2

Waring's problem: history

In his *Meditationes algebraicae* (1770), Edward Waring made the statement that every number is expressible as a sum of 4 squares, or 9 cubes, or 19 biquadrates, 'and so on'. By the last phrase, it is presumed that he meant to assert that for every $k \geq 2$ there is some s such that every positive integer N is representable as

$$N = x_1^k + x_2^k + \cdots + x_s^k, \tag{2.1}$$

for $x_i \geq 0$. This assertion was first proved by Hilbert in 1909. Hilbert's proof was a very great achievement, though some of the credit should go also to Hurwitz, whose work provided the starting point. Hurwitz had already proved that if the assertion is true for any exponent k, then it is true for $2k$. I shall not discuss Hilbert's method of proof here; for this one may consult papers by Stridsberg [79], Schmidt [72] or Rieger [71]. It is usual to denote the least value of s, such that every N is representable, by $g(k)$. The exact value of $g(k)$ is now known for all values of k.

The work of Hardy and Littlewood appeared in several papers of the series 'On Partitio Numerorum' (P.N.), the other papers of the series being concerned mainly with Goldbach's ternary problem. In P.N. I [37] they obtained an asymptotic formula for $r(N)$, the number of representations of N in the form (2.1) with $x_i \geq 1$, valid provided $s \geq s_0(k)$, a certain explicit function of k. The asymptotic formula was of the following form:

$$r(N) = C_{k,s} N^{s/k-1} \mathfrak{S}(N) + O(N^{s/k-1-\delta}), \tag{2.2}$$

where $\delta > 0$ and

$$C_{k,s} = \frac{\Gamma(1+1/k)^s}{\Gamma(s/k)} > 0.$$

In the above formula, $\mathfrak{S}(N)$ is an infinite series of a purely arithmetical nature, which Hardy and Littlewood called the *singular series*. They proved further that

$$\mathfrak{S}(N) \geq \gamma > 0, \tag{2.3}$$

for some γ independent of N, provided that $s \geq s_1(k)$. However they did not at that stage give any explicit value for $s_1(k)$. Thus the formula implies that

$$r(N) \sim C_{k,s} N^{s/k-1} \mathfrak{S}(N) \tag{2.4}$$

as $N \to \infty$, provided $s \geq \max(s_0(k), s_1(k))$, and thereby provided an independent proof of Hilbert's theorem.

Hardy and Littlewood introduced the notation $G(k)$ for the least value of s such that every sufficient large N is representable in the form (2.1); this function is really of more significance than $g(k)$, since the latter is affected by the difficulty of representing one or two particular numbers N. In P.N. II [38] and P.N. IV [39], Hardy and Littlewood proved that the asymptotic formula and the lower bound for $\mathfrak{S}(N)$ both hold for $s \geq (k-2)2^{k-1} + 5$, which implies that

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

In P.N. VI [40] they found a better upper bound for $G(k)$, though not for the validity of the asymptotic formula, and in particular they proved that $G(4) \leq 19$. The last paper of the series, P.N. VIII [41], was entirely concerned with the singular series and with the congruence problem to which it gives rise.

Hardy and Littlewood took as their starting point the generating function for $r(N)$, that is, the power series

$$\sum_{N=0}^{\infty} r(N) z^N = \left(\sum_{n=0}^{\infty} z^{n^k} \right)^s.$$

They expressed $r(N)$ in terms of this function by means of Cauchy's formula for the coefficients of a power series, using a contour integral taken along the circle $|z| = \rho$, where ρ is slightly less than 1. A helpful technical simplification was introduced by Vinogradov in 1928; this consists of replacing the power series by a finite exponential sum, and the effect is to eliminate a number of unimportant complications that occurred in the original presentation of Hardy and Littlewood.

Write $e(t) = e^{2\pi it}$. We define $T(\alpha)$, for a real variable α, by

$$T(\alpha) = \sum_{x=1}^{P} e(\alpha x^k), \tag{2.5}$$

where P is a positive integer. Then

$$(T(\alpha))^s = \sum_m r'(m)e(m\alpha), \tag{2.6}$$

where $r'(m)$ denotes the number of representations of m as

$$x_1^k + \cdots + x_s^k, \quad (1 \le x_i \le P).$$

If $P \ge [N^{1/k}]$, where $[\lambda]$ denotes the integer part of any real number λ, then $r'(N)$ is the total number of representations of N in the form (2.1) with $x_i \ge 1$. Consequently $r'(N) = r(N)$. If we multiply both sides of (2.6) by $e(-N\alpha)$ and integrate over the unit interval $[0, 1]$ (or over any interval of length 1), we get

$$r(N) = \int_0^1 (T(\alpha))^s e(-N\alpha)d\alpha. \tag{2.7}$$

This is the starting point of our work on Waring's problem. It corresponds to the contour integral for $r(N)$ used by Hardy and Littlewood, with z replaced by $e^{2\pi i\alpha}$.

Our first aim will be to establish the validity of the asymptotic formula (2.2) for $r(N)$ as $N \to \infty$, subject to the condition $s \ge 2^k + 1$. It is possible to do this in a comparatively simple manner by using an inequality found by Hua in 1938 (Lemma 3.2 below). It may be of interest to observe that no improvement on the condition $s \ge 2^k + 1$ has yet been made for small values of k, as far as the asymptotic formula itself is concerned. For large k it has been shown by Vinogradov that a condition of the type $s > Ck^2 \log k$ is sufficient.

If we prove that the asymptotic formula holds for a particular value of s, say $s = s_1$, it will follow that every large number is representable as a sum of s_1 kth powers, whence $G(k) \le s_1$. But to prove this it is not essential to prove the asymptotic formula for the total number of representations; it would be enough to prove it for some special type of representation as a sum of s_1 kth powers. This makes it possible to get better estimates for $G(k)$ than one can get for the validity of the asymptotic formula. In 1934 Vinogradov proved that $G(k) < Ck \log k$ for large k, and we shall give a proof in Chapter 9. The best known results for small k were found by Davenport in 1939–41 [19].

A new 'elementary' proof of Hilbert's theorem was given by Linnik in 1943 [58], and was selected by Khintchine as one of his 'three pearls' [53]. The underlying ideas of this proof were undoubtedly suggested by certain features of the Hardy–Littlewood method, and in particular by Hua's inequality.