The Diophantine equations
$$x_1^n + x_2^n + \ldots + x_{r_1}^n = y_1^n + y_2^n + \ldots + y_{r_2}^n$$

Michael A. Ivanov

Physics Dept.,

Belarus State University of Informatics and Radioelectronics,

6 P. Brovka Street, BY 220027, Minsk, Republic of Belarus.

E-mail: ivanovma@gw.bsuir.unibel.by.

Abstract

The aim of this paper is to prove the possibility of linearization of such equations by means of introduction of new variables. For $n = 2$ such a procedure is well known, when new variables are components of spinors and they are widely used in mathematical physics. For example, parametrization of Pythagoras threes $a^2 + b^2$, $a^2 - b^2$, $2ab$ may be cited as an example in number theory where two independent variables form a spinor which can be obtained by solution of a system of two linear equations.

We also investigate the combinatorial estimate for the smallest sum $r(n) = r_1 + r_2 - 1$ for solvable equations of such a type as $r(n) \leqslant 2n + 1$ (recently the better one with $r(n) \leqslant 2n - 1$ was received by L. Habsieger (J. of Number Theory 45 (1993) 92)). Apart from that we consider two conjectures about $r(n)$ and particular solutions for $n \leqslant 11$ which were found with the help of the algorithm that is not connected with linearization.

1   Introduction

A great interest was displayed to the particular case of such equations - $x^n + y^n = z^n$ [1, 2], and it seems rather strange, that the general case of

equality of sums of the same powers

$$\sum_1^{r_1} x_i^n = \sum_1^{r_2} y_j^n \qquad (1)$$

was left in the shadow, although for large $r_1 + r_2$ such equations are solvable and for small $r_1 + r_2$ a set of assertions of the type of Fermat's last theorem can be formulated for them: $r(n) > R(n) \mid n > n_0$, where $R(n)$ and n_0 are fixed, $r(n)$ is the smallest sum $r_1 + r_2 - 1$ for which Eqs.(1) have at least a non-trivial solution.

A geometrical approach [3] was used by G. Faltings to prove Mordell's conjecture [4], that gave a new push to the investigations of Fermat's equations. On the other hand, an interesting algebraic fact, concerning (1), was recently discovered : they can be linearized by introducing new variables [5, 6, 7, 8] with the help of which assumed solutions of Eqs.(1) can be parametrized. A different approach to linearization is described in [9].

In this paper we give a constructive proof of the theorem about the possibility of linearization of (1) (when n is a prime integer), the technique of which consists in adding a system of linear equations in which some new variables enter (see [7]):

Theorem 1 *A. If we take the equation*

$$\sum_1^k x_i^n = y^n, k \geqslant 2, \qquad (2)$$

a linear matrix equation

$$\sum_1^k x_i A_i \Psi = y E \Psi, \qquad (3)$$

can be juxtaposed with it, where A_i are the square matrices of order mn, for which the following condition is necessary

$$\forall\{x_i\}: \quad \left(\sum_1^k x_i A_i\right)^n = E \sum_1^k x_1^n, \qquad (4)$$

where E is the unit matrix, Ψ is the column-vector of new variables $(\Psi_1, \ldots, \Psi_{mn})^T$. B. If (2) is fulfilled, the determinant of the system (3) with respect to unknowns Ψ_i is equal to zero.

2

The same matrices, which are built to linearize (2), permit to linearize (1), juxtaposing the equation

$$\sum_{1}^{r_1} x_i A_i \Psi = \sum_{1}^{r_2} y_j B_j \Psi, \tag{5}$$

with (1), where $B_j = \varepsilon_j A_{r_1+j}$, $\varepsilon_j^n = -1$, $r_1 + r_2 = k$.

For composite powers n the linearization of (1) can be carried out step by step for all prime factors of n using Theorem 1 (see also [8]).

The procedure of linearization is generalization of the algebraic part of the Dirac procedure [10] to the equations

$$\varepsilon = \sum_{m=1}^{N} (p_{m1}^n + p_{m2}^n + \ldots + p_{mr}^n)^{1/l},$$

and apart from that the particular case of $N = 1$, $l = n$ is considered in Theorem 1 [7]. Special case of $N = l = n = 2$, $r = 4$, concerns the physical models of composite particles [11, 12] . A more complicated case of it is considered in [8].

In this paper, we also consider the parametrization of solutions of (1) for $n = 3$. Furthermore we add some remarks concerning solvability of Eqs.(1) and some estimates of $r(n)$ for small n.

The matrices A_i , which will be built below, are the permutation matrices with the elements ζ^m, where ζ is a primitive root of degree n of unity.

2 An example of linearization, definitions, proof of lemmas, auxiliary and main theorems

Let us consider the equation $x_1^2 + x_2^2 = y^2$ as the simplest example and juxtapose the linear matrix equation $x_1 A_1 \Psi + x_2 A_2 \Psi = y E \Psi$ with it, where

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 1 \\ 1 & -0 \end{pmatrix}$$

and $\Psi = (\Psi_1, \ \Psi_2)^T$. The solutions of this linear equation have the form $x_i = (y/\delta)\delta_i$, $\delta_1 = \Psi_1^2 - \Psi_2^2$, $\delta_2 = 2\Psi_1^2\Psi_2^2$, $\delta = \Psi_1^2 + \Psi_2^2$. If we suppose that $y = \delta$, then one gets all Pythagoras threes for integers Ψ_i .

Definition 1 *Such a combination as*

$$(ABC\ldots)_+ \equiv \sum_P ABC\ldots, \tag{6}$$

where the summation has been made over all different permutations of $A, B, C, \ldots$, is the generalized anticommutator of matrices $A, B, C, \ldots$.

For instance: $(AB)_+ = AB + BA$, but $(A)^2{}_+ = A^2$,

$$(A^k BC\ldots)_+ = (1/k!)(A_1\ldots A_k BC\ldots)|_{A_i=A}. \tag{7}$$

Definition 2 *A set of k matrices A_i for some n is a concerted one, if*

$$\forall\{n_1,\ldots,n_k|\ 0 \leqslant n_i < n, n_1 + n_2 + \ldots + n_k = n\}:$$
$$(A_1^{n_1} A_2^{n_2}\ldots A_k^{n_k})_+ = 0. \tag{8}$$

Suggestion 1 *If k matrices A_i are concerted and $A_i^n = E$, the condition (4) is fulfilled.*

Definition 3 *A set of positions of non-zero elements is a structure of a permutation matrix.*

To prove the part A of Theorem 1, we need two lemmas:

Lemma 1 *There exists a concerted pair of $n \times n$ matrices.*

Lemma 2 *$m + 1$ concerted matrices can be built from m ones, $m > 1$.*

The order of matrices A_i can be reduced thanks to Lemma 3:

Lemma 3 *There exist three concerted $n \times n$ matrices.*

(There are only three matrices of this kind [8]).

Lemma 2 is proved by the author and Lemmas 1 and 3 are proved for the prime $n > 2$ (see [8] for a more general case).

Proof of lemmas. Let $S_n \equiv \{0, 1, \ldots, n-1\}$ be the full system of residue classes modulo n. Then we can speak about the natural numbers of some set S that they are distributed uniformly over the classes of S_n, if the same number of elements of S belongs to every class of S_n. Let $S_{nj}^m \subset S_n$ be some subset of S_n, which includes m elements, with a number j, where $j = 1, C_n^m$ (C_n^m is the combination number), then $S_{nj}^m \neq S_{ni}^m|\ i \neq j$; let us denote $b_j^m \equiv \sum_1^m k_i,\ k_i \in S_{nj}^m$. To prove Lemmas 1 and 3 for prime $n > 2$, we shall use

4

Theorem 2 *For every m, $1 \leqslant m < n$, the numbers b_j^m are distributed uniformly over the classes of S_n.*

Proof for the prime $n > 2$ (by induction).

1. The theorem is evident for $m = 1$.

2. Let for $\forall m'$, $m' \leqslant m - 1$, the numbers $b_j^{m'}$ be distributed uniformly over residue classes. Then we prove, that b_j^m are distributed uniformly. We have $\{(a + k_i)| \ k_i \in S_n\} = S_n$ for integer a. By the assumption of induction b_j^{m-1} are distributed uniformly; if we add all possible $k_l \in S_n$ to each sum b_j^{m-1}, these numbers will be distributed uniformly too. To get $\{b_j^m\}$, we should exclude $m - 1$ numbers from n ones. Then the excluded numbers will have the following view:

$$a_j = 2k_j + b_l^{m-2}| \ k_j \in S_{nl}^{m-2},$$

with $\{2k_j (mod \ n)\} = S_n$ for the prime $n > 2$, where one part of numbers a_j is

$$\{a_j| \ \exists k_i \equiv 2k_j (mod \ n), \ k_i \in S_{nl}^{m-2}\} = \{b_j^{m-1}\}$$

which is distributed uniformly; the other part of a_j is

$$\{a_j| \ \exists k_i \equiv 2k_j (mod \ n), \ k_i \in S_{nl}^{m-2}\} = \{2k_r + b_j^{m-3}| \ k_r \in S_{nj}^{m-3}\}$$

and from it one can again pick out a uniformly distributed set. Finally, the remainder will have the view $\{b_j^{m'}\}$ where $m' \leqslant m - 1$, thus it will be distributed uniformly too. $\blacksquare$

Proof of Lemma 1 for the prime $n > 2$. Let $\{k_i\} = S_n$ and the structures of matrices A and B coincide, let $A = (a_{ik} = \zeta^{k_i})$, $B = (b_{ik} = \zeta^{2k_i})$. Then $(A^{n-m}B^m)_+ = E\zeta^c \sum_{j=1}^{C_n^m} \zeta^{a_j}$, where c is a constant and $a_j = b_j^m$. According to Theorem 2, a_j are distributed uniformly over the classes of S_n; then $(A^{n-m}B^m)_+ = E\zeta^c \sum_0^{n-1} \zeta^i / n = 0$. $\blacksquare$

Proof of Lemma 2. Let $\{A_i| \ i = 1, \ldots, m\}$ be concerted, $B_i = A_l \times A_i$, $i = 1, \ldots, m$, $B_{m+1} = A_l \times E$, $l \neq 1$. We prove, that $\{B_i| \ i = 1, \ldots, m + 1\}$ is also concerted. In

$$(B_1^{n_1} B_2^{n_2} \ldots)_+,$$

including B_{m+1}, we divide terms into groups with the same right part relative to product sign $\times$; every such a group contains C_n^s different terms with the

5

left part being $P(A_1^{n-s}A_l^s)$, where P is some permutation. Hence, the sum of terms of the group is $(A_1^{n-s}A_l^s)_+ \times c = 0$. ∎

The order of constructed matrices may be reduced due to the fact that the set $\{B_i = A_l \times A_i', \ B_l = A_l \times E'| \ l \neq 1\}$ is concerted if $\{A_i| \ i = 1, \ldots, k\}$ and $\{A_j'| \ j = 1, \ldots, k'\}$ are concerted too; the orders of A_i and A_j' may be different (compare with [8]).

Proof of Lemma 3 for the prime $n > 2$. Let specifically $A = (a_{i,i+1} = \zeta^{k_i})$, $B = (b_{i,i+1} = \zeta^{2k_i})$, where $k_i \equiv i(mod \ n)$, the indices are also residues modulo n. We prove that (A, B, AB) is concerted. One has $AB = \zeta BA$, as

$$AB = (c_{i,i+2} = \zeta^{k_i+2k_{i+1}}), \ BA = (c_{i,i+2}' = \zeta^{2k_i+k_{i+1}}).$$

Instead of

$$(A^{n_1}B^{n_2}(AB)^{n_3})_+, \ n_1 + n_2 + n_3 = n,$$

one can consider the expression

$$(A^{n_1}B^{n_2}(AB)^{n_3})_+ + (A^{n_1}B^{n_2}(BA)^{n_3})_+,$$

turning into zero simultaneously with the first one. In the last, all terms can be divided into groups, in which n_3 numbers of k_i from the variable set of $(n_2 + n_3)$ elements are fixed, and, besides, every group has $C_n^{n_2}$ terms of the view

$$\zeta^c \zeta^{\sum_1^{n_2} k_i''},$$

where $c = \sum_1^{n+n_3} k_i'$ and is not changed, $k_i' \in \{k_i\}$, $\{k_i''\} = S_n$ and $\sum_1^{n_2} k_i''$ are distributed uniformly by Theorem 2. Thus $(A^{n_1}B^{n_2}(AB)n_3)_+ = 0$. ∎

Proof of Theorem 1 for the prime $n > 2$. The validity of part A of the theorem for the prime $n > 2$ is secured by Lemmas 1 and 2. In proofs of Lemmas it is $A_i^n = E$ by construction. We prove now that part B is true. Let $r_1 = k$ and the condition (4) be satisfied.

Rewrite (3) as $A\Psi = B\Psi$, where $A = \sum_1^k A_i x_i$, $B = Ey$.

As $AB - BA = 0$, therefore, one gets $A^n\Psi = B^n\Psi$, multiplying by A^{n-1} from the left, where we have $A^n = E\sum_1^k x_i^n$ by (4).

As $|A| \neq 0$, by the equalities $|A^{n-1}||A - B| = |A^n - B^n| = 0$ it turns out, that if (2) takes place, a determinant of the system (3) is equal to zero. ∎

6

3 The particular case of $n = 3$

For $n = 3$ the following representation of three concerted matrices can be taken:

$$A_1 = \begin{pmatrix} 0 & \zeta & 0 \\ 0 & 0 & \zeta^2 \\ 1 & 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & 0 & 0 \\ \zeta & 0 & 0 \\ 0 & \zeta^2 & 0 \end{pmatrix}.$$

Then solutions of the equation $x_1^n + x_2^n + x_3^n = y^n$ can be parametrized with complex quantities Ψ_i (δ is a determinant of the system of linear equations relative to x_i) in such the manner:

$$y \equiv \delta = \Psi_1 \Psi_3^2 (1 - \zeta^2) - \Psi_3 \Psi_2^2 (1 - \zeta) - \Psi_2 \Psi_1^2 (\zeta - \zeta^2);$$

$$x_1 = \zeta \Psi_1^3 + \zeta^2 \Psi_2^3 + \Psi_3^3;$$

$$x_2 = -(\zeta \Psi_1^3 + \Psi_2^3 + \zeta^2 \Psi_3^3);$$

$$x_3 = \Psi_3 \Psi_1^2 (1 - \zeta^2) - \Psi_1 \Psi_2^2 (1 - \zeta) - \Psi_2 \Psi_3^2 (\zeta - \zeta^2).$$

If Ψ_i are integer numbers of the 3-circular field K_3, x_i and y should be the same ones.

This example shows that search for integer solutions of (1), after parametrization of it with the help of quantities Ψ_i, meets with another problem: it is necessary to study a possibility of such a choice of complex parameters, that $\Im x_i = \Im y_j = 0$.

For (1), Euler's complete rational parametrization for $n = 3$ is known (see [13]). Also Ramanujan has considered this problem [14]. Our parametrization is a general one: for any complex values of x_i, y (not only for integer or rational), for which the equation is true, a set of parameters Ψ_j exists. But there are similar problems for both parametrizations: it is difficult to ascertain for which values of the parameters we will find integer solutions.

4 Remarks on a solvability of Eqs. (1).

A combinatorial consideration leads to the restriction:

Theorem 3

$$r(n) \leqslant 2n + 1.$$

Proof. Let us consider a set $\{(a_1, a_2, \ldots, a_k)\}$ with $0 \leqslant a_i \leqslant A$, $a_i \leqslant a_{i+1}$. The number of elements of the set [15] is

$$N(A, k) = C_{A+k}^k = (A + k)!/k!A!,$$

and the number of different non-zero values of $\sum_1^k a_i^n$ is kA^n. If for $k = k_0$ some A exists such that $N(A, k_0) > k_0 A^n$, then $r(n) \leqslant 2k_0 - 1$. As $(A + k)!/A! > A^k$, and for $A^{k_0} > k_0!k_0 A^n$, the inequality $N(A, k_0) > k_0 A^n$ will be fulfilled; k_0 is fixed, therefore $k_0 = n + 1$ is enough. ∎

The bound $r(n) \leqslant 2n - 1$ has been recently obtained by Habsieger [16]; I am grateful to the referee who indicated me this fact.

There are two conjectures about $r(n)$:

$$1. \ r(n) \leqslant n, \ n > 1.$$

$$2. \ r(n) > e \ln(n),$$

where $e > 0$ is fixed.

In my unpublished work [7], the conjecture: $r(p) = p$ for the prime p was mentioned, but it is not correct because of the particular solution for $n = 5$ [17] (see below).

I give also the identities which are just for arbitrary numbers a_i (the proof was given in [7]):

$$\sum_{k=0}^{n} \sum_{j=1}^{C_n^k} (g_{jk})^n (-1)^k \equiv 2^n n! \prod_1^n a_i,$$

where $g_{jk} = \sum_{i=1}^{n} \sigma_{jki} a_i$ and σ_{jki} is the element of a sign matrix: $|\sigma_{jki}| = 1$, $\sum_{i=1}^{n} \sigma_{jki} = n - 2k$, $\sigma_{jki} = \sigma_{lmi} \delta_j^l \delta_k^m$. The identities have 2^{n-1} different terms on the left side .

Let us consider now the restrictions on $r(n)$ for $n \leqslant 12$, obtained by the author under search of particular solutions of (1) with PC AT 386. I describe briefly the used algorithm which is not connected with linearization of Eqs. (1). It is based on the idea of the most compact filling or taking up of some priming volume V_0 by single n-cubes, while V_0 is picked out by means of

8

running over on a few cycles. If $L \equiv \sum_1^m x_i^n$, $R \equiv \sum_1^k y_j^n$, $V = L - R$ and $x_i \in [0, A]$ with the fixed A, $y_j \in [f_1(L - \sum_1^{j-1} y_l^n), f_2(L - \sum_1^{j-1} y_l^n)]$, I shall refer to such an algorithm as $LmRk$, for instance $L1R3$ etc. Comparing $||V_0| - y^n|$ with $||V_0| - (y + 1)^n|$, we choose the smallest of them; then a value of y or $y + 1$, corresponding to it, will be chosen at the first step as a new element, V_0 will be replaced by $|V_0| - y^n$ or $|V_0| - (y + 1)^n$ and the procedure will be repeated. If after a fixed number of iterating a final volume is not equal to zero, V_0 must be replaced. Functions f_1 and f_2 were chosen empirically, usually as $((L - \sum_1^{j-1} y_l^n)/c)^{1/n}$ with some c. I do not describe the technical problem of representation of big integers by computations.

Instead of the evident record

$$\sum_1^{r_1} x_i^n = \sum_1^{r_2} y_j^n$$

let us describe solutions of (1) as

$$(x_1, x_2, \ldots, x_{r_1}; y_1, y_2, \ldots, y_{r_2})^n = 0.$$

It is known that $r(3) = 3$. In [7] the solutions were adduced for $n = 4, 5, 6$: $(3, 5, 8; 7, 7)^4 = 0$, $(4, 10, 20, 28; 3, 29)^5 = 0$, $(3, 19, 22; 10, 15, 23)^6 = 0$. The referee has indicated me Euler's solution for $n = 4$: $(133, 134; 158, 59)^4 = 0$ [13] and the very important one for $n = 5$: $(27, 84, 110, 133; 144)^5 = 0$ [17]. Hence, it follows that $r(4) \leqslant 3$, $r(5) \leqslant 4$, $r(6) \leqslant 5$.

I give the best restriction on $r(n)$ for $n = 7$:

$$(149, 123, 14, 10; 146, 129, 90, 15)^7 = 0, \; r(7) \leqslant 7, \;\; L3R3;$$

and the same for $n = 8$:

$$(43, 20, 11, 10, 1; 41, 35, 32, 28, 5)^8 = 0, \; r(8) \leqslant 9, \;\; L1R3.$$

For $n = 9$ we have:

$$(73, 38, 29, 9, 1; 68, 67, 45, 21, 18, 11, 6, 4)^9 = 0, \; r(9) \leqslant 12, \;\; L2R1.$$

Restrictions on $r(n)$ for $n = 10,\; 11$ are weaker, and the ones are the best:

$$(149, 42, 37, 30, 25, 20, 8, 5; 145, 128, 100, 73, 48, 13, 6, 1)^{10} = 0, \; r(10) \leqslant 15, \;\; L2R1;$$

$$(18, 6, 6, 6, 4, 4, 4; 17, 16, 15, 13, 13, 10, 9, 9, 8, 1, 1, 1, 1)^{11} = 0, \;\; r(11) \leqslant 19$$

(V_0 was chosen by randomization).

Thus, the particular solutions show that $r(n) \leqslant n| \; n \leqslant 7$.

5   Conclusion

A possibility of described linearization of Eqs.(1) sets a number of algebraic problems: the study of algebras of matrices $\{A_i\}$, construction of concerted sets of matrices with the smallest dimension (see [8]) etc.

New quantities Ψ for $n > 2$, also as spinors for $n = 2$, permit to build new algebraic objects with different transformation laws by transformations of coordinates $\{x_i,\ y_j\}$; research of such objects will be probably of a great interest.

A fundamental fact for the Diophantine equations (1) can be the possibility of parametrization of their solutions.

6   Acknowledgements

References

[1] H.M. Edwards. *Fermat's last theorem*, Springer-Verlag, New York Heidelberg Berlin, 1977.

[2] M.M. Postnikov. *Fermat's theorem*, Nauka, Moscow, 1978 (in Russian).

[3] S. Leng. *Fundamentals of Diophantine Geometry*, Springer-Verlag, New York Berlin Heidelberg Tokyo, 1983.

[4] G. Faltings. Invent. Math. **73** (1983), 349-366; Erratum: **75** (1984), 381.

[5] A.K. Kwasniewski. J.Math. Phys., **26** (1985), 2234.

[6] R.M. Jamaleev. JINR Report P5-87-766, Dubna, 1987 (in Russian).

[7] M.A. Ivanov. VINITI No 1844-B90, Moscow, 1990 (in Russian).

[8] N. Fleury, M. Rausch de Traubenberg. J. Math. Phys., **33** (1992), 3356.

[9] A. Granville. Acta Arithmetica **60** (1992), 203.

[10] P.A.M. Dirac. Proc. R. Soc. London, **A117** (1928), 610; A118 (1928), 351.

[11] E.E. Salpeter and H.A. Bethe. Phys. Rev., **84** (1951), 1232.

[12] M.A. Ivanov. Nuovo Cimento, **A105** (1992), 77 [hep-th/0207210].

[13] G.H. Hardy and E.M. Wright. *An Introduction to the theory of Numbers*, Clarendon Press, 1975.

[14] B.C. Brendt and S. Bhargava. Amer. Math. Month. **100** (1993), 644.

[15] A. Kaufmann. *Introduction a la combinatorique en vue des applications*, Dunod, Paris, 1968.

[16] L. Habsieger. J. of Number Theory **45** (1993), 92.

[17] C.J. Lander and T.R. Parkin. Math. Comp. **21** (1967), 101.