

# A Six Generalized Squares Theorem, with Applications to Polynomial Identity Algebras

Paula B. Cohen<sup>1</sup>

*School of Mathematics, Institute for Advanced Study, Princeton, New Jersey 08540 and  
UMR AGAT au CNRS, Université des Sciences et Technologies de Lille,  
59655 Villeneuve d'Ascq cedex, France*

E-mail: [pcohen@math.ias.edu](mailto:pcohen@math.ias.edu), [Paula.Cohen@univ-lille1.fr](mailto:Paula.Cohen@univ-lille1.fr)

and

Amitai Regev<sup>2</sup>

*Department of Mathematics, The Weizmann Institute of Science, Rehovot 76100, Israel*

E-mail: [regev@wisdom.weizmann.ac.il](mailto:regev@wisdom.weizmann.ac.il)

*Communicated by Susan Montgomery*

Received March 3, 2000

The theories of superalgebras and of P.I. algebras lead to a natural  $\mathbb{Z}_2$ -graded extension of the integers. For these generalized integers, a “six generalized squares” theorem is proved, which can be considered as a  $\mathbb{Z}_2$ -graded analogue of the classical “four squares” theorem for the natural numbers. This theorem was conjectured by A. Berele and A. Regev (“Exponential Growth of Some P.I. Algebras,” [BR2]) and has applications to p.i. algebras. © 2001 Academic Press

*Key Words:* additive number theory; four squares theorem; exponents of P.I. algebras; Amitsur–Cappelli polynomials.

## 1. INTRODUCTION

The celebrated *four squares theorem* from number theory says that every positive integer is a sum of four squares. The study of superalgebras leads to a natural generalization  $B = \mathbb{Z}[t]$ ,  $t^2 = 1$ , of the integers, for which we

<sup>1</sup> The first author thanks the Weizmann Institute of Science where this research was initiated. She also thanks the Institute for Advanced Study, Princeton, NJ, where this research was completed with the support of the Ellentuck Fund.

<sup>2</sup> The second author is partially supported by ISF Grant 6629/1.



prove a *six generalized squares theorem*. Writing  $t = (0, 1)$ , we can identify  $B$  with  $\mathbb{Z} \times \mathbb{Z}$ , endowed with the addition

$$(a, b) + (c, d) = (a + c, b + d), \quad (a, b), (c, d) \in \mathbb{Z} \times \mathbb{Z}$$

and the multiplication

$$(a, b)(c, d) = (ac + bd, ad + bd), \quad (a, b), (c, d) \in \mathbb{Z} \times \mathbb{Z}.$$

We adopt throughout the convention that  $0 \in \mathbb{N}$ . We call

$$\mathcal{P} = \{(r^2, r^2), (r^2 + s^2, 2rs) \mid r, s \in \mathbb{N}\}$$

the set of the generalized squares in  $\mathbb{Z}[t]$ . Justification of this terminology is given in the sequel.

In this paper we prove,

**THE SIX GENERALIZED SQUARES THEOREM.** *Given  $r \geq s \geq 0$  in  $\mathbb{N}$ , the pair  $(r, s)$  is always the sum of at most six elements in  $\mathcal{P}$ .*

Note that, for example,  $(10, 3)$  is not a sum of any five elements of  $\mathcal{P}$ .

It is possible to formulate the six generalized squares theorem in terms of  $\mathbb{N} \times \mathbb{N}$  (endowed with the usual componentwise addition and multiplication). Under the map  $(a, b) \mapsto (a + b, a - b)$  the generalized squares  $\mathcal{P}$  are mapped bijectively onto the set

$$\mathcal{S} = \{(2n^2, 0), (n^2, m^2), n \equiv m \pmod{2}, n \geq m\},$$

and the six generalized squares theorem can be restated as follows.

**SIX GENERALIZED SQUARES THEOREM RESTATED.** *Given  $u \geq v \geq 0$  in  $\mathbb{N}$  with  $u \equiv v \pmod{2}$ , the pair  $(u, v)$  is always the sum of at most six elements in  $\mathcal{S}$ .*

We now explain how representation theory and Young diagrams suggest that  $B$  is the right  $\mathbb{Z}_2$  (or “super”) generalization of  $\mathbb{Z}$  and that  $\mathcal{P}$  is the right generalization of the squares in  $\mathbb{Z}$ . Applications of the six generalized squares theorem, which appear in [BR2], are described below.

Squares of integers are the dimensions of the matrix algebras. In the classical (Jacobson) ring theory, arbitrary algebras are, in a sense, approximated by simple (i.e., matrix) algebras. Thus, for example, the classical four square theorem in number theory implies that the dimension of any finite dimensional algebra is the sum of the dimensions of four or less matrix algebras.

The theory of algebras satisfying polynomial identities (P.I. algebras) suggests a natural generalisation of the above. In Kemer’s structure theory [K], the passage from the general to the finitely generated case uses

superalgebras, and it introduces new classes of “simple” superalgebras, called *verbally prime*. Through the theory of the cocharacters of P.I. algebras—and their corresponding Young diagrams—the dimension  $\dim A$  is replaced here by a certain pair  $\text{hook}(A) = (k, l) \in \mathbb{N} \times \mathbb{N}$ . When  $A$  is a matrix algebra,  $\text{hook}(A) = (r^2, 0)$  where  $r^2 = \dim A$ . The hook pairs of the verbally prime superalgebras are the above set

$$\mathcal{P} = \{\text{hook}(A) \mid A \text{ verbally prime}\} = \{(r^2, r^2), (r^2 + s^2, 2rs) \mid r, s \in \mathbb{N}\}.$$

We call  $\mathcal{P}$  the set of the generalized squares. The approximation of a superalgebra by verbally prime superalgebras leads to expressing a given  $(k, l) \in \mathbb{N} \times \mathbb{N}$  as a sum of such generalized squares, where addition in  $\mathbb{N} \times \mathbb{N}$  is coordinatewise. Conversely, expressing  $(k, l)$  as a sum of such generalized squares leads to the corresponding approximation of a given P.I. algebra by verbally prime algebras [BR2, Sect. 4]. It is natural to ask what the minimal necessary number of such generalized squares is in such a presentation. Moreover, as explained below, the elements of  $\mathcal{P}$  are indeed squares under a natural new multiplication in  $\mathbb{N} \times \mathbb{N}$ , once we admit  $(1, 1)$  as a square.

We describe this now in some more detail. In the character theory of  $S_n$  (the  $n$ th symmetric group), the theory of Lie superalgebras, and the theory of P.I. algebras, an integer  $k \in \mathbb{N}$  corresponds to an infinite strip of the Young diagrams of height at most  $k$ . Such strips of Young diagrams arise in the representation theory of classical Lie algebras, while the more general  $(k, l)$  hooks of Young diagrams arise in the study of Lie superalgebras.

The irreducible  $S_n$  characters are given by  $\text{Par}(n)$ , the partitions of  $n$ . An  $S_n$  character  $\chi_n$  is *supported* by  $K \subseteq \text{Par}(n)$  if  $\chi_n = \sum_{\lambda \in K} m_\lambda \chi_\lambda$ , where  $\chi_\lambda$ 's are the  $S_n$  irreducible characters, and the  $m_\lambda \in \mathbb{N}$  are their multiplicities in  $\chi_n$ .

DEFINITION. Let  $\text{Par} = \bigcup_n \text{Par}(n)$ . Denote

$$H(k, l) = \{\lambda = (\lambda_1, \lambda_2, \dots) \in \text{Par} \mid \lambda_{k+1} \leq l\}, \quad k, l \in \mathbb{N}.$$

We shall often identify  $(k, l)$  with  $H(k, l)$ .

In P.I. theory, one associates with a given P.I. algebra  $A$  a certain sequence  $\chi_n(A)$ ,  $n = 0, 1, 2, \dots$ , of  $S_n$  characters: these are the so-called *cocharacters* of  $A$ , and  $c_n(A) = \deg(\chi_n(A))$  are its *codimensions*.

THEOREM 1 [AR]. *Given a P.I. algebra  $A$ , there exist  $k, l$  such that for all  $n$  the cocharacter  $\chi_n(A)$  is supported by  $H(k, l)$ .*

Here, the “strip” case  $H(k, 0)$  is characterized by the so-called Capelli identities [R1]. In the above theorem, denote:  $\text{hook}(A) \subseteq H(k, l)$ . If in

addition  $k, l$  are minimal and unique, denote:  $\text{hook}(\mathcal{A}) = H(k, l)$  (also, the notation  $\text{hook}(\chi_n) \subseteq H(k, l)$  is clear).

In Kemer's structure theory for P.I. algebras over a field  $F$  of characteristic zero [K], the basic algebras are three classes of superalgebras, called *verbally prime*:

$$M_r(F), M_r(E), M_{r,s}.$$

Here  $M_r(F)$  are the  $r \times r$  matrices over  $F$ , and  $E$  is the infinite Grassmann (exterior) algebra, while  $M_{r,s}$  is a certain subalgebra of  $M_{r+s}(E)$ . The algebra  $F[t] = F \cdot 1 \oplus F \cdot t$ ,  $t^2 = 1$ , is instrumental in this classification of the verbally prime algebras [K, p. 21].

PROPOSITION. *The following have been proved in [Br1, R2]:*

- (1)  $\text{hook}(M_r(F)) = H(r^2, 0)$  (the "strip" case),
- (2)  $\text{hook}(M_r(E)) = H(r^2, r^2)$ ,
- (3)  $\text{hook}(M_{r,s}) = H(r^2 + s^2, 2rs)$  ((2) and (3) are the "hook" case).

We call the above (1), (2), and (3) "the verbally prime hooks."

The set of the verbally prime hooks is now identified with

$$\mathcal{P} = \{(r^2, r^2), (r^2 + s^2, 2rs) \mid r, s \in \mathbb{N}\}.$$

These are the hook generalizations of the strips  $H(r^2, 0) \equiv (r^2, 0)$  of the matrix algebras. For that reason we call  $\mathcal{P}$  the set of *generalized squares*. In fact, we show below that with a natural multiplication of pairs that is induced by the Kronecker product, these are indeed squares, once the pair  $(1, 1) = \text{hook}(E)$  is admitted as a generalized square.

Two natural operations in the character theory of  $S_n$  [JK] that frequently arise in the study of cocharacters of P.I. algebras are:

- (a) The *outer* product  $\hat{\otimes}$ : if  $\text{hook}(\chi_i) \subseteq H(k_i, l_i)$ ,  $i = 1, 2$ , then

$$\text{hook}(\chi_1 \hat{\otimes} \chi_2) \subseteq H(k_1 + k_2, l_1 + l_2).$$

- (b) The *inner* (Kronecker) product  $\otimes$ : if  $\text{hook}(\chi_i) \subseteq H(k_i, l_i)$ ,  $i = 1, 2$ , then

$$\text{hook}(\chi_1 \otimes \chi_2) \subseteq H(k_1 k_2 + l_1 l_2, k_1 l_2 + k_2 l_1).$$

Identifying  $H(k, l)$  with  $(k, l)$ , we see that the operation  $\hat{\otimes}$  induces the coordinatewise addition

$$(a, b) + (c, d) = (a + c, b + d),$$

while  $\otimes$  induces the following multiplication:

$$(a, b)(c, d) = (ac + bd, ad + bc).$$

This gives a commutative ring structure on  $B = \mathbb{Z} \times \mathbb{Z} = \mathbb{Z}[t]$ , where  $t = (0, 1)$ . We remark that on  $B$  there is a natural conjugation  $\overline{(a, b)} = (a, -b)$  with corresponding “hyperbolic” norm

$$|(a, b)|^2 = (a, b)(a, -b) = (a^2 - b^2, 0) \equiv a^2 - b^2.$$

Therefore  $B$  is also called *the hyperbolic integers*. In a sense, the elements of  $B$  are the (generalized) integers in  $F[t]$ , where  $\mathbb{Z} \subseteq F$  is a field.

Note that  $(r^2 + s^2, 2rs) = (r, s)^2$ , and also  $(r^2, r^2) = (1, 1)(r, 0)^2$ . Thus the elements of  $\mathcal{P}$  are indeed squares, provided we agree to call  $(1, 1) \equiv H(1, 1) = \text{hook}(E)$  a (generalized) square. Note also that  $(1, 1)$  is almost a square: let  $\mu = \frac{1}{2}(1, 1)$ ; then  $\mu^2 = \mu$ . Trivially, if  $(r, s) \in B$  is a sum of elements of  $\mathcal{P}$  then  $r \geq s \geq 0$ .

**THEOREM 2 [GZ].** *Given any P.I. algebra  $A$  in characteristic zero, with the corresponding sequence of codimensions  $c_n(A)$ , the limit*

$$\exp(A) := \lim_{n \rightarrow \infty} (c_n(A))^{1/n}$$

*exists and is always an integer. It is called the exponent of  $A$ .*

Given  $A$ , it is natural to wish to compute the invariant  $\exp(A)$ . For certain polynomials  $C_{k+1}[x; y]$ , called Capelli polynomials,  $A$  satisfies  $C_{k+1}[x; y] = 0$  if and only if  $\chi_n(A) \subseteq H(k, 0)$  for all  $n$  [R1]. For these Capelli identities (i.e., the “strip” case),  $\exp(A)$  is calculated in [MRZ]. Denote by  $U_k$  the universal algebra of the  $k$ th Capelli identity. We have

$$0 \leq k - \exp(U_k) \leq 3.$$

The proof utilizes a construction involving the verbally prime algebras, which in this “strip” case are just the matrix algebras  $M_r(F)$ . Consequently it depends on expressing integers as sums of squares  $r^2 (\equiv (r^2, 0))$ . The *four square theorem* from number theory implies the above upper bound 3 (which cannot be improved).

In the  $(k, l)$  hook case there are analogous “Amitsur–Capelli” type polynomials  $C_{k+1, l+1}[x; y] = 0$  [AR], with the corresponding universal algebra  $U_{k, l}$  replacing  $U_k$ . The computation of  $\exp(A)$  is done in [BR2], where it is shown that when  $k \geq l \geq 0$ ,

$$0 \leq (k + l) - \exp(U_{k, l}) \leq 4.$$

In this case, the proof depends on expressing elements  $(k, l) \in B$ ,  $k \geq l \geq 0$ , as sums of generalized squares. For these generalized squares, a seven (generalized) squares theorem was proved, and a *six (generalized) squares theorem* is conjectured in [BR2]. As a consequence of the proof of this conjecture given in this paper we have the sharp bound

$$0 \leq (k + l) - \exp(U_{k,l}) \leq 3.$$

We express our special thanks to Umberto Zannier for his interest in our work, especially in connection with the remarks of Section 3.

## 2. PROOF OF THE SIX GENERALIZED SQUARES THEOREM

In this section we prove the six generalized squares theorem by a series of lemmata. We shall often use the result of Legendre [Di, Vol. II, p. 261], which states that a number is a sum of three squares with no common factor unless it is of the form  $4^u(8l + 7)$ ,  $u, l \in \mathbb{N}$ , in which case it is a sum of four squares. We also use a number of corollaries to this result, like the fact that every odd integer can be written in the form  $x^2 + y^2 + 2z^2$ . A full treatment of such results can be found in [Di] and when our statements are not found in that reference we give details.

**DEFINITION.** Call any of the following presentations of  $s \in \mathbb{N}$  a quadratic-ternary presentation:

- (1)  $s = \varepsilon x^2 + \eta y^2 + \gamma z^2$ ,  $\varepsilon, \eta, \gamma \in \{0, 1, 2\}$ ,
- (2)  $s = \varepsilon x^2 + \eta y^2 + 2z(z + l)$ ,  $\varepsilon, \eta \in \{0, 1, 2\}$ ,
- (3)  $s = \varepsilon x^2 + 2y(y + k) + 2z(z + l)$ ,  $\varepsilon \in \{0, 1, 2\}$ .

We define the length of such a presentation to be the number of its non-zero summands and its shift to be 0 in case (1),  $l^2$  in case (2), and  $k^2 + l^2$  in case (3).

**LEMMA 1.** *In  $\mathbb{N}$ , let  $r \geq s \geq 0$ . Assume  $s$  has a quadratic-ternary presentation of length  $\leq 2$  with shift  $\leq r - s$ . Then  $(r, s) \in B$  can be written as a sum of six or fewer elements in  $\mathcal{P}$ ; that is,  $(r, s)$  is a sum of six generalized squares.*

*Similarly, if  $s$  admits any quadratic-ternary presentation with shift  $= d \leq r - s$ , such that  $r - s - d$  is a sum of three squares in  $\mathbb{N}$ , then  $(r, s)$  is a sum of six generalized squares.*

*Proof.* The proof follows from the four squares theorem in  $\mathbb{N}$ , applied to  $r - s - \text{shift} \geq 0$ . Here are the cases.

(1)  $s = \varepsilon x^2 + \eta y^2$ ,  $\varepsilon, \eta \in \{0, 1, 2\}$  (hence  $\text{shift} = 0$ ). In  $\mathbb{N}$  let  $r - s = q_1^2 + \dots + q_4^2$ ; then

$$(r, s) = (\varepsilon x^2, \varepsilon x^2) + (\eta y^2, \eta y^2) + (q_1^2, 0) + \dots + (q_4^2, 0),$$

a sum of six terms from  $\mathcal{S}$ .

(2)  $s = \varepsilon x^2 + 2y(y + k)$ ,  $\varepsilon \in \{0, 1, 2\}$ . Let  $r - s - \text{shift} = r - s - k^2 = q_1^2 + \dots + q_4^2$ ; then

$$(r, s) = (\varepsilon x^2, \varepsilon x^2) + (y^2 + (y + k)^2, 2y(y + k)) \\ + (q_1^2, 0) + \dots + (q_4^2, 0)$$

(since  $y^2 + (y + k)^2 = 2y(y + k) + k^2$ ).

(3)  $s = 2y(y + k) + 2z(z + l)$ . Now  $r - s - \text{shift} = r - s - (k^2 + l^2) = q_1^2 + \dots + q_4^2$ , and

$$(r, s) = (y^2 + (y + k)^2, 2y(y + k)) + (z^2 + (z + l)^2, 2z(z + l)) \\ + (q_1^2, 0) + \dots + (q_4^2, 0).$$

LEMMA 2. Let  $5 \leq M \in \mathbb{N}$  and assume that none of  $M$ ,  $M - 1$ , and  $M - 4$  is a sum of three squares.

Then  $M = 4^u(8l + 7)$ , with  $0 \leq l$  and  $3 \leq u$ . In particular,  $448 = 4^{37} \leq M$ .

*Proof.* By assumption,  $M = 4^u(8l + 7) = (M - 1) + 1 = 4^u(8k + 7) + 1$ . If  $1 \leq v$  then  $u = 0$  so l.h.s.  $\equiv 3 \pmod{4}$  while r.h.s.  $\equiv 1 \pmod{4}$ , a contradiction. Thus  $v = 0$ , so  $M = 4^u(8l + 7) = 8(k + 1)$ , implying  $2 \leq u$ .

Similarly  $M = (M - 4) + 4 = 4^w(8m + 7) + 4$ , so  $4^u(8l + 7) = 4^w(8m + 7) + 4$ ; hence  $1 \leq w$ . This implies  $4^{u-1}(8l + 7) = 4^{w-1}(8m + 7) + 1$  which implies that  $w - 1 = 0$  since  $2 \leq u$ . Deduce that  $4^{u-1}(8l + 7) = 8(m + 1)$ , which implies that  $2 \leq u - 1$ .

LEMMA 3. As in Lemma 2, let  $5 \leq M \in \mathbb{N}$ , and assume that none of  $M$ ,  $M - 1$  and  $M - 4$  is a sum of three squares. Let  $t \in \mathbb{N}$  be an integer that modulo 8 is congruent to either 2 or 5. Then  $M - t$  is a sum of three squares.

*Proof.* Assume this is not the case. With the above notations, deduce that  $M = 4^u(8l + 7) = (M - t) + t = 4^q(8r + 7) + t$ ,  $0 \leq l$ ,  $3 \leq u$ . If  $q = 0$ , deduce that  $t \equiv 1 \pmod{8}$ , a contradiction. If  $q = 1$ , then 4 divides  $t$ , so  $\pmod{8}$ ,  $t$  is congruent to either 0 or 4, a contradiction. Finally, if  $2 \leq q$ , deduce that  $t \equiv 0 \pmod{8}$ , again a contradiction, and the proof follows.

LEMMA 4. *As in Lemma 2, let  $9 \leq M$  be an integer and assume that none of  $M$ ,  $M - 1$ , and  $M - 4$  is a sum of three squares. Then  $M - 8$  is a sum of three squares.*

*Proof.* Assume the contrary. Deduce that  $M$  is of the form

$$M = 4^u(8l + 7) = (M - 8) + 8 = 4^y(8s + 7) + 8$$

for certain integers  $u \geq 3$ ,  $l \geq 0$ ,  $y \geq 0$ ,  $s \geq 0$ . Clearly, by parity,  $y \geq 1$ , so that on dividing both sides of the above equation by 4 we have

$$4^{u-1}(8l + 7) = 4^{y-1}(8s + 7) + 2.$$

Again by parity, as  $u \geq 3$ , we must have  $y \geq 2$ . Reducing the above equation modulo 4 gives a contradiction and the proof follows.

LEMMA 5. *Every  $s \in \mathbb{N}$  can be represented as  $s = x^2 + y^2 + \varepsilon z^2$ ,  $x, y, z \in \mathbb{N}$ ,  $\varepsilon \in \{0, 1, 2\}$ .*

*Proof.* Every odd integer  $s$  can be written in the form

$$s = x^2 + y^2 + 2z^2.$$

Suppose now that  $s$  is even. If  $s \equiv 2 \pmod{4}$ , then  $s$  is a sum of three squares. If  $s \equiv 0 \pmod{4}$  then either  $s$  is a sum of three squares or it is of the form  $s = 4^u(8l + 7)$  with  $u \geq 1$ , so that we can write it as

$$s = 2 \cdot 4^{u-1}(8l + 7) + 2 \cdot 4^{u-1}(8l + 7).$$

Clearly  $2 \cdot 4^{u-1}(8l + 7)$  can be written as a sum of three squares  $x^2 + y^2 + z^2$  so that

$$s = 2x^2 + 2y^2 + 2z^2 = (x + y)^2 + (x - y)^2 + 2z^2.$$

This completes the proof.

LEMMA 6. *Every  $s \in \mathbb{N}$  can be represented as  $s = \varepsilon x^2 + y^2 + 2z(z + 1)$ , where  $\varepsilon \in \{0, 1, 2\}$ ,  $x, y, z \in \mathbb{N}$ .*

*Proof.* We consider first the case where  $s$  is a positive even integer. Now, every odd integer is the sum of four squares of which two are consecutive [Di, Vol. II, p. 293]. Hence,

$$s + 1 = p^2 + q^2 + z^2 + (z + 1)^2$$

and so

$$s = p^2 + q^2 + 2z(z + 1).$$

This deals with the case where  $s$  is even.

If  $s$  is odd, then  $s$  is congruent to either 1 or 3 mod 4. Now suppose that  $s$  is congruent to 1 mod 4, and write  $s = 4k + 1$ . We know that for any integer  $k \geq 0$  the integer  $8k + 3$  is a sum of three squares:

$$8k + 3 = a^2 + b^2 + c^2.$$

Reducing mod 4 we see that  $a, b, c$  are all odd. Therefore we may rewrite

$$8k + 4 = 1 + a^2 + b^2 + c^2$$

as

$$8k + 4 = \frac{1}{2}((a-1)^2 + (a+1)^2 + (b-c)^2 + (b+c)^2).$$

Thus,

$$4k + 2 = \left(\frac{a-1}{2}\right)^2 + \left(\frac{a+1}{2}\right)^2 + \left(\frac{b-c}{2}\right)^2 + \left(\frac{b+c}{2}\right)^2$$

and

$$s = 4k + 1 = 2\left(\frac{a-1}{2}\right)\left(\frac{a+1}{2}\right) + \left(\frac{b-c}{2}\right)^2 + \left(\frac{b+c}{2}\right)^2.$$

If  $s$  is congruent to 3 mod 4, we write  $s = 4k + 3$  and argue in a similar way as above but with the representation, for any integer  $k \geq 0$ , of  $8k + 7$  as

$$8k + 7 = a^2 + b^2 + 2c^2,$$

where we have  $a \geq 1$  odd and  $b \geq 0$  even. We deduce from this that

$$s = 4k + 3 = 2\left(\frac{a-1}{2}\right)\left(\frac{a+1}{2}\right) + 2\left(\frac{b}{2}\right)^2 + c^2.$$

**LEMMA 7.** *Every  $s \in \mathbb{N}$  can be represented as  $s = \varepsilon x^2 + y^2 + 2z(z+2)$ , where  $\varepsilon \in \{0, 1, 2\}$ ,  $x, y, z \in \mathbb{N}$ .*

*Proof.* Assume first that  $s$  is odd. Then, we may write

$$s + 2 = a^2 + b^2 + 2c^2.$$

Clearly, if  $c \neq 0$ , then

$$s = a^2 + b^2 + 2(c-1)(c+1)$$

and we are done. So assume that  $c = 0$ . If  $a, b \neq 0$  then we can suppose that  $a$  is even, as  $a$  and  $b$  have opposite parity. Write  $a = 2k$ ,  $k \geq 1$ . Then

$$s + 2 = (2k)^2 + b^2 = 2k^2 + 2k^2 + b^2$$

so that

$$s = b^2 + 2k^2 + 2(k-1)(k+1),$$

as required. We can use previous arguments to see that we can suppose  $ab \neq 0$ . Alternatively, argue as follows. Suppose that

$$s + 2 = b^2$$

with  $b$  odd. Then  $b^2$  is congruent to  $1 \pmod{8}$  and  $2b^2$  is congruent to  $2 \pmod{8}$  and so is not of the form  $8n + 7$  or  $4n$ . We can therefore represent  $2b^2$  as a sum of three squares with no common factor,

$$2b^2 = x^2 + y^2 + z^2.$$

We may suppose that  $x, y$  are odd and that  $z$  is even. If  $z \neq 0$  we write

$$p = \frac{1}{2}(x+y), \quad q = \frac{1}{2}(x-y), \quad r = \frac{1}{2}z$$

so that

$$b^2 = p^2 + q^2 + 2r^2, \quad r \geq 1$$

and we conclude that

$$s = b^2 - 2 = p^2 + q^2 + 2(r-1)(r+1).$$

If  $z = 0$ , then  $x$  and  $y$  are coprime so that we may assume  $pq \neq 0$  and

$$b^2 = p^2 + q^2,$$

with  $p = 2k$  even and  $q$  odd. This enables us to write

$$b^2 = 2k^2 + 2k^2 + q^2$$

and

$$s = b^2 - 2 = q^2 + 2k^2 + 2(k-1)(k+1).$$

Now suppose that  $s$  is even. Consider first the case where  $s + 2 = 2^{2e+1}v$ , some  $e \geq 0$ , and  $v$  odd. As  $v$  is odd, it is of the form

$$v = p^2 + q^2 + 2r^2,$$

where one of  $p$  or  $q$  is non-zero, say  $p \neq 0$ . Therefore

$$2^{2e}v = (2^e p)^2 + (2^e q)^2 + 2(2^e r)^2 = A^2 + B^2 + 2C^2$$

for  $A = 2^e p$ ,  $B = 2^e q$ ,  $C = 2^e r$ , and

$$s + 2 = 2 \cdot 2^{2e} v = 2A^2 + 2B^2 + (2C)^2,$$

with  $A \neq 0$  as  $p \neq 0$ . Therefore,

$$s = 2(A - 1)(A + 1) + 2B^2 + (2C)^2.$$

Now, consider the case where  $s + 2 = 2^{2e} v$ , some  $e \geq 1$  and  $v$  odd. We write again

$$v = p^2 + q^2 + 2r^2.$$

We have

$$s + 2 = (2^e p)^2 + (2^e q)^2 + 2(2^e r)^2$$

and if  $r \neq 0$  then

$$s = (2^e p)^2 + (2^e q)^2 + 2(2^e r - 1)(2^e r + 1).$$

If  $r = 0$ , then  $v = p^2 + q^2$  and we can suppose that  $p \neq 0$ . Then

$$s + 2 = (2^e p)^2 + (2^e q)^2$$

and as  $e \geq 1$ , with  $A = 2^{e-1} p \neq 0$  and  $B = 2^e q$ , we can write

$$s + 2 = 2A^2 + 2A^2 + B^2$$

and hence

$$s = 2(A - 1)(A + 1) + 2A^2 + B^2,$$

as required.

**COROLLARY 1.** *Let  $r \geq s \geq 0$  and  $r - s < 448$ . Then  $(r, s)$  is a sum of six hyperbolic squares.*

*Proof.* Let  $M = r - s$ . If  $M \leq 4$  then  $M$  is a sum of three squares in  $\mathbb{N}$ . If  $5 \leq M$  then by Lemma 2, either  $M$  or  $M - 1$  or  $M - 4$  is a sum of three squares in  $\mathbb{N}$ . If  $M$  is a sum of three squares, apply Lemma 5; if  $M - 1$  is a sum of three squares, apply Lemma 6; if  $M - 4$  is a sum of three squares, apply Lemma 7.

For example, assume  $M - 4 = q_1^2 + q_2^2 + q_3^2$ . By Lemma 7, we have  $s = \varepsilon x^2 + y^2 + 2z(z + 2)$ ,  $\varepsilon \in \{0, 1, 2\}$ , and we have

$$\begin{aligned} (r, s) &= (\varepsilon x^2, \varepsilon x^2) + (y^2, y^2) + (z^2 + (z + 2)^2, 2z(z + 2)) \\ &\quad + (q_1^2, 0) + (q_2^2, 0) + (q_3^2, 0). \end{aligned}$$

*Remark.* Lemma 8 below will show that most  $s \in \mathbb{N}$  admit the presentation

$$(2.1) \quad s = \varepsilon x^2 + 2y(y + 1) + 2z(z + 2), \quad \varepsilon \in \{0, 1, 2\}.$$

We comment in Section 3 on the difficulty of establishing this fact for all integers  $\neq 3, 23$ . We have

**COROLLARY 2.** *Let  $s \in \mathbb{N}$  admit the presentation (2.1), and let  $r \geq s \geq 0$ ; then  $(r, s)$  is a sum of six hyperbolic squares.*

*Proof.* This follows by similar arguments to those of Corollary 1, using Lemmas 2 and 3.

**LEMMA 8.** *Every  $s \in \mathbb{N}$  for which  $2s + 5$  is not of the form  $a^2 + 2c^2$ , with  $a$  odd and  $c$  odd, can be written in the form  $s = \varepsilon x^2 + 2y(y + 1) + 2z(z + 2)$ ,  $\varepsilon \in \{0, 1, 2\}$ ,  $x, y, z \in \mathbb{N}$ .*

*Proof.* Write

$$2s + 5 = a^2 + b^2 + 2c^2,$$

with  $a \geq 1$  odd and  $b \geq 0$  even.

*Step 1.* We show we can assume that either  $b > 0$  or  $c > 0$ .

This is because we can represent  $2(2s + 5)$  as a sum of three squares with no common factor:

$$2(2s + 5) = x^2 + y^2 + z^2.$$

We can assume that  $z$  is even and that  $x$  and  $y$  are odd,  $x \geq y$ . Substituting

$$a = \frac{1}{2}(x + y), \quad b = \frac{1}{2}(x - y), \quad c = \frac{1}{2}z,$$

we recover

$$2s + 5 = a^2 + b^2 + 2c^2.$$

We can assume that  $a$  is odd and non-zero and  $b$  is even. If  $b = c = 0$  this would mean that  $z = 0$  and  $x = y$ , contradicting the fact that  $x, y$  and  $z$  have no common factor.

*Step 2.* Assume  $b \neq 0$ , so  $b \geq 2$  since it is even. From

$$2s + 5 = a^2 + b^2 + 2c^2$$

deduce that

$$2s + 6 = 1 + a^2 + b^2 + 2c^2 = \frac{1}{2}((a - 1)^2 + (a + 1)^2 + 2b^2 + (2c)^2)$$

and so

$$s + 3 = \left(\frac{a-1}{2}\right)^2 + \left(\frac{a+1}{2}\right)^2 + 2\left(\frac{b}{2}\right)^2 + c^2.$$

We conclude that

$$s = 2\left(\frac{a-1}{2}\right)\left(\frac{a+1}{2}\right) + 2\left(\frac{b}{2} - 1\right)\left(\frac{b}{2} + 1\right) + c^2,$$

as required.

*Step 3.* In this case  $b = 0$ . Then  $c > 0$  and is even by the assumptions. So let

$$2s + 5 = a^2 + 2c^2,$$

with  $a \geq 1$  odd and  $c \geq 2$  even (thus  $s$  is congruent to 2 modulo 4). We then have

$$2s + 6 = \frac{1}{2}((a-1)^2 + (a+1)^2 + 4c^2).$$

It follows that

$$s + 3 = \left(\frac{a-1}{2}\right)^2 + \left(\frac{a+1}{2}\right)^2 + 4\left(\frac{c}{2}\right)^2.$$

We deduce that

$$s = 2\left(\frac{a-1}{2}\right)\left(\frac{a+1}{2}\right) + 2\left(\frac{c}{2} - 1\right)\left(\frac{c}{2} + 1\right) + 2\left(\frac{c}{2}\right)^2,$$

as required.

This completes the proof. Note that in the remaining cases where

$$2s + 5 = a^2 + 2c^2,$$

with  $a, c \geq 1$  odd,  $s$  is congruent to 3 modulo 4.

*Remark.* Notice that the integers  $s = 3, 23$  cannot be written in the form  $\varepsilon x^2 + 2y(y+1) + 2z(z+2)$ ,  $\varepsilon \in \{0, 1, 2\}$ .

It follows from Corollary 2 that for  $r \geq s \geq 0$ , if  $s$  satisfies the assumptions of Lemma 8, then  $(r, s)$  is a sum of six elements in  $\mathcal{P}$ .

We now complete the proof of the six generalized squares theorem.

By Corollary 2 and Lemma 8, it remains to prove this theorem in the case where  $2s + 5 = a^2 + 2c^2$ , where both  $a, c \geq 1$  are odd. By Corollary 1, we can assume that  $r - s \geq 448$ . There remain a number of cases to treat.

*Case 1.* Assume  $a \geq 7$ . Then

$$2s - 44 = a^2 - 49 + 2c^2 = (a - 7)(a + 7) + 2c^2,$$

so that on dividing by 2 we have

$$s - 22 = 2\left(\frac{a - 7}{2}\right)\left(\frac{a + 7}{2}\right) + c^2,$$

and finally we deduce the representation

$$s = 2 \cdot 1 \cdot 11 + 2\left(\frac{a - 7}{2}\right)\left(\frac{a + 7}{2}\right) + c^2.$$

This representation has the shift  $10^2 + 7^2 = 149$  which is congruent to 5 modulo 8 and is less than 448. This enables us to appeal to Lemmas 3 and 1.

Notice that this gives the representation

$$23 = 2 \cdot 1 \cdot 11 + 1,$$

using

$$51 = 7^2 + 2 \cdot 1^2.$$

*Case 2.* Assume  $a = 5$ . Then

$$2s + 5 = 25 + 2c^2 = 3^2 + 4^2 + 2c^2$$

and we have

$$s = 2 \cdot 1 \cdot 2 + 2 \cdot 1 \cdot 3 + c^2,$$

giving rise to a shift of 5; hence we are done by Lemmas 3 and 1.

*Case 3.* Assume  $a = 3$ . Then

$$2s + 5 = 9 + 2c^2,$$

so that

$$2s = 4 + 2c^2$$

and

$$s = 2 \cdot 1 + c^2,$$

which is a representation of length 2 and hence suffices—by Lemma 1.

For example, when  $s = 3$  we have  $c = 1$ .

*Case 4.* Assume  $a = 1$ . Then

$$2s + 5 = 1 + 2c^2.$$

As  $s \geq 1$  we may assume that  $c \geq 3$ . We have

$$2s + 4 = 2c^2$$

and

$$s = c^2 - 2.$$

Recalling that  $c$  is odd, we may write

$$s = c^2 - 2 = 1 + 2\left(\frac{c-1}{2}\right)\left(\frac{c+3}{2}\right) + 2\left(\frac{c+1}{2}\right)\left(\frac{c-3}{2}\right).$$

This is a representation of the form

$$s = 1 + 2y(y+2) + 2z(z+2),$$

which gives rise to a shift of 8, which will enable us to appeal to Lemmas 4 and 1.

This completes the proof of the six generalized squares theorem.

### 3. CONCLUDING REMARKS

During the 19th and earth 20th centuries, various mathematicians obtained results containing additional information about the four squares theorem for natural numbers (see [Di, Vols. II and III]), and this motivated our present treatment of the six generalized squares theorem which relies on some of these very old classical observations.

The discussion of Section 2 makes it clear that we would have a shorter proof of the six generalized squares theorem if we knew that Lemma 8 held for all integers except an explicit finite list. A computer search leads us to ask the following.

QUESTION. *Is it true that every integer  $s \neq 3, 23$  can be written in the form*

$$s = \varepsilon x^2 + 2y(y+1) + 2z(z+2), \quad \varepsilon \in \{0, 1, 2\}?$$

By the treatment of Section 2, it is clear that the answer to the question is “yes” given that we can settle it for integers  $s$  congruent to 3 mod 4, of which  $s = 3, 23$  are special cases. In Section 2 we reduced this question to the study of representations of  $2s + 5$  in the form  $a^2 + b^2 + 2c^2$  with  $a$

and  $c$  odd (we can assume  $a, c \geq 1$ ) and  $b \geq 0$  even. We relied on being able to take  $b \neq 0$ : for example for  $s = 23 \cdot 2 \cdot 23 + 5 = 51 = 1^2 + 2 \cdot 5^2 = 7^2 + 2 \cdot 1^2$  are the only representations of this form for 51 and both have  $b = 0$ . As recalled at the beginning of Section 2, the representation of an odd integer  $t$  in the form  $a^2 + b^2 + 2c^2$  follows from the representation of  $2t$  as a sum of three squares  $x^2 + y^2 + z^2$ , with  $x, y$  odd and  $z$  even. We have  $b = 0$  precisely when  $x = y$ . Therefore, we are faced with comparing the number of representations of  $2t$  as  $x^2 + y^2 + z^2$  with  $x \neq y$ , to the number of such representations with  $x = y$ . We have the following result [Di, Vol. II, p. 262].

PROPOSITION 1. *The number  $\varphi(m)$  of proper representations  $x^2 + y^2 + z^2$  of  $m$  as a sum of three squares with no common factor is given by*

$$\begin{aligned}\varphi(m) &= 12 \cdot 2^\mu \cdot h'(-4m), & m &\equiv 1, 2, 5, 6 \pmod{8} \\ &= 2^{\mu+2} \cdot h'(-4m), & m &\equiv 3 \pmod{8}.\end{aligned}$$

Here  $h'(-4m)$  is the number of classes in the principal genus of the properly primitive binary quadratic forms of discriminant  $-4m$  and  $\mu$  is the number of distinct odd prime factors of  $m$  (see [Ca] for definitions).

On the other hand we have [Di, Vol. III, p. 38]

PROPOSITION 2. *Let  $m$  be a natural number which we factorize into its primes as  $m = 2^\alpha \prod_p p^r \prod_q q^s$ , where the primes  $p$  are those congruent to 1 or 3 mod 8 and the primes  $q$  are those congruent to 5 or 7 mod 8. The number  $\psi(m)$  of proper representations  $z^2 + 2x^2$  of  $m$ , where  $x$  and  $z$  have no common factor, is given by  $\psi(m) = 2 \prod_p (r+1) \prod_q ((1 + (-1)^s)/2)$ .*

Consider the case where  $m = 2t$  and  $t$  is a product of  $\mu$  distinct prime factors  $p$  congruent to either 1 or 3 mod 8. Then, by Proposition 1 we have  $\varphi(2t) = 12 \cdot 2^\mu \cdot h'(-8t)$  and by Proposition 2 we have  $\psi(2t) = 2 \cdot 2^\mu$ . The number of automorphs [Ca, Chap. 9, p. 127] of the form  $x^2 + y^2 + z^2$  is 6 times the number of automorphs of  $z^2 + 2x^2$ . On the other hand, for the  $t$  under consideration we have  $\varphi(2t)/\psi(2t) = 6 \cdot h'(-8t)$  and therefore, taking into account the automorphs, there is a representation  $2t = x^2 + y^2 + z^2$  with  $x \neq y$  precisely when  $h'(-8t) > 1$ . It is an open problem to list explicitly those  $m$  for which  $h'(-4m) = 1$ , although it is known that there are 65 such  $m$  with at most 1 extra  $m$  about which there is effectively no information. These  $m$  are the *numeri idonei* of Euler (also called “convenient” or “suitable” numbers, listed, for example, in [BoSh, p. 427]). From the Riemann hypothesis, it would follow that the list of the 65 *numeri idonei* of Euler is complete. Inspection of Euler’s list shows that  $m = 22, 102$  are its only entries which are of the form  $m = 2(2s + 5)$ , with  $2s + 5$  a product of distinct prime factors congruent to either 1 or 3 mod 8,

so this checks with the fact  $m = 2(2s + 5) = 22, 102$  cannot be written as a sum of three distinct squares and hence the cases  $s = 3, 23$  cannot be treated by the arguments of Lemma 8. As we can verify the question above by computer for all numbers  $s \neq 3, 23$  well beyond those of the *numeri idonei* list, the only obstacle to answering the question in the affirmative is the knowledge that this list is complete.

## REFERENCES

- [AR] S. A. Amitsur and A. Regev, P.I. algebras and their cocharacters, *J. Algebra* **78** (1982), 248–254.
- [BR1] A. Berele and A. Regev, On the codimensions of the verbally prime P.I. algebras, *Israel J. Math.* **91** (1995), 239–247.
- [BR2] A. Berele and A. Regev, Exponential growth of some P.I. algebras, *J. Algebra*, to appear.
- [BoSh] Z. I. Borevich and I. R. Shafarevich, “Number Theory,” Academic Press, New York, 1966.
- [Ca] J. W. S. Cassels, “Rational Quadratic Forms,” Academic Press, San Diego, 1978.
- [Di] L. E. Dickson, “History of the Theory of Numbers,” Vols. I–III, Chelsea, New York, 1992 (original publication 1874).
- [GZ] A. Giambruno and M. V. Zaicev, Exponential codimension growth of PI algebras: An exact estimate, *Adv. Math.* **142** (1999), 221–243.
- [JK] G. James and A. Kerber, “The Representation Theory of the Symmetric Group,” Encyclopedia of Mathematics, Vol. 16, Addison–Wesley, Reading, MA, 1981.
- [K] A. R. Kemer, “Ideals of Identities of Associative Algebras,” Am. Math. Soc. of Mathematical Monographs, Vol. 87, Am. Math. Soc., Providence, 1991.
- [MRZ] S. Mishchenko, A. Regev, and M. Zaicev, The exponential growth of codimensions for Capelli identities, *Israel J. Math.*, to appear.
- [R1] A. Regev, Algebras satisfying Capelli identities, *Israel J. Math.* **33** (1979), 149–154.
- [R2] A. Regev, On the identities of subalgebras of matrices over the Grassmann algebra, *Israel J. Math.* **58** (1987), 351–369.