When does a polynomial ideal contain a positive polynomial?

Manfred Einsiedler [a] and Selim Tuncel [b]

[a] *Institut für Mathematik, Strudlhofgasse 4, A-1090 Vienna, Austria*

[b] *Department of Mathematics, University of Washington, Seattle, WA 98195, USA*

Abstract

We use Gröbner bases and a theorem of Handelman to show that an ideal I of $\mathbb{R}[x_1, \ldots, x_k]$ contains a polynomial with positive coefficients if and only if no initial ideal $in_v(I)$, $v \in \mathbb{R}^k$, has a positive zero.

Let $R = \mathbb{R}[x_1, \ldots, x_k]$, $R^+ = \mathbb{R}^+[x_1, \ldots, x_k]$ and, considering Laurent polynomials, let $\widetilde{R} = \mathbb{R}[x_1^{\pm}, \ldots, x_k^{\pm}]$, $\widetilde{R}^+ = \mathbb{R}^+[x_1^{\pm}, \ldots, x_k^{\pm}]$. For $a = (a_1, \ldots, a_k) \in \mathbb{Z}^k$, write $x^a = x_1^{a_1} \cdots x_k^{a_k}$ and denote the coefficient of x^a in $p \in \widetilde{R}$ by p_a. Then $p = \sum_{a \in \mathbb{Z}^k} p_a x^a$ and the Newton polytope $N(p)$ of p is the convex hull of the finite set $\mathrm{Log}(p) = \{a \in \mathbb{Z}^k : p_a \neq 0\}$. For $v \in \mathbb{R}^k$, let $in_v(p)$ be the sum of $p_a x^a$ over those $a \in \mathrm{Log}(p)$ for which the dot product $a \cdot v$ is maximal.

For an ideal $I \subset R$ and $v \in \mathbb{R}^k$ we have the initial ideal $in_v(I) = \langle in_v(p) : p \in I \rangle \subset R$ and the corresponding variety $\mathcal{V}(in_v(I)) = \{z \in \mathbb{C}^k : in_v(p)(z) = 0 \; \forall \, p \in I\}$. Observe that in the case $v = 0$ the ideal $in_v(I)$ equals I. We write $\mathbb{R}^{++}$ for the positive reals.

Theorem. *An ideal I of R contains a nonzero element of R^+ if and only if $(\mathbb{R}^{++})^k \cap \mathcal{V}(in_v(I)) = \emptyset$ for all $v \in \mathbb{R}^k$.*

It will be clear that there are analogous statements for ideals of $\widetilde{R}$, as well as for ideals of polynomial (or Laurent polynomial) rings over $\mathbb{Q}$ or $\mathbb{Z}$ instead of $\mathbb{R}$.

The question "When does a submodule M of R^n contain an element of $(R^+ \setminus \{0\})^n$?" will be answered in a longer sequel. The present paper, in dealing with the simpler case $n = 1$, highlights the utility of Gröbner bases in positivity problems.

One ingredient of our proof will be the following theorem of Handelman which deals with the case of a principal ideal.

Handelman's Theorem [3]. *For $p \in R$ the following are equivalent.*

(a) There exists $q \in R$ such that $qp \in R^+ \setminus \{0\}$.

(b) We have $in_v(p)(z) \neq 0$ for every $v \in \mathbb{R}^k$ and $z \in (\mathbb{R}^{++})^k$.

A short and self-contained account of the proof of Handelman's theorem may be found in [2].

The other ingredient we need is the basic theory of Gröbner bases. Everything we use from this theory can be found in the first 50 pages of [1].

Monomials of R are in bijective correspondence with $(\mathbb{Z}^+)^k$. A *term order* on $(\mathbb{Z}^+)^k$ is a total order $\prec$ satisfying the following two conditions:

(i) $0 \prec a$ for all nonzero $a \in (\mathbb{Z}^+)^k$,

(ii) $a \prec b$ implies $a + c \prec b + c$ for all $a, b, c \in (\mathbb{Z}^+)^k$.

Fix an ideal $I \subset R$. For a term order $\prec$, we let $in_\prec(p)$ denote the unique initial (or leading) monomial of $p \in R$ and have the initial ideal $in_\prec(I) = \langle in_\prec(p) : p \in I \rangle \subset R$. Elements $f_1, \ldots, f_l \in I$ form a Gröbner basis for I with respect to $\prec$ if and only if $in_\prec(I) = \langle in_\prec(f_i) : i = 1, \ldots, l \rangle$. Though there are infinitely many term orders, it is shown on p. 1–2 of [4] that I has finitely many initial ideals and, therefore, a universal Gröbner basis. That is, there exist $f_1, \ldots f_l \in I$ which form a Gröbner basis of I with respect to every term order. (The existence of universal Gröbner bases was originally established in [5].)

Since we are interested in $in_v(I)$ for arbitrary $v \in \mathbb{R}^k$, we need to work around the fact that the dot product with v yields term orders on R only when v is positive. One way to do this is to introduce new variables $y_1, \ldots, y_k$ and obtain from I an ideal that is homogeneous in each pair x_i, y_i. We will take another tack: Let $\delta = (\delta_1, \ldots, \delta_k) \in \{-1, 1\}^k$. Pick $a \in \mathbb{Z}^k$ so that $x^a f_1, \ldots, x^a f_l \in \mathbb{R}[x_1^{\delta_1}, \ldots, x_k^{\delta_k}]$, and let $f_{\delta,1}, \ldots, f_{\delta,l_\delta}$ be a universal Gröbner basis for the ideal $\langle x^a f_1, \ldots, x^a f_l \rangle$ of $\mathbb{R}[x_1^{\delta_1}, \ldots, x_k^{\delta_k}]$. List the union of $f_{\delta,1}, \ldots, f_{\delta,l_\delta}$ over $\delta \in \{-1, 1\}^k$ as $g_1, \ldots, g_m$. Let $\prec$ be an arbitrary term order on $(\mathbb{Z}^+)^k$.

Lemma. *Let $p \in I$ and $v \in \mathbb{R}^k$. There exist $\alpha_i \in \widetilde{R}$ such that $p = \sum_{i=1}^m \alpha_i g_i$ and*

$$\max(N(p) \cdot v) = \max_i \{\max(N(\alpha_i g_i) \cdot v)\}. \qquad (*)$$

Putting $S(v) = \{1 \le i \le m : \max(N(\alpha_i\, g_i) \cdot v) = \max(N(p) \cdot v)\}$*, we have*

$$in_v(p) = \sum_{i \in S(v)} in_v(\alpha_i)\, in_v(g_i)\,. \qquad (**)$$

PROOF. The second statement follows easily from the first. For the first statement, define $\delta \in \{-1,1\}^k$ by letting $\delta_j = 1$ if $v_j \ge 0$ and $\delta_j = -1$ if $v_j < 0$. For $a, b \in (\mathbb{Z}^+)^k$, put $(x_1^{\delta_1})^{a_1} \cdots (x_k^{\delta_k})^{a_k} \prec_v (x_1^{\delta_1})^{b_1} \cdots (x_k^{\delta_k})^{b_k}$ if

(i) $\sum_{j=1}^{k} v_j \delta_j a_j < \sum_{j=1}^{k} v_j \delta_j b_j$, or

(ii) $\sum_{j=1}^{k} v_j \delta_j a_j = \sum_{j=1}^{k} v_j \delta_j b_j$ and $a \prec b$.

This defines a term order $\prec_v$ on the monomials of $\mathbb{R}[x_1^{\delta_1}, \ldots, x_k^{\delta_k}]$. Now consider that $a \in \mathbb{Z}^k$ involved in the definition of $\{f_{\delta,1}, \ldots, f_{\delta,l_\delta}\}$. Find $b \in \mathbb{Z}^k$ so that $x^b p$ lies in the ideal $\langle x^a f_1, \ldots, x^a f_l \rangle$ of $\mathbb{R}[x_1^{\delta_1}, \ldots, x_k^{\delta_k}]$. Apply the division algorithm [1] to $x^b p$ and the subset $\{f_{\delta,1}, \ldots, f_{\delta,l_\delta}\}$ of $\{g_1, \ldots, g_m\}$ to find $\alpha_i \in x^{-b}\, \mathbb{R}[x_1^{\delta_1}, \ldots, x_k^{\delta_k}]$ such that $p = \sum_{i=1}^m \alpha_i g_i$, we have $\alpha_i = 0$ if $g_i \notin \{f_{\delta,1}, \ldots, f_{\delta,l_\delta}\}$, and

$$\mathrm{in}_{\prec_v}(x^b p) = \max\{\mathrm{in}_{\prec_v}(x^b \alpha_i g_i) : i = 1, \ldots, m\}.$$

The last equation straightforwardly implies the desired equality $(*)$. $\square$

Remark. *It is evident from the definition of* $g_1, \ldots, g_m$ *that a monomial multiple of each* g_i *lies in* I*. This fact and the above lemma imply that*

$$\mathcal{V}(in_v(I)) \setminus \{0\} = \{z \in \mathbb{C}^k : z \ne 0 \text{ and } in_v(g_i)(z) = 0 \text{ for } i = 1, \ldots, m\}.$$

Hence, an equivalent formulation of the theorem is that $I \cap R^+$ *contains a nontrivial polynomial if and only if for every* $v \in \mathbb{R}^k$ *the set* $\{in_v(g_1), \ldots, in_v(g_m)\}$ *has no common zero in* $(\mathbb{R}^{++})^k$*. One easily finds a finite set of vectors* v *which is sufficient for checking the last condition. In fact, if we let* $G = \prod_{i=1}^m g_i$ *and for each face* F *of* $W(G)$ *pick a vector* v_F *such that* $W(in_{v_F}(G)) = F$*, it suffices to check the condition for the finite set of vectors* $\{v_F\}$*.*

Proof of the theorem. Suppose $p \in I \cap R^+$ and $p \ne 0$. For $v \in \mathbb{R}^k$ and $z \in (\mathbb{R}^{++})^k$ we have $in_v(p) \in R^+$ and, therefore, $in_v(p)(z) > 0$. Considering $(**)$, we see that $in_v(g_i)(z) \ne 0$ for some $i \in S(v)$.

Conversely, suppose that $in_v(g_1), \ldots, in_v(g_m)$ do not have a common root in $(\mathbb{R}^{++})^k$ for any $v \in \mathbb{R}^k$. Let $\tilde{g}_i$ be a monomial multiple of g_i such that $\tilde{g}_i \in I$.

Then $in_v(\tilde{g}_1), \ldots, in_v(\tilde{g}_m)$ do not have a common root in $(\mathbb{R}^{++})^k$ for any $v \in \mathbb{R}^k$. Let h_i be the sum of x^a over all a in the set

$$\mathrm{Log}\left(\prod_{\substack{j\in\{1,\ldots,m\},\\ j\neq i}} \tilde{g}_j^2\right).$$

Note that, on $(\mathbb{R}^{++})^k$, each $in_v(h_i\, \tilde{g}_i^2)$ is nonnegative and has the same roots as $in_v(\tilde{g}_i)$. As $N(h_i\, \tilde{g}_i^2)$ is independent of i, we conclude that $p \equiv \sum_{i=1}^m h_i\, \tilde{g}_i^2$ satisfies (b) of Handelman's theorem. By Handelman's theorem, $qp \in R^+ \cap I$ for some nonzero $q \in R$. (In fact, letting $f = \sum_{a\in \mathrm{Log}(p)} x^a$, the proof of Handelman's theorem reveals that we can take $q = f^n$ for some positive integer n.) □

We end the paper with some examples.

Examples. (i) Take $k = 2$, and write $x = x_1, y = x_2$ and $w = (0, 1)$. Consider $p = 1 + y - 2xy + x^2y$, $q = 1 + y - xy + x^2y$ and the principal ideals $I = \langle p\rangle$, $J = \langle q\rangle$. Note that $in_v(p) = in_v(q) \in R^+$ for $v \neq w$, while $in_w(p) = (x-1)^2y$ and $in_w(q) = ((x-1)^2 + x)y$. By Handelman's theorem, $I \cap R^+ = \{0\}$, while $J \cap R^+$ contains a nonzero element. (For instance, $(1 + x)q \in R^+$.)

(ii) Now take $k = 3$ and write $x = x_1, y = x_2, z = x_3$. Consider $p = 1 + (2x + 2y)z + (1-x)^2(1-y)^2z^2$, $q = 1 + (x + y)z + (1-x)^2(1-y)^2z^2$ and $s = 2 + x^2 + y^2 + (1-x)^2(1-y)^2z$, and the ideals $I = \langle p, s\rangle$, $J = \langle q, s\rangle$. Let D be the subset of $\mathbb{R}^3$ consisting of vectors of the form $(0, a, b)$ and $(a, 0, b)$ for $a \geq 0$, $b > 0$. Observe that we have $in_v(s)(x, y, z) > 0$ for all $x, y, z > 0$, provided $v \notin D$. In the case of J and $v \in D$, the polynomial

$$in_v(sz-q) = in_v(-1+(1-x+x^2+1-y+y^2)z) = in_v((1-x)^2+(1-y)^2+x+y)z$$

is (numerically) positive for all $x, y, z > 0$. By our theorem, $J \cap R^+$ contains a nonzero element. Turning to I, let $w = (0, 0, 1)$ and consider the lexicographic order $\prec$ with $y \prec x \prec z$. A Gröbner basis of I with respect to $\prec$ is given by

$$f = 5 - 6x - 6y - 2y^3 + y^4 - 2x^3 + x^4 + 4xy - 4xy^2 - 4x^2y + 3x^2y^2 + 5x^2 + 5y^2,$$

$$g = -3 - x^2 + 2y - 2y^2 + (1-y)^4z,$$
$$h = -1 + ((1-x)^2 + (1-y)^2)z.$$

It follows (see the lemma above and its proof) that

$$in_w(I) = \langle in_w(f), in_w(g), in_w(h)\rangle.$$

Since $in_w(f) = f$, $in_w(g) = (1-y)^4z$ and $in_w(h) = ((1-x)^2 + (1-y)^2)z$ vanish when $x = y = z = 1$, we have $I \cap R^+ = \emptyset$.

We thank Klaus Schmidt and the Erwin Schroedinger Institute, Vienna for their hospitality, and Doug Lind and Bernd Sturmfels for discussions. We were partially supported by FWF grant P12250-MAT and NSF grant DMS-9622866, respectively.

References

[1] W. Adams and P. Loustaunau, *An Introduction to Gröbner Bases*, Amer. Math. Soc., Providence, R. I., 1994.

[2] V. De Angelis and S. Tuncel, Handelman's theorem on polynomials with positive multiples, to appear in *Codes, Systems, and Graphical Models*, IMA Volumes in Mathematics and its Applications, Springer.

[3] D. Handelman, Positive polynomials and product type actions of compact groups, *Mem. Amer. Math. Soc.*, vol. 320, 1985.

[4] B. Sturmfels, *Gröbner Bases and Convex Polytopes*, Amer. Math. Soc., Providence, R. I., 1996.

[5] V. Weispfenning, Contructing universal Gröbner bases, in Proceedings AAECC 5, Menorca 1987, *Springer Lecture Notes in Computer Science* **356**, pp. 408–417.