger $k$ (of course, with the natural extension of the definition of $g(k)$, Lagrange proved in 1770 that $g(2) = 4$). However, starting with Hilbert [69], who showed that $g(k)$ does indeed exist for every $k$, the twentieth century has seen an almost complete solution of this problem. Let $[x]$ denote the greatest integer not exceeding $x$ and write $\{x\}$ for $x - [x]$. As the result of the work of many mathematicians we now know that

$$g(k) = 2^k + \lfloor (3/2)^k \rfloor - 2,$$

provided that

$$2^k \{(3/2)^k\} + \lfloor (3/2)^k \rfloor \leq 2^k. \quad (1.1)$$

If this fails, then

$$g(k) = 2^k + \lfloor (3/2)^k \rfloor + \lfloor (4/3)^k \rfloor - \theta$$

---

1 Research supported by NSF grant DMS-9970632.
2 Packard Fellow, and supported in part by NSF grant DMS-9970440.
where $\theta$ is 2 or 3 according as
\[
[(4/3)^k][(3/2)^k] + [(4/3)^k] + [(3/2)^k]
\]
equals or exceeds $2^k$.

The condition (1.1) is known to hold (Kubina & Wunderlich [85]) whenever $k \leq 471, 600, 000$, and Mahler [91] has shown that there are at most a finite number of exceptions. To complete the proof for all $k$ it would suffice to know that $\{(3/2)^k\} \leq 1 - (3/4)^{k-1}$. Beukers [3] has shown that whenever $k > 5,000$ one has $\{(3/2)^k\} \leq 1 - a^k$, where $a = 2^{-0.9} = 0.5358...$, and this has been improved slightly by Dubitskas [49] to $a = 0.5769...$, so long as $k$ is sufficiently large (see also Bennett [1] for associated estimates). A problem related to the evaluation of $g(k)$ now has an almost definitive answer. Let $g_r(k)$ denote the smallest integer $s$ with the property that every natural number is the sum of at most $s$ elements from the set $\{1^k, r^k, (r+1)^k, \ldots \}$. Then Bennett [2] has shown that for $4 \leq r \leq (k + 1)^{1-1/k} - 1$, one has $g_r(k) = r^k + [(1 + 1/r)^k] - 2$.

By the way, before turning to the modern form of Waring’s problem, it is perhaps worth observing that in the 1782 edition of Meditationes Algebraicae, Waring makes an addition:

“confimilia etiam affirmari possunt (exceptis excipiendis) de eodem numero quantitatum earundem dimensionum.”

Waring [151, p. 349].

“similar laws may be affirmed (exceptis excipiendis) for the correspondingly defined numbers of quantities of any like degree.”

Waring [152, p. 336].

It would be interesting to know exactly what Waring had in mind. This, taken with some of the observations which immediately follow the remark, suggest that for more general polynomials than the $k$-th powers he was aware that some kind of local conditions can play a rôle in determining when representations occur.

2 The Modern Problem

The value of $g(k)$ is determined by the peculiar behaviour of the first three or four $k$-th powers. A much more challenging question is the value, for $k \geq 2$, of the function $G(k)$, the smallest number $t$ such that every sufficiently large number is the sum of at most $t$ $k$-th powers of positive integers. The function $G(k)$ has only been determined for two values of $k$, namely $G(2) = 4$, by Lagrange in 1770, and $G(4) = 16$, by Davenport [30]. The bulk of what is known about $G(k)$ has been obtained through the medium
of the Hardy-Littlewood method. This has its genesis in a celebrated paper of Hardy and Ramanujan [64] devoted to the partition function. In this paper (section 7.2) there is also a brief discussion about the representation of a natural number as the sum of a fixed number of squares of integers, and there seems little doubt that it is the methods described therein which inspired the later work of Hardy and Littlewood.

Our knowledge concerning the function \( G(k) \) currently leaves much to be desired. If, instead of insisting that all sufficiently large numbers be represented in a prescribed form, one rather asks that almost all numbers (in the sense of natural density) be thus represented, then the situation is somewhat improved. Let \( G_1(k) \) denote the smallest number \( u \) such that almost every number \( n \) is the sum of at most \( u \) \( k \)-th powers of positive integers. The function \( G_1(k) \) has been determined in five non-trivial instances as follows:

- Davenport [29], \( G_1(3) = 4 \),
- Hardy and Littlewood [62], \( G_1(4) = 15 \),
- Vaughan [121], \( G_1(8) = 32 \),
- Wooley [155], \( G_1(16) = 64 \),
- Wooley [155], \( G_1(32) = 128 \)

(of course, the conclusion \( G_1(2) = 4 \) is classical).

3 General Upper Bounds for \( G(k) \)

The first explicit general upper bound for \( G(k) \), namely

\[
G(k) \leq (k - 2)2^{k-1} + 5,
\]

was obtained by Hardy and Littlewood [61] (in [58] and [59], only the existence of \( G(k) \) is stated, although it is already clear that in principle their method gave an explicit upper bound). In Hardy and Littlewood [62] this is improved to

\[
G(k) \leq (k - 2)2^{k-2} + k + 5 + [\zeta_k],
\]

where

\[
\zeta_k = \frac{(k - 2)\log 2 - \log k + \log(k - 2)}{\log k - \log(k - 1)}.
\]

There has been considerable activity reducing this upper bound over the years, and Table 1 below presents upper bounds for \( G(k) \) that were probably the best that were known at the time they appeared, at least for
Vinogradov [136], \(32(k \log k)^2\),
Vinogradov [137][139], \(k^2 \log 4 + (2 - \log 16)k\) \((k \geq 3)\),
Vinogradov [135][138][140][143], \(6k \log k + 3k \log 6 + 4k\) \((k \geq 14)\),
Vinogradov [147], \(k(3 \log k + 11)\),
Tong [114], \(k(3 \log k + 9)\),
Jing-Run Chen [24], \(k(3 \log k + 5.2)\),
Vinogradov [148], \(2k(\log k + 2 \log \log k + O(\log \log \log k))\),
Vaughan [124], \(2k(\log k + \log \log k + O(1))\),
Wooley [155], \(k(\log k + \log \log k + O(1))\).

Table 1. General upper bounds for \(G(k)\)

\(k\) sufficiently large. This list is not exhaustive. In particular, there is a long sequence of papers by Vinogradov between 1934 and 1947, and for further details we refer the reader to the Royal Society obituary of I. M. Vinogradov (see Cassels and Vaughan [23]).

The last entry on this list has been refined further by Wooley [159], and this provides the estimate

\[ G(k) \leq k(\log k + \log \log k + 2 + O(\log \log k/ \log k)) \]

that remains the sharpest available for larger exponents \(k\).

4 Cubes

For small values of \(k\) there are many special variants of the Hardy-Littlewood method that have been developed. However, in the case of cubes, until recently the best upper bounds were obtained by rather different methods that related cubes to quadratic forms, especially sums of squares. Thus Landau [86] had shown that \(G(3) \leq 8\), and this bound was reduced by Linnik [87][88] to \(G(3) \leq 7\), with an alternative and simpler proof given by Watson [153]. Only with the advent of refinements to the circle method utilising efficient differencing did it become feasible (Vaughan [119]) to give a proof of the bound \(G(3) \leq 7\) via the Hardy-Littlewood method. Subsequent developments involving the use of smooth numbers (see Vaughan [124][125] and Wooley [158]) have provided a more powerful approach to this problem that, from a practical point of view, is more direct than earlier treatments. Complicated nonetheless, these latter proofs yield much more
information concerning Waring’s problem for cubes. We can illustrate the latter observation with two examples which, in the absence of foreseeable progress on the upper bound for $G(3)$, provide the problems central to current activity surrounding Waring’s problem for cubes.

When $X$ is a large real number, denote by $E(X)$ the number of positive integers not exceeding $X$ that cannot be written as a sum of four positive integral cubes. Then the conclusion $G_1(3) = 4$, attributed above to Davenport [29], is an immediate consequence of the estimate $E(X) \ll X^{29/30+\epsilon}$ established in the latter paper. Following subsequent work of Vaughan [119], Brüdern [10][12], and Wooley [158], the sharpest conclusion currently available (see Wooley [162]) shows that $E(X) \ll X^{1-\beta}$ for any positive number $\beta$ smaller than

$$(422 - 6\sqrt{2833})/861 = 0.119215\ldots$$

It is conjectured that $G(3) = 4$ (see §10 below), and this would imply that $E(X) \ll 1$.

Consider next the density of integers represented as a sum of three positive integral cubes. When $X$ is a large real number, let $N(X)$ denote the number of positive integers of the latter type not exceeding $X$. It is conjectured that $N(X) \gg X$, and following work of Davenport [29][33], Vaughan [118][119], Ringrose [100], Vaughan [124] and Wooley [158], the sharpest currently available conclusion due to Wooley [162] establishes that $N(X) \gg X^{1-\alpha}$ for any real number $\alpha$ exceeding

$$(\sqrt{2833} - 43)/123 = 0.083137\ldots$$

We remark that, subject to the truth of an unproved Riemann Hypothesis concerning certain Hasse-Weil $L$-functions, one has the conditional estimate $N(X) \gg X^{1-\epsilon}$ due to Hooley [73][74], and Heath-Brown [67]. Unfortunately, the underlying $L$-functions are not yet known even to have an analytic continuation inside the critical strip.

We finish our discussion of Waring’s problem for cubes by noting that, while Dickson [47] was able to show that 23 and 239 are the only positive integers not represented as the sum of eight cubes of natural numbers, no such conclusion is yet available for sums of seven or fewer cubes (but see McCurley [92] for sums of seven cubes, and Deshouillers, Hennecart and Landreau [44] for sums of four cubes).

5 Biquadrates

Davenport’s definitive statement that $G(4) = 16$ is not the end of the story for sums of fourth powers (otherwise known as biquadrates). Let
$G^\#(4)$ denote the least integer $s_0$ such that whenever $s \geq s_0$, and $n \equiv r \pmod{16}$ for some integer $r$ with $1 \leq r \leq s$, then $n$ is the sum of at most $s$ biquadrates. Then Davenport [30] showed that $G^\#(4) \leq 14$, and this has been successively reduced by Vaughan [121][124] to $G^\#(4) \leq 12$. In an ironic twist of fate, the polynomial identity

$$x^4 + y^4 + (x + y)^4 = 2(x^2 + xy + y^2)^2,$$

reminiscent of identities employed in the nineteenth century, has recently been utilised to make yet further progress. Thus, when $n \equiv r \pmod{16}$ for some integer $r$ with $1 \leq r \leq 10$, Kawada and Wooley [82] have shown that $n$ is the sum of at most 11 biquadrates. This and allied identities have also permitted the proof of an effective version of Davenport’s celebrated theorem. Thus, as a consequence of work of Deshouillers, Hennecart and Landreau [45] and Deshouillers, Kawada and Wooley [46], it is now known that all integers exceeding 13,792 may be written as the sum of at most sixteen biquadrates. A detailed history of Waring’s problem for biquadrates is provided in Deshouillers, Hennecart, Kawada, Landreau and Wooley [43].

6 Upper Bounds for $G(k)$ when $5 \leq k \leq 20$

Although we have insufficient space to permit a comprehensive account of the historical evolution of available upper bounds for $G(k)$ for smaller values of $k$, in Table 2 we have recorded many of the key developments, concentrating on the past twenty-five years. Each row in this table presents the best upper bound known for $G(k)$, for the indicated values of $k$, at the time of publication of the cited work. We note that the claimed bound $G(7) \leq 52$ of Sambasiva Rao [99] is based on an arithmetical error, and hence we have attributed the bound $G(7) \leq 53$ parenthetically to Davenport’s methods [28] [32]. Also, it is worth remarking that the work of Vaughan and Wooley [130] [131] and [132] appeared in print in an order reversed from its chronological development (indeed, this work was first announced in 1991). The bounds parenthetically attributed to Vaughan and Wooley [132] follow directly from the methods therein, and were announced on that occasion, though details (with additional refinements) appeared only in Vaughan and Wooley [134]. Meanwhile, the bounds recorded in Wooley [159] were an immediate consequence of the methods of Wooley [155] combined with the new estimates for smooth Weyl sums obtained in the former work (no attempt was made therein to exploit the methods of Vaughan and Wooley [132]).
Table 2. Upper bounds for $G(k)$ when $5 \leq k \leq 20$

<table>
<thead>
<tr>
<th>$k$</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$G(k)$</td>
<td>109</td>
<td>112</td>
<td>117</td>
<td>125</td>
<td>131</td>
<td>142</td>
<td>143</td>
<td>145</td>
<td>151</td>
<td>157</td>
<td>163</td>
<td>169</td>
<td>175</td>
<td>181</td>
<td>187</td>
<td>193</td>
</tr>
<tr>
<td>$G(k)$</td>
<td>109</td>
<td>112</td>
<td>117</td>
<td>125</td>
<td>131</td>
<td>142</td>
<td>143</td>
<td>145</td>
<td>151</td>
<td>157</td>
<td>163</td>
<td>169</td>
<td>175</td>
<td>181</td>
<td>187</td>
<td>193</td>
</tr>
<tr>
<td>$G(k)$</td>
<td>109</td>
<td>112</td>
<td>117</td>
<td>125</td>
<td>131</td>
<td>142</td>
<td>143</td>
<td>145</td>
<td>151</td>
<td>157</td>
<td>163</td>
<td>169</td>
<td>175</td>
<td>181</td>
<td>187</td>
<td>193</td>
</tr>
<tr>
<td>$G(k)$</td>
<td>109</td>
<td>112</td>
<td>117</td>
<td>125</td>
<td>131</td>
<td>142</td>
<td>143</td>
<td>145</td>
<td>151</td>
<td>157</td>
<td>163</td>
<td>169</td>
<td>175</td>
<td>181</td>
<td>187</td>
<td>193</td>
</tr>
<tr>
<td>$G(k)$</td>
<td>109</td>
<td>112</td>
<td>117</td>
<td>125</td>
<td>131</td>
<td>142</td>
<td>143</td>
<td>145</td>
<td>151</td>
<td>157</td>
<td>163</td>
<td>169</td>
<td>175</td>
<td>181</td>
<td>187</td>
<td>193</td>
</tr>
<tr>
<td>$G(k)$</td>
<td>109</td>
<td>112</td>
<td>117</td>
<td>125</td>
<td>131</td>
<td>142</td>
<td>143</td>
<td>145</td>
<td>151</td>
<td>157</td>
<td>163</td>
<td>169</td>
<td>175</td>
<td>181</td>
<td>187</td>
<td>193</td>
</tr>
<tr>
<td>$G(k)$</td>
<td>109</td>
<td>112</td>
<td>117</td>
<td>125</td>
<td>131</td>
<td>142</td>
<td>143</td>
<td>145</td>
<td>151</td>
<td>157</td>
<td>163</td>
<td>169</td>
<td>175</td>
<td>181</td>
<td>187</td>
<td>193</td>
</tr>
<tr>
<td>$G(k)$</td>
<td>109</td>
<td>112</td>
<td>117</td>
<td>125</td>
<td>131</td>
<td>142</td>
<td>143</td>
<td>145</td>
<td>151</td>
<td>157</td>
<td>163</td>
<td>169</td>
<td>175</td>
<td>181</td>
<td>187</td>
<td>193</td>
</tr>
<tr>
<td>$G(k)$</td>
<td>109</td>
<td>112</td>
<td>117</td>
<td>125</td>
<td>131</td>
<td>142</td>
<td>143</td>
<td>145</td>
<td>151</td>
<td>157</td>
<td>163</td>
<td>169</td>
<td>175</td>
<td>181</td>
<td>187</td>
<td>193</td>
</tr>
</tbody>
</table>
7 The Hardy-Littlewood Method

Practically all of the above conclusions have been obtained via the Hardy-
Littlewood method. Here is a quick introduction. Let $n$ be a large natural
number, and write $P = n^{1/k}$ and
\[ f(\alpha) = \sum_{x \leq P} e(\alpha x^k) \]
(here we follow the standard convention of writing $e(z)$ for $e^{2\pi iz}$). Then
on writing $R(n)$ for the number of representations of $n$ as the sum of $s$ $k$th
powers of natural numbers, it follows from orthogonality that
\[ R(n) = \int_0^1 f(\alpha)^s e(-\alpha n) d\alpha. \]

When $\alpha$ is “close” to a rational number $a/q$ with $(a, q) = 1$ and $q$ “small”,
we expect that
\[ f(\alpha) \sim q^{-1} S(q, a) v(\alpha - a/q), \]
where
\[ S(q, a) = \sum_{r=1}^q e(ar^k/q) \quad \text{and} \quad v(\beta) = \int_0^P e(\beta \gamma^k) d\gamma. \]

This relation is straightforward to establish in an interval about $a/q$, so
long as “close” and “small” are interpreted suitably. Now put
\[ R_A(n) = \int_A f(\alpha)^s e(-\alpha n) d\alpha. \]

For a suitable union $\mathcal{M}$ of such intervals centred on $a/q$ (the major arcs),
and for $s$ sufficiently large in terms of $k$, one can establish that as $n \to \infty$,
the asymptotic relation
\[ R_{\mathcal{M}}(n) \sim \frac{\Gamma(1 + 1/k)^s}{\Gamma(s/k)} n^{s/k - 1} \mathcal{S}(n) \]
holds, where $\mathcal{S}(n)$ is the singular series
\[ \mathcal{S}(n) = \sum_{q=1}^\infty T(q; n) \]
and
\[ T(q; n) = \sum_{\substack{a=1 \\ (a,q)=1}}^q q^{-s} S(q, a)^s e(-an/q). \]
8 The Necessary Congruence Condition

For each prime $p$, let

$$U(p; n) = \sum_{k=0}^{\infty} T(p^k; n).$$

The function $T(q; n)$ is multiplicative. Thus, when the singular series converges absolutely, one has

$$\mathcal{S}(n) = \prod_p U(p; n).$$

It is helpful to the success of the circle method that the singular series should satisfy $\mathcal{S}(n) \gg 1$. With this observation in mind, Hardy and Littlewood [61] defined $\Gamma(k)$ to be the least integer $s$ with the property that, for every prime number $p$, there is a positive number $C(p)$ such that $U(p; n) \geq C(p)$ uniformly in $n$. Subsequently, Hardy and Littlewood [62] showed that indeed $\mathcal{S}(n) \gg 1$ whenever $s \geq \max\{\Gamma(k), 4\}$. Next let $\Gamma_0(k)$ be the least number $s$ with the property that the equation

$$x_1^k + \cdots + x_s^k = n \quad (8.1)$$

has a non-singular solution in $\mathbb{Q}_p$ (or rather, that the corresponding congruence modulo $q$ always has a solution with $(x_1, q) = 1$). Hardy and Littlewood [63] were able to show that $\Gamma_0(k) = \Gamma(k)$ (see Theorem 1 of the aforementioned paper). Thus one sees that the singular series reflects the local properties of sums of $k$-th powers. In particular, the singular series is zero whenever the equation (8.1) fails to have a $p$-adic solution, for some prime $p$, and this reflects the trivial observation that the equation can be soluble over $\mathbb{Z}$ only if it is soluble everywhere locally.

Hardy and Littlewood [63] conjecture that $\Gamma(k) \to \infty$ as $k \to \infty$, but it is not even known whether or not one has

$$\liminf_{k \to \infty} \Gamma(k) \geq 4.$$

When $k > 2$, they showed that $\Gamma(k) = 4k$ when $k$ is a power of 2 and that $\Gamma(k) \leq 2k$ otherwise. They also computed $\Gamma(k)$ exactly when $3 \leq k \leq 36$, and established that $\Gamma(k) \geq 4$ when $3 \leq k \leq 3000$. Here they showed that equality occurs only when $k = 3, 7, 19$, and possibly (but improbably) when $k = 1163, 1637, 1861, 1997, 2053$. These values of $k$ can probably be settled by modern computing methods, and doubtless the calculations could be carried a good deal further. As far as we are aware, nothing has been done in this direction.
For a more detailed exposition of the Hardy-Littlewood method and the analysis of the major arcs and the singular series, see Vaughan [128] (especially Chapters 2 and 4).

9 The Minor Arcs

In order to establish an asymptotic formula for $R(n)$ it suffices to show that $R_m(n) = o(n^{s/k-1})$, where $m = [0, 1) \setminus \mathfrak{M}$ (the minor arcs). One needs to show that the minor arc contribution $R_m(n)$ is smaller by a factor $o(n^{-1})$, or equivalently $o(P^{-k})$, than the trivial estimate $P^s$. Routinely this is established via an inequality of the kind

$$\int_m |f(\alpha)|^s d\alpha \leq \left(\sup_m |f(\alpha)|\right)^{s-2t} \int_0^1 |f(\alpha)|^{2t} d\alpha.$$ 

The integral on the right hand side of this inequality may be interpreted as the number of solutions of an underlying diophantine equation, and it is from here that most of the savings usually come. On the other hand, non-trivial estimates for $|f(\alpha)|$, when $\alpha \in m$, may be obtained from estimates stemming from work of Weyl [154] and Vinogradov [144] (see Vaughan [128] for more modern estimates). When successful, this leads to the relation

$$R(n) \sim \frac{\Gamma(1 + 1/k)^s}{\Gamma(s/k)} n^{s/k-1} \tilde{\mathcal{S}}(n). \quad (9.1)$$

It is in finding ways of dealing with the minor arcs, or in modifying the method so as to make the minor arcs more amenable, that most of the research has concentrated in the eighty years that have elapsed since the pioneering investigations of Hardy and Littlewood.

10 The Asymptotic Formula

Define $\tilde{G}(k)$ to be the smallest natural number $s_0$ such that whenever $s \geq s_0$, the asymptotic relation (9.1) holds. Work of Hardy and Littlewood [62], described already in §3, established the general upper bound

$$\tilde{G}(k) \leq (k - 2)2^{k-1} + 5.$$ 

Progress on upper bounds for $\tilde{G}(k)$ has since been achieved on two fronts. In Table 3 we present upper bounds for $\tilde{G}(k)$ relevant for small values of $k$.

When $k$ is large, bounds stemming from Vinogradov’s mean value theorem provide dramatic improvements over the estimates recorded in Table 3.
Hua [75], $2^k + 1$,
Vaughan [119][122], $2^k$ ($k \geq 3$),
Heath-Brown [65][66], $7 \cdot 2^{k-3} + 1$ ($k \geq 6$),
Boklan [5], $7 \cdot 2^{k-3}$ ($k \geq 6$).

Table 3. Upper bounds for $\tilde{G}(k)$: smaller $k$

Vinogradov [142], $183k^9(\log k + 1)^2 + 1$,
Vinogradov [141], $91k^8(\log k + 1)^9 + 1$ ($k \geq 20$),
Vinogradov [147], $10k^2 \log k$,
Hua [77], $4k^2(\log k + \frac{1}{2} \log \log k + 8)$,
Wooley [156], $2k^2(\log k + O(\log \log k))$,
Ford [53], $k^2(\log k + \log \log k + O(1))$.

Table 4. Upper bounds for $\tilde{G}(k)$: larger $k$

In Table 4 we present upper bounds for $\tilde{G}(k)$ of use primarily when $k$ is large.

We note that the methods underlying the last two bounds can be adapted to give explicit bounds for $\tilde{G}(k)$ when $k$ is of moderate size. Thus the method of Ford yields a bound for $\tilde{G}(k)$ that is superior to the best recorded in Table 3 as soon as $k \geq 9$, and indeed unpublished work of Boklan and Wooley pushes this transition further to $k \geq 8$.

The bounds recorded in Tables 3 and 4 are likely to be a long way from the truth. One might expect that $G(k) = \max\{k+1, \Gamma_0(k)\}$, and, with an appropriate interpretation of the asymptotic formula when $\Theta(n) = 0$, that $\tilde{G}(k) = k + 1$.

One curiosity is that when $k = 3$ and $s = 7$, it can be shown that $R(n) \gg n^{4/3}$ (see Vaughan [125]), yet we are unable to show that $R(n) \ll n^{4/3}$. Indeed, it is currently the case that, quite generally, when $s$ lies in the range between the known upper bounds for $G(k)$ and $\tilde{G}(k)$, we can show that $R(n) \gg n^{s/k-1}$, but not $R(n) \ll n^{s/k-1}$.
11 Diminishing Ranges

In the Hardy-Littlewood method as outlined above the main problem is that of obtaining a suitable estimate for the mean value

$$\int_0^1 |f(\alpha)|^2 d\alpha,$$

that is, the number of integral solutions of the equation

$$x_1^k + \cdots + x_t^k = y_1^k + \cdots + y_t^k,$$

with $1 \leq x_j, y_j \leq P$ $(1 \leq j \leq t)$. Available estimates can be improved significantly if the variables are restricted, an idea already present in Hardy and Littlewood [62]. Define

$$P_1 = \frac{1}{2}P, \quad P_j = \frac{1}{2}P_{j-1}^{1-1/k} \quad (2 \leq j \leq t),$$

and consider the equation (11.1) subject to the constraints $P_j < x_j, y_j \leq 2P_j$ $(1 \leq j \leq t)$. Inspecting the expression $|x_j^k - y_j^k|$ successively for $j = 1, 2, \ldots$ when $x_j \neq y_j$, we find that only the diagonal solutions in which $x_j = y_j$ $(1 \leq j \leq t)$ occur. Thus we find that the number of solutions of this type is at most $O(P_1 \ldots P_t)$. This saves $P_1 \ldots P_t$ over the trivial bound, which is of order $(P_1 \ldots P_t)^2$. Now $P_1 \ldots P_t \approx P^\lambda$, where

$$\lambda = 1 + (1 - 1/k) + \cdots + (1 - 1/k)^{t-1} = k - k(1 - 1/k)^t.$$

Already when $t \sim Ck \log k$, for a suitable positive constant $C$, this exponent is close to $k$. Vinogradov and Davenport have exploited and developed this idea in a number of ways (see, for example, Davenport [27][34], and Vinogradov [147][149]; see also Davenport and Erdős [35]).

There is a “$p$-adic” analogue of this idea, first exploited by Davenport [31], in which one considers expressions of the kind

$$x_1^k + p_2^k(x_2^k + p_3^k(x_3^k + \cdots ))$$

on each side of the equation. Here the $p_i$ denote suitably chosen prime numbers. The analysis rests on congruences of the type $x_1 \equiv y_1 \pmod{p^k}$. When $p > P^{1/k}$ and $1 \leq x_1, y_1 \leq P$, this congruence implies that $x_1 = y_1$, and so on, just as in the diminishing ranges device. This idea has the merit of returning the various $k$-th powers $p_2^k \cdots p_t^k x_t^k$ to being in comparable size ranges. However in each of these methods the variables in (11.1) have varying natures and the homogeneity is essentially lost.
12 Smooth Numbers and Efficient Differences

In modern variants of the circle method as applied to Waring’s problem, starting with the work of Vaughan [124], homogeneity is restored by considering the number of solutions $S_s(P,R)$ of the equation (11.1) with $x_j, y_j \in A(P,R)$, where $A(P,R)$ denotes the set of $R$-smooth numbers up to $P$, namely

$$A(P,R) = \{n \in [1,P] \cap \mathbb{Z} : p|n \implies p \leq R\}.$$

In applications, one takes $R$ to be a suitably small, but positive power of $P$. The set $A(P,R)$ has the extremely convenient property that, given any positive number $M$ with $M \leq P$, and an element $x \in A(P,R)$ with $x > M$, there is always an integer $m$ with $m \in [M, MR]$ for which $m|x$. Moreover, this integer $m$ can be coaxed into playing the rôle of the prime $p$ in the $p$-adic argument mentioned above. Finally, and of great importance in what follows, the set $A(P,R)$ has positive density whenever $R$ is no smaller than a positive power of $P$.

The objective now is to find good exponents $\lambda_s$ with the property that whenever $\varepsilon > 0$, there exists a positive number $\eta_0 = \eta_0(s,k,\varepsilon)$ such that whenever $R = P^\eta$ with $0 < \eta \leq \eta_0$, one has

$$S_s(P,R) \ll P^{\lambda_s+\varepsilon}.$$

Such exponents are established via an iterative process in which a sequence of sets of exponents $\lambda_s = (\lambda_1^{(s)}, \lambda_2^{(s)}, \ldots)$ is constructed by finding an expression for each $\lambda_s^{(n+1)}$ in terms of the elements of $\lambda_s^{(n)}$. Boundedness is trivial, so there is always a convergent subsequence. In fact, our arguments produce monotonicity, and the convergence is fairly rapid. For a more detailed introduction and motivation for the underlying ideas in using smooth numbers in Waring’s problem, see the survey article Vaughan [127].

Beginning with the work of Wooley [155], a key element in the iterations is the repeated use of efficient differencing, and this procedure is fully exploited in subsequent work of Vaughan and Wooley [130] [131] [132] [134]. For each $s \in \mathbb{N}$, we take $\phi_i = \phi_{i,s}$ ($i = 1, \ldots, k$) to be real numbers with $0 \leq \phi_i \leq 1/k$. For $0 \leq j \leq k$, we then define

$$P_j = 2^j P, \quad M_j = P^{\phi_j}, \quad H_j = P_j M_j^{-k}, \quad Q_j = P_j (M_1 \ldots M_j)^{-1},$$

$$\tilde{H}_j = \prod_{i=1}^j H_i \quad \text{and} \quad \tilde{M}_j = \prod_{i=1}^j M_i R.$$

Define the modified forward difference operator, $\Delta^*_s$, recursively by taking

$$\Delta^*_s(f(x); h; m) = m^{-k}(f(x + hm^k) - f(x)),$$
and when \( j \geq 1 \), by inductively defining

\[
\Delta_{j+1}^*(f(x); h_1, \ldots, h_{j+1}; m_1, \ldots, m_{j+1}) = \Delta_1^*(f(x); h_1, \ldots, h_j; m_1, \ldots, m_j; h_{j+1}; m_{j+1}).
\]

For \( 0 \leq j \leq k \), let

\[
f(z) = (z - h_1 m_1^k - \cdots - h_j m_j^k)^k,
\]

and define the difference polynomial

\[
\Psi_j = \Psi_j(z; h_1, \ldots, h_j; m_1, \ldots, m_j)
\]

by taking

\[
\Psi_j = \Delta_j^*(f(z); 2 h_1, \ldots, 2 h_j; m_1, \ldots, m_j).
\]

Here we adopt the convention that \( \Psi_0(z; h; m) = z^k \). We write

\[
f_j(\alpha) = \sum_{x \in A(Q_j, R)} e(\alpha x^k),
\]

and

\[
F_j(\alpha) = \sum_{z, h, m} e(\alpha \Psi_j(z; h; m)),
\]

where the summation is over \( z, h, m \) with

\[
1 \leq z \leq P_j, \quad 1 \leq h_i \leq 2^{j-i} H_i,
\]

\[
M_i < m_i \leq M_i R, \quad m_i \in A(P, R),
\]

for \( 1 \leq i \leq j \). Finally, we define

\[
T(j, s) = \int_0^1 |F_j(\alpha)f_j(\alpha)^{2s}|d\alpha.
\]

Now, on considering the underlying diophantine equation, we have

\[
S_{s+1}(P, R) \leq \int_0^1 |F_0(\alpha)^2 f_0(\alpha)^{2s}|d\alpha.
\]

The starting point in the iterative process is to bound the latter expression in terms of \( S_s(Q_1, R) \) and \( T(1, s) \). This corresponds to taking the first difference in the classical Weyl differencing argument, and extracting the contribution arising from those terms with \( x_1 = y_1 \). Thus one obtains

\[
S_{s+1}(P, R) \ll P^s M_1^{2s-1} (P M_1 S_s(Q_1, R) + T(1, s)),
\]
and this inequality we write symbolically as

\[ F_0^2 f_0^{2s} \rightarrow F_1 f_1^{2s} \]

One way to proceed is by means of a repeated efficient differencing step. In principle this is based on the Cauchy-Schwarz inequality, applied in the form

\[
\int_0^1 |F_j f_j^{2s}| \, d\alpha \leq \left( \int_0^1 |f_j|^{2t} \, d\alpha \right)^{1/2} \left( \int_0^1 |F_j^2 f_j^{4s-2t}| \, d\alpha \right)^{1/2},
\]

where for the sake of concision we have written \( f_j \) for \( f_j(\alpha) \) and likewise \( F_j \) for \( F_j(\alpha) \). Thus, for \( j = 1, 2, \ldots \), the mean value \( T(j, s) \) can be related to \( S_t(Q_j, R) \) and \( T(j+1, 2s-t) \), where \( t < 2s \) is a parameter at our disposal, via inequalities of the shape

\[
T(j, s) \ll P^e (S_t(Q_j, R))^{1/2} (\tilde{H}_j \tilde{M}_j M_{j+1}^{4s-2t-1} \Xi_{j+1})^{1/2},
\]

where we write

\[
\Xi_{j+1} = P \tilde{H}_j \tilde{M}_{j+1} S_{2s-t}(Q_{j+1}, R) + T(j + 1, 2s - t).
\]

This is the \((j+1)\)-th step in the differencing process and can be portrayed by

\[
F_j f_j^{2s} \rightarrow F_{j+1} f_{j+1}^{4s-2t} \quad \downarrow \quad \Xi_{2t}
\]

There are more sophisticated variants of this procedure wherein it may be useful to restrict some of the variables to a range \( (\frac{1}{2} Q_j R^{-j}, Q_j] \), or to replace the set \( A(Q_j, R) \) by \( \mathbb{N} \cap [1, Q_j] \) (see §2 of Vaughan and Wooley [134] for details, and a more complete discussion).

Another option is to use Hölder’s inequality to bound \( T(j, s) \). Thus we obtain an inequality of the type

\[
T(j, s) = \int_0^1 |F_j(\alpha) f_j(\alpha)^{2s}| \, d\alpha \ll I_l^a I_{l+1}^b U_v^c U_w^d,
\]

where

\[
I_m = \int_0^1 |F_j(\alpha)|^{2m} \, d\alpha \quad (m = l, l + 1),
\]

\[
U_u = \int_0^1 |f_j(\alpha)|^{2u} \, d\alpha \quad (u = v, w),
\]
and $l, v, w, a, b, c, d$ are non-negative numbers satisfying the equations
\[ a + b + c + d = 1, \quad 2^la + 2^{l+1}b = 1, \quad vc + wd = s. \]
There is clearly great flexibility in the possible choices of the parameters here. We can summarise this process by
\[ F_j f_j^{2s} \implies (F_j^2)^a (F_j^{2^{l+1}})^b (f_j^{2v})^c (f_j^{2w})^d. \]
Yet another option is to apply the Hardy-Littlewood method to $T(j, s)$. In practice we expect that the minor arc contribution dominates, although this is not guaranteed. But if it does, then
\[ T(j, s) \ll \left( \sup_{\alpha \in \mathbb{R}} |F_j(\alpha)| \right) S_s(Q_j, R), \quad (12.1) \]
and this we abbreviate to
\[ F_j f_j^{2s} \implies (F_j)(f_j^{2s}). \]

By optimising choices for the parameters in order to obtain the sharpest estimates at each stage of the iteration process, one ultimately obtains relations describing $\lambda^{(n+1)}$ in terms of $\lambda^{(n)}$. The sharpest permissible exponents $\lambda$ attainable by these methods are in general not easy to describe, and require substantial computations to establish (see, for example, Vaughan and Wooley [134]). However, one can describe in general terms the salient features of the permissible exponents $\lambda_s$. When $s$ is rather small compared to $k$, it transpires that permissible exponents $\lambda_s = s + \delta_s$ can be derived with $\delta_s$ positive but small (see the next section for a consequence of this fact). Further, the simplest versions of the repeated efficient differencing method (see Wooley [155][157]) establish that the exponent $\lambda_s = 2s - k + ke^{-2s/k}$ is permissible for every natural number $s$. Roughly speaking, therefore, one may compare the respective strengths of the diminishing ranges argument, and the repeated efficient differencing method, by comparing how rapidly the respective functions $k(1 - 1/k)^s \sim ke^{-s/k}$ and $ke^{1-2s/k}$ tend to zero as $s$ increases.

The improvements in the most recent work (Vaughan and Wooley [134]) come about mostly through the following technical improvements:

- Better use of the Hardy-Littlewood method to estimate
\[ T(j, s) = \int_0^1 |F_j(\alpha) f_j(\alpha)^{2s}| \, d\alpha. \]

In particular, tighter control is exercised in mean on the behaviour of the exponential sum $F_j(\alpha)$ on the major arcs, and this permits the assumption of (12.1) for a larger range of parameters $\phi$ than previously available.
Waring’s Problem: A Survey

• Better estimates for

\[ I_m = \int_0^1 |F_j(\alpha)|^{2m} d\alpha, \]

established largely via estimates for the number of integral points on certain affine plane curves.

• Better estimates (see Wooley [159]) for

\[ \sup_{\alpha \in m} |f_0(\alpha)|. \]

Such estimates might be described as playing a significant rôle in the estimation of \( G(k) \) in the final stages of the analysis.

13 Breaking Classical Convexity

All of the methods thus far described depend, in a fundamental manner, on the natural interpretation of even moments of exponential sums in terms of the number of solutions of certain underlying diophantine equations. In §12, for example, one is limited to permissible exponents \( \lambda_s \) corresponding to integral values of \( s \), and in this setting the most effective method for bounding odd and fractional moments of smooth Weyl sums is to apply Hölder’s inequality to interpolate between even moments. With the natural extension of the notion of a permissible exponent \( \lambda_s \) from integral values of \( s \) to arbitrary positive numbers \( s \), the resulting exponents form the convex hull of the set of permissible exponents \( \{\lambda_s : s \in \mathbb{N}\} \). A perusal of §12 reveals that extra flexibility in choice of parameters, and therefore the potential for further improvements, will be achieved by the removal of this “classical convexity” barrier, and such has recently become available.

In Wooley [158], a method is established which, loosely speaking, enables one to replace the inequality

\[ S_{s+1}(P, R) \ll P^{s} M_1^{2s-1}(P M_1 S_s(Q_1, R) + T(1, s)) \]

that occurred in §12 with \( s \) restricted to be a natural number, by the new inequality

\[ S_{s+t}(P, R) \ll P^{s} M_1^{2s-t}(P^t M_1^t S_s(Q_1, R) + T^*), \]

where

\[ T^* = \int_0^1 |F(\alpha)^t f_1(\alpha)^{2s}| d\alpha \]
and
\[ F(\alpha) = \sum_{u \in A(M_1, R, R)} \sum_{\substack{z_1, z_2 \in A(P, R) \atop z_1 \equiv z_2 \pmod{u^k} \atop z_1 \neq z_2}} e(\alpha u^{-k}(z_1^k - z_2^k)). \]

In this latter estimate, the parameter \( s \) is no longer restricted to be integral, and the parameter \( t \) may be chosen freely with \( 0 < t \leq 1 \). Moreover, the mean value \( T^* \) is very much reminiscent of \( T(1, s) \), with \( F(\alpha) \) substituted for \( F_1(\alpha) \) and exhibiting similar properties. Thus, in addition to removing the integrality constraint on \( s \), one may also iterate with a fractional number \( 2t \) of variables.

As might be expected, the additional flexibility gained in this way leads to improved permissible exponents \( \lambda_s \) even for integral \( s \), since our methods are so highly iterative. The overall improvements are usually quite small and are largest for smaller values of \( k \). Such progress has not yet delivered sharper bounds for \( G(k) \), but this work provides the sharpest results available concerning sums of cubes (see \( \S \), and also Br"udern and Wooley [18]), and has also permitted new conclusions to be derived in certain problems involving sums of mixed powers (see Br"udern and Wooley [16] [19] [17], and also \( \S \) and below). Also, this “breaking convexity” device provides the best available lower bounds for the number \( N_{k,s}(X) \) of natural numbers not exceeding \( X \) that are the sum of \( s \) \( k \)th powers of positive integers, at least when \( s \) is small compared to \( k \). Thus, when \( 2 < s \leq 2e^{-1/\sqrt{k}} \), one has
\[ N_{k,s}(X) \gg X^{s/k - e^{-\gamma_s k}}, \]
where \( \gamma_s = 16/(es)^2 \) (see Theorems 1.3 and 1.4 of Wooley [158], and the associated discussion). For comparison, one conjecturally has the lower bound \( \hat{N}_{k,s}(X) \gg X^{s/k} \) whenever \( s < k \).

Much remains to be investigated for fractional moments, in part owing to the substantial increase in complexity of the underlying computations (see Wooley [161] for more on this). However, such developments presently appear unlikely to have a large impact on the central problem of bounding \( G(k) \) in the classical version of Waring’s problem.

14 Variants of Waring’s Problem: Primes

Much work has been devoted to various generalisations of the classical version of Waring’s problem, and it seems appropriate to discuss some of the more mainstream variants.

We begin with the Waring-Goldbach problem, in which one seeks to represent integers as sums of \( k \)th powers of prime numbers. In order to
describe the associated local conditions, suppose that \( k \) is a natural number and \( p \) is a prime number. We denote by \( \theta = \theta(k, p) \) the integer with \( p^\theta \mid k \) and \( p^{\theta+1} \notmid k \), and then define \( \gamma = \gamma(k, p) \) by

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

Finally, we put \( K(k) = \prod_{(p-1)\mid k} p^\gamma \).

Denote by \( H(k) \) the least integer \( s \) such that every sufficiently large positive integer congruent to \( s \) modulo \( K(k) \) may be written as a sum of \( s \) \( k \)th powers of prime numbers. Note that when \( (p-1)\mid k \), one has \( p^\theta (p-1)\mid k \), whence \( a^k \equiv 1 \pmod{p^\gamma} \) whenever \( (p, a) = 1 \). This explains our seemingly awkward definition of \( H(k) \), since whenever \( n \) is the sum of \( s \) \( k \)th powers of primes exceeding \( k + 1 \), then necessarily \( n \equiv s \pmod{K(k)} \). Naturally, further congruence conditions may arise from primes \( p \) with \( (p-1) \notmid k \).

Following the pioneering investigations of Vinogradov [145][146] (see also Vinogradov [147]), Hua comprehensively investigated additive problems involving prime numbers in his influential book (see Hua [79], but also Hua [78]). Thus, it is known that for every natural number \( k \) one has

\[
H(k) \leq 2^k + 1,
\]

and, when \( k \) is large, that

\[
H(k) \leq 4k (\log k + \frac{1}{2} \log \log k + O(1)).
\]

In the conventional plan of attack on the Waring-Goldbach problem, one applies the Hardy-Littlewood method in a manner similar to that outlined above, but in interpreting the number of solutions of an analogue of the equation (11.1) over prime numbers, one obtains an upper bound by discarding the primality condition. With sufficiently many variables employed to save a factor of \( n \) via such an approach, one additional variable suffices to save the extra power of \( \log n \) required by primality considerations. Although this strategy evidently prohibits the use of smooth numbers, the diminishing ranges technology perfected by Davenport, and refined by Vaughan [121] and Thanigasalam [109]–[113] plays a prominent role in establishing the best available upper bounds for \( H(k) \) when \( k \) is small. We should also mention that recent progress depends on good estimates of Weyl-type for exponential sums over primes, and allied sums, available from the use of Vaughan’s identity (see Vaughan [117]), combined
with the linear sieve equipped with a switching principle (see Kawada and Wooley [83]). Thus, for \(4 \leq k \leq 10\), the best known upper bounds for \(H(k)\) are as follows:

\[
\begin{align*}
\text{Kawada and Wooley [83],} & \quad H(4) \leq 14, \ H(5) \leq 21, \\
\text{Thanigasalam [111],} & \quad H(6) \leq 33, \ H(7) \leq 47, \ H(8) \leq 63, \\
& \quad H(9) \leq 83, \ H(10) \leq 107.
\end{align*}
\]

Despite much effort on the Waring-Goldbach problem for exponents 1, 2 and 3, further progress remains elusive. Improvements are feasible, however, if one is prepared to accept almost-primes in place of prime numbers (see, in particular, Chen [25], Brüdern [13][14], and Brüdern and Fouvry [15]). Difficulties related to those associated with the Waring-Goldbach problem are encountered when other sequences are substituted for prime numbers. For Waring’s problem with smooth variables, see Harcos [57] and Brüdern and Wooley [18]. Also, see Nechaev [97] for work on Waring’s problem with polynomial summands (Wooley [160] and Ford [55] have restricted improvements employing smooth numbers).

15 Variants of Waring’s Problem: Sums of Mixed Powers

Suppose that \(k_1, k_2, \ldots, k_s\) are natural numbers with \(2 \leq k_1 \leq k_2 \leq \cdots \leq k_t\). Then an optimistic counting argument suggests that whenever the equation

\[x_1^{k_1} + x_2^{k_2} + \cdots + x_s^{k_s} = n\]  \hspace{1cm} (15.1)

has \(p\)-adic solutions for each prime \(p\), and

\[k_1^{-1} + k_2^{-1} + \cdots + k_s^{-1} > 1,\]  \hspace{1cm} (15.2)

then \(n\) should be represented as the sum of mixed powers of positive integers (15.1) whenever it is sufficiently large in terms of \(k\). When \(s = 3\) such an assertion may fail in certain circumstances (see Jagy and Kaplansky [80], or Exercise 5 of Chapter 8 of Vaughan [128]), but a heuristic application of the Hardy-Littlewood method suggests, at least, that the condition (15.2) should ensure that almost all integers in the expected congruence classes are thus represented. Moreover, subject instead to the condition

\[k_1^{-1} + k_2^{-1} + \cdots + k_s^{-1} > 2,\]  \hspace{1cm} (15.3)

a formal application of the circle method suggests that all integers in the expected congruence classes should be represented in the form (15.1). Meanwhile, a simple counting argument shows that in circumstances in which
the condition (15.2) does not hold, then arbitrarily large integers are not represented in the desired form.

The investigation of such analogues of Waring’s problem for mixed powers has, since the early days of the Hardy-Littlewood method, stimulated progress in technology of use even in the classical version of Waring’s problem. Additive problems in which the summands are restricted to be squares, cubes or biquadrates are perhaps of greater interest than those with higher powers, and here the current situation is remarkably satisfactory. We summarise below the current state of knowledge in the simpler problems of this nature. In Tables 5 and 6 we list constellations of powers whose sum represents, respectively, almost all, and all, integers subject to the expected congruence conditions. The tables are arranged, roughly speaking, starting with predominantly smaller exponents, and ending with predominantly larger exponents, and therefore not in chronological order of the results.

We have been unable to trace the origin in the literature of the conclusion on a square, two biquadrates and a \( k \)th power, but refer the reader to Exercise 6 of §2.8 of Vaughan [128] for related ideas (see also Roth

\begin{itemize}
\item Davenport & Heilbronn [36], two squares, one \( k \)th power,
\item Davenport & Heilbronn [37], one square, two cubes,
\item Roth [101], one square, one cube, one biquadrate,
\item Vaughan [116], one square, one cube, one fifth power,
\item Folklore (?) one square, two biquadrates,
\item Hooley [70], one square, one cube, one sixth power,
\item Davenport [29], one cube, four biquadrates,
\item Brüdern [9][8], three cubes, one biquadrate,
\item Brüdern [8], Lu [89], three cubes, one fifth power,
\item Brüdern & Wooley [19], three cubes, one sixth power
\item Kawada & Wooley [82], one cube, four biquadrates,
\item Vaughan [124], six biquadrates,
\item Kawada & Wooley [82], five biquadrates, one \( k \)th power (\( k \) odd).
\end{itemize}

\textbf{Table 5.} Representation of almost all integers
Gauss [56], three squares,
Hooley [71], two squares, three cubes,
Hooley [70], two squares, assorted powers,
Vaughan [123], one square, five cubes,
Brüdern & Wooley [16], one square, four cubes, one biquadrate,
Ford [54], one square, one cube, one biquadrate,...,
one fifteenth power,
Linnik [87][88], seven cubes,
Brüdern [9], six cubes, two biquadrates,
Brüdern [9], five cubes, three biquadrates,
Kawada & Wooley [82], three cubes, six biquadrates,
Brüdern & Wooley [17], two cubes, seven biquadrates,
Kawada & Wooley [82], one cube, nine biquadrates
Vaughan [124], twelve biquadrates,
Kawada & Wooley [82], ten biquadrates, one $k$th power
($k$ odd).

Table 6. Representation of all integers

[101]). We remark that all ternary problems of interest have been solved, since for non-trivial triples $(k_1, k_2, k_3)$ not accounted for in Table 5, one has $k_1^{-1} + k_2^{-1} + k_3^{-1} \leq 1$. Also, energetic readers may be interested in tackling a problem which presently defies resolution only by the narrowest of margins, namely the problem of showing that almost all integers are represented as the sum of two cubes and two biquadrates of positive integers.

Note that although the three square theorem is commonly ascribed to Legendre, his “proof” depended on an unsubstantiated assumption only later established by Dirichlet, and the first complete proof is due to Gauss. We finish by noting that in problems involving sums of two squares, methods more effective than the circle method can be brought into play (see especially Hooley [70][71] and Brüdern [7]).

16 Variants of Waring’s Problem: Beyond $\mathbb{Z}$

Given the considerable energy expended on the investigation of Waring’s problem over the rational integers, it seems natural to extend this work
to algebraic number fields. Here one encounters the immediate difficulty of deciding what precisely Waring’s problem should mean in this broader context. It is possible, for example, that an algebraic integer in a number field $K$ may not be a sum of any finite number of $k$th powers of algebraic integers of that field (consider, say, the parity of the imaginary part of $a^2$ when $a \in \mathbb{Z}[i]$). With this in mind, when $K$ is a number field and $\mathcal{O}_K$ is its ring of integers, we define $J_k$ to be the subring of $\mathcal{O}_K$ generated by the $k$th powers of integers of $K$. We must also provide an analogue of the positivity of the $k$th powers inherent in the classical version of Waring’s problem. Thus we define $G_K(k)$ to be the smallest positive integer $s$ with the property that, for some positive number $c = c(k, K)$, and for all totally positive integers $\nu \in J_k$ of sufficiently large norm, the equation

$$\nu = \lambda_1^k + \lambda_2^k + \cdots + \lambda_s^k$$

(16.1)

is always soluble in totally non-negative integers $\lambda_j$ of $K$ with $N(\lambda_j) \leq cN(\nu)^{1/k}$ ($1 \leq j \leq s$).

Following early work of Meissner [93] and Mordell [95] for a restricted class of number fields, Siegel [102][103] was the first to obtain quite general conclusions for sums of squares, and hence, via the method of Hilbert [69], for sums of $k$th powers. Siegel later developed a proper generalisation of the Hardy-Littlewood method to number fields, and here the dissection into major and minor arcs is a particular source of difficulty. In this way, Siegel [104][105] obtained the upper bound

$$G_K(k) \leq dk(2^{k-1} + d) + 1,$$

where $d$ denotes the degree of the field $K$. If one were to break the equation (16.1) into components with respect to an integral basis for $\mathcal{O}_K$, then one would obtain $d$ equations of degree $k$ in $ds$ variables, and so one might optimistically expect the analytic part of the circle method to apply with a number of variables roughly the same in both $K$ and $\mathbb{Q}$. Perhaps motivated by such considerations, Siegel asked for reasonable bounds on $G_K(k)$ independent of $d$. This question was ultimately addressed through the work of Birch [4] and Ramamujam [98], who provided the upper bound

$$G_K(k) \leq \max\{8k^5, 2^k + 1\}.$$  

(16.2)

It is evident that the uniform bound (16.2) is far from the above cited bound $G(k) \leq (1 + o(1))k\log k$ of Wooley [155], and the only slightly weaker precursors of Vinogradov. An important desideratum, therefore, is the reduction of the bound (16.2) to one of similar order to that presently available for $G(k)$, or at least a reduction to a bound polynomial in $k$. 


Failing this, effort has been expended in pursuit of bounds of order $k \log k$, but with a modest dependence on the degree $d$ of the field $K$. Progress towards this objective has mirrored developments in the classical version of Waring’s problem. Thus, building on work generalising that of Vinogradov to the number field setting by Körner [84] and Eda [50], methods employing smooth numbers and repeated efficient differences have recently been applied by Davidson [39] to establish the bound

$$G_K(k) \leq (3 + o(1))k \log k + c_d k,$$

where

$$c_d = 4d + 3 \log \left( \frac{1}{2} (d^2 + 1) \right) + 7,$$

and the term $o(1)$ is independent of $d$ (apparently, the term $3 + o(1)$ here can be replaced by $2 + o(1)$ with only modest effort). Moreover, when $K$ has class number one, Davidson [40][42] has obtained the bound

$$G_K(k) \leq k (\log k + \log \log k + c_d),$$

where $c_d$ is approximately $4d$. Finally, Davidson [41] has improved on earlier work of Tatzuawa [106] to establish the strikingly simple conclusion that, for every number field $K$, one has

$$G_K(k) \leq 2d (\hat{G}(k) + 2k),$$

where $\hat{G}(k)$ denotes the least number $s$ satisfying the property that the set of rational integers that can be expressed as the sum of $s$ $k$th powers of natural numbers has positive density (in particular, of course, $\hat{G}(k) \leq G_1(k)$).

More exotic still than the variants of Waring’s problem over number fields are those in which one works over the polynomial rings $\mathbb{F}_q[t]$. Here one must again impose restrictions on the size of polynomials employed in the representation (and in this situation, the degree of a polynomial provides a measure of its size). Analogues of the Hardy-Littlewood method have been devised in this polynomial ring setting (see, for example, Effinger and Hayes [51]). Unfortunately, however, Weyl differencing proves ineffective whenever $k$th powers are considered over $\mathbb{F}_q[t]$ with $k \geq \text{char}(\mathbb{F}_q)$, for in such circumstances the factor $k!$, introduced into the argument of the exponential sum over $k$th powers via the differencing argument, is equal to zero in $\mathbb{F}_q[t]$. Consequently, one frequently restricts attention to $k$th powers with $k$ smaller than the characteristic (but see Car and Cherly [22] for results on sums of 11 cubes in $\mathbb{F}_{2^h}[t]$). With this restriction, Car [20][21] has shown that every polynomial $M$, with $M \in \mathbb{F}_q[t]$, of sufficiently large degree, can be written in the form

$$M = M_1^k + \cdots + M_s^k,$$
with $M_i \in \mathbb{F}_q[t]$ of degree smaller than $1 + (\deg(M))/k$ ($1 \leq s \leq k$), provided that

$$s \geq \min\{2^k + 1, 2k(k - 1)\log 2 + 2k + 3\}.$$ 

### 17 Open Problems and Conjectures

Returning temporarily to the methods of §12, there are a number of problems connected with the mean value

$$I_m = \int_0^1 |F_j(\alpha)|^{2^m} d\alpha$$

which suggest some interesting questions. We will concentrate on the situation with $m = 1$, and remark only that the available results become less satisfactory as $m$ increases.

A simple combinatorial argument reveals that the difference polynomial $\Psi_j = \Psi_j(z; h; m)$ is given explicitly by the formula

$$\Psi_j = k!2^{j}h_1 \ldots h_j \sum_{u \geq 0} \sum_{v_1 \geq 0} \cdots \sum_{v_j \geq 0} z^u (h_1 m_1^k)^{2v_1} \cdots (h_j m_j^k)^{2v_j} u!(2v_1 + 1)\cdots(2v_j + 1)!,$$

where the summation is subject to the condition $u + 2v_1 + \cdots + 2v_j = k - j$. Consequently, one has

$$\Psi_j = h_1 \ldots h_j z^d \sum_{r=0}^{\frac{1}{2}(k-j-d)} c_r(h_1 m_1^k, \ldots, h_j m_j^k) z^{2r},$$

where

$$d = \begin{cases} 0, & \text{when } k - j \text{ is even,} \\ 1, & \text{when } k - j \text{ is odd,} \end{cases}$$

and for $0 \leq r \leq (k - j - d)/2$, the coefficients $c_r(\xi)$ are polynomials with positive integral coefficients that are symmetric in $\xi_1^2, \ldots, \xi_j^2$, and have total degree $k - j - 2r - d$. Now the mean value

$$I_1 = \int_0^1 |F_j(\alpha)|^2 d\alpha$$

is equal to the number of solutions of the diophantine equation

$$\Psi_j(z; h; m) = \Psi_j(z'; h'; m'),$$

(17.1)
with the variables in ranges discernible from the definition of \( F_j(\alpha) \) in §12, and one might hope that the total number of solutions is close to the number of diagonal solutions, which is to say that

\[
I_1 \ll P^{1+\varepsilon}M_j \tilde{H}_j.
\]

When \( k-j \) is odd (so \( d = 1 \)), the presence of the term \( z^d \) makes it especially easy to deal with equation (17.1) by exploiting the inherent multiplicative structure, and indeed one can achieve the desired bound provided also that \( j \leq (k-d)/3 \). The cases \( j = 1 \) and \( k-j = 2 \) or 4 are also doable. However, when \( k-j \) is even and \( 1 < j \leq k-6 \), the situation is not so easy. By the way, this difficulty already occurs in Davenport’s work [31]. To illustrate this situation, the simplest special case that we cannot handle directly corresponds to \( k = 8 \) and \( j = 2 \), and here one has

\[
\Psi_j = h_1h_2(224z^6 + 1120z^4(h_1^2m_1^{16} + h_2^2m_2^{16}) \\
+ z^2(672h_1^4m_1^{32} + 2240h_1^2h_2^2m_1^{16}m_2^{16} + 672h_2^4m_2^{32}) \\
+ 32h_1^6m_1^{48} + 224h_1^4h_2^4m_1^{32}m_2^{16} + 224h_1^2h_2^4m_1^{16}m_2^{32} + 32h_2^6m_2^{48}).
\]

This suggests various general questions.

- Suppose that \( f, g \in \mathbb{Z}[x] \). Are there simple conditions on \( f, g \) such that the number \( N \) of integral points \( (x, y) \in [-P, P]^2 \) for which \( f(x) = g(y) \) satisfies

\[
N \ll (P\mathcal{H}(f)\mathcal{H}(g))^{\varepsilon}?
\]

Here \( \mathcal{H}(h) \) denotes the height of \( h \). A qualitative version of this has already been considered by Davenport, Lewis and Schinzel [38], and if \( f(x) - g(y) \) is irreducible over \( \mathbb{C} \), then a celebrated theorem of Siegel shows that the number of solutions is finite unless there is a rational parametric solution of special form.

By the way, in view of the above, it is perhaps not surprising that in our treatment of \( I_m \) the bound of Bombieri and Pila [6] plays a rôle.

- Suppose that \( A \subset \mathbb{Z}^k \cap [-X, X]^k \). Let \( R(n; A) \) denote the number of solutions of the equation

\[
a_1x + a_2x^2 + \cdots + a_kx^k = n
\]

with \( x \in \mathbb{Z} \cap [-P, P] \) and \( a \in A \). Are there any simple conditions under which it is true that

\[
\sum_n R(n; A)^2 \ll \text{card}(A)P(XP)^{\varepsilon}?
\]
• A well-known conjecture in connection with Waring’s problem is Hypothesis K (Hardy and Littlewood [62]). Let
\[ R_{k,s}(n) = \text{card}\{x \in \mathbb{N}^s : x_1^k + \cdots + x_s^k = n\} \]
Then Hypothesis K asserts that for each natural number \( k \), one has
\[ R_{k,k}(n) \ll n^{\varepsilon}. \]  
(17.2)
From this it would follow that
\[ G(k) \leq \max\{2k + 1, \Gamma_0(k)\} \]  
(17.3)
and
\[ G_1(k) = \max\{k + 1, \Gamma_0(k)\}. \]  
(17.4)
The conjecture (17.2) was later shown by Mahler [90] to be false for \( k = 3 \), and indeed his counter-example shows that, infinitely often, one has \( R_{3,3}(n) > 9^{-1/3}n^{1/12} \). However, the conjecture is still open when \( k \geq 4 \), and for (17.3) and (17.4), it suffices to know that
\[ \sum_{n \leq N} R_{k,k}(n)^2 \ll N^{1+\varepsilon}. \]

Hooley [74] established this when \( k = 3 \), under the assumption of the Riemann Hypothesis for a certain Hasse-Weil L-function. As far as we know, no simple conjecture of this kind is known from which it would follow that \( G(k) = \max\{k + 1, \Gamma_0(k)\} \).

• It may well be true that, when \( k \geq 3 \), one has
\[ \sum_{n \leq N} R_{k,k}(n)^2 \sim CN. \]
However Hooley [72] has shown, at least when \( k = 3 \), that the constant \( C \) here is larger than what would arise simply from the major arcs. This leads to some interesting speculations. The number of solutions of the equation
\[ x_1^k + \cdots + x_s^k = y_1^k + \cdots + y_s^k \leq N, \]
in which the variables on the right hand side are a permutation of those on the left hand side, is asymptotic to \( C_1 N^{s/k} \), for a certain positive number \( C_1 = C_1(k,s) \), and when \( s < k \) the contribution arising from the major arcs is smaller. Maybe one should think of these solutions as being “trivial”, “parametric”, or as arising from some “degenerate” property of the geometry of the surface. Anyway, their contribution is mostly concentrated
on the minor arcs. It seems rather likely that this phenomenon persists for \( s \geq k \) and explains Hooley’s discovery. This leads to the philosophy that the major arcs correspond to non-trivial solutions, and the minor arcs to trivial solutions. There is an example of this phenomenon in Vaughan and Wooley [133].

- Recall the definitions of \( f(\alpha) \), \( S(q,a) \) and \( v(\beta) \) from §7. One can conjecture that, whenever \((a,q) = 1\), one has

\[
f(\alpha) - q^{-1} S(q,a)v(\alpha - a/q) \ll (q + P^k|q\alpha - a|)^{1/k}. \tag{17.5}
\]

Possibly the exponent has to be weakened to \( 1/k + \varepsilon \), but any counterexamples would be interesting.

From (17.5) it would follow that

\[
\sum_{n \leq N} R_{k,k}(n)^2 \ll N.
\]

Also it is just conceivable that (17.5), in combination with a variant of the Hardy-Littlewood-Kloosterman method, would achieve the bound \( G(k) \leq \max\{\mathfrak{G}(k), \Gamma_0(k)\} \), where \( \mathfrak{G}(k) < 2k + 1 \).

- The inequality (17.5) is a special case of conjectures that can be made about the exponential sum

\[
f(\alpha) = \sum_{x \leq P} e(\alpha_1 x + \cdots + \alpha_k x^k)
\]

that would have many consequences in analytic number theory. For example one can ask if something like

\[
f(\alpha) - q^{-1} S(q,a)v(\alpha - a/q) \ll \sum_{j=2}^{k} \left( q + P^j|a_jq - a_j| \right)^{1/j} \tag{17.6}
\]

is true. Here

\[
S(q,a) = \sum_{r=1}^{q} e((a_1 r + \cdots + a_k r^k)/q),
\]

\[
v(\beta) = \int_{0}^{P} e(\beta_1 \gamma + \cdots + \beta_k \gamma^k) d\gamma,
\]

and \( a_1 = c(q,a_2,\ldots,a_k) \), where \( c = c(q,a) \) is the unique integer with

\[-\frac{1}{2} < c - \alpha_1 q/(q,a_2,\ldots,a_k) \leq \frac{1}{2}.
\]
The inequality (17.6) is known to hold for $k = 2$ (Vaughan unpublished) with the right hand side weakened slightly to

$$
\left( \frac{q}{(q, a_2)} \right)^{1/2} \left( \log \frac{2q}{(q, a_2)} + P|\alpha_2 - a_2/q_2|^{1/2} \right).
$$

- One way of viewing the Hardy-Littlewood method is that we begin by considering the Fourier transform with respect to Lebesgue measure on $[0, 1)$ for an appropriate generating function defined in terms of the additive characters, and then approximate to it by a product of discrete measures. Since part of the problem when one has relatively few variables is that the geometry genuinely intrudes, one can ask whether we are using the best measure for the problem at hand. In this situation something more closely related to the underlying geometry might be more useful.

In conclusion, it is clear that although we have come a long way in the twentieth century, there remains plenty still to be done!

References


[60] ______, Some problems of “Partitio Numerorum”: II. Proof that every large number is the sum of at most 21 biquadrates, Math. Z. 9 (1921), 14–27.


[88] , On the representation of large numbers as sums of seven cubes, Mat. Sb. 12 (1943), 218–224.


[91] , On the fractional parts of the powers of a rational number (II), Mathematika 4 (1957), 122–124.


K. F. Roth, *Proof that almost all positive integers are sums of a square, a positive cube and a fourth power*, J. London Math. Soc. 24 (1949), 4–13.


Waring’s Problem: A Survey


Waring's Problem: A Survey


