MATHEMATICS

MASTER'S THESIS

WARING'S PROBLEM

JANNE SUOMALAINEN

HELSINKI 2016
UNIVERSITY OF HELSINKI

HELSINGIN YLIOPISTO — HELSINGFORS UNIVERSITET — UNIVERSITY OF HELSINKI

Tiedekunta/Osasto — Fakultet/Sektion — Faculty	Laitos — Institution — Department
Faculty of Science	Department of Mathematics and Statistics

Tekijä — Författare — Author
Janne Suomalainen

Työn nimi — Arbetets titel — Title

Waring's Problem

Oppiaine — Läroämne — Subject
Mathematics

Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages
Master's Thesis	9/2016	36 p.

Tiivistelmä — Referat — Abstract

Waring's problem is one of the two classical problems in additive number theory, the other being Goldbach's conjecture. The aims of this thesis are to provide an elementary, purely arithmetic solution of the Waring problem, to survey its vast history and to outline a few variations to it.

Additive number theory studies the patterns and properties, which arise when integers or sets of integers are added. The theory saw a new surge after 1770, just before Lagrange's celebrated proof of the four-square theorem, when a British mathematician, Lucasian professor Edward Waring made the profound statement nowadays dubbed as Waring's problem: for all integers n greater than one, there exists a finite integer s such that every positive integer is the sum of s nth powers of non-negative integers. Ever since, the problem has been taken up by many mathematicians and state of the art techniques have been developed — to the point that Waring's problem, in a general sense, can be considered almost completely solved.

The first section of the thesis works as an introduction to the problem. We give a profile of Edward Waring, state the problem both in its original form and using present-day language, and take a broad look over the history of the problem. The main emphasis is on the classical version of the problem, whereas the modern version is described in Section 5 with numerous other variations. In addition, generalizations to integer-valued polynomials and to general algebraic fields are described. Goldbach's conjecture is also briefly illustrated.

The elementary solution of Waring's problem is presented in Sections 2 to 4. Historical perspective is carried through the thesis with the profiles of the key mathematicians to the solution. The proof presented is an improved and simplified version of Yuri Linnik's solution of Waring's problem. The second section provides the groundwork, an ingenious density argument by Lev Shnirelman, which is applied to the problem in the so called Fundamental lemma presented in Section 3. The proofs of the intermediate results needed to prove the lemma are presented in the following sections. The third section reduces the proof to an estimation of the number of solutions of a certain system of Diophantine equations. The final argument, longish induction is given at the end of the fourth section.

Even though Waring's problem is solved, the progress made in the field is far from being idle. The plethora of variations and generalizations, especially Ideal Waring's problem, Modern Waring's problem and Waring–Goldbach problem are actively studied today. It is surprising how deep a problem with such a seemingly simple assertion can be. Conclusively, the challenge in this branch of mathematics is to develop new mathematical methods to prove and explain what seems so obvious.

Avainsanat — Nyckelord — Keywords
Additive number theory, Waring's problem, History of mathematics

Säilytyspaikka — Förvaringsställe — Where deposited
Kumpula Campus Library

Muita tietoja — Övriga uppgifter — Additional information

Advisor prof. Eero Saksman

Contents

1. Additive number theory and Waring's problem

A good deal is known about the multiplicative properties of the integers. We have, for example, the *Fundamental theorem of arithmetic*, every integer has a unique prime decomposition up to the order of the factors. Decomposing integers additively is a much less studied problem. In additive number theory we are interested in the patterns and properties, which arise when integers or sets of integers are added. For instance, how many ways are there to write an integer as the sum of two squares? What about four squares? In how many ways can we write the number one as the sum of three cubes? Is it true that every number is the sum of two primes?

This thesis focuses on a famous problem in additive number theory, dubbed *Waring's problem*. Can number 45 be written as as a sum of at most four squares? Yes, indeed $45 = 3^2 + 6^2$. What about 399? For example, $399 = 3^2 + 5^2 + 13^2 + 14^2$. How about 1963? No problem, $1963 = 9^2 + 19^2 + 39^2$. How many cubes does it take to represent the same numbers?

$$45 = 1^3 + 1^3 + 2^3 + 2^3 + 3^3$$

$$399 = 1^3 + 1^3 + 3^3 + 3^3 + 7^3$$

$$1963 = 3^3 + 5^3 + 5^3 + 7^3 + 7^3 + 10^3$$

We could extend this question to any non-negative integer and to any positive exponent we choose; Waring's problem is about the possibility to represent any non-negative integer as a sum of a finite number of positive integer powers. Additive number theory is nowadays an ample and blooming subject, which grew from Waring's problem and various generalizations of it.

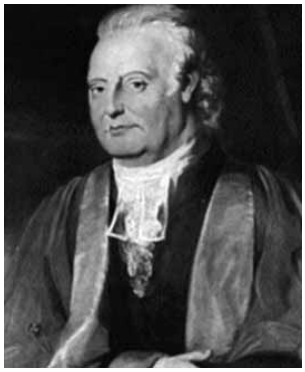
This section comprises of an introduction to the man after whom Waring's problem is named, the explicit statement of the problem, and a broad overview of its history. The main emphasis is on the classical version of the problem, whereas the modern version is described in Section 5 with numerous other variations of the problem. The other classical problem in additive number theory, Goldbach's conjecture is also briefly illustrated.

An elementary solution to Waring's problem is presented in Sections 2 to 4. Herein 'elementary' does not mean simple but a type of a solution, which requires no concepts or methods transcending the limits of basic arithmetic. A historical perspective is carried through the thesis with the profiles of the key mathematicians to the solution. The proof presented is an improved and simplified version of Linnik's solution of Waring's

problem. Our main reference for the exposition is the methodological paper by Nesterenko (2006). Section 2 provides the groundwork, an ingenious density argument by Shnirelman, which is applied to the problem in the Fundamental lemma of Section 3. The intermediate results needed in the proof of the Fundamental lemma are proved in Sections 3 and 4, after which the solution is complete.

Finally, we provide an extensive list of bibliographic references at the end of the thesis. On a few occasions the reader is directed to the references contained therein. Portraits of Waring and Shnirelman are from the great online gallery maintained by Swetz (2007), while Hua's and Linnik's in turn are from (Wikimedia Commons 2015) and (Russian Academy of Sciences 2002) respectively. All photos used are in the public domain.

1.1. **Edward Waring.** Edward Waring was a British mathematician born in Shropshire, England in 1736. The eldest son born and raised on a farm, he was educated in Shrewsbury school before entering Magdalene College, Cambridge on a scholarship at the age of seventeen. Waring's mathematical talents impressed the teachers and he graduated with a bachelor's degree as a senior wrangler (the top mathematics undergraduate) in 1757. One year later he was elected a fellow of the college.



Portrait of Edward Waring

Before graduating, Waring worked on his *Meditationes Algebraicae*, covering topics around the theory of equations, number theory and geometry. He submitted the first chapter of the book to the Royal Society but was bluntly ignored. Only after Waring had been nominated for the Lucasian Chair of Mathematics, one of the highest positions in Cambridge, the first chapter of the book was issued as *Miscellanea Analytica* in 1759. The publication worked as a qualification, proof that Waring was competent enough for the post in spite of his young age.

The change of decade was comprised of quick-tempered exchange of pamphlets between Waring and William Powell of St John's College, Cambridge, who doubted Waring's mathematical abilities and tried to prevent him being appointed to the position. It was John Wilson's supporting letter and a master's degree granted by the royal mandate that finalized Waring's confirmation as Lucasian professor at the age of just 23, holding the chair until his death.

Miscellanea Analytica was published as a complete work in 1762, after which Waring was elected a fellow of the Royal Society. A new version of the book was published in 1776 and further extended in 1785. Surprisingly, Waring graduated with a postgraduate degree as a doctor of medicine in 1767 and managed to practise a few years in various hospitals in Eastern England.

Most notably, *Meditationes Algebraicae* published in 1770 and expanded twelve years later (translated as Waring 1991) had the greatest distribution of Waring's books. He worked on the theory of symmetric functions; Waring can be considered as the earliest contributors to Galois theory. He proved a generalization of Bézout's theorem and was the first to publish numerous conjectures in number theory, such as Goldbach's conjecture, Goldbach's weak conjecture, Waring's problem and Wilson's theorem.

Even though Waring had been fairly acknowledged and even awarded — among other honours the Copley Medal of the Royal Society in 1784 —, he resigned his position in the society in 1795 claiming poverty and old age. Waring died because of a violent cold just three years later. Considered as a somewhat misunderstood, 'a good though not great mathematician' (Batchelder 1936, p. 21) with poor communication skills and inadequate algebraic notation, Waring never lectured during his years at the university, and his books did not achieve the audience they deserved. For these reasons he have been compared to Paolo Ruffini. (O'Connor and E. F. Robertson 2015; J. C. Robertson and Byerley 1821)

1.2. **Waring's problem.** Waring makes in his *Meditationes Algebraicae*, without a proof, the thoughtful statement that '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'. In a later edition he adds cautiously that 'similar laws may be affirmed (*exceptis excipiendis*) for the correspondingly defined numbers of quantities of any like degree'. (Waring 1991, p. 336)

It was the comments 'and so forth' and 'similar laws may be affirmed' that gave rise to the problem nowadays dubbed as the Waring's problem. Apparently, based on numerical evidence, Waring conjectured that, for each exponent $n \geqslant 2$, some fixed number of non-negative nth powers is sufficient to represent all positive integers. The smallest such integer that suffices is usually denoted by $g(n)$ to emphasize its only dependence on n.

Definition 1.1. Let n be a natural number greater than one. Now $g = g(n)$ is the smallest number such that every positive integer is a sum of at most g nth powers. If such a finite g does not exist, we set $g = \infty$.

Using this notation, Waring conjectured that $g(3) = 9$ and $g(4) = 19$. In addition, he stated that the sequence 4, 9, 19,... went on.

For example,

$$\begin{aligned} 454 &= 3^2 + 11^2 + 18^2, \\ &= 1^3 + 1^3 + 1^3 + 3^3 + 3^3 + 3^3 + 3^3 + 7^3, \\ &= 1^4 + 2^4 + 2^4 + 2^4 + 3^4 + 3^4 + 3^4 + 3^4 + 3^4 \\ &= 1^5 + 1^5 + 1^5 + 1^5 + 1^5 + 1^5 + 2^5 + 2^5 + 2^5 + 2^5 + 2^5 \\ &\qquad + 2^5 + 2^5 + 2^5 + 2^5 + 2^5 + 2^5 + 2^5 + 2^5 + 2^5 \end{aligned}$$

are representations of number 454 as a sum of 2nd, 3rd, 4th and 5th integer powers, from which we can conclude at once that $g(2) \geqslant 3$, $g(3) \geqslant 8$, $g(4) \geqslant 9$ and $g(5) \geqslant 20$. Naturally, we get to the following questions — the first one being, in a restricted sense, the famous Waring's problem.

Question 1.2. Is $g(n)$ finite for all n?

Question 1.3. Can we find interesting lower and upper bounds for $g(n)$?

Question 1.4. **i)** For a fixed n, can we determine the exact value of $g(n)$?
ii) Can we find an explicit formula for $g(n)$ that works for all n?

Question 1.5. Can we find a formula for $g(n)$ that works for sufficiently large values of n?

Let us start with the first questions 1.2 and 1.3, the problem of existence, and the numerical problem 1.4. The last question, the asymptotic problem 1.5 will be answered much later in Section 5.1.

1.2.1. *The case $n = 2$.* The problem for squares is very old. The fact that every natural number can be presented as a sum of four squares of non-negative integers was already hinted by Diophantus around third century. The explicit statement was done by Bachet in 1621 and later by Fermat in 1640. In general, the case $n = 2$ is well-known and completely solved in the light of the following eminent theorems proved in 1640, 1797 and 1770 respectively.

Theorem 1.6 (Fermat's theorem on sums of two squares)**.** *Every prime number p can be written as the sum of two squares of integers, if and only if $p = 2$ or p is of the form $4n + 1$, where n is an integer, that is $n \in \mathbf{Z}$.*

Theorem 1.7 (Legendre's three-square theorem)**.** *Every positive integer a can be written as the sum of three squares of integers, if and only if a is not of the form $a = 4^n(8m + 7)$, where n, $m \in \mathbf{Z}$.*

Theorem 1.8 (Lagrange's four-square theorem)**.** *Every positive integer can be written as the sum of at most four squares.*

For proofs, see for example (Hardy, Wright et al. 2008, pp. 395–408; Pollack 2004, pp. 91–103). For Legendre's three-square theorem in particular, the paper by Wagstaff (1975) is very interesting, since it uses a similar density argument that we introduce in the next section. All the same, we can conclude that $g(2) = 4$.

According to Nathanson, Lagrange's four-square theorem is 'the most important result in additive number theory' (Nathanson 1996, p. 5). Lagrange posed his proof of the theorem just few months after Waring's conjecture, starting the long road of different results and new techniques around the conjecture. As a base for the following overview of the historical evolution towards the solution of Waring's problem, we use the great surveys by Calderón (2011), Dickson (1920), Ellison (1971), Hardy, Wright et al. (2008) and Vaughan and Wooley (2002).

1.2.2. *First bounds for $g(n)$.* Very shortly after Waring's conjecture around 1772, J. A. Euler found a lower bound for $g(n)$. The argument (see Euler 1862, pp. 203–204) is very neat and exceptionally elementary, which is why we reproduce it below. What is remarkable about the bound achieved is that by current knowledge it is the best possible!

Theorem 1.9. *For $n \geqslant 2$*

$$g(n) \geqslant \left\lfloor \left(\frac{3}{2}\right)^n \right\rfloor + 2^n - 2,$$

where $\lfloor \cdot \rfloor$ denotes the floor function*, that is $\lfloor a \rfloor = \max\{m \in \mathbf{Z} \mid m \leqslant a\}$.*

Proof. Denote $q = \lfloor (3/2)^n \rfloor$ and consider the number $k = 2^n q - 1 < 3^n$. Clearly, only the terms 1^n and 2^n can sum up to k. To minimize the needed number of summands,

we use as many 2^ns as possible. The smallest number of summands is given in

$$k = (q-1)2^n + (2^n - 1)1^n,$$

that is, k requires $(q-1)$ 2^ns and (2^n-1) 1^ns. Thus $g(n) \geqslant q + 2^n - 2$. □

Denote $\underline{g}(n) = \lfloor (3/2)^n \rfloor + 2^n - 2$. If we tabulate values of $\underline{g}(n)$ we get the following.

TABLE 1. Values of $\underline{g}(n)$ (OEIS Foundation Inc. 2016).

n	2	3	4	5	6	7	8	9	10	11	12	...
$\underline{g}(n)$	4	9	19	37	73	143	279	548	1079	2132	4223	...

Now that $g(n)$ is bounded from below we can focus on finding some upper bounds for it — essentially finding an answer to our Question 1.3. For almost a century after Euler's estimate there were particularly no published attempts at solving Waring's problem; the problem seemed to be insoluble. The first results were achieved by brute force, using huge tabulations of compositions of numbers. Following Jacobi's idea, Zornow (1835) tabulated all integers up to 3000 as sums of as few cubes as possible. He concluded that all positive integers not exceeding 3000 can be expressed as a sum of at most 9 cubes, confirming Waring's assertion for $a \leqslant 3000$. Dase extended Zornow's table to 12 000, which was published by Jacobi (1851). Similarly for higher powers, tables by Bretschneider (1853) verified that 19 biquadrates is sufficient to represent all positive integers not exceeding 4100. As far as 4096, he confirmed that 37 fifth powers and that 73 sixth powers is needed. All of these and many other tabulations have later been vastly extended by the use of more sophisticated algorithms and with the help of powerful computers. Note how well the aforementioned limits correspond the values of $\underline{g}(n)$ (Table 1).

Tables were by no means the only results achieved in the 19th century; the observations made from the tables were — one by one — backed up by proofs. The situation is well described by Hardy, according to whom 'in the Theory of Numbers it is singularly easy to speculate, though often terribly difficult to prove; and it is only proof that counts' (Hardy 2011, p. 16). Liouville, a great French mathematician found a concrete upper bound for $g(4)$ during his years at the Collège de France: every positive integer is the sum of at most 53 biquadrates. The solution was apparently presented in his lectures, and the argument is printed in (Lebesgue 1859, pp. 112–115). The proof uses Lagrange's four-square theorem and the *Liouville polynomial identity*. For short, the identity

$$\begin{aligned} 6\left(x^2+y^2+z^2+t^2\right)^2 = {} & (x+y)^4 + (x+z)^4 + (y+z)^4 + (x+t)^4 + (y+t)^4 + (z+t)^4 \\ & + (x-y)^4 + (x-z)^4 + (y-z)^4 + (x-t)^4 + (y-t)^4 + (z-t)^4 \end{aligned}$$

verifiable at once, gives that any square times 6 is a sum of 12 biquadrates. Yet any number can be written in the form $6p + r$, where p is a natural number, that is $p \in \mathbf{N}$, and $r \in \{0, 1, 2, 3, 4, 5\}$, r is expressible by at most five ones. By Lagrange's four-square theorem, we can write $p = n_1^2 + n_2^2 + n_3^2 + n_4^2$. Thus $6p$ is a sum of $4 \cdot 12 = 48$ biquadrates and we get $g(4) \leqslant 48 + 5$.

To be fair, this limit was still far above the ideal limit 19 suggested by tables and, as will be seen, the progress towards the ideal limit was painstakingly slow. The bound was improved in stages by the use of more refined identities and larger tables — by steadily revising the previous argument; Réalis (1878) and Lucas (1878a,b) achieved $g(4) \leqslant 47$, $g(4) \leqslant 45$ and $g(4) \leqslant 41$ respectively.

Towards the end of the century Maillet (1895) attacked the case of cubes. Application of the identity

$$(r+x)^3 + (r-x)^3 = 2r^3 + 6rx^2$$

made it possible to translate the problem of writing an integer as the sum of a certain number of cubes to writing a related integer as the sum of a smaller number of squares. Using this idea Maillet concluded that $g(3) \leqslant 21$. One year later he improved upon this by showing $g(3) \leqslant 12$ (Maillet 1896). The same paper considered also the case $g(5) \leqslant 192$. The bound for $g(4)$ was reduced to 39 by Fleck (1906), who, as well, made a remark (Fleck 1907) that Maillet's limit for $g(5)$ can easily be reduced by about 36. At the same time Fleck proved the finiteness of $g(6) \leqslant 184g(3) + 59$. Furthermore, the bound for $g(4)$ was improved by Landau (1907) and Wieferich (1908b) to 38 and 37 respectively.

Even though the ideal limits were still a far cry, generally speaking, the beginning of the 20th century saw drastically increased activity around Waring's problem. Maillet (1908) gave an elementary proof for $g(8) < \infty$, for which Hurwitz (1908) gave an explicit bound of 36 119. The proof of Hurwitz uses an identity with a right-hand side as large as 184 terms! The identity with which $g(10) < \infty$ can be obtained was found by Schur and it is printed in (Landau 1908, p. 105).

Around the same time the case of cubes was solved. Exploiting Maillet's idea (1895), Wieferich (1908a) achieved that all integers exceeding $2{,}25 \cdot 10^9$ can be written as sums of nine cubes. Greedy algorithm and the extension of Dase's table to 40 000 by von Sterneck (1903) completed the proof: every natural number is the sum of at most nine cubes, that is $g(3) \leqslant 9$, which is in accordance with tables.

Since number 23 requires 9 cubes, it must be that $g(3) = 9$. The lower bound by Euler (Theorem 1.9) could also be used. As a side note, this did not mean that all numbers required nine cubes, since we defined g as the smallest number with the property needed. To illustrate this, consider the following example:

$$\begin{aligned}
8 &= 2^3 \\
16 &= 2^3 + 2^3 \\
17 &= 2^3 + 2^3 + 1^3 \\
18 &= 2^3 + 2^3 + 1^3 + 1^3 \\
19 &= 2^3 + 2^3 + 1^3 + 1^3 + 1^3 \\
20 &= 2^3 + 2^3 + 1^3 + 1^3 + 1^3 + 1^3 \\
21 &= 2^3 + 2^3 + 1^3 + 1^3 + 1^3 + 1^3 + 1^3 \\
22 &= 2^3 + 2^3 + 1^3 + 1^3 + 1^3 + 1^3 + 1^3 + 1^3 \\
23 &= 2^3 + 2^3 + 1^3 + 1^3 + 1^3 + 1^3 + 1^3 + 1^3 + 1^3
\end{aligned}$$

$$239 = 4^3 + 4^3 + 3^3 + 3^3 + 3^3 + 3^3 + 1^3 + 1^3 + 1^3.$$

Seemingly some integers can be written as the sum of nine cubes, some as the sum of eight cubes, seven cubes,..., or one cube.

Wieferich's proof did overlook a case, which was considered later by Kempner (1912). Kempner also lowered the known limit of $g(6)$ to 920. At once Waring's problem seemed more approachable than ever, even though the question of the finiteness of $g(n)$ for all n was still unsolved. As a recapitulation, by 1909 it had been shown that $g(n) < \infty$ for a few special cases, namely $n \in \{2, 3, 4, 5, 6, 8, 10\}$. The case $n = 7$ was proved by Wieferich (1909), who achieved $g(7) \leqslant 3806$ and $g(5) \leqslant 59$ in his paper.

1.2.3. *The solution of Waring's problem.* The breaking point in solving the problem came in 1909, when the first solution to the problem was posed by one of the most influential mathematicians of the 20th century, David Hilbert — no less than 139 years after Waring's original conjecture.

Theorem 1.10 (Hilbert–Waring theorem)**.** *For all natural numbers $n \geqslant 2$ there exists a finite integer $g = g(n)$, which depends only on n, such that every $a \in \mathbf{N}$ can be represented as the sum of at most g nth powers of positive integers, that is $g(n) < \infty$ for all n.*

In other words, for every $a \geqslant 1$, there exist non-negative integers $x_1, x_2, \ldots, x_g$ such that

$$a = x_1^n + x_2^n + \ldots + x_g^n.$$

This solves Waring's problem, our Question 1.2. Note that g is chosen to be the smallest number with the property above. Especially, at least one $a \geqslant 1$ can be found, which cannot be written as a sum of $(g-1)$ nth powers.

The original proof was published in a paper (see Hilbert 1909) dedicated to the memory of Minkowski. Later on several authors simplified the proof, different versions of which are described by, for example, Ellison (1971, pp. 23–29) and Pollack (2011), and expounded by Nathanson (1996, pp. 86–93) and Nesterenko (2006, pp. 4699–4705). The proof itself is considered cumbersome, since it uses Lagrange's four-square theorem as the base of a difficult induction, relying on complicated multiple integrals and polynomial identities, such as

Theorem 1.11 (Hilbert's identity)**.** *For every pair of integers $m > 1$ and $r > 1$, there are $M = (2m+1)^r$, positive rationals b_i and natural numbers $a_{i,j}$, such that*

$$\left(x_1^2 + \ldots + x_r^2\right)^m = \sum_{j=1}^{M} b_j \left(a_{1,j}x_1 + \ldots + a_{r,j}x_r\right)^{2m}.$$

The identity is based on an integral identity and it was first conjectured by Hurwitz (1908) but proved by Hilbert (1909).

In addition, Hilbert's original version of the proof is purely an existence proof, yielding no respectable upper bound for $g(n)$. With a suitable modification such a bound can be given (Rieger 1953). As well as reviewing the progress made earlier on the topic, Pollack (2011) shows that, for every $n \geqslant 2$,

$$g(n) < (2n+1)^{1808n^5},$$

which is — as Pollack himself admits — fairly weak by today's standards (see Theorem 1.13). Nonetheless, the original proof was ground-breaking in the sense that it introduced completely new ideas to apply analysis to additive number theory. Hardy went on and praised that

> it would hardly be possible for me to exaggerate the admiration which I feel for the solution of this historic problem – – it is absolutely and triumphantly successful, and it stands with the work of Hadamard and de la Vallée-Poussin, in the theory of primes, as one of the landmarks in the modern history of the theory of numbers. (Hardy 2011, p. 24)

In the following years of Hilbert's solution the proof itself — the proof of Hilbert's identity in particular — was, for example, simplified to algebraic expressions. The first completely elementary version of the proof was published by 1911. Several known bounds for special cases of $g(n)$ were also improved. However, nothing drastic was achieved around Waring's problem until Hardy and Littlewood succeeded in applying the theory of analytic functions on the problem; like the distinguished French mathematician Poincaré had predicted, after the details in Hilbert's proof were fully understood, highly important arithmetical results would follow inevitably (Hilbert 1912, p. 10).

1.2.4. *New methods.* The general question (Question 1.2) was undertaken also by Hardy and Littlewood in a long series of papers published in the 1920s titled *On Some Problems of Partitio Numerorum I, II,…, VIII*. The theory developed was so influential that *Partitio Numerorum* has since become a synonym for additive number theory — like *Analysis situs* has become topology.

The Hardy–Littlewood method (see Hardy and Littlewood 1919, 1920) improved Hilbert's proof by offering a concrete upper bound for $g(n)$ and thus obtaining a quantitative understanding of the problem. The method originated from the study of the partition function done by Hardy and Ramanujan (1918), and it is described by Hardy (2011) and expounded by, for example, Ellison (1971, pp. 15–23) and Davenport (2005, pp. 1–66).

The starting point of a refined version of the method is the realization that the sum

$$r_{n,g}(a) = \sum_{\substack{x_1^n+\ldots+x_g^n=a \\ x_i \geqslant 0}} 1,$$

that is, the number of ways to represent an integer $a \geqslant 1$ as the sum of g nth powers of non-negative integers, can be expressed as an integral. Especially, with a *trigonometric polynomial*

$$f(\alpha) = \sum_{m=0}^{N} \exp(2\pi \mathrm{i} \alpha m^n), \quad \alpha \in \mathbf{R},$$

where $N = \lfloor a^{1/n} \rfloor$, $\exp(z)$ denotes the *exponential function* e^z and $\mathbf{R}$ the set of real numbers, it can be shown that

$$r_{n,g}(a) = \int_0^1 f(\alpha)^g \exp(-2\pi \mathrm{i} \alpha a)\, \mathrm{d}\alpha.$$

For a fixed n and large enough g, the behaviour of $r_{n,g}(a)$ can be analysed with the help of the behaviour of f. Conclusively, the integral can be shown to be positive for all sufficiently large g.

The so-called circle method developed has proved out to be a powerful analytic tool, which is still an influential force in additive number theory; starting, in a sense, a new era in the theory of numbers. Nathanson considers it as one of the two best analytical tools to attack the classical problems in additive number theory (Nathanson 1996, p. vii). The method was later improved in 1928 by Vinogradov (1985) proving the following result, a weaker form of Goldbach's weak conjecture.

Theorem 1.12 (Vinogradov's theorem)**.** *Every sufficiently large odd number is a sum of three primes.*

In addition, the method has an important role in so-called Modern Waring's problem described in Section 5.1. Very soon mathematicians learned new ways to apply the Hardy–Littlewood–Vinogradov method to finding better bounds for $g(n)$. An excellent description of the work done on determining these bounds can be found in (Hardy, Wright et al. 2008, pp. 444–450).

By the 1930s the method was adequately improved for some serious application on the determination of the formula for $g(n)$. Dickson (1936) managed to prove a formula (see Theorem 1.13) for exponents 7–180, showing specifically that $g(7) = 143$. The proof omitted a few cases, which were filled in by multiple authors, including Dickson. Not too long, a few years later Pillai (1940) established $g(6) = 73$ by an extension of the argument.

1.2.5. *Elementary and original.* With all respect to Hardy and Littlewood's deep analytical solution, it would be very interesting to find an elementary and original solution to Waring's problem, since the problem itself is fairly simple to state. We already mentioned the modification of Hilbert's argument. In turn, Linnik found another, surprisingly elementary proof (see Linnik 1943a), which we will follow in this thesis. The details of the proof are presented in the following sections. Since the proof uses results due to Shnirelman, it is dubbed occasionally the Shnirelman–Linnik approach.

The solution itself is described by Ellison (1971, pp. 12–15), expounded by Khinchin (1998, pp. 18–64), and further improved and simplified by Hua (1982, pp. 494–534), Nesterenko (2006, pp. 4706–4714) and Jameson (2015). That is to say, the presentation of the proof deviate from the original one, because it has incorporated various simplifications and enhancements introduced afterwards by many authors.

One of the shortest and maybe the most elegant known solution to Waring's problem is the so-called Linnik–Newman approach (see Newman 1960, 1997, pp. 49–56) introduced in 1960, which is refined even further by Pollack (2004, pp. 275–286). The proof uses the same density argument as Linnik's and the groundwork is very similar to that of Hardy–Littlewood method. What is ingenious about it, is the application of Weyl sums to estimating the number of ways to write an integer a as a sum of pre-defined number of nth powers. In fact, this number is the integral over the closed interval from 0 to 1 of the gth power of a Weyl sum times $\exp(-2ax\pi\mathrm{i})$, which can be shown to be bounded from above.

In addition to these new proofs — returning back to the historical evolution of solving the problem —, at last in 1964, Chen (1964) tackled the case $n = 5$ and showed $g(5) = 37$, again by a reinforcement of Dickson's argument. However, the most complex problem on determining the values of $g(n)$, was the case $n = 4$, which defied the attempts of many mathematicians until the late 20th century. In 1986, over a hundred years after Liouville's first limit, Balasubramanian, Deshouillers and Dress (1986a,b) obtained $g(4) = 19$. In the light of this last result the classical Waring's problem is completely solved and our Question 1.4.i is answered. This completes our overview.

LeVeque (1996, p. 189) does note that Waring's problem turned out to be a 'splendid problem' in a sense that its manageable enough to be appealing, but the proof of which is challenging enough — as we have seen — to encourage the development of new techniques; Waring's problem is the epitome of number-theoretic problems. Citing Small, 'it is one of those nasty gems, like Fermat's Last Theorem, which begins with a simply-stated assertion about natural numbers, and leads quickly into deep water' (Small 1977b, p. 13). Indeed, both in Waring's problem and in Fermat's theorem the question is about basic properties of nth powers of integers. In Fermat's theorem one must prove that the sum of two nth powers of natural numbers cannot be the nth power of a third.

1.2.6. *Ideal Waring's problem.* Recall Theorem 1.9. The problem of showing the equality $g(n) = \underline{g}(n)$ for all n is known as Ideal Waring's theorem; comparing the values of $\underline{g}(n)$ (Table 1) and already achieved values of $g(n)$, we can perceive the level of accuracy in Euler's theorem. Thanks to many years of work of many mathematicians, Dickson, Pillai, Chen, Rubugunday and Niven to name a few, we are very close to the explicit formula — an answer to our Question 1.4.ii. Full references can be found in the bibliography of (Vaughan 1997).

Theorem 1.13. *Denote* $q = \lfloor (3/2)^n \rfloor$ *and* $p = \lfloor (4/3)^n \rfloor$. *For* $n \geqslant 2$

$$g(n) = \begin{cases} q + 2^n - 2, & \text{if} \quad 2^n \left(\left(\frac{3}{2} \right)^n - q \right) + q \leqslant 2^n \\ 2^n + q + p - \theta, & \text{otherwise,} \end{cases} \tag{1}$$

where

$$\theta = \begin{cases} 2, & \text{if} \quad pq + p + q = 2^n \\ 3, & \text{if} \quad pq + p + q > 2^n \end{cases}$$

The first alternative in (1) has been checked to hold for $n \leqslant 471\,600\,000$ by Kubina and Wunderlich (1990), and, in fact, there is at most a finite number of exceptions, if any (Mahler 1957). Thus it is reasonable that Euler's estimate of $g(n)$ is generally believed to be the exact value of $g(n)$.

To complete the proof of Ideal Waring's theorem, it is sufficient to show that the inequality

$$\left(\frac{3}{2} \right)^n - \left\lfloor \left(\frac{3}{2} \right)^n \right\rfloor \leqslant 1 - \left(\frac{3}{4} \right)^{n-1}$$

holds for all n. The most recent advancement on this is by Pupyrev (2009), who — as well as describes many of the earlier results on the topic — proves that

$$\left(\frac{3}{2}\right)^n - \left\lfloor \left(\frac{3}{2}\right)^n \right\rfloor \leqslant 1 - a^n,$$

where $a = 0{,}5795$, as long as $n \geqslant 871\,387\,440\,264$. Conclusively, despite we having gotten very far in 246 years, there is much yet to be done before Ideal Waring's problem is completely solved. Without entering too much into it, and now that all questions 1.2, 1.3 and 1.4 are answered, we move on to the reasoning itself, Linnik's elementary proof on the finiteness of $g(n)$.

2. The density of a sequence

In this section we construct the principal argument, with which the proof of Hilbert–Waring theorem can be formulated using only basic techniques. This same argument is used in all known solutions of Waring's problem, apart from those of Hilbert's and Hardy–Littlewood–Vinogradov's. The fundamental definition, Shnirelman density is due to Lev Shnirelman (see Shnirel'man 1933), who developed it originally to attack Goldbach's conjecture.

2.1. **Shnirelman density.** We will pursue the following notation throughout this whole section unless otherwise noted. Let $(A) = (a_0, a_1, \ldots)$ and $(B) = (b_0, b_1, \ldots)$ be infinite, strictly monotonically increasing sequences of integers beginning with zero. For every $n \geqslant 0$ we denote by $A(n)$ the *counting function* of sequence (A); $A(n)$ is the number of positive members of sequence (A) that do not exceed n. For example, if $(A) = (0, 2, 4, 6, 8, \ldots)$, that is the even numbers, then $A(0) = 0$, $A(2) = 1$, $A(10) = 5$ and $A(231) = 115$.

By definition $0 \leqslant A(n) \leqslant \lfloor n \rfloor \leqslant n$ holds so we get

$$0 \leqslant \frac{A(n)}{n} \leqslant 1. \tag{2}$$

The fraction in (2) has different values for different n. Following Shnirelman's example, we interpret it as a measure of sequence's density in the segment from 1 to n of the sequence of natural numbers.

Definition 2.1. *Shnirelman density* of sequence (A) is the number

$$\sigma(A) = \inf_{n \geqslant 1} \frac{A(n)}{n},$$

where inf denotes *infimum* and $A(n)$ the counting function of (A).

Since the density $\sigma(A)$ — sometimes also called the natural density of (A) — is the greatest lower bound of all values of the fraction, (2) gives

$$0 \leqslant \sigma(A) \leqslant 1. \tag{3}$$

Let us make other useful observations about $\sigma(A)$. After getting well acquainted with it but before continuing with the theory of Shnirelman density, we will take a brief look at the man after whom the concept is named.

Theorem 2.2. **i)** *If sequence* (A) *does not contain* 1*, then* $\sigma(A) = 0$.

ii) $\sigma(A) = 1$ *if and only if* (A) *coincides with* **N**.

Proof. **i)** Let us assume that 1 is not a member of (A). We get $A(1) = 0$ so $A(1)/1 = 0$, and by (3) $\sigma(A) \geqslant 0$. Thus $\sigma(A)$ must be zero.

ii) Suppose first that $\sigma(A) = 1$. Let us assume to the contrary that (A) does not coincide with **N**; let k be the smallest positive integer, which is not a member of (A). We get $A(k) \leqslant k - 1$ and thus $\sigma(A) \leqslant A(k)/k \leqslant 1 - 1/k < 1$, which is a contradiction.

Suppose then that (A) coincides with **N**. Now (A) contains every positive integer, so $A(n) = n$ for all $n \geqslant 1$. Thus (A) must have Shnirelman density of one. □

It follows by contraposition from Theorem 2.2.i that if $\sigma(A) > 0$, then number one is a member of sequence (A).

2.2. **Lev Shnirelman.** Lev Genrikhovich Shnirelman was a Soviet mathematician born in Gomel, Belarus in 1905. The son of a school teacher, Shnirelman studied himself the complete school course of mathematics in just one year at the age of eleven. Only five years later he continued to impress with his academic skills by entering the University of Moscow, where he was taught by outstanding mathematicians such as Khinchin, Luzin and Urysohn. Shnirelman started research in algebra, geometry and topology, and soon graduated in 1925.

Portrait of Lev Shnirelman

Shnirelman was assigned to the chair of mathematics at the Don Polytechnic Institute in Novocherkassk in 1929. One year later he went back to the University of Moscow and continued from there to study at Göttingen in 1931. Shnirelman taught a few years at the university before being elected to the Soviet Academy of Sciences in 1933, after which he worked at the Mathematical Institute of the Academy. Tragically, shortly after being elected to the academy, Shnirelman died in Moscow, USSR in 1938. It has been speculated that he committed suicide or that he was assassinated by the Soviet intelligence service.

Shnirelman managed, however, to contribute on several fields of mathematics. Together with Lazar Lyusternik they, for example, developed the Lyusternik–Shnirelman category and proved Theorem of the three geodesics. Shnirelman's most famous research, however, was done in additive number theory; in 1930 he was able to prove a weak form of the Goldbach's conjecture (see Theorem 5.5) using ideas of compactness of a sequence of natural numbers. (O'Connor and E. F. Robertson 2015)

2.3. **Shnirelman's lemmas.** Let (A) and (B) be two sequences. We denote by $(A+B)$ the *sum of sequences* (A) and (B), the sequence consisting of all integers of the form a, b or $a + b$, where a is a member of (A) and b a member of (B), each counted only once, in order of magnitude. For example, if $(A) = (1, 2, 5, 8)$ and $(B) = (3, 7, 10, 11)$, then $(A + B) = (1, 2, 3, 4, 5, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 18, 19)$.

Using the notation above, Waring's problem can be stated in another way. Let (A_1), (A_2),..., (A_g) be g strictly monotonically increasing sequences of kth powers of integers, all of which begin with zero. Waring's problem now asserts simply that there is $g \geqslant 1$,

which depends only on k, such that the sequence $(A_1 + A_2 + \ldots + A_g)$ contains all natural numbers, that is, $(A_1 + A_2 + \ldots + A_g)$ coincides with $\mathbf{N}$.

It was Shnirelman who first discovered that there is a natural definition of the density of a sequence. He asked, 'to what extent is this density of the sum of several sequences determined solely by the density of the summands, irrespective of their arithmetical nature' (Khinchin 1998, p. 21). These ideas led, among other things, to Linnik's elementary solution of Waring's problem that we pursue here. The two following lemmas 2.3 and 2.4 are due to Shnirelman.

Lemma 2.3. *If* $\sigma(A) \geqslant 1/2$*, then sequence* $(A + A)$ *coincides with* $\mathbf{N}$.

Proof. Clearly all the members of $(A + A)$ belong to $\mathbf{N}$. Let us assume that $N \in \mathbf{N}$. Now either N is a member of (A) or not. If it is, then by definition N is a member of $(A + A)$ and we are done. Thus suppose that N is not a member of (A).

Let $B = \{a_1, \ldots, a_n\}$ be the set of all positive members of the sequence (A) not exceeding N. By assumption $N \notin B$. Since $\sigma(A) \geqslant 1/2$, we get $n = A(N) \geqslant N/2$. Define

$$C = \{N - a_i \mid a_i \text{ is a member of } (A) \text{ and } i \in [n]\},$$

where $[n]$ denotes the set $\{k \in \mathbf{N} \mid k \leqslant n\} = \{1, \ldots, n\}$. We have that, for the cardinalities of the sets B and C, $|B| \geqslant N/2$ and $|C| \geqslant N/2$. In addition, all the elements of the sets B and C belong to the closed interval from 1 to N. Thus, by the *Pigeon-hole principle*, the sets B and C intersect, which implies that there exists members a and b in (A) such that $a = N - b$. It follows that $N = a + b$ is a member of $(A + A)$, and hence $(A + A)$ contains all natural numbers. □

Shnirelman's inequality, the density of the sum of any two sequences is not smaller than the sum of their densities subtracted by the product of these densities, makes it possible to estimate the density of a sum from the densities of the summands. Let us state this formally.

Lemma 2.4 (Shnirelman's inequality)**.** *For sequences* (A) *and* (B)

$$\sigma(A + B) \geqslant \sigma(A) + \sigma(B) - \sigma(A)\,\sigma(B).$$

Proof. Suppose first that $\sigma(A) = 0$ or $\sigma(B) = 0$, say $\sigma(A) = 0$. Now clearly

$$\sigma(A) + \sigma(B) - \sigma(A)\,\sigma(B) = \sigma(B) \leqslant \sigma(A + B).$$

Suppose then that $\alpha = \sigma(A) > 0$ and $\beta = \sigma(B) > 0$. Especially, this means that $a_1 = 1$ (Theorem 2.2.i) and, for every $m \geqslant 1$, $A(m) \geqslant \alpha m$ and $B(m) \geqslant \beta m$. Let $N \in \mathbf{N}$ and n be the greatest index such that $a_n \leqslant N$. Now by the definition of the counting function $n = A(N)$.

In the following we translate the members of (B) by certain members of (A). The intention of this construction is to obtain a part of members of $(A + B)(N)$, with which we achieve the required estimate. Denote by i_k the greatest index for every $1 \leqslant k < n$, for which $a_k + b_{i_k} < a_{k+1}$. Moreover, let i_n be the greatest index satisfying $a_n + b_{i_n} \leqslant N$. Observe that all the numbers $a_k + b_j$, where $1 \leqslant k < n$ and $0 \leqslant j \leqslant i_k$, appear in $(A + B)$, are distinct and do not exceed N. Let us estimate.

$$\begin{aligned}(A + B)(N) &\geqslant n + i_1 + \ldots + i_n \\ &= A(N) + B(a_2 - a_1 - 1) + \ldots + B(a_n - a_{n-1} - 1) + B(N - a_n)\end{aligned}$$

$$\begin{aligned}
&\geqslant A(N) + \beta(a_2 - a_1 - 1) + \ldots + \beta(a_n - a_{n-1} - 1) + \beta(N - a_n) \\
&= A(N) + \beta\left(a_2 - a_1 - 1 + \ldots + a_n - a_{n-1} - 1 + N - a_n\right) \\
&= A(N) + \beta\left(a_2 - a_1 + \ldots + a_n - a_{n-1} + N - a_n - (n-1)\right),
\end{aligned}$$

where the sum $a_2 - a_1 + \ldots + a_n - a_{n-1} = -a_1 + a_n = -1 + a_n$ telescopes, and we get

$$\begin{aligned}
(A+B)(N) &\geqslant A(N) + \beta(N - n) \\
&= A(N) + \beta N - \beta A(N) \\
&= (1-\beta)A(N) + \beta N \\
&\geqslant (1-\beta)\alpha N + \beta N \\
&= (\alpha + \beta - \alpha\beta)N
\end{aligned}$$

or equivalently

$$\frac{(A+B)(N)}{N} \geqslant \alpha + \beta - \alpha\beta. \tag{4}$$

Now since $\sigma(A+B)$ is the infimum of all values of the fraction in (4), we get the claim:

$$\sigma(A+B) = \inf_{n \geqslant 1} \frac{(A+B)(N)}{N} \geqslant \alpha + \beta - \alpha\beta. \qquad \square$$

The inequality can be sharpened as demonstrated by Mann (1942). For the proof, see for example (Khinchin 1998, pp. 28–36; Pollack 2004, pp. 203–205). We do not need the result for this survey but note it for those readers, who are after a challenge.

Theorem 2.5 (Mann's theorem)**.** *For sequences* (A) *and* (B)

$$\sigma(A+B) \geqslant \min(1, \sigma(A) + \sigma(B)).$$

We already noted that lemmas 2.3 and 2.4 work as rudimentary tools for the estimation of the densities of certain type of sequences. Now we give two evident corollaries to them, which completes our exposition of Shnirelman density. From these corollaries especially the latter proves out to be highly useful. First, generalizing Shnirelman's inequality to the sum of any finite number of sequences of integers, is a simple inductive exercise.

Corollary 2.6. *For r sequences* $(A_1), \ldots, (A_r)$

$$1 - \sigma(A_1 + \ldots + A_r) \leqslant \prod_{i=1}^{r} (1 - \sigma(A_i)).$$

Proof. If $r = 1$, then there is nothing to prove. Thus let us induct on $r \geqslant 2$. First, note that equivalently to Shnirelman's inequality one could write

$$1 - \sigma(A+B) \leqslant (1 - \sigma(A))(1 - \sigma(B)),$$

where (A) and (B) are any two sequences. This gives the base case $r = 2$.

Suppose $r \geqslant 3$ and that the claim is true for $r-1$ sequences. Let $(A) = (A_1 + \ldots + A_{r-1})$ and $(B) = (A_r)$. Then by Shnirelman's inequality and induction hypothesis

$$\begin{aligned}
1 - \sigma(A_1 + \ldots + A_r) &= 1 - \sigma(A+B) \\
&\leqslant (1 - \sigma(A))(1 - \sigma(B))
\end{aligned}$$

$$\leqslant (1-\sigma(A_r))\prod_{i=1}^{r-1}(1-\sigma(A_i))$$

$$=\prod_{i=1}^{r}(1-\sigma(A_i)).$$

By *mathematical induction* the statement holds for all $r \geqslant 2$. □

Corollary 2.7. *If* $\sigma(A) > 0$*, then there exists* $r \in \mathbf{N}$ *such that sequence* (rA) *coincides with* $\mathbf{N}$.

Proof. Suppose that $\sigma(A) > 0$. If $\sigma(A) = 1$, we are done (Theorem 2.2.ii). If not, $0 < \sigma(A) < 1$ and further $0 < 1 - \sigma(A) < 1$. We may choose a positive integer s large enough so that $(1-\sigma(A))^s < 1/2$. By Corollary 2.6

$$1-\sigma(sA) \leqslant \prod_{i=1}^{s}(1-\sigma(A)) = (1-\sigma(A))^s < \frac{1}{2}$$

or equivalently

$$\sigma(sA) > \frac{1}{2}.$$

Applying Lemma 2.3 we get that $(sA+sA)$ coincides with $\mathbf{N}$. Choosing $r = 2s$ finishes the proof. □

3. Reduction to Diophantine equations

One of the most persisting topics in mathematics is — without a doubt — the theory of solving equations. At the same time, it can be interpreted as an important driving force of the science. This section focuses on Diophantine equations, polynomial equations over the integers, the study of which saw a new surge in early modern age thanks to Waring's conjecture.

Let $n \geqslant 2$ be a fixed natural number, $r \in \mathbf{N}$, (A) be the sequence of non-negative integers raised to the power n, that is $(A) = (0, 1, 2^n, 3^n, 4^n, \ldots)$, and (B) be the sequence

$$(rA) = (\underbrace{A + A + \ldots + A}_{r}).$$

Observe that for all $k \geqslant 1$

$$\frac{1}{k} \leqslant \frac{A(k)}{k} \leqslant \frac{k^{\frac{1}{n}}}{k}.$$

Hence it must be that

$$0 \leqslant \sigma(A) \leqslant \lim_{k\to\infty}\frac{k^{\frac{1}{n}}}{k} = 0,$$

that is $\sigma(A) = 0$.

Our goal is to show that, for a positive integer r, the sequence (B) has positive Shnirelman density. This is the crux of Linnik's elementary proof. In fact, this would imply that Waring's problem is true by Corollary 2.7; if we can show that $\sigma(B) > 0$ for some r, then we can find an integer s such that (sB) coincides with $\mathbf{N}$. This would imply that $g(n) \leqslant rs$.

To achieve our objective, first, we will translate the problem of showing $\sigma(B) > 0$ to estimating the number of solutions of Diophantine equations by using the so-called Fundamental lemma. Verifying the Fundamental lemma, in turn, is easier with the help of a few intermediate results. At the end of this section we present two of these results, lemmas 3.4 and 3.6, after introducing Hua, the mathematician behind them.

3.1. **Number of solutions of Diophantine equations.** Let N be any non-negative integer. Now, by our definition, $B(N)$ is the number of integers in $0 \leqslant m \leqslant N$, for which the equation with respect to the variables x_i

$$x_1^n + \ldots + x_r^n = m, \quad \text{where} \quad x_i \geqslant 0 \quad \text{for every} \quad i \in [r], \tag{5}$$

is solvable in integers. Denote the number of solutions of (5) by $R(m)$. Now specifically, m is a member of $(B) = (rA)$, if and only if $R(m) > 0$. The *Cauchy–Bunyakovsky inequality* gives

$$\left(\sum_{m=0}^{N} R(m)\right)^2 \leqslant \sum_{\substack{0\leqslant m\leqslant N\\ R(m)\neq 0}} 1 \cdot \sum_{m=0}^{N} R(m)^2. \tag{6}$$

Note that the sum $\sum_{m=0}^{N} R(m)$ equals to the number of solutions of the inequality

$$0 \leqslant x_1^n + \ldots + x_r^n \leqslant N. \tag{7}$$

Every group of non-negative integers $x_1, \ldots, x_r$, for which $x_i \leqslant (N/r)^{1/n}$ for all $i \in [r]$, clearly satisfies (7). There is r variables, each of which can be chosen in $1 + \left\lfloor (N/r)^{1/n} \right\rfloor$ different ways. Thus we can estimate

$$\sum_{m=0}^{N} R(m) \geqslant \left(1 + \left\lfloor \left(\frac{N}{r}\right)^{\frac{1}{n}} \right\rfloor\right)^r \geqslant \left(\frac{N}{r}\right)^{\frac{r}{n}}.$$

Plugging the above into (6) gives

$$\left(\frac{N}{r}\right)^{\frac{2r}{n}} \leqslant B(N) \sum_{m=0}^{N} R(m)^2. \tag{8}$$

As for the sum $\sum_{m=0}^{N} R(m)^2$, it equals to the number of solutions of the system

$$\begin{cases} x_1^n + \ldots + x_r^n = y_1^n + \ldots + y_r^n \\ x_1^n + \ldots + x_r^n \leqslant N, \end{cases}$$

where x_i and y_i are both non-negative integers. In addition, the sum is less than the number of solutions of the system

$$\begin{cases} x_1^n + \ldots + x_r^n = y_1^n + \ldots + y_r^n \\ 0 \leqslant x_i \leqslant N^{\frac{1}{n}} \\ 0 \leqslant y_i \leqslant N^{\frac{1}{n}}. \end{cases} \tag{9}$$

With all these preliminaries we are ready to state the fundamental lemma, the claim in which most of the work in Linnik's proof is hidden.

Theorem 3.1 (Fundamental lemma)**.** *There exists a positive integer $r = r(n)$ such that, for every positive integer N, the number of solutions of (9) does not exceed*

$$cN^{\frac{2r}{n}-1},$$

where c is a positive constant depending only on r and n.

The proof of the Fundamental lemma is considered to be very long and complicated by many authors. We will postpone the proof until we have gathered a few intermediate results to help with the task. In a sense, the proof of the Hilbert–Waring theorem is divided into sections of a chain; the result of the Fundamental lemma makes solving Waring's problem quite trivial with the help of the results from previous section.

Proof of the Hilbert–Waring theorem. The estimate in (8) and the Fundamental lemma give

$$\left(\frac{N}{r}\right)^{\frac{2r}{n}} \leqslant B(N)cN^{\frac{2r}{n}-1}$$

or equivalently

$$\frac{B(N)}{N} \geqslant \frac{\left(\frac{N}{r}\right)^{\frac{2r}{n}}}{cN^{\frac{2r}{n}}} = \frac{1}{c}\left(\frac{1}{r}\right)^{\frac{2r}{n}},$$

which is clearly positive. Thus $\sigma(B)$ is positive and by Corollary 2.7 we can find an integer s such that (sB) includes all natural numbers, that is $g(n) \leqslant rs$ for all n. □

Hence to solve Waring's problem it is sufficient to prove the Fundamental lemma. Furthermore, the latter can be deduced from the following result.

Theorem 3.2. *Let $f(x) = a_0x^n + \ldots + a_{n-1}x + a_n$ be an integer polynomial that satisfies*

$$n \geqslant 2, \quad 0 < |a_0| < \lambda \quad \text{and} \quad |a_1| \leqslant \lambda P, \ldots, |a_{n-1}| \leqslant \lambda P^{n-1} \tag{10}$$

for some $\lambda \geqslant 1$ and $P \geqslant 1$. Then there is $\lambda_0(n) < \infty$ such that, if $\lambda > \lambda_0(n)$, then the number of integer solutions of the equation

$$\sum_{j=1}^{s}(-1)^j f(x_j) = 0, \quad \text{where} \quad s = 8^{n-1}, \quad x_j \in \mathbf{Z} \quad \text{and} \quad 0 \leqslant x_j \leqslant P$$

is no greater than $\lambda^{s-3}P^{s-n}$.

We will next verify the Fundamental lemma using Theorem 3.2, the latter of which, in a way, translates the problem to a more general context of polynomials. Indeed Waring's problem can be generalized to polynomials (see Section 5.2), but we settle for using them for a mere notational convenience.

The rest of this thesis is dedicated to the proof of Theorem 3.2, which is achieved, in the end, by induction on the degree of the polynomial $f(x)$ in Section 4.3. The following intermediate results are tailored to suit this inductive argument.

Derivation of the Fundamental lemma from Theorem 3.2 is fairly straightforward. Note that the result itself gives actually far more than we need.

Proof of the Fundamental lemma. Re-indexing the variables in (9), let x_is represent the even indexes of z_j and y_is the odd indexes of z_j, that is $x_1 = z_2$, $x_2 = z_4$, …, and

$y_1 = z_1, y_2 = z_3, \ldots$. This way (9) can be written in the form

$$\sum_{j=1}^{2r} (-1)^j z_j^n = 0, \quad \text{where} \quad 0 \leqslant z_j \leqslant N^{\frac{1}{n}}. \tag{11}$$

Choose $f(z) = z^n$ in Theorem 3.2. In addition, set $2r = 8^{n-1} = s$ and $P = N^{1/n}$. Now by Theorem 3.2 the number of solutions of (11) does not exceed $\lambda^{s-3} P^{s-n} = \lambda^{s-3} N^{2r/n-1}$, where λ^{s-3} depends only on r and n. □

Remark 3.3. The value of s in Theorem 3.2 obtained originally by Linnik was 2^{4^n}, whereas $s = 8^{n-1}$ is due to Hua (Hua 1982, p. 529).

3.2. **Loo-Keng Hua.** Loo-Keng Hua was a Chinese mathematician born in Jintan, China in 1910. Son of a local shopkeeper, Hua attended elementary and middle school in the neighbourhood. He was a frail child, who almost became disabled due to complications of a disease. However, his joyful and positive personality carried him over the illnesses and many other trials to come. Early on Hua spent his free time engaging in non-trivial mathematical problems directly from first principles — without books or scientific literature. His self-study spanned easily the entire high school and early undergraduate mathematics curricula.

Portrait of Loo-Keng Hua

Hua's tertiary education was interrupted by financial problems. Hua gained admission to the Chinese Vocational College in Shanghai, where he won a national abacus competition. Due to high living costs and failing to find a job in Shanghai, Hua dropped out and went back home in 1927. However, he was already committed to mathematics and managed to get his first publication out in 1929. Next year he published corrections to an earlier article, which caught the eye of professor Qinglai at Quing Hua University in Beijing. One more year and Hua was invited to the university's mathematics department.

In just eight years of time from a clerk in the library to a lecturer and finally a full professor — despite not having any degree —, Hua kept publishing around Waring's problem, Diophantine analysis and function theory. In the following years he contributed to the university's objective to raise China's mathematics and science alongside of knowledge in the West. Visitations and invitations led Hua to spend two years 1936–1938 in Cambridge, England, where he met, for example, Hardy. In England Hua was encouraged to pursue Doctor of Philosophy degree but he refused because of high registration fees.

Hua's research interests extended to analysis and algebra, never forgetting number theory in general. Most notable were his seminal work on the estimation of trigonometric sums, singly or on average, initiated by his Cambridge years; he refined and reformulated the Hardy–Littlewood–Vinogradov method of estimating trigonometric sums — better known as Vinogradov's mean value theorem. Around in the middle of the century Hua's life and work were significantly impeded by Japanese invasion of China, the Chinese Civil War, World War II, Cultural Revolution and Great Leap Forward. Despite the difficulties, Hua's work in applied mathematics and his oratorical

talent made him seen as a hero by the public; mathematical optimization by Hua had an enormous effect on the economy of China.

Hua's first degree was an honorary doctorate from the University of Nancy given in 1980 — one of several other honoraries. Hua died of a heart attack at the end of a lecture he gave in Tokyo, Japan five years later. Nowadays there is, for example, a high school named after him and a memorial building in Jintan celebrating his achievements. Hua has been compared to Shiing-shen Chern and, as a scholar and a teacher, he is credited for popularity and innovativeness of present-day Chinese mathematicians. (Halberstam 2002)

3.3. **Lemmas due to Hua.** In order to prove Theorem 3.2 we have a considerably non-trivial road ahead. Thanks to Hua, the number of needed intermediate results is fairly small. Nevertheless, we encourage the reader to focus on each of the results individually. Attention to the problem posed will be called upon separately in the final section of the proof.

Lemma 3.4. *Let X and Y be real numbers such that $1 \leqslant X \leqslant Y$, and let a be an integer. The number of solutions of the Diophantine equation*

$$x_1 y_1 + x_2 y_2 = a, \tag{12}$$

where $0 < |x_i| \leqslant X$ and $|y_i| \leqslant Y$ for all $i \in \{1, 2\}$, is no greater than

$$\begin{cases} 12X^2Y, & \text{if } a = 0 \\ 40XY \displaystyle\sum_{d \mid a} \frac{1}{d}, & \text{if } a \neq 0. \end{cases}$$

Note that in the sum above the number d ranges over all positive divisors of a.

Proof. Suppose that $a = 0$. Consider the triples (x_1, x_2, y_2), where $x_i \neq 0$ for $i \in \{1, 2\}$. Since each x_i can be chosen in $2\lfloor X \rfloor$ and y_2 in $2\lfloor Y \rfloor + 1$ different ways, there are in total

$$2\lfloor X \rfloor \cdot 2\lfloor X \rfloor \cdot (2\lfloor Y \rfloor + 1) = 8\lfloor X \rfloor^2 \lfloor Y \rfloor + 4\lfloor X \rfloor^2 \leqslant 12X^2Y$$

variants of the triples. Now since $x_1 \neq 0$, at most one value of y_1 satisfies (12) for each triple, and the required inequality follows.

Suppose then that $a \neq 0$. First, let us consider the case $|x_2| \leqslant |x_1|$. Decompose the set of all solutions (x_1, x_2, y_1, y_2) of (12) into a disjoint union of subsets with the same value of $d = \gcd(x_1, x_2)$, that is the greatest common divisor of x_1 and x_2.

Consider the case $d = 1$ and fix co-prime non-negative integers x_1 and x_2. Now, using a major result in linear Diophantine equations (LeVeque 1996, Theorem 2.9), since $1 \mid a$, (12) has integer solutions in y_1, y_2, and, for every two solutions (y_1', y_2') and (y_1'', y_2''), we have

$$\begin{cases} y_1'' = y_1' + tx_2 \\ y_2'' = y_2' - tx_1, \end{cases}$$

where $t \in \mathbf{Z}$, and

$$|t| = \frac{|y_2'' - y_2'|}{|x_1|} \leqslant \frac{2Y}{|x_1|}.$$

Since $|x_1| \leqslant X \leqslant Y$, the number of possible t does not exceed

$$2\left\lfloor \frac{2Y}{|x_1|} \right\rfloor + 1 \leqslant \frac{4Y + X}{|x_1|} \leqslant \frac{5Y}{|x_1|}.$$

Thus the number of solutions for which $|x_2| \leqslant |x_1|$ and $d = 1$ is no greater than

$$\sum_{\substack{1 \leqslant |x_1| \leqslant X \\ 1 \leqslant |x_2| \leqslant |x_1|}} \frac{5Y}{|x_1|} \leqslant 5Y \sum_{1 \leqslant |x_1| \leqslant X} \frac{2\,|x_1|}{|x_1|} \leqslant 20XY. \tag{13}$$

Now, for the more general case without the restriction $|x_2| \leqslant |x_1|$, the number of solutions is clearly dominated by the number $2 \cdot 20XY = 40XY$. Furthermore, if $d \neq 1$ but $d \mid a$, then we let $x_1' = x_1/d$ and $x_2' = x_2/d$, seeking the number of solutions of the equation

$$x_1' y_1 + x_2' y_2 = \frac{a}{d},$$

where $0 < \left|x_i'\right| \leqslant X/d$ and $|y_i| \leqslant Y$ for all $i \in \{1, 2\}$. Now again $\gcd(x_1', x_2') = 1$, and we see from the above — division by d coming into play in (13) — that the number of solutions does not exceed $(40/d)XY$.

Summing these estimates over all d, we get that the number of solutions of (12) does not exceed

$$40XY \sum_{d \mid a} \frac{1}{d}. \qquad \square$$

For clarity, we present two, more calculational intermediate results needed in the proof of Lemma 3.6. They are of some interest in their own right.

Lemma 3.5. *For $T \geqslant 1$*

i) $\displaystyle\sum_{1 \leqslant d \leqslant T} d^{-\frac{3}{2}} \leqslant 2 + \sqrt{2}$ *and*

ii) $\displaystyle\sum_{1 \leqslant k \leqslant T} \left(\sum_{d \mid k} \frac{1}{d} \right)^2 \leqslant \left(2 + \sqrt{2}\right)^2 T.$

Proof. **i)** The estimation of the sum can be done elementarily — following the ideas of Oresme's proof that the harmonic series diverges and the Cauchy condensation test. Clearly

$$\sum_{1 \leqslant d \leqslant T} d^{-\frac{3}{2}} \leqslant \sum_{d=1}^{\infty} d^{-\frac{3}{2}}.$$

Write

$$\sum_{d=1}^{\infty} d^{-\frac{3}{2}} = 1^{-\frac{3}{2}} + 2^{-\frac{3}{2}} + 3^{-\frac{3}{2}} + 4^{-\frac{3}{2}} + \ldots + \left(2^{k-1}\right)^{-\frac{3}{2}} + \ldots + \left(2^k\right)^{-\frac{3}{2}} + \ldots$$

There is in total $2^k - 2^{k-1} = 2^{k-1}$ summands in $\left(2^{k-1}\right)^{-\frac{3}{2}} + \ldots + \left(2^k - 1\right)^{-\frac{3}{2}}$, from which the term $\left(2^{k-1}\right)^{-\frac{3}{2}}$ is the greatest. Thus it must be that

$$\begin{aligned}\sum_{d=1}^{\infty} d^{-\frac{3}{2}} &\leqslant 1^{-\frac{3}{2}} + 2^{-\frac{3}{2}} + 2^{-\frac{3}{2}} + 4^{-\frac{3}{2}} + 4^{-\frac{3}{2}} + 4^{-\frac{3}{2}} + 4^{-\frac{3}{2}} + 8^{-\frac{3}{2}} + \ldots \\ &= \sum_{k=1}^{\infty} 2^{k-1} \left(2^{k-1}\right)^{-\frac{3}{2}} \\ &= \sum_{k=0}^{\infty} \left(\frac{1}{\sqrt{2}}\right)^k \\ &= 2 + \sqrt{2}.\end{aligned}$$

ii) For every $T \geqslant 1$

$$\begin{aligned}\sum_{1 \leqslant k \leqslant T} \left(\sum_{d|k} \frac{1}{d}\right)^2 &= \sum_{1 \leqslant k \leqslant T} \sum_{\substack{d_1|k \\ d_2|k}} \frac{1}{d_1 d_2} \\ &= \sum_{\substack{1 \leqslant d_1 \leqslant T \\ 1 \leqslant d_2 \leqslant T}} \frac{1}{d_1 d_2} \sum_{\substack{1 \leqslant k \leqslant T \\ d_1|k \\ d_2|k}} 1,\end{aligned}$$

where the latter sum counts the number of $k \in [T]$, which are divisible by both d_1 and d_2. Thus

$$\begin{aligned}\sum_{1 \leqslant k \leqslant T} \left(\sum_{d|k} \frac{1}{d}\right)^2 &= \sum_{\substack{1 \leqslant d_1 \leqslant T \\ 1 \leqslant d_2 \leqslant T}} \frac{1}{d_1 d_2} \left\lfloor \frac{T}{\operatorname{lcm}(d_1, d_2)} \right\rfloor \\ &\leqslant T \sum_{\substack{1 \leqslant d_1 \leqslant T \\ 1 \leqslant d_2 \leqslant T}} \frac{\gcd(d_1, d_2)}{(d_1 d_2)^2}.\end{aligned}$$

Since $\gcd(d_1, d_2) \leqslant \min(d_1, d_2) \leqslant \sqrt{d_1 d_2}$, we can continue

$$\begin{aligned}\sum_{1 \leqslant k \leqslant T} \left(\sum_{d|k} \frac{1}{d}\right)^2 &\leqslant T \sum_{\substack{1 \leqslant d_1 \leqslant T \\ 1 \leqslant d_2 \leqslant T}} \frac{1}{(d_1 d_2)^{\frac{3}{2}}} \\ &= T \left(\sum_{1 \leqslant d \leqslant T} \frac{1}{d^{\frac{3}{2}}}\right)^2.\end{aligned}$$

Using the previous result to this estimate proves the lemma. □

The estimation done above is very crude but sufficient for the application needed; for future arguments, the exact value of constant c in the following lemma is of no

relevance. However, the estimate can be sharpened to $(5/2)\zeta(3)T$ (Jameson 2015, pp. 3–4), where ζ denotes the Riemann zeta function. According to Jameson, actually

$$\sum_{1\leqslant k\leqslant T}\left(\sum_{d|k}\frac{1}{d}\right)^2 = \frac{5}{2}\zeta(3)T + O\left(\ln^2 T\right),$$

where $O(\cdot)$ is the *big O notation*.

Lemma 3.6. *Let X and Y be real numbers such that $1 \leqslant X \leqslant Y$. The number of solutions of the Diophantine equation*

$$x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4 = 0, \tag{14}$$

where $0 < |x_i| \leqslant X$ and $|y_i| \leqslant Y$ for all $i \in [4]$, does not exceed $c(XY)^3$, where $c = 8\cdot 10^4$.

Proof. Denote by $q(a)$ the number of solutions of (12) and by M the number of solutions of (14). Now by Lemma 3.4

$$\begin{aligned} M &\leqslant \sum_{|k|\leqslant 2XY} q(k)q(-k) \\ &\leqslant \left(12X^2Y\right)^2 + \sum_{1\leqslant|k|\leqslant 2XY} q(k)q(-k) \\ &\leqslant 144X^4Y^2 + 2\cdot 40^2(XY)^2 \sum_{1\leqslant k\leqslant 2XY}\left(\sum_{d|k}\frac{1}{d}\right)^2. \end{aligned}$$

Applying Lemma 3.5.ii for $T = 2XY$ gives

$$M \leqslant 144X^4Y^2 + 2\cdot 40^2(XY)^2\cdot\left(2+\sqrt{2}\right)^2\cdot 2XY \leqslant 8\cdot 10^4(XY)^3. \qquad \square$$

4. Combinatorics and the final induction

We are almost ready to prove Theorem 3.2, with which it is possible to deduce the Fundamental lemma and further solve Waring's problem, as was seen in Section 3. Linnik's line of reasoning requires two more lemmas, which we present right after introducing the man himself.

4.1. **Yuri Linnik.**

Portrait of Yuri Linnik

Yuri Vladimirovich Linnik was a talented Soviet mathematician born in Belaya Tserkov, Ukraine, in 1915. Linnik's parents were school teachers and later his father became a highly successful academician — like his son after him. Linnik's education begun in a secondary school in Leningrad, after which he started at Leningrad University in 1932. At first Linnik majored in physics but he transferred to mathematics at Leningrad State University in 1935 feeling 'an irresistible bent for higher arithmetic'.

Graduating in 1938, Linnik's doctor's degree was based on research in quadratic forms and advised by Vladimir Tartakovski. The studies were impeded, Linnik fighting twice in World War II. Presumably, he was spared from Leningrad Blockade being evacuated due to sickness. The quality

of his work, however, was not affected and after submitting his thesis, Linnik was awarded the higher degree of Doctor of Science in Mathematics and Physics. He enlisted the Leningrad branch of the Steklov Institute for Mathematics, where he worked until his death. Linnik wrote over 240 research papers and published several books during his career as a professional mathematician. Despite his outstanding talents, Linnik has been described as a person, who believes that anyone with a similar background could make comparable breakthroughs.

After the war Linnik was appointed as professor of mathematics at Leningrad State University. Probability theory, mathematical statistics and number theory were his main research topics — lending ideas fluently from one area to another. His main contributions were Linnik's theorem in analytic number theory, use of dispersion and ergodic method in number theory, and introduction of 'large sieve'. In addition, in probability theory and statistics he proved a generalization of Cramér's theorem and solved the Behrens–Fisher problem. Linnik received many honours, begin elected to multitude of high-level scientific organisations, awarded multiple prizes and admitted an honorary doctorate. (O'Connor and E. F. Robertson 2015)

4.2. **Combinatorial lemmas.** The two following lemmas are abstract combinatorial results, whose idea and form are quite simple. However, the difficulty lies in the notation. The latter one will be applied many times in the proof of Theorem 3.2 and thus is of high importance.

Lemma 4.1. *Let $F(\boldsymbol{x})$ be a polynomial in d variables $x_1, \dots, x_d$ with coefficients in $\mathbf{Z}$, and let A be a finite subset of $\mathbf{Z}^d$. Now, for every $k \in \mathbf{Z}$, the number of solutions of the equation*

$$F(\boldsymbol{x}) - F(\boldsymbol{y}) = k, \quad \text{where} \quad \boldsymbol{x}, \boldsymbol{y} \in A, \tag{15}$$

does not exceed the number of solutions of the equation

$$F(\boldsymbol{x}) - F(\boldsymbol{y}) = 0, \quad \text{where} \quad \boldsymbol{x}, \boldsymbol{y} \in A. \tag{16}$$

Proof. Let $\lambda(a)$ be the number of solutions of the equation

$$F(\boldsymbol{x}) = a, \quad \text{where} \quad \boldsymbol{x} \in A,$$

N the number of solutions of (15), and M the number of solutions of (16). Then summing the product $\lambda(a)\lambda(a-k)$ over $a \in \mathbf{Z}$ gives N. Similarly, $M = \sum_{a\in\mathbf{Z}} \lambda(a)^2$. Thus

$$N = \sum_{a\in\mathbf{Z}} \lambda(a)\lambda(a-k) \leqslant \frac{1}{2} \sum_{a\in\mathbf{Z}} \left(\lambda(a)^2 + \lambda(a-k)^2\right) = \sum_{a\in\mathbf{Z}} \lambda(a)^2 = M,$$

where the inequality follows from simple $2nm \leqslant n^2 + m^2$, which is equivalent to $(n-m)^2 \geqslant 0$, where $n, m \in \mathbf{N}$. □

Lemma 4.2. *Let $F(\boldsymbol{x})$ be a polynomial in d variables $x_1, \dots, x_d$ with coefficients in $\mathbf{Z}$, $k \in \mathbf{N}$, and $A_1, \dots, A_{2k}$ be finite subsets of $\mathbf{Z}^d$. Denote by $N(A_1, \dots, A_{2k})$ the number of solutions of the equation*

$$\sum_{j=1}^{2k} (-1)^j F(\boldsymbol{x}_j) = 0, \quad \text{where} \quad \boldsymbol{x}_i \in A_i \quad \text{for all} \quad i \in [2k].$$

Now there exists an index $r \in [2k]$ such that $N(A_1, \dots, A_{2k}) \leqslant N(A_r, \dots, A_r)$.

Proof. Note that the set of different collections $(A_{j_1}, \dots, A_{j_{2k}})$, $0 \leqslant j_1, \dots, j_{2k} \leqslant 2k$, is finite. Let $(B_1, \dots, B_{2k})$ be such a collection, satisfying both

i) $N(B_1, \dots, B_{2k})$ equals the maximum of $N(A_{j_1}, \dots, A_{j_{2k}})$, and
ii) $(B_1, \dots, B_{2k})$ contains the least number of different sets.

Denote by q the number of different sets in $(B_1, \dots, B_{2k})$. Now it is sufficient to show that $q = 1$.

Assume to the contrary that $q \geqslant 2$. We may divide $(B_1, \dots, B_{2k})$ into two subcollections so that each collection consists of k sets, $B_{j_1}, \dots, B_{j_k}$ and $B_{j_{k+1}}, \dots, B_{j_{2k}}$, and the first collection contains no more than $q-1$ different sets.

Denote, for every $T \in \mathbf{Z}$, by $\lambda(T)$ and $\mu(T)$, respectively, the numbers of solutions of the equations

$$\sum_{\ell=1}^{k} (-1)^{j_\ell} F(\boldsymbol{x}_{j_\ell}) = T \qquad \text{and} \qquad \sum_{\ell=k+1}^{2k} (-1)^{j_\ell} F(\boldsymbol{x}_{j_\ell}) = -T,$$

where $\boldsymbol{x}_{j_\ell} \in B_{j_\ell}$ for all $\ell \in [2k]$. Using this notation, we can estimate — in a manner similar to that in the proof of Lemma 4.1 — that

$$N(B_1, \dots, B_{2k}) = \sum_{T \in \mathbf{Z}} \lambda(T)\mu(T) \leqslant \frac{1}{2} \sum_{T \in \mathbf{Z}} \left(\lambda(T)^2 + \mu(T)^2 \right). \tag{17}$$

Note that the sum $\sum_T \lambda(T)^2$ equals the number of solutions of the equation

$$\sum_{\ell=1}^{k} (-1)^{j_\ell} F(\boldsymbol{x}_{j_\ell}) - \sum_{\ell=1}^{k} (-1)^{j_\ell} F(\boldsymbol{y}_{j_\ell}) = 0, \tag{18}$$

where $\boldsymbol{x}_{j_\ell}, \boldsymbol{y}_{j_\ell} \in B_{j_\ell}$ for all $\ell \in [k]$. In the set $\{j_1, \dots, j_k, j_1 + 1, \dots, j_k + 1\}$ there are exactly k even and k odd numbers. Hence

$$\sum_{T \in \mathbf{Z}} \lambda(T)^2 = N(B_{i_1}, \dots, B_{i_{2k}}),$$

where all the sets B_{i_ℓ} belong to the first collection $(B_{j_1}, \dots, B_{j_k})$ defined earlier.

Correspondingly, the sum $\sum_{T \in \mathbf{Z}} \mu(T)^2$ can be interpreted as the number of solutions of an equation similar to (18). Thus, by the definition of the collection $(B_1, \dots, B_{2k})$, we get the inequalities

$$\sum_{T \in \mathbf{Z}} \lambda(T)^2 < N(B_1, \dots, B_{2k}) \qquad \text{and} \qquad \sum_{T \in \mathbf{Z}} \mu(T)^2 \leqslant N(B_1, \dots, B_{2k}),$$

which contradicts (17). □

4.3. **Proof of Theorem 3.2.** We are ready to prove Theorem 3.2. Recall, if needed, the assumptions (10) of the theorem. The statement is that the number of integer solutions of the equation

$$\sum_{j=1}^{s} (-1)^j f(x_j) = 0, \quad \text{where} \quad s = 8^{n-1} \quad \text{and} \quad 0 \leqslant x_j \leqslant P, \tag{19}$$

is no greater than $\lambda^{s-3} P^{s-n}$.

Let $n \geqslant 2$ be arbitrary and $f(x)$ be a corresponding polynomial of degree n. Set $t = 8^{n-2} = s/8$ and note that, with a suitable re-indexing of the terms, (19) is equivalent to

$$\sum_{i=1}^{4}(-1)^i \sum_{j=1}^{t}(-1)^j \left(f(x_{i,t+j}) - f(x_{i,j})\right) = 0, \quad \text{where} \quad 0 \leqslant x_{i,k} \leqslant P, \tag{20}$$

for $i \in [4]$ and $k \in [2t]$, the number of positive and negative terms being equal.

Lemma 4.2 allows us to estimate the number of solutions of equations like the above. Set $k = 2t$, $d = 2$, $F(x, y) = f(x) - f(y)$,

$$A_0 = \left\{(x, y) \in \mathbf{Z}^2 \mid 0 \leqslant x \leqslant P, 0 \leqslant y \leqslant P, \text{ and } x = y\right\}, \quad \text{and}$$
$$A_1 = \left\{(x, y) \in \mathbf{Z}^2 \mid 0 \leqslant x \leqslant P, 0 \leqslant y \leqslant P, \text{ and } x \neq y\right\}.$$

Consider the vectors

$$\left(\ell_{1,1}, \dots, \ell_{t,4}\right), \quad \text{where} \quad \ell_{j,i} \in \{0, 1\} \quad \text{for} \quad i \in [4] \quad \text{and} \quad j \in [t].$$

There are in total 2^{4t} variants of them. Pair each vector with a class of solutions of (20), consisting of all solutions for which

$$\left(x_{i,j}, x_{i,t+j}\right) \in A_{\ell_{j,i}}, \quad \text{where} \quad i \in [4] \quad \text{and} \quad j \in [t].$$

As a result the set of all solutions of (20) is decomposed into a union of 2^{4t} classes.

Now, Lemma 4.2 states that, for every collection $(i_1, \dots, i_{4t})$, the inequality,

$$N(A_{i_1}, \dots, A_{i_{4t}}) \leqslant N(A_r, \dots, A_r)$$

holds for $r = 0$ or $r = 1$. It follows that the required number of solutions is no greater than $2^{4t} N(A_r, \dots, A_r)$, where $r = 0$ or $r = 1$.

First, let us consider the case $N(A_1, \dots, A_1) \leqslant N(A_0, \dots, A_0)$. We have

$$N(A_0, \dots, A_0) = (\lfloor P \rfloor + 1)^{4t},$$

and thus the number of solutions of (19) is no greater than

$$2^{4t} (\lfloor P \rfloor + 1)^{4t} \leqslant 2^{4t} (P + P)^{4t} \leqslant 4^{4t} P^{4t}. \tag{21}$$

Estimating the exponents, since the inequality $4 \cdot 8^{n-2} \leqslant 8 \cdot 8^{n-2} - 3$ is equivalent to $48 \leqslant 8^n$, which is true for $n \geqslant 2$, we may write $4^{4t} \leqslant \lambda^{s-3}$ for $\lambda \geqslant 4$. Similarly $4 \cdot 8^{n-2} \leqslant 8 \cdot 8^{n-2} - n$ is equivalent to $16n \leqslant 8^n$, which is also true, and we get $P^{4t} \leqslant P^{s-n}$. Thus the estimate (21) can be further increased to $\lambda^{s-3} P^{s-n}$. All in all, the case $r = 0$ is in line with Theorem 3.2.

Hence we may assume that $N(A_0, \dots, A_0) \leqslant N(A_1, \dots, A_1)$. We need to deduce that the number $N(A_1, \dots, A_1)$ also has a good upper bound. Proceed with induction on n. The base case $n = 2$ is simple enough. Now $s = 8^1 = 8$ and $t = 1$. To estimate $N(A_1, \dots, A_1)$ we re-formulate (20) to

$$\sum_{i=1}^{4}(-1)^i \left(f(y_i) - f(x_i)\right) = 0, \quad \text{where} \quad x_i \neq y_i \quad \text{and} \quad 0 \leqslant x_i, y_i \leqslant P. \tag{22}$$

Furthermore, since

$$f(y) - f(x) = a_0 y^2 + a_1 y + a_2 - \left(a_0 x^2 + a_1 x + a_2\right) = (y - x)(a_0 y + a_0 x + a_1),$$

(22) can be given in the form

$$z_1h_1 - z_2h_2 + z_3h_3 - z_4h_4 = 0, \tag{23}$$

where $h_i = y_i - x_i$ and $z_i = a_0(x_i + y_i) + a_1$ for all $i \in [4]$. We get that, for all i,

$$0 < |h_i| \leqslant P \quad \text{and} \quad |z_i| \leqslant 3\lambda P. \tag{24}$$

By assumption $a_0 \neq 0$, so, for any collection $(z_1, h_1, \dots, z_4, h_4)$ satisfying both (23) and (24), there exists at most one solution $(x_1, \dots, x_4, y_1, \dots, y_4)$ of (22). Thus, by Lemma 3.6,

$$N(A_1, \dots, A_1) \leqslant c(P \cdot 3\lambda P)^3 \leqslant \lambda^4 P^6, \quad \text{when} \quad \lambda > 27c,$$

which implies that the number of solutions of (19) is bounded from above by $\lambda^5 P^6$, when $\lambda > 2^{4\cdot 8^{n-2}}$, in accordance with the theorem. This proves the base case $n = 2$.

Moving on to the inductive step, assume that $n \geqslant 3$ and that the statement holds for $n - 1$. Again, in order to estimate $N(A_1, \dots, A_1)$, re-formulate (20):

$$\sum_{j=1}^{t}(-1)^j \sum_{i=1}^{4}(-1)^i \left(f(x_{i,t+j}) - f(x_{i,j})\right) = 0, \quad x_{i,t+j} \neq x_{i,j}, \quad 0 \leqslant x_{i,j} \leqslant P. \tag{25}$$

Let us apply Lemma 4.2 for $d = 8$, $k = t/2$ and

$$F(x_1, \dots, x_4, y_1, \dots, y_4) = \sum_{i=1}^{4}(-1)^i \left(f(y_i) - f(x_i)\right).$$

We have to decompose the set of solutions of (25) into a union of classes. To do this, define a set

$$M(\boldsymbol{u}) = \left\{(x_1, \dots, y_4) \in \mathbf{Z}^8 \mid 0 \leqslant x_i, y_i \leqslant P \text{ and } y_i - x_i = u_i \text{ for all } i \in [4]\right\}$$

for every quadruple of integers $\boldsymbol{u} = (u_1, \dots, u_4)$ such that $0 < \left|u_j\right| \leqslant P$ for all $j \in [4]$, and let $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_t)$ be an arbitrary collection of such vectors $\boldsymbol{u}_i = (u_{i,1}, \dots, u_{i,4})$. A class $K(\boldsymbol{u}_1, \dots, \boldsymbol{u}_t)$ of solutions of (25) consists of those solutions that satisfy

$$(x_{1,j}, \dots, x_{4,j}, y_{1,j}, \dots, y_{4,j}) \in M(\boldsymbol{u}_j) \quad \text{for all} \quad j \in [t].$$

Now, by Lemma 4.2, we may choose a vector $\boldsymbol{h} = (h_1, \dots, h_4)$, where $0 < \left|h_j\right| \leqslant P$ for all $j \in [4]$, such that $\boldsymbol{h}$ is a member of $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_t)$ and

$$|K(\boldsymbol{u}_1, \dots, \boldsymbol{u}_t)| \leqslant |K(\boldsymbol{h}, \dots, \boldsymbol{h})|.$$

We get

$$N(A_1, \dots, A_1) \leqslant \sum_{(\boldsymbol{u}_1, \dots, \boldsymbol{u}_t)} |K(\boldsymbol{u}_1, \dots, \boldsymbol{u}_t)| \leqslant \sum_{\boldsymbol{h}} r(\boldsymbol{h}) \, |K(\boldsymbol{h}, \dots, \boldsymbol{h})|,$$

where $r(\boldsymbol{h})$ denotes the number of collections $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_t)$ containing $\boldsymbol{h}$.

There is t vectors in the vector $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_t)$. The vector $\boldsymbol{h}$ can be any of them. For this reason,

$$r(\boldsymbol{h}) \leqslant t \cdot \left((2P)^4\right)^{t-1} \leqslant \lambda P^{4t-4}, \quad \text{when} \quad \lambda > 16^{8^{n-2}-1} \cdot 8^{n-2},$$

and

$$N(A_1,\dots,A_1) \leqslant \lambda P^{4t-4} \sum_{\boldsymbol{h}} |K(\boldsymbol{h},\dots,\boldsymbol{h})|\,. \tag{26}$$

If we assume that $0 \leqslant x_{i,j} \leqslant P$ and $0 < |h_i| \leqslant P$, then the sum above does not exceed the number of solutions of the equation

$$\sum_{i=1}^{4}(-1)^i \sum_{j=1}^{t}(-1)^j \left(f(x_{i,j}+h_i) - f(x_{i,j})\right) = 0 \tag{27}$$

with respect to the variables h_i and $x_{i,j}$.

For every integer $h \neq 0$, define

$$\varphi_h(x) = \frac{1}{h}\left(f(x+h) - f(x)\right) \qquad \text{and} \qquad z_i = \sum_{j=1}^{t}(-1)^j \varphi_{h_i}(x_{i,j}) \quad \text{for} \quad i \in [4].$$

Now (27) is equivalent to the equation

$$z_1h_1 - z_2h_2 + z_3h_3 - z_4h_4 = 0. \tag{28}$$

In addition, using the *Binomial theorem* and changing the order of summation, we can write

$$\varphi_h(x) = \frac{1}{h}\sum_{\ell=0}^{n} a_{n-\ell}\left((x+h)^\ell - x^\ell\right) = \frac{1}{h}\sum_{\ell=1}^{n}\sum_{i=0}^{\ell-1} a_{n-\ell}\binom{\ell}{i}x^i h^{\ell-i} = \sum_{i=0}^{n-1} b_{n-1-i}x^i,$$

where

$$b_{n-1-i} = \sum_{i<\ell\leqslant n} a_{n-\ell}\binom{\ell}{i}h^{\ell-i-1}.$$

Since clearly

$$\sum_{i<\ell\leqslant n}\binom{\ell}{i} \leqslant \sum_{i<\ell\leqslant n} 2^\ell \leqslant 2^{n+1},$$

the estimation of b_{n-1-i} and z_i gives

$$|b_{n-1-i}| \leqslant \sum_{i<\ell\leqslant n} \lambda P^{n-\ell}\binom{\ell}{i}P^{\ell-i-1} \leqslant 2^{n+1}\lambda P^{n-i-1}, \quad 0 \leqslant i \leqslant n-1 \tag{29}$$

and

$$|z_i| \leqslant t\sum_{i=0}^{n-1} 2^{n+1}\lambda P^{n-i-1}\cdot P^i \leqslant \lambda^2 P^{n-1}, \quad \text{when} \quad \lambda > 8^{n-2}\cdot 2^{n+1}\cdot n.$$

Under the assumptions $0 < |h_i| \leqslant P$ and $|z_i| \leqslant \lambda^2 P^{n-1}$, the number of solutions of (28) does not exceed the number

$$c\left(P\cdot\lambda^2P^{n-1}\right)^3 \leqslant \lambda^7 P^{3n}, \quad \text{when} \quad \lambda > c, \tag{30}$$

by Lemma 3.6.

Note that $\varphi_h(x)$ is a polynomial of degree $n-1$ satisfying the conditions of the theorem. By Lemma 4.1, the number of solutions of the equation

$$\sum_{j=1}^{t}(-1)^j\varphi_{h_i}(x_{i,j}) = z_i, \quad \text{where} \quad 0 \leqslant x_{i,j} \leqslant P,$$

does not exceed the number of solutions of the equation

$$\sum_{j=1}^{t}(-1)^j\varphi_{h_i}(x_{i,j}) = 0, \quad \text{where} \quad 0 \leqslant x_{i,j} \leqslant P.$$

By induction hypothesis and (29), for sufficiently large λ, this number is bounded from above by the number $(2^{n+1}\lambda)^{t-3}P^{t-(n-1)}$, which is less or equal to $\lambda^{t-2}P^{t-n+1}$, when $\lambda > 2^{(n+1)(8^{n-2}-3)}$. Combining this and estimates in (26) and (30) gives eventually

$$N(A_1,\dots,A_1) \leqslant \lambda P^{4t-4}\cdot\lambda^7P^{3n}\left(\lambda^{t-2}P^{t-n+1}\right)^4 = \lambda^{4t}P^{8t-n}.$$

Estimating the exponents, since the inequality $4\cdot 8^{n-2} \leqslant 8\cdot 8^{n-2}-4$ is equivalent to $64 \leqslant 8^n$, which is true for $n \geqslant 3$, we may write $\lambda^{4t} \leqslant \lambda^{s-4}$. We obtain that $N(A_1,\dots,A_1) \leqslant \lambda^{s-4}P^{s-n}$.

Finally, the number of solutions of (19) is bounded from above by $\lambda^{s-3}P^{s-n}$, when $\lambda > 2^{4\cdot 8^{n-2}}$. By *mathematical induction*, the theorem holds for all $n \geqslant 2$. The proof is complete. □

5. Variants and generalizations

In 246 years Waring's problem has taken many forms. This section works as a brief overview of the different variations to and generalizations of the problem.

5.1. **Variations to Waring's problem.** Recall Question 1.5. Returning to the case of cubes, we mentioned that $g(3) = 9$. Examining the argumentation of Wieferich, Landau (1908) noted that, actually, only finitely many numbers required nine cubes. It seemed that the greatest number of summands is required by certain fairly small integers. We also mentioned the large tabulations done in the 19th century; Jacobi (1851) had tabulated all numbers up to 12 000 as sums of as few cubes as possible. He found out, for example, that only two numbers, namely 23 and 239 required nine cubes — 'an entertaining arithmetical fluke' (Hardy 2011, p. 18). Further, Jacobi observed also that only fifteen numbers, 15, 22, 50, 114, 167, 175, 186, 212, 231, 238, 303, 364, 420, 428, and 454, required eight cubes, and only 121 numbers required seven cubes. The first two of these observations were confirmed later by Dickson (1939) and Siksek (2015) respectively. The last is still an open question. Nevertheless, we can conclude that every sufficiently large positive integer is a sum of seven cubes or fewer, which was first explicitly demonstrated by Linnik (1942, 1943b). As a pair for Definition 1.1 we define $G(n)$.

Definition 5.1. Let n be a natural number greater than one. Now $G = G(n)$ is the smallest number such that every sufficiently large positive integer is a sum of at most G nth powers.

In other words, $G(n)$ is the smallest integer, for which there exists a finite $N \in \mathbf{Z}$ such that, for every natural number $a \geqslant N$ and for every $s \geqslant G(n)$, the equation

$$a = x_1^n + x_2^n + \ldots + x_s^k$$

has a solution $(x_1, \ldots, x_s)$ in non-negative integers. Clearly $G(n) \leqslant g(n)$, $G(n)$ is finite for all n, and it has been shown quite elementarily by Hurwitz (1908) and Maillet (1908) that $G(n) \geqslant n + 1$ for all $n \geqslant 2$ (see also Hardy, Wright et al. 2008, pp. 426–427). A proof of $G(n) = O(n \ln n)$ can be found in (Bredikhin and Grišina 1978). Results like these are called asymptotic theorems.

The definition of $G(n)$ gives rise to the so-called Modern Waring's problem, around which there is plenty of research going on. In general, the problem is much more involved and complex than the classical Waring's problem; fairly little is known on $G(n)$. The number itself is considered far more fundamental than $g(n)$. In fact, the Hardy–Littlewood–Vinogradov method leads straight to upper bounds for $G(n)$. If such a bound $\bar{G}(n)$ is found, and if it is less than or equal to the lower bound $\underline{g}(n)$ for $g(n)$, then it can be used in the computation of $g(n)$: using the method, just extract a N_0 such that every $N > N_0$ is a sum of $\bar{G}(n)$ nth powers. This determines $g(n)$ unambiguously. For example, the original Hardy–Littlewood method gives

$$\bar{G}(n) \leqslant 2^{n-1} n + 1,$$

with which the finiteness of $g(n)$ can easily be established.

Even the determination of the values of $G(n)$ have proven out to be fairly difficult. The discussion above implies that $4 \leqslant G(3) \leqslant 7$; it is conjectured that every large integer is the sum of four non-negative cubes. $G(2) = 4$ follows from Legendre's three-square theorem and Lagrange's four-square theorem. Apart from the case $n = 2$, the only exact estimate achieved is $G(4) = 16$ by Davenport (1939b).

Most recent upper bounds for $G(n)$, achieved by Vaughan and Wooley (1994, 1995) and Wooley (2016), are $G(5) \leqslant 17$, $G(6) \leqslant 24$, $G(7) \leqslant 31$, $G(8) \leqslant 39$, $G(9) \leqslant 47$, $G(10) \leqslant 55$, $G(11) \leqslant 63$, $G(12) \leqslant 72$, $G(13) \leqslant 81$, $G(14) \leqslant 90$, $G(15) \leqslant 99$ and $G(16) \leqslant 108$. As with the most of the latest bounds for $G(n)$, these were obtained by exploiting the Hardy–Littlewood–Vinogradov method, more precisely, Vinogradov's mean-value theorem. Again, we refer the reader to (Vaughan and Wooley 2002; Hardy, Wright et al. 2008, pp. 444–450) for accounts of the historical evolution of the bounds of $G(n)$.

For large n, the sharpest general upper bound for $G(n)$ is

$$G(n) \leqslant n \left(\ln n + \ln \ln n + 2 + O\left(\frac{\ln \ln n}{\ln n} \right) \right) \tag{31}$$

by Wooley (1995). It may be reasonable to conjecture that $G(n) = O(n)$ or even $G(n) = n + 1$, if there are no 'local obstructions'.

Instead of asking the minimal number of summands for large enough integers, we could investigate another asymptotic behaviour of $g(n)$. Let $G'(n)$ be the smallest number such that *almost every* number a is the sum of at most $G'(n)$ nth powers of positive integers. This means that $\sigma(A) = 1$, where (A) is the sequence of integers with the property mentioned. In other words, if (B) is the sequence of integers, which cannot be expressed as sums of $G'(n)$ nth powers, then $\sigma(B) = 0$. The problem of

determining $G'(n)$ is easier than that of $G(n)$; the value of $G'(n)$ is determined for six non-trivial cases: $G'(2) = 4$, $G'(3) = 4$ (Davenport 1939a), $G'(4) = 15$ (Hardy and Littlewood 1925), $G'(8) = 32$ (Vaughan 1986), $G'(16) = 64$ (Wooley 1992), and $G'(32) = 128$ (Wooley 1992).

Wright (1934) formulated an interesting variant to Waring's problem while intending to weaken the usual problem. Instead of restricting ourselves to only summation of integer powers, so-called Easier Waring's problem asks whether there exists a finite $v = v(n)$ such that for all $a \in \mathbf{N}$ the equation

$$a = \pm x_1^n \pm x_2^n \pm \ldots \pm x_v^n$$

has a solution $(x_1, x_2, \ldots, x_v)$ for some choice of signs. The finiteness of $v(n)$ follows from $g(n) < \infty$. For small n, a couple of values of $v(n)$ has been found with the help of elementary identities. For example, $v(2) = 3$, $4 \leqslant v(3) \leqslant 5$ and $v(5) \leqslant 18$. In addition, it has been shown that $v(n) = O(n \ln(n))$. As with $G(n)$, the conjecture is that $v(n) = O(n)$.

Of course we can also consider sums of mixed powers, that is the non-homogeneous Waring equations. Can all integers be represented as the sum of one square, four cubes and one biquadrate of positive integers? What about almost every integer as the sum of two cubes and two biquadrates? The answer to the first one is positive, while the latter is still an unsolved problem. Such variations to Waring's problem were first studied when the Hardy–Littlewood method were in its infancy; the method is central to practical treatment of problems like these but also other, more effective methods have been used. Nonetheless, new technology was developed, which helped mathematicians to make progress even in the classical version of the problem. Most interesting cases are perhaps those, in which the summands are restricted to be squares, cubes or biquadrates at most. A listing of achieved cases can be found in (Vaughan and Wooley 2002, pp. 21–22).

Introduction of coefficients is another interesting variation. We have, for example, that every positive integer can be expressed in the form

$$x^2 + 2y^2 + 3z^2 + 6w^2.$$

Considering the numbers expressible in the form $x^2 + y^2$ leads us quickly to analyse representations of the form

$$ax^2 + bxy + cy^2, \quad a, b, c \in \mathbf{Z},$$

essentially, the theory of quadratic forms. Further, $x^3 + y^3 + z^3$ brings us to ternary quadratic forms, and so on.

All in all, the list of different variations is abundant. As the last one we mention $g(n, m)$ defined and solved by Small (1977a,c). The number $g(n, m)$ is the smallest positive integer s such that for all $a \in \mathbf{N}$ the equation

$$a \equiv x_1^n + x_2^n + \ldots + x_s^n \pmod{m}$$

has a solution in non-negative integers. Here one assumes that $n > 1$ and $m > 1$.

5.2. **Generalizations of Waring's problem.** As was already hinted in Section 3, Waring's problem generalizes naturally to integer-valued polynomials (see Kamke 1921). Let us state this assertion formally. Let $f(x)$ be an integer-valued polynomial, that is,

for all $x \in \mathbf{Z}$, $f(x) \in \mathbf{Z}$, coefficients of which need not be integers, but with a positive leading coefficient. Define the set

$$A(f) = \{f(k) \mid k \in \mathbf{N} \cup \{0\}\}\,.$$

If k is large enough, that is after some K, the term of highest degree in $f(x)$ dominates the others. Thus, for all $k \geqslant K$, $f(k)$ is monotonically increasing sequence of integers. Moreover, the polynomial $f_K(x) = f(x + K)$ has the same degree and leading coefficient as $f(x)$. Therefore it is all the same whether we consider f or f_K, or just assume that $f(k)$ is monotonically increasing.

In addition, let d be the greatest common divisor of $A(f)$. Clearly, the polynomial $f(x)/d$ is also integer-valued, of the same degree, and the greatest common divisor of $A(f(x)/d)$ is one. Thus we may assume without loss of generality that $d = 1$.

Theorem 5.2 (Waring's problem for polynomials)**.** *Let $f(x) = \sum_{i=0}^{n} a_i x^i$ be an integer-valued polynomial with $a_n > 0$. If $\gcd(A(f)) = 1$ and $0, 1 \in A(f)$, then there exists $h \in \mathbf{N}$ such that every non-negative integer can be written as the sum of at most h elements of $A(f)$.*

For a complete proof see for example (Nathanson 2000, pp. 355–373). In the light of this result, the classical Waring's problem is the special case $f(x) = x^n$ (recall the proof of Fundamental lemma). In a sense this problem dates back to Fermat, who claimed in 1640 that 'every number is either triangular or the sum of 2 or 3 triangular numbers; every number is either a square or the sum of 2, 3 or 4 squares; either pentagonal or the sum of 2, 3, 4 or 5 pentagonal numbers; and so on' (Dickson 1920, p. 6). Coincidence or not, the expression

$$P_r(n) = \frac{1}{2}(r-2)\left(n^2 - n\right) + n, \quad \text{where} \quad r > 2,$$

for the nth r-gonal number occurs in Waring's book (Waring 1991) just preceding the statement of Waring's problem!

Another way to generalize the problem is the question about algebraic number fields or even about arbitrary fields or rings. Compared to the classical problem, the problems around expressing the elements of a field K as the sum of nth powers are far from complete. In addition, one must overcome the difficulty of precisely translating Waring's problem to the broader context. The following is one possible interpretation. Let K be a number field. The *ring of integers* of an algebraic number field is the ring of all integral elements contained in the field. Let R_K be the ring of integers of K. Define S_n as the subring of R_K generated by the nth powers of integers of K. To give an analogue to the positivity of the summands, define $G_K(n)$ to be the smallest positive integer s such that, for some $c = c(n, K) > 0$ and for all *totally positive* integers $\nu \in S_n$ with a sufficiently large norm $\mathrm{N}(\nu)$, the equation

$$\nu = \lambda_1^n + \lambda_2^n + \ldots + \lambda_s^n$$

has always a solution in totally non-negative integers $\lambda_i \in K$, for which the inequality $\mathrm{N}(\lambda_i) \leqslant c\mathrm{N}(\nu)^{1/n}$ holds for $i \in [s]$.

Both Hilbert's, and Hardy and Littlewood's methods have been developed to work partly in number fields in general. However, even the best results achieved with these

methods are far from (31); it has been shown, for example, that

$$G_K(n) \leqslant \max\left(8n^5, 2^n + 1\right).$$

New methods employing smooth numbers and repeated efficient differences have proven out to be more successful.

5.3. **On Goldbach's conjecture.** For many years Waring's problem possessed a similar status to that of Goldbach's conjecture as the most famous unsolved problem in additive number theory. Now that Waring's problem can be considered more or less solved, we find it fitting to give a brief introduction to the latter. In its original form the conjecture was posed by Goldbach in a letter to Euler in 1740s.

Conjecture 5.3 (Goldbach's conjecture)**.** Every even integer greater than 2 can be expressed as the sum of two primes.

If we assume that Goldbach's conjecture holds, then every even integer greater than 4 is the sum of two odd primes. For example,

$$332 = 313 + 19.$$

Adding 3 on both sides we get

$$335 = 313 + 19 + 3,$$

that is, 335 is the sum of three primes. Generally, adding 3 to each even number greater than 4 will give all the odd numbers greater than 7. This implies

Conjecture 5.4 (Goldbach's weak conjecture)**.** Every odd integer greater than 5 can be expressed as the sum of three primes.

We already mentioned Vinogradov's theorem, a weaker form of the above, proved in 1928 by the use of Vinogradov's mean value theorem. A few years later Shnirelman succeeded in applying the theory of Shnirelman density to the problem.

Theorem 5.5. *Any natural number greater than one can be written as the sum of not more than C prime numbers.*

The estimation of the number C, dubbed the *Shnirelman constant* has been since targeted by a legion of mathematicians. Shnirelman himself achieved $C < 800\,000$ whereas the best current estimates are $C < 7$ by Ramaré (1995) and $C < 6$ by Tao (2012). Goldbach's conjecture has been checked to hold for all integers up through $4 \cdot 10^{18}$ but no rigorous proof for it have been found.

We end this thesis with a tantalizing question combining both Waring's problem and Goldbach's conjecture.

Question 5.6 (Waring–Goldbach problem). For all natural numbers n, is there a finite g such that every sufficiently large $a \in \mathbf{N}$ can be represented as the sum of at most g nth powers of prime numbers?

Hua (1938) confirmed that the answer to Waring–Goldbach problem is affirmative, which leaves us with the numerical problem. To illustrate Hua's conclusion and the most recent progress made, we must introduce $H(n)$, the associated local conditions of which are embedded in its definition; the Hardy–Littlewood method applied precisely

indicates that if $s \geqslant n+1$ for natural s and n, then all large $a \in \mathbf{N}$ satisfying some congruence conditions can be written as the sum of s nth powers of prime numbers.

For a natural number n and a prime number p, define $\theta = \theta(n,p)$ as the (unique) integer so that $p^\theta \mid n$ but $p^{\theta+1} \nmid n$. In addition, set

$$\gamma = \gamma(n,p) = \begin{cases} \theta+2, & \text{if } p = 2 \text{ and } \theta > 0, \\ \theta+1, & \text{otherwise,} \end{cases} \quad \text{and} \quad K(n) = \prod_{(p-1)\mid n} p^\gamma.$$

Then $H(n)$ is the smallest integer s such that every sufficiently large positive integer a congruent to $s \pmod{K(n)}$ can be written in the form

$$a = p_1^n + p_2^n + \ldots + p_s^n,$$

where $p_1, p_2, \ldots, p_s$ are prime numbers.

Motivation for the unwieldy definition is the following. Particularly, $K(1) = 2$. If $(p-1) \mid n$, then $p^\theta(p-1) \mid n$. Thus $b^n \equiv 1 \pmod{p^\gamma}$ as long as $\gcd(p,b) = 1$. We get that if a is the sum of s nth powers of prime numbers $p_i > n+1$, then it must be that $a \equiv s \pmod{K(n)}$. The problem without these kind of restrictions is described by, for example, Buttcane (2010) and Chubarikov (2009).

Basically, Hua (1938) achieved the general bound

$$H(n) \leqslant 2^n + 1, \quad n \geqslant 1,$$

which is the best known for $n \in \{1,2,3\}$, by a generalization of Vinogradov's theorem. For large n, however, Hua (1959, 1965) sharpened the bound to

$$H(n) \leqslant n(4\ln n + 2\ln\ln n + O(1)), \quad \text{as} \quad n \to \infty.$$

Not until recently was this bound improved further to

$$H(n) \leqslant (4n-2)\ln n - (2\ln 2 - 1)n - 3$$

by Kumchev and Wooley (2016). The conjecture that $H(n) = n+1$ for all $n \geqslant 1$ looms far on the horizon.

References

Balasubramanian, R., J.-M. Deshouillers and F. Dress (1986a). 'Problème de Waring pour les bicarrés. I. Schéma de la solution'. In: *Comptes Rendus de l'Académie des Sciences. Mathématique*. 1st ser. 303.4, pp. 85–88.

– (1986b). 'Problème de Waring pour les bicarrés. II. Résultats auxiliaires pour le théorème asymptotique'. In: *Comptes Rendus de l'Académie des Sciences. Mathématique*. 1st ser. 303.5, pp. 161–163.

Batchelder, P. M. (1936). 'Waring's Problem'. In: *The American Mathematical Monthly* 43.1, pp. 21–27.

Bredikhin, B. M. and T. I. Grišina (1978). 'An elementary estimate of $G(n)$ in Waring's problem'. In: *Mathematical notes of the Academy of Sciences of the USSR* 24.1, pp. 507–513.

Bretschneider, C. A. (1853). 'Tafeln für die Zerlegung der Zahlen bis 4100 in Biquadrate'. In: *Journal für die reine und angewandte Mathematik* 46, pp. 1–23.

Buttcane, J. (2010). 'A note on the Waring–Goldbach problem'. In: *Journal of Number Theory* 130.1, pp. 116–127.

Calderón, C. (2011). 'On the classical Waring problem'. In: *Revista de la Real Academia de Ciencias Exactas, Físicas, Químicas y Naturales de Zaragoza*. 2nd ser. 66, pp. 105–125.

Chen, J.-R. (1964). 'Waring's problem for $g(5) = 37$'. In: *Scientia Sinica* 13, pp. 1547–1568.

Chubarikov, V. N. (2009). 'On the Waring-Goldbach problem'. In: *Doklady Mathematics* 80.1, pp. 470–473.

Davenport, H. (1939a). 'On Waring's problem for cubes'. In: *Acta Mathematica* 71.1, pp. 123–143.

– (1939b). 'On Waring's problem for fourth powers'. In: *Annals of Mathematics*. 2nd ser. 40.4, pp. 731–747.

Davenport, H. (2005). *Analytic Methods for Diophantine Equations and Diophantine Inequalities*. Ed. by T. D. Browning. With a forew. by R. C. Vaughan, D. R. Heath–Brown and D. E. Freeman. 2nd ed. Cambridge, United Kingdom: Cambridge University Press.
Dickson, L. E. (1920). *History of the theory of numbers. Diophantine analysis*. Vol. 2. Washington D.C., United States: Carnegie Institution of Washington.
– (1936). 'Proof of the ideal Waring theorem for exponents 7–180'. In: *American Journal of Mathematics* 58.3, pp. 521–529.
– (1939). 'All integers except 23 and 239 are sums of eight cubes'. In: *Bulletin of the American Mathematical Society* 45.8, pp. 588–591.
Ellison, W. J. (1971). 'Waring's problem'. In: *The American Mathematical Monthly* 78.1, pp. 10–36.
Euler, L. P. (1862). *Opera postuma. Mathematica et physica*. Vol. 1. Saint Petersburg, Russia: Eggers.
Fleck, A. (1906). 'Über die Darstellung ganzer Zahlen als Summen von positiven Kuben und als Summen von Biquadraten ganzer Zahlen'. In: *Sitzungsberichte der Berliner Mathematischen Gesellschaft* 5, pp. 2–9.
– (1907). 'Über die Darstellung ganzer Zahlen als Summen von sechsten Potenzen ganzer Zahlen'. In: *Mathematische Annalen* 64.4, pp. 561–566.
Halberstam, H. (2002). *Loo-Keng Hua*. A biographical memoir. Washington D.C., United States: National Academy of Sciences.
Hardy, G. H. (2011). *Some famous problems of the theory of numbers and in particular Waring's problem*. An inaugural lecture delivered before the University of Oxford. Oxford, United Kingdom: The Clarendon Press.
Hardy, G. H. and J. E. Littlewood (1919). 'A new solution of Waring's problem'. In: *Quarterly Journal of Mathematics* 48, pp. 272–293.
– (1920). 'Some problems of 'Partitio numerorum' (I). A new solution of Waring's problem'. In: *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, pp. 33–54.
– (1925). 'Some problems of 'Partitio numerorum' (VI). Further researches in Waring's problem'. In: *Mathematische Zeitschrift* 23.1, pp. 1–37.
Hardy, G. H. and S. Ramanujan (1918). 'Asymptotic formulaæ in combinatory analysis'. In: *Proceedings of the London Mathematical Society* 17.1, pp. 75–115.
Hardy, G. H., E. M. Wright et al. (2008). *An Introduction to the Theory of Numbers*. With a forew. by A. Wiles. 6th ed. Oxford mathematics. Oxford, United Kingdom: Oxford University Press.
Hilbert, D. (1909). 'Beweis für die Darstellbarkeit der ganzen Zahlen durch eine feste Anzahl *n*ter Potenzen (Waringsches Problem)'. In: *Mathematische Annalen* 67.3, pp. 281–300.
– (1912). 'Prix Bolyai. Procès verbal des séances de la commission Internationale de 1910'. In: *Acta Mathematica* 35.1, pp. 1–28.
Hua, L.-K. (1938). 'Some results in the additive prime-number theory'. In: *The Quarterly Journal of Mathematics* 9.1, pp. 68–80.
– (1959). *Additive Primzahltheorie*. Leipzig, Germany: B. G. Teubner.
– (1965). *Additive theory of prime numbers*. Translations of Mathematical Monographs 13. Providence, Rhode Island, United States: American Mathematical Society.
– (1982). *Introduction to Number Theory*. Trans. by P. Shiu. Berlin, Germany: Springer.
Hurwitz, A. (1908). 'Über die Darstellung der ganzen Zahlen als Summen von *n*ten Potenzen ganzer Zahlen'. In: *Mathematische Annalen* 65.3, pp. 424–427.
Jacobi, C. G. J. (1851). 'Über die Zusammensetzung der Zahlen aus ganzen positiven Cuben; nebst einer Tabelle für die kleinste Cubenanzahl, aus welcher jede Zahl bis 12000 zusammengesetzt werden kann'. In: *Journal für die reine und angewandte Mathematik* 42, pp. 41–69.
Jameson, T. (2015). 'Linnik's proof of the Waring–Hilbert theorem from Hua's book (with a correction)'. URL: http://www.maths.lancs.ac.uk/∼jameson/warlin.pdf (visited on 20/12/2015).
Kamke, E. (1921). 'Verallgemeinerungen des Waring-Hilbertschen Satzes'. In: *Mathematische Annalen* 83.1–2, pp. 85–112.
Kempner, A. J. (1912). 'Bemerkungen zum Waringschen Problem'. In: *Mathematische Annalen* 72.3, pp. 387–399.
Khinchin, A. Y. (1998). *Three pearls of number theory*. Trans. by F. Bagemihl, H. Komm and W. Seidel. Mineola, New York, United States: Dover Publications.
Kubina, J. M. and M. C. Wunderlich (1990). 'Extending Waring's conjecture to 471,600,000'. In: *Mathematics of computation* 55.192, pp. 815–820.
Kumchev, A. V. and T. D. Wooley (2016). 'On the Waring-Goldbach problem for seventh and higher powers'. Submitted.
Landau, E. (1907). 'Über die Darstellung einer ganzer Zahl als Summe von Biquadraten'. In: *Rendiconti del Circolo matematico di Palermo* 23, pp. 91–96.
– (1908). 'Über eine Anwendung der Primzahltheorie auf das Waringsche Problem in der elementaren Zahlentheorie'. In: *Mathematische Annalen* 66.1, pp. 102–105.

Lebesgue, V.-A. (1859). *Exercices d'analyse numérique*. Paris, France.
LeVeque, W. J. (1996). *Fundamentals of number theory*. Mineola, New York, United States: Dover Publications.
Linnik, Y. V. (1942). 'On the representation of large numbers as sums of seven cubes'. In: *Proceedings of the USSR Academy of Sciences* 35, p. 162.
– (1943a). 'An elementary solution of the problem of Waring by Schnirelman's method'. In: *Matematicheskii Sbornik* 12.2 (54), pp. 225–230.
– (1943b). 'On the representation of large numbers as sums of seven cubes'. In: *Matematicheskii Sbornik* 12.2 (54), pp. 218–224.
Lucas, M. É. (1878a). 'Sur la décomposition des nombres en bicarrés'. In: *Nouvelle correspondance mathématique* 4, pp. 323–325.
– (1878b). 'Sur un théorème de M. Liouville, concernant la décomposition des nombres en bicarrés'. In: *Nouvelles annales de mathématiques, journal des candidats aux écoles polytechnique et normale*. 2nd ser. 17, pp. 536–537.
Mahler, K. (1957). 'On the fractional parts of the powers of a rational number (II)'. In: *Mathematika* 4.2, pp. 122–124.
Maillet, E. (1895). 'Sur la décomposition d'un nombre entier en une somme de cubes d'entiers positifs'. In: *Association française pour l'avancement des sciences* 24, pp. 242–247.
– (1896). 'Quelques extensions du théorème de Fermat sur les nombres polygones'. In: *Journal de mathématiques pures et appliquées*. 5th ser. 2, pp. 363–380.
– (1908). 'Sur la décomposition d'un entier en une somme de puissances huitièmes d'entiers (Problème de Waring)'. In: *Bulletin de la Société Mathématique de France* 36, pp. 69–77.
Mann, H. B. (1942). 'A proof of the fundamental theorem on the density of sums of sets of positive integers'. In: *Annals of Mathematics*. 2nd ser. 43.3, pp. 523–527.
Nathanson, M. B. (1996). *Additive Number Theory. The Classical Bases*. Graduate Texts in Mathematics 164. New York, United States: Springer.
– (2000). *Elementary methods in number theory*. Graduate Texts in Mathematics 195. New York, United States: Springer.
Nesterenko, Y. V. (2006). 'On Waring's problem (elementary methods)'. Trans. by A. V. Yakovlev. In: *Journal of Mathematical Sciences* 137.2, pp. 4699–4715. Trans. of 'On Waring's problem (elementary methods)'. In: *Proceedings on number theory*. Vol. 322: *Zapiski Nauchnykh Seminarov POMI*. (St. Petersburg). Ed. by Y. V. Nesterenko and A. I. Vinogradov. 2005, pp. 149–175.
Newman, D. J. (1960). 'A simplified proof of Waring's conjecture'. In: *The Michigan Mathematical Journal* 7.3, pp. 291–295.
– (1997). *Analytic Number Theory*. Graduate Texts in Mathematics 177. New York, United States: Springer.
O'Connor, J. J. and E. F. Robertson (2015). *MacTutor History of Mathematics archive*. School of Mathematics and Statistics, University of St Andrews. URL: http://www-history.mcs.st-andrews.ac.uk/index.html (visited on 18/12/2015).
OEIS Foundation Inc. (2016). *The On-Line Encyclopedia of Integer Sequences. A002804*. URL: http://oeis.org/A002804 (visited on 12/06/2016).
Pillai, S. S. (1940). 'On Waring's problem $g(6) = 73$'. In: *Proceedings of the Indian Academy of Sciences. Section A* 12.1, pp. 30–40.
Pollack, P. (2004). *Not Always Buried Deep. Selections from Analytic and Combinatorial Number Theory*. Course Notes. URL: http://staff.polito.it/danilo.bazzanella/PhD_files/Not%20always%20buried%20deep%20(Pollack).pdf (visited on 21/12/2015).
– (2011). 'On Hilbert's solution of Waring's problem'. In: *Open Mathematics* 9.2, pp. 294–301.
Pupyrev, Y. A. (2009). 'Effectivization of a lower bound for $\|(4/3)^k\|$'. In: *Mathematical Notes* 85.5, pp. 877–885.
Ramaré, O. (1995). 'On Šnirel'man's Constant'. In: *Annali della Scuola Normale Superiore di Pisa - Classe di Scienze* 22.4, pp. 645–706.
Réalis, M. S. (1878). 'Note sur un théorème d'arithmétique'. In: *Nouvelle correspondance mathématique* 4, pp. 209–210.
Rieger, G. J. (1953). 'Zur Hilbertschen Lösung des Waringschen Problems: Abschätzung von $g(n)$'. In: *Archiv für Mathematische Logik und Grundlagenforschung* 4, pp. 275–281.
Robertson, J. C. and T. Byerley, eds. (1821). *The Percy anecdotes. Original and select*. The Percy Anecdotes 6. Ludgate Hill, London: T. Boys, pp. 83–84.
Russian Academy of Sciences (2002). *Linnik Yuri Vladimirovich*. URL: http://www.ras.ru/win/db/show_per.asp?P=.id-51108.ln-en (visited on 07/06/2016).
Shnirel'man, L. G. (1933). 'Über additive Eigenschaften von Zahlen'. In: *Mathematische Annalen* 107.1, pp. 649–690.
Siksek, S. (2015). 'Every integer greater than 454 is the sum of at most seven positive cubes'. Submitted.

Small, C. (1977a). 'Solution of Waring's Problem mod *n*'. In: *The American Mathematical Monthly* 84.5, pp. 356–359.
– (1977b). 'Waring's Problem'. In: *Mathematics Magazine* 50.1, pp. 12–16.
– (1977c). 'Waring's Problem mod *n*'. In: *The American Mathematical Monthly* 84.1, pp. 12–25.
von Sterneck, R. D. (1903). 'Über die kleinste Anzahl Kuben, aus welchen jede Zahl bis 40000 zusammengesetzt werden kann'. In: *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften in Wien, mathematisch-naturwissenschaftliche Classe* 112, pp. 1627–1666.
Swetz, F. J. (2007). *Portrait Gallery*. Mathematical Association of America. URL: http://www.maa.org/press/periodicals/convergence/portrait-gallery (visited on 21/12/2015).
Tao, T. (2012). 'Every odd number greater than 1 is the sum of at most five primes'. Forthcoming.
Vaughan, R. C. (1986). 'On Waring's problem for smaller exponents'. In: *Proceedings of the London Mathematical Society* 3.52 (3), pp. 445–463.
– (1997). *The Hardy–Littlewood Method*. 2nd ed. Cambridge Tracts in Mathematics 125. Cambridge, United Kingdom: Cambridge University Press.
Vaughan, R. C. and T. D. Wooley (1994). 'Further improvements in Waring's problem. II: Sixth powers'. In: *Duke Mathematical Journal* 76.3, pp. 683–710.
– (1995). 'Further improvements in Waring's problem'. In: *Acta Mathematica* 174.2, pp. 147–240.
– (2002). 'Waring's Problem. A Survey'. In: *Number Theory for the Millennium* 3, pp. 301–340.
Vinogradov, I. M. (1985). 'On Waring's theorem'. In: *Selected Works*. Ed. by L. D. Faddeev et al. Trans. by P. S. V. Naidu. Springer Collected Works in Mathematics. Berlin, Germany: Springer, pp. 101–106.
Wagstaff Jr., S. S. (1975). 'The Schnirelmann density of the sums of three squares'. In: *Proceedings of the American Mathematical Society* 52, pp. 1–7.
Waring, E. M. (1991). *Meditationes Algebraicae*. Trans. by D. Weeks. Providence, Rhode Island, United States: American Mathematical Society.
Wieferich, A. J. A. (1908a). 'Beweis des Satzes, daß sich eine jede ganze Zahl als Summe von höchstens neun positiven Kuben darstellen läßt'. In: *Mathematische Annalen* 66, pp. 95–101.
– (1908b). 'Über die Darstellung der Zahlen als Summen von Biquadraten'. In: *Mathematische Annalen* 66, pp. 106–108.
– (1909). 'Zur Darstellung der Zahlen als Summen von 5ten und 7ten Potenzen positiver ganzer Zahlen'. In: *Mathematische Annalen* 67, pp. 61–75.
Wikimedia Commons (2015). *Hua Luogeng 1956*. URL: https://commons.wikimedia.org/wiki/File:Hua_Luogeng_1956.jpg (visited on 07/06/2016).
Wooley, T. D. (1992). 'Large improvements in Waring's problem'. In: *Annals of mathematics*. 2nd ser. 135.1, pp. 131–164.
– (1995). 'New estimates for smooth Weyl sums'. In: *Journal of the London Mathematical Society*. 2nd ser. 51.1, pp. 1–13.
– (2016). 'On Waring's problem for intermediate powers'. Submitted.
Wright, E. M. (1934). 'An Easier Waring's Problem'. In: *The Journal of the London Mathematical Society* 9.4, pp. 267–272.
Zornow, A. R. (1835). 'De compositione numerorum e cubis integris positivis'. In: *Journal für die reine und angewandte Mathematik* 14, pp. 276–280.