

On a Frank-Wolfe Type Theorem in Cubic Optimization

Diethard Klatte *

4 May 2018, revised 15 May 2018

Abstract. A classical result due to Frank and Wolfe (1956) says that a quadratic function f attains its supremum on a nonempty polyhedron M if f is bounded from above on M . In this note, we present a stringent proof of the extension of this result to cubic optimization (known from Andronov, Belousov and Shironin (1982)). Further, we discuss related results. In particular, we bring back to attention Kummer's (1977) generalization of the Frank-Wolfe theorem to the case that f is quadratic, but M is the Minkowski sum of a compact set and a polyhedral cone.

Keywords. Existence of maxima, cubic optimization, quadratic optimization, Frank-Wolfe theorem, continuity of optimal values.

2010 Mathematics Subject Classification. 90C26, 90C20, 90C31.

1 Introduction

In 1982, Andronov, Belousov and Shironin [1] published the following extension of the famous Frank-Wolfe theorem [2] (cf. also [3, 4]) in quadratic optimization to the case of cubic polynomials:

Theorem 1 *Let $f: \mathbb{R}^n \rightarrow \mathbb{R}$ be a polynomial of degree ≤ 3 and $M \subset \mathbb{R}^n$ be a nonempty polyhedral set. If f is bounded from above on M , then f attains its supremum at some point of M .*

If a statement of this form is true for a given class of functions f and sets M , we will speak of a *Frank-Wolfe type theorem*. In discussions with colleagues interested in such theorems, the author noticed certain doubts about the correctness of the proof of Theorem 1, as it was given in [1] or [5, §5.1]. Indeed, the proof is rather compactly written, in particular, some arguments are missing or have a lack of clarity. On the other hand, the idea of the proof in [1, 5] is very nice and elegant, and the gaps are not difficult to fill. To top it all, one finds sometimes even the (unverified) claim that the result is not true, cf. e.g. [2, p. 109] or [6, pp. 35-36]. So we think it is desirable to give a stringent proof of Theorem 1; this will be done below. We would like to present this short note also as a tribute to E.G. Belousov who strongly contributed to optimization theory, in particular to parametric optimization; he passed away in February 2018.

Another impetus to write this note comes from a recent paper by Martínez-Legaz, Noll and Sosa [7] (2018) who answer the interesting question which sets are so-called *Frank-and-Wolfe sets*, that is, convex sets M with the property that every quadratic function which is bounded from above on M attains its supremum on M . One of their main theorems ([7, Thm. 1]) says that the Minkowski sum of a convex compact set and a polyhedral cone is a Frank-and-Wolfe set. However, the authors of this paper were not aware that this is a classical result due to Kummer [8, §6] (1977).

*IBW, Universität Zürich, Moussonstrasse 15, CH-8044 Zürich, Switzerland, Email: diethard.klatte@uzh.ch.

In Sect. 2, Theorem 1 will be proved, where we essentially follow the approach of [1], but add more detailed arguments. In Sect. 3, we recall Kummer's generalization of the Frank-Wolfe theorem, discuss a few consequences of the main statements and refer to related work. Though the results presented here are not new, we think that our paper is of some value for methodological and bibliographical reasons.

Finishing this section, let us recall some notation, following Rockafellar [9]. A *polyhedral set* (or *polyhedron*) $M \subset \mathbb{R}^n$ is defined as the intersection of finitely many closed halfspaces of \mathbb{R}^n , so it can be represented by $M = \{x \mid Ax \leq b\}$ with some suitable matrix A and vector b . F is called a *face of M* if $F \neq \emptyset$ and $F = \{x \in M \mid A'x = b'\}$, where $A'x \leq b'$ is a subsystem of $Ax \leq b$. By $\dim M$ and $\text{ri } M$ we denote the dimension and the relative interior of M , respectively, while $O^+M = \{d \mid x + td \in M \ \forall t \geq 0 \ \forall x \in M\}$ is the *recession cone of M* and has in the representation of M above the form $O^+M = \{d \mid Ad \leq 0\}$. A *polyhedral cone* $D \subset \mathbb{R}^n$ is both a cone and a polyhedral set, where D is said to be a *cone* if $0 \in D$ and $td \in D$ for all $t \geq 0$, $d \in D$. A bounded polyhedron is also called a *polytope*. Recall M is a polyhedral set if and only if it can be represented as the Minkowski sum of a polytope and a polyhedral cone [9, Thm. 19.1].

2 Proof of the main result

Let us start with a standard lemma, for completeness we adapt the proof from [1].

Lemma 1 *Let the set $X \subset \mathbb{R}^n$ be nonempty and closed, let $f : X \rightarrow \mathbb{R}$ be continuous on X and suppose that $s = \sup_{\xi \in X} f(\xi) > f(x)$ for all $x \in X$ (the case $s = +\infty$ is not excluded here). Then there exists a sequence $\{x^j\} \subset X$ such that*

- (i) $f(x^{j-1}) < f(x^j) \rightarrow s$, $\|x^{j-1}\| < \|x^j\| \rightarrow +\infty$ and
- (ii) for each $x \in X$, one has $\|x\| < \|x^j\| \Rightarrow f(x) < f(x^j)$.

Proof. For $\bar{x} \in X$ and $j = 1, 2, \dots$, using the continuity of f , we define the (nonempty) compact sets

$$X_j = X \cap \{x \in \mathbb{R}^n \mid \|x\| \leq \|\bar{x}\| + j\}, \quad Y_j = \{x \in X_j \mid f(x) = \max_{\xi \in X_j} f(\xi)\},$$

and a sequence of points $x^j \in Y_j$ satisfying $\min_{x \in Y_j} \|x\| = \|x^j\|$. Obviously,

$$f(x^{j-1}) \leq f(x^j) < s \quad \text{and} \quad f(x^j) \rightarrow s \quad \text{for } j \rightarrow +\infty,$$

which entails $\|x^j\| \rightarrow +\infty$ since otherwise there is some cluster point $\bar{x} \in X$ of $\{x^j\}$ with $s = f(\bar{x})$, contradicting the assumption. This implies (i) by choosing a monotonously increasing subsequence of $\{f(x^j)\}$ if necessary. Further, by construction, $x \in X$ with $\|x\| < \|x^j\|$ implies $x \in X_j$, and, since x^j is by definition norm-minimal in Y_j , one has $f(x) < f(x^j)$. Thus (ii) is shown. \square

Proof of Theorem 1. (Indirect proof) If the theorem is not true, there is some constraint polyhedron of *smallest dimension* m such that the statement of the theorem fails for some polynomial of degree ≤ 3 . Let M be such a polyhedron and f such a polynomial. Then, obviously, $m = \dim M > 0$ and

$$\begin{aligned} f(x) \text{ does not attain its (finite) supremum on } M, \text{ i.e.,} \\ \sup_{\xi \in M} f(\xi) > f(x) \quad \forall x \in M. \end{aligned} \tag{1}$$

Then, by Lemma 1, there are sequences of points $x^j \in M$ and directions $d^j \in \mathbb{R}^n$ with $\|d^j\| = 1$, $j \rightarrow \infty$, such that

$$x^j = \tau_j d^j, \quad d^j \rightarrow d, \quad 0 < \tau_{j-1} < \tau_j \rightarrow +\infty, \quad f(x^{j-1}) < f(x^j) \rightarrow \sup_{\xi \in M} f(\xi), \quad (2)$$

and for all $x \in M$,

$$\|x\| < \|x^j\| \Rightarrow f(x) < f(x^j). \quad (3)$$

First we observe

$$d \in O^+M.$$

Indeed, using a representation $M = \{x \mid Ax \leq b\}$, then (2) and $x^j \in M$ imply that $Ad^j = (\tau_j)^{-1}Ax^j \leq (\tau_j)^{-1}b$ and (by taking the limits) $Ad \leq 0$, i.e., $d \in O^+M$.

Further, there is some j' such that

$$x^j \in \text{ri } M \text{ for } j \geq j'. \quad (4)$$

This is trivial if M is an affine set. Otherwise, if infinitely many x^j are points of $M \setminus \text{ri } M$, we can choose an infinite subsequence $\{x^{j_k}\}$ belonging to the same face F (of the finitely many faces) of M , where $\dim F < m$. Therefore, the theorem is true for the problem $\max f(x)$ s.t. $x \in F$, and one finds $\bar{x} \in F$ such that $f(\bar{x}) \geq f(x^{j_k}) \rightarrow s = \sup_{\xi \in M} f(\xi)$ which contradicts $s > f(\bar{x})$ according to (1). Hence, for each $j \geq j'$, one has

$$x^j - \mu_j d \in M \quad \text{for sufficiently small } \mu_j > 0. \quad (5)$$

Let $\mu_j \in]0, 1]$ fulfill (5). Further, let j' be already large enough such that for all $j \geq j'$, also

$$\tau_j - \mu_j > 0 \quad \text{and} \quad \|d^j\| - \|d^j - d\| > 0$$

are true (which is implied by $\tau_j \rightarrow +\infty$, $d^j \rightarrow d$ and $\|d^j\| = 1$), and therefore

$$\begin{aligned} \|x^j - \mu_j d\| = \|\tau_j d^j - \mu_j d\| &\leq \|\tau_j d^j - \mu_j d^j\| + \|\mu_j d^j - \mu_j d\| \\ &= \tau_j \|d^j\| - \mu_j \|d^j\| + \mu_j \|d^j - d\| \\ &< \tau_j \|d^j\| = \|x^j\|. \end{aligned}$$

This gives for $j \geq j'$, by applying (3) and (5),

$$f(x^j - \mu_j d) < f(x^j). \quad (6)$$

Using $d \in O^+M$, we see $R_j = \{x^j + td \mid t \geq 0\} \subset M$. By (1), the polynomial $f(x)$ is not constant, but bounded from above on the ray R_j , hence $\lim_{t \rightarrow +\infty} f(x^j + td) = -\infty$ for all j .

After fixing $j = i \geq j'$ we then have

$$\exists t_i > 0 : \quad f(x^i + t_i d) < f(x^i). \quad (7)$$

Now we define for each $k = i + 1, i + 2, \dots$ the vector $v^k = \|x^k - x^i\|^{-1}(x^k - x^i)$, the line

$$L_k = \{x^i + tv^k \mid -\infty < t < +\infty\}$$

and the restriction $F_k(t)$ of $f(x)$ to this line, i.e.,

$$F_k(t) = f(x^i + tv^k), \quad t \in \mathbb{R}.$$

Since $d^k \rightarrow d$, $\tau_k \rightarrow +\infty$ and $\|\tau_k d^k - \tau_i d^i\| \rightarrow +\infty$ if $k \rightarrow +\infty$ (recall i is fixed), it follows

$$v^k = \frac{\tau_k d^k - \tau_i d^i}{\|\tau_k d^k - \tau_i d^i\|} = \frac{\tau_k d^k}{\|\tau_k d^k - \tau_i d^i\|} - \frac{\tau_i d^i}{\|\tau_k d^k - \tau_i d^i\|} \rightarrow \lim_{k \rightarrow \infty} \frac{d^k}{\|d^k - (\tau_i/\tau_k) d^i\|} = d.$$

Hence, there is some k' such that for all $k \geq k'$, d in (6) and (7) can be replaced by v^k , i.e.,

$$\begin{aligned} F_k(-\mu_i) &= f(x^i - \mu_i v^k) < f(x^i) = F_k(0), \\ F_k(t_i) &= f(x^i + t_i v^k) < f(x^i) = F_k(0), \end{aligned} \quad (8)$$

where k' has to be large enough such that also $t_i < \|x^k - x^i\|$ holds for $k \geq k'$ (note that $\|x^k - x^i\| \rightarrow +\infty$). Applying (2) and $k > i$, one has also

$$F_k(\|x^k - x^i\|) = f(x^k) > f(x^i) = F_k(0). \quad (9)$$

Note that (8) already yields a contradiction if $f(x)$ is linear, while (8) and (9) together are contradictory if $f(x)$ is quadratic. So, we are done in these cases. From now let $f(x)$ and hence $F_k(t)$ be cubic polynomials.

Let $k \geq k'$, and put $t_k = \|x^k - x^i\|$. By construction, $-\mu_i < 0 < t_i < t_k$. So (8) and (9) entail that $F_k(t)$ increases somewhere in the interval $[-\mu_i, 0]$, then decreases somewhere in $[0, t_i]$, and again increases somewhere in $[t_i, t_k]$. Since $F_k(t)$ is cubic, $F_k(t)$ thus also increases on the interval $[t_k, +\infty)$, i.e., $F_k(t)$ is unbounded from above on the ray $S_k = \{x^i + t(x^k - x^i) \mid t \geq 0\}$. Note that the segment between x^i and x^k belongs to M , and one has immediately a contradiction to $s = \sup_{x \in M} f(x) < +\infty$ if M is an affine set. If M is not affine, then again by $s < +\infty$, S_k has to leave the polyhedral set M at some point $y^k \in M \setminus \text{ri } M$ with $f(y^k) > f(x^k)$. Now consider the sequence $\{y^k\}$, $k \rightarrow \infty$. Then, arguing as in the proof of (4), an infinite subsequence of $\{y^k\}$ belongs to the same face of M with dimension $< m$, leading to a contradiction. This completes the proof. \square

3 Related results

It is well-known that Theorem 1 cannot be extended to polynomials of degree 4 or higher, see e.g. the counterexample $\max -x_1^2 - (1 - x_1 x_2)^2$ s.t. $(x_1, x_2) \in \mathbb{R}^2$ by Frank and Wolfe [2]. Further, Theorem 1 is no longer true if M is a more general set, namely the Minkowski sum of a convex compact set and a polyhedral cone, the counterexample $\max 2x_1 x_3 - x_1^2 x_2$ s.t. $(x_2 - 1)^2 + x_3^2 \leq 1$ is due to Andronov, Belousov and Shironin [1]. Obviously, the constraint set has the mentioned form; for a detailed discussion of this example we refer to [1] or [10, §4].

In contrast to the cubic case, the following Frank-Wolfe type theorem due to Kummer [8] holds true; it was later generalized to mixed-integer problems in [11, §6.2].

Theorem 2 *Consider the optimization problem*

$$\max f(x) = x^T Q x + p^T x \quad \text{s.t. } x \in M = C + D, \quad (10)$$

where Q is a symmetric (n, n) -matrix, $p \in \mathbb{R}^n$, and M is the Minkowski sum of a compact set $C \subset \mathbb{R}^n$ and a polyhedral cone D in \mathbb{R}^n . Suppose M is nonempty and $f(x)$ is bounded from above on M . Then the supremum of $f(x)$ is attained at some point of M .

Proof. (Kummer [8, §6]) It is well-known that the supremum function

$$\lambda \mapsto \varphi(\lambda) = \sup_{d \in D} (d^T Q d + \lambda^T d) \quad (11)$$

is continuous on $\text{dom } \varphi = \{\lambda \in \mathbb{R}^n \mid \varphi(\lambda) < +\infty\}$, see e.g. [8, Satz 3.2] or [12, Thm. 4.5.1]. For each $y \in C$, let $\tilde{\varphi}(y) = \sup_{d \in D} f(y + d)$, thus $\tilde{\varphi}(y) \leq \sup_{x \in C+D} f(x) < +\infty$ by assumption. Then we obtain

$$\tilde{\varphi}(y) = f(y) + \sup_{d \in D} \left(d^T Q d + (p + 2Qy)^T d \right) = f(y) + \varphi(p + 2Qy),$$

where the supremum $\varphi(p + 2Qy)$ is attained at some point in D , by Theorem 1. Continuity of $y \mapsto \varphi(p + 2Qy)$ implies that $\tilde{\varphi}$ is continuous on the compact set C . So, $\max_{y \in C, d \in D} f(y + d) = \max_{y \in C} \tilde{\varphi}(y)$ exists by the Weierstrass theorem. This completes the proof. \square

It is clear that Theorem 1 can be immediately extended to the case that M is the union of finitely many polyhedral sets. Similarly, Theorem 2 obviously also holds if M is the union of finitely many sets of the form $C_i + D_i$, with compact sets C_i and polyhedral cones D_i . This allows, for example, the direct application to problems with M being the solution set of a linear complementarity problem or being the set of optimal solutions (or stationary solutions satisfying the Karush-Kuhn-Tucker conditions) of a quadratic program of the form (10).

Let us note that the function φ defined in (11) is even Lipschitzian on bounded subsets of $\text{dom } \varphi$ ([13], see also [6, §16]), where $\text{dom } \varphi$ is a polyhedral cone ([12, Lemma 4.5.2]). Further note that in the setting of Theorem 2, continuity of optimal values holds in a general sense ([8, Satz 2.1]): $C_\nu \rightarrow C$ (in the Painlevé-Kuratowski convergence) and $p_\nu \rightarrow p_0$ imply $\max_{y \in C_\nu, d \in D} (f(y + d) + p_\nu^T(y + d)) \rightarrow \max_{y \in C, d \in D} (f(y + d) + p_0^T(y + d))$ provided that the maxima exist for $\{(C_\nu, p_\nu)\}, (C, p_0)$. For representations of a closed convex set M in the form $M = C + O^+M$, with a convex, compact set C , we refer e.g. to [11, 14, 15]. Further, it is known that a set $M = \{x \in \mathbb{R}^n \mid g_i(x) \leq 0, i = 1, \dots, m\}$ defined by convex polynomials g_1, \dots, g_m , has a polyhedral recession cone O^+M (cf. e.g. [15, Chapt. 2, §3]). The same holds if M is given by quasiconvex polynomials f_i , see e.g. [11, §4.2].

Beyond the Theorems 1 and 2, there is a broad literature on other extensions of the Frank-Wolfe [2] result. For example, generalizations to mixed-integer problems are handled in [11, 12, 16, 17]. Let us discuss in more detail a few related results for optimization problems with real variables.

A generalization to a class of non-quadratic objective functions on a polyhedral constraint set is derived in [18], however this class of functions does not contain cubic polynomials. As shown by Belousov [15] (see also [10]), a Frank-Wolfe type theorem (for maximization) holds if f is a concave polynomial and M is defined by $M = \{x \in \mathbb{R}^n \mid g_i(x) \leq 0, i = 1, \dots, m\}$ via convex polynomials g_1, \dots, g_m . In [19] a version of the latter result is presented, namely, for the case of quadratic functions f, g_i under additional requirements: for example, if g_1 is convex and g_2, \dots, g_m are affine-linear, or if f is quasiconcave and all g_i are convex. This covers [20], where convex quadratic-quadratic programs are handled. Recently, Frank-Wolfe type theorems for more general polynomial programs have been published (cf. [21, 22]), where so-called non-degenerate polynomials are considered and the results are presented in terms of the Newton polyhedra of the polynomials defining the problem.

How to deal with the unbounded is a basic subject of convex analysis, when asking for the existence of optimal solutions, see e.g. [9, 14, 15]. Various results of this type allow consequences in form of a Frank-Wolfe type theorem. In §4.2 of Chapt. II of his book [15], Belousov studies systematically classes of convex programs under the viewpoint whether a Frank-Wolfe type theorem holds, or not. A positive answer is not only given for convex polynomial programs (see above), but also, for example, if f is a concave polynomial and M is the Minkowski sum of a convex compact set and a polyhedral cone. For a comprehensive study of asymptotic cones and functions and their role in existence results, both in convex and non-convex settings, we refer

to the book by Auslender and Teboulle [14], but also to [23, 24, 25] and the references therein. However, it is open, for example, whether it is possible to derive Theorem 1 from these results, or not, cf. [25, p. 541].

We finish this section with a reference to [7]. In that paper, the problem of characterizations of Frank-and-Wolfe sets (cf. Sect. 1 for the definition) is handled. In the case of quadratic functions being quasiconcave on M , this problem has been completely solved ([7, Thm. 1]). In addition, our Theorem 2 was rediscovered. In a recent private communication, D. Noll informed the author that the converse of Theorem 2 is also true, i.e., if a set M of the form $M = C + D$ with compact set C and closed convex cone D is a Frank-and-Wolfe set, then the cone D is necessarily polyhedral. In our opinion, to continue this research and to ask similar questions if f is a cubic polynomial would be of great value.

Acknowledgement. Thanks to Immanuel Bomze and Bernd Kummer for various suggestions improving the clarity of presentation.

References

- [1] V.G. Andronov, E.G. Belousov and V.M. Shironin, On Solvability of the Problem of Polynomial Programming (In Russian). *Izvestija Akadem. Nauk SSSR, Tekhnicheskaja Kibernetika* No.4, 194–197, 1982. Translation appeared in *News of the Academy of Science of USSR, Dept. of Technical Sciences, Technical Cybernetics*.
- [2] M. Frank and P. Wolfe, An algorithm for quadratic programming. *Naval Research Logistics Quarterly*, 3: 95–110, 1956.
- [3] E. Blum and W. Oettli, Direct proof of the existence theorem in quadratic programming. *Operations Research*, 20: 165–167, 1972.
- [4] B.C. Eaves, On quadratic programming. *Management Science*, 17: 698–711, 1971.
- [5] E.G. Belousov and V.G. Andronov, *Solvability and Stability of Problems of Polynomial Programming*, Moscow University Publ., Moscow, 1993 (In Russian).
- [6] G.M. Lee, N.N. Tam and N.D. Yen, *Quadratic Programming and Affine Variational Inequalities*, Springer, New York, 2005.
- [7] J.E. Martínez-Legaz, D. Noll and W. Sosa, Minimization of quadratic functions on convex sets without asymptotes. *Journal of Convex Analysis*, 25: 623–641, 2018.
- [8] B. Kummer, Globale Stabilität quadratischer Optimierungsprobleme. *Wissenschaftliche Zeitschrift der Humboldt-Universität zu Berlin, Math.-Nat. R.*, XXVI, 5: 565–569, 1977.
- [9] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, N.J., 1970.
- [10] E.G. Belousov and D. Klatte, A Frank–Wolfe type theorem for convex polynomial programs. *Computational Optimization and Applications*, 22: 3748, 2002.
- [11] B. Bank and R. Mandel, *Parametric Integer Optimization*, Mathematical Research Vol. 39, Akademie-Verlag Berlin, 1988.
- [12] B. Bank, J. Guddat, D. Klatte, B. Kummer and K. Tammer, *Non-Linear Parametric Optimization*, Akademie-Verlag, Berlin, 1982, and Birkhäuser, Basel–Boston–Stuttgart, 1983.
- [13] D. Klatte, On the Lipschitz continuity of optimal solutions in parametric problems of quadratic optimization and linear complementarity. *Optimization*, 16: 819–831, 1985.
- [14] A. Auslender and M. Teboulle, *Asymptotic Cones and Functions in Optimization and Variational Inequalities*, Springer, New York, 2003.

- [15] E.G. Belousov, *Introduction to Convex Analysis and Integer Programming* (in Russian), Moscow University Publisher, 1977.
- [16] B. Bank and R. Hansel, Stability of mixed-integer quadratic programming problems. *Mathematical Programming Study*, 21: 1–17, 1984.
- [17] R. Hansel, Ein Existenzsatz für gemischt-ganzzahlige quadratische Optimierungsprobleme. *Wissenschaftliche Zeitschrift der Humboldt-Universität zu Berlin, Math.-Nat. R.*, XXX, 5: 377–379, 1981.
- [18] A.F. Perold, A generalization of the Frank–Wolfe Theorem. *Mathematical Programming*, 18: 215–227, 1980.
- [19] Z.-Q. Luo and S. Zhang, On extensions of the Frank-Wolfe Theorems. *Computational Optimization and Applications*, 13: 87–110, 1999.
- [20] T. Terlaky, On l_p programming. *European Journal of Operations Research*, 22: 70–100, 1985.
- [21] S.T. Dinh, H.V. Ha and T.S. Pham, A Frank-Wolfe type theorem for nondegenerate polynomial programs. *Mathematical Programming, Series A*, 147: 519–538, 2014.
- [22] T.S. Pham, Optimality conditions for minimizers at infinity in polynomial programming. arXiv:1706.00234v2 [math.OC] 30 Jun 2017.
- [23] A. Auslender, How to deal with the unbounded in optimization: Theory and algorithms. *Mathematical Programming, Series B*, 79: 3–18, 1997.
- [24] A. Auslender, Existence of optimal solutions and duality results under weak conditions. *Mathematical Programming, Series A*, 88: 45–59, 2000.
- [25] A.E. Ozdaglar and P. Tseng, Existence of global minima for constrained optimization. *Journal of Optimization Theory and Applications*, 128: 523–546, 2006.