

SUMS OF SQUARES IN PLANAR NASH RINGS

GUSTAVE A. EFROYMSON

In a previous paper, the author developed the ideal theory of Nash rings on planar domains. Here that theory is applied to show that a Nash function on a planar domain D which takes on only non-negative values on D is a sum of two squares of Nash functions on D . Examples are given which clarify the situation for higher dimensions.

0. Introduction. In the paper, "Nash Rings on Planar Domains" [2], the ideal structure of planar Nash rings was discussed. Using the results of this paper, it is possible to show that for D a planar semi-algebraic domain, that if $f \in N_D$, the Nash ring on D , then $f \geq 0$ on D implies that there exist f_1, f_2 in N_D with $f = f_1^2 + f_2^2$. Recall that for polynomials, i.e. for $R[x, y]$, it is known that there exist polynomials $p(x, y)$ for which there is no representation as a sum of any number of squares of polynomials (e.g. Motzkin's example [3]) and that if we use elements in $R(x, y)$ four squares may be necessary [1]. As can be seen, the situation for Nash rings is merely shifted by one dimension since by "homogenizing" Motzkin's example, one obtains a Nash function which can't be represented as a sum of squares of any number of Nash functions, and if quotients of Nash functions are allowed, at least four are necessary to get this function as a sum of squares.

1. The Main Theorem. We first recall some definitions.

DEFINITION 1. A semi-algebraic planar domain is a semi-algebraic domain $D \subseteq R^2$. In addition to D being open, and connected, we require that D can be defined by a finite number of polynomial inequalities.

DEFINITION 2. A Nash function on D is a real valued function $f(x, y)$ such that f is analytic on D , and there exists a polynomial $p_f(x, y, z)$ such that $p_f(x, y, f(x, y)) = 0$ identically on D . This is just to say f is algebraic and analytic on D .

DEFINITION 3. For D as above, $N_D =$ the ring of Nash functions on D .

The main result of this paper can now be stated.

THEOREM 1. *Let N_D be as above and suppose f is an element of*

N_D such that $f \geq 0$ on D . Then there exist f_1 and f_2 in N_D so that $f = f_1^2 + f_2^2$.

Recall that for analytic functions, one has the exact sequence of sheaves of germs of analytic functions:

$$0 \longrightarrow \mathcal{O}_X \xrightarrow{f \rightarrow e^f} \mathcal{O}_X^* \longrightarrow \mathbb{Z}_2 \longrightarrow 0.$$

Here \mathcal{O}_X is the sheaf of germs of analytic functions on X , and \mathcal{O}_X^* is the sheaf of non-zero germs under multiplication. From the long exact sequence for cohomology, one finds:

$$H^1(X, \mathcal{O}_X) \longrightarrow H^1(X, \mathcal{O}_X^*) \longrightarrow H^1(X, \mathbb{Z}_2) \longrightarrow H^2(X, \mathcal{O}_X)$$

but since $H^1(X, \mathcal{O}_X) = H^2(X, \mathcal{O}_X) = 0$, we obtain $H^1(X, \mathcal{O}_X^*) \cong H^1(X, \mathbb{Z}_2)$. We also have the exact sequence of sheaves:

$$1 \longrightarrow \mathcal{O}_X^* \longrightarrow \mathcal{M}^* \longrightarrow \mathcal{M}^*/\mathcal{O}_X^* = \mathcal{D} \longrightarrow 1$$

where \mathcal{M}^* is the sheaf of germs of meromorphic functions and \mathcal{D} is the sheaf of divisors. Again the long exact sequence can be taken

$$\begin{array}{ccccccc} H^0(X, \mathcal{O}_X^*) & \longrightarrow & H^0(X, \mathcal{M}^*) & \longrightarrow & H^0(X, \mathcal{D}) & \longrightarrow & H^1(X, \mathcal{O}_X^*) \\ & & \parallel & & \parallel & & \\ & & K & & \Gamma(\mathcal{D}) & & \\ & \longrightarrow & H^1(X, \mathcal{M}^*) & & & & \\ & & \parallel & & & & \\ & & 0 & & & & \end{array}$$

so $H^1(X, \mathbb{Z}_2) \cong H^1(X, \mathcal{O}_X^*) \cong \Gamma(\mathcal{D})/K =$ class group of $\Gamma(\mathcal{O}_X)$.

If D is a simply connected region in \mathbb{R}^2 , we find $H^1(D, \mathbb{Z}_2) = 0$. This implies by the above result that the class group is trivial. In [2] we showed that if $H^1(D, \mathbb{Z}_2) = 0$, then the class group for N_D , the ring of Nash functions on D , is also trivial even though one no longer has $H^1(X, N_D) = 0$. This means that to give a Nash divisor (locally) on \mathbb{R}^2 is to define up to a unit a Nash function which will generate the principal ideal so defined. Also, since every Nash function only vanishes on parts of algebraic sets which therefore must be unions of analytic components of algebraic sets, we see that the support of a Nash divisor on N_D will be a union of analytically connected components of an algebraic set. A minimal such set will be either a curve C or a real point P . We call the associated prime \mathcal{P}_C or \mathcal{P}_P . In the case of a point there is some ambiguity since the same point may be the support of different primes; for example, $(x^2 + y^2)$ and $(x^2 + 2y^2)$ are different primes with support at the origin. Of course, the complex support is much larger.

For general semi-algebraic regions $D \subset \mathbb{R}^2$, there will be analytically connected arcs of algebraic curves which run from one hole of D to another. So we have 3 possibilities for irreducible primes in N_D . First we have those primes \mathcal{P}_C which have support a curve C which passes through no hole of D . For such a prime, there exists a global Nash function f on \mathbb{R}^2 with $(f) = \mathcal{P}_C$. Then there are the primes associated with a curve running from one hole of D to another, and the third possibility is a prime with real support on one point.

Now take a function $f \in N_D$. Then $(f) = \prod_{i,j} \mathcal{P}_{C_i} \mathcal{P}_{P_j}$. We claim that each \mathcal{P}_{C_i} must occur to an even power if $f \geq 0$ on D . For, let t be a local parameter for \mathcal{P}_{C_i} at a non-singular point P of C_i where no other divisor of f has support. Then if $f = ht^{2m+1}$ locally, where h is a unit at P , we would find that f changes sign near P and so we would get a contradiction.

So if C_i is of the first type, \mathcal{P}_{C_i} is principal and $=(f_i)$, and so f_i^{2m} divides f for some integer m . And if C_i is of the second type, $\mathcal{P}_{C_i}^2$ again divides (f) , but $\mathcal{P}_{C_i}^2$ is principal since the class group is $H^1(D, \mathbb{Z}_2)$ as shown in [2], and so $\mathcal{P}_{C_i}^2 = (g_i)$ and g_i^m will divide f for some integer m . We now have reduced the proof of Theorem 1 to the following cases, using the fact that if f_1 and f_2 are sums of two squares, then so is $f_1 f_2$.

Case 1. The support of f is one point and f is irreducible in N_D . There exists an irreducible polynomial $p(x, y)$ such that f divides p in N_D . We can assume that the support of f is the point $(0, 0)$, using translation. It was shown in [2] that, by a change of variables, one can get p so that it factors $p = \prod_j (y - \alpha_j(x))$ and each $\alpha_j(x)$ is analytic except possibly at $x = 0$. The factors which divide f at $(0, 0)$ are thus only those for which $\alpha_j(0) = 0$. By multiplying the factors which divide f in a complex neighborhood of $(0, 0)$, we get f back, up to a factor which is a unit at $(0, 0)$. Note also that since f is real, if $y - \alpha_j(x)$ is a factor of f so is $y - \bar{\alpha}_j$ the complex conjugate.

Now expand $\alpha_j(x)$ in Puiseux series and get $\sum a_k x^k$, $k \in \mathbb{Q}$, with bounded denominators. Suppose all these are in $\mathbb{C}((x^{1/n}))$. Then, since the Galois group of $\mathbb{C}((x^{1/n}))/\mathbb{C}((x))$ is generated by the substitution $x^{1/n} \rightarrow \omega_n x^{1/n}$, for ω_n a primitive n th root of unity, we see that the conjugates of $y_1 = \sum a_k x^k$ are just $y_j = \sum a_k \omega_n^{kj} x^k$, (over $\mathbb{C}((x))$, $j = 1, \dots, n$). As long as no two of the y_j are complex conjugates, we are done since after multiplying together all the $(y - y_j)$, we will get a function $h(x, y)$ in $N_D[i]$. So $h\bar{h}$ will be a sum of two squares in N_D .

So now suppose two of the y_j are complex conjugates. We can

suppose one is y_1 and reparameterizing by letting $x = t^n$, we obtain $y_1 = \sum_{i=1}^{\infty} a_i t^i$. So suppose $y_k = \bar{y}_1 = \sum \bar{a}_j t^j = \sum a_j \omega^{jk} t^j$, $a_j = R_j e^{i\theta_j}$ and $\omega = e^{i\alpha}$. Then $\bar{a}_j = R_j \bar{e}^{i\theta_j} = R_j e^{i\theta_j} e^{i\alpha j}$. From this it follows that $j\alpha + \theta_j \equiv -\theta_j \pmod{2\pi}$ and so $\theta_j \equiv -j(\alpha/2) \pmod{\pi}$. Let $u = e^{-(i\alpha/2)} t$ and $\sum a_j t^j = \sum a_j e^{i(\alpha j/2)} u^j = \sum \varepsilon_j R_j u^j$ where $\varepsilon_j = \pm 1$ and then $x = \pm u^n$. Now we have a real parameterization of the curve and so don't have an isolated point.

Case 2. The support of f is an arc C which runs from one hole of D to another. Next, we send one end of C to infinity and the other end to $(0, 0)$. Then find half lines L_1, L_2 going from $(0, 0)$ to infinity so that $L_1 \neq L_2$ and neither L_1 nor L_2 passes through any singular point of C in D . Let $V_i = L_i \cup C$. Now $I(L_i \cup C) = I(L_i)I(C)$. By the Corollary on p. 444 of [2], we find that if the C_i are arcs which run from one hole of D to another, then $\prod_i I(C_i)$ is principal iff the number of ends of the C_i at each hole is even. This is, of course, just another way of saying that the class group of N_D is isomorphic to $H^1(D, \mathbb{Z}_2)$. So since both L_i and C have one end at $(0, 0)$ and the other at infinity, the above result implies that $I(L_i \cup C)$ is principal. So there exists f_i in N_D with $I(V_i) = (f_i)$. Now I claim that $f/(f_1^n + f_2^n)$ is a unit in N_D . Just look locally at each P of D . We only have to worry about P on C and then have two subcases. First, when P lies on C but not on L_1 or L_2 , then the fact that f vanishes on C implies that f is a multiple (locally) of f_i , $i = 1, 2$. But $f = u f_i$, for u a local unit, is impossible since f doesn't change sign near P . So $f = g_i f_i^2$ for some g_i , again locally at P . Now, since P is not on L_1 or L_2 , each f_i locally generates $I(C)$, so $f_1 = u f_2$ for u unit at P . So $f_1^2 + f_2^2 = (1 + u^2) f_2^2$ and so divides f at P . Now assume that P is at an intersection of L_1 and C (not the origin which isn't even in D), then C is non-singular at P and we can assume the intersection of L_1 and C is transversal (moving L_1 a little if necessary). Now let t_C and t_{L_1} be local parameters for C and L_1 respectively. Then $f_1 = t_C t_{L_1} u_1$, $f_2 = t_C u_2$, for u_1 and u_2 units at P . So we see $f_1^2 + f_2^2 = t_C^2 (u_2^2 + u_1^2 t_{L_1}^2)$, which implies that $f_1^2 + f_2^2$ divides f at P . Thus $f/(f_1^2 + f_2^2)$ is analytic on D . Since f is irreducible, $f = u(f_1^2 + f_2^2)$, for u a unit in N_D . Then $v = u^{1/2}$ is in N_D and $f = (v f_1)^2 + (v f_2)^2$. Q.E.D.

2. An example.

EXAMPLE. Let $f = x + (x^2 + y^2)^{1/2}$ in N_D , where $D = \mathbb{R}^2 - (0, 0)$. Here C is the negative y -axis. We can let L_1 be the positive x -axis and L_2 be the negative x -axis. Using the methods in [2], we find $f_1 = x - y + (x^2 + y^2)^{1/2}$, $f_2 = x + y + (x^2 + y^2)^{1/2}$. So we find $f_1^2 + f_2^2 =$

$2(x + (x^2 + y^2)^{1/2})^2 + 2y^2 = 4(x^2 + y^2 + x(x^2 + y^2)^{1/2}) = 4(x^2 + y^2)^{1/2}(x + (x^2 + y^2)^{1/2})$. This means v should be $2(x^2 + y^2)^{1/4}$ and finally $f_1 = (x - y + (x^2 + y^2)^{1/2})/2(x^2 + y^2)^{1/4}$, $f_2 = (x + y + (x^2 + y^2)^{1/2})/2(x^2 + y^2)^{1/4}$. Of course other representations are also possible.

3. No higher dimensional result. Finally, consider Motzkin's example, $p(x, y) = 1 + x^2(x^2 - 3)y^2 + x^2y^4$ in $R[x, y]$. Motzkin [3] showed that this polynomial is ≥ 0 everywhere but is not the sum any number of squares in $R[x, y]$ or see [4]. Moreover, Cassels, Ellison, and Pfister [1] have shown that p is not the sum of 3 squares in $R(x, y)$. For our purposes, we homogenize $p(x, y)$ to get $q(x, y, z) = z^6 + x^2(x^2 - 3z^2)y^2 + x^2y^4$. Let D be any domain R^3 containing $(0, 0, 0)$. We find with the above notation.

THEOREM 2. *The homogeneous polynomial $q(x, y, z)$ is not the sum of any number of squares elements of N_D . Moreover q is not the sum of three squares in the quotient field of N_D .*

Proof. Assume $q = \sum h_i^2$ where the h_i are in N_D . Since each h_i is analytic at $(0, 0, 0)$ we can expand in power series, and looking at the lowest degree terms which appear in any h_i (and which must be of degree 3) we get a new equation $\sum g_i^2 = q(x, y, z)$ homogeneous of degree 6. Divide by z^6 and dehomogenize and we get a contradiction to Motzkin's result. Finally, let $q = \sum_{i=1}^3 h_i^2/g^2$ where the h_i and g are in N_D . Then $g^2q = \sum h_i^2$. Again look at the lowest degree terms on each side and we get a homogeneous equation which when dehomogenized gives a contradiction to Cassels, Ellison, and Pfister, loc. cit. One should note that $q \geq 0$ for all x, y, z which is clear since when $z \neq 0$ we can dehomogenize and when $z = 0$, we get $x^4y^2 + x^2y^4$.

REFERENCES

1. J. W. S. Cassels, W. J. Ellison and A. Pfister, *On sums of squares and elliptic curves over function fields*, J. Number Theory, **3** (1971), 125-149.
2. G. Efroymsen, *Nash rings on planar domains*, Trans. Amer. Math. Soc. **249** (1979), 435-445.
3. T. S. Motzkin, *The arithmetic-geometric inequality*, in "Inequalities" (O. Shisha, editor) pp. 205-224, Academic Press, NY, 1967.
4. A. Pfister, *Quadratic forms over fields*, Proc. of Symposia in Pure Math. XX, 150-160, AMS, Providence, (1971).

Received July 10, 1980.

UNIVERSITY OF NEW MEXICO
ALBUQUERQUE NM 87131

