JOURNAL OF APPROXIMATION THEORY 76, 326–350 (1994)

Linear Approximation in l_n^∞

M. FINZEL

Mathematical Institute,
University of Erlangen-Nuremberg, 91054 Erlangen, Germany

Communicated by A. L. Brown

Received June 26, 1991; accepted in revised form December 4, 1992

The paper studies the approximation behavior of a linear subspace U in l_n^∞; i.e., in $\mathbb{R}^n$ equipped with the maximum norm. As a principal tool the Plücker–Graßmann coordinates of U are used; they allow a classification of the index set $\{1, ..., n\}$ through which we determine the extremal points of the intersection of the orthogonal complement $U^\perp$ of U and the closed l_n^1-unit ball in $\mathbb{R}^n$, leading to the dual problem. As a consequence, we describe the metric complement $U^{(0)}$ of U and give a decomposition of $\mathbb{R}^n \setminus U$ into a finite set of pairwise disjoint convex cones on which the metric projection P_U has some characteristic properties. In the Chebyshev case, e.g., the metric projection is linear on these cones and, consequently, globally Lipschitz continuous. A refinement allows an analogous statement for the strict approximation, proving a conjecture of Wu Li. Besides the strict approximation, we are studying continuous selections of P_U with and without the Nulleigenschaft, and characterize those subspaces U which admit a linear selection. © 1994 Academic Press, Inc.

1. A CLASSIFICATION BY PLÜCKER–GRAßMANN COORDINATES

Let us denote the Euclidean space $\mathbb{R}^n$, $n \in \mathbb{N}$, endowed with the maximum norm $|\cdot|_\infty$ by l_n^∞; its elements are considered to be column vectors; in particular, e_ν denotes the νth standard basic vector. Let U be an r-dimensional subspace of l_n^∞, $0 < r = \dim U \leqslant n-1$, to exclude the trivial cases[1], and let $u^{(1)}, ..., u^{(r)}$ be a basis of U. For the matrix

$$\begin{bmatrix} u_{\nu_1}^{(1)} & u_{\nu_1}^{(2)} & \cdots & u_{\nu_1}^{(r)} \\ \vdots & \vdots & & \vdots \\ u_{\nu_m}^{(1)} & u_{\nu_m}^{(2)} & \cdots & u_{\nu_m}^{(r)} \end{bmatrix}, \qquad \nu_1, ..., \nu_m \in \{1, ..., n\},$$

let us introduce the abbreviation

$$\begin{bmatrix} u^{(1)} & \cdots & u^{(r)} \\ \nu_1 & \cdots & \nu_m \end{bmatrix}.$$

[1] At one point we need to deal with these cases, but this will not cause any difficulties.

0021-9045/94 $6.00

The values

$$p(\nu_1, \nu_2, ..., \nu_r) = \det \begin{bmatrix} u^{(1)} & \cdots & u^{(r)} \\ \nu_1 & \cdots & \nu_r \end{bmatrix}, \qquad \nu_1, ..., \nu_r \in \{1, ..., n\},$$

are known as the *Plücker–Graßmann coordinates* of U. Since

$$\operatorname{rank} \begin{bmatrix} u^{(1)} & \cdots & u^{(r)} \\ 1 & \cdots & n \end{bmatrix} = r,$$

not all the coordinates vanish. Regarded as homogeneous coordinates they are determined by the subspace U; indeed, they are independent of the particular choice of the basis. Conversely, they determine the subspace uniquely. Furthermore, they satisfy the relations

(P) if σ is a permutation of the index set $\{\nu_1, ..., \nu_r\}$ in $\{1, ..., n\}$, then $p(\sigma(\nu_1), ..., \sigma(\nu_r)) = \operatorname{sgn} \sigma \cdot p(\nu_1, ..., \nu_r)$,

(R) for all choices of $2r$ indices $\mu_1, ..., \mu_{r-1}, \nu_1, ..., \nu_{r+1}$ in $\{1, ..., n\}$,

$$\sum_{\rho=1}^{r+1} (-1)^{\rho+1} p(\mu_1, ..., \mu_{r-1}, \nu_\rho)\, p(\nu_1, ..., \nu_{\rho-1}, \nu_{\rho+1}, ..., \nu_{r+1}) = 0.$$

These are known as the Plücker relations, see, e.g., B. L. van der Waerden [23].

It is further known that the vectors $\tilde{u}^{(\rho)}$, $\rho = 1, ..., r$, defined by $\tilde{u}_j^{(\rho)} = p(\nu_1, ..., \nu_{\rho-1}, j, \nu_{\rho+1}, ..., \nu_r)$, $j = 1, ..., n$, form a basis of U, where $\{\nu_1, ..., \nu_r\}$ is chosen in such a way that $p(\nu_1, ..., \nu_r) \neq 0$.

Let $p^\perp(\nu_{r+1}, ..., \nu_n)$, $\nu_{r+1}, ..., \nu_n \in \{1, ..., n\}$, be the Plücker–Graßmann coordinates of the orthogonal complement $U^\perp$ of U. If $\langle \nu_1, ..., \nu_n \rangle$ denotes the permutation $j \mapsto \nu_j$ of the set $\{1, ..., n\}$, then the Plücker–Graßmann coordinates of U and $U^\perp$, respectively, are related by the equation $p^\perp(\nu_{r+1}, ..., \nu_n) = \operatorname{sgn}\langle \nu_1, ..., \nu_n \rangle \cdot p(\nu_1, ..., \nu_r)$.

Henceforth U will always be an r-dimensional subspace of $\mathbb{R}^n$, $1 \leq r \leq n-1$, and $\{u^{(1)}, ..., u^{(r)}\}$ will form a basis of U. Let us introduce the *classes of index sets* I_m, $0 \leq m \leq r+1$,

$$I_1 := \{(\nu) : \nu \in \{1, ..., n\},\ p(\nu, N) = 0 \text{ for all } N \subset \{1, ..., n\},\ \# N = r-1\};$$

$$I_2 := \{(\nu_1, \nu_2) : \nu_1, \nu_2 \in \{1, ..., n\},\ \nu_1 \neq \nu_2,\ (\nu_1), (\nu_2) \notin I_1,\ p(\nu_1, \nu_2, N) = 0 \\ \text{for all } N \subset \{1, ..., n\},\ \# N = r-2\}.$$

In general, for $2 \leq m \leq r$,

$$I_m := \{(\nu_1, ..., \nu_m) : \nu_1, ..., \nu_m \in \{1, ..., n\}, \text{ pairwise distinct}, \\ p(\nu_1, ..., \nu_m, N) = 0 \text{ for all } N \subset \{1, ..., n\}, \\ \# N = r-m, \text{ and for all } (\nu_1', ..., \nu_\mu') \in I_\mu, \\ 1 \leq \mu < m,\ \{\nu_1', ..., \nu_\mu'\} \not\subset \{\nu_1, ..., \nu_m\}\};$$

while

$$I_{r+1} := \{(\nu_1, ..., \nu_{r+1}) : \nu_1, ..., \nu_{r+1} \in \{1, ..., n\},$$

pairwise distinct and for all

$$(\nu'_1, ..., \nu'_\mu) \in I_\mu,\ 1 \leqslant \mu \leqslant r,\ \{\nu'_1, ..., \nu'_\mu\} \not\subset \{\nu_1, ..., \nu_{r+1}\}\}.$$

Finally,

$$I_0 := \{\nu : \nu \in \{1, ..., n\},$$

for all $(\nu_1, ..., \nu_m) \in I_m$, $1 \leqslant m \leqslant r+1$, $\nu \notin \{\nu_1, ..., \nu_m\}\}$.

For the notation of an index set we use round brackets instead of braces although the order of the elements is of no importance.

In his paper [2] on subspaces of l_n^p, F. Bohnenblust introduced the first two classes for a subspace U of l_n^p, $1 < p < \infty$, $p \neq 2$, in order to determine those subspaces which are the range of a contractive projection.

THEOREM 1. *Let U be an r-dimensional subspace of $\mathbb{R}^n$ with the basis $u^{(1)}, ..., u^{(r)}$, and let $1 \leqslant m \leqslant r+1 \leqslant n$. The following conditions on an index set $(\nu_1, ..., \nu_m)$ are equivalent:*

(i) *$(\nu_1, ..., \nu_m)$ belongs to I_m;*

(ii) *The matrix*

$$\begin{bmatrix} u^{(1)} & \cdots & u^{(r)} \\ \nu_1 & \cdots & \nu_m \end{bmatrix}$$

is of rank $m-1$ and any $m-1$ rows are linearly independent;

(iii) *$p_{U'}(\nu_1, ..., \nu_m) = 0$ for any m-dimensional subspace U' of U and for all pairwise disjoint indices $\nu'_1, ..., \nu'_{m-1} \in \{\nu_1, ..., \nu_m\}$ there exists an $(m-1)$-dimensional subspace U'' of U satisfying $p_{U''}(\nu'_1, ..., \nu'_{m-1}) \neq 0$.*

(iv) *$\dim U^\perp \cap \operatorname{span}\{e_{\nu_1}, ..., e_{\nu_m}\} = 1$ and if $v \in U^\perp \cap \operatorname{span}\{e_{\nu_1}, ..., e_{\nu_m}\} \setminus \{0\}$ then $v_\nu \neq 0$ for all $\nu \in \{\nu_1, ..., \nu_m\}$.*

Each index set determines a minimal collection of linearly dependent row vectors of the matrix

$$\begin{bmatrix} u^{(1)} & \cdots & u^{(r)} \\ 1 & \cdots & n \end{bmatrix}.$$

For $(\nu_1, ..., \nu_m) \in I_m$, $1 \leqslant m \leqslant r+1$, the theorem guarantees the existence of $r+1-m$ indices $\nu_{m+1}, ..., \nu_{r+1}$ in $\{1, ..., n\}$ such that $p(\nu_1, ..., \nu_{\mu-1}, \nu_{\mu+1}, ..., \nu_{r+1}) \neq 0$, $1 \leqslant \mu \leqslant m$. Let $\langle \nu_1, ..., \nu_n \rangle$ be a permutation of $\{1, ..., n\}$

with the first $r+1$ indices determined as above, and let $\varepsilon = \operatorname{sgn}\langle \nu_1, ..., \nu_n \rangle$. Then the vector v in Theorem 1(iv) is determined as follows:

$$\begin{aligned}\forall\, 1 \leqslant \mu \leqslant m \qquad v_{\nu_\mu} &= p^\perp(\nu_\mu, \nu_{r+2}, ..., \nu_n)\\ &= \operatorname{sgn}\langle \nu_1, ..., \nu_{\mu-1}, \nu_{\mu+1}, ..., \nu_{r+1}, \nu_\mu, \nu_{r+2}, ..., \nu_n \rangle\\ &\quad \cdot p(\nu_1, ..., \nu_{\mu-1}, \nu_{\mu+1}, ..., \nu_{r+1})\\ &= \varepsilon \cdot (-1)^{r+1-\mu} \cdot p(\nu_1, ..., \nu_{\mu-1}, \nu_{\mu+1}, ..., \nu_{r+1}).\end{aligned}$$

Proof. (i) $\Rightarrow$ (ii). From the very definition of the index sets it follows that (i) is equivalent to

(i)′ For each μ-tuple of indices $\{\nu'_1, ..., \nu'_\mu\} \subset \{\nu_1, ..., \nu_m\}$, $1 \leqslant \mu \leqslant m-1$, there is an $(r-\mu)$-tuple of indices N' in $\{1, ..., n\}$ such that $p(\nu'_1, ..., \nu'_\mu, N') \neq 0$, and $p(\nu_1, ..., \nu_m, N) = 0$ for all $N \subset \{1, ..., n\}$, $\# N = r - m$.

Let us assume that (i)′ holds, and let $\mu = m-1$. Then the row vectors of the matrix

$$\begin{bmatrix} u^{(1)} & \cdots & u^{(r)} \\ \nu'_1 & \cdots & \nu'_{m-1} \end{bmatrix}$$

are linearly independent; i.e.,

$$m - 1 = \operatorname{rank}\begin{bmatrix} u^{(1)} & \cdots & u^{(r)} \\ \nu'_1 & \cdots & \nu'_{m-1} \end{bmatrix} \leqslant \operatorname{rank}\begin{bmatrix} u^{(1)} & \cdots & u^{(r)} \\ \nu_1 & \cdots & \nu_m \end{bmatrix} \leqslant m.$$

Assume,

$$\operatorname{rank}\begin{bmatrix} u^{(1)} & \cdots & u^{(r)} \\ \nu_1 & \cdots & \nu_m \end{bmatrix} = m.$$

Since $\dim U = r$, there exist $r - m$ further indices $\nu_{m+1}, ..., \nu_r$ in $\{1, ..., n\}$ such that $p(\nu_1, ..., \nu_m, \nu_{m+1}, ..., \nu_r) \neq 0$ which contradicts the fact that $p(\nu_1, ..., \nu_m, N) = 0$ for all $N \subset \{1, ..., n\}$, $\# N = r - m$.

(ii) $\Rightarrow$ (iii). Let U' be an m-dimensional subspace of U; without loss of generality we may assume that the first m basis vectors $u^{(1)}, ..., u^{(m)}$ of U form a basis of U'. From the fact that

$$\operatorname{rank}\begin{bmatrix} u^{(1)} & \cdots & u^{(r)} \\ \nu_1 & \cdots & \nu_m \end{bmatrix} = m - 1$$

it follows that

$$\operatorname{rank}\begin{bmatrix} u^{(1)} & \cdots & u^{(m)} \\ \nu_1 & \cdots & \nu_m \end{bmatrix} \leqslant m - 1;$$

i.e., $p_{U'}(\nu_1, ..., \nu_m) = 0$. On the other hand, for each subset $\{\nu'_1, ..., \nu'_{m-1}\} \subset \{\nu_1, ..., \nu_m\}$ there exist basis vectors $u^{(\rho_1)}, ..., u^{(\rho_{m-1})}$ of U such that

$$\det\begin{bmatrix} u^{(\rho_1)} & \cdots & u^{(\rho_{m-1})} \\ \nu'_1 & \cdots & \nu'_{m-1} \end{bmatrix} \neq 0;$$

i.e., the subspace $U'' = \operatorname{span}\{u^{(\rho_1)}, ..., u^{(\rho_{m-1})}\}$ has dimension $m-1$ and satisfies the condition $p_{U''}(\nu'_1, ..., \nu'_{m-1}) \neq 0$.

(iii)$\Rightarrow$(iv). As seen above, condition (iii) implies that there are $(m-1)$ basis vectors $u^{(\rho_1)}, ..., u^{(\rho_{m-1})}$ of U such that

$$\operatorname{rank}\begin{bmatrix} u^{(\rho_1)} & \cdots & u^{(\rho_{m-1})} \\ \nu_1 & \cdots & \nu_m \end{bmatrix} = m-1.$$

Consequently, there is, up to a scalar multiple, a unique vector $v \in \mathbb{R}^m$, $v \neq 0$, which is perpendicular to all the column vectors of the matrix above. Extending v to $\mathbb{R}^n$ by setting $v_\nu = 0$ for $\nu \in \{1, ..., n\} \setminus \{\nu_1, ..., \nu_m\}$, we have

$$U^\perp \cap \operatorname{span}\{e_{\nu_1}, ..., e_{\nu_m}\} = \operatorname{span}\{v\}.$$

It remains to show that $v_{\nu_\mu} \neq 0$ for all $1 \leqslant \mu \leqslant m$. Assume $v_{\nu_1} = 0$. Then the vector $(v_{\nu_2}, ..., v_{\nu_m})^T$ is perpendicular to each of the column vectors of the matrix

$$\begin{bmatrix} u^{(1)} & \cdots & u^{(r)} \\ \nu_2 & \cdots & \nu_m \end{bmatrix}.$$

But since the rank of the matrix above is equal to $m-1$, all the coefficients v_{ν_μ}, $1 \leqslant \mu \leqslant m$, of v have to vanish.

(iv)$\Rightarrow$(i). The vector $(v_{\nu_1}, ..., v_{\nu_m})^T \neq 0$ is perpendicular to each of the column vectors of the matrix

$$\begin{bmatrix} u^{(1)} & \cdots & u^{(r)} \\ \nu_1 & \cdots & \nu_m \end{bmatrix}.$$

Thus the rank of this matrix is not greater than $m-1$; i.e., the Plücker–Graßmann coordinates satisfy the condition $p(\nu_1, ..., \nu_m, N) = 0$ for each choice of an $(r-m)$-tuple of indices $N \subset \{1, ..., n\}$. Suppose $\{\nu'_1, ..., \nu'_\mu\} \subset \{\nu_1, ..., \nu_m\}$ and $(\nu'_1, ..., \nu'_\mu) \in I_\mu$, $1 \leqslant \mu \leqslant m-1$. Then there exists a vector $w \in \mathbb{R}^n$, $w \neq 0$, such that $w_\nu = 0$ for all indices $\nu \in \{1, ..., n\} \setminus \{\nu'_1, ..., \nu'_\mu\}$ and $(w_{\nu'_1}, ..., w_{\nu'_\mu})^T$ is perpendicular to each of the column vectors of the matrix given above. In particular, $w \notin \operatorname{span}\{v\}$, but this is a contradiction to $w \in U^\perp \cap \operatorname{span}\{e_{\nu_1}, ..., e_{\nu_m}\} = \operatorname{span}\{v\}$. It follows that

$$\operatorname{rank}\begin{bmatrix} u^{(1)} & \cdots & u^{(r)} \\ \nu'_1 & \cdots & \nu'_\mu \end{bmatrix} = \mu,$$

and consequently, we can complete the μ row vectors of this matrix with $r-\mu$ further row vectors of the matrix $\left[\begin{smallmatrix} u^{(1)} & \cdots & u^{(r)} \\ 1 & \cdots & n \end{smallmatrix}\right]$ to form a basis of $\mathbb{R}^r$; that is, there is an $(r-\mu)$-tuple of indices $N \subset \{1, ..., n\}$ such that $p(\nu'_1, ..., \nu'_\mu, N) \neq 0$. ∎

If an index ν belongs to the class I_0 then the νth row vector of U is linearly independent of each collection of the remaining row vectors of U;

i.e., the νth coordinate of every vector in $U^\perp$ vanishes. Thus the vector e_ν belongs to U; i.e., there is a decomposition of U as follows:

$$U = U' \oplus \mathbb{R}^{\# I_0},$$

where $U' \subset \mathbb{R}^{n - \# I_0}$ denotes the $(r - \# I_0)$-dimensional orthogonal projection of U onto span $\{e_\nu : \nu \in \{1, ..., n\} \backslash I_0\}$.

There is a geometric interpretation of Theorem 1(iv). Let v be the vector in $U^\perp$ associated with an index set $(\nu_1, ..., \nu_m)$ in I_m, $1 \leqslant m \leqslant r+1$,—we assume v to be normalized by the l^1-norm; i.e., $|v|_1 = \sum_{\mu=1}^m |v_{\nu_\mu}| = 1$. Let $l = \sum_{m=1}^{r+1} \# I_m$, and let us enumerate the vectors v from 1 to l. If $\bar{b}_1^1(0)$ denotes the closed unit ball of l_n^1, and if $Q := U^\perp \cap \bar{b}_1^1(0)$, then Q is a compact convex symmetric polytope. Indeed, we have the following theorem.

THEOREM 2. *We have*

$$\text{ext}\, Q = \{\pm v^1, ..., \pm v^l\};$$

i.e., the normalized vectors v in $U^\perp$ determined by the classification of $\{1, ..., n\}$ with respect to U are the extremal points of Q.

Proof. Let v be the normalized vector in $U^\perp$ associated with the index set $(\nu_1, ..., \nu_m)$ in I_m, $1 \leqslant m \leqslant r+1$. Assume v is not extremal in Q; i.e., there are vectors w and w' in Q and $0 < t < 1$ such that $v = tw + (1-t)w'$. Since

$$\begin{aligned} 1 = |v|_1 = \sum |v_\nu| &= \sum |tw_\nu + (1-t)\, w'_\nu| \\ &\leqslant \sum t\, |w_\nu| + \sum (1-t)\, |w'_\nu| \\ &\leqslant t\, |w|_1 + (1-t)\, |w'|_1 \leqslant 1, \end{aligned}$$

where the sum is taken over the index set $\{\nu_1, ..., \nu_m\}$, it follows that w as well as w' have support on this index set; i.e., $w, w' \in U^\perp \cap$ span $\{e_{\nu_1}, ..., e_{\nu_m}\}$. But then by Theorem 1(iv) the equations $w = \lambda v$ and $w' = \lambda' v$ with $|\lambda| = |\lambda'| = 1$ are true. It follows trivially that $\lambda = \lambda' = 1$, and $w = w' = v$. Conversely, let v be an extremal point of Q, and let $\{\nu_1, ..., \nu_m\}$ be its support on $\{1, ..., n\}$, $1 \leqslant m \leqslant n$. We claim that $m \leqslant r+1$ and that

$$U^\perp \cap \text{span}\, \{e_{\nu_1}, ..., e_{\nu_m}\} = \text{span}\, \{v\}.$$

Assume there is a $w \in U^\perp \cap$ span $\{e_{\nu_1}, ..., e_{\nu_m}\}$, normalized and linearly independent of v. Because of linearity we may assume that both v and w belong to the $(m-1)$-dimensional closed face

$$\left\{ y \in \mathbb{R}^n : y = \sum_{\nu \in \{\nu_1, ..., \nu_m\}} \beta_\nu \,\text{sgn}\, v_\nu e_\nu, \beta_\nu \geqslant 0, \text{and} \sum_{\nu \in \{\nu_1, ..., \nu_m\}} \beta_\nu = 1 \right\}$$

of the unit ball of l_n^1; where v even belongs to its relative interior. But then v can never be an extremal point of Q. Since $U^\perp \cap \operatorname{span}\{e_{v_1}, ..., e_{v_m}\} = \operatorname{span}\{v\}$, clearly, m is less than or equal to $r+1$. ∎

The aim is to describe the approximation behavior of U in l_n^∞; i.e., the behavior of the metric projection P_U on l_n^∞, as a set-valued mapping of $\mathbb{R}^n$ into U, as well as the behavior of the best approximants $P_U(x)$ for a single element $x \in \mathbb{R}^n$. The basic tool will be the classification of the index set $\{1, ..., n\}$ by means of the Plücker–Graßmann coordinates. Theorem 1 gives various equivalent characterizations of those index sets $(v_1, ..., v_m)$ which belong to I_m, $1 \leqslant m \leqslant r+1$. From the point of view of approximation the characterization (iv) is the most useful one. It is related to a theorem of T. J. Rivlin and H. S. Shapiro on the characterization of the elements of best approximation to a point in l_n^∞. Theorem 2, on the other hand, allows us to rewrite the problem of best approximation of a point as a problem in linear optimization over the compact convex polytope Q in l_n^1.

2. The Metric Projection

Let U be an r-dimensional subspace of $\mathbb{R}^n$, $1 \leqslant r \leqslant n-1$, and let $P_U : l_n^\infty \to U$ denote the *metric projection* of l_n^∞ onto U. By definition, for all $x \in \mathbb{R}^n$,

$$P_U(x) = \{u \in U : |x-u|_\infty = \operatorname{dist}(x; U)\},$$

where $\operatorname{dist} : l_n^\infty \to \mathbb{R}$ is the *distance function* on l_n^∞. P_U is a set-valued mapping and $P_U(x)$ is compact and convex for each x in $\mathbb{R}^n$. Indeed, $P_U(x) = U \cap \bar{b}_d(x)$, where $\bar{b}_d(x)$ is the closed ball with center at x and of radius $d = \operatorname{dist}(x; U)$. Moreover, for all $x \in \mathbb{R}^n$, for all $u \in U$, and for all $\lambda \in \mathbb{R}$

$$P_U(x+u) = P_U(x) + u \qquad \text{and} \qquad P_U(\lambda x) = \lambda P_U(x).$$

Because of these properties P_U is said to be quasi-linear.

It is further known that P_U is upper as well as lower semi-continuous, see e.g., [22, pp. 58, 62]. While the first property is an immediate consequence of the finite dimensional setting, the lower s.c. was first observed by A. L. Brown [4] in 1964; we shall come back to this fact in Section 4.

If $P_U(x)$ is single-valued for each $x \in \mathbb{R}^n$, U is called a *Chebyshev subspace*. It is well-known that U is Chebyshev exactly when each vector $u \in U$, $\neq 0$, vanishes at at most $r-1$ indices (Haar's theorem). Clearly, this is satisfied exactly when for all r-tuples of indices $(v_1, ..., v_r)$

$$p(v_1, ..., v_r) = \det \begin{bmatrix} u^{(1)} & \cdots & u^{(r)} \\ v_1 & \cdots & v_r \end{bmatrix} \neq 0;$$

i.e., if and only if all the classes $I_1, ..., I_r$ are empty (and I_{r+1} contains exactly all possible $(r+1)$-tuples of $\{1, ..., n\}$). More generally, following G. S. Rubinstein [20] a linear subspace U of l_n^∞ has *Chebyshev rank* less than or equal to t, $0 \leqslant t \leqslant r$, if for all choices of $(r-t)$-tuples of indices $(v_1, ..., v_{r-t})$ of $\{1, ..., n\}$

$$\operatorname{rank} \begin{bmatrix} u^{(1)} & \cdots & u^{(r)} \\ v_1 & \cdots & v_{r-t} \end{bmatrix} = r - t.$$

This is true exactly when the classes $I_1, ..., I_{r-t}$ are empty, see also [24].

Let us denote by

$$U^{(0)} = \{x \in \mathbb{R}^n : 0 \in P_U(x)\}$$

the *metric complement* of U in l_n^∞; and let us set

$$U_1^{(0)} = U^{(0)} \cap S_1(0),$$

the intersection of $U^{(0)}$ with the unit sphere $S_1(0)$ in l_n^∞. The latter set can be identified with the Blaschke boundary of U on $S_1(0)$. (W. Blaschke introduced the notation *Schattengrenze* of $S_1(0)$ w.r.t. U.) For this reason the metric complement is also called the *Blaschke cone* of U in l_n^∞; obviously, $U^{(0)}$ is a cone with vertex at the origin.

Our first aim is to characterize $U_1^{(0)}$ by use of the classification of the index set $\{1, ..., n\}$ w.r.t. U. To do this, let us introduce the following notation. For $1 \leqslant m \leqslant n$, let

$$\begin{aligned} S &:= S(\varepsilon_{v_1} e_{v_1}, ..., \varepsilon_{v_m} e_{v_m}) \\ &= \{x \in S_1(0) : x_{v_\mu} = \varepsilon_{v_\mu} \text{ for } 1 \leqslant \mu \leqslant m, \text{ and } |x_j| < 1 \text{ otherwise}\} \end{aligned}$$

denote a *relatively open face* of $S_1(0)$ of dimension $n-m$, where $v_1, ..., v_m$ are pairwise disjoint indices of $\{1, ..., n\}$ and $\varepsilon_{v_\mu} = \pm 1$. Clearly, the faces are pairwise disjoint and $S_1(0)$ is equal to their union; moreover, a vector x in $S_1(0)$ uniquely determines the face S to which it belongs.

Let v in $\mathbb{R}^n$, $\neq 0$, be a vector with support $\{v_1, ..., v_m\}$. We shall use the following abbreviations:

$$\operatorname{supp} v := \{v_1, ..., v_m\} \qquad \text{and} \qquad S_v = S(\operatorname{sgn} v_{v_1} e_{v_1}, ..., \operatorname{sgn} v_{v_m} e_{v_m}).$$

Consider a face S of $S_1(0)$ and an $x \in S$ such that x belongs to its Schattengrenze with respect to U, then S is contained in $U_1^{(0)}$, see the proof below. That is why we shall loosely call S a face of $U_1^{(0)}$. The face S of $U_1^{(0)}$ is said to be *maximal*, if it is not contained in the closure of a face of $U_1^{(0)}$ of higher dimension.

THEOREM 3. *Let U be an r-dimensional subspace of $\mathbb{R}^n$ with basis vectors $u^{(1)}, ..., u^{(r)}$, and let I_m, $1 \leqslant m \leqslant r+1$, be the classification of the index set $\{1, ..., n\}$ w.r.t. U. If $(\nu_1, ..., \nu_m) \in I_m$ and if v is the associated vector in $U^\perp$, then $\pm S_v$ are maximal faces of $U_1^{(0)}$. And all maximal faces of $U_1^{(0)}$ are determined this way, moreover,*

$$U_1^{(0)} = \bigcup_{v \in \operatorname{ext} Q} \overline{S_v},$$

where $\overline{S_v}$ is the closure of a maximal face and Q denotes the polytope $U^\perp \cap \bar{b}_1^1(0)$.

Proof. Let $x \in S = S(\varepsilon_{\nu_1} e_{\nu_1}, ..., \varepsilon_{\nu_m} e_{\nu_m})$ belong to $U_1^{(0)}$. By the Variational Lemma of Rivlin and Shapiro [18, Theorem 1], there exist weights $\rho_{\nu_\mu} \geqslant 0$, $\sum \rho_{\nu_\mu} = 1$, such that

$$\sum_{\mu=1}^{m} \rho_{\nu_\mu} \varepsilon_{\nu_\mu} \cdot u_{\nu_\mu} = 0, \qquad \forall u \in U.$$

Within this setting, the lemma has probably older roots than the theorem referred to above, but it seems to be difficult to point out a precise reference. Since the equation above does not depend on the components x_ν of x, $\nu \in \{1, ..., n\} \setminus \{\nu_1, ..., \nu_m\}$, each $x' \in S$ belongs to $U_1^{(0)}$, Also, if S is a maximal face of $U_1^{(0)}$, then the weights have to be strictly positive on $\{\nu_1, ..., \nu_m\}$.

Let the vector v be defined by $v = \rho_{\nu_1} \varepsilon_{\nu_1} \cdot e_1 + \cdots + \rho_{\nu_m} \varepsilon_{\nu_m} \cdot e_m$. Obviously, $v \in U^\perp$ and $|v|_1 = 1$. We claim

$$\operatorname{span}\{v\} = U^\perp \cap \operatorname{span}\{e_\nu : \nu \in \operatorname{supp} v\},$$

or $(\nu_1, ..., \nu_m) \in I_m$ by Theorem 1(iv). For $m = 1$ there is nothing to prove. Assume there is a second vector w in $U^\perp \cap \operatorname{span}\{e_{\nu_1}, ..., e_{\nu_m}\}$, linearly independent of v. A linear combination of w and v will lead to a new vector w' which vanishes in at least one index of $\{\nu_1, ..., \nu_m\}$; i.e., supp w' is a real subset of supp v. But then $S_{w'}$ is a face in $S_1(0)$ which contains S in its closure, a contradiction to the maximality of S.

Conversely, the vector $v \in \operatorname{ext} Q$ satisfies the equation,

$$0 = \sum_{\nu \in \operatorname{supp} v} v_\nu u_\nu = \sum_{\nu \in \operatorname{supp} v} \operatorname{sgn} v_\nu \, |v_\nu| \, u_\nu \qquad \text{for all} \quad u \in U.$$

From the Variational Lemma, it follows again that the face S_v belongs to the Blaschke boundary of U on $S_1(0)$. Clearly, since the support of v is minimal, S_v has to be maximal. ∎

Let S_v be a maximal open face of $U_1^{(0)}$ and let x be a point of S_v. Setting $T = \text{span}\,\{e_\nu : \nu \in \{1, ..., n\} \setminus \text{supp}\, v\}$, it is not difficult to see that

$$\dim P_U(x) = \dim (T \cap U) = r + 1 - m.$$

Indeed, while the first statement follows from the fact that S_v is maximal and open, the second one follows because $(T + U)^\perp = T^\perp \cap U^\perp = \text{span}\,\{v\}$ and, consequently $n - 1 = \dim (T + U) = \dim T + \dim U - \dim (T \cap U) = n - m + r - \dim (T \cap U)$.

If U has Chebyshev rank $\leqslant t$, $0 \leqslant t \leqslant r$, then $m \geqslant r - t + 1$, or $\dim P_U(x) \leqslant t$. Conversely, if for all $x \in \mathbb{R}^n$ $\dim P_U(x) \leqslant t$, then the classes $I_1, ..., I_{r-t}$ are empty, reproving G. S. Rubinstein's results [20] in the discrete setting.

Let us conclude the section with two simple examples:

(1) Let U be a hyperplane in $\mathbb{R}^n$, and let v be the up to a multiplicative factor uniquely defined normal vector of U. There is just one index tuple, say $(\nu_1, ..., \nu_m) \in I_m$, $1 \leqslant m \leqslant n$, and an index ν belongs to the tuple if and only if $\nu \in \text{supp}\, v$. Clearly, I_0 is the complement of $\{\nu_1, ..., \nu_m\}$ relative to $\{1, ..., n\}$ and U can be decomposed into

$$U = U' \oplus \mathbb{R}^{\# I_0},$$

where $U' = U \cap \text{span}\,\{e_{\nu_1}, ..., e_{\nu_m}\}$. In $\text{span}\,\{e_{\nu_1}, ..., e_{\nu_m}\}$, considered as an l_m^∞ in its own right, U' is a Chebyshev hyperplane.

(2) Let U be an one-dimensional subspace in $\mathbb{R}^n$, say $U = \text{span}\,\{u\}$, $u \neq 0$. There is the following classification:

$$I_1 = \{(\nu) : u_\nu = 0\}, \qquad I_2 = \{(\nu, \mu) : u_\nu \cdot u_\mu \neq 0\}, \qquad \text{and} \qquad I_0,$$

where I_0 is not empty exactly when $U = \text{span}\,\{e_\nu\}$ for some index $\nu \in \{1, ..., n\}$.

3. The Characteristic Index Set

In this section we study the set of best approximations in U of an individual vector x in l_n^∞.

The following duality relation

$$\forall x \in l_n^\infty, \qquad \text{dist}(x; U) := \min_{u \in U} |x - u|_\infty = \max_{v \in Q} \langle x, v \rangle \qquad (*)$$

is well-known in functional analysis. R. C. Buck [5] attributes it to M. G. Krein and to S. Banach. The following statement is an immediate consequence of Theorem 2.

THEOREM 4. *For each $x \in l_n^\infty$*

$$\operatorname{dist}(x; U) = \max_{v \in \{\pm v^1, \ldots, \pm v^l\}} \langle x, v \rangle.$$

In almost all books on optimization (and approximation) one will find the discrete Chebyshev approximation as an application of linear programming; in particular, one will find the statement that

$$\forall x \in l_n^\infty, \qquad \operatorname{dist}(x; U) = \max_{v \in Q} \langle x, v \rangle, \qquad \text{where} \quad Q = U^\perp \cap \bar{b}_1^1(0),$$

see for example that of L. Collatz and W. Wetterling [7, Sect. 16]. Clearly, the maximum of the linear form $\langle x, v \rangle$ over Q is assumed at an extremal point of Q. The point here is that Theorem 1 gives a way of determining the extremal points of Q explicitly. On the other hand, the simplex method does not make an explicit use of all the extremal points to calculate the maximum.

In the following we shall assume without loss of generality that for the given subspace U the class I_0 is empty. If $x \in \mathbb{R}^n \setminus U$, and if

$$\operatorname{dist}(x; U) = d = \langle x, v^{(1)} \rangle = \cdots = \langle x, v^{(k)} \rangle > \langle x, v' \rangle,$$

$v^{(1)}, \ldots, v^{(k)} \in \operatorname{ext} Q$ and $v' \in \operatorname{ext} Q \setminus \{v^{(1)}, \ldots, v^{(k)}\}$, then, choosing any $u \in P_U(x)$, the equations

$$d = \langle x, v^{(\kappa)} \rangle = \langle x - u, v^{(\kappa)} \rangle \leqslant |x - u|_\infty \, |v^{(\kappa)}|_1 = |x - u|_\infty, \qquad 1 \leqslant \kappa \leqslant k,$$

imply that for each $1 \leqslant \kappa$, $\lambda \leqslant k$ and for all $1 \leqslant \nu \leqslant n$

$$v_\nu^{(\kappa)} v_\nu^{(\lambda)} \geqslant 0.$$

In particular, the extremal vectors $v^{(1)}, \ldots, v^{(k)}$ define a face of $\operatorname{bd} Q$; in other words, their arithmetical mean determines a vector v in the boundary of Q, known as the *center of gravity* of the face. It follows that

$$\forall u \in P_U(x) \text{ and } \forall \nu \in \operatorname{supp} v, \qquad u_\nu + d \cdot \operatorname{sgn} v_\nu = x_\nu;$$

i.e., all elements of best approximation of x in U are equal at the indices ν in $\operatorname{supp} v$ and the error $x_\nu - u_\nu$ is maximal. Following J. Descloux, we call

$$C_x := \{\nu \in \{1, \ldots, n\} : |x_\nu - u_\nu| = \operatorname{dist}(x; U) \; \forall u \in P_U(x)\}$$

the *characteristic index set of x with respect to U*, and denote by C_x' its complement in $\{1, \ldots, n\}$. Clearly, $\operatorname{supp} v \subset C_x$. Moreover, we have

THEOREM 5. *Let* $x \in l_n^\infty \backslash U$, *and let* $\operatorname{dist}(x; U) = \langle x, v^{(1)} \rangle = \cdots = \langle x, v^{(k)} \rangle = \langle x, v \rangle > \langle x, v' \rangle$, *for* $v^{(1)}, ..., v^{(k)} \in \operatorname{ext} Q$ and $v' \in \operatorname{ext} Q \backslash \{v^{(1)}, ..., v^{(k)}\}$, *and for* $v = \sum_{\kappa=1}^{k} v^{(\kappa)}/k$, *then*

$$C_x = \bigcup_{\kappa=1}^{k} \operatorname{supp} v^{(\kappa)} = \operatorname{supp} v.$$

The theorem is crucial for a thorough investigation of the metric projection. It goes back to J. Descloux and is stated and proved in his doctoral thesis [8]. We shall give a proof by making use of a result of R. T. Rockafellar [19, Sect. 22, Theorem 22.6] on linear inequalities:

THEOREM. *Let L be a subspace of $\mathbb{R}^N$, and let $J_1, ..., J_N$ be real intervals. Then one and only one of the following alternatives holds*:

(a) *There exists a vector* $z = (\zeta_1, ..., \zeta_N) \in L$ *such that*

$$\zeta_1 \in J_1, ..., \zeta_N \in J_N;$$

(b) *There exists a vector* $z^* = (\zeta_1^*, ..., \zeta_N^*) \in L^\perp$ *such that*

$$\zeta_1^* J_1 + \cdots + \zeta_N^* J_N > 0.$$

If alternative (b) *holds, z^* can actually be chosen to be an elementary vector of $L^\perp$.*

The intervals are considered to be nonempty; no further restrictions are assumed. R. T. Rockafellar defines a vector of a subspace to be *elementary* if its support is minimal. In our notation a vector in $L^\perp$ is elementary if it is up to normalization equal to an extremal vector of $L^\perp \cap \bar{b}_1^1(0)$. Thus a classification of the index set $\{1, ..., N\}$ w.r.t. L determines all elementary vectors in $L^\perp$, as proved in Theorem 2.

Proof of Theorem 5. It remains to prove $C_x \subset \operatorname{supp} v$. It follows from the convexity of $P_U(x)$ that for any index $\nu \in C_x$ and for all $u \in P_U(x)$ either $x_\nu - u_\nu = \operatorname{dist}(x; U)$ or $= -\operatorname{dist}(x; U)$.

Let us assume without loss of generality that $x \in U_1^{(0)}$, and let $\nu_0 \in C_x$. Then $x_{\nu_0} = 1$ or $= -1$. We restrict ourselves further to $x_{\nu_0} = 1$. Hence, for all elements u of best approximation of x $u_{\nu_0} = 0$ holds. Set $N = r + n$. With respect to the equation $u + (\delta - x) = 0$ in $\mathbb{R}^n$, δ being the error, the following system (A) has a solution, while system (B) does not.

$$(\mathrm{A}) \quad \begin{cases} \sum_{\rho=1}^{r} \alpha_\rho u_\nu^{(\rho)} + \alpha_{r+\nu} = 0, & \nu \in \{1, ..., n\}; \\ \alpha_\rho \in \mathbb{R} =: J_\rho, & \rho \in \{1, ..., r\}; \\ \alpha_{r+\nu} \in \{0\} =: J_{r+\nu}, & \nu = \nu_0; \\ \alpha_{r+\nu} \in [-x_\nu - 1, -x_\nu + 1] =: J_{r+\nu}, & \nu \in \{1, ..., n\} \backslash \{\nu_0\}. \end{cases}$$

$$
\text{(B)}\quad \begin{cases} \sum_{\rho=1}^{r} \alpha_\rho u_\nu^{(\rho)} + \alpha_{r+\nu} = 0, & \nu \in \{1, ..., n\}; \\ \alpha_\rho \in \mathbb{R} =: J_\rho, & \rho \in \{1, ..., r\}; \\ \alpha_{r+\nu} \in [-2, 0) =: J_{r+\nu}, & \nu = \nu_0; \\ \alpha_{r+\nu} \in [-x_\nu - 1, -x_\nu + 1] =: J_{r+\nu}, & \nu \in \{1, ..., n\} \setminus \{\nu_0\}. \end{cases}
$$

We let L be the r-dimensional subspace of $\mathbb{R}^N$ the orthogonal complement $L^\perp$ of which is spanned by the column vectors of the matrix

$$
\begin{bmatrix} u_1^{(1)} & \cdots & u_n^{(1)} \\ \vdots & & \vdots \\ u_1^{(r)} & \cdots & u_n^{(r)} \\ 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} \in \mathbb{R}^{N \times n}.
$$

Since system (B) has no solution, by R. T. Rockafellar's alternative there exists a vector $z^* = (\zeta_1^*, ..., \zeta_N^*) \in L^\perp$ such that

$$
\langle z^*, z \rangle = \zeta_1^* \zeta_1 + \cdots + \zeta_N^* \zeta_N > 0
$$

for each vector $z = (\zeta_1, ..., \zeta_N)$, $\zeta_1 \in J_1, ..., \zeta_N \in J_N$.

Let us look at the vector z^* more closely. At first we remark that $\zeta_{r+\nu_0}^* \neq 0$; the coefficient being equal to zero is in contradiction to the solvability of system (A). Moreover, since the coefficients $\zeta_1, ..., \zeta_r$ of z are unrestricted, the coefficients $\zeta_1^*, ..., \zeta_r^*$ vanish. Hence, we can rewrite z^* as

$$
\zeta_\rho^* = 0, \qquad \rho = 1, 2, ..., r; \qquad \zeta_{r+\nu}^* = v_\nu, \qquad \nu = 1, 2, ..., n,
$$

where v is a vector in $U^\perp$. We choose $|v|_1 = 1$. If $v_\nu = 0$ for all $\nu \neq \nu_0$, then $\langle z^*, z \rangle = v_{\nu_0} \zeta > 0$, $-2 \leqslant \zeta < 0$; i.e., $v_{\nu_0} = -1$, and $v = -e_{\nu_0}$ is an extremal point of Q; as before, Q being the closed convex polytope $U^\perp \cap \bar{b}_1^1(0)$. Hence $\langle x, -v \rangle = 1$, and consequently $\nu_0 \in \operatorname{supp} v$.

Assume $\mu \in C_x \setminus \{\nu_0\}$. As seen above, $\operatorname{sgn} v_{\nu_0} = -1$. Setting $\zeta_{r+\nu} = 0$ for $\nu \neq \nu_0$, μ, and setting $\zeta_{r+\mu}$ in turn equal to $-x_\mu$, $-x_\mu + 1$, and $-x_\mu - 1$, we obtain

$$
-x_\mu v_\mu + \zeta v_{\nu_0} > 0 \qquad \text{and} \qquad (-x_\mu \pm 1) v_\mu + \zeta v_{\nu_0} > 0, \qquad -2 \leqslant \zeta < 0.
$$

By assumption, $|x_\mu| \leqslant 1$. For the three inequalities to be true, we have to have $|x_\mu| = 1$ and $\operatorname{sgn} v_\mu = -x_\mu$. Hence $\langle x, -v \rangle = 1$.

Let $-v$ be the convex combination of the extremal vectors $v^{(1)}, ..., v^{(k)}$ of Q, say $-v = \sum_{\kappa=1}^{k} \beta_\kappa v^{(\kappa)}$, $\beta_1, ..., \beta_k > 0$, $\sum_{\kappa=1}^{k} \beta_\kappa = 1$, then

$$1 = \langle x, -v \rangle = \sum_{\kappa=1}^{k} \beta_\kappa \langle x, v^{(\kappa)} \rangle \leqslant \sum_{\kappa=1}^{k} \beta_\kappa = 1.$$

Consequently, $\langle x, v^{(\kappa)} \rangle = 1$ for $1 \leqslant \kappa \leqslant k$, and $v_0 \in \operatorname{supp} v = \bigcup_{1 \leqslant \kappa \leqslant k} \operatorname{supp} v^{(\kappa)}$. ∎

The proof uses methods of linear optimization. R. T. Rockafellar's theorem allows one to consider a single index; the "traditional" methods in approximation theory such as the Kolmogorov criterion, and its derivation, the Variational Lemma, seem not to be sufficiently sensitive to yield Theorem 5.

J. Descloux in his proof makes use of a theorem of H. Weyl on the solutions of homogeneous linear inequalities. He calls the index sets of the classification *cadres* and defines them via characterization (ii) of Theorem 1.

In the fifties S. I. Zuhovickii [24] investigated the approximation of real-valued functions in the sense of P. L. Chebyshev on compact point sets and particularly on *minimal subsets* of these compacta, which are index sets in our notation. His definition, however, is still tied to a vector $x \in \mathbb{R}^n \setminus U$.

J. R. Rice in his famous AMS Bulletin note [16] on the strict approximation denotes these index sets *critical points sets* of best approximation to x; see also [17, Chap. 12-7]. He uses the phrase *set of critical point sets* denoting C_x. Rice's approach, however, is more difficult to follow than Descloux's one.

In 1961, T. J. Rivlin and H. Shapiro [18] introduced what they called an *extremal signature* (on a compact Hausdorff space T), see also H. Shapiro [21, Chap. 2.6]. B. Brosowski [3] picked up their notion and, more recently, W. Li [13] again; both define *primitive extremal signatures*. Using our notation a primitive extremal signature σ of U is a mapping from $\{1, ..., n\}$ to $\{-1, 0, 1\}$ which corresponds to an extremal vector v in $U^\perp$; i.e., $\sigma(\nu) = \operatorname{sgn} v_\nu$, for all $\nu \in \{1, ..., n\}$. Their approach is closer to that of J. R. Rice.

F. Bohnenblust's paper [2] on *Subspaces of $l_{p,n}$ Spaces* and his use of Plücker–Graßmann coordinates of the linear subspace U in $\mathbb{R}^n$ was the starting point of my investigation on discrete linear Chebyshev approximation. When I saw the connections between my considerations and the results of Zuhovickii, Rice, Rivlin and Shapiro, Descloux, and Brosowski, which date back more than 30 years, I was at first puzzled, later they strengthened my confidence. The approach via Plücker–Graßmann coordinates does exceed theirs and does lead to new results as we proved so far and will continue to prove.

4. Decomposition of $l_n^\infty \setminus U$

As before we assume that the index class I_0 is empty. Let us choose a face of $\operatorname{bd} Q$ with extremal points $v^{(1)}, ..., v^{(k)}$, which is uniquely determined by their arithmetical mean $v = \sum_{\kappa=1}^{k} v^{(\kappa)}/k$ and which will henceforth be denoted by face_v, and let us define

$$K_v := \{x \in \mathbb{R}^n : \operatorname{dist}(x; U) = \langle x, v^{(1)} \rangle = \cdots = \langle x, v^{(k)} \rangle = \langle x, v \rangle \geqslant \langle x, v' \rangle$$
$$\forall v' \in \operatorname{ext} Q \setminus \{v^{(1)}, ..., v^{(k)}\}\},$$

and

$$K_v^{(0)} := \operatorname{cone} \{S_v\}$$
$$= \{x \in U^{(0)} : x_\nu = d \cdot \operatorname{sgn} v_\nu,\ \nu \in C_x, \text{ and } |x_\nu| < d,\ \nu \in C'_x,\ d > 0\}.$$

For all $x \in K_v$ its characteristic index set C_x is equal to $\operatorname{supp} v$, for this reason we shall write C_v whenever we are discussing the approximation behavior of U on the cone. Since Q has only finitely many faces, there are finitely many characteristic index sets, hence finitely many cones K and $K^{(0)}$, respectively.

By the definition of these cones and by use of Theorem 5, we can extend the statement of Theorem 3.

Theorem 6. *By use of the notation given above, the cones K_v, face_v being a face of* $\operatorname{bd} Q$, *are convex, relatively open, and pairwise disjoint. They satisfy the relation*

$$K_v = K_v^{(0)} + U$$

and decompose $\mathbb{R}^n \setminus U$.

In particular, for each $v \in \operatorname{ext} Q$, K_v is an open convex cone and $\bigcup_{v \in \operatorname{ext} Q} K_v$ is dense in $\mathbb{R}^n$.

It follows that for all $x \in K_v$

$$(x + U) \cap K_v^{(0)} \neq \varnothing;$$

i.e., there exists an element of best approximation u of x such that $|x_\nu - u_\nu| < d$, $\nu \in C'_v$, where d is the distance of x from U. Furthermore, for all $x \in K_v$

$$x - P_U(x) = \overline{(x + U) \cap K_v^{(0)}} = (x + U) \cap \overline{K_v^{(0)}}.$$

Moreover, the lower semi-continuity of P_U follows easily. To indicate the proof, we may restrict ourselves to $x^{(0)} \in K_v^{(0)}$ and may select the origin as element of best approximation.

The vector v determines a subspace

$$U_v = \{u \in U : u_\nu = 0, \nu \in C_v\}$$

of U of dimension

$$r' := r - \operatorname{rank} \begin{bmatrix} u^{(1)} \cdots u^{(r)} \\ C_v \end{bmatrix};$$

let us assume that the first r' basis vectors of U, $u^{(1)}, ..., u^{(r')}$, are a basis of U_v.

Choose a sequence $\{x^{(k)}\}$, converging to $x^{(0)}$ as $k \to \infty$, and choose any sequence $\{u^{(k)}\}$ in U, where $u^{(k)} = \sum_{\rho=1}^{r} \alpha_\rho^{(k)} u^{(\rho)}$ belongs to $P_U(x^{(k)})$. Because of the upper semi-continuity of P_U, $u_\nu^{(k)} \to 0$ as $k \to \infty$ for $\nu \in C_v$ and since

$$\operatorname{rank} \begin{bmatrix} u^{(r'+1)} \cdots u^{(r)} \\ C_v \end{bmatrix} = r - r',$$

the coefficients $\alpha_{r'+1}^{(k)}, ..., \alpha_r^{(k)}$ converge to zero as $k \to \infty$ as well.

We claim, that for k large, the element

$$\tilde{u}^{(k)} = \sum_{\rho = r'+1}^{r} \alpha_\rho^{(k)} u^{(\rho)}$$

belongs to $P_U(x^{(k)})$. Indeed, for $\nu \in C_v$

$$|x_\nu^{(k)} - \tilde{u}_\nu^{(k)}| = |x_\nu^{(k)} - u_\nu^{(k)}| \leqslant d_k,$$

where d_k and d_0 are the distances of $x^{(k)}$ and $x^{(0)}$ from U, respectively, while for $\nu \in C_v'$

$$|x_\nu^{(k)} - \tilde{u}_\nu^{(k)}| \leqslant |x_\nu^{(0)}| + |x_\nu^{(k)} - x_\nu^{(0)}| + |\tilde{u}_\nu^{(k)}|,$$

the last two terms on the right-hand side converge to zero as $k \to \infty$, and the first one is strictly less than d_0. Since $d_k \to d_0$ as $k \to \infty$, the left-hand side is strictly less than d_k for k large.

Selecting for $x \in \mathbb{R}^n \setminus U$ the element in $P_U(x)$ which has minimal Euclidean norm, one obtains a continuous selection of the metric projection which possesses the property called *Nulleigenschaft*; i.e., for $x \in U^{(0)}$ the origin is selected. It was G. Nürnberger [15] who pointed out that the existence of a continuous selection with Nulleigenschaft is sufficient for lower semi-continuity of the set-valued metric projection. Two years later H. Krüger [12] proved the necessity for complete subspaces.

Further selections are easy to construct: Taking the center of $P_U(x)$ w.r.t. the Euclidean norm, or the *center of gravity* of $P_U(x)$, by continuity of the metric projection one obtains a continuous selection, respectively; see, e.g., [11].

5. The Metric Projection onto a Chebyshev Subspace

In this case I_0 as well as all index classes I_m, $0 \leqslant m \leqslant r$, are empty, while I_{r+1} includes all possible $(r+1)$-tuples of pairwise disjoint indices. Let us take face$_v$ of bd$Q$ and let K_v be the corresponding relatively open convex cone in $\mathbb{R}^n \setminus U$, as defined in the previous section.

Theorem 7. *On $\overline{K_v}$ the metric projection is linear.*

Proof. Let x and $\tilde{x}$ belong to K_v, let u and $\tilde{u}$ be their elements of best approximation in U with distances d and $\tilde{d}$, respectively, and let $x^{(0)}$ and $\tilde{x}^{(0)}$ be their respective projections in $K_v^{(0)}$. On $C_v = \{v_1, ..., v_{r+1}\}$

$$x_v^{(0)} = d \operatorname{sgn} v_v \qquad \text{and} \qquad \tilde{x}_v^{(0)} = \tilde{d} \operatorname{sgn} v_v,$$

while

$$|x_v^{(0)}| < d \qquad \text{and} \qquad |\tilde{x}_v^{(0)}| < \tilde{d} \qquad \text{on } C_v'.$$

Clearly $x + \tilde{x} \in K_v$, $x^{(0)} + \tilde{x}^{(0)}$ belongs to $K_v^{(0)}$, and by the uniqueness of the representation,

$$x + \tilde{x} = (x^{(0)} + \tilde{x}^{(0)}) + (u + \tilde{u}) = (x + \tilde{x})^{(0)} + P_U(x + \tilde{x}).$$

It follows that

$$P_U(x + \tilde{x}) = u + \tilde{u} \qquad \text{and} \qquad \operatorname{dist}(x + \tilde{x}; U) = d + \tilde{d},$$

proving additivity. Since the metric projection is homogeneous, we have trivially

$$P_U(\lambda x) = \lambda u \qquad \text{and} \qquad \operatorname{dist}(\lambda x; U) = |\lambda| \operatorname{dist}(x; U), \qquad \lambda \in \mathbb{R}.$$

Clearly, these arguments carry over to $\overline{K_v} = \overline{K_v^{(0)}} + U$. ∎

Because of uniqueness, P_U is continuous, and because of P_U being linear on the closed convex cones $\overline{K_v}$, $v \in \operatorname{ext} Q$, we reproved the following corollary which goes back to A. K. Cline [6] and M. Bartelt [1].

Corollary. *P_U is a globally Lipschitz-continuous projection of $\mathbb{R}^n$ onto U.*

For a cone K_v, $v \in \operatorname{ext} Q$ with support $\{v_1, ..., v_{r+1}\}$, we can explicitly determine the metric projection. Indeed, for $x \in \overline{K_v}$ the system of linear equations

$$\sum_{\rho=1}^{r} \alpha_\rho u_v^{(\rho)} + \delta_v = x_v, \qquad 1 \leqslant v \leqslant n,$$

subject to the constrains

$$\delta_v = d \cdot \operatorname{sgn} v_v, \qquad \forall v \in C_v, \qquad \text{and} \qquad |\delta_v| \leqslant d, \qquad \forall v \in C'_v,$$

is uniquely solvable. Choosing an index $v \in C'_v$ and considering the subsystem determined by the indices $v_1, ..., v_{r+1}$ and v, the column vectors of the corresponding coefficient matrix

$$H = \begin{bmatrix} u_{v_1}^{(1)} & \cdots & u_{v_1}^{(r)} & \operatorname{sgn} v_{v_1} \\ \vdots & & \vdots & \vdots \\ u_{v_{r+1}}^{(1)} & \cdots & u_{v_{r+1}}^{(r)} & \operatorname{sgn} v_{v_{r+1}} \\ u_v^{(1)} & \cdots & u_v^{(r)} & 0 \end{bmatrix} \in \mathbb{R}^{(r+2)\times(r+1)}$$

span a hyperplane in $\mathbb{R}^{r+2}$. A non-zero vector perpendicular to H can be determined by use of the Plücker–Graßmann coordinates of H, its coefficients are given by

$$p_H(v_1, ..., v_{r+1}) = \sum_{\mu=1}^{r+1} (-1)^{r+1+\mu} p(v_1, ..., v_{\mu-1}, v_{\mu+1}, ..., v_{r+1}) \cdot \operatorname{sgn} v_{v_\mu},$$

for $j = r+2$, while for $1 \leqslant j \leqslant r+1$,

$$\begin{aligned} &p_H(v_1, ..., v_{j-1}, v_{j+1}, ..., v_{r+1}, v) \\ &\quad = \sum_{\mu=1}^{j-1} (-1)^{r+1+\mu} \\ &\qquad \cdot p(v_1, ..., v_{\mu-1}, v_{\mu+1}, ..., v_{j-1}, v_{j+1}, ..., v_{r+1}, v) \cdot \operatorname{sgn} v_{v_\mu} \\ &\qquad + \sum_{\mu=j+1}^{r+1} (-1)^{r+\mu} \\ &\qquad \cdot p(v_1, ..., v_{j-1}, v_{j+1}, ..., v_{\mu-1}, v_{\mu+1}, ..., v_{r+1}, v) \cdot \operatorname{sgn} v_{v_\mu}. \end{aligned}$$

For $x \in K_v$, the element of best approximation $u^{(0)}$ of x in U is then given by

$$u^{(0)} = P_v \cdot x,$$

where the projection matrix $P_v = (\pi_{v\mu}^{(v)}) \in \mathbb{R}^{n \times n}$ is determined by

$$\pi_{\nu\nu_\mu}^{(v)} = \begin{cases} \delta_{\nu\nu_\mu} - \operatorname{sgn} v_\nu \cdot v_{\nu_\mu}, & \nu_\mu \in C_v, \ \nu \in C_v, \\ (-1)^{r-\mu} \dfrac{p_H(\nu_1, ..., \nu_{\mu-1}, \nu_{\mu+1}, ..., \nu_{r+1}, \nu)}{p_H(\nu_1, ..., \nu_{r+1})}, & \nu_\mu \in C_v, \ \nu \in C'_v, \\ 0, & \text{otherwise,} \end{cases}$$

$\delta_{\nu\mu}$ being the Kronecker symbol.

6. The Strict Approximation

Here we want to extend the result of the previous section to the non-Chebyshev case. The role of the best approximation which in the Chebyshev case is a continuous point-valued projection of $\mathbb{R}^n$ onto U, is played by the strict approximation, the "best" of the best approximation. The concept of strict approximation, i.e., of selecting for a vector in $\mathbb{R}^n$ an element of best approximation such that the error is minimal in each component, was introduced by J. R. Rice [16]. We follow J. Descloux's construction [9] which differs somewhat from Rice's one and which in our notation reads as follows:

Let $I_0, ..., I_{r+1}$ be the classification of the index set $\{1, ..., n\}$ w.r.t. U.

Let $x \in \mathbb{R}^n \backslash U$. If ν is an index belonging to I_0, we define the νth component of its strict approximation to be equal to x_ν. So we may as well assume that $I_0 = \varnothing$.

It follows from Theorem 6 that there exists a face of $\operatorname{bd} Q$, say face_v, such that x belongs to K_v. As in Section 4, let U_v denote the linear subspace of U defined by

$$U_v = \{u \in U : u_\nu = 0, \ \nu \in C_v\};$$

and let us recall that U_v has dimension

$$r' = r - \operatorname{rank} \begin{bmatrix} u^{(1)} \cdots u^{(r)} \\ C_v \end{bmatrix}.$$

If $r' = 0$, then x has a unique element of best approximation which we define to be its strict approximation, too.

Let $0 < r' (\leqslant r)$. We may select a basis of U such that the first r' vectors $u^{(1)}, ..., u^{(r')}$, form a basis of U_v. With this convention we obtain for each $u = \sum_{\rho=1}^{r} \alpha_\rho u^{(\rho)} \in P_U(x)$ the representation

$$\sum_{\rho=r'+1}^{r} \alpha_\rho u_\nu^{(\rho)} + d \operatorname{sgn} v_\nu = x_\nu, \qquad \nu \in C_v,$$

where d is the distance and where the coefficients α_ρ, $r'+1 \leqslant \rho \leqslant r$, are uniquely determined, while

$$\sum_{\rho=1}^{r'} \alpha_\rho u_\nu^{(\rho)} + \sum_{\rho=r'+1}^{r} \alpha_\rho u_\nu^{(\rho)} + \delta_\nu = x_\nu, \qquad \nu \in C'_v.$$

For the error $\delta = x - u$, $|\delta_\nu| \leqslant d$ on C'_v, there exists, however, an element of best approximation of x for which $|\delta_\nu| < d$ on C'_v holds true.

Next we restrict x to the index set C'_v and write x', and similarly, we write $U' = \text{span}\,\{u'^{(1)}, ..., u'^{(r')}\}$ for the restriction of U_v. In case $x' - \sum_{\rho=r'+1}^{r} \alpha_\rho u'^{(\rho)}$ belongs to U'; i.e.,

$$\sum_{\rho=1}^{r'} \alpha'_\rho u_\nu'^{(\rho)} + \sum_{\rho=r'+1}^{r} \alpha_\rho u_\nu'^{(\rho)} = x_\nu, \qquad \nu \in C',$$

for some $\alpha'_1, ..., \alpha'_{r'}$, then the strict approximation of x is defined by

$$u^s = \sum_{\rho=1}^{r'} \alpha'_\rho u^{(\rho)} + \sum_{\rho=r'+1}^{r} \alpha_\rho u^{(\rho)};$$

again, we are done. Otherwise, let $I'_0, ..., I'_{r'+1}$ be the classification of the index set C'_v with respect to U'. I'_0 has to be empty, since an index in I'_0 will also belong to I_0 which is assumed to be empty. Thus U' does not degenerate to $\{0\}$ and also to $\mathbb{R}^{n'}$, $n' = \# C'_v$. But then there exists a face of bd$\,Q'$, say face$_{v'}$, such that $x' - \sum_{\rho=r'+1}^{r} \alpha_\rho u'^{(\rho)}$ belongs to $K_{v'}$. Let

$$U_{v'} = \{u' \in U' : u'_\nu = 0, \nu \in C_{v'}\},$$

and let $u'^{(1)}, ..., u'^{(r'')}$ be its basis. For each $u' \in P_{U'}(x' - \sum_{\rho=r'+1}^{r} \alpha_\rho u'^{(\rho)})$

$$\sum_{\rho=r''+1}^{r'} \alpha_\rho u_\nu^{(\rho)} + \sum_{\rho=r'+1}^{r} \alpha_\rho u_\nu^{(\rho)} + d' \,\text{sgn}\, v'_\nu = x_\nu, \qquad \nu \in C_{v'},$$

where the distance $d' < d$, and where the coefficients α_ρ, $r''+1 \leqslant \rho \leqslant r'$, are again uniquely determined.

If $r'' = 0$, we are done; otherwise we continue the process. After s steps, $s \leqslant n - r$, one obtains a uniquely defined element u^s in $P_U(x)$, the so-called *strict approximation* of x.

In other words, there exist a sequence of v-vectors, say $v^1, ..., v^s$, a corresponding sequence of pairwise disjoint characteristic sets $C^1, ..., C^s$, a sequence of subspaces $(U=)\,U^1 \supset \cdots \supset U^s$ of dimensions $r^1, ..., r^s$, respectively, and a sequence of distances $(d=)\ d_1 > \cdots > d_s > 0$, such that u^s is uniquely determined and the error $x - u^s$ satisfies

$$\delta_\nu = d_\sigma \,\text{sgn}\, v_\nu^\sigma \qquad \text{for} \quad \nu \in C^\sigma, 1 \leqslant \sigma \leqslant s,$$

or, for some $1 \leqslant s' \leqslant s-1$

$$\delta_\nu = \begin{cases} d_\sigma \operatorname{sgn} v_\nu^\sigma & \text{for } \nu \in C^\sigma,\ 1 \leqslant \sigma \leqslant s', \\ 0 & \text{for } \nu \text{ otherwise;} \end{cases}$$

to simplify notation we write C^σ instead of C_{v^σ}. Let us, in addition, write C^{s+1} for the indices of $\{1, ..., n\}$ not belonging to any C^σ, $1 \leqslant \sigma \leqslant s$. Take such a sequence of v-vectors, say $v^1, ..., v^s$, and the corresponding sequence of characteristic sets, $C^1, ..., C^s$, and C^{s+1}, the convex cones

$$\begin{aligned} K^{(0)}_{v^1 \cdots v^s} := \{x \in \mathbb{R}^n : & \text{ there exists } d_1 > \cdots > d_s > 0 \text{ such that} \\ & x_\nu = d_\sigma \operatorname{sgn} v_\nu^\sigma \text{ for all } \nu \in C^\sigma, \text{ and } 1 \leqslant \sigma \leqslant s, \\ & \text{and } |x_\nu| < d_s \text{ for all } \nu \in C^{s+1}\} \end{aligned}$$

and, for $1 \leqslant s' \leqslant s-1$,

$$\begin{aligned} L^{(0)}_{v^1 \cdots v^{s'}} := \{x \in \mathbb{R}^n : & \text{ there exists } d_1 > \cdots > d_{s'} > 0 \text{ such that} \\ & x_\nu = d_\sigma \operatorname{sgn} v_\nu^\sigma \text{ for all } \nu \in C^\sigma, \text{ and } 1 \leqslant \sigma \leqslant s', \\ & \text{and } x_\nu = 0 \text{ otherwise}\} \end{aligned}$$

belong to the metric complement $U^{(0)}$ of U and the origin is the strict approximation to each of its elements. Clearly, if for two different sequences of v-vectors the first vectors $v^1 \cdots v^{s'}$ are equal, so are the associated cones $L^{(0)}_{v^1 \cdots v^{s'}}$, $1 \leqslant s' \leqslant s-1$.

It follows from above that for each $x \in \mathbb{R}^n \backslash U$ there exists a sequence of v-vectors, $v^1, ..., v^s$, such that x belongs to

$$K_{v^1 \cdots v^s} := K^{(0)}_{v^1 \cdots v^s} + U \qquad \text{or} \qquad L_{v^1 \cdots v^{s'}} := L^{(0)}_{v^1 \cdots v^{s'}} + U, \qquad 1 \leqslant s' \leqslant s-1.$$

Moreover, the finite family of K- and L-cones partitions $\mathbb{R}^n \backslash U$; i.e., for all $x \in \mathbb{R}^n \backslash U$ there exists a unique K-cone or L-cone such that

$$x = x^0 + u^s.$$

As in the Chebyshev case, the strict approximation is linear on each cone. The strict approximation is known to be continuous, see, e.g. [17, Sects. 12–7] or [8, 9]; it also follows almost immediately from the considerations given above. Indeed, the following stronger statement holds true, which was conjectured by W. Li [13] in 1990:

THEOREM 8. *The strict approximation is globally Lipschitz continuous.*

Let us consider a sequence of v-vectors, say $v^1, ..., v^s$, and let $K^{(0)}_{v^1 \cdots v^s}$ and $L^{(0)}_{v^1 \cdots v^{s'}}$, $1 \leqslant s' \leqslant s-1$, be the corresponding cones in $U^{(0)}$. The cones are relatively open. In case all vectors v^σ, $1 \leqslant \sigma \leqslant s$, are extremal, the cone $K^{(0)}_{v^1 \dots v^s}$ is of dimension $n-r$, and, consequently,

$$K_{v^1 \dots v^s} = K^{(0)}_{v^1 \dots v^s} + U$$

is open in $\mathbb{R}^n$. Indeed, let $C^1, ..., C^s$ be the associated sequence of characteristic sets, and let $m^1, ..., m^s$ be their respective cardinality.

If v^σ is extremal, $r^{\sigma+1} = r^\sigma - m^\sigma + 1$, $1 \leqslant \sigma < s$, while $0 = r^s - m^s + 1$. It follows that $m^1 + \cdots + m^s = r + s$ and that the cardinality of C^{s+1} is equal to $n-r-s$. Hence the dimension of $K^{(0)}_{v^1 \dots v^s}$ equals $n-r$. On the other hand, if v^σ is a proper mean of extremal vectors, then $r^\sigma - m^\sigma + 1 < r^{\sigma+1}$, $1 \leqslant \sigma < s$, while $r^s - m^s + 1 < 0$. Hence the dimension of $K^{(0)}_{v^1 \dots v^s}$ is strictly less than $n-r$. Clearly, the cones $L^{(0)}_{v^1 \dots v^{s'}}$ are of dimension strictly less than $n-r$. The union of all cones $K_{v^1 \dots v^s}$, where the v-vectors are all extremal, form an open and dense subset of $\mathbb{R}^n$.

We actually do not need these considerations. Since there are only finitely many relatively open cones which decompose $\mathbb{R}^n \setminus U$, there have to be open ones the union of which is dense. The argument given above characterizes these open cones.

In some way the strict approximation is a "maximally" linear continuous selection.

7. Linear Selection

It is a trivial fact that the metric projection onto an r-dimensional subspace U of l_n^∞ admits a linear selection if and only if the metric complement $U^{(0)}$ contains a subspace of dimension $n-r$, see, e.g., [10] for a general discussion of this matter as well as [14]. The classification of the index set w.r.t. U characterizes $U^{(0)}$ and as a consequence allows further characterizations of linear selections.

THEOREM 9. *The metric projection P_U admits a linear selection, exactly when*

(i) *the union of the classes $I_1, ..., I_{r+1}$ contains $n-r$ pairwise disjoint index sets, or*

(ii) *there exists a basis in the orthogonal complement of U the vectors of which have pairwise disjoint supports.*

The two characterizations are obviously equivalent, we only have to prove one. The proof rests on the following lemma which is of some interest in itself.

LEMMA. *An s-dimensional subspace L of $\mathbb{R}^n$, $1 \leqslant s \leqslant n$, has a basis the vectors of which attain their maximum norm on pairwise disjoint index sets.*

Proof. For $s=1$ and $s=n$ there is nothing to prove.

At first we claim the existence of a relatively open face S of the unit sphere $S_1(0)$ in l_n^∞ for which $\dim(L \cap S) = s-1$. Otherwise all relatively open faces S of $S_1(0)$ satisfy $\dim(L \cap S) < s-1$, and, consequently, $L \cap S_1(0)$ has dimension less than $s-1$, which is a contradiction.

To construct the basis, let S^1 be a face in $S_1(0)$ with $\dim(S^1 \cap L) = s-1$; i.e., $S^1 \cap L$ contains linearly independent vectors $w^1, w^1+y^1, \ldots, w^1+y^{s-1}$ which attain their maximum norm at the index set J_1 determined by S^1. The vectors $y^1, \ldots, y^{s-1}$ belong to L and have their support on $\{1, \ldots, n\} \setminus J_1$; they determine an $(s-1)$-dimensional subspace L_1 of L. We repeat this process on L_1, and so forth, and obtain a basis $w^1, \ldots, w^{s-1}$ of L with the desired property. ∎

Proof of Theorem 9. Let L be an $(n-r)$-dimensional subspace of $U^{(0)}$. The lemma guarantees a basis of L the vectors of which attain their maximum norm at $n-r$ pairwise disjoint index sets. By the characterization of $U_1^{(0)}$, each of this l_n^∞-normed basis vectors of L belongs to a face S_v of $U_1^{(0)}$, v being a vector in $U^\perp$ with support in one of those $n-r$ pairwise disjoint index sets.

Obviously, each of these index sets contains a sub(index)set belonging to one of the classes $I_1, \ldots, I_{r+1}$.

Conversely, let us denote the $n-r$ disjoint index sets of the classification by $J_1, \ldots, J_{n-r}$, and let $v^{(1)}, \ldots, v^{(n-r)}$ be the corresponding extremal vectors of Q. The vectors

$$w^{(\rho)} = \sum_{\nu \in J_\rho} \operatorname{sgn} v_\nu^{(\rho)} e_\nu, \qquad 1 \leqslant \rho \leqslant n-r,$$

belong to $U^{(0)}$ and define an $(n-r)$-dimensional subspace of $U^{(0)}$. ∎

It follows from the proof that the vectors $u^{(1)}, \ldots, u^{(r)}, w^{(1)}, \ldots, w^{(n-r)}$ form a basis of $\mathbb{R}^n$; i.e.,

$$\text{for all } x \in \mathbb{R}^n \qquad x = \sum_{\rho=1}^{r} \alpha_\rho u^{(\rho)} + \sum_{\rho=1}^{n-r} \beta_\rho w^{(\rho)}, \qquad \alpha_\rho, \beta_\rho \in \mathbb{R}.$$

Taking into account that the extremal vectors $v^{(\sigma)}$, $1 \leqslant \sigma \leqslant n-r$, satisfy the orthogonality relations

$$\langle v^{(\sigma)}, u^{(\rho)} \rangle = 0, \qquad 1 \leqslant \rho \leqslant r,$$

and

$$\langle v^{(\sigma)}, w^{(\rho)} \rangle = \delta_{\sigma\rho}, \qquad 1 \leqslant \rho \leqslant n-r,$$

we obtain $\beta_\rho = \langle x, v^{(\rho)} \rangle$ for all $1 \leqslant \rho \leqslant n-r$. The linear selection L_U of P_U determined by $L = \text{span}\,\{w^{(1)}, ..., w^{(n-r)}\}$ is given by

$$L_U(x) = x - \sum_{\rho=1}^{n-r} \langle x, v^{(\rho)} \rangle\, w^{(\rho)}.$$

In case the index class I_0 is empty, it follows from above that the linear selection is uniquely determined and equal to the strict approximation.

Finally, in his paper on linear selections P.-K. Lin [14] proved that P_U admits a linear selection if and only if there exists a basis $u^{(1)}, ..., u^{(r)}$ of U such that $\#\text{supp}\, u^{(\rho)} \leqslant 2$, $1 \leqslant \rho \leqslant r$.

Indeed, his characterization is equivalent to the ones in Theorem 9: If $U^\perp$ has the basis $v^{(1)}, ..., v^{(n-r)}$ satisfying property (ii) of Theorem 9, it is straightforward to construct a basis of U the vectors of which are supported by at most two indices. The converse is proved similarly.

Acknowledgment

I thank Professor Berens for his helpful suggestions during the preparation of this paper.

References

1. M. Bartelt, On Lipschitz conditions, strong unicity and a theorem of A. K. Cline, *J. Approx. Theory* **14** (1975), 245–250.
2. F. Bohnenblust, Subspaces of $l_{p,n}$ spaces, *Amer. J. Math.* **63** (1941), 64–72.
3. B. Brosowski, "Nicht-Lineare Tschebyscheff-Approximation," B. I. Hochschulskripten 808/808a, Mannheim, 1968.
4. A. L. Brown, Best n-dimensional approximation to sets of functions, *Proc. London Math. Soc. (3)* **14** (1964), 577–594.
5. R. C. Buck, Applications of duality in approximation theory, *in* "Approx. of Funct., Proc. Symp. Warren 1964," pp. 27–42, 1965.
6. A. K. Cline, Lipschitz conditions on uniform approximation operators, *J. Approx. Theory* **8** (1973), 169–172.
7. L. Collatz and W. Wetterling, Optimierungsaufgaben, *in* "Heidelberger Taschenbücher" Vol. 15, Springer-Verlag, New York, 1971.
8. J. Descloux, "Contribution au Calcul des Approximations de Tschebicheff," thèse présentée à l'Ecole Polytechnique de Zurich, 1961.
9. J. Descloux, Approximation in L^p and chebyshev approximations, *Soc. Indust. Appl. Math.* **11**, No. 4 (1963), 1017–1026.
10. F. Deutsch, Linear selections for the metric projection, *J. Funct. Anal.* **49** (1982), 269, 191.
11. F. S. Garnir–Monjoie, Selections continues de multifonctions, *Bull. Soc. Roy. Sci. Liège* **46**, No. 1–2 (1977), 51–56.
12. H. Krüger, A remark on the lower semi-continuity of the set-valued metric projection, *J. Approx. Theory* **28** (1980), 83–86.
13. W. Li, Lipschitz continuous metric selections in $C_0(T)$, *SIAM J. Math. Anal.* **21** (1990), 205–220.

14. P.-K. LIN, Remarks on linear selections for the metric projection, *J. Approx. Theory* **43** (1985), 64–74.
15. G. NÜRNBERGER, Schnitte für die metrische Projektion, *J. Approx. Theory* **20** (1977), 673–681.
16. J. R. RICE, Tchebycheff approximation in a compact metric space, *Bull. Amer. Soc.* **68** (1962), 405–410.
17. J. R. RICE, The approximation of functions, *in* "Nonlinear and Multivariate Theory," Vol. 2, Addison–Wesley, London, 1969.
18. T. J. RIVLIN AND H. S. SHAPIRO, A unified approach to certain problems of approximation and minimization, *J. Soc. Indust. Appl. Math.* **9** (1961), 670–699.
19. R. T. ROCKAFELLAR, "Convex Analysis," Princeton Univ. Press, Princeton, NJ, 1972.
20. G. S. RUBINSTEIN, On a method of investigation of convex sets, *Dokl. Akad. Nauk. SSSR* **102** (1955), 451–454. [Russian]
21. H. S. Shapiro, Topics in approximation theory, *in* "Lecture Notes in Math.," Vol. 187, Springer-Verlag, New York, 1971.
22. I. SINGER, The theory of best approximation and functional analysis, *in* "Regional Conference Series in Applied Mathematics," Vol. 13, 1974.
23. B. L. VAN DER WAERDEN, "Einführung in die algebraische Geometrie," Springer-Verlag, Berlin, 1939.
24. S. I. ZUHOVICKIĬ, On the approximation of real functions in the sense of P. L. Čebyšev, *Amer. Math. Soc. Transl. Ser. 2* **19** (1956), 221–252.