POLYNOMIALS NON-NEGATIVE ON A STRIP

M. MARSHALL

(Communicated by Ted Chinburg)

Abstract. We prove that if \( f(x, y) \) is a polynomial with real coefficients which is non-negative on the strip \([0, 1] \times \mathbb{R}\), then \( f(x, y) \) has a presentation of the form

\[
f(x, y) = \sum_{i=1}^{k} g_{i}(x, y)^{2} + \sum_{j=1}^{l} h_{j}(x, y)^{2}(1-x),
\]

where the \( g_{i}(x, y) \) and \( h_{j}(x, y) \) are polynomials with real coefficients.

1. Introduction

In [2] Hilbert showed that there are polynomials \( f(x, y) \in \mathbb{R}[x, y] \) (necessarily of degree \( \geq 6 \)) which are non-negative on all of \( \mathbb{R}^2 \) but are not expressible as a sum of squares in \( \mathbb{R}[x, y] \). The best-known example is the polynomial \( f(x, y) = 1 - 3x^2 y^2 + x^4 y^2 + x^2 y^4 \). In contrast to this result, we prove:

**Theorem 1.1.** Suppose \( f(x, y) \in \mathbb{R}[x, y] \) is non-negative on the strip \([0, 1] \times \mathbb{R}\). Then \( f(x, y) \) is expressible as

\[
f(x, y) = \sigma(x, y) + \tau(x, y)(1-x),
\]

where \( \sigma(x, y), \tau(x, y) \) are sums of squares in \( \mathbb{R}[x, y] \).

This answers questions in [3] and [10] arising from the solution of the moment problem for cylinders with compact cross-section; see [3] and [16]. In [10] the authors claimed to know a proof of the result, but this claim was later withdrawn. Certain weak versions of the result were known already; see [3], [8] and [9].

A preordering of a ring \( A \) (commutative with 1) is a subset \( T \) of \( A \) satisfying \( T+T \subseteq T, TT \subseteq T \) and \( f^2 \in T \) for all \( f \in A \). The unique smallest preordering of \( A \) is \( \sum A^2 := \) the set of all (finite) sums of squares of elements of \( A \). The preordering of \( A \) generated by finitely many elements \( g_1, \cdots, g_s \) of \( A \) consists of all elements of the form \( \sum \sigma_i g_i^i \), \( \sigma_i \in \sum A^2 \), \( g^i := g_1^{i_1} \cdots g_s^{i_s}, i := (i_1, \cdots, i_s) \) running through the set \([0, 1]^s\).

A finitely generated preordering \( T \) of the polynomial ring \( \mathbb{R}[x_1, \cdots, x_n] \) is said to be **saturated** if, for all \( f \in \mathbb{R}[x_1, \cdots, x_n], f \geq 0 \) on \( K_T \Rightarrow f \in T \). Here, \( K_T := \{ a \in \mathbb{R}^n | \forall g \in T, g(a) \geq 0 \} \). If \( g_1, \cdots, g_s \) are generators of \( T \), then \( K_T \) is the subset of \( \mathbb{R}^n \) defined by the polynomial inequalities \( g_i \geq 0, i = 1, \cdots, s \).
By [12] Prop. 6.1, a finitely generated preorder $T$ of $R[x_1, \ldots, x_n]$ cannot be saturated if $\dim(K_T) \geq 3$. By [14] Th. 5.4, the same is true for $\dim(K_T) = 2$, if $T$ is stable. The preordering of $R[x, y]$ consisting of sums of squares is stable, so the result of Hilbert referred to earlier can be seen as a special case of this latter result. Theorem [13] asserts that the preordering of $R[x, y]$ generated by $x(1 - x)$ is saturated. The identities $x = x^2 + (1 - x)^2$ and $1 - x = x(1 - x) + (1 - x)^2$ imply that the preordering generated by $x(1 - x)$ coincides with the preordering generated by $x$ and $1 - x$. Before the present paper was written, the only example of a finitely generated saturated preordering in the 2-dimensional non-compact case was the rather artificial example given in [15] Rem. 3.14 (the preordering of $R[x, y]$ generated by $x, 1 - x, y$ and $1 - xy$). It is hoped that the techniques employed in the present paper will yield additional examples of this sort in the future. See [1] and [15] for examples of finitely generated saturated preorderings in the 2-dimensional compact case. See [3, 7, 12] and [13] for 1-dimensional examples.

2. Preliminary reductions

We assume $f \in R[x, y], f \geq 0$, on the strip $[0, 1] \times R$. We want to show $f$ has a presentation $f = \sigma + \tau x(1 - x)$, with $\sigma, \tau$ sums of squares in $R[x, y]$. By considering the behavior of $f$ as $|y| \to \infty$, we see that $f$ has even degree $2d$, as a polynomial in $y$, and that the leading coefficient is $\geq 0$ on $[0, 1]$, i.e., $f$ has the form $f(x, y) = \sum_{i=0}^{2d} a_i(x)y^i, a_i(x) \in R[x], a_{2d}(x) \geq 0$ on $[0, 1]$. If $d = 0$ the result is well-known, e.g., by [3, Th. 2.2] or [4, Prop. 2.7.3], so we assume always that $d \geq 1$.

Lemma 2.1. We may assume $a_{2d}(x) > 0$ on $[0, 1]$.

Proof. Factor $a_{2d}$ as $a_{2d} = \bar{a} a$, where $\bar{a}, a \in R[x], \bar{a} > 0$ on $[0, 1]$ and $\bar{a}$ is $\pm 1$ times a product of linear factors $x - r, r \in [0, 1]$. Then

\[ a^{2d-1}f = \bar{a}a^d \bar{a}y^{2d} + a_{2d-1}^d \bar{a}y^{2d-1} + \cdots + a_0 \bar{a}a^{2d-1}. \]

Let $g := \bar{a}y^{2d} + a_{2d-1}^d \bar{a}y^{2d-1} + \cdots + a_0 \bar{a}a^{2d-1}$. Using the fact that $\bar{a} \geq 0$ on $[0, 1]$ and the set of points $(r, s)$ in the strip $[0, 1] \times R$ satisfying $\bar{a}(r) \neq 0$ is dense in the strip, one sees that $g \geq 0$ on the strip. If we are able to show that $g = \sigma + \tau x(1 - x)$, with $\sigma, \tau$ sums of squares in $R[x, y]$, then $\bar{a}(x)^{2d-1}f(x, y) = \sigma(x, \bar{a}(x)y) + \tau(x, \bar{a}(x)y)x(1 - x)$. Thus we are reduced to showing that if $b(x)f(x, y)$ has a presentation $b(x)f(x, y) = \sigma(x, y) + \tau(x, y)x(1 - x)$ for some sums of squares $\sigma(x, y), \tau(x, y)$, where $b(x) \geq 0$ on the interval $[0, 1]$ and $b(x)$ is $\pm 1$ times a product of linear factors $x - r, r \in [0, 1]$, then $f(x, y)$ also has such a presentation. The proof is by induction on the degree of $b(x)$. Suppose $x - r$ is a factor of $b(x), 0 \leq r \leq 1$. First suppose $0 < r < 1$. Then $b(x) = b(x)(x - r)^2$. Also, $\sigma(x, y)$ and $\tau(x, y)$ vanish at $x = r$, so $\sigma(x, y) = \sigma(x, y)(x - r)^2, \tau(x, y) = \tau(x, y)(x - r)^2$, with $\sigma(x, y), \tau(x, y)$ sums of squares in $R[x, y]$, and $f(x, y) = \sigma(x, y) + \tau(x, y)x(1 - x)$. If $r = 0$, then $b(x) = b(x)x$ and $\sigma(x, y) = \sigma(x, y)x^2, \tau(x, y) = \tau(x, y)x$ a sum of squares in $R[x, y]$, and $f(x, y) = \sigma(x, y)x + \tau(x, y)(1 - x)$. Using $x = x^2 + x(1 - x)$ and $1 - x = (1 - x)^2 + x(1 - x)$ this yields $b(x)f(x, y) = \sigma'(x, y) + \tau'(x, y)(1 - x)$, where $\sigma'(x, y)$ and $\tau'(x, y)$ are sums of squares in $R[x, y]$. The argument for $r = 1$ is basically the same as that for $r = 0$. □

Lemma 2.2. We may assume $f$ has only finitely many zeros in the strip $[0, 1] \times R$. 

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
Proof: If \( f = g^2h \), then, using the fact that the set of points \((a, b)\) in the strip satisfying \( g(a, b) \neq 0 \) is dense in the strip \([1\) Prop. 1.1.2], we see that \( h \geq 0 \) on the strip. So we are reduced to showing the result for \( h \). Thus we are reduced to the case where \( f \) is square-free. Since \( x \) and \( 1 - x \) do not divide \( f \) (because they do not divide \( a_{2d} \)), \( f \) has only finitely many zeros on the boundary of the strip. If some irreducible factor \( p \) of \( f \) has infinitely many zeros in the interior of the strip, then \( p \) has a non-singular zero \((a, b)\) in the interior of the strip which is not a zero of any other irreducible factor of \( f \) \([1\) Lem. 9.4.1]. Then \( p \) changes sign at \((a, b)\), but each other irreducible factor of \( f \) has constant sign in a neighborhood of \((a, b)\), contradicting the fact that \( f \) is \( \geq 0 \) on the strip. It follows that \( f \) has only finitely many zeros in the strip. \( \square \)

If \( f \) is square-free, then no irreducible factor of \( f \) can change sign in the interior of the strip, so each irreducible factor has constant sign on the strip. Replacing \( p \) by \(-p \) if necessary, for each irreducible factor \( p \) we may assume each irreducible factor of \( f \) is \( \geq 0 \) on the strip. In this way we are reduced further to the case where \( f \) itself is irreducible. But this does not seem to help us much in the proof.

3. The idea of the proof

Consider the case where the polynomial \( f(x, y) := \sum_{i=0}^{2d} a_i(x)y^i \) is positive on \([0, 1] \times \mathbb{R} \), and \( a_{2d}(x) > 0 \) on \([0, 1] \). The form \( F(x, y, z) := \sum_{i=0}^{2d} a_i(x)y^i z^{2d-i} \) is positive for \( 0 \leq x \leq 1 \), \((y, z) \neq (0, 0)\) (since \( F(x, y, z) = z^{2d} f(x, \frac{y}{z}) \) if \( z \neq 0 \)), so it achieves a positive minimum \( \epsilon \) on the compact set

\[ \{ (x, y, z) \mid 0 \leq x \leq 1 \text{ and } y^2 + z^2 = 1 \} \]

Then, on the strip \([0, 1] \times \mathbb{R} \),

\[ f(x, y) = F(x, y, 1) = F(x, \frac{y}{\sqrt{1 + y^2}}, \frac{1}{\sqrt{1 + y^2}})(1 + y^2)^d \geq \epsilon(1 + y^2)^d. \]

Using this, the argument in \([3\) Th. 5.1\] shows that \( f(x, y) \) has the required presentation. See \([3\) Th. 2.2\] for additional explanation.

In the general case, one cannot possibly have such an \( \epsilon \). The idea is to replace \( \epsilon \) by a polynomial \( \epsilon(x) \). Specifically, we look for a polynomial \( \epsilon(x) \in \mathbb{R}[x] \) such that

\[ f(x, y) \geq \epsilon(x)(1 + y^2)^d \]

holds on the strip, \( \epsilon(x) \geq 0 \) on \([0, 1] \), and, \( \forall x \in [0, 1] \), \( \epsilon(x) = 0 \) iff \( f(x, y) = 0 \) for some \( y \in \mathbb{R} \). It is always possible to find such a polynomial \( \epsilon(x) \), assuming that \( a_{2d}(x) > 0 \) on \([0, 1] \) and \( f(x, y) \) has only finitely many zeros in the strip. Once this is established, one can show that a modified version of the argument in \([3\) Th. 5.1\] carries through, with \( \epsilon \) replaced by \( \epsilon(x) \). This latter part of the argument is a bit technical: it is necessary to take pains to ensure that the continuous functions considered are analytic at the points where \( \epsilon(x) = 0 \) and to use a refined version of the Weierstrass Approximation Theorem.
4. The additional ingredients

We establish the additional results that we need in the proof of Theorem \[\text{[1.1]}\]

Lemma 4.1. Suppose \(f \in \mathbb{R}[x, y]\) is non-negative on a strip of the form \([0, \epsilon) \times \mathbb{R}\), \(\epsilon > 0\), \(f\) has only finitely many zeros in this strip, and the leading coefficient of \(f\) is positive on the interval \([0, \epsilon)\). Then there exists a real constant \(C > 0\), an even integer \(m \geq 0\), and a real number \(\delta\), \(0 < \delta \leq \epsilon\), such that \(f(x, y) \geq C x^m\) holds on the strip \([0, \delta) \times \mathbb{R}\). If \(f\) has no real zeros on the \(y\)-axis, we may take \(m = 0\).

We denote by \(k((x))\) the formal power series field over a field \(k\), i.e., the field of fractions of the formal power series ring \(k[[x]]\).

Proof. The leading coefficient of \(f\) is positive at zero; in particular, it is a unit in \(\mathbb{R}[x]\). By Puiseux’s Theorem, \(f\) factors into linear factors in \(\mathbb{C}((x^{\frac{1}{n}}))[y]\), for some \(n \geq 1\). Each root has the form \(z = \sum_{i=0}^{\infty} a_i x^{\frac{i}{n}}\), \(a_i \in \mathbb{C}\), and is a complex analytic function of \(x^{\frac{1}{n}}\) in a neighborhood of zero. \(\exists \) another root of \(f\), and \(\tau \neq z\). (If all the \(a_i\) were real, then the equation \(y = \sum_{i=0}^{\infty} a_i x^{\frac{i}{n}}\) would define a real half-branch in the zero set of \(f\), for \(x\) close to zero, \(x > 0\), contradicting our assumption that \(f\) has only finitely many zeros in the strip \([0, \epsilon) \times \mathbb{R}\).) Write \(a_i = b_i + c_i \sqrt{-1}\), \(b_i, c_i \in \mathbb{R}\), so \(z = z_1 + z_2 \sqrt{-1}\), \(z = z_1 - z_2 \sqrt{-1}\), where \(z_1 = \sum b_i x^{\frac{i}{n}}\), \(z_2 = \sum c_i x\frac{i}{n}\), so \((y - z)(y - \tau) = (y - z_1)^2 + (\frac{z_2}{n})^2\). Let \(k = \frac{m}{n}\) be the order of \(z_2\) at zero, i.e., the least \(\frac{1}{n}\) such that \(c_i \neq 0\). For any real \(x > 0\) close to zero and any \(y \in \mathbb{R}\),

\[
(\frac{y - z}{x^{\frac{k}{2}}})(\frac{y - \tau}{x^{\frac{k}{2}}}) = (\frac{y - z_1}{x^{\frac{k}{2}}})^2 + (\frac{z_2}{x^{\frac{k}{2}}})^2,
\]

and \(\frac{1}{x^{\frac{k}{2}}}\) is a real number close to the non-zero real constant \(c_{i_0}\). This implies there exists a real constant \(C > 0\) such that \(\frac{(y - z)(y - \tau)}{x^{\frac{k}{2}}} > C\) for all real \(x > 0\) sufficiently close to zero and all real \(y\). Note: If \(f\) has no real zeros on the \(y\)-axis, then \(a_0\) is not real; i.e., \(a_0 \neq 0\), so \(k = 0\). Factoring \(f\) as \(f(x, y) = a(x) \prod_{i=1}^{d} (y - z_i)(y - \tau_i)\), where \(a(x)\) is the leading coefficient, this yields rationals \(k_i \geq 0\) and real constants \(C_i > 0\)

\[
f(x, y) = a(x) \prod_{i=1}^{d} (y - z_i)(y - \tau_i) = a(x) \prod_{i=1}^{d} C_i
\]

for all real \(y\) and all real \(x > 0\) sufficiently close to zero. Finally, \(f(x, y) \geq C x^m\) for any real \(x \geq 0\) sufficiently close to zero and any \(y \in \mathbb{R}\), where \(C\) is \(\prod_{i=1}^{d} C_i\) times the minimum value of \(a(x)\) on \([0, \frac{\epsilon}{2}]\) and \(m\) is the least even integer \(\geq 2 \sum_{i=1}^{d} k_i\). \(\square\)

Lemma 4.2. Suppose \(f(x, y) = \sum_{i=0}^{2d} a_i(x)y^i\) is non-negative on the strip \([0, 1] \times \mathbb{R}\), \(f(x, y)\) has only finitely many zeros in the strip, and \(a_{2d}(x)\) is positive on the interval \([0, 1]\). Then there exists a polynomial \(\epsilon(x) \in \mathbb{R}[x]\), \(\epsilon(x) \geq 0\) on \([0, 1]\), such that \(f(x, y) \geq \epsilon(x)(1 + y^2)^d\) holds on the strip and, for each \(x \in [0, 1]\), \(\epsilon(x) = 0\) if there exists \(y \in \mathbb{R}\) such that \(f(x, y) = 0\).

Proof. For each \(r \in [0, 1]\), by Lemma 4.1 (applied to the new variables \(t = x - r\) and \(t = r - x\), or just to \(t = x - r\), resp., just to \(t = x - r\), if \(r = 0\), resp., if \(r = 1\)) we have a real constant \(C > 0\) and an even integer \(m \geq 0\) (with \(m = 0\) if \(f\) has no real zeros on the line \(x = r\)), such that \(f(x, y) \geq C(x - r)^m\) holds for all \((x, y)\) in the strip, with \(x\) sufficiently close to \(r\). By compactness of the interval \([0, 1]\), there are finitely many \(0 \leq r_1 < \cdots < r_k \leq 1\) and finitely many positive constants \(C_i\) and
even integers \( m_i \geq 0 \) (with \( m_i = 0 \) if \( f \) has no real zeros on the line \( x = r_i \)) such that, for each \((x, y)\) in the strip, \( f(x, y) \geq C_i(x - r_i)^{m_i} \), for some \( i \). We may assume each \( C_i \) is \( \leq 1 \). Then \( C_i(x - r_i)^{m_i} \geq \prod_{j=1}^{k} C_j(x - r_j)^{m_j} \) for each \( x \in [0, 1] \) and each \( i \). Thus \( f(x, y) \geq \epsilon_1(x) \) holds on the strip, where \( \epsilon_1(x) := \prod_{j=1}^{k} C_j(x - r_j)^{m_j} \).

Since \( f(x, y) \) has only finitely many zeros in the strip, there exists a real constant \( M > 0 \) such that \( f(x, y) > 0 \) if \( |y| \geq M \), \( 0 \leq x \leq 1 \). Arguing with the form \( F(x, y, z) := \sum_{i=0}^{2d} a_i(x) y^i z^{2d-i} \) as in Section 3, but with \( S^1 = \{(y, z) \mid y^2 + z^2 = 1\} \) replaced by the set \( \{(y, z) \mid |y|^2 + z^2 = 1, |y| \geq M|z|\} \), we see there exists a positive constant \( C \) such that \( f(x, y) \geq C(1 + y^2)^d \) for all \((x, y)\) in the strip satisfying \(|y| \geq M\). If \(|y| \leq M\), then \( 1 + y^2 \leq 1 + M^2 \), and \( \frac{f(x, y)}{(1+y^2)^d} \geq \frac{f(x, y)}{(1+M^2)^d} \geq C \epsilon_1(x) \). If \(|y| \geq M\), \( \frac{f(x, y)}{(1+y^2)^d} \geq C \geq \frac{C}{\epsilon_1(x)} \), where \( D := \max\{\epsilon_1(x) \mid x \in [0, 1]\} \). So, in any case, \( \frac{f(x, y)}{(1+y^2)^d} \geq C \epsilon(x) \) holds on the strip, where

\[
\epsilon(x) := \min\left\{ \frac{1}{(1+M^2)^d} \frac{C}{D} \epsilon_1(x) \right\}.
\]

Lemma 4.3. Suppose \( f \in \mathbb{R}[x, y] \) and there exists \( \epsilon > 0 \) such that \( f \) is non-negative on the strip \((-\epsilon, \epsilon) \times \mathbb{R}\) and the leading coefficient of \( f \) is positive on the interval \((-\epsilon, \epsilon)\). Then there exist \( g_1, g_2 \) polynomials in \( y \) whose coefficients are analytic functions of \( x \) defined in a neighborhood of zero, such that \( f = g_1^2 + g_2^2 \), for \( x \) sufficiently close to zero.

Proof. Let \( p \) be an irreducible factor of \( f \) in \( \mathbb{C}((x))[y] \) which is monic. By Puiseux’s Theorem, \( p \) factors in \( \mathbb{C}((x))[(y^\frac{1}{m})] \), where \( m \) is the degree of \( p \), as \( p = \prod_{\omega \in \mu_n} (y - z_\omega) \), where \( \mu_n \) denotes the group of complex \( n \)-th roots of 1, and \( z_\omega = \sum_{i=0}^{\infty} a_i \omega^i x^\frac{1}{m} \) for each \( \omega \in \mu_n \), where the \( a_i \) are complex numbers. The \( z_\omega \) are complex analytic functions of \( x^\frac{1}{m} \) in a neighborhood of zero [11, Sect. 12.3]. The coefficients of \( p \) are elementary symmetric functions of the roots and so are complex analytic functions of \( x \) in some neighborhood of zero. Denote by \( \overline{p} \) the polynomial in \( \mathbb{C}((x))[y] \) obtained from \( p \) by conjugating coefficients in the obvious way. \( \overline{p} \) is an irreducible factor of \( f \). If \( \overline{p} = p \), then \( z_1 \) coincides with one of the \( \overline{z}_\omega := \sum_{i=0}^{\infty} a_i \omega^i x^\frac{1}{m} \). This implies, in turn, that there are (two) real half-branches of \( f \) coming from \( p \). Since \( p \) changes sign at any such half-branch, \( p \) must appear in \( f \) with even multiplicity in this case. Thus \( f \) has a factorization of the form \( f = a(x) \prod_{i=1}^{k} p_i \overline{p}_i \), where each \( p_i \) is irreducible and \( a(x) \) is the leading coefficient. Then \( f = g_1^2 \), where \( g = \sqrt{a(x)} p_1 \cdots p_k \). Decomposing \( g \) as \( g = g_1 + g_2 \sqrt{-1}, g_1, g_2 \in \mathbb{R}((x))[y] \), this yields \( f = g_1^2 + g_2^2 \).

Lemma 4.4. Suppose \( f \in \mathbb{R}[x, y] \) is non-negative on the strip \([0, 1] \times \mathbb{R}\) and the leading coefficient of \( f \) is positive on the interval \([0, 1]\). Then:

1. For each \( r \in (0, 1) \), there exist \( g_1, g_2 \) polynomials in \( y \) with coefficients analytic functions in \( x \) in some neighborhood of \( r \) such that \( f = g_1^2 + g_2^2 \) holds for \( x \) sufficiently close to \( r \).
(2) There exist \( g_{ij}, i, j = 1, 2 \), polynomials in \( y \) with coefficients analytic functions in \( x \) in some neighborhood of 0 such that \( f = \sum_{i=1}^{2} g_{i1}^2 + \sum_{i=1}^{2} g_{i2}^2 x \) holds for \( x \) sufficiently close to 0.

(3) There exist \( g_{ij}, i, j = 1, 2 \), polynomials in \( y \) with coefficients analytic functions in \( x \) in some neighborhood of 1 such that \( f = \sum_{i=1}^{2} g_{i1}^2 + \sum_{i=1}^{2} g_{i2}^2 (1 - x) \) holds for \( x \) sufficiently close to 1.

Proof. For (1), apply Lemma 4.3 viewing \( f \) as a polynomial in \( x - r \) and \( y \). For (2), apply Lemma 4.3 viewing \( f \) as a polynomial in \( \sqrt{x} \) and \( y \), to obtain \( f = g_1^2 + g_2^2 \) with \( g_i \) a polynomial in \( y \) with coefficients analytic in \( \sqrt{x} \), \( i = 1, 2 \). Decomposing each of the coefficients, using \( \sum_{k} a_k x^{k} = \sum_{k} a_{2k} x^{k} + \sum_{k} a_{2k+1} x^{k} \sqrt{x} \), yields \( g_i = g_{i1} + g_{i2} \sqrt{x} \), where the \( g_{ij} \) are polynomials in \( y \) with coefficients analytic functions of \( x \) near \( x = 0 \). Expanding \( g_i^2, i = 1, 2 \), then yields \( f = \sum_{i=1}^{2} g_{i1}^2 + \sum_{i=1}^{2} g_{i2}^2 x + 2 \sum_{i=1}^{2} g_{i1} g_{i2} \sqrt{x} \), so \( f = \sum_{i=1}^{2} g_{i1}^2 + \sum_{i=1}^{2} g_{i2}^2 x \) and \( \sum_{i=1}^{2} g_{i1} g_{i2} = 0 \). The proof of (3) is similar to the proof of (2).

\[ \square \]

Proposition 4.5. Suppose \( \phi, \psi : [0, 1] \to \mathbb{R} \) are continuous functions, \( \phi(x) \leq \psi(x) \) for all \( x \in [0, 1] \), and \( \phi(x) < \psi(x) \) for all but finitely many \( x \in [0, 1] \). If \( \phi \) and \( \psi \) are analytic at each point \( a \in [0, 1] \) where \( \phi(a) = \psi(a) \), then there exists a polynomial \( p(x) \in \mathbb{R}[x] \) such that \( \phi(x) \leq p(x) \leq \psi(x) \) holds for all \( x \in [0, 1] \).

Proof. Induct on the number of points \( a \in [0, 1] \) satisfying \( \phi(a) = \psi(a) \). If there are no such points, existence of \( p(x) \) follows from the Weierstrass Approximation Theorem. Suppose \( a \in [0, 1] \) is such that \( \phi(a) = \psi(a) \). Let \( k \) be the vanishing order of \( \psi - \phi \) at \( a \). If \( a \in (0, 1) \), then \( k \) is even. In this case, \( \phi(x) = f(x) + (x - a)^k \phi_1(x) \), \( \psi(x) = f(x) + (x - a)^k \psi_1(x) \), where \( f(x) \in \mathbb{R}[x], \phi_1(x), \psi_1(x) \) are analytic at \( a \), and \( \phi_1(a) < \psi_1(a) \). Extend \( \phi_1, \psi_1 \) to continuous functions \( \phi_1, \psi_1 : [0, 1] \to \mathbb{R} \) by defining \( \phi_1(x) = \frac{\phi(x) - \phi(a)}{(x-a)^k} \), \( \psi_1(x) = \frac{\psi(x) - \psi(a)}{(x-a)^k} \) for \( x \neq a \). Then \( \phi_1(x) \leq \psi_1(x) \) for all \( x \in [0, 1] \), and, \( \forall b \in [0, 1], \phi_1(b) = \psi_1(b) \iff \phi(b) = \psi(b) \) and \( b \neq a \). By induction we have \( p_1(x) \in \mathbb{R}[x] \) such that \( \phi_1(x) \leq p_1(x) \leq \psi_1(x) \) on \( [0, 1] \). Take \( p(x) = f(x) + (x - a)^k p_1(x) \). The case where \( a = 0 \) and the case where \( a = 1 \) are dealt with in a similar fashion. \( \square \)

5. The End of the Proof

Let \( f(x, y) = \sum_{i=0}^{2d} a_i(x)y^i, d \geq 1 \). By Lemmas 2.1 and 2.2 we can assume \( a_{2d}(x) > 0 \) on the interval and \( f(x, y) \) has only finitely many zeros in \([0, 1] \times \mathbb{R}\). By Lemma 4.2 there exists a polynomial \( \epsilon(x) \in \mathbb{R}[x] \) such that \( f(x, y) \geq \epsilon(x)(1 + y^2)^d \) on the strip, \( \epsilon(x) \geq 0 \) on \([0, 1]\), and \( \epsilon(x) = 0 \) iff \( \exists y \in \mathbb{R} \) with \( f(x, y) = 0 \). Let \( f_1(x, y) := f(x, y) - \epsilon(x)(1 + y^2)^d \). Then \( f_1 \) is \( \geq 0 \) on the strip. Replacing \( \epsilon(x) \) by \( \frac{\epsilon(x)}{N} \), \( N > 1 \), if necessary, we can assume \( f_1 \) has degree \( 2d \) (as a polynomial in \( y \)) and the leading coefficient of \( f_1 \) is positive on \([0, 1]\).

By Lemma 4.3 for each \( r \in [0, 1] \), there exists an open neighborhood \( U(r) \) of \( r \) in \( \mathbb{R} \) such that \( f_1 \) decomposes as

\[
 f_1 = \sum_{j=1}^{2} g_{0j}(r)^2 + \sum_{j=1}^{2} g_{1j}(r)^2 x + \sum_{j=1}^{2} g_{2j}(r)^2 (1 - x)
\]

\[ \text{Proposition 4.5 is probably well-known. The author only became aware of Proposition 4.5 and its simple proof through reading an unpublished manuscript of V. Powers.} \]
on \(U(r) \times \mathbb{R}\), where the \(g_{ij}(r)\) are polynomials in \(y\) (of degree \(\leq d\)) whose coefficients are analytic functions of \(x\), for \(x \in U(r)\). By compactness of \([0,1]\), finitely many of the \(U(r)\) cover \([0,1]\), say \(U(r_1), \ldots, U(r_k)\) cover \([0,1]\). Choose a continuous partition of unity \(1 = \nu_1 + \cdots + \nu_k\) on \([0,1]\), with \(0 \leq \nu_k \leq 1\) on \([0,1]\) and \(\text{supp}(\nu_k) \subseteq U(r_k)\) for \(k = 1, \ldots, \ell\), having the additional property that, for each root \(r\) of \(\epsilon(x)\) in \([0,1]\), there is just one \(k\) such that \(\nu_k(x) \neq 0\) close to \(r\) (so \(\nu_k(x) = 1\) for \(x\) close to \(r\)). One way to ensure the last property is to shrink the covering sets \(U(r_k)\) ahead of time so that each root \(r\) of \(\epsilon(x)\) in \([0,1]\) lies in some unique \(U(r_k)\). Then \(f_1\) decomposes as

\[
f_1 = \sum_{k=1}^{\ell} \nu_k f_1 = \sum_{k=1}^{\ell} \left( \sum_{j=1}^{2} \phi_{0jk}^2 + \sum_{j=1}^{2} \phi_{1jk}^2 x + \sum_{j=1}^{2} \phi_{2jk}^2 (1-x) \right)
\]

on \([0,1] \times \mathbb{R}\), where \(\phi_{ijk}\) denotes the polynomial of degree \(\leq d\) in \(y\) whose coefficients are the functions from \([0,1]\) to \(\mathbb{R}\) obtained by extending the corresponding coefficients of \(\sqrt{\nu_k} g_{ij}(r_k)\) by zero off \(U(r_k)\). The coefficients of the \(\phi_{ijk}\) are continuous on \([0,1]\) and analytic at each of the roots of \(\epsilon(x)\) in \([0,1]\) (since \(\nu_k\) is constantly 0 or 1 in a neighborhood of each of these roots).

By Proposition 4.5, for each real \(N > 0\) and each triple \(i, j, k\), there exists a polynomial \(h_{ijk}\) of degree \(\leq d\) in \(y\) with coefficients in \(\mathbb{R}[x]\) such that, for each coefficient \(u\) of \(\phi_{ijk}\), the corresponding coefficient \(w\) of \(h_{ijk}\) satisfies

\[
u(x) - \frac{\epsilon(x)}{N} \leq w(x) \leq u(x) + \frac{\epsilon(x)}{N}, \text{ for each } x \in [0,1].
\]

At this point we proceed as in the proof of [3, Th. 5.1], approximating the coefficients of the \(\phi_{ijk}\) closely by polynomials (by taking \(N\) sufficiently large), to obtain polynomials \(h_{ijk}\) of degree \(\leq d\) in \(y\) with coefficients in \(\mathbb{R}[x]\) such that

\[
f_1(x, y) = \sum_{k=1}^{\ell} \left( \sum_{j=1}^{2} h_{0jk}(x, y)^2 + \sum_{j=1}^{2} h_{1jk}(x, y)^2 x + \sum_{j=1}^{2} h_{2jk}(x, y)^2 (1-x) \right) + \sum_{i=0}^{2d} b_i(x) y^i,
\]

\(b_i(x) \in \mathbb{R}[x]\), \(|b_i(x)| \leq \frac{2\epsilon(x)}{N}\) on \([0,1]\), \(i = 0, \ldots, 2d\). Combining this with \(f(x, y) = f_1(x, y) + \epsilon(x)(1 + y^2)^d\) yields \(f(x, y) = s_1(x, y) + s_2(x, y) + s_3(x, y)\), where

\[
s_1(x, y) := \sum_{k=1}^{\ell} \left( \sum_{j=1}^{2} h_{0jk}(x, y)^2 + \sum_{j=1}^{2} h_{1jk}(x, y)^2 x + \sum_{j=1}^{2} h_{2jk}(x, y)^2 (1-x) \right),
\]

\[
s_2(x, y) := \frac{\epsilon(x)}{5}(2 + y + 3y^2 + y^3 + 3y^4 + \cdots + y^{2d-1} + 2y^{2d}) + \sum_{i=0}^{2d} b_i(x) y^i.
\]

\[
s_3(x, y) := \epsilon(x)[(1 + y^2)^d - \frac{2}{5}(2 + y + 3y^2 + y^3 + 3y^4 + \cdots + y^{2d-1} + 2y^{2d})].
\]

Let \(T\) denote the preordering of \(\mathbb{R}[x, y]\) generated by \(x(1-x)\). As pointed out earlier, \(x, 1-x \in T\). Clearly \(s_1(x, y) \in T\). The argument in [3, Th. 5.1] shows that

\footnote{Applying Lemma 4.4 we can choose the \(g_{ij}(r)\) so that \(g_{2j}(r) = 0, j = 1, 2, \text{if } r = 0; g_{1j}(r) = 0, j = 1, 2, \text{if } r = 1; \text{and } g_{1j}(r) = g_{2j}(r) = 0, j = 1, 2, \text{if } 0 < r < 1.}
$s_2(x, y) \in T$. In more detail, since $|b_i(x)| \leq \frac{2}{5} \epsilon(x)$ on $[0, 1]$, \( \frac{2}{5} \epsilon(x) \pm b_i(x) \in T \), by [3 Th. 2.2] or [4 Prop. 2.7.3], for $i = 0, \ldots, 2d$. This yields

\[(5.1) \quad \frac{2}{5} \epsilon(x)y^i + b_i(x)y^i \in T, \quad \text{for } i \text{ even.}
\]

For $i$ odd, say $i = 2m + 1$, the identity $y^{2m+1} = \frac{2}{5}y^{2m}(y+1)^2 - y^2 - 1$ plus the fact that $\frac{2}{5} \epsilon(x)y^{2m}(y+1)^2 + b_i(x)y^{2m}(y+1)^2$, $\frac{2}{5} \epsilon(x)y^{2m}y - b_i(x)y^{2m}y^2$ and $\frac{2}{5} \epsilon(x)y^{2m} - b_i(x)y^{2m}$ all belong to $T$ to obtain

\[(5.2) \quad \frac{2}{5} \epsilon(x)(y^{i+1} + y^i + y^{i-1}) + b_i(x)y^i \in T, \quad \text{for } i \text{ odd.}
\]

Adding together the various terms of type (5.1) and (5.2), for $i = 0, \ldots, 2d$, we see that $s_2(x, y) \in T$. The fact that $s_3(x, y)$ belongs to $T$ follows from the identity

\[
(1 + y^2)^d - \frac{2}{5}(2 + y + 3y^2 + y^3 + 3y^4 + \cdots + y^{2d-1} + 2y^{2d})
\]

\[
= (1 + y^2)^d + \frac{1}{5}(1 + y^2 + \cdots + y^{2d-2})(1 - y)^2
\]

\[
- \frac{8}{5}(y^2 + y^4 + \cdots + y^{2d-2}) - (1 + y^{2d})
\]

\[
= \frac{1}{5}(1 + y^2 + \cdots + y^{2d-2})(1 - y)^2 + \sum_{i=1}^{d-1} \binom{d}{i} - \frac{8}{5}y^{2i}.
\]

This means, finally, that $f(x, y) = s_1(x, y) + s_2(x, y) + s_3(x, y) \in T$. \hfill \Box

\begin{thebibliography}{10}


\end{thebibliography}

DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF SASKATCHEWAN, 106 WIGGINS
ROAD, SASKATOON, SK, CANADA, S7N 5E6
E-mail address: marshall@math.usask.ca