DEMONSTRATIO MATHEMATICA
Vol.IX No 3 1976

Tadeusz Kreid

ON REPRESENTATIONS OF NATURAL NUMBERS AS SUMS OF TWO SQUARES

Introduction

The aim of this study is stating some theorems concerning representations of natural numbers as sums of two squares. The study concludes with the derivation of the formula for function r, well known in the theory of numbers. We can obtain these theorems easy if we apply the arithmetic of complex integers. But we do not use this arithmetic here.

Let I - the set of integers, N - the set of natural numbers, $N_0 = N \cup \{0\}$.

D e f i n i t i o n. A representation of number $n \in N$ as a sum of two squares is a pair $(a,b) \in N \times N_0$ such that

$$a^2 + b^2 = n .$$

We denote for $n \in N$:

$$\alpha_n = \{(a,b) \in N \times N_0 : a^2 + b^2 = n\}$$

- the set of all representations of number n. The family $\{\alpha_n : n \in N\}$ is a partition of $N \times N_0$.

We define the function r as follows

$$r(n) = \overline{\overline{\alpha_n}}$$

i.e. as the number of representations of n as a sum of two squares.

R e m a r k. This definition differs from the one given in [1], where all pairs $(a,b) \in I \times I$ are considered. Treating pairs of numbers as points on the plane, the restriction to pairs of $N \times N_0$ is the restriction to one quarter of the plane: $0 \leq \varphi < \frac{\pi}{2}$ (φ - the polar coordinate). With each $(a,b) \in N \times N_0$ also theree other points: $(-b,a)$, $(-a,-b)$, $(b,-a)$ belonging to other quarters are representations of n. Thus the introduced restriction gives in effect omitting superfluous and constant factor 4 in the formula for function τ. This simplification has one more virtue: in this way the function τ becomes multiplicative (see [2]).

1. Multiplying of representations

In the following considerations we shall apply the geometric interpretation of representations. A pair from $N \times N_0$ can be considered as a point or vector with integer coordinates on the Cartesian plane. For vector $\vec{v} = [v_x, v_y] \in N \times N_0$ we have: $v_x > 0$, $v_y \geq 0$, $|\vec{v}| \geq 1$, $0 \leq \sphericalangle (Ox, \vec{v}) < \frac{\pi}{2}$ ($\sphericalangle$ denotes the orientated angle).

In the set of all representations $N \times N_0$ we define the following operation $\times$ (multiplying of representations)

$$(1) \qquad (a,b) \times (c,d) = \begin{cases} (ac + bd,\ ad - bc) & ad-bc \geq 0 \\ & \text{if} \\ (bc - ad,\ ac + bd) & ad-bc < 0. \end{cases}$$

This operation can be also written as follows

$$(2) \qquad \vec{v} \times \vec{w} = \begin{cases} [\vec{v}.\vec{w},\ \vec{v} \wedge \vec{w}\,] & \operatorname{sign}(\vec{v},\vec{w}) = +1,\ 0 \\ & \text{if} \\ [-\vec{v} \wedge \vec{w},\ \vec{v}.\vec{w}] & \operatorname{sign}(\vec{v},\vec{w}) = -1 \end{cases}$$

(sign$(\vec{v},\vec{w})$ - the orientation of the pair $(\vec{v},\vec{w})$)or

$$(3)\quad \vec{v}\times\vec{w}=\begin{cases}|\vec{v}|\cdot|\vec{w}|\left[\cos\sphericalangle(\vec{v},\vec{w}),\ \sin\sphericalangle(\vec{v},\vec{w})\right] & \sphericalangle(\vec{v},\vec{w})\geqslant 0\\ |\vec{v}|\cdot|\vec{w}|\left[-\sin\sphericalangle(\vec{v},\vec{w}),\ \cos\sphericalangle(\vec{v},\vec{w})\right] & \sphericalangle(\vec{v},\vec{w})<0.\end{cases}\quad \text{if}$$

Hence we obtain the formula

$$(4)\qquad |\vec{v}\times\vec{w}| = |\vec{v}|\cdot|\vec{w}|.$$

C o n c l u s i o n. If $\vec{v}\in\alpha_n$ and $\vec{w}\in\alpha_m$ then $\vec{v}\times\vec{w}\in\alpha_{nm}$.

2. Preliminary theorems

T h e o r e m 1. For $n\in N$

$$\tau(2n)=\tau(n).$$

T h e o r e m 2. Let $n\in N$. If p is a prime of the form $4k+3$, then

$$\tau(np^2)=\tau(n).$$

T h e o r e m 3. If $n,m>1$ - numbers representable as a sum of two squares and relatively prime, then the mapping

$$\times : (\alpha_n\times\alpha_m)\to\alpha_{nm}$$

is a one-to-one function.

P r o o f. Let $\vec{v}, \vec{v}'\in\alpha_n$; $\vec{w}, \vec{w}'\in\alpha_m$ and $\vec{v}\times\vec{w}=\vec{v}'\times\vec{w}'$. We apply (3) and notice that $|\vec{v}'|=|\vec{v}|$ and $|\vec{w}'|=|\vec{w}|$. After considering the cases which derive from (3) we come to the following conclusion

$$\bigvee_{\varepsilon\in I}\ \varphi'=\varphi+\varepsilon\frac{\pi}{2}$$

where $\varphi = \sphericalangle(\vec{v},\vec{w})$, $\varphi' = \sphericalangle(\vec{v}',\vec{w}')$.

Let $\alpha = \sphericalangle(\vec{v},\vec{v}')$. We have

$$\sphericalangle(\vec{w},\vec{w}') = \sphericalangle(\vec{w},\vec{v}) + \sphericalangle(\vec{v},\vec{v}') + \sphericalangle(\vec{v}',\vec{w}') = -\varphi+\alpha+\varphi' = \alpha + \varepsilon\frac{\pi}{2}.$$

Let θ be the rotation of the plane through the angle α. Thus $\theta(\vec{v}) = \vec{v}'$. Let $\vec{w}'' = \theta(\vec{w})$. We have

$$\sphericalangle(\vec{w}',\vec{w}'') = \sphericalangle(\vec{w}',\vec{w}) + \sphericalangle(\vec{w},\vec{w}'') = -\alpha + \alpha + \varepsilon\frac{\pi}{2} = \varepsilon\frac{\pi}{2}.$$

The vector $\vec{w}''$ can be obtained as an image of $\vec{w}'$ after rotation through a multiple of right angle, so its coordinates are integers too. We will find them.

Let $\vec{v} = [a,b]$, $\vec{v}' = [a',b']$, $\vec{w} = [c,d]$. We have $a^2 + b^2 = a'^2 + b'^2 = n$, $c^2 + d^2 = m$,

$$\cos\alpha = \frac{\vec{v}\cdot\vec{v}'}{|\vec{v}|^2} = \frac{aa'+bb'}{n}, \qquad \sin\alpha = \frac{\vec{v}\wedge\vec{v}'}{|\vec{v}|^2} = \frac{ab'-a'b}{n}.$$

Therefore

$$\vec{w}'' = \theta(\vec{w}) = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}\begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} \frac{1}{n}\left[c(aa'+bb')-d(ab'-a'b)\right] \\ \frac{1}{n}\left[c(ab'-a'b)+d(aa'+bb')\right] \end{bmatrix}.$$

The coordinates of $\vec{w}''$ are integers, thus there exist $i,j \in I$ such that

$$c(aa' + bb') - d(ab' - a'b) = in,$$

$$c(ab' - a'b) + d(aa' + bb') = jn.$$

Hence

$$n(jc-id) = c(jn) - d(in) = c^2(ab' - a'b) + d^2(ab' - a'b) =$$

$$= (c^2 + d^2)(ab' - a'b) = m(ab' - a'b),$$

and thus it follows that $n \mid m(ab'-a'b)$. But n,m are relatively prime, so we get $n \mid (ab'-a'b)$. In a similar way we obtain

$$n(jd+ic) = d(jn) + c(in) = m(aa' + bb')$$

hence $n \mid (aa'+bb')$.

This implies that there exist $k,l \in I$ such that

$$ab' - a'b = kn, \quad aa' + bb' = ln.$$

Thus we obtain

$$(k^2+l^2)n^2 = (kn)^2 + (ln)^2 = (ab'-a'b)^2 + (aa'+bb')^2 =$$

$$= (a^2 + b^2)(a'^2 + b'^2) = n^2 ,$$

and $k^2 + l^2 = 1$.
Since $ln = aa'+bb' > 0$ we infer that $l = 1$, $k = 0$. From $ab'-a'b = 0$ we get immediately $a' = a$ and $b' = b$, i.e. $\vec{v}' = \vec{v}$. Now it follows that $\alpha = 0$, $\vec{w}'' = \vec{w}$, $\sphericalangle(\vec{w},\vec{w}') = \sphericalangle(\vec{w}'',\vec{w}') = \varepsilon\frac{\pi}{2}$, but $-\frac{\pi}{2} < \sphericalangle(\vec{w},\vec{w}') < \frac{\pi}{2}$, so we have $\sphericalangle(\vec{w},\vec{w}') = 0$ and $\vec{w}' = \vec{w}$.

3. Theorems on representations

We shall call a representation (a,b) relatively prime if a,b are relatively prime. For an arbitrary $n \in N$, $n > 1$, an element x of the ring of residues $(Z_n; \oplus, \odot, \ominus)$ will be called a root of -1 if $x^2 = \ominus 1$ (i.e. $x^2 \oplus 1 = 0$) in Z_n.

R e m a r k 1. If $p^2 + q^2 = v > 1$ and p is relatively prime to v, then also q is relatively prime to v and the representation (p,q) is relatively prime.

R e m a r k 2. If (p,q) is a relatively prime representation of number $n > 1$, then $p,q \in Z_n$ are inversible and $p^{-1} \odot q$ is a root of -1 (because $p^2 \oplus q^2 = 0$ and $(p^{-1} \odot q)^2 = p^{-2} \odot q^2 = p^{-2} \odot (\ominus p^2) = \ominus 1$).

Thus with every relatively prime representation (p,q) of number n we may connect a root of -1: $p^{-1} \odot q \in Z_n$.
We will show that this mapping is one-to-one.

L e m m a 1. Let (i,j), (k,l) be relatively prime representations of number $n > 1$. If $(i,j) \neq (k,l)$ then $i^{-1} \odot j \neq k^{-1} \odot l$.

P r o o f. The elements $i,j,k,l \in Z_n$ are inversible. Let $jk \geq il$ (if not, we exchange the denotation of pairs). Assuming that $i^{-1} \odot j = k^{-1} \odot l$, i.e. $j \odot k = i \odot l$, we get

$$\bigvee_{t \geq 0} jk = il + tn,$$

and

$$n(2ilt + nt^2) = 2iltn + n^2t^2 = (il + nt)^2 - i^2l^2 =$$

$$= j^2k^2 - i^2l^2 = k^2(n - i^2) - i^2(n - k^2) = n(k^2 - i^2),$$

which implies that $2ilt + nt^2 = k^2 - i^2$. But $nt^2 = k^2 - i^2 - 2ilt < k^2 < n$, hence $t^2 < 1$ and $t = 0$, so $jk = il$. We notice that

$$i^2n = i^2(k^2 + l^2) = (ik)^2 + (il)^2 =$$

$$= (ik)^2 + (jk)^2 = (i^2 + j^2)k^2 = nk^2.$$

Hence $i = k$ and $j = l$.

L e m m a 2. Every number of the form p^k, where $k \in N$ and p - prime of the form $4l+1$, has at least two relatively prime representations.

P r o o f (by induction on k). From Fermat's theorem we know that p can be expressed as a sum of two squares $p = i^2 + j^2$, so (i,j), (j,i) are two $(i \neq j)$ relatively prime representations of p.

Let (s,t) be a relatively prime representation of p^k and let $s' = s \bmod p$, $t' = t \bmod p$ $(s',t' \in Z_p)$. We have

$$s'^2 \oplus t'^2 = (s^2 + t^2) \bmod p = p^k \bmod p = 0,$$

$$(s'^{-1} \odot t')^2 = (s')^{-2} \odot (\ominus s'^2) = \ominus 1.$$

There exist exactly two numbers $x \in Z_p$ such that $x^2 = \ominus 1$, namely $i^{-1} \bullet j$ and $j^{-1} \bullet i$, so

$$s'^{-1} \odot t' = i^{-1} \odot j \quad \text{or} \quad s'^{-1} \odot t' = j^{-1} \bullet i.$$

Let

$$(i',j') = \begin{cases} (i,j) \\ \\ (j,i) \end{cases} \text{ if } s'^{-1} \odot t' = \begin{cases} j^{-1} \odot i \\ \\ i^{-1} \odot j \end{cases}.$$

We notice that $i'^{-1} j' s'^{-1} t' = 1$. The pair $(A,B) = (s,t) \times (i',j')$ is a representation of p^{k+1}, so according to (1) (A,B) equals either $(si'+tj', sj'-ti')$ or $(ti'-sj', si'+tj')$. We observe that

$$(si'+tj') \bmod p = s' \odot i' \odot t' \odot j' = s'i'(1 \bullet s'^{-1} t' i'^{-1} j') = 2 s' i' \neq 0,$$

hence $(si'+tj')$ is relatively prime to p^{k+1}. So the representation (A,B) is relatively prime. The pairs (A,B) and (B,A) are two relatively prime representations of number p^{k+1}.

L e m m a 3. Let $v,w > 1$. If (s,t) is a relatively prime representation of v, (p,q) is a relatively prime representation of w and v,w are relatively prime, then the representation $(s,t) \times (p,q)$ of number vw is relatively prime.

P r o o f. First we will show that $sp+tq$ is relatively prime to v. Let us consider the ring Z_v and let $p' = p \bmod v$, $q' = q \bmod v$. Then we have

$$p'^2 \bullet q'^2 = (p^2 + q^2) \bmod v = w \bmod v.$$

Since w is relatively prime to v, also $p'^2 \bullet q'^2$ is and $p'^2 \bullet q'^2 \in Z_v$ is an inversible element. Also $s,t \in Z_v$ are inversible. We notice that

$$(p' \ominus s'^{-1} t q') \odot (p' \bullet s'^{-1} t q') = p'^2 \ominus s'^{-2} t^2 q'^2 = p'^2 \ominus (\ominus 1) q'^2 =$$

$$= p'^2 \bullet q'^2.$$

Hence $p' \circledast s^{-1}tq'$ is inversible and also $sp' \circledast tq' = s(p' \circledast s^{-1}tq')$ is inversible, i.e. relatively prime to v. So sp+tq is also relatively prime to v.

Exchanging the roles of v and w we see that sp+tq is relatively prime to w as well. Hence sp+tq is relatively prime to vw. The number sp+tq is one of two elements of the representation $(s,t) \times (p,q)$ of the number vw. By virtue of Remark 1 this representation is relatively prime.

L e m m a 4. Let $m > 1$ be a number having a relatively prime representation. The mapping defined for relatively prime representations of m as follows

$$(a,b) \longrightarrow a^{-1} \circ b$$

is a one-to-one function onto the set $Z = \{x \in Z_m : x^2 = \circ 1\}$.

P r o o f. By virtue of Lemma 1 this function is one-to--one.

Let (p,q) be a relatively prime representation of m. If a prime of the form $r = 4l+3$ were a divisor of $m = p^2 + q^2$, then it would not be a divisor of p nor q, but in this case it would be a sum of two squares, which is a contradiction. If p,q are both odd, then m mod 4 = 2, if one of them is even and the other odd, then m mod 4 = 1, so 4 does not divide m. We conclude that m is of the form

$$m = 2^{\varepsilon} p_1^{\alpha_1} \ldots p_k^{\alpha_k},$$

where $k \in N_0$; $\varepsilon = 0,1$; $p_1, p_2, \ldots, p_k$ – primes of the form $4l+1$; $\alpha_1, \alpha_2, \ldots, \alpha_k \in N$. We will show by induction on k that every number of this form has at least 2^k relatively prime representations.

By virtue of Lemma 2 the number $p_1^{\alpha_1}$ has two relatively prime representations. If the number $p_1^{\alpha_1} \ldots p_k^{\alpha_k}$ has 2^k different relatively prime representations $(a_1,b_1), \ldots, (a_{2^k}, b_{2^k})$, then taking two different relatively prime representations (i_1,j_1), (i_2,j_2) of $p_{k+1}^{\alpha_{k+1}}$ we obtain by multiplying

2^{k+1} representations of number $p_1^{\alpha_1} \ldots p_k^{\alpha_k} p_{k+1}^{\alpha_{k+1}}$, namely

$$(a_t,b_t) \times (i_\varepsilon,j_\varepsilon), \qquad \text{where} \quad 1 \leqslant t \leqslant 2^k; \quad \varepsilon = 1,2.$$

These representations are different and relatively prime (by virtue of Theorem 3 and Lemma 3). For a number of the form $2p_1^{\alpha_1} \ldots p_k^{\alpha_k}$ the formula $(1,1) \times (a,b)$ gives 2^k relatively prime representations when (a,b) runs through 2^k relatively prime representations of number $p_1^{\alpha_1} \ldots p_k^{\alpha_k}$.

Let M be the number of relatively prime representations of m. We have stated that $M \geqslant 2^k$. The number of elements of the image of the mapping

$$A = \left\{ a^{-1} \odot b : \ (a,b) \text{ - relatively prime representation} \right\}$$

is also equal M (because this mapping is one-to-one). But $A \subset Z$, and from the theory of congruences with composite modulus we know that for $m = 2^\varepsilon p_1^{\alpha_1} \ldots p_k^{\alpha_k}$ the number of elements of the set Z is equal 2^k. So $M \leqslant 2^k$, $\overline{\overline{A}} = M = 2^k = \overline{\overline{Z}}$, $A = Z$ and the mapping $(a,b) \to a^{-1} \odot b$ is one-to-one and onto.

C o n c l u s i o n. The number m has exactly 2^k relatively prime representations.

T h e o r e m 4. If (p,q) is a relatively prime representation of number nm, then there exist representations (a,b) of n and (c,d) of m such that

$$(p,q) = (a,b) \times (c,d).$$

P r o o f. Let $m > 1$. The numbers p,q are both relatively prime to nm, so to m too. In the ring Z_m we have

$$(p \bmod m)^2 \odot (q \bmod m)^2 = (p^2 + q^2) \bmod m = 0,$$

$$((p \bmod m)^{-1} \odot (q \bmod m))^2 = e1.$$

The number nm is of the form $2^\varepsilon p_1^{\alpha_1} \ldots p_k^{\alpha_k}$, because it has a relatively prime representation. Hence m, as its divisor, is also of this form, and in consequence it has relatively prime representations.

By virtue of Lemma 4 there exists a representation (c,d) of m such that

$$c^{-1} \odot d = (p \bmod m)^{-1} \odot (q \bmod m).$$

Now let us consider the representation $(p,q) \times (c,d)$ of the number nm^2. By definition $(p,q) \times (c,d)$ equals either (pc+qd, pd-cq) or (cq-pd, pc+qd). We notice that

$$(pc+qd) \bmod m = (p \bmod m) \odot c \oplus (q \bmod m) \odot d =$$

$$= (p \bmod m)c^{-1} \left[c^2 \oplus (p \bmod m)^{-1}(q \bmod m)cd \right] =$$

$$= (p \bmod m)c^{-1} \left[c^2 \oplus c^{-1}dcd \right] = (p \bmod m)c^{-1}(c^2 \oplus d^2) = 0$$

(since $c^2 + d^2 = m$). Hence $m \mid (pc+qd)$ and $m \mid (pd-cq)$, thus the numbers $\frac{pc+qd}{m}$ and $\frac{pd-cq}{m}$ are integers. The pair

$$(a,b) = \frac{1}{m} \left[(p,q) \times (c,d) \right]$$

is a representation of n. Besides we have

$$(a,b) \times (c,d) = \left(\frac{pc+qd}{m}, \frac{pd-cq}{m} \right) \times (c,d) \quad \left(= \left(\frac{cq-pd}{m}, \frac{pc+qd}{m} \right) \times (c,d) \right)$$

$$= \left(\frac{p(c^2 + d^2)}{m}, \frac{q(c^2 + d^2)}{m} \right) = (p,q).$$

If $m = 1$ then $(p,q) = (q,p) \times (0,1)$.

T h e o r e m 5. If $n,m > 1$ are numbers representable as a sum of two squares and are relatively prime, then the mapping

$$\times : (\alpha_n \times \alpha_m) \to \alpha_{nm}$$

is onto.

P r o o f. Let $(a,b) \in \alpha_{nm}$, i.e. $a^2 + b^2 = nm$ and let $d = GCD(a,b)$, $p = \frac{a}{d}$, $q = \frac{b}{d}$. So p,q are relatively prime. Moreover $d^2 \mid nm$ and n,m are relatively prime, so $GCD(d^2,n)$ and $GCD(d^2,m)$ are squares. Let us denote

$$i^2 = \mathrm{GCD}(d^2,n), \quad j^2 = \mathrm{GCD}(d^2,m),$$

thus the numbers i,j are relatively prime and $i^2j^2 = d^2$ so ij = d. Since

$$p^2 + q^2 = \frac{a^2 + b^2}{d^2} = \frac{nm}{i^2j^2} = \frac{n}{i^2} \cdot \frac{m}{j^2},$$

by virtue of Theorem 4 there exist representations (k,l) of $\frac{n}{i^2}$ and (s,t) of $\frac{m}{j^2}$ such that $(k,l) \times (s,t) = (p,q)$. Then for the pairs (ik,il), (js,jt) we have

$$(ik)^2 + (il)^2 = i^2(k^2+l^2) = i^2 \frac{n}{i^2} = n,$$

$$(js)^2+(jt)^2 = j^2(s^2+t^2) = j^2 \frac{m}{j^2} = m,$$

which means that $(ik,il) \in \alpha_n$, $(js,jt) \in \alpha_m$, and

$$(ik,il) \times (js,jt) = ij\left[(k,l) \times (s,t)\right] = ij(p,q) =$$

$$= d(p,q) = (dp,dq) = (a,b).$$

C o n c l u s i o n. If $n,m > 1$ are relatively prime and representable as sums of squares then the function

$$\times : (\alpha_n \times \alpha_m) \to \alpha_{nm}$$

is one-to-one and onto. Hence

$$\tau(n \cdot m) = \overline{\overline{\alpha_{nm}}} = \overline{\overline{\alpha_n \times \alpha_m}} = \overline{\overline{\alpha_n}} \cdot \overline{\overline{\alpha_m}} = \tau(n) \cdot \tau(m).$$

T h e o r e m 6. If p is a prime number of the form 4k+1 and $\alpha \in N_0$, then

$$\tau(p^\alpha) = \alpha + 1.$$

P r o o f (by induction on α with step 2). It is evident that $\tau(1) = 1$ and $\tau(p) = 2$. If $\tau(p^\alpha) = \alpha + 1$, then let $(x_1,y_1),\dots,(x_{\alpha+1}, y_{\alpha+1})$ be all representations of p^α. Then $(px_1,py_1),\dots,(px_{\alpha+1},py_{\alpha+1})$ are all different not relatively prime representations of $p^{\alpha+2}$. We know that there are exactly two relatively prime representations of $p^{\alpha+2}$. So there are $(\alpha + 2) + 1$ of them all.

C o n c l u s i o n. For an arbitrary natural number

$$n = 2^\alpha p_1^{\alpha_1}\dots p_k^{\alpha_k} q_1^{\beta_1}\dots q_l^{\beta_l},$$

where $p_1,\dots,p_k$ are primes of the form $4t+1$; $q_1,\dots,q_l$ are primes of the form $4t+3$ and $\beta_1,\dots,\beta_l$ are even numbers, we get by virtue of Theorems 1,2, the previous conclusion and Theorem 6

$$\tau(n) = \tau(p_1^{\alpha_1}\dots p_k^{\alpha_k}) = \tau(p_1^{\alpha_1})\dots\tau(p_k^{\alpha_k}) = (\alpha_1 + 1)\dots(\alpha_k+1).$$

So we have obtained the formula

$$\tau(n) = (\alpha_1+1)\dots(\alpha_k+1).$$

R e m a r k. The proofs of all existence theorems stated here are effective. So for an arbitrary number, when we have the representations of its prime factors, we can find by methods given here all its representations (or only relatively prime ones).

REFERENCES

[1] W. S i e r p i ń s k i: Teoria liczb. Warszawa-Wrocław 1950.
[2] I. W i n o g r a d o w: Elementy teorii liczb. Warszawa 1954.

INSTITUTE OF MATHEMATICS, UNIVERSITY OF WARSAW
Received June 2nd, 1975