REVUE FRANÇAISE D'AUTOMATIQUE, D'INFORMATIQUE ET DE RECHERCHE OPÉRATIONNELLE. RECHERCHE OPÉRATIONNELLE

P. D'ALESSANDRO
M. DALLA MORA

# Fast projection method for a special class of polytopes with applications

*Revue française d'automatique, d'informatique et de recherche opérationnelle. Recherche opérationnelle*, tome 22, nº 4 (1988), p. 347-361.

<http://www.numdam.org/item?id=RO_1988__22_4_347_0>

Recherche opérationnelle/Operations Research
(vol. 22, n° 4, 1988, p. 347 à 361)

# FAST PROJECTION METHOD FOR A SPECIAL CLASS OF POLYTOPES WITH APPLICATIONS (*)

by P. D'ALESSANDRO (1) and M. DALLA MORA (1)

Abstract. — *In this paper some known results of convex analysis are exploited to derive a fast method to project a point of $R^n$ onto a polytope of a special class. As an application a new technique of smoothing is illustrated. Numerical results are included.*

Keywords : Optimization; convex analysis; polytopes.

Résumé. — *Dans ce document, certains résultats biens connus d'analyse convexe sont exploités pour dériver une méthode rapide qui permet de projeter un point de $R^n$ sur un polytope d'une certaine catégorie. Comme application, une nouvelle technique de "smoothing" est illustrée. Quelques résultats numériques sont également inclus.*

Mots clés : Optimisation; analyse convexe; polytopes.

## 1. INTRODUCTION

There are already many methods to solve the problem of projecting a point onto a polytope (*see* [1], [2]). In this paper a special class of polytopes is considered, and it is remarked that they allow the application of a fast projection method. This method has some practical applications. In [3] it is exploited to introduce a new technique of smoothing, called optimal smoothing, which is briefly recalled here. Beside the interest on its own right, optimal smoothing is the main ingredient of a new technique of seasonal adjustment, which is introduced in [4].

The numerical performance experienced so far in the use of the fast projection method is excellent.

(*) Received in November 1987.
(1) Department of Electrical Engineering, University of L'Aquila, Monteluco, 67100 L'Aquila, Italy.

Recherche opérationnelle/Operations Research, 0399-0559/88/04 347 15/$ 3.50

## 2. PROJECTIONS AND POLYTOPES

Let $H$ be a real Hilbert space with inner product $(.,\ .)$ and norm $\|\ .\ \|$. For any set $A$ in $H$ denote by $\mathscr{C}_0(A)$ the convex conical extension of $A$, $\mathscr{C}_0^-(A)$ the closed convex conical extension of $A$; $\mathscr{C}(A)$ the convex extension of $A$ and $\mathscr{C}^-(A)$ the closed convex extension of $A$ and finally by $Af(A)$ the affine extension of $A$.

For a closed convex set $A$ and $x$ in the boundary of $A$ define:

$$S(x,\ A)=\mathscr{C}_0^-(A-x)+x$$

(which is called the support cone to $A$ at $x$) and

$$N(x,\ A)=(S(x,\ A)-x)^0$$

(which is called the normal cone to $A$ at $x$) where if $B\subset H$, $B^0$ denotes the polar of $B$ defined by:

$$B^0=\{h:\ h\in H,\ (h,\ x)\leqq 1 \text{ for all } x\in B\}$$

(*see* e. g. [5]) $A$ vector $h$ is in $N(x,\ A)$ if and only if

$$A\subset\{z:\ z\in H \text{ and } (h,\ z)\leqq(h,\ x)\}.$$

They are called normals to $A$ at $x$.

If $A$ is closed and convex, $P_A$ denotes the projection of $H$ onto $A$. Some major properties of this function are collected in the following theorem (*see* e. g. [6]).

THEOREM 1: *Let $A$ be a closed convex subset of $H$. Then*

(i) *If $x$ is any point in $H$, there exists a unique element $P_A(x)$ in $F$ such that* $\|x-P_A(x)\|=\inf\{\|x-z\|:\ z\in A\}$.
(*The map $P_A$ is called the projection of $H$ on $A$*).

(ii) *$z\in A$ is the projection of $x$ onto $A$ if and only if*

$$(x-z,\ z-y)\leqq 0$$

*for each $y$ in $A$.*

(iii) *For all $x$ and $y$ in $H$*

$$\|P_Ax-P_Ay\|\leqq\|x-y\|.$$

*It is clear for Theorem* 1 (*ii*) *that*

$$P_A(x+h)=x \text{ for all } x\in A \text{ and } h\in N(x, A).$$

In what follows it is assumed that $H=R^n$.

A polytope is the convex extension of a finite set [7], consequently [8] it is compact and hence closed. Any convex set has nonvoid relative (in the sense of [7]) interior and a polytope is a convex body if and only if the dimension of the affine extension of the finite set is $n$.

The description of compact convex sets by means of convex extensions can be reduced to minimal form. In fact it is immediate consequence of definitions, that if any such set $C$ is given by $\mathscr{C}(A)$, for some set $A$, then $A$ contains the set of extreme points of $C$ (for the definition of extreme point *see* e. g. [6]) and, on the other hand, the celebrate Krein-Milman theorem [6] insures that any compact convex subset of a locally convex linear topological Haussdorff space is the convex extension of the set of its extreme points. It follows in particular that it admits a unique minimal generating set which is precisely the set of its extreme points (and is obviously finite, in view of the above argument, if the set is a polytope).

Recall that a set is symmetric if the opposite of any member of the set is a member of the set. It is clear that a convex compact set is symmetric if and only if such is the set of its extreme points.

Notice also that if $A$ is any set and $t$ any vector then $\mathscr{C}(t+A)=t+\mathscr{C}(A)$; and if $A$ is convex and $S$ is the set of its extreme points, $t+A$ is convex and its set of extreme points is $t+S$.

Sometimes polytopes are spheres of some norm (pseudonorms may be handled in similar fashion, *see* [9]). More precisely, denoting by $b(.)$ the barycentre of a finite set, the following theorem can be stated.

THEOREM 2: *Let $P$ be a polytope and $S$ the set of its extreme points, then $P$ is the closed unit sphere of a norm if and only if $S$ contains a set of $n+1$ affinely independent points and $S$-$b$ $(S)$ is symmetric.*

*Proof:* Sufficiency: because $P-b(S)=\mathscr{C}(S)-b(S)=\mathscr{C}(S-b(S))$ and translation is an homeomorfism, $P-b(S)$ has the following properties: it is a compact convex set and it is symmetric. Since in a real space a set is convex and circled if and only if it is convex and symmetric, $P-b(S)$ is convex and circled. It is now shown that $P$ [and hence $P-b(S)$] are convex bodies: for example $b(S)$ is an interior point of $P$, so that 0 is an interior point of $P-b(S)$. To this purpose it is convenient to invoke a powerful topological principle because it makes the point immediate (without obsuring the ideas

involved in this matter). Let $S$ be $\{x_1, \ldots, x_m\}$ and, without any restriction of generality, assume that $x_m$ belongs to the affinely independent subset of $S$.

Then $P$ is the image of the closed unit sphere in $R^{m-1}$ relative to the norm $\|\cdot\|_1$ intersected with the nonnegative orthant under the composed map $T_{x_m} \circ T$ where $T$ is a linear map defined by

$$T(\lambda) = \lambda_1 (x_1 - x_m) + \ldots + \lambda_{m-1} (x_{m-1} - x_m)$$

for each $\lambda$ in $R^{m-1}$ and $T_{x_m}$ in the translation by $x_m$. In view of the Open Mapping Theorem ([6], pp. 99), $T$ is open (because its range is $R^n$, which, by the Baire theorem, is of the second category) so that being the translation an homeomorphism, $T_{X_m} \circ T$ is open, whence the desired conclusion is immediately shown.

It follows that $P - b(S)$ is also radial at zero. The proof of sufficiency can be therefore concluded showing that the Minkowski functional $p$ of $P - b(S)$ is a norm and the corresponding closed unit sphere about the origin coincides with $P - b(S)$. Denote by $B_p$ this sphere. Since $P - b(S)$ is a bounded neighborhood of 0, its positive multiples form local base for the topology of $R^n$. On the other hand for each real $\varepsilon > 0$

$$P - b(S) \subset B_p \subset (1+\varepsilon)(P - b(S)) \subset (P - b(S)) + \varepsilon(P - b(S))$$

hence

$$P - b(S) \subset B_p \subset \bigcap \{(P - b(S)) + \varepsilon(P - b(S)) : \varepsilon > 0\} = (P - b(S))^- = P - b(S)$$

(where an elementary computation for vector topologies has being applied). Because $(1+\varepsilon)(P - b(S))$ is compact $p$ cannot be a pseudonorm thus it is a norm and the latter relation concludes the proof of sufficiency.

Necessity: By hypothesis for some norm $p$ and some vector $t$, $B_p = P - t$ so that the set of extreme points of $B_p$ is $S - t$. But this set is symmetric and therefore

$$\sum_{i=1}^{m} x_i - mt = 0 \quad \text{or} \quad t = b(S).$$

Thus $S - b(S)$ is actually symmetric. If there were at most $N$ affinely independent points in $S$ then because $\mathscr{C}(S) \subset Af(S)$ it would follows

$$\mathscr{C}(S) - b(S) = B_p \subset Af(S) - b(S).$$

But then $0 \in Af(S) - b(S)$, or this latter set is the linear subspace parallel to $Af(S)$ and has at most dimension $N-1$. This contradicts the assumption that $p$ is a norm. $\lhd$

Note that for the case of spheres about the origin the condition would be that $S$ be symmetric and contains a subset of $n$ linearly independent points. Thus it is clear from the theorem that there is no hope of describing the topology with less then $2N$ points because it is impossible to specify a symmetric set containing a base with less that $2N$ points. For future reference we state the following.

THEOREM 3: *Let $B$ any base for $R^n$ and let $S$ be the set $B \cup (-B)$. Then the polytope $\mathscr{C}(S)$ is the closed unit sphere about the origin of a norm and $S$ is the set of its extreme points. Conversely if the closed unit sphere about the origin of a norm has $2N$ extreme points, then the set of its extreme points has the form $B \cup (-B)$ where $B$ is a base for $R^n$.*

The proof is rather straightforward in view of the preceding work and is therefore omitted. Of course the norm in question is the Minkowski functional of $\mathscr{C}(S)$. Any norm whose closed spheres are polytopes with $2N$ extreme points is called a $M$-norm.

Note that if $p$ is any norm for $R^n$ and $T$ any linear isomorphism of $R^n$ onto itself, then $p \circ T$ is also a norm, moreover if $B_p$ is the closed unit sphere (about the origin) of $p$ then $T^{-1}(B_p)$ is the closed unit sphere of $p \circ T$. If $S$ is the set of extreme points of a convex set $C$, $T^{-1}(S)$ is the set of extreme points of the convex set $T^{-1}(C)$, as is immediately verified. Two norms $p_1$ and $p_2$ are called linearly equivalent if for some linear isomorphism $T$, $p_1 = p_2 \circ T$ (of course this is an equivalence relation for the set of all norms). Note also that if $C$ is a polytope i.e. $C = \mathscr{C}(A)$ for some finite set $A$, and $T$ is a linear isomorphism then $T(C) = \mathscr{C}(T(A))$. At this point, in view of the invariance of the cardinal of the set of the extreme points of a convex set under a linear isomorphism, the existence of norms, that are not linearly equivalent, is obvious. In view of Theorem 3 the set of $M$-norms for $R^n$ is an equivalence class under linear equivalence.

Faces of convex sets are defined as in [7]. However to avoid repeated exclusion of trivial cases it is stipulated that the whole convex set in question is not a face. Trivially the singletons of extreme points are faces. Any convex subset of $R^n$ contains a polytope with the same affine extension. Recall that any convex subset of $R^n$ has relative interior (Theorem 6.2 in [7]), thus in view of Theorem 18.1 in [7] a face in the present sense is a face in the sense

of [7], if and only if it is contained in the relative boundary of the convex set in question (which is Corollary 18.1.2 in [7]).

Note that, since the union of a chain of faces is a face, each face is contained in a maximal face. In view of the Theorem 18.3 in [7], any face of a polytope is the convex extension of a subset of the set of the extreme points of the polytope. Thus any face of a polytope is in turn a polytope. A direct verification shows that any face of a face is also a face of the original polytope. The relative boundary of a polytope is entirely made up of faces:

THEOREM 4: *The relative boundary of any polytope is the union of the family of its faces.*

*Proof:* Because translation is a homeomorphism it is easy to see that is suffices to make the proof for the case where the affine extension of the polytope, say $P$, is a linear subspace, say $F$, of $R^n$. With reference to the topological subspace $F$, $P$ has nonvoid interior $P^i$, which in addition is convex and coincides with its radial kernel (Theorem 13.1 in [6]), so that, if $x$ belongs to the boundary of $P$ in $F$, by Theorem 3.8 in [6] the existence of a linear functional, necessarily of the form $(h, .)|_F$ for some $h$ in $F$, separating $x$ and $P$ is ensured. Thus $(h, .)$ separates $x$ and $P$ in $R^n$. The rest of the proof amounts to the straightforward verification that the intersection of the hyperplane $\{z: z \in R^n, (h, z) = (h, x)\}$ with $P$ is a face of $P$. $\lhd$

A further important result is immediately scored arguing similarly to the last proof.

THEOREM 5: *For any face of a polytope there exists a hyperplane containing the face and such that the polytope is contained in a closed half space defined by the hyperplane.*

*Proof:* Again it is readily seen that it suffices to make the proof for the case where the affine extension of the polytope $P$ is a linear subspace $F$ of $R^n$. Let $D$ be a face of $P$. Then in the topological space $F$, $D$ is disjoint from $P^i$ and therefore there exist $h \in F$ such that $(h, .)$ separates $P$ and $D$.

Still reasoning in $F$ and again by Theorem 13.1 in [6] $P^{i^-} = P$. Therefore if $x$ belongs to $D$ and hence also to $P^-$ there exists a sequence $\{x_n\}$ in $P^i$ that converges to $x$ in $F$ and therefore also in $R^n$. But this implies that $(h, .)$ is constant in $D$, thereby concluding the proof.

It is convenient at this point to introduce some terminology. The extreme points forming the generating set of a face are called vertices of the face. Varying slightly the terminology of [7] if $x$ is in the relative boundary of a convex set $C$, and, for some $h$ in $R^n$, the convex set is contained in the half

space $\{z: z \in R^n, (h, z) \leqq (h, x)\}$ then $h$ is called a normal to $C$ at $x$, and also normal to any face $F$ of $C$ which is contained in the hyperplane $\{z: z \in R^n, (h, z) = (h, x)\}$. This hyperplane is called tangent hyperplane to $C$ at $x$ or tangent hyperplane to $F$. Moreover the following obvious remark is nevertheless particularly useful: if the elements of a set $A$ are normal to $C$ at $x$, such are the elements of $\mathscr{C}_0(A)$. Finally it is easy to verify that all points in the relative interior of a face have the same normal cone, that will be called the normal cone of the face. Furthermore if a face $D_1$ contains a face $D_2$, then the normal cone of $D_2$ contains that of $D_1$.

Before concluding it is stated a result that exploits the properties of $M$-norms to characterize boundary points of their unit (without restriction of generality) closed spheres, and that is particularly useful in applications.

THEOREM 6: *Let $P$ be the closed unit sphere of a $M$-norm about the origin and $S$ the set of its extreme points. Then a point $x$ belongs to the boundary of the sphere if and only if it can be expressed as a convex combination of a subset of $S$, where there are no pairs of opposite points.*

*Proof:* The already mentioned fact that $M$-norms form an equivalence class under linear equivalence is first exploited to reduce the proof to a most easy case.

As proved earlier $P = \mathscr{C}(B \cup (-B))$ for some base $B$. Now if with selfexplanatory notations $\tilde{B}$ is the base $\{(1, \ldots, 0) \ldots (0, \ldots, 1)\}$ there exists a linear isomorphism $T$ that takes $\tilde{B}$ onto $B$. Thus

$$T(\mathscr{C}(\tilde{B} \cup (-\tilde{B}))) = \mathscr{C}(T(\tilde{B} \cup (-\tilde{B}))) = \mathscr{C}(B \cup (-B)) = P,$$

where clearly $\mathscr{C}(\tilde{B} \cup (-\tilde{B}))$ is just the closed unit sphere of the norm $\| \cdot \|_1$. Since in $R^n$ a linear isomorphism is a topological isomorphism it suffices to make the proof for the polytope $\tilde{P} = \mathscr{C}(\tilde{B} \cup (-\tilde{B}))$. Now let $\tilde{b}_1 \ldots \tilde{b}_n$ denote the elements of $\tilde{B}$. If

$$x = \sum_{i=1}^{n} \alpha_i \tilde{b}_i \qquad \text{with} \quad \alpha_i \in (0, 1],$$

$$\forall i, \quad \sum_{i=1}^{n} \alpha_i = 1, \qquad n \in \{1, \ldots, N\},$$

then clearly $\|x\|_1 = 1$. Conversely if it is not possible to express $x$ in this way then $x$ will have the form

$$x = \sum_{i=1}^{n_1} \alpha_i \tilde{b}_i + \sum_{j=1}^{n_2} (\alpha_j^+ - \alpha_j^-) \tilde{b}_j$$

where:

$$\{\tilde{b}_i\} \cap \{\tilde{b}_j\} \text{ is void}, \qquad \alpha_j^+ \neq \alpha_j^-, \ \forall j,$$

$$\sum_{i=1}^{n_1} \alpha_i + \sum_{j=1}^{n_2} \alpha_j^+ + \sum_{j=1}^{n_2} \alpha_j^- = 1.$$

$n_1$ and $n_2$ are nonnegative integers with $n_2 > 0$ and $n_1 + 2n_2 \leqq 2N$. It follows that $\|x\|_1 < 1$, and hence the proof is concluded. $\lhd$

It is obvious that the convex extension of each subset of $S$, made up of $N$ points and satisfying the condition of the theorem is a maximal face of $P$ and hence the number of such maximal faces is $2N$. Notice also that the vertices of any face form a linearly independent set and those of a maximal face a base.

## 3. FAST PROJECTIONS

It is considered the following special class $\mathscr{P}$ of polytopes. A member of $\mathscr{P}$ is the closed unit sphere about the origin of a $M$-norm (in this respect notice that if the center of the sphere is not the origin and/or the radius is not 1 by the translation and/or scaling the projection problem can be reduced to the case considered here). Moreover it is assumed that each face of $P$ admits a nonzero normal belonging to the convex cone generated by its vertices. An example of $M$-norms for which this condition hold (that will also be called of class $\mathscr{P}$) are $\|\,.\,\|_1$ and $\|\,.\,\|_\infty$. A less trivial example is given in the next section.

Now consider any such polytope $P$ and let $p$ the corresponding norm. Consider $x \in H$ with $p(x) > 1$ (to avoid trivial cases). Because $x/p(x)$ is in the boundary of $P$, it belongs to a face $F_1$ and actually it may be assumed it is in the relative interior of $F_1$, for otherwise it would be in a face of $F_1$ which is a face of $P$.

Let $\{v_1, \ldots, v_K\}$ be the vertices of $F_1$ and $n_1$ be a normal to $F_1$ in $\mathscr{C}_0(\{v_1, \ldots, v_K\})$. Thus

$$x = \sum_{i=1}^{K} \lambda_i v_i: \qquad \lambda_i > 0, \quad \sum \lambda_i > 1$$

$$n_1 = \sum_{i=1}^{K} \gamma_i v_i: \qquad \gamma_i \geqq 0, \quad \sum \gamma_i > 0$$

thus the following two positive real numbers can be defined

$$\beta_1 = \min\{\lambda_i/\gamma_i : \gamma_i \neq 0\}$$

$$\delta_1 = \left(\left(\sum_{i=1}^{k} \lambda_i\right) - 1\right) \Big/ \sum_{i=1}^{k} \gamma_i.$$

The meaning of these two parameters is as follows. If $\alpha \leqq \beta_1$ then $x - \alpha n_1 \in \mathscr{C}_0$ $(\{v_i\})$ and if $\alpha > \beta_1$ then $X - \alpha n_1 \notin \mathscr{C}_0(\{v_i\})$. If $\alpha < \delta_1$ then $\| x - \alpha n_1 \|_1 > 1$; if $\alpha = \delta_1$ then $\| x - \alpha n_1 \| = 1$.

Let $\alpha_1$ be min $(\{\beta_1, \delta_1\})$ and $x_1$ be $x - \alpha_1 n_1$.

If $\alpha_1 = \delta_1$ then $p(x_1) = 1$ and the procedure is arrested.

Otherwise $p(x_1) > 1$ and $x_1/p(x_1)$ is in the relative boundary of $F_1$ and hence in the relative interior of a face $F_2 \subset F_1$. Now repeat the above step with $x_1$ in lieu of $x$.

By construction the procedure stops in at most $k \leqq n$ steps, at a point $x_j$, with $p(x_j) = 1$ and

$$x = x_j + \sum_{i=1}^{j-1} \alpha_i n_i.$$

Because by construction $\sum_{i=1}^{j} \alpha_i n_i \in N(x_j, P)$, $x_j$ is the projection of $x$ onto $P$.

In the example of the next section, the algorithm will be applied to a sphere for which the linear isomorphism transforming the sphere into that of the norm $\| \cdot \|_1$ is known. In this case the extreme points are known *a priori* and hence the speed of the algorithms depends essentially on the method used to compute normals, that belong to the cone defined by the vertices of the face. In the example we use an *ad hoc* method that takes advantage of the special structure of the problem.

Regarding the computation of normal cones to a face the following general observation may turn out to be suseful.

Consider a sphere $B \in \mathscr{P}$, a linear isomorphism $T$ and a face $F$ of $B$. Then if $N$ is the normal cone of $F$, $T^* N$ is the normal cone to the face $T^{-1} F$ of $T^{-1} B$.

## 4. OPTIMAL SMOOTHINGS

The optimal smoothing approach is recalled briefly. For a complete discussion and details the interested reader is referred to [3].

Consider a finite time series of signal $f \in R^n$. Let $v$ a functional on $R^n$, such that $v(f)$ represents the variation of the time series and, in addition, assume that $v$ is a norm.

To smooth the signal one should seek another signal $\hat{f}$ with lower variation $v(\hat{f}) < v(f)$ that approximates $f$. Suppose that for some $c < 1$, $v(\hat{f}) \leqq cv(f)$ is adequate, then the following convex programming problem define the optimal smoothing problem

$$\text{find } \hat{f} \text{ such that } \| \hat{f} - f \| = \text{minimum, under the constraints } \hat{f} \in S^v_{cv(f)}$$

where $S^v_{cv(f)}$ is the closed sphere centered in the origin defined by the norm $v$ and with radius $cv(f)$.

If the sphere is in the class $\mathscr{P}$ then the fast projection method can be used, since the solution $\hat{f}$ is the projection of $f$ onto $S^v_{cv(f)}$.

The role of $v$ will be played by the norm $w_1$ defined by

$$w_1(f) = | f(1) | + \sum_{i=1}^{n-1} | f(i+1) - f(i) |.$$

Denote by $e_i$ the vector in $R^n$ with the ith component equal to one and the other components equal to 0.

Let $E$ be the set

$$E = \{e_1, -e_1, \ldots, e_n, -e_n\}.$$

Consider the $n \times n$ matrix

$$T = \begin{pmatrix} 1 & 0 & \ldots & 0 & 0 \\ -1 & 1 & \ldots & 0 & 0 \\ \ldots & . & \ldots & \ldots & . \\ 0 & 0 & \ldots & -1 & 1 \end{pmatrix}.$$

Then clearly, if, to simplify notations, we denote $\| \cdot \|_1$ by $p_1$,

$$w_1 = p_1 \circ T$$

and hence

$$S^1_{w_1} = w_1^{-1}([0, 1]) = T^{-1} p_1^{-1}([0, 1]) = T^{-1}(S^1_{p_1}) = T^{-1}(\mathscr{C}(E)) = \mathscr{C}(T^{-1}(E))$$

where the matrix $T^{-1}$ is given by

$$T^{-1} = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 1 & 1 & \dots & 0 & 0 \\ . & . & \dots & . & . \\ 1 & 1 & \dots & 1 & 1 \end{pmatrix}.$$

The norm $w_1$ is a non-trivial example of norm with spheres in class $\mathscr{P}$. Actually it is possible to state the following

THEOREM 7: *The sphere $S^1_{w_1}$ belongs to class $\mathscr{P}$.*

*Proof:* Constructive proofs are usually deemed to be more valuable. In the present case the proof consists essentially of a MODULA – 2 procedure, called normal, that given the signal and hence a face that generates a cone containing the signal in its relative interior, produces a normal to the face, that belongs to the same cone. For the sake of brevity we only outline the ideas underlying the algorithm. This should make easy to understand how the code works and hence how a proof of the theorem is obtained.

The signal should be viewed as combination of the above vertices of $S^1_{w_1}$ (that are the column of the matrix $T^{-1}$ given right above, each column is to be taken with a plus and a minus sign), in the same order as they appear in the matrix. Each vertex corresponds to a variation in level of the signal and appears with a plus sign if the variation is positive, a minus sign if it is negative, or does not appear at all if the signal remains constant. The normality condition is equivalent to the fact that the normal is orthogonal to each difference of two subsequent vertices (having the matrix of these differences handy for various examples of signals is helpful to understand the arguments of this proof). This condition is satisfied starting with the last of these differences and going on in decreasing order determining, (also in decreasing order and starting from the zero normal) the components of the normal. This is a diagonalization technique because once each of these conditions is satisfied, it remains true (thanks to the zero's appearing in the differences of each pair of subsequent vertices) for whatever determination

of the other entries of the normal. At each step the component corresponding to indices between two variations of the signal are settled via normality to a new difference of vertices. If the inner product of this difference with the already determined normal is zero, those components are left to zero, otherwise they are determined by the two required conditions that.

(*a*) The inner product be zero, and hence the new component must compensate a possible value given by the already determined components.

(*b*) The normal be in the cone of vertices.

The procedure normal is now appended below with some explicative comments.

```
MODULE timse;
TYPE rve=ARRAY [0...nvmax] OF REAL;
VAR fve, dve, nrl : rve;
PROCEDURE normal;
VAR c, i, p1, p2, p3, d : INTEGER;
    som, en : REAL;
BEGIN
c:=0; WHILE fve [nv-c]=0. DO c:=c+1 END; pl:=nv-c;
c:=0; WHILE fve [p1-c-1]=0. DO c:=c+1 END; p3:=p1-c-1; p2:=1; d:=p2-p3;
FOR i:=p1 TO nv DO nrl [i]:=fve [p1] END;
IF p3=0 THEN
IF p1 > 1 THEN FOR i:=1 TO p1-1 DO nrl [i]:=0. END END
ELSE
WHILE p3 > 0 DO
FOR i:=1 TO p3-1 DO dve [i]:=0. END;
FOR i:=p3 TO p1-1 DO dve [i]:=fve [p3] END;
FOR i:=p1 TO nv DO dve [i]:=fve [p3]-fve [p1] END;
som:=0.; FOR i:=p2 TO nv DO som:=som+nrl [i] * dve [i] END;
en:=-som * fve [p3]/FLOAT (d); FOR i:=p3 TO p2-1 DO nrl [i]:=en END;
p2:=p3; c:=0; WHILE fve [p2-c-1]=0. DO c:=c+1 END; p3:=p2-c-1; d:=p2-p3
END;
IF p2 > 1 THEN FOR i:=1 TO p2-1 DO nrl [i]:=0. END END
END END normal;
```

## Remarks on the procedure

*nv* is the dimension of the space $R^n$. The vector *nrl* is the sought normal. Note that the vectors have a component of index zero which is added for coding reasons and has no role in the problem. Such component is always equal to 1.

The only variable which is essentially external to the procedure normal is the variable *fve*. Let the vector *fsi* represent the current point. Let the vector $v$ be defined as follows:

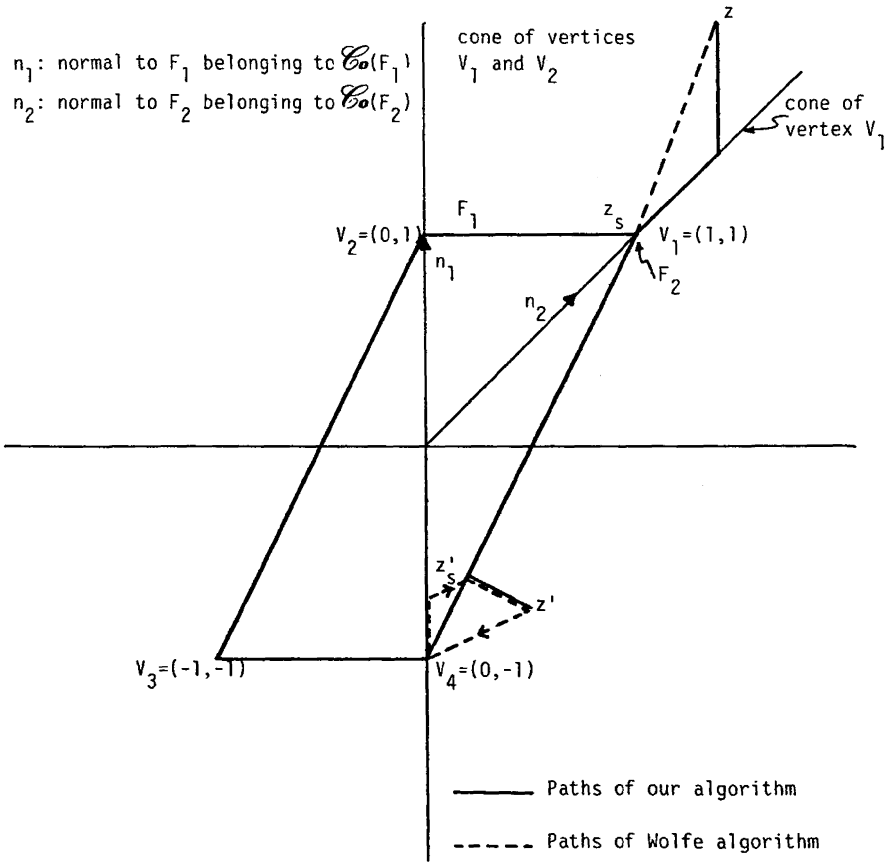
$$v(1)=fsi(1), \qquad v(i)=fsi(i)-fsi(i-1), \qquad i=2, \ldots, nv$$

then

$$fve(i)=\begin{cases} -1 & \text{if } v(i)<0 \\ 0 & \text{if } v(i)=0 \\ 1 & \text{if } v(i)>0 \end{cases}$$

*fve* represents the sign of the variations in the signal. The vector *dve* is used to represent differences of subsequent vertices (to which the normal must be orthogonal).

It may be helpful to give the picture of the unit sphere of $w_1$ in the two dimensional case. Such a spher is represented in figure, where it is also




**Figure**

depicted the path along which our algorithm obtains from two exemplar points $z$ and $z'$ their projections $z_s$ and $z'_s$ respectively. We also give (represented by a dashed line) the path that Wolfe's algorithm determines. The path of Wolfe's algorithm is understood to be the sequence of $x$ points (*see* [2]) to which the stop condition is referred. Note that the path defined by Wolfe algorithm are not unique in general, and for the case of $z'$ all possible paths are shown. These examples show that Wolfe algorithm and ours are not equivalent.

However the major difference between the algorithms is not much in the path followed to reach the solution but rather in the numerical computations. The reader will notice how trivial are the computations required to obtain the solution, in our specific example, for our algorithm compared to that of Wolfe. This speed si obtained in trade of the lack of generality of our algorithm. However the class $\mathscr{P}$ of polytopes is not believed to be the largest class of polytopes for which algorithms based on a similar technique might be developed.

Of course $w_1$ is not an ideal norm to represent variation in view of the presence of the first term in its definition, which is added to the variation just for the convenience of obtaining a norm of class $\mathscr{P}$. However there is some practical evidence that such term is not much disturbing and the ensuing filter work satisfactorily as seen from the numerical example below and in particular from the graphs of the given signal vs the filtered signal (this example is reproduced from [3]). In the table Filsig stands for filtered signal. Here the percents refer to the cut in the variation term of the functional $w_1$.

Finally in [4] it is illustrated how the optimal smoothing approach can be exploited to derive new techniques of seasonal adjustment, with a numerical example.

## REFERENCES

1. R. Mifflin, *A Feasible Descent Algorithm for Linearly Constrained Least Squares Problems*, Nonsmooth Optimization, C. Lemarechal and R. Mifflin Eds, Pergamon Press, 1978.
2. P. Wolfe, *Finding the Nearest Point in a Polytope*, Mathematical Programming, Vol. 11, 1976, pp. 128-149.
3. P. d'Alessandro, *Optimal Smoothings*, Proceedings of the 7th Symposium uber Operations research, St. Gallen, Aug. 1982, Athenaum Printers Konigstein, Germany.
4. P. d'Alessandro and E. de Luca, *A new Technique of Seasonal Adjustement with an Application to an Epidemiology Data*, Report No. 12-82, Institute of Electrical Engineering, University of L'Aquila, Oct. 1982.

5. RICHARD B. HOLMES, *A Course on Optimization and Best Aproximation*, Springer Verlag, 1972.
6. J. L. KELLEY and I. NAMIOKA *et al.*, *Linear Topological Spaces*, Springer Verlag, 1963.
7. R. T. ROCKAFELLAR, *Convex Analysis*, Princeton, 1982.
8. H. H. SCHAEFER, *Topological Vector Spaces*, Springer Verlag, 1971.
9. P. D'ALESSANDRO, *Finite Convergence Method to Solve Certain Projection Problems*, Report No. 2-81, Feb. 81, Istituto di Elettrotecnica, Università dell'Aquila.