Ann. Scient. Éc. Norm. Sup.,
4^e série, t. 38, 2005, p. 751 à 772.

ON THE PYTHAGORAS NUMBERS OF REAL ANALYTIC SURFACES ☆

BY FRANCESCA ACQUISTAPACE, FABRIZIO BROGLIA, JOSÉ F. FERNANDO AND JESÚS M. RUIZ

Dedicated to Professor Enrique Outerelo, on the occasion of his 65th anniversary

ABSTRACT. – We show that (i) every positive semidefinite meromorphic function germ on a surface is a sum of 4 squares of meromorphic function germs, and that (ii) every positive semidefinite global meromorphic function on a normal surface is a sum of 5 squares of global meromorphic functions.

RÉSUMÉ. – Nous montrons que : (i) tout germe de fonction méromorphe semi-définie positive sur une surface réelle est une somme de quatre carrés de germes de fonctions méromorphes, et que : (ii) toute fonction méromorphe globale semi-définie positive sur une surface normale est une somme de cinq carrés de fonctions méromorphes globales.

1. Introduction

The famous 17th Hilbert Problem asks whether positive semidefinite (= psd) functions are always sums of squares, and in that case of how many. The two parts of this question are distinguished as the qualitative and the quantitative aspects of the problem. The specialists have studied them for different types of functions: polynomial, regular, Nash, analytic and smooth, and found full or partial solutions in most cases (see [5,3] or [17]). But of them all, analytic functions remain by far the most defying type. Indeed, although the qualitative aspect has been solved locally, i.e. for analytic germs, it is still open globally: the solution is only known for global analytic functions *on normal surfaces* ([2], see also [9]). Even worse is our quantitative information. Recall that the *Pythagoras number* of a ring A is the smallest integer $p \geqslant 1$ such that every sum of squares of A is a sum of p squares, or infinity if such an integer does not exist. In our setting, A is the ring $\mathcal{M}(X_x)$ of meromorphic function germs on a real analytic surface germ X_x, or the ring $\mathcal{M}(X)$ of global meromorphic functions on a normal real analytic surface X; we shorten the notation to

$$p(X_x) = p\big(\mathcal{M}(X_x)\big), \qquad p(X) = p\big(\mathcal{M}(X)\big).$$

☆ All authors supported by European RAAG HPRN-CT-2001-00271; first and second named authors also by Italian GNSAGA of INdAM and MIUR, third and fourth by Spanish GAAR BFM-2002-04797.

With this terminology, the quantitative problem is to estimate the Pythagoras numbers $p(X_x)$ and $p(X)$. We recall here that both Pythagoras numbers are always > 1.

Concerning germs, we have readily that $p(X_x) \leqslant 8$. To get this, one embeds X_x in $\mathbb{R}^3$ through a birational model. Then, any sum of squares on X_x is the restriction of one on $\mathbb{R}^3$, which is a sum of 8 squares of meromorphic function germs by [10]. Finally this sum of 8 squares restricts well to X_x: the equation of X_x in $\mathbb{R}^3$ is real, hence it can be factored out from the poles of all 8 addends. Thus we have a universal bound for $p(X_x)$, but it is not sharp. In fact, we will here prove the following result:

THEOREM 1.1. – *The Pythagoras number of the ring of meromorphic function germs on a real analytic surface germ X_x is $p(X_x) \leqslant 4$.*

In the global case, the situation is rather worse. As far, we only knew that the Pythagoras number is finite. The bound comes from the qualitative solution itself, and it is some non explicit function of the embedding dimension (see [2]). Unfortunately, the only true interest of such a bound is to confirm finiteness. In this paper we will improve much on this finiteness information as follows:

THEOREM 1.2. – *The Pythagoras number of the ring of global meromorphic functions on a normal real analytic surface X is $p(X) \leqslant 5$.*

This second theorem relies heavily on the way we prove the first. In fact, the easy bound 8 for $p(X_x)$ described earlier is of little use to deduce anything like 1.2: one needs the very delicate description of the sums of squares constructed for 1.1. Indeed, when a psd function is represented as a sum of squares of meromorphic functions, these meromorphic functions may have poles. Then, some of these poles can be eliminated by combining different representations, but others always remain: these form the so-called *bad set*. However new representations may require additional squares, which is not at all convenient when bounding Pythagoras numbers. What we will do here is keep bad sets *under control*, which means that *the poles of the summands of the sum of squares are among the zeros of the represented psd function*. And we should recall here that the standard control through the Positivstellensatz *gives no information on the number of squares*.

Thus, we will prove the following stronger theorems:

THEOREM 1.3. – *Let X be a real analytic surface germ, and $f : X \to \mathbb{R}$ a positive semidefinite analytic function germ. Then, there are analytic function germs $g, h_1, h_2, h_3, h_4 \in \mathcal{O}(X)$ such that*

$$g^2 f = h_1^2 + h_2^2 + h_3^2 + h_4^2$$

and g is a sum of squares with $\{g = 0\} \subset \{f = 0\}$.

THEOREM 1.4. – *Let X be a normal real analytic surface, and $f : X \to \mathbb{R}$ a positive semidefinite analytic function. Then, there are analytic functions $g, h_1, h_2, h_3, h_4, h_5 \in \mathcal{O}(X)$ such that*

$$g^2 f = h_1^2 + h_2^2 + h_3^2 + h_4^2 + h_5^2$$

and g is a sum of squares whose zero set $\{g = 0\}$ is a discrete subset of the zero set $\{f = 0\}$ of f.

In case X is non singular, one can get rid of the denominator [9], and in general, one can get rid of non singular points in the bad set. To do it, one finds two different representations whose bad sets only share singular points of X, and add them both. This is quite technically demanding, but no new idea is behind. Furthermore, the number of squares worsen to the double, hence we will not dive here into more details.

Anyway, our proofs require a careful preparatory job. In Section 2 we discuss sums of squares of totally positive elements, much inspired by Mahe's results in [14]. Section 3 is devoted to the proof of Theorem 1.3, which besides Section 2 needs a relative algebrization lemma in the style of classification of singularities, based on Tougeron's Implicit Function Theorem. The global bound for normal surfaces is given in Section 4, as a globalization of the local one. This involves, on the one hand, some techniques that will be further developed in [1], and, on the other, removing real analytic divisors as in [2].

One final word is in order concerning the most general application of our arguments. In fact, and to discard a little the technical toll of some of them, we have restricted our global statement to normal surfaces, in accordance with [2]. But while in that paper the restriction was relevant to prove the Artin–Lang property, here we could quite straightforwardly obtain Theorem 1.2 *for real coherent surfaces with isolated singularities*.

2. Totally positive sums of squares

The purpose here is to study the representation of totally positive elements as sums of squares in certain relative polynomial rings. This will be used later to control bad sets. The idea is that: (i) a psd element $f \in A$ is totally positive in $A[1/f]$, (ii) a sum of squares in $A[1/f]$ becomes a sum of squares in A after multiplying by an even power of f, and (iii) this multiplication does not add zeros other than those of f. This is inspired in [13, 7.3], and we follow the notation and terminology introduced there.

Consider the ring of power series $\mathbb{R}\{t\}$ in one variable t and its field of fractions $\mathbb{R}(\{t\})$, as well as the ring $\mathbb{C}\{t\}$ and the field $\mathbb{C}(\{t\})$. We are interested in rings A which are finitely generated algebras over $\mathbb{R}\{t\}$, that is, $A = \mathbb{R}\{t\}[z]/\mathfrak{a}$ for some ideal $\mathfrak{a} \subset \mathbb{R}\{t\}[z]$, with additional variables $z = (z_1, \ldots, z_m)$. Given such a presentation of A, let us denote by $\mathfrak{p}_i$ the minimal primes of $\mathfrak{a}$ in $\mathbb{R}\{t\}[z]$, so that $\sqrt{\mathfrak{a}} = \bigcap_i \mathfrak{p}_i$. Then, the minimal primes of (0) in A are $\mathfrak{a}_i = \mathfrak{p}_i/\mathfrak{a}$, that is: $\sqrt{(0)} = \bigcap_i \mathfrak{a}_i$. Let K be the total ring of fractions of the reduction $A/\sqrt{(0)}$, and for each i, let K_i be the field of fractions of $A_i = A/\mathfrak{a}_i = \mathbb{R}\{t\}[z]/\mathfrak{p}_i$. We have:

$$\operatorname{ht}(\mathfrak{a}) = \min_i \operatorname{ht}(p_i), \qquad \dim(A) = \max_i \dim(A_i), \qquad K = \prod_i K_i.$$

We call the A_i's *the reduced branches of* A, and use systematically the notations above.

COHEIGHT 2.1. – Let A be a finitely generated algebra over $\mathbb{R}\{t\}$, say $A = \mathbb{R}\{t\}[z]/\mathfrak{a}$. We define the *coheight of* A by

$$\delta(A) = m + 1 - \operatorname{ht}(\mathfrak{a}).$$

In terms of the reduced branches A_i of A we have:

$$\delta(A) = m + 1 - \operatorname{ht}(\mathfrak{a}) = \max_i \bigl\{m + 1 - \operatorname{ht}(\mathfrak{p}_i)\bigr\} = \max_i \delta(A_i).$$

For instance, $\delta(\mathbb{R}(\{t\})) = \delta(\mathbb{R}(\{t\})[z]/(zt-1)) = 1$.

This invariant $\delta(A)$ will be essential to deal with sums of squares with controlled bad sets. But first of all we must check that δ does not depend on the chosen presentation $\mathbb{R}\{t\}[z]/\mathfrak{a}$ of A. For this we need the following:

LEMMA 2.2. – *Let* $\mathfrak{m} \subset \mathbb{R}\{t\}[z]$ *be a maximal ideal.*
(1) *If* $t \in \mathfrak{m}$, *then* $\operatorname{ht}(\mathfrak{m}) = m + 1$ *and* $\mathbb{R}\{t\}[z]/\mathfrak{m}$ *is isomorphic to* $\mathbb{R}$ *or* $\mathbb{C}$.
(2) *If* $t \notin \mathfrak{m}$, *then* $\operatorname{ht}(\mathfrak{m}) = m$ *and* $\mathbb{R}\{t\}[z]/\mathfrak{m}$ *is isomorphic to* $\mathbb{R}(\{t\})$ *or* $\mathbb{C}(\{t\})$.

Proof. – (1) If $t \in \mathfrak{m}$ then $\mathbb{R}\{t\}[z]/\mathfrak{m} = \mathbb{R}[z]/\mathfrak{m} \cap \mathbb{R}[z]$, which is a field isomorphic to $\mathbb{R}$ or $\mathbb{C}$. Moreover, since all maximal ideals of $\mathbb{R}[z]$ have height m, we conclude:

$$m = \mathrm{ht}\big(\mathfrak{m} \cap \mathbb{R}[z]\big) = \mathrm{ht}\big(\mathfrak{m}/t \cdot \mathbb{R}\{t\}[z]\big) = \mathrm{ht}(\mathfrak{m}) - \mathrm{ht}\big(t \cdot \mathbb{R}\{t\}[z]\big) = \mathrm{ht}(\mathfrak{m}) - 1.$$

(2) Suppose $t \notin \mathfrak{m}$. Then t is a unit in $\mathbb{R}\{t\}[z]/\mathfrak{m}$, and

$$\mathbb{R}\{t\}[z]/\mathfrak{m} = \mathbb{R}\big(\{t\}\big)[z]/\mathfrak{m}\mathbb{R}\big(\{t\}\big)[z].$$

Now, the field $\mathbb{R}\{t\}[z]/\mathfrak{m}$ is a finitely generated algebraic extension of $\mathbb{R}(\{t\})$, and there exists an integer $p \geqslant 1$ such that $\mathbb{R}\{t\}[z]/\mathfrak{m} \cong \mathbb{K}(\{t^{1/p}\}) \cong \mathbb{K}(\{s\})$, where $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. Since $\mathfrak{m}\mathbb{R}(\{t\})[z]$ is a maximal ideal of $\mathbb{R}(\{t\})[z]$, it has height m. Moreover, since $\mathbb{R}\{t\}[z]_{\mathfrak{m}} = \mathbb{R}(\{t\})[z]_{\mathfrak{m}\mathbb{R}(\{t\})[x]}$ we conclude that $\mathrm{ht}(\mathfrak{m}) = \mathrm{ht}(\mathfrak{m}\mathbb{R}(\{t\})[x]) = m$. $\square$

This leads to the following computation, which shows the coheight does not depend on the presentation:

PROPOSITION 2.3. – *Consider the algebra $A = \mathbb{R}\{t\}[z]/\mathfrak{a}$ and its reduced branches A_i. Then*:

$$\delta(A_i) = \begin{cases} \dim(A_i), & \textit{if } t \textit{ is not a unit in } A_i, \\ \dim(A_i) + 1, & \textit{otherwise.} \end{cases}$$
$$= \begin{cases} \dim(A_i), & \textit{if some residue field of } A_i \textit{ is } \mathbb{R} \textit{ or } \mathbb{C}, \\ \dim(A_i) + 1, & \textit{otherwise.} \end{cases}$$

In particular, $\delta(A) = \max_i \delta(A_i)$ does not depend on the presentation of A.

Proof. – First suppose that t is not a unit mod $\mathfrak{p}_i$. This means that some maximal ideal $\mathfrak{m}$ of $\mathbb{R}\{t\}[z]$ containing $\mathfrak{p}_i$ must contain t, and, by 2.2, have height $m+1$ and residue field $\mathbb{R}$ or $\mathbb{C}$. Hence,

$$\mathrm{ht}(\mathfrak{m}/\mathfrak{p}_i) = \mathrm{ht}(\mathfrak{m}) - \mathrm{ht}(\mathfrak{p}_i) = m + 1 - \mathrm{ht}(\mathfrak{p}_i) = \delta(A_i).$$

As the height of all maximal ideals is $\leqslant m+1$, we conclude:

$$\dim\big(\mathbb{R}\{t\}[z]/\mathfrak{p}_i\big) = \sup_{\mathfrak{m} \supset \mathfrak{p}_i} \mathrm{ht}(\mathfrak{m}/\mathfrak{p}_i) = \delta(A_i).$$

Contrarily, if t is a unit mod $\mathfrak{p}_i$, then no maximal ideal $\mathfrak{m} \supset \mathfrak{p}_i$ contains t, hence all have height m, and, by 2.2, residue field $\mathbb{R}(\{t\})$ or $\mathbb{C}(\{t\})$. Thus, $\dim(\mathbb{R}\{t\}[z]/\mathfrak{p}_i) = m - \mathrm{ht}(\mathfrak{p}_i) = \delta(A_i) - 1$. $\square$

Once presentations can be disregarded, the elementary properties of δ follow readily from the definition. We will need these two bounds:

PROPOSITION 2.4. – *Let A be as above. We have*:
(i) *If $v \in A$ is neither a unit in A nor a zero divisor in $A/\sqrt{(0)}$, then $\delta(A/vA) \leqslant \delta(A) - 1$.*
(ii) $\delta(A[T]) \leqslant \delta(A) + 1$.

Proof. – By the hypotheses in (i) v generates a proper ideal, and $\mathrm{ht}((v) + \mathfrak{a}) > \mathrm{ht}(\mathfrak{a})$, and the assertion is clear. On the other hand, (ii) follows readily from the good dimension properties of the extension $A \subset A[T]$. $\square$

We come now to the crucial link between coheight and sums of squares:

PROPOSITION 2.5. – *Let A be a finitely generated algebra over $\mathbb{R}\{t\}$ and K the total ring of fractions of its reduction $A/\sqrt{(0)}$. Then*

$$p(K) \leqslant 2^{\delta(A)}.$$

Proof. – We consider the reduced branches A_i of A and their fields of quotients K_i. As $p(K) = \max_i p(K_i)$ and $\delta(A) = \max_i \delta(A_i)$, it suffices to see that

$$p(K_i) \leqslant 2^{\delta(A_i)}.$$

Firstly, suppose $t \in \mathfrak{p}_i$. Then $A_i = \mathbb{R}\{t\}[z]/\mathfrak{p}_i = \mathbb{R}[z]/\mathfrak{p}_i \cap \mathbb{R}[z]$ is a finitely generated $\mathbb{R}$-algebra, and as is well known, $p(K_i) \leqslant 2^{d_i}$, where $d_i = \dim(A_i)$. Moreover, in this case, t is not a unit in A_i, hence $\dim(A_i) = \delta(A_i)$ by 2.3, and we are done.

Next, suppose $t \notin \mathfrak{p}_i$. In this case, K_i contains the field $\mathbb{R}(\{t\})$, and we can easily compute the transcendence degree d_i of this extension. Indeed, note that K_i is also the quotient field of $\mathbb{R}(\{t\})[z]/\mathfrak{p}_i\mathbb{R}(\{t\})[z]$, and $\mathfrak{p}_i\mathbb{R}(\{t\})[z]$ is a proper prime ideal of $\mathbb{R}(\{t\})[z]$ of height $\mathrm{ht}(\mathfrak{p}_i)$. Consequently

$$d_i = \dim\big(\mathbb{R}\big(\{t\}\big)[z]/\mathfrak{p}_i\mathbb{R}\big(\{t\}\big)[z]\big) = m - \mathrm{ht}\big(\mathfrak{p}_i\mathbb{R}\big(\{t\}\big)[z]\big) = m - \mathrm{ht}(\mathfrak{p}_i) = \delta(A_i) - 1.$$

Hence, $L_i = K_i[\sqrt{-1}\,]$ has transcendence degree $d_i = \delta(A_i) - 1$ over $\mathbb{C}(\{t\})$.

Recall now that a field L is C_k if every homogeneous polynomial over L of degree d in more than d^k variables has some non trivial solution in L [7, 1.4]. For instance, $\mathbb{C}(\{t\})$ is a C_1 field (this is a straightforward consequence of [7, 4.8] and M. Artin's Approximation Theorem, [11]). Furthermore, this implies, by [7, 3.6], that L_i is a C_{d_i+1} field. Once we know this, we conclude by Pfister's theorem ([16], [12, XI.1.9]) that any sum of squares of K_i can be represented as a sum of 2^{d_i+1} squares of K_i. But $d_i + 1 = \delta(A_i)$, which completes the proof. □

After the preceding preparation, consider the *real spectrum* $\mathrm{Spec}_r(A)$ *of* A, and say as usual that an element $f \in A$ is *positive semidefinite* if $f(\alpha) \geqslant 0$ (respectively *totally positive* if $f(\alpha) > 0$) for every prime cone $\alpha \in \mathrm{Spec}_r(A)$. Thus we are ready to obtain the main result of this section:

THEOREM 2.6. – *Let A be a finitely generated algebra over $\mathbb{R}\{t\}$. Let $f \in A$ be totally positive. Then there exist a sum of squares $a = a_1^2 + \cdots + a_r^2$ in A such that $(1+a)^2 f$ is a sum of 2^δ squares in A, where $\delta = \delta(A)$.*

In order to ease the writing of what follows we will use the standard notation due to Pfister: $f = \boxed{r}$ means that f is a sum of r squares in A; when several $\boxed{r}$'s appear in the same formula, they need not be the same. For instance, the well-known fact that in a field a product of sums of 2^d squares is again a sum of 2^d squares can be formulated as

$$\boxed{2^d}\,\boxed{2^d} = \boxed{2^d}.$$

Theorem 2.6 will follow from the following variation:

PROPOSITION 2.7. – *Let $f \in A$ and $\delta = \delta(A)$ be as above. Then there exists a totally positive element $u \in A$ such that*

$$\boxed{2^\delta}\, f = u^2 + \boxed{2^{\delta-1}}.$$

Proof. – We first show that the assertion follows for A if it holds for $A/\sqrt{(0)}$. Notice here that since the real spectrum does not change $\operatorname{mod} \sqrt{(0)}$, $h \in A$ is totally positive if and only if it is totally positive $\operatorname{mod} \sqrt{(0)}$. Also recall that $\delta(A) = \delta(A/\sqrt{(0)}\,)$. Now suppose that

$$\boxed{2^\delta} f = u^2 + \boxed{2^{\delta-1}} \mod \sqrt{(0)}.$$

for some totally positive element $u \in A$. Then

$$\boxed{2^\delta} f = u^2 + \boxed{2^{\delta-1}} - \theta$$

for some nilpotent element $\theta \in A$.

Now, we have the following identity:

$$(x+y)^2(x+y/4) = x^3 + y\left(\frac{3x+y}{2}\right)^2$$

(just expand both sides), which setting $x = u^2$ and $y = \boxed{2^{\delta-1}} - \boxed{2^\delta} f$ gives

$$\theta^2 h = u^6 + \left(\boxed{2^{\delta-1}} - \boxed{2^\delta} f\right) g^2 = u^6 + \boxed{2^{\delta-1}} - \boxed{2^\delta} f,$$

hence

$$\boxed{2^\delta} f = v^2 + \boxed{2^{\delta-1}} - \theta^2 h,$$

where $v = u^3$ is totally positive and $h \in A$. Since θ is nilpotent, after several applications of the same trick, the θ addend becomes 0, and we get the required inequality in A.

After this, we can suppose A reduced, and will prove the statement by induction on δ. We use the usual notations: $\mathfrak{p}_i$, A_i, K_i, and recall that $\delta = \max_i \delta(A_i)$.

If $\delta = 0$, by 2.3 $\dim(A_i) = 0$ and $A_i = K_i$ is either $\mathbb{R}$ or $\mathbb{C}$. As A is reduced, $A = \prod_i K_i$, hence f is in fact a square in A.

Suppose now $\delta \geqslant 1$. By [14, 2.3], there exists a nonzero divisor $g \in A$ such that

$$A[1/g] \cong \prod_i B_i, \quad B_i = A_i[1/g].$$

Note that the quotient field of the domain $B_i = A_i[1/g]$ is the same K_i, and by 2.5 f is a sum of 2^δ squares in K_i. Hence we can write $f = \boxed{2^{\delta-1}} + \boxed{2^{\delta-1}}$, and multiplying by the first sum of squares

$$\boxed{2^{\delta-1}} f = v_i^2 + \boxed{2^{\delta-1}}, \quad 0 \neq v_i \in K_i$$

(recall that in K_i it holds $\boxed{2^d}\,\boxed{2^d} = \boxed{2^d}$). Clearing denominators we can suppose the above equation holds in B_i. Consequently, in $A[1/g] = \prod_i B_i$ we have

$$\boxed{2^{\delta-1}} f = v^2 + \boxed{2^{\delta-1}},$$

where $v \in A[1/g]$ is not a zero divisor. Multiplying by a big enough even power of g, we obtain a similar formula in A

$$(\bullet) \qquad \boxed{2^{\delta-1}}_1 f = v^2 + \boxed{2^{\delta-1}}_2,$$

where $v \in A$ is not a zero divisor.

Now, if v is a unit in A, dividing by v^2 the equation becomes $\boxed{2^{\delta-1}}_1 f = 1 + \boxed{2^{\delta-1}}_2$, and we are done. Hence, we may assume that v is not a unit in A, and by 2.4(i), $\delta(A/vA) \leqslant \delta - 1$. Then, by induction,

(••)
$$\boxed{2^{\delta-1}} f = w^2 + \boxed{2^{\delta-2}} \bmod v,$$

where $w \in A$ is totally positive mod v. This can be arranged for w to be totally positive in A. Indeed, as w is totally positive in A/vA, the Positivstellensatz gives an expression

$$\boxed{p} w = 1 + \boxed{q} \bmod v,$$

and multiplying (••) by the square of $\boxed{p}$ we can replace w by $1 + \boxed{q}$, which is clearly totally positive in A.

Once this is settled, we have:

$$\lambda v = w^2 + \boxed{2^{\delta-2}} - \boxed{2^{\delta-1}} f = a - bf,$$

for some $\lambda \in A$, $a = w^2 + \boxed{2^{\delta-2}}$ totally positive, $b = \boxed{2^{\delta-1}}$. Multiplying (•) by λ^2 and substituting λv by its value we get

$$\boxed{2^{\delta-1}}_1 f = (a - bf)^2 + \boxed{2^{\delta-1}}_2 = (a + bf)^2 - 4abf + \boxed{2^{\delta-1}}_2.$$

Modifying a little this equation we get:

$$\left(\boxed{2^{\delta-1}}_1 + 4ab\right) f = u^2 + \boxed{2^{\delta-1}}_2,$$

where $u = a + bf$ is totally positive. In order to complete the argument we must still modify the term $\boxed{2^{\delta-1}}_1 + 4ab$ to have a sum of 2^δ squares. To that end, it is enough to show the following: *there is a totally positive element $\gamma \in A$ such that $\gamma^2 ab = \boxed{2^{\delta-1}}$.*

This is in turn a statement about matrices. Indeed, as $b = \boxed{2^{\delta-1}}$, we can write:

$$\gamma^2 ab = (b_1, \ldots, b_r)\left(\gamma^2 a I\right) \begin{pmatrix} b_1 \\ \vdots \\ b_r \end{pmatrix} \quad \text{where } r = 2^{\delta-1}$$

and we only need that $\gamma^2 a I = M^t M$ for some $r \times r$ matrix M with coefficients in A. This we prove by induction on $d = \delta - 1$.

If $d = 1$, $a = w^2 + \theta^2$, and the solution is $\gamma = 1$ and $M = \left(\begin{smallmatrix} w & -\theta \\ \theta & w \end{smallmatrix}\right)$.

Assume $d \geqslant 2$, and let $a_1 = (\frac{3}{5}w)^2 + \boxed{2^{d-2}}$ and $a_2 = (\frac{4}{5}w)^2 + \boxed{2^{d-2}}$ such that $a = a_1 + a_2$. Note that since w is totally positive, a_1 and a_2 are totally positive too. By induction, there exist totally positive elements γ_1, γ_2 and matrices M_1 and M_2 (of suitable order) such that $M_i^t M_i = \gamma_i^2 a_i I$. Take

$$M = \begin{pmatrix} \gamma_2^2 a_2 M_1 & -\gamma_1 \gamma_2 a_2 M_2 \\ \gamma_1 \gamma_2 a_2 M_2 & M_2 M_1^t M_2 \end{pmatrix}$$

and $\gamma = \gamma_1\gamma_2^2 a_2$ which is a totally positive element of A. A straightforward computation shows these are the M and γ we sought. $\square$

Now, we are ready for the

Proof of Theorem 2.6. – We must find a formula of the type

$$\big(1+\boxed{r}\big)f = \boxed{2^\delta},$$

and what we have by Proposition 2.7 is

$$\boxed{2^\delta}\,f = u^2 + \boxed{2^{\delta-1}}.$$

We write

$$af = \boxed{2^\delta}_1,$$

where $a = \boxed{2^\delta}$ is totally positive, as so are f and u. Now, arguing as at the end of the latter proof, we find a totally positive element γ, such that $\gamma^2 aI = M^tM$ for a suitable $2^\delta \times 2^\delta$ matrix M. Hence

$$\gamma^2 a\boxed{2^\delta}_1 = \boxed{2^\delta} \quad \text{and} \quad \gamma^2a^2f = \gamma^2a(af) = \gamma^2a\boxed{2^\delta}_1 = \boxed{2^\delta}.$$

Here the element γa is totally positive, and by the Positivstellensatz, we can write

$$\boxed{r}\,\gamma a = 1 + \boxed{r}.$$

Consequently,

$$\big(1+\boxed{r}\big)^2 f = \big(\boxed{r}\,\gamma a\big)^2 f = \boxed{r}^2\gamma^2a^2f = \boxed{r}^2\,\boxed{2^\delta} = \boxed{2^\delta},$$

as wanted. $\square$

3. Analytic surface germs

The purpose of this section is to prove Theorem 1.3, crucial for the proof of Theorem 1.4. The arguments, somehow inspired in [10], rely heavily on the previous section.

We denote by $\mathbb{R}\{x\}$ the ring of convergent power series in $x = (x_1,\dots,x_n)$ with real coefficients, seen also as the ring of analytic function germs at the origin in $\mathbb{R}^n$; its maximal ideal is $(x) = (x_1,\dots,x_n)\mathbb{R}\{x\}$. Let $X \subset \mathbb{R}^n$ be an analytic set germ (at the origin always), and consider the ring $\mathcal{O}(X)$ of analytic function germs on X. Explicitly, $\mathcal{O}(X) = \mathbb{R}\{x\}/J$, where J is the ideal of (all analytic function germs vanishing on) X. Of course, *positive semidefinite on* X means $\geqslant 0$ on X. Any ideal $I \subset \mathbb{R}\{x\}$ defines a *zero set germ* $X = \mathcal{Z}(I)$, and the real Nullstellensatz says that the ideal J of X is the *real radical* $\sqrt[r]{I}$ of I; in particular, J is a radical ideal. Similarly, the ring $\mathbb{C}\{x\}$ of convergent complex power series with complex coefficients is seen as the ring of holomorphic function germs at the origin in $\mathbb{C}^n$. As above, every ideal $I \subset \mathbb{C}\{x\}$ defines a complex analytic set germ $Z \subset \mathbb{C}^n$, but here the Nullstellensatz is simpler: the ring $\mathbb{C}\{x\}/J$ of germs of holomorphic functions on Z is defined by the radical $J = \sqrt{I}$. We will resource to *complexification* via the canonical inclusion $\mathbb{R}\{x\} \subset \mathbb{C}\{x\}$. Any element $h \in \mathbb{C}\{x\}$ can be uniquely written as $h = f + \sqrt{-1}g$, with $f, g \in \mathbb{R}\{x\}$, and its *conjugate* is

$\bar{h} = f - \sqrt{-1}g$; f and g are respectively the *real* and the *imaginary* part of h. Given an ideal $I \subset \mathbb{R}\{x\}$, we denote $\widetilde{I} = I\mathbb{C}\{x\}$; these extended ideals are *invariant* by conjugation. Given an analytic set germ X, we denote $\widetilde{X} = \mathcal{Z}(\widetilde{J})$, where J is the ideal of X. Note that since J is a radical, so is $\widetilde{J}$, and this is essential: if $X = \mathcal{Z}(I)$, it may well happen that $Z = \mathcal{Z}(\widetilde{I})$ is not $\widetilde{X}$. For generalities concerning all of this, we refer to [15,11,18].

After this standard introduction to fix notations and terminology, we come to our fundamental algebrization result:

PROPOSITION 3.1. – *Let $X \subset \mathbb{R}^n$ be a singular surface germ at the origin whose ideal we denote by J. Let $f \in \mathbb{R}\{x\}$ be positive semidefinite on the germ of $\mathbb{R}^n$ at the origin and such that $f(0) = 0$. Suppose furthermore that f does not vanish on any irreducible component of X of dimension 2. Then after an analytic change of coordinates there are*:

(i) *A sum of squares of analytic function germs $h \in J$,*
(ii) *$f' \in \mathbb{R}\{x_1\}[x_2,\dots,x_n]$, and*
(iii) *$Q_3,\dots,Q_n \in \mathbb{R}\{x_1\}[x_2,\dots,x_n] \cap J$,*

such that

(1) $\operatorname{ht}((Q_3,\dots,Q_n)\mathbb{R}\{x\}) = n-2$, *and*
(2) $(f+h) - f' \in (Q_3,\dots,Q_n)\mathbb{R}\{x\}$ (*hence*, $f' = f \bmod J$).

Proof. – Let $X_1,\dots,X_s$ be the irreducible components of dimension 2 of X, so that $X = X_1 \cup \cdots \cup X_s \cup Y$, where Y is an analytic curve germ. The ideal J has height $n-2$, and its associated primes of height $n-2$ are the ideals of the X_i's. Then, $\widetilde{J} = J\mathbb{C}\{x\}$ is the ideal of the complexification $\widetilde{X}$ of X, and $\widetilde{X} = \widetilde{X}_1 \cup \cdots \cup \widetilde{X}_s \cup \widetilde{Y}$.

Step I. First of all, after a linear change of coordinates, we find square free Weierstrass polynomials $P_k \in \mathbb{R}\{x_1,x_2\}[x_k] \cap J$, $k = 3,\dots,n$, such that $\operatorname{ht}(P_3,\dots,P_n) = n-2$ (Rückert's Parametrization, [18, II.2.3]). In particular, the discriminant $\Delta_k \in \mathbb{R}\{x_1,x_2\}$ is not zero. We denote $J' = (P_3,\dots,P_n)\mathbb{R}\{x\}$ and consider the extension $\widetilde{J}' = J'\mathbb{C}\{x\}$. The ideal J' needs not be real, but we look at its complex zero set germ $Z = \mathcal{Z}(\widetilde{J}') \subset \mathbb{C}^n$; clearly $Z \supset \widetilde{X}$, but these two complex germs need not coincide. Since $\operatorname{ht}(J') = n-2$, also $\operatorname{ht}(\widetilde{J}') = n-2$, and $\dim(Z) = 2$. Consequently, the complexifications $\widetilde{X}_i$ are irreducible components of Z, but Z may very well have other irreducible components Z_ℓ of dimension 2. What we know is that no such Z_ℓ is contained in $\widetilde{X}$, so that there is $g_\ell \in \widetilde{J}$ which does not vanish on Z_ℓ. As $\widetilde{J}$ is an extended ideal, we can choose $g_\ell \in J$.

On the other hand, as the P_k's are monic polynomials, the holomorphic map germ

$$\pi_k : D_k = \left\{ \frac{\partial P_k}{\partial x_k} = P_3 = \cdots = P_n = 0 \right\} \to \mathbb{C}^2$$

induced by the linear projection $x \mapsto (x_1,x_2)$ is a finite map germ, so that $\dim(D_k) = \dim(\pi_k(D_k))$. But $\pi_k(D_k) \subset \{\Delta_k = 0\}$; note that $\{\Delta_k = 0\}$ may be empty. We conclude

$$\dim\left(\left\{ \frac{\partial P_k}{\partial x_k} = 0 \right\} \cap Z\right) \leqslant 1 \quad \text{(this set may be empty).}$$

Step II. Now we construct a sum of squares $h \in J$ such that $g = f + h$ does not vanish on any irreducible component Z_ℓ. Note that since f is psd and does not vanish on any X_i, the germ $g = f + h$ cannot vanish on any $\widetilde{X}_i$ either. We proceed by induction and construct a sum of squares $h_1^2 + \cdots + h_\ell^2$ with $h_i \in J$ such that $f_\ell = f + h_1^2 + \cdots + h_\ell^2$ does not vanish on any irreducible component $Z_1,\dots,Z_\ell$. Of course $f_0 = f$. Assume $\ell \geqslant 1$ and that $f_{\ell-1}$ has been constructed. If $f_{\ell-1}$ does not vanish on Z_ℓ, let $h_\ell = 0$. Otherwise, we take $g_\ell \in J$ which does not vanish on Z_ℓ (step I) and set $h_\ell = g_\ell^{m_\ell}$. We can choose m_ℓ large enough so that $f_\ell = f_{\ell-1} + h_\ell^2$ does not vanish on $Z_1, \ldots, Z_{\ell-1}$. Indeed, by Krull's theorem

$$\widetilde{J}'_i = \bigcap_m \widetilde{J}'_i + (g_\ell)^{2m}, \quad \text{where } \widetilde{J}'_i \text{ is the ideal of the complex germ } Z_i.$$

Step III. In order to apply Tougeron's Implicit Functions Theorem, consider the matrix

$$\lambda = \left(\begin{array}{ccc|ccc|ccc|c|ccc}
\frac{\partial g}{\partial x_1} & \cdots & \frac{\partial g}{\partial x_n} & P_3 & \cdots & P_n & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\
\frac{\partial P_3}{\partial x_1} & \cdots & \frac{\partial P_3}{\partial x_n} & 0 & \cdots & 0 & P_3 & \cdots & P_n & \cdots & 0 & \cdots & 0 \\
\vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots & & \vdots & & \vdots \\
\frac{\partial P_n}{\partial x_1} & \cdots & \frac{\partial P_n}{\partial x_n} & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & P_3 & \cdots & P_n
\end{array}\right)$$

and let $I \subset \mathbb{R}\{x\}$ be the ideal generated by the $(n-1) \times (n-1)$ minors of λ. We claim that $\operatorname{ht}(I) \geqslant n-1$. Since heights do not change by complexification, it is enough to see that $\operatorname{ht}(\widetilde{I}) \geqslant n-1$, or that the complex analytic set germ $Z' = \mathcal{Z}(\widetilde{I})$ has dimension $\leqslant 1$. We argue by way of contradiction.

Since $P_3^{n-1}, \ldots, P_n^{n-1} \in I$, we have $Z = \{P_3 = \cdots = P_n = 0\} \supset Z'$, and $\dim(Z') \leqslant \dim(Z) = 2$. Suppose $\dim(Z') = 2$. Then Z and Z' share some irreducible component T of dimension 2 (either one $\widetilde{X}_i$ or one Z_ℓ). By step II, we know that g does not vanish on T; since $g(0) = 0$, g is not constant on T. But T is irreducible, hence g is not constant on any nonempty open subset U of the regular locus T^0 of T, and we conclude that

$$\mathrm{d}g = \frac{\partial g}{\partial x_1}\mathrm{d}x_1 + \cdots + \frac{\partial g}{\partial x_n}\mathrm{d}x_n$$

cannot vanish on the tangent bundle τU. Contrarily, since all P_k's vanish on T,

$$\mathrm{d}P_k = \frac{\partial P_k}{\partial x_1}\mathrm{d}x_1 + \cdots + \frac{\partial P_k}{\partial x_n}\mathrm{d}x_n$$

do vanish on τU. We know from step I that $\dim(\{\frac{\partial P_k}{\partial x_k} = 0\} \cap Z) \leqslant 1$, so that

$$U = T^0 \Big\backslash \Big\{ \prod_k \frac{\partial P_k}{\partial x_k} = 0 \Big\}$$

is open and nonempty. As P_k only has the variables x_1, x_2 and x_k, it holds

$$\prod_k \frac{\partial P_k}{\partial x_k} = \det \begin{pmatrix} \frac{\partial P_3}{\partial x_3} & \cdots & \frac{\partial P_3}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial P_n}{\partial x_3} & \cdots & \frac{\partial P_n}{\partial x_n} \end{pmatrix}$$

and, consequently, the $\mathrm{d}P_k$'s are independent on U. On the other hand, on $Z' \supset U$ all $(n-1)\times(n-1)$ minors of the matrix λ vanish, so that in particular its submatrix

$$\begin{pmatrix} \frac{\partial g}{\partial x_1} & \cdots & \frac{\partial g}{\partial x_n} \\ \frac{\partial P_3}{\partial x_1} & \cdots & \frac{\partial P_3}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial P_n}{\partial x_1} & \cdots & \frac{\partial P_n}{\partial x_n} \end{pmatrix}$$

has rank $n-2$. But on U the $\mathrm{d}P_k$'s are independent, hence $\mathrm{d}g$ depends on them, and must vanish where they do, namely on τU.

This contradiction shows that Z' must have dimension $\leqslant 1$, as wanted.

Step IV. Consider the ideal $(x)I^2$. Since g is psd and $g(0)=0$, its derivatives $\partial g/\partial x_i$ vanish all at 0, and so does the first row of the matrix λ. Hence $I \subset (x)$, and we have $I^3 \subset (x)I^2 \subset I$, so that $\mathrm{ht}((x)I^2) \geqslant n-1$. Furthermore, since $P_k^{n-1} \in I$, we have $P_k^{3(n-1)} \in (x)I^2$, and we see that the homomorphism $\mathbb{R}\{x_1,x_2\} \to \mathbb{R}\{x\}/(x)I^2$ is finite. Since $\mathrm{ht}((x)I^2) \geqslant n-1$, the homomorphism cannot be injective, and $\mathfrak{a} = (x)I^2 \cap \mathbb{R}\{x_1,x_2\} \neq 0$. Next, we look at the ring $\mathbb{R}\{x_1,x_2\}/\mathfrak{a}$, and after a linear change of the variables x_1,x_2 (which does not modify all preceding constructions), the homomorphism $\mathbb{R}\{x_1\} \to \mathbb{R}\{x_1,x_2\}/\mathfrak{a}$ is finite. By composition, also the homomorphism $\mathbb{R}\{x_1\} \to \mathbb{R}\{x\}/(x)I^2$ is finite, and each class $x_j \bmod (x)I^2$, $j \geqslant 2$, verifies a monic equation with coefficients in $\mathbb{R}\{x_1\}$. Thus we find monic polynomials

$$\Phi_k(x_1,x_j) \in \mathbb{R}\{x_1\}[x_j] \cap (x)I^2.$$

Each Φ_j is a regular power series of some order with respect to x_j, hence after successive Weierstrass divisions of g and $P_3,\dots,P_n$ by the Φ_j's, we find $f', Q_3,\dots,Q_n \in \mathbb{R}\{x_1\}[x_2,\dots,x_3]$ such that

$$\begin{aligned} g &\equiv f' \bmod (\Phi_2,\dots,\Phi_n), \\ P_k &\equiv Q_k \bmod (\Phi_2,\dots,\Phi_n), \quad k=3,\dots,n. \end{aligned}$$

Now add to the x_i's new variables y_i, t_k and z_{jk}, and consider the system of equations

$$\begin{cases} 0 = F_0(x_i,y_i,t_k,z_{jk}) = g(x+y) + \sum_{k=3}^n t_k\big(P_k(x+y) + \sum_{j=3}^n z_{jk}P_j(x+y)\big) - f'(x), \\ 0 = F_3(x_i,y_i,t_k,z_{jk}) = P_3(x+y) + \sum_{j=3}^n z_{j3}P_j(x+y) - Q_3(x), \\ \quad\vdots \\ 0 = F_n(x_i,y_i,t_j,z_{kj}) = P_n(x+y) + \sum_{j=3}^n z_{jn}P_j(x+y) - Q_n(x). \end{cases}$$

One sees immediately that the Jacobian matrix of this system at $y_i = t_k = z_{jk} = 0$ is the matrix λ in step III, and it holds

$$\begin{aligned} F_0(x,0) &= g - f' \in (\Phi_2,\dots,\Phi_n) \subset (x)I^2 \\ F_k(x,0) &= P_k - Q_k \in (\Phi_2,\dots,\Phi_n) \subset (x)I^2, \quad k=3,\dots,n. \end{aligned}$$

Whence, we can apply Tougeron's Implicit Functions Theorem ([19], [18, V.1]) to find a solution $y_i(x), t_j(x), z_{kj}(x) \in (x)I$ of the system $F_0 = F_3 = \cdots = F_n = 0$. This gives:

$$\begin{cases} f'(x) = g\big(x + y(x)\big) + \sum_{k=3}^n t_k(x)\Big(P_k\big(x + y(x)\big) + \sum_{j=3}^n z_{jk}(x)P_j\big(x + y(x)\big)\Big), \\ Q_3(x) = P_3\big(x + y(x)\big) + \sum_{j=3}^n z_{jk}(x)P_j\big(x + y(x)\big), \\ \qquad \vdots \\ Q_n(x) = P_n\big(x + y(x)\big) + \sum_{j=3}^n z_{jk}(x)P_j\big(x + y(x)\big). \end{cases}$$

Now, since $y_i(x) \in (x)I \subset (x)^2$, the series $x_i + y_i(x)$ define a change of variables, after which we have

$$g - f' \in (Q_3, \ldots, Q_n).$$

Furthermore, since the $z_{kj}(x)$'s are in $(x)I \subset (x)$, after the change we also have:

$$(Q_3, \ldots, Q_n) + (x)(P_3, \ldots, P_n) = (P_3, \ldots, P_n).$$

Hence, by Nakayama's Lemma, the ideals $(Q_3, \ldots, Q_n)$ and $(P_3, \ldots, P_n)$ coincide, and $\mathrm{ht}(Q_3, \ldots, Q_n) = n - 2$.

This completes Step IV and the proof of the proposition. $\square$

Now we are ready for Theorem 1.3, but we prove first a more technical statement. This is obtained combining the previous algebrization procedure with the quantitative refinements of Section 2.

PROPOSITION 3.2. – *Let $X \subset \mathbb{R}^n$ be a surface germ at the origin and let J denote its ideal. Let $f \in \mathbb{R}\{x\}$ be positive semidefinite on the germ of $\mathbb{R}^n$ at the origin and suppose it does not vanish on any irreducible component of dimension 2 of X. Then there exist analytic function germs $g, h_1, h_2, h_3, h_4 \in \mathbb{R}\{x\}$ such that*

$$g^2 f \equiv h_1^2 + h_2^2 + h_3^2 + h_4^2 \bmod J$$

and g is a sum of squares with $\{g = 0\} \subset \{f = 0\}$.

Proof. – The case $f(0) > 0$ is clear, so we suppose $f(0) = 0$. After a change of coordinates we find the germs h, f' and $Q_3, \ldots, Q_n$ as in Proposition 3.1. We are to move the problem to a suitable finitely generated algebra over $\mathbb{R}\{x_1\}$, but this requires some work.

First of all, consider the ideal $\mathfrak{a} = (Q_3, \ldots, Q_n)\mathbb{R}\{x_1\}[x_2, \ldots, x_n]$ and the algebra $A = \mathbb{R}\{x_1\}[x_2, \ldots, x_n]/\mathfrak{a}$. Its minimal primes split into some $\mathfrak{p}_i$ contained in the maximal ideal $\mathfrak{m} = (x) \bmod \mathfrak{a}$, and some others $\mathfrak{q}_j$ not contained: choose $f_0 \in \bigcap_j \mathfrak{q}_j \setminus \mathfrak{m}$, which is not nilpotent in A. Then, in the localization $A_0 = A[1/f_0]$ only the $\mathfrak{p}_i$'s remain, and by 2.3 and 2.2 and flatness, we get:

$$\delta(A_0) = \dim(A_0) = \dim(A_{\mathfrak{m}}) = \dim\big(\mathbb{R}\{x\}/(Q_3, \ldots, Q_n)\big) = 2.$$

Next, consider f'. We claim it is not nilpotent in A_0. Indeed, otherwise, it would belong to all the $\mathfrak{p}_i$'s, and since the ideals $(Q_3, \ldots, Q_n)\mathbb{R}\{x\} \subset J$ have the same height $n-2$, f' would belong to some minimal prime of height $n-2$ of J. Thus, f' would vanish on some irreducible component of dimension 2 of X. Since $f' = f + h \bmod (Q_3, \ldots, Q_n)$, and f, h are both psd, we would conclude that f vanishes on that same component, which is not the case by hypothesis.

Thus, we can properly consider the localization $A' = A_0[1/f'] = A_0[T]/(1 - f'T)$, and by 2.4, $\delta(A') \leqslant \delta(A_0) = 2$.

Next, since f' is psd on the germ $Y = \{Q_3 = \cdots = Q_n = 0\}$, we can choose $\varepsilon > 0$ such that $f', Q_3, \ldots, Q_n$ converge on $U = \{|x_1| < 2\varepsilon, \ldots, |x_n| < 2\varepsilon\}$ and $f' \geqslant 0$ on $U \cap Y$. Consider the algebra

$$B = A'[T_2, \ldots, T_n] / \big(T_2^2 - \big(\varepsilon^2 - x_2^2\big), \ldots, T_n^2 - \big(\varepsilon^2 - x_n^2\big)\big),$$

which is finitely generated over $\mathbb{R}\{x_1\}$. As, by 2.4, $\delta(C[T]/(T^2 - c)) \leqslant \delta(C)$, we see that $\delta(B) \leqslant \delta(A') \leqslant 2$. We claim that

(•) The element f' is totally positive in B.

If not, there exists $\beta \in \mathrm{Spec}_r(B)$ such that $f'(\beta) \leqslant 0$; in fact, $f'(\beta) < 0$ since f' is a unit in B. As is well known, β can be seen as a homomorphism $\beta : B \to R$ into a real closed field R such that $\beta(f') < 0$. Immediately, we get a homomorphism $\alpha : \mathbb{R}\{x_1\}[x_2, \ldots, x_n, T_2, \ldots, T_n] \to R$ such that

$$\alpha(f') < 0, \quad \alpha(Q_k) = 0, \quad \alpha\big(T_j^2 - \big(\varepsilon^2 - x_j^2\big)\big) = 0.$$

We set $\alpha(x_i) = \alpha_i, \alpha(T_j) = \tau_j$, and distinguish two cases:

(1) If $\alpha_1 = 0$, then $\alpha|_{\mathbb{R}\{x_1\}}$ is evaluation at $x_1 = 0$, and we get:

$$f'(0, \alpha_2, \ldots, \alpha_n) < 0, \quad Q_k(0, \alpha_2, \ldots, \alpha_n) = 0, \quad \tau_j^2 = \varepsilon^2 - \alpha_j^2.$$

Thus we can apply Tarski's principle, and suppose $\alpha_i, \tau_j \in \mathbb{R}$. Now note that the condition $\tau_j^2 = \varepsilon^2 - \alpha_j^2$ implies $\alpha_j^2 \leqslant \varepsilon^2$, so that $(0, \alpha_2, \ldots, \alpha_n) \in U$, and this is in fact a point of $U \cap Y$ at which f' is < 0. Contradiction.

(2) If $\alpha_1 \neq 0$, then $\alpha|\mathbb{R}\{x_1\}$ is injective, and we may assume R contains $\mathbb{R}(\{x_1\})$. Then we have:

$$f'(x_1, \alpha_2, \ldots, \alpha_n) < 0, \quad Q_k(x_1, \alpha_2, \ldots, \alpha_n) = 0, \quad \tau_j^2 = \varepsilon^2 - \alpha_j^2.$$

We can again apply Tarski's principle, and get the α_i's and the τ_j's in the real closure of $\mathbb{R}(\{x_1\})$. This real closure is the field of convergent Puiseux series on the variable $t = \pm x_1$ according to the sign of $x_1 = \alpha_1$, so that $x_1 \mapsto (x_1, \alpha_2, \ldots, \alpha_n)$ is a well defined analytic map at least for $x_1 \neq 0$ small enough. But again the condition $\tau_j^2 = \varepsilon^2 - \alpha_j^2$ guarantees that the image of that map is contained in U, and thus, we get points in $Y \cap U$ at which f' is negative. Impossible.

Thus we have proved our claim (•) that f' is a totally positive element in B. Then, since $\delta(B) \leqslant 2$, by Theorem 2.6 we can write in B:

$$\big(1 + a_1^2 + \cdots + a_r^2\big)^2 f' = b_1^2 + b_2^2 + b_3^2 + b_4^2.$$

Now we remark that the inclusion $\mathbb{R}\{x_1\}[x_2, \ldots, x_n] \subset \mathbb{R}\{x\}$ induces a homomorphism $A = \mathbb{R}\{x_1\}[x_2, \ldots, x_n]/\mathfrak{a} \to \mathbb{R}\{x\}/J$, which extends to another $B \to (\mathbb{R}\{x\}/J)[1/f']$ (recall that $f_0 \notin \mathfrak{m}$, hence f_0 is a unit in $\mathbb{R}\{x\}$, and all the $\varepsilon_j^2 - x_j^2$'s have square roots in $\mathbb{R}\{x\}$). Consequently, we can suppose the above formula holds in $(\mathbb{R}\{x\}/J)[1/f']$, and clearing denominators we get a similar formula in $\mathbb{R}\{x\}/J$:

$$\big(f'^{2m} + a_1^2 + \cdots + a_r^2\big)^2 f' = b_1^2 + b_2^2 + b_3^2 + b_4^2.$$

Finally, since $f = f' \bmod J$, we get

$$\big(f^{2m} + g_1^2 + \cdots + g_r^2\big)^2 f = h_1^2 + h_2^2 + h_3^2 + h_4^2 \bmod J$$

with $g_k, h_\ell \in \mathbb{R}\{x\}$. Clearly, the denominator $g = f^{2m} + g_1^2 + \cdots + g_r^2$ cannot vanish off $\{f = 0\}$, and we have finished. □

As said before, Theorem 1.3 follows from the latter result.

Proof of Theorem 1.3. – We are given a psd analytic function germ $f : X \to \mathbb{R}$ on the surface germ $X \subset \mathbb{R}^n$. By the Positivstellensatz $g'^2 f$ is a sum of squares for a suitable denominator g' such that $\{g' = 0\} \subset \{f = 0\}$. In particular, $g'^2 f$ can be extended to a psd analytic function germ $f' : \mathbb{R}^n \to \mathbb{R}$. Consequently, after substituting f' for f, we simply suppose that f is defined and psd on $\mathbb{R}^n$. Now, decompose $X = X' \cup X''$, so that f does not vanish on any irreducible component of X' and $f|X'' \equiv 0$. By Proposition 3.2 we find $g, h_1, h_2, h_3, h_4 \in \mathbb{R}\{x\}$ such that $g^2 f = h_1^2 + h_2^2 + h_3^2 + h_4^2$ on X', and g is a sum of squares with $\{g = 0\} \subset \{f = 0\}$. We are done, because on the whole of X we can write

$$\left(gf^2\right)^2 f = \left(h_1 f^2\right)^2 + \left(h_2 f^2\right)^2 + \left(h_3 f^2\right)^2 + \left(h_4 f^2\right)^2$$

(we use the factor f^2 to preserve the fact that g is a sum of squares). $\square$

4. Normal real analytic surfaces

In this last section we are to prove Theorem 1.4. To that end, we will use a particular case of a result further extended in [1]. We include here this particular case with a direct condensed proof for the convenience of the reader:

LEMMA 4.1. – *Let $\theta : \mathbb{R}^n \to \mathbb{R}$ be a fixed analytic function. Let $\xi : \mathbb{R}^n \to \mathbb{R}$ be an analytic function with isolated zeros. Suppose that at every zero x, the germ ξ_x is a sum of q squares of analytic function germs, one of them divisible by θ. Then there are analytic functions $f_1, \dots, f_q : \mathbb{R}^n \to \mathbb{R}$, one of them divisible by θ, such that*

$$\left(f_1^2 + \cdots + f_q^2\right)\mathcal{O}_{\mathbb{R}^n,x} = \xi \mathcal{O}_{\mathbb{R}^n,x}$$

at every zero x of ξ.

Proof. – We will resource to complexification and holomorphic functions, for which we refer the reader to the classical [8]. Take coordinates $z = (z_1, \dots, z_n)$ in $\mathbb{C}^n$, with $z_i = x_i + \sqrt{-1}y_i$, $x_i, y_i \in \mathbb{R}$. Consider then the conjugation $\sigma : z \mapsto \overline{z} = (\overline{z_1}, \dots, \overline{z_n})$, whose fixed points are $\mathbb{R}^n$. A subset $Y \subset \mathbb{C}^n$ is *invariant* if $\sigma(Y) = Y$. We will denote by Int and Cl topological interiors and closures, respectively.

An holomorphic function $F : \mathcal{U} \to \mathbb{C}$ defined on an invariant open set $\mathcal{U} \subset \mathbb{C}^n$ is *invariant* if $F(z) = \overline{F(\overline{z})}$. This implies that F restricts to a real analytic function on $\mathcal{U} \cap \mathbb{R}^n$. In general, we have the *real* and the *imaginary* parts of F

$$\Re(F)(z) = \frac{1}{2}\left(F(z) + \overline{F(\overline{z})}\right), \qquad \Im(F)(z) = \frac{1}{2\sqrt{-1}}\left(F(z) - \overline{F(\overline{z})}\right)$$

which satisfy $F = \Re(F) + \sqrt{-1}\,\Im(F)$; both are invariant holomorphic functions.

Now, we split the proof of the lemma into several steps.

Step I: Globalization of the sums of squares. Let $x_k, k \geqslant 1$, be the zeros of ξ, and consider an open neighborhood $\mathcal{V}$ of $\mathbb{R}^n$ in $\mathbb{C}^n$ on which ξ and θ have invariant holomorphic extensions Ξ and Θ. By hypothesis, for each k there are invariant holomorphic functions $F_{ki} : \mathcal{V}_k \to \mathbb{C}$, $1 \leqslant i \leqslant q$, defined on an open neighborhood $\mathcal{V}_k \subset \mathcal{V}$ of x_k in $\mathbb{C}^n$, such that $\Xi|_{\mathcal{V}_k} = \sum_k F_{ki}^2$, and $\Theta|_{\mathcal{V}_k}$ divides F_{kq}, say

$$F_{kq} = F_{kq}^* \Theta,$$

for a suitable invariant holomorphic function $F^*_{kq}:\mathcal{V}_k\to\mathbb{C}$. Clearly, the $\mathcal{V}_k$'s may be chosen disjoint each other. The open set

$$\mathcal{V}'=\big(\mathcal{V}\setminus\{\Xi=0\}\big)\cup\bigcup_k\mathcal{V}_k$$

is a neighborhood of $\mathbb{R}^n$ in $\mathbb{C}^n$, and we can choose an invariant open Stein neighborhood $\mathcal{U}\subset\mathcal{V}'$ of $\mathbb{R}^n$ in $\mathbb{C}^n$, such that $\mathbb{R}^n$ is a deformation retract of $\mathcal{U}$ [4]. We restrict all functions to $\mathcal{U}$, and shrink $\mathcal{V}_k$ inside $\mathcal{U}$ so that the connected component S_k of $\{\Xi=0\}$ that contains x_k is the only one that meets $\mathcal{V}_k$, and it is in fact contained in $\mathcal{V}_k$. Now, by the condition on $\mathcal{V}_k$ and the connected components of $\{\Xi=0\}$, each function $\zeta=F^*_{kq},F_{ki},1\leqslant i<q$, defines a global cross section of the sheaf $\mathcal{O}_{\mathcal{U}}/\Xi^2$ as follows:

$$\begin{cases}\zeta \bmod \Xi^2\mathcal{O}_{\mathbb{C}^n,x} & \text{if } x\in\mathcal{V}_k\\ 0 & \text{if } x\in\mathcal{U}\setminus S_k.\end{cases}$$

By Cartan's Theorem B, these sections are just holomorphic functions $\Phi^*_{kq},\Phi_{ki}:\mathcal{U}\to\mathbb{C}$, $1\leqslant i<q$, such that Ξ^2 divides $\Phi^*_{kq}-F^*_{kq}$ and $\Phi_{ki}-F_{ki}$, $1\leqslant i<q$; set $\Phi_{kq}=\Phi^*_{kq}\Theta$. Replacing them by their real parts, we may assume that they all are invariant.

On $\mathcal{V}_k$ we have:

$$\sum_i\Phi^2_{ki}-\Xi=\sum_i\Phi^2_{ki}-\sum_i F^2_{ki}=\sum_i(\Phi_{ki}+F_{ki})(\Phi_{ki}-F_{ki})=\Psi_k\Xi^2,$$

for some holomorphic function $\Psi_k:\mathcal{V}_k\to\mathbb{C}$. Hence

$$\sum_i\Phi^2_{ki}=\Xi+\Psi_k\Xi^2=(1+\Psi_k\Xi)\Xi.$$

Step II: Auxiliary construction. Let $\{L_k\}_{k\geqslant1}$ be a family of invariant compact subsets of $\mathcal{U}$ such that $L_1\cap\mathbb{R}^n\neq\emptyset$, $L_1\not\subset\bigcup_\ell\mathcal{V}_\ell$, $L_k\subset\operatorname{Int}_{\mathbb{C}^n}(L_{k+1})$ for all k, and $\bigcup_k L_k=\mathcal{U}$; we replace each L_k by $L_k\setminus\bigcup_{\ell\geqslant k}\mathcal{V}_\ell$ to have in addition $S_k\cap L_k=\emptyset$.

We are to construct invariant holomorphic functions $\Lambda_k:\mathcal{U}\to\mathbb{C}$ such that

(i) $S_k=\{\Lambda_k=0\}$ is the connected component of $\{\Xi+\Lambda_k^2=0\}$ that contains x_k,

(ii) the meromorphic function $w_k=\Xi/(\Xi+\Lambda_k^2)$ is a holomorphic unit on a neighborhood of S_k, which we may suppose to be $\mathcal{V}_k$, and

(iii) $\Xi+\Lambda_k^2$ has no zero in L_k.

Indeed, fix k, and let $\mathcal{J}$ be the sheaf of ideals of holomorphic function germs on $\mathcal{U}$ defined by

$$\mathcal{J}_x=\begin{cases}\Xi_x\mathcal{O}_{\mathbb{C}^n,x} & \text{if } x\in S_k,\\ \mathcal{O}_{\mathbb{C}^n,x} & \text{if } x\in\mathcal{U}\setminus S_k.\end{cases}$$

The open set $\mathcal{U}$ is a Stein manifold, hence $H^1(\mathcal{U},\mathcal{O}^*_{\mathbb{C}})=H^2(\mathcal{U},\mathbb{Z})$, and this group is trivial because $\mathbb{R}^n$ is a deformation retract of $\mathcal{U}$. Consequently, all locally principal coherent sheaves of ideals on $\mathcal{U}$ are in fact globally principal. In particular, $\mathcal{J}$ is generated by a holomorphic function $H:\mathcal{U}\to\mathbb{C}$. We can write $H=A+\sqrt{-1}B$, where $A=\Re(H)$ and $B=\Im(H)$; note that $x_k\in\{A=B=0\}\subset\{H=0\}=S_k$.

Let $\Lambda_k=\mu(A^2+B^2)$ for a certain positive real number $\mu>0$ that we will choose later; this is clearly an invariant holomorphic function. Since $\Lambda_k(z)=\mu H(z)\overline{H(\overline{z})}$ for all $z\in\mathcal{U}$, we have

$\Lambda_k(z)=0$ if and only if $H(z)=0$ or $H(\overline{z})=0$, that is, $z\in S_k$ or $\overline{z}\in S_k$. But S_k is invariant (because Ξ and $\mathcal{U}$ are so), hence $z\in S_k$. Thus, $\{\Lambda_k=0\}=S_k$.

Now, by construction, we have $\Xi=CH$ for some holomorphic unit C on an open neighborhood of S_k, hence:

$$\Xi+\Lambda_k^2=\Xi+\mu^2H^2\overline{H^2\circ\sigma}=\left(1+\frac{\mu^2H\overline{H^2\circ\sigma}}{C}\right)\Xi.$$

Obviously $v_k=1+(\mu^2H\overline{H^2\circ\sigma})/C$ is a well defined holomorphic unit in a neighborhood of S_k, say $\mathcal{V}_k$ after shrinking, and $w_k=1/v_k$ is a unit too.

Next, we choose μ. Since the zeros of the holomorphic function A^2+B^2 are all in S_k and $L_k\cap S_k=\emptyset$, we can take

$$\mu=\sqrt{\frac{1+\max_{L_k}|\Xi|}{\min_{L_k}|A^2+B^2|^2}}>0$$

so that $|\Xi|<\mu^2|A^2+B^2|^2$ on L_k. Hence, $\Xi+\Lambda_k^2$ has no zero in L_k.

Let us check that the connected component T of $\{\Xi+\Lambda_k^2=0\}$ that contains x_k is S_k. Clearly $x_k\in S_k\subset T$. Suppose that $S_k\neq T$, say $a\in T\setminus S_k$. Since T is connected there is a path $\gamma:[0,1]\to T$ such that $\gamma(0)=a$ and $\gamma(1)=x_k$. Let $0<s=\min\{t\in[0,1]:\ \gamma(t)\in S_k\}$. Since $z=\gamma(s)\in S_k\subset\mathcal{V}_k$, the germs at z of $\Xi+\Lambda_k^2$ and Ξ differ by a unit, hence the set germs T_z and $S_{k,z}$ coincide. But this is impossible because $\gamma[0,s)\subset T\setminus S_k$.

Step III: Gluing of sums of squares. As far, we have that x_k is the unique real zero of $\Xi+\Lambda_k^2$, hence the connected components of $\{\Xi+\Lambda_k^2=0\}$ other than S_k do not meet $\mathbb{R}^n$, and dropping them, we get an open neighborhood $\mathcal{W}_k$ of $L_k\cup\mathbb{R}^n$ on which

$$w_k=\frac{\Xi}{\Xi+\Lambda_k^2}$$

is holomorphic, and $\{\Xi+\Lambda_k^2=0\}\cap\mathcal{W}_k=S_k$. As a matter of fact, *there is a common open neighborhood $\mathcal{W}\subset\mathcal{U}$ of $\mathbb{R}^n$ on which all the above quotients w_k are holomorphic, and $\{\Xi+\Lambda_k^2=0\}\cap\mathcal{W}\subset S_k$.*

Indeed, it is enough to find for each $x\in\mathbb{R}^n$ an open neighborhood $\mathcal{W}^x$ in $\mathbb{C}^n$, on which the required properties hold true, and the union of these $\mathcal{W}^x$'s will be the $\mathcal{W}$ we seek. But $x\in\operatorname{Int}_{\mathbb{C}^n}(L_{k_0})$ for some k_0, hence $x\in L_k$ for all $k\geqslant k_0$. Consequently, all w_k's are holomorphic in $\mathcal{W}^x=\mathcal{W}_1\cap\cdots\cap\mathcal{W}_{k_0-1}\cap\operatorname{Int}_{\mathbb{C}^n}(L_{k_0})$, and if $z\in\mathcal{W}^x$ is a zero of $\Xi+\Lambda_k^2$, then $k<k_0$, hence $z\in\mathcal{W}_k$ and $z\in S_k$.

Once we have this $\mathcal{W}$, we can paste the sums of squares $\sum_i\Phi_{ki}^2$ to get a single one. Define, for each k:

$$M_k=\max_i\max_{L_k}\bigl|w_k^2\Phi_{ki}\bigr|,\qquad \gamma_k=\frac{1}{2^kM_k}.$$

On L_k we have $|\gamma_kw_k^2\Phi_{ki}|\leqslant\frac{1}{2^k}$ for all i.

Now, let L be a compact subset of the $\mathcal{W}$ found above, where all the functions $\gamma_kw_k^2\Phi_{ki}$ are holomorphic. As $\mathcal{W}\subset\bigcup_{k\geqslant1}\operatorname{Int}_{\mathbb{C}^n}(L_k)$, L is contained in some L_{k_0}, hence in all L_k for $k\geqslant k_0$, and so:

$$\sum_k\sup_L\bigl|\gamma_kw_k^2\Phi_{ki}\bigr|=\sum_{k=1}^{k_0-1}\sup_L\bigl|\gamma_kw_k^2\Phi_{ki}\bigr|+\sum_{k\geqslant k_0}\sup_{L_k}\bigl|\gamma_kw_k^2\Phi_{ki}\bigr|$$

$$\leqslant \sum_{k=1}^{k_0-1} \sup_L \bigl|\gamma_k w_k^2 \Phi_{ki}\bigr| + \sum_{k\geqslant k_0} \frac{1}{2^k} < +\infty.$$

Consequently, each infinite sum

$$F_i = \sum_k \gamma_k w_k^2 \Phi_{ki}, \quad i = 1, \ldots, q,$$

converges uniformly on compact sets, hence defines a holomorphic function on $\mathcal{W}$. Notice also that since Θ divides each Φ_{kq}, it divides F_q. Fix now k. As each $\Xi + \Lambda_\ell^2, \ell \neq k$, is a unit on $\mathcal{W} \cap \mathcal{V}_k$, we can write there

$$\gamma_\ell w_\ell^2 \Phi_{\ell i} = \gamma_\ell \left(\frac{\Xi}{\Xi + \Lambda_\ell^2}\right)^2 \Phi_{\ell i} = \Delta_{k\ell i} \Xi^2$$

so that

$$F_i = \gamma_k w_k^2 \Phi_{ki} + \sum_{\ell \neq k} \gamma_\ell w_\ell^2 \Phi_{\ell i} = \gamma_k w_k^2 \Phi_{ki} + \Delta_{ki} \Xi^2$$

where $\Delta_{ki} = \sum_{\ell \neq k} \Delta_{k\ell i}$ is a holomorphic function. From this and step II, we get

$$\sum_i F_i^2 = \gamma_k^2 w_k^4 \sum_i \Phi_{ki}^2 + \Delta \Xi^2 = \bigl(\gamma_k^2 w_k^4 (1 + \Psi_k \Xi) + \Delta \Xi\bigr)\Xi,$$

where Δ is holomorphic. But w_k is a unit at x_k, and we deduce:

$$\sum_i F_i^2 \mathcal{O}_{\mathbb{C}^n, x_k} = \Xi \mathcal{O}_{\mathbb{C}^n, x_k}.$$

After restriction to $\mathbb{R}^n$ we get $\sum_i f_i^2 \mathcal{O}_{\mathbb{C}^n, x_k} = \xi \mathcal{O}_{\mathbb{R}^n, x_k}$, where each $f_i = F_i|_{\mathbb{R}^n}$ is a real analytic function. As Θ divides F_q, θ divides f_q. $\square$

Once the preceding result is available, we can turn to the

Proof of Theorem 1.4. – We have a normal real analytic surface X and a psd analytic function $f: X \to \mathbb{R}$, which we must represent as a sum of squares.

First of all, we recall that X can be embedded as a closed subset of $\mathbb{R}^n$, which we suppose henceforth. On the other hand, since a normal surface is locally irreducible, the irreducible components of X are its connected components, and working separately on each we may assume X is irreducible; thus the ring $\mathcal{O}(X)$ is a normal domain. Also, we know that all singularities of X are isolated. For a point $x \in X$, we denote $\mathcal{O}(X)_x$ the localization at its corresponding maximal ideal $\mathfrak{m}_x$: x is a regular point if and only if $\mathcal{O}(X)_x$ is a regular ring. Recall as well that normal surfaces are coherent, and we can use sheaf theory on X without restrictions.

After this, we split our argument in several steps.

Step I: Construction of suitable equations for the codimension 1 *part of the zero set* $\{f = 0\}$.

We split $\{f = 0\} = D \cup Y$, where D is a discrete set and $Y = \bigcup_i Y_i$ is the union of the irreducible components of dimension 1. Then, the ideal $\mathfrak{p}_i \subset \mathcal{O}(X)$ of all functions vanishing on Y_i is a prime ideal of height 1, and, $\mathcal{O}(X)$ being normal, the localization $V_i = \mathcal{O}(X)_{\mathfrak{p}_i}$ is a discrete valuation ring. We will use freely the so-called multiplicity along Y_i, which is the real valuation m_{Y_i} associated to the discrete valuation ring V_i (see [2, §§1,2] for full details). Pick any uniformizer $g_i \in \mathfrak{p}_i$ of V_i, so that $m_{Y_i}(g_i) = 1$. Since f is psd, and the valuation is real, $m_{Y_i}(f) = 2m_i$, and $f/g_i^{2m_i}$ is a unit in V_i. From this it follows that *at all points of Y_i off a discrete set* the following three properties hold true:

(i) $f/g_i^{2m_i}$ is analytic,

(ii) $f/g_i^{2m_i} > 0$, and

(iii) g_i generates the ideal of Y_i.

We are to modify g_i still a little, keeping these properties. To that end, consider any point $c \notin Y$, and denote $d_i = \operatorname{dist}(Y_i, c) > 0$. Let θ_i be an equation for Y_i, that is $\{\theta_i = 0\} = Y_i$. Scaling θ_i we may assume that

$$\|g_i(x)\| < \|\theta_i^2(x)\| \quad \text{on } \|x - c\| \leqslant \frac{1}{2} d_i.$$

Now, since $\theta_i \in \mathfrak{p}_i$,

$$m_{Y_i}(\theta_i^2) = 2m_{Y_i}(\theta_i) \geqslant 2 > m_{Y_i}(g_i),$$

and $g_i + \theta_i^2$ is also a uniformizer of V_i with the same three properties above. But in addition, the zero sets $Z_i = \{g_i + \theta_i^2 = 0\}$ form a locally finite family.

Indeed, it is enough to show that for every radius $\rho > 0$, only finitely many Z_i's meet the ball $\{\|x - c\| < \rho\}$. To see this, notice that, the Y_i's being the irreducible components of the analytic curve Y, they form a locally finite family, hence for i large, $Y_i \cap \{\|x - c\| < 2\rho\} = \emptyset$, so that $\frac{1}{2} d_i \geqslant \rho$ and

$$\|g_i(x)\| < \|\theta_i^2(x)\|, \quad \text{hence } g_i(x) + \theta_i^2(x) \neq 0,$$

for $\|x - c\| < \rho$, as wanted.

Finally, we replace each g_i by $g_i + \theta_i^2$, but keep the notation g_i.

Step II: Reduction to the case of a discrete zero set.

Set $Z = \bigcup_i Z_i$, and consider the analytic sheaf of ideals given by

$$\mathcal{I}_x = \begin{cases} \prod_{i|x \in Z_i} g_i^{m_i} \mathcal{O}_{X,x} & \text{for } x \in Z, \\ \mathcal{O}_{X,x} & \text{otherwise.} \end{cases}$$

This is well defined and coherent at any $a \in Z$: on a neighborhood U of a where all the finitely many Z_i's that meet U pass through a, the ideal $\mathcal{I}$ is generated by $\prod_{i|x \in Z_i} g_i^{m_i}$. By [6], since $\mathcal{I}$ is locally principal, $\mathcal{I}$ is globally generated by three sections $h_1, h_2, h_3 \in \mathcal{O}(X)$.

In this situation, on Y_i off a discrete set, $\mathcal{I} = (h_1, h_2, h_3)\mathcal{O}_X$ is generated by $g_i^{m_i}$, which readily implies that all the quotients $h_j/g_i^{m_i}$ for $j = 1, 2, 3$, are analytic there and at least one is a unit. Denote $h_4 = f$. As $f/g_i^{2m_i} = h_4/g_i^{2m_i}$ is a unit on Y_i off another discrete set, we deduce that

$$\frac{f}{h_1^2 + h_2^2 + h_3^2 + h_4^2} = \frac{f}{g^{2m_i}} \Big/ \frac{h_1^2 + h_2^2 + h_3^2 + h_4^2}{g^{2m_i}}$$

is an analytic unit on Y_i off a (bigger) discrete set $D_i \subset Y \subset \{f = 0\}$. As the Y_i's form a locally finite family, we conclude that the zeros and poles of this meromorphic function form a discrete subset of $\{f = 0\}$.

Write $h = h_1^2 + h_2^2 + h_3^2 + h_4^2$ and consider the coherent sheaf $(h : f)\mathcal{O}_X$. This sheaf is generated in a neighborhood of each pole x of f/h by finitely many sections $\delta_1, \ldots, \delta_r$. By the standard sum of squares trick, $f_x/h_x = g/\delta$ for $\delta = \sum_k \delta_k^2$ and some g. Furthermore, x is an isolated zero of δ. For that, suppose that there is $y' \neq x$ arbitrarily close to x with $\delta(y) = 0$.

Then, all δ_k's vanish at y, and since the ideal $(h:f)\mathcal{O}_{X,y}$ is generated by them, it contains no unit. This means that f/h is not analytic at y, a contradiction. Adding the square of an equation of X in $\mathbb{R}^n$, we extend δ to a sum of squares $\tilde{\delta}$ of analytic functions in a neighborhood of x in $\mathbb{R}^n$ that vanishes only at x; denote $\mathcal{I}_x = \tilde{\delta}\mathcal{O}_{X,x}$. These ideals $\mathcal{I}_x$ glue to define a locally principal sheaf of ideals $\mathcal{I}$ on $\mathbb{R}^n$, whose zero set consists of the poles of f/h. Since $H^1(\mathbb{R}^n, \mathbb{Z}_2) = 0$, all locally principal sheaves are globally principal, and $\mathcal{I}$ has a global generator Δ. This Δ is a non-negative analytic function on $\mathbb{R}^n$ whose zeros are the poles x of f/h. In fact, we have just glued the local denominators δ to get a global denominator: $\Delta f/h$ is an analytic function. Then $\Delta^2 f/h$ is also analytic, and its zeros are either poles or zeros of f/h, hence a discrete subset of $\{f=0\}$. Summing up, $f' = \Delta^2 f/h$ is psd with discrete zero set $\{x_k\colon k \geqslant 1\} \subset \{f=0\}$.

Step III. Construction of analytic functions $g, f_1, f_2, f_3, f_4 \in \mathcal{O}(X)$ *such that*

$$g^2 f'\mathcal{O}_{X,x} = \big(f_1^2 + f_2^2 + f_3^2 + f_4^2\big)\mathcal{O}_{X,x}$$

at every zero $x \in \{f'=0\} = \{g=0\}$.

To start with, by Theorem 1.3, in a small enough neighborhood U_k of every zero x_k of f' we have a formula

$$g_k^2 f' = f_{k1}'^2 + f_{k2}'^2 + f_{k3}'^2 + f_{k4}'^2 \quad \text{on } U_k \cap X,$$

where $g_k, f'_{ki} : U_k \to \mathbb{R}$ are analytic functions, and g_k is a sum of squares whose single zero in $U_k \cap X$ is x_k; in fact, replacing g_k by $g_k + \theta^2$ for some equation θ of X, we can suppose x_k is the unique zero of g_k in U_k. Then the ideals $g_k\mathcal{O}_{\mathbb{R}^n}|_{U_k}$ define a locally principal sheaf of ideals on $\mathbb{R}^n$, which is globally principal, say generated by $g : \mathbb{R}^n \to \mathbb{R}$. Thus g is a psd analytic function that vanishes exactly at the x_k's, and on each U_k the function g_k/g is an analytic unit. Hence, we can replace g_k in all the above formulas by g. In other words, we have already found the global denominator g.

Next, consider again the above equation $\theta : \mathbb{R}^n \to \mathbb{R}$ of X in $\mathbb{R}^n$. Then, $\xi_k = f_{k1}'^2 + f_{k2}'^2 + f_{k3}'^2 + f_{k4}'^2 + \theta^2$ only vanishes at x_k, and the ideals $\xi_k\mathcal{O}_{\mathbb{R}^n}|_{U_k}$ define a locally principal sheaf of ideals $\mathcal{I}_i$ on $\mathbb{R}^n$, which as usual is globally generated by some analytic function $\xi : \mathbb{R}^n \to \mathbb{R}$. Clearly, $\{x_k\colon k \geqslant 1\}$ is the zero set of ξ, and ξ is a sum of five squares of analytic functions on a neighborhood of that zero set, with the condition that the fifth function is always (divisible by) θ. We thus can apply Lemma 4.1, and find a sum $f_1^2 + f_2^2 + f_3^2 + f_4^2 + f_5^2$ of 5 squares of analytic functions on $\mathbb{R}^n$, such that f_5 is divisible by θ, and

$$\xi_k\mathcal{O}_{\mathbb{R}^n,x} = \xi\mathcal{O}_{\mathbb{R}^n,x} = \big(f_1^2 + f_2^2 + f_3^2 + f_4^2 + f_5^2\big)\mathcal{O}_{\mathbb{R}^n,x}$$

at every zero $x = x_k$. Since f_5 is divisible by θ, which vanishes on X, we conclude:

$$\big(f_1^2 + f_2^2 + f_3^2 + f_4^2\big)\mathcal{O}_{X,x} = \xi_k\mathcal{O}_{X,x} = \big(f_{k1}'^2 + f_{k2}'^2 + f_{k3}'^2 + f_{k4}'^2\big)\mathcal{O}_{X,x} = g^2 f'\mathcal{O}_{X,x}.$$

Step IV: Further control on the zero set.

Recall that $\{x_k\colon k \geqslant 1\} = \{f'=0\} = \{g=0\} \subset \{f=0\} = D \cup \bigcup_i Y_i$. Pick a real number a such that $h_1' = f_1 + ag^4 f'^2$ does not vanish identically on any Y_i, so that the set $\{f=0, h_1'=0\}$ is discrete. Then, let τ be an analytic function whose zero set is $\{f=0, h_1'=0, f_2 \neq 0\}$, and put

$$h_2' = f_2 + \tau g^4 f'^2, \qquad h_3' = f_3, \qquad h_4' = f_4.$$

We claim that the sum of squares $h_1'^2 + h_2'^2 + h_3'^2 + h_4'^2$ does not vanish on $\{f=0, f' \neq 0\}$.

In fact, suppose $f(y) = h_1'(y) = h_2'(y) = 0$ for some y with $f'(y) \neq 0$, hence $g(y) \neq 0$. Since

$$0 = h_2'(y) = f_2(y) + \tau(y)g(y)^4 f'(y)^2,$$

we deduce that $f_2(y) = 0$ if and only if $\tau(y) = 0$, against the definition of τ.

The final remark is now that at every zero x of f, the ideals

$$\begin{aligned} I_x &= \left(h_1^2 + h_2^2 + h_3^2 + h_4^2\right)\left(h_1'^2 + h_2'^2 + h_3'^2 + h_4'^2\right)\mathcal{O}_{X,x} \quad \text{and} \\ J_x &= g^2\Delta^2 f \mathcal{O}_{X,x} \end{aligned}$$

coincide.

Indeed, we consider first $x \in \{f = 0, f' \neq 0\}$. By the discussion above, the sum of squares $h_1'^2 + h_2'^2 + h_3'^2 + h_4'^2$ is a unit at x, and I_x is generated by $h = h_1^2 + h_2^2 + h_3^2 + h_4^2$. But on the other hand, $x \in \{f' \neq 0\} = \{g \neq 0\}$, hence g is a unit at x, and J_x is generated by $\Delta^2 f$. Finally, $f' = \Delta^2 f/h$ is a unit at x, again because $f'(x) \neq 0$, so that $I_x = J_x$.

Next, we pick a zero $x = x_k$ of f', and compute in $\mathcal{O}_{X,x}$. By step III there is a unit u_x such that $f_1^2 + f_2^2 + f_3^2 + f_4^2 = u_x g^2 f'$, and by definition of the h_i''s we have:

$$\sum_i h_i^2 \sum_i h_i'^2 = h\left(\sum_i f_i^2 + \mu g^4 f'^2\right) = hg^2 f'\left(u_x + \mu g^2 f'\right) = g^2\Delta^2 f v_x,$$

where $v_x = u_x + \mu g^2 f'$ is a unit, always in $\mathcal{O}_{X,x}$. Once again, $I_x = J_x$.

Step V: Conclusion.

From the preceding step we see that the function

$$u = \frac{(h_1^2 + h_2^2 + h_3^2 + h_4^2)(h_1'^2 + h_2'^2 + h_3'^2 + h_4'^2)}{g^2\Delta^2 f}$$

is analytic and a unit in a neighborhood of $\{f = 0\}$. Then, the function

$$v = \frac{g^2\Delta^2 f^2 + (h_1^2 + h_2^2 + h_3^2 + h_4^2)(h_1'^2 + h_2'^2 + h_3'^2 + h_4'^2)}{g^2\Delta^2 f} = f + u$$

is a well defined strictly positive analytic function on X: both addends in the right-hand side are $\geqslant 0$, and the second one does not vanish on the zero set of the first. Thus, v has a strictly positive analytic square root w, and on X we get:

$$w^2 g^2\Delta^2 f = g^2\Delta^2 f^2 + \left(h_1^2 + h_2^2 + h_3^2 + h_4^2\right)\left(h_1'^2 + h_2'^2 + h_3'^2 + h_4'^2\right).$$

Since products of sums of four squares are again sums of four squares, the right-hand side is a sum of five squares. We are done.

One final remark is that Theorem 1.4 also asks for $wg\Delta$ to be a sum of squares. This can be amended easily. By our construction, $wg\Delta$ is psd with discrete zero set contained in $\{f = 0\}$. Thus it can be represented by a sum of squares with controlled bad set, and multiplying by the denominator of that representation we obtain a new representation of f whose denominator is indeed a sum of squares. □

REFERENCES

[1] ACQUISTAPACE F., BROGLIA F., FERNANDO J.F., RUIZ J.M., On the 17th Hilbert Problem for global analytic functions, Preprint RAAG 2004.

[2] ANDRADAS C., DÍAZ-CANO A., RUIZ J.M., The Artin–Lang property for normal real analytic surfaces, *J. reine angew. Math.* **556** (2003) 99–111.

[3] BOCHNAK J., COSTE M., ROY M.F., Real Algebraic Geometry, Ergeb. Math. Grenzgeb., vol. **36**, Springer, Berlin, 1998.

[4] CARTAN H., Variétés analytiques réelles et variétés analytiques complexes, *Bull. Soc. Math. France* **85** (1957) 77–99.

[5] CHOI M.D., DAI Z.D., LAM T.Y., REZNICK B., The Pythagoras number of some affine algebras and local algebras, *J. reine angew. Math.* **336** (1982) 45–82.

[6] COEN S., Sul rango dei fasci coerenti, *Boll. Univ. Mat. Ital.* **22** (1967) 373–383.

[7] GREENBERG M.J., Lectures on Forms in Many Variables, W.A. Benjamin, New York, 1969.

[8] GUNNING R., ROSSI H., Analytic Functions of Several Complex Variables, Prentice-Hall, Englewood Cliffs, NJ, 1965.

[9] JAWORSKI P., Positive definite analytic functions and vector bundles, *Bull. Acad. Pol. Sci.* **30** (1982) 501–506.

[10] JAWORSKI P., About estimates on number of squares necessary to represent a positive-semidefinite analytic function, *Arch. Math.* **58** (1992) 276–279.

[11] DE JONG T., PFISTER G., Local Analytic Geometry, Basic Theory and Applications, Advanced Lectures in Mathematics, Vieweg, Braunschweig, 2000.

[12] LAM T.Y., The Algebraic Theory of Quadratic Forms, Mathematics Lecture Notes Series, W.A. Benjamin, Massachusetts, 1973.

[13] MAHÉ L., Level and Pythagoras number of some geometric rings, *Math. Z.* **204** (4) (1990) 615–629.

[14] MAHÉ L., Théorème de Pfister pour les variétés et anneaux de Witt réduits, *Invent. Math.* **85** (1) (1986) 53–72.

[15] NARASIMHAN R., Introduction to the Theory of Analytic Spaces, Lecture Notes in Math., vol. **25**, Springer, Berlin, 1966.

[16] PFISTER A., Quadratic Forms with Applications to Algebraic Geometry and Topology, London Math. Soc. Lecture Note Ser., vol. **217**, Cambridge University Press, Cambridge, 1995.

[17] PRESTEL A., DELZELL C.N., Positive Polynomials, Monographs in Mathematics, Springer, Berlin, 2001.

[18] RUIZ J.M., The Basic Theory of Power Series, Advanced Lectures in Mathematics, Vieweg, Braunschweig, 1993.

[19] TOUGERON J.-C., Idéaux de fonctions différentiables, Ergeb. Math. Grenzgeb., vol. **71**, Springer, Berlin, 1972.

(Manuscrit reçu le 30 juin 2004 ;
accepté, après révision, le 7 avril 2005.)

Francesca ACQUISTAPACE
Dipartimento di Matematica,
Università degli Studi di Pisa,
Via Filippo Buonarroti, 2,
56127 Pisa, Italy
E-mail: acquistf@dm.unipi.it

Fabrizio BROGLIA
Dipartimento di Matematica,
Università degli Studi di Pisa,
Via Filippo Buonarroti, 2,
56127 Pisa, Italy
E-mail: broglia@dm.unipi.it

José F. FERNANDO
Departamento de Matemáticas,
Facultad de Ciencias,
Universidad Autónoma de Madrid,
Cantoblanco, 28040 Madrid, Spain
E-mail: josefrancisco.fernando@uam.es

Jesús M. RUIZ
Departamento de Geometría y Topología,
Facultad de Ciencias Matemáticas,
Universidad Complutense de Madrid,
28040 Madrid, Spain
E-mail: jesusr@mat.ucm.es