Macroeconomic Dynamics, **10**, 2006, 77–99. Printed in the United States of America.
DOI: 10.1017.S1365100505050108

MD SURVEY

CONTINUOUS PIECEWISE LINEAR FUNCTIONS

CHARALAMBOS D. ALIPRANTIS
Purdue University

DAVID HARRIS
The University of Melbourne

RABEE TOURKY
Purdue University
and
The University of Melbourne

The paper studies the function space of continuous piecewise linear functions in the space of continuous functions on the m-dimensional Euclidean space. It also studies the special case of one dimensional continuous piecewise linear functions. The study is based on the theory of Riesz spaces that has many applications in economics. The work also provides the mathematical background to its sister paper Aliprantis, Harris, and Tourky (2006), in which we estimate multivariate continuous piecewise linear regressions by means of Riesz estimators, that is, by estimators of the the Boolean form

$$\hat{Y} = \bigvee_{j \in J} \bigwedge_{i \in E_j} \left(r_i^0 + r_i^1 X_1 + r_i^2 X_2 + \cdots + r_i^m X_m \right),$$

where $\mathbf{X} = (X_1, X_2, \ldots, X_m)$ is some random vector, $\{E_j\}_{j \in J}$ is a finite family of finite sets.

1. INTRODUCTION

The purpose of this paper is twofold: first, to study the function space of continuous piecewise linear functions in the space of continuous functions; and second, to provide the necessary mathematical background to our paper, Aliprantis, Harris, and Tourky (2006), which studies statistical estimators that we dub ***Riesz estimators***. In that paper, we envisage a situation in which we seek to estimate a random variable Y based on some observed random vector $\mathbf{X} = (X_1, X_2, \ldots, X_m)$ using estimators of the Boolean form:

$$\hat{Y}_t = \bigvee_{j \in J} \bigwedge_{i \in E_j} f_i \circ \mathbf{X_t}. \qquad (\mathcal{R})$$

The research of Aliprantis is supported by the NSF Grants EIA-0075506, SES-0128039 and DMI-0122214 and the DOD Grant ACI-0325846. The research of R. Tourky is funded by the Australian Research Council Grant A00103450. Address correspondence to Charalambos D. Aliprantis, Department of Economics, Purdue University, West Lafayette, IN 47907-1310, USA; e-mail: aliprantis@mgmt.purdue.edu.

In this equation, each $f_i : \mathbf{R}^m \to \mathbf{R}$ is an affine function of the form $f_i(x) = a_i \cdot x + \alpha_i$, where $a_i \in \mathbf{R}^m$ and $\alpha_i \in \mathbf{R}$. Furthermore, $\{E_j\}_{j \in J}$ is a finite family of finite sets and $\vee$ and $\wedge$ are the vector lattice operations *almost sure supremum* and *almost sure infimum*, respectively.

One application of Riesz estimators is the parametric estimation of continuous piecewise linear functions from data. That is, a situation in which the conditional expectations $\mathbf{E}(Y_t|\mathbf{X_t})$ is equal to $f \circ \mathbf{X_t}$, where the function $f : \mathbf{R}^m \to \mathbf{R}$ is a continuous function that agrees with a finite number of affine functions. In other words, the estimated function is continuous and there exist regions $S_1, S_2, \ldots, S_p \subseteq \mathbf{R}^m$ and parameters $\beta_1, \beta_2, \ldots, \beta_p \in \mathbf{R}^{m+1}$ such that (in matrix notation)

$$\mathbf{E}(Y_t|\mathbf{X_t}) = \begin{cases} (\mathbf{1}, \mathbf{X_t})\beta_1 & \text{if } \mathbf{X_t} \in S_1, \\ (\mathbf{1}, \mathbf{X_t})\beta_2 & \text{if } \mathbf{X_t} \in S_2, \\ \quad \vdots & \\ (\mathbf{1}, \mathbf{X_t})\beta_p & \text{if } \mathbf{X_t} \in S_p. \end{cases} \tag{$\star$}$$

In this paper, we explore in a deterministic setting the relationship between continuous piecewise linear functions that induces the functional form ($\star$) and functions that are Max-Min of a finite number of affine functions that induce the form ($\mathcal{R}$). We also study in a deterministic setting the very special case of one-dimensional piecewise linear functions.

We briefly summarize the work in the present paper: we denote the function space of affine functions from $\mathbf{R}^m$ to $\mathbf{R}$ by **Aff**. A continuous piecewise linear function is a continuous function from $f : \mathbf{R}^m \to \mathbf{R}$ that agrees with a finite number of affine functions $f_1, f_2, \ldots, f_p$. These affine functions are the ***components*** of the piecewise linear function. Now, for each $i = 1, 2, \ldots, p$, let

$$S_i = \{x \in \mathbf{R}^m : f(x) = f_i(x)\}.$$

The sets $S_1, S_2, \ldots, S_p$ are the "regions" of the function. (For a complete definition, see Section 4, in which we require that each S_i be the closure of its interior.)

Following the recent work of Ovchinnikov (2002), we establish that the space of continuous piecewise linear functions is the linear lattice hull of **Aff**, that is, it is the smallest lattice subspace containing **Aff**. In particular, there exists a family of subsets $E_1, E_2, \ldots, E_J$ of $\{1, 2, \ldots, p\}$, such that

$$f(x) = \bigvee_{j=1}^{J} \bigwedge_{i \in E_j} f_i(x),$$

for every $x \in \mathbf{R}^m$. We note several things about this Max-Min representation. First, we can compute the Max-Min representation of a piecewise linear function using information about the function f and its components $f_1, f_2, \ldots, f_p$. Second, we can compute the regions of the functions starting from Max-Min representations. Third, given a set of affine functions $F = \{f_1, f_2, \ldots, f_p\}$, we can enumerate

https://doi.org/10.1017/S1365100506050103 Published online by Cambridge University Press

through a finite combination of Max-Min operations the finite family of continuous piecewise linear functions generated by the set F. This third property is exploited in Aliprantis, Harris, and Tourky (2006).

Continuous piecewise linear functions are often used in computational economics. For instance, in the computation of Nash equilibrium of two-person finite games and fixed points approximation. Therefore, the ideas studied in the present paper may be useful in computational economics. This avenue of research has not been explored by the authors.

2. RIESZ SPACES AND BANACH LATTICES

The objective of this section is to present a brief discussion of the basic mathematical background in Riesz space theory needed for the present work and to study the Riesz estimators in Aliprantis, Harris, and Tourky (2006). The mathematics behind the theory of Riesz estimators are those of Riesz spaces and Banach lattices. We recall here some basic properties of Riesz spaces, and for details and terminology we refer to Abramovich and Aliprantis (2002a), Aliprantis and Border (1999), Aliprantis and Burkinshaw (2003), Luxemburg and Zaanen (1971), and Schaefer (1974).

An ***ordered vector space*** is a real vector space L equipped with an order relation $\geq$ that is compatible with the algebraic structure of L in the sense that if $x \geq y$, then:

$$\text{(a) } x + z \geq y + z \quad \text{for each } z \in L, \quad \text{and}$$

$$\text{(b) } \alpha x \geq \alpha y \quad \text{for all } \alpha \geq 0.$$

An ordered vector space L is said to be a ***Riesz space*** (or a ***vector lattice***) if L is also a lattice in the sense that every nonempty finite subset of L has a supremum (least upper bound) and an infimum (greatest lower bound). Following the standard terminology from lattice theory, we shall denote the supremum and infimum of a set $\{x_1, \ldots, x_n\}$ by

$$\bigvee_{i=1}^{n} x_i \quad \text{and} \quad \bigwedge_{i=1}^{n} x_i,$$

respectively. In particular, the supremum and infimum of any pair of vectors x and y are denoted by $x \vee y$ and $x \wedge y$, respectively. The simplest example of a Riesz space is $\mathbf{R}$ with the usual order. Here $x \vee y$ and $x \wedge y$ are the largest and smallest numbers of the set $\{x, y\}$; for instance, $2 \vee 3 = 3$, $1 \wedge 0 = 0$, and $3 \wedge 3 = 3$.

For an element x of a Riesz space L, the ***positive part*** of x is defined by $x^+ = x \vee 0$, the ***negative part*** by $x^- = (-x) \vee 0$, and the ***absolute value*** by $|x| = x \vee (-x)$.

The following is a simple but very useful result.

LEMMA 2.1. *An ordered vector space is a Riesz space if and only if x^+ exists for each vector x.*

For an illustration of the above notions, let $L = C[0, 1]$, the vector space of all continuous real valued functions defined on $[0, 1]$. With the pointwise ordering and algebraic operations $C[0, 1]$ is a Riesz space such that for each $x \in L$ and each $t \in [0, 1]$ we have

$$x^+(t) = \max\{x(t), 0\}, \quad x^-(t) = \max\{-x(t), 0\}, \quad \text{and} \quad |x|(t) = |x(t)|.$$

Similarly, if $x \in L$ and $r \in \mathbf{R}$, then for each $t \in [0, 1]$ we have

$$(x-r)^+(t) = \begin{cases} x(t)-r & \text{if } x(t) \geq r \\ 0 & \text{if } x(t) < r, \end{cases} \quad \text{and} \quad (x-r)^-(t) = \begin{cases} r-x(t) & \text{if } x(t) \leq r \\ 0 & \text{if } x(t) > r. \end{cases}$$

Also, notice that if $\{x_1, \ldots, x_n\} \in C[0, 1]$, then for each $t \in [0, 1]$ we have

$$\left[\bigvee_{i=1}^{n} x_i\right](t) = \max\{x_1(t), \ldots, x_n(t)\} \quad \text{and} \quad \left[\bigwedge_{i=1}^{n} x_i\right](t) = \min\{x_1(t), \ldots, x_n(t)\}.$$

Since $C(\mathbf{R}^n)$ with the pointwise ordering is a Riesz space, the above formulas are also true for functions of $C(\mathbf{R}^n)$.

Our interest here is in the structure of the Riesz subspaces of a Riesz space. A vector subspace M of a Riesz space L is said to be a **Riesz subspace** (or a **vector sublattice**) if $x, y \in M$ imply that both $x \vee y$ and $x \wedge y$ belong to M. If we consider the product vector space $\mathbf{R}^\Omega$ (where Ω is any nonempty set) and order it pointwise, then (with the above lattice operations) $\mathbf{R}^\Omega$ is a Riesz space. Moreover, if Ω is a topological space, then $C(\Omega)$ (the vector space of all continuous real-valued functions on Ω) and $C_b(\Omega)$ (the vector space of all uniformly bounded continuous real-valued functions on Ω) are both Riesz subspaces of $\mathbf{R}^\Omega$.

It should be clear that arbitrary intersections of Riesz subspaces are Riesz subspaces. This implies that every nonempty subset A of a Riesz space L is included in a smallest Riesz subspace, called the **Riesz subspace** (or the **vector sublattice**) **generated** by A and denoted $\mathcal{R}(A)$.

Next, we shall briefly describe the Riesz subspace $\mathcal{R}(A)$, an important subspace for our work. For every nonempty subset A of a Riesz space L, the symbol $A^\wedge$ will denote the collection of all vectors that can be written as infima of finite subsets of A. That is, a vector $a \in L$ belongs to $A^\wedge$ if there exist vectors $a_1, a_2, \ldots, a_k \in A$ such that $a = \bigwedge_{i=1}^k a_i$. Similarly, $A^\vee$ is the set consisting of all suprema of finite subsets of A. We write $A^{\vee\wedge}$ for $(A^\vee)^\wedge$ and $A^{\wedge\vee}$ for $(A^\wedge)^\vee$. So, a vector a belongs to $A^{\vee\wedge}$ if and only if there exists a finite family $\{E_j\}_{j\in J}$ of nonempty finite subsets of L such that $a = \bigvee_{j\in J} \bigwedge E_j$. It turns out that $A^{\vee\wedge} = A^{\wedge\vee}$ is always true.

Now we can describe the Riesz subspace generated by a set as follows. For proofs and more discussion, see Section 5 of Abramovich and Aliprantis (2002a,b).

LEMMA 2.2. *The Riesz subspace* $\mathcal{R}(A)$ *generated by a vector subspace* A *of a Riesz space coincides with* $A^{\wedge\vee}$ *and also with* $A^{\vee\wedge}$. *That is,* $\mathcal{R}(A) = A^{\wedge\vee} = A^{\vee\wedge}$.

COROLLARY 2.3. *The Riesz subspace generated by a nonempty subset* A *of a vector lattice is precisely the vector space* $\mathcal{R}(A) = [A]^{\wedge\vee}$, *where* $[A]$ *is the linear span of* A.

When a Riesz space L is equipped with a norm that is compatible with the order structure of the space in the sense that $|x| \le |y|$ implies $\|x\| \le \|y\|$, then L is called a normed Riesz space.[1] A **Banach lattice** is a Riesz space that is a Banach space under a lattice norm. It is not difficult to see that in a Banach lattice the closure of a Riesz subspace is likewise a Riesz subspace.

The two classical examples of Banach lattices are the $C(\mathcal{X})$-spaces, where $\mathcal{X}$ is a compact topological space and the norm is the sup norm $\|\cdot\|_\infty$, that is,

$$\|f\|_\infty = \sup_{x\in\mathcal{X}} |f(x)|,$$

and the $L_p(\mu)$-spaces, where $1 \le p \le \infty$, and the norm is given by

$$\|f\|_p = \left[\int |f|^p \, d\mu\right]^{\frac{1}{p}}, \quad \text{if } 1 \le p < \infty \text{ and } \|f\|_\infty = \text{ess sup} f, \quad \text{if } p = \infty.$$

3. ONE-DIMENSIONAL PIECEWISE LINEAR FUNCTIONS

We present here a few properties and formulas dealing with continuous piecewise linear functions defined on **R** or on a closed interval of **R**.

DEFINITION 3.1. *A function* $f\colon \mathbf{R} \to \mathbf{R}$ *is called* ***piecewise linear*** (***affine***) *if there exist real numbers* $-\infty < a_0 < a_1 < \cdots < a_k < \infty$ *and pairs of real numbers* (m_i, b_i), $i = 0, 1, \ldots, k, k+1$, *such that*

$$f(t) = \begin{cases} m_i t + b_i & \text{if } a_{i-1} \le t \le a_i \quad \text{for some } 1 \le i \le k, \\ m_0 t + b_0 & \text{if } t \le a_0, \\ m_{k+1} t + b_{k+1} & \text{if } t \ge a_{k+1}. \end{cases}$$

The parameters $\{a_0, a_1, \ldots, a_k\}$ *and the pairs* (m_i, b_i), $i = 0, 1, \ldots, k, k+1$, *are referred to as a* ***representation*** *of* f *and the functions* $f_i(t) = m_i t + b_i$ *as the* ***components*** *of the representation.*

Similarly, a function $f\colon [a, b] \to \mathbf{R}$, *where* $[a, b]$ *is a closed interval of* **R**, *is* ***piecewise linear*** *if there exist a partition* $a = a_0 < a_1 < \cdots < a_k = b$ *of the interval* $[a, b]$ *and pairs of real numbers* (m_i, b_i), $i = 1, \ldots, k$, *such that* $f(t) = m_i t + b_i$ *for all* $a_{i-1} \le t \le a_i$.

Notice that, according to these definitions, piecewise linear functions are automatically continuous. The following result should be obvious.

https://doi.org/10.1017/S1365100506050103 Published online by Cambridge University Press

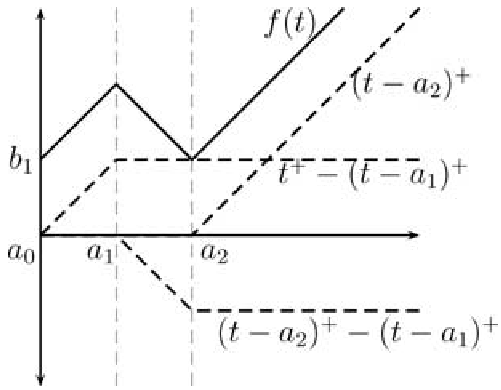

FIGURE 1. Notice that $f(t) = b_1 + t^+ - 2(t - a_1)^+ + 2(t - a_2)^+$.

LEMMA 3.2. *If* $f: \mathbf{R} \to \mathbf{R}$ *is piecewise linear, then its restriction to any closed interval of* $\mathbf{R}$ *is likewise piecewise linear. Moreover, if* $[a, b]$ *is any closed subinterval of* $\mathbf{R}$, *then the components of the piecewise linear function* $f: [a, b] \to \mathbf{R}$ *are among the components of* $f: \mathbf{R} \to \mathbf{R}$.

In addition, every piecewise linear function on a closed interval of $\mathbf{R}$ *can be extended to a piecewise linear function to all of* $\mathbf{R}$.

The piecewise linear functions on a closed interval are characterized as follows. The idea is depicted in Figure 1.

LEMMA 3.3. *Let* $f: [a, b] \to \mathbf{R}$ *be a piecewise linear function. If* $\{a_0, a_1, \ldots, a_k\}$ *and* (m_i, b_i), $i = 1, \ldots, k$, *is any representation of* f, *then for each* $t \in \mathbf{R}$ *we have*

$$f(t) = b_1 + m_1 t + \sum_{i=1}^{k-1} (m_{i+1} - m_i)(t - a_i)^+.$$

In particular, a function $f: [a, b] \to \mathbf{R}$ *is piecewise linear if and only if there exist a partition* $a = a_0 < a_1 < \cdots < a_k = b$ *of* $[a, b]$ *and constants* $c, c_0, c_1, \ldots, c_k$ *such that for each* $t \in [a, b]$ *we have* $f(t) = c + \sum_{i=0}^{k} c_i (t - a_i)^+$.

Proof. Let $a \leq t \leq b$. If $a_0 \leq t \leq a_1$, then note that

$$b_1 + m_1 t + \sum_{i=1}^{k-1} (m_{i+1} - m_i)(t - a_i)^+ = b_1 + m_1 t = f(t).$$

So, we can assume that $a_{j-1} \leq t \leq a_j$ for some $1 < j \leq k$. Notice that for each $1 < i \leq k - 1$ we have $m_i a_i + b_i = m_{i+1} a_i + b_{i+1}$ or $(m_{i+1} - m_i) a_i =$

https://doi.org/10.1017/S1365100506050103 Published online by Cambridge University Press

$-(b_{i+1} - b_i)$. Consequently, we have

$$\begin{aligned}
& b_1 + m_1 t + \sum_{i=1}^{k-1} (m_{i+1} - m_i)(t - a_i)^+ \\
& \quad = b_1 + m_1 t + \sum_{i=1}^{j-1} (m_{i+1} - m_i)(t - a_i)^+ \\
& \quad = b_1 + m_1 t + \sum_{i=1}^{j-1} (m_{i+1} - m_i)(t - a_i) \\
& \quad = b_1 + m_1 t + \left[\sum_{i=1}^{j-1} (m_{i+1} - m_i) \right] t - \sum_{i=1}^{j-1} (m_{i+1} - m_i) a_i \\
& \quad = b_1 + m_1 t + \left[\sum_{i=1}^{j-1} (m_{i+1} - m_i) \right] t + \sum_{i=1}^{j-1} (b_{i+1} - b_i) \\
& \quad = b_1 + m_1 t + (m_j - m_1) t + (b_j - b_1) = m_j t + b_j = f(t),
\end{aligned}$$

and the proof is finished. ∎

COROLLARY 3.4 [Brown, Huijsmans, and de Pagter (1991)]. *The vector subspace generated in* $C[0, 1]$ *by the collection* $\{\mathbf{1}, t\} \cup \{(\alpha - t)^+ \colon \ , \alpha \in \mathbf{R}\}$ *coincides with the Riesz subspace of all piecewise linear functions on* $[0, 1]$.

COROLLARY 3.5. *Let* $g \colon \mathbf{R} \to \mathbf{R}$ *be a piecewise linear function. If* $\{a_0, a_1, \ldots, a_k\}$ *and* (m_i, b_i), $i = 0, 1, \ldots, k, k+1$, *is an arbitrary representation of* g, *then for each* $t \in \mathbf{R}$ *we have*

$$g(t) = b_0 + m_0 t + \sum_{i=0}^{k} (m_{i+1} - m_i)(t - a_i)^+.$$

In particular, a function $f \colon \mathbf{R} \to \mathbf{R}$ *is piecewise linear if and only if there exist real constants* $m_0, b_0, a_0, a_1, \ldots, a_k$ *and* $c_0, c_1, \ldots, c_k$ *such that for each* $t \in \mathbf{R}$ *we have*

$$f(t) = b_0 + m_0 t + \sum_{i=0}^{k} c_i (t - a_i)^+.$$

Proof. Consider the function $h \colon \mathbf{R} \to \mathbf{R}$ defined by

$$h(t) = b_1 + m_1 t + \sum_{i=1}^{k-1} (m_{i+1} - m_i)(t - a_i)^+.$$

https://doi.org/10.1017/S1365100506050103 Published online by Cambridge University Press

As in the proof of Lemma 3.3, it is easy to see that $h(t) = f(t)$ for all $a_0 \le t \le a_k$. Moreover, $h(t) = m_1 t + b_1$ for all $t \le a_0$ and $h(t) = m_k t + b_k$ for all $t \ge a_k$. Since

$$m_1 t + b_1 + m_1(a_0 - t)^+ - m_0(a_0 - t)^+ = b_0 + m_0 t \text{ for all } t \le a_0$$
$$m_1 t + b_1 + m_1(a_0 - t)^+ - m_0(a_0 - t)^+ = m_1 t + b_1 \text{ for all } t \ge a_0$$
$$m_k t + b_k + (m_{k+1} - m_k)(t - a_k)^+ = m_{k+1} + b_{k+1} \text{ for all } t \ge a_k, \quad \text{and}$$
$$m_k t + b_k + (m_{k+1} - m_k)(t - a_k)^+ = m_k t + b_k \text{ for all } t \le a_k,$$

it follows that

$$\begin{aligned} g(t) &= m_1(a_0 - t)^+ - m_0(a_0 - t)^+ h(t) + (m_{k+1} - m_k)(t - a_k)^+ \\ &= m_1(a_0 - t)^+ - m_0(a_0 - t)^+ + b_1 + m_1 t + \cdots \\ &\quad \cdots + \sum_{i=1}^{k-1} (m_{i+1} - m_i)(t - a_i)^+ + (m_{k+1} - m_k)(t - a_k)^+ \\ &= m_1(a_0 - t)^+ - m_0(a_0 - t)^+ + b_1 + m_1 t + \sum_{i=1}^{k} (m_{i+1} - m_i)(t - a_i)^+ \\ &= b_0 + m_0 t + (m_1 - m_0)(t - a_0)^+ + \sum_{i=1}^{k} (m_{i+1} - m_i)(t - a_i)^+ \\ &= b_0 + m_0 t + \sum_{i=0}^{k} (m_{i+1} - m_i)(t - a_i)^+, \end{aligned}$$

as desired. ■

We close the section with two results that will be useful for our study later.

LEMMA 3.6. *Let* $f : [a, b] \to \mathbf{R}$ *be a piecewise linear function and let* $\{a_0, a_1, \ldots, a_k\}$ *and* (m_i, b_i), $i = 1, \ldots, k$, *be a representation of* f. *Also let* $m = \frac{f(b) - f(a)}{b - a}$, *the slope of the line segment joining the points* $(a, f(a))$ *and* $(b, f(b))$.

Then there exist some $1 \le i \le k$ *with* $m_i \ge m$ *and some* $a_{i-1} \le \xi \le a_i$ *satisfying* $f(\xi) = m(\xi - a) + f(a)$.

Proof. Assume by way of contradiction that if $m_i \ge m$, then we have $f(t) \ne m(t - a) + f(a)$ for all $a_{i-1} \le t \le a_i$. In particular, we have $m_1 < m$. Given that for $a \le t \le a_1$ we have $f(t) = m_1 t + b_1 = m_1(t - a) + f(a)$, the latter implies $f(t) < m(t - a) + f(a)$ for all $a < t \le a_1$. Notice that for each $a_1 \le t \le a_2$ we have

$$f(t) = m_2 t + b_2 = m_2(t - a_1) + f(a_1).$$

So, if $m_2 < m$, then for each $a_1 \le t \le a_2$ we have $f(t) < m(t - a) + f(a)$. On the other hand, if $m_2 \ge m$, then for each $a_1 \le t \le a_2$ we must have $f(t) <$

https://doi.org/10.1017/S1365100506050103 Published online by Cambridge University Press

$m(t-a)+f(a)$; otherwise (by the intermediate value theorem) there should exist some $a_1 \le \xi \le a_2$ with $f(\xi) = m(t-a) + f(a)$, which contradicts our assumption. The same argument yields $f(t) < m(t-a)+f(a)$ for all $a_2 \le t \le a_3$. Continuing this way we see that $f(a_k) = f(b) < m(b-a)+f(a) = f(b)$, which is impossible. ■

As an immediate consequence we get the following result.

COROLLARY 3.7 [Ovchinnikov (2002)]. *Let $f: [a,b] \to \mathbf{R}$ be a piecewise linear function and let $\{a_0, a_1, \ldots, a_k\}$ and (m_i, b_i), $i = 1, \ldots, k$, be the parameters of a representation of f. Then there exists some $1 \le i \le k$ such that $f(a) \ge m_i a + b_i$ and $f(b) \le m_i b + b_i$.*

Proof. According to Lemma 3.6 there exist some $1 \le i \le k$ and some $a_{i-1} \le \xi \le a_i$ satisfying $m_i \ge m = \frac{f(b)-f(a)}{b-a}$ and $f(\xi) = m(t-a) + f(a)$. Note that for each $a_{i-1} \le t \le a_i$ we have $m_i t + b_i = m_i(t-\xi) + f(\xi)$ and that for all $a \le t \le b$ we have $m(t-a) + f(a) = m(t-\xi) + f(\xi)$. This implies $m_i t + b_i \le m(t-a) + f(a)$ for all $a \le t \le \xi$ and $m_i t + b_i \ge m(t-a) + f(a)$ for all $\xi \le t \le b$, and our conclusion follows. ■

4. MULTIVARIATE PIECEWISE LINEAR FUNCTIONS

Recall that any function $f: \mathbf{R}^m \to \mathbf{R}$ of the form $f(x) = \alpha + a \cdot x$, where $\alpha \in \mathbf{R}$ is a constant and $a \in \mathbf{R}^m$ is a fixed vector, is called an **affine function**. As usual, an affine function f is **linear** if $\alpha = 0$, i.e., $f(x) = a \cdot x$. A function $f: S \to \mathbf{R}$, where S is a subset of $\mathbf{R}^m$, is said to be an **affine function** if it is the restriction of an affine function defined on $\mathbf{R}^m$. Let **Aff** denote the collection of all affine functions on $\mathbf{R}^m$ and note that **Aff** is a vector subspace of $C(\mathbf{R}^m)$.

LEMMA 4.1. *Regarding affine functions we have the following*:

1. *The vector space* **Aff** *of all affine functions is the linear span in $C(\mathbf{R}^m)$ of the functions $\{\mathbf{1}, e_1, e_2, \ldots, e_m\}$, where $\mathbf{1}(x) = 1$ and $e_i(x) = x_i$ for all $x \in \mathbf{R}^m$. That is, we have* $\mathbf{Aff} = \mathrm{Span}\,\{\mathbf{1}, e_1, e_2, \ldots, e_m\}$; *and so* **Aff** *is an $(m+1)$-dimensional vector space.*[2]
2. *Two affine functions $f, g \in$* **Aff** *coincide if and only if $f(x) = g(x)$ for all x in a nonempty open subset of $\mathbf{R}^m$. In particular, if a subset S of $\mathbf{R}^m$ has an interior point, then any affine function on S is the restriction of a unique affine function defined on $\mathbf{R}^m$.*

Proof. The proof of part (1) is obvious. The proof of part (2) follows easily from the following simple property: *If a nonzero linear functional f satisfies $f(x) \ge \alpha$ for all x in a nonempty open set $\mathcal{O}$, then $f(x) > \alpha$ must be the case for all $x \in \mathcal{O}$.*

To see this, fix $x \in \mathcal{O}$ and assume that $f(x) = \alpha$. Since $\mathcal{O}$ is an open set, there exists some $\epsilon > 0$ such that $x + B(0,\epsilon) \subseteq \mathcal{O}$. So, for each $y \in B(0,\epsilon)$ we have $\alpha + f(y) = f(x+y) \ge \alpha$ or $f(y) \ge 0$. This implies $f(y) = 0$ for all $y \in B(0,\epsilon)$ and so $f = 0$, which is impossible. ■

We are now ready to introduce the general concept of piecewise linear function.

DEFINITION 4.2. *A function* $f: \mathbf{R}^m \to \mathbf{R}$ *is called* ***piecewise linear*** (*or* ***piecewise affine***) *if there exist distinct affine functions* $f_1, f_2, \ldots, f_p$ *and subsets* $S_1, S_2, \ldots, S_p$ *of* $\mathbf{R}^m$ *such that*:

(1) *Each* S_i *is closed with nonempty interior and* $\overline{\mathrm{Int}(S_i)} = S_i$.[3]
(2) *If* $i \neq j$, *then* $\mathrm{Int}(S_i) \cap \mathrm{Int}(S_j) = \emptyset$.
(3) $\bigcup_{i=1}^{p} S_i = \mathbf{R}^m$.
(4) *If* $x \in S_i$, *then* $f(x) = f_i(x)$.

We also introduce the following terminology and notation.

(a) *The sets* S_i *are called the* ***regions*** *of* f *and the functions* f_i *will be referred to as the* ***components*** *of* f.
(b) *The pairs* $(S_1, f_1), \ldots, (S_p, f_p)$ *are the* ***characteristic pairs*** *of* f.
(c) *The set of all piecewise linear functions will be denoted by* **PL**.

A remark is in order here. The same definition of a piecewise linear function can be given for **solid** domains, that is, for closed convex subsets of $\mathbf{R}^m$ with nonempty interior. All results in this section hold true for piecewise linear functions with solid domains. We assume that our functions have domain $\mathbf{R}^m$ for the sole purpose of simplifying the exposition. The reader can verify directly that when $m = 1$ the definitions for piecewise linear functions given in Definitions 3.1 and 4.2 are equivalent; see also Corollary 4.10.

Here is an example of an piecewise linear function with a solid domain in $\mathbf{R}^2$.

Example 4.3. Let $Q = [0, 12] \times [0, 12] = \{(x, y) \in \mathbf{R}^2\colon 0 \leq x \leq 12;\ 0 \leq y \leq 12\}$. Consider the piecewise linear function $f: Q \to \mathbf{R}$ defined by

$$f(x_1, x_2) = \begin{cases} x_1 - 5 & \text{if } x_2 \geq x_1 \ \&\ 2x_1 \geq 17 - x_2, \\ x_2 - 5 & \text{if } x_2 \leq x_1 \ \&\ x_1 \geq 17 - 2x_2, \\ -x_1 - x_2 - 12 & \text{if } x_2 \geq x_1 \ \&\ 2x_1 \leq 17 - x_2 \ \&\ 2x_1 \geq 17 - 2x_2, \\ & \quad \text{or } x_2 \leq x_1 \ \&\ x_1 \leq 17 - 2x_2 \ \&\ 2x_1 \geq 17 - 2x_2, \\ x_1 + x_2 - 5 & \text{if } 2x_1 \leq 17 - 2x_2. \end{cases}$$

The regions of this function are shown in Figure 2 and its graph is depicted in Figure 3.

Notice that the regions cannot be specified by separate thresholds on the variables x_1 and x_2. This would be the case only when the function f is itself separable. ■

The rest of the discussion in this section is devoted to the properties of piecewise linear functions. The fundamental result for our work will be obtained in the sequel

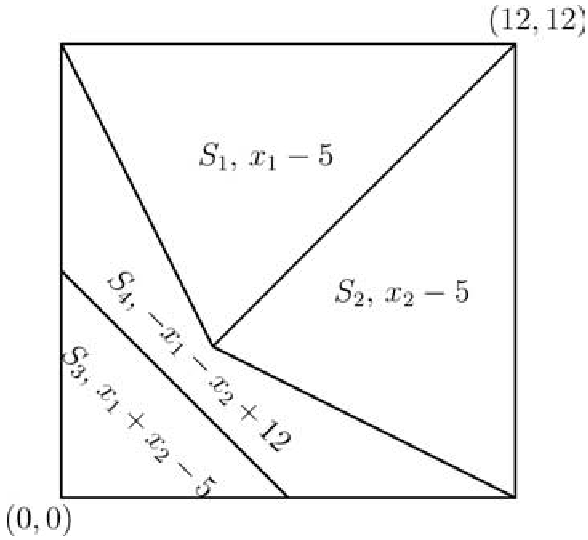

FIGURE 2. The regions of the function $f\colon \mathbf{R}^2 \to \mathbf{R}$.

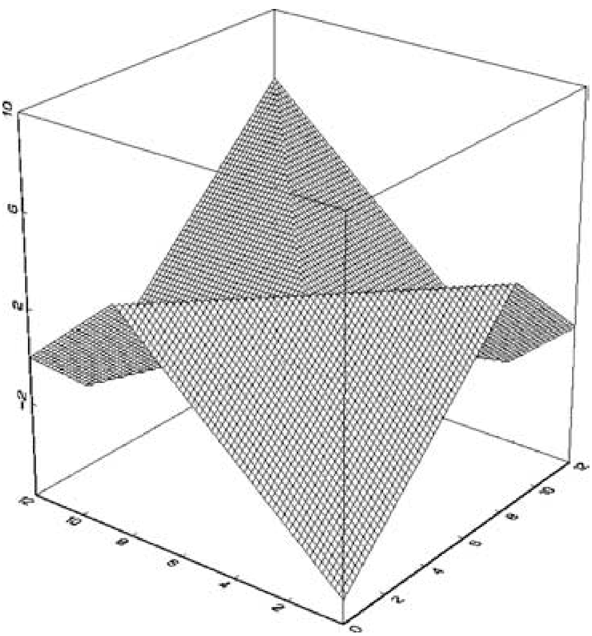

FIGURE 3. The graph of $f\colon \mathbf{R}^2 \to \mathbf{R}$.

https://doi.org/10.1017/S1365100506050103 Published online by Cambridge University Press

(see Theorem 4.15) and it states that the collection of all piecewise linear functions is precisely the Riesz subspace generated in $C(\mathbf{R}^m)$ by the affine functions.

LEMMA 4.4. *Every piecewise linear function is continuous.*

Proof. Let $f: \mathbf{R}^m \to \mathbf{R}$ be piecewise linear and let $x_n \to x$. If $f(x_n) \not\to f(x)$, then (by passing to a subsequence) we can assume without loss of generality that there exists some $\epsilon > 0$ such that $|f(x_n) - f(x)| \geq \epsilon$ for each n. Now notice that there exist some i and a subsequence $\{y_n\}$ of $\{x_n\}$ satisfying $y_n \in S_i$ for each n. But then we have $\epsilon \leq |f(y_n) - f(y)| = |f_i(y_n) - f_i(y)| \to 0$, which is impossible. This shows that f is continuous. ■

The following result presents an extremely simple characterization of piecewise linear functions.

THEOREM 4.5. *A continuous function $f: \mathbf{R}^m \to \mathbf{R}$ is piecewise linear if and only if there exist affine functions $f_1, \ldots, f_k$ such that for each $x \in \mathbf{R}^m$ there exists some $1 \leq i \leq k$ satisfying $f(x) = f_i(x)$.*

Moreover, the set of components of f is a subcollection of the collection of affine functions $\{f_1, \ldots, f_k\}$.

Proof. If f is piecewise linear, then the condition is trivially true. So, for the converse, assume that there exist affine functions $f_1, \ldots, f_k$ such that for each $x \in \mathbf{R}^m$ there exists some $1 \leq i \leq k$ such that $f(x) = f_i(x)$. We can assume that the affine functions $f_1, \ldots, f_k$ are distinct. We claim the following:

- *For each nonempty open subset V of $\mathbf{R}^m$ there exists a nonempty open subset W of V and some $1 \leq i \leq k$ such that $f = f_i$ on W.*

To see this, assume by way of contradiction that the claim is false. This implies that $f \neq f_1$ on V, that is, $f_1(v) \neq f(v)$ for some $v \in V$. Since f and f_1 are continuous, there exists some nonempty open subset V_1 of V such that $f_1(x) \neq f(x)$ for all $x \in V_1$. Similarly, since (by our hypothesis) $f \neq f_2$ on V_1 there exists some nonempty open subset V_2 of V_1 such that $f_2(x) \neq f(x)$ for all $x \in V_2$. Continuing this way, we see that there exist nonempty open sets $V_k \subseteq V_{k-1} \subseteq \cdots \subseteq V_1 \subseteq V$ such that for each $1 \leq i \leq k$ we have $f_i(x) \neq f(x)$ for all $x \in V_i$. But then for each $x \in V_k$ we have $f(x) \neq f_i(x)$ for all $1 \leq i \leq k$, which is impossible, and our claim has been established.

Now for each $1 \leq i \leq k$ let $\mathcal{O}_i = \bigcup\{U \subseteq \mathbf{R}^m\colon\ U \text{ is open and } f = f_i \text{ on } U\}$. That is, $\mathcal{O}_i$ is the largest open set on which $f = f_i$. By the preceding discussion $\mathcal{O}_i \neq \emptyset$ for at least one i. (To see this take $V = \mathbf{R}^m$ and apply (•).) Deleting the $\mathcal{O}_i$ with $\mathcal{O}_i = \emptyset$, we can assume that $\mathcal{O}_i \neq \emptyset$ for each i. Put $S_i = \overline{\mathcal{O}_i}$, and note that $f = f_i$ on S_i. We shall verify that the closed sets $S_1, \ldots, S_k$ satisfy the conditions of Definition 4.2. Start by observing that condition (4) is obvious.

For (1) note that from $\mathcal{O}_i \subseteq S_i$, we get that $\text{Int}(S_i) \neq \emptyset$ and that $\mathcal{O}_i \subseteq \text{Int}(S_i)$. Moreover, $\mathcal{O}_i = \text{Int}(S_i)$ must be the case, since otherwise the maximality property of $\mathcal{O}_i$ will be violated. The condition $\mathcal{O}_i \cap \mathcal{O}_j = \emptyset$ for $i \neq j$ should be obvious and the validity of (2) follows. If $\bigcup_{i=1}^{k} S_i \neq \mathbf{R}^m$, then by the above discussion there exists some nonempty open subset Q of $\mathbf{R}^m \setminus \bigcup_{i=1}^{k} S_i$ and some $1 \leq \ell \leq k$ such that $f = f_\ell$ on Q. But then the open set $\mathcal{O}_\ell \cup Q$ violates the maximality property of $\mathcal{O}_\ell$. Hence, $\bigcup_{i=1}^{k} S_i = \mathbf{R}^m$.

That the components of f are among the affine functions $f_1, \ldots, f_k$ should be obvious from the above discussion. ■

An immediate consequence of the preceding result is that **PL** is a Riesz subspace.

COROLLARY 4.6. *The collection of all piecewise linear functions on* $\mathbf{R}^m$ *is a Riesz subspace of* $C(\mathbf{R}^m)$. *In particular,* $\mathcal{R}(\mathbf{Aff}) = \mathbf{Aff}^{\vee\wedge} = \mathbf{Aff}^{\wedge\vee} \subseteq \mathbf{PL}$.

Recall that an **affine transformation** from $\mathbf{R}^k$ to $\mathbf{R}^m$ is any function $T\colon \mathbf{R}^k \to \mathbf{R}^m$ of the form $T(t) = At + b$, where A is an $m \times k$ real matrix and $b \in \mathbf{R}^m$ is a fixed vector. Now if T is an affine transformation and $f\colon \mathbf{R}^m \to \mathbf{R}$ is an affine function, then the function $f \circ T\colon \mathbf{R}^k \to \mathbf{R}$ is also an affine function. To see this, assume that f is defined as $f(x) = \alpha + u \cdot x$ and note that for each $t \in \mathbf{R}^k$ we have

$$[f \circ T](t) = f(T(t)) = \alpha + u \cdot (At + b) = (\alpha + u \cdot b) + (A'u) \cdot t.$$

This conclusion in connection with Theorem 4.5 yields the following result.

COROLLARY 4.7. *If* $f\colon \mathbf{R}^m \to \mathbf{R}$ *is an arbitrary piecewise linear function and* $T\colon \mathbf{R}^k \to \mathbf{R}^m$ *is an affine transformation, then the function* $f \circ T\colon \mathbf{R}^k \to \mathbf{R}$ *is piecewise linear. Moreover, if* f *has the components* $f_1, \ldots, f_p$, *then the components of* $f \circ T$ *are among the affine functions* $f_1 \circ T, \ldots, f_p \circ T$.

In particular, for any two fixed vectors $a, b \in \mathbf{R}^m$ *the function* $\theta\colon \mathbf{R} \to \mathbf{R}$, *defined via the formula* $\theta(t) = f(ta + (1-t)b)$, *is (one-dimensional) piecewise linear.*

A **hyperplane** of $\mathbf{R}^m$ is any subset of the form $H = \{x \in \mathbf{R}^m\colon a \cdot x = \alpha\}$, where $a \in \mathbf{R}^m$ is a nonzero vector and $\alpha \in \mathbf{R}$ is a constant. Clearly, every hyperplane is a closed set and has Lebesgue measure zero. Notice that two affine functions $f, g\colon \mathbf{R}^m \to \mathbf{R}$ either do not agree at any point or the set that they agree is a hyperplane, that is, the set $[f = g] = \{x \in \mathbf{R}^m\colon\ f(x) = g(x)\}$ is either empty or a hyperplane.

The boundaries of the regions of a piecewise linear function are parts of hyperplanes.

LEMMA 4.8. *Let* $(S_1, f_1), \ldots, (S_p, f_p)$ *be the characteristic pairs of a piecewise linear function* $f\colon \mathbf{R}^m \to \mathbf{R}$. *For each* i *let* $\mathcal{I}_i = \{j \in \{1, \ldots, p\}\colon\ j \neq i$ *and* $S_i \cap S_j \neq \emptyset\}$. *Then the boundary of the region* S_i *has the following property:*

$$\partial S_i = \bigcup_{j \in \mathcal{I}_i} S_i \cap S_j \subseteq \bigcup_{j \in \mathcal{I}_i} [f_i = f_j].$$

In particular,

(a) *each boundary* ∂S_i *has Lebesgue measure zero and consists of "parts" of hyperplanes*, *and*
(b) *if* $x \in \mathrm{Int}(S_i)$ *for some* i, *then* $x \notin S_j$ *for all* $j \neq i$.

Proof. Let $x \in \partial S_i$. Since $B(x, \frac{1}{n}) \cap (\mathbf{R}^m \backslash S_i) \neq \emptyset$, there exists for each n some $x_n \in \bigcup_{r \neq i} S_r$ such that $x_n \in B(x, \frac{1}{n})$. It follows that for some $j \neq i$ we have $x_n \in S_j$ for infinitely many n. This implies $x \in \overline{S_j} = S_j$ and so $x \in S_i \cap S_j$.

Now assume that $x \in S_i \cap S_j$ for some $j \neq i$. If $x \notin \partial S_i$, then $x \in \mathrm{Int}(S_i)$ and so there exists some $\delta > 0$ such that $B(x, \delta) \subseteq \mathrm{Int}(S_i)$. Since $\mathrm{Int}(S_i) \cap \mathrm{Int}(S_j) = \emptyset$, we infer that $x \in \partial S_j$. From $\overline{\mathrm{Int}(S_j)} = S_j$, it follows that there exists some $y \in \mathrm{Int}(S_j)$ such that $y \in B(x, \delta)$. This implies $y \in \mathrm{Int}(S_i) \cap \mathrm{Int}(S_j)$, which is impossible. Consequently, $x \in \partial S_i$, and the proof is finished. ■

The characteristic pairs of a piecewise linear function are uniquely determined.

LEMMA 4.9. *The regions and the components of an arbitrary piecewise linear function* $f\colon \mathbf{R}^m \to \mathbf{R}$ *are uniquely determined in the following sense*: *If another collection of pairs* $\{(S'_1, g_1), \ldots, (S'_q, g_q)\}$ *satisfies properties* (1)–(4) *of Definition* 4.2, *then* $q = p$ *and* $\{(S'_1, g_1), (S'_2, g_2), \ldots, (S'_q, g_q)\}$ *is a permutation of the collection of pairs* $\{(S_1, f_1), (S_2, f_2), \ldots, (S_p, f_p)\}$.

Proof. Fix some $1 \leq i \leq p$. Because $\mathrm{Int}(S_i)$ is nonempty (and hence it has positive Lebesgue measure), it follows from Lemma 4.8 that there exists some $1 \leq j \leq q$ such that the open set $V = \mathrm{Int}(S_i) \cap \mathrm{Int}(S'_j)$ is nonempty. In particular, as $f_i(x) = g_j(x) = f(x)$ holds true for each $x \in V$, it follows from part (2) of Lemma 4.1 that $f_i = g_j$.

Now let $x \in \mathrm{Int}(S_i)$. Fix $\delta > 0$ such that $B(x, \delta) \subseteq \mathrm{Int}(S_i)$ and let $0 < \epsilon < \delta$. As above, $B(x, \epsilon) \cap \mathrm{Int}(S'_r) \neq \emptyset$ must hold true for some index $1 \leq r \leq q$. But then (as above again) $g_j = f_i = g_r$ must be the case. Because the affine functions $g_1, \ldots, g_q$ are all distinct, we infer that $r = j$. Therefore, $B(x, \epsilon) \cap \mathrm{Int}(S'_j) \neq \emptyset$ for all $0 < \epsilon < \delta$. This implies $x \in \overline{S'_j} = S'_j$, and so $\mathrm{Int}(S_i) \subseteq S'_j$. Consequently, $S_i = \overline{\mathrm{Int}(S_i)} \subseteq S'_j$.

By the symmetry of the situation, there exists some $1 \leq m \leq p$ such that $S'_j \subseteq S_m$. This implies $\mathrm{Int}(S_i) \cap \mathrm{Int}(S_m) = \mathrm{Int}(S_i) \neq \emptyset$, from which it follows that $m = i$. Therefore, $S_i = S'_j$ and so $(S_i, f_i) = (S'_j, g_j)$. From the last result, the desired conclusion now easily follows. ■

Another consequence of Theorem 4.5 is that for real functions defined on $\mathbf{R}$ the definitions for piecewise linear functions given in Definitions 3.1 and 4.2 are equivalent.

COROLLARY 4.10. *A function* $f\colon \mathbf{R} \to \mathbf{R}$ *is piecewise linear according to Definition* 3.1 *if and only if it is piecewise linear according to Definition* 4.2.

Proof. Let $f: \mathbf{R} \to \mathbf{R}$ be a function. If f is piecewise linear according to Definition 3.1, then f is clearly piecewise linear according to Definition 4.2.

For the converse, assume that f is piecewise linear according to Definition 4.2. Let $\{(S_1, f_1), (S_2, f_2), \ldots, (S_p, f_p)\}$ be the collection of characteristic pairs of f. Notice that every f_i is of the form $f_i(t) = m_i t + b_i$. So, every nonempty set of the form $[f_i = f_j]$ is simply a point of $\mathbf{R}$. This is connection with Lemma 4.8 shows that the boundary of each S_i is a finite set. Now each $\mathrm{Int}(S_i)$ is the union of an at most countable collection of pairwise disjoint open intervals. Because ∂S_i is a finite set, a moment's thought reveals that $\mathrm{Int}(S_i)$ is a union of a finite number of pairwise disjoint open intervals. From this it follows that S_i is the union of the closures of these intervals. Now it is easy to see that f is a piecewise linear function according to Definition 3.1. ■

In order to further study piecewise linear functions, we shall need the theory of arrangements of hyperplanes, which are well-studied combinatorial constructions that are closely related to vector lattices and the simplex methods in linear programming; see Chapter 4 of Björner, Las Vergnas, Sturmfels, White, and Ziegler (1999).

Recall once more that any subset of $\mathbf{R}^m$ of the form $H = \{x \in \mathbf{R}^m\colon a \cdot x = \alpha\}$, where $a \in \mathbf{R}^m$ is a nonzero fixed vector and $\alpha \in \mathbf{R}$ is a constant, is called a **hyperplane** of $\mathbf{R}^m$. We can assume without loss of generality that $\|a\| = 1$ and refer to a as a (unit) vector **normal** to H. Since $H = \{x \in \mathbf{R}^m\colon (-a) \cdot x = -\alpha\}$, we see that $-a$ is also another (unit) normal vector to H. In other words, H has essentially two unit normal vectors, each of which defines an **orientation** in the sense that it divides $\mathbf{R}^m$ into three parts: a "positive" part $\{x \in \mathbf{R}^m\colon a \cdot x > \alpha\}$, a "zero" part $\{x \in \mathbf{R}^m\colon a \cdot x = \alpha\}$, and a "negative" part $\{x \in \mathbf{R}^m\colon a \cdot x < \alpha\}$. Of course, if we let $H = \{x \in \mathbf{R}^m\colon (-a) \cdot x = -\alpha\}$, then the orientation changes: the positive part is now negative and the negative part is positive. Thus, writing H in the form $H = \{x \in \mathbf{R}^m\colon a \cdot x = \alpha\}$, the vector a defines automatically an orientation, and H is called an **oriented hyperplane**.

Now let E be a finite index set and let $(H_e)_{e \in E}$, where $H_e = \{x \in \mathbf{R}^m\colon a_e \cdot x = \alpha_e\}$, be a family of (oriented) hyperplanes in $\mathbf{R}^m$. The family $(H_e)_{e \in E}$, is called an ***oriented arrangement of hyperplanes*** (or simply an ***arrangement***). Every arrangement of hyperplanes $(H_e)_{e \in E}$ "almost" subdivides $\mathbf{R}^m$ into a finite number of nonempty convex regions. The subdivisions are obtained by means of the "sign" mapping $x \mapsto \sigma_x$, from $\mathbf{R}^m$ to $\{+, -, 0\}^E$, that is defined by

$$\sigma_x(e) = \begin{cases} + & \text{if } a_e \cdot x > \alpha_e, \\ - & \text{if } a_e \cdot x < \alpha_e, \\ 0 & \text{if } a_e \cdot x = \alpha_e, \end{cases}$$

that is, $\sigma_x = (\mathrm{Sign}(a_e \cdot x - \alpha_e))_{e \in E}$. Let $\mathcal{M}$ denote the range of the function σ, that is, $\mathcal{M} = \sigma(\mathbf{R}^m) \subseteq \{+, -, 0\}^E$.

A vector $T \in \mathcal{M}$ satisfying $T(e) \neq 0$ for all $e \in E$ is called a ***tope*** of $\mathcal{M}$. Note that σ_x is a tope if and only if $x \notin \bigcup_{e \in E} H_e$. Let $T_1, T_2 \ldots, T_J$ be an enumeration

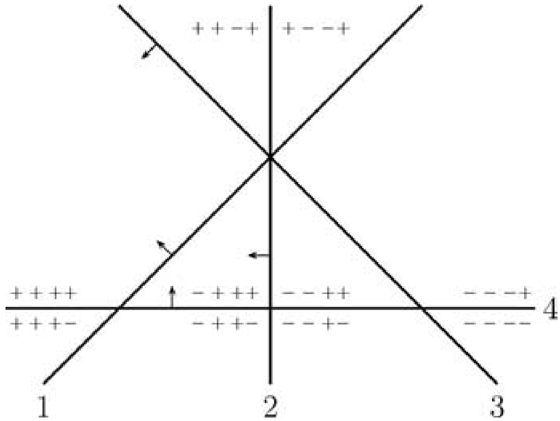

FIGURE 4. An arrangement of four oriented hyperplanes in $\mathbf{R}^2$.

of the topes of $\mathcal{M}$. For each $1 \le h \le J$ let

$$K_h = \{x \in \mathbf{R}^m \colon \sigma_x = T_h\} = \sigma^{-1}(\{T_h\}).$$

Obviously, each K_h is a nonempty, open, and convex set. Moreover, from the identity $\bigcup_{h=1}^{J} K_h = \mathbf{R}^m \setminus \bigcup_{e \in E} H_e$, we see that $\overline{\bigcup_{h=1}^{J} K_h} = \mathbf{R}^m$. The sets K_1, $K_2, \dots, K_J$ are called the ***cells*** induced by the arrangement of the hyperplanes $(H_e)_{e \in E}$. It should not be difficult to see that the collection of cells $\{K_1, K_2, \dots, K_J\}$ is independent of the orientation of the planes H_e, and so we can refer to $\{K_1, K_2, \dots, K_J\}$ as the **collection of cells generated** (or **induced**) by the family of hyperplanes $(H_e)_{e \in E}$. For an example of an arrangement of hyperplanes, see Figure 4.

Now let $\{f_1, \dots, f_p\}$, where $p \ge 2$, be a collection of distinct affine functions on $\mathbf{R}^m$. If for each $1 \le i < j \le p$ we let $H_{i,j} = [f_i = f_j]$, then the set

$$E = \{(i, j) \colon 1 \le i < j \le p \text{ and } H_{i,j} \ne \emptyset\}$$

is a finite set. Letting $H_e = [f_i = f_j] = \{x \in \mathbf{R}^m \colon a_e \cdot x = \alpha_e\}$ for each $e = (i, j) \in E$, we see that the family $(H_e)_{e \in E}$ is an arrangement of hyperplanes, called **an arrangement generated** by $\{f_1, \dots, f_p\}$. The collection of cells generated by $(H_e)_{e \in E}$ is called the **collection of cells generated** (or **induced**) by $\{f_1, \dots, f_p\}$.

With this terminology at hand, we are now ready to state several extra properties of piecewise linear functions.

LEMMA 4.11. *Let $F = \{f_1, \dots, f_k\}$ be a finite collection of distinct affine functions of $\mathbf{R}^m$ and let $\{K_1, K_2, \dots, K_J\}$ be the collection of cells induced by F. Assume also that $f \colon \mathbf{R}^m \to \mathbf{R}$ is a continuous function such that for each $x \in \mathbf{R}^m$ there exists some $1 \le i \le k$ satisfying $f(x) = f_i(x)$.*[4] *Then for a vector $x \in K_h$ we have the following*:

https://doi.org/10.1017/S1365100506050103 Published online by Cambridge University Press

(1) *If* $f(x) = f_i(x)$, *then* $f(y) = f_i(y)$ *for all* $y \in K_h$.
(2) *If* $f(x) > f_i(x)$, *then* $f(y) > f_i(y)$ *for all* $y \in K_h$.
(3) *If* $f(x) < f_i(x)$, *then* $f(y) < f_i(y)$ *for all* $y \in K_h$.

Moreover, for each $1 \le h \le J$ *there is a unique* $1 \le i_h \le k$ *with* $f = f_{i_h}$ *on* K_h.

Proof. We shall prove (1) first. To this end, suppose that some $x \in K_h$ satisfies $f(x) = f_i(x)$.

Let $X = \bigcup_{h=1}^{J} K_h$ and note that X is an open dense subset of $\mathbf{R}^m$. Notice that for each $z \in X$ any pair of distinct functions $f_i, f_j \in F$ we have $f_i(z) \neq f_j(z)$. So for each $z \in X$ there exists a unique $1 \le i_z \le k$ such that $f(z) = f_{i_z}(z)$. Because f and the f_i are continuous functions and $f(z) = f_{i_z}(z) \neq f_j(z)$ for each $j \neq i_z$, there exists an open neighborhood $N_z \subseteq X$ of z such that for each $y \in N_z$ and all $j \neq i_z$ we have $f(y) \neq f_j(y)$ and $f_{i_z}(y) \neq f_j(y)$. This implies that for each $y \in N_z$ we have $f(y) = f_{i_z}(y)$, that is, $i_y = i_z$.

Now fix $y \in K_h$. Let $L(x, y)$ be the line segment joining x and y and notice that $L(x, y) \subseteq K_h$ as K_h is convex. Because $L(x, y)$ is compact, there exists a finite set $Z = \{z_1, \dots, z_r\} \subseteq L(x, y)$ such that $L(x, y) \subseteq \bigcup_{z \in Z} N_z$. We can assume that the neighborhoods $\{N_z : z \in Z\}$ form a chain, that is, $N_{z_t} \cap N_{z_{t+1}} \neq \emptyset$ for each $t = 1, \dots, r-1$; see (Abramovich and Aliprantis, 2002b, Problem 1.5.7, p. 50). This easily implies that for each $z \in L(x, y)$ we have $f(z) = f_{i_x}(z) = f_{i_y}(z)$. In particular, $i_x = i_y$.

Therefore, we have shown that for each K_h there exists a unique index $1 \le i_h \le k$ such that $y \in K_h$ implies $f(y) = f_{i_h}(y)$. This proves (1) and the last part of the lemma.

To establish (2), assume that $f(x) > f_i(x)$ holds true for some $x \in K_h$ and that some other $y \in K_h$ satisfies $f(y) \le f_i(y)$. If $f(y) = f_i(y)$, then according to (1) we must have $f(x) = f_i(x)$, which is impossible. If $f(y) < f_i(y)$, then there exists some z in the line segment joining x and y (and hence $z \in K_h$) satisfying $f(z) = f_i(z)$. But then (according to (1) again) we get $f(x) = f_i(x)$, a contradiction. This establishes (2) and the validity of (3) can be proven in a similar fashion. ∎

From Theorem 4.5 we know that if for a continuous function $f : \mathbf{R}^m \to \mathbf{R}$ and affine functions $f_1, \dots, f_k$ for each $x \in \mathbf{R}^m$ there exists some $1 \le i \le k$ satisfying $f(x) = f_i(x)$, then f is piecewise linear. The next result constructs the characteristic pairs of such a piecewise linear function from a given collection of affine functions.

THEOREM 4.12. *Assume that a continuous function* $f : \mathbf{R}^m \to \mathbf{R}$ *and a finite set of distinct affine functions* $F = \{f_1, \dots, f_k\}$ *are such that for each* $x \in \mathbf{R}^m$ *there exists some* $1 \le i \le k$ *satisfying* $f(x) = f_i(x)$. *Let* $\{K_1, K_2, \dots, K_J\}$ *be the cells generated by* F. *For each* $1 \le i \le k$ *let*

$$E_i = \{h \in \{1, \dots, J\} : f = f_i \text{ on } K_h\},$$

https://doi.org/10.1017/S1365100506050103 Published online by Cambridge University Press

and then define $S_i = \overline{\bigcup_{h \in E_i} K_h}$. *We have the following.*

(a) *If* $\{E_i\}_{i \in \mathcal{I}}$ *is the family of nonempty* E_i, *then the family* $\{(S_i, f_i)\}_{i \in \mathcal{I}}$ *is precisely the family of characteristic pairs of the piecewise linear function* f.
(b) *For each* $1 \le h \le J$ *there exists exactly one* $i \in \mathcal{I}$ *such that* $K_h \subseteq \mathrm{Int}(S_i)$.
(c) *For each* $i \in \mathcal{I}$ *the nonempty set* $\mathrm{Int}(S_i)$ *is a union of a finite collection of pairwise disjoint nonempty open and connected subsets of* $\mathbf{R}^m$.

Proof. (a) We know from Theorem 4.5 that the function f is piecewise linear whose components are among the $f_1, \dots, f_k$. The proof here will present also an alternate constructive proof of Theorem 4.5. Let $\{K_1, \dots, K_J\}$ be the collection of cells generated by F and for each $1 \le i \le k$ define E_i and S_i as in the statement of the lemma.

According to Lemma 4.11 at least one of the E_i is nonempty; relabeling, we can assume that $E_1, \dots, E_p$ are the nonempty E_i, that is, $\mathcal{I} = \{1, \dots, p\}$. Clearly, $f = f_i$ on S_i. Because the affine functions $f_1, \dots, f_k$ are distinct, it follows from part (2) of Lemma 4.1 that $E_r \cap E_s = \emptyset$ for $r \neq s$ and from Lemma 4.11 we see that $\bigcup_{i=1}^p E_i = \{1, \dots, J\}$. The latter yields

$$\bigcup_{i=1}^p S_i = \bigcup_{i=1}^p \overline{\bigcup_{h \in E_i} K_h} = \overline{\bigcup_{i=1}^p \bigcup_{h \in E_i} K_h} = \overline{\bigcup_{h=1}^J K_h} = \mathbf{R}^m.$$

Next notice that because for each $1 \le i \le p$ we have $\bigcup_{h \in E_i} K_h \subseteq \mathrm{Int}(S_i)$, it follows, on one hand, that $\mathrm{Int}(S_i) \neq \emptyset$ and, on the other hand, that $\overline{\mathrm{Int}(S_i)} = S_i$. Moreover, using part (2) of Lemma 4.1, it is easy to see that $\mathrm{Int}(S_r) \cap \mathrm{Int}(S_s) = \emptyset$ for $r \neq s$. Because $f = f_i$ holds true for each $1 \le i \le p$, it follows from Definition 4.2 that f is a piecewise linear function with characteristic pairs $(S_1, f_1), \dots, (S_p, f_p)$.

(b) Now let $1 \le h \le J$. According to Lemma 4.11 there exists a unique $1 \le i_h \le k$ such that $f = f_{i_h}$ on K_h. This implies that $E_{i_h} \neq \emptyset$ and $K_h \subseteq \mathrm{Int}(S_{i_h})$.

(c) Observe that for each $i \in \mathcal{I}$ every component of $\mathrm{Int}(S_i)$, that is, every maximal (with respect to $\supseteq$) nonempty and connected subset of $\mathrm{Int}(S_i)$, is open. Now notice that every $K_h \subseteq \mathrm{Int}(S_i)$ is open and connected (as being a convex set) and so is included in some component of $\mathrm{Int}(S_i)$. Moreover, from the definition of S_i, it is not difficult to see that every component of $\mathrm{Int}(S_i)$ includes some K_h. Thus, the number of components of $\mathrm{Int}(S_i)$ is at most J, and the proof is finished. ■

To continue our study, we need one more property of piecewise linear functions.

LEMMA 4.13 [Ovchinnikov (2002)]. *If* $f: \mathbf{R}^m \to \mathbf{R}$ *is a piecewise linear function with components* $f_1, \dots, f_p$, *then for any pair* $a, b \in \mathbf{R}^m$ *there exists a component* f_i *of* f *satisfying* $f_i(a) \le f(a)$ *and* $f_i(b) \ge f(b)$.

Proof. Fix $a, b \in \mathbf{R}^m$ and consider the continuous function $g: \mathbf{R} \to \mathbf{R}$ defined via the formula by $g(t) = f[tb + (1 - t)a]$. By Corollary 4.7, g is a one-dimensional piecewise linear function whose components are among the affine functions $g_1, g_2, \ldots, g_p$, where $g_i(t) = f_i(tb + (1-t)a)$. Consider g restricted to [0, 1] and then use Lemma 3.2 in conjunction with Corollary 3.7 to see that there exists a component g_i satisfying $f(a) = g(0) \geq g_i(0) = f_i(a)$ and $f(b) = g(1) \leq g_i(1) = f_i(b)$. ■

The next result presents the basic structural properties of piecewise linear functions. Its proof is based on the discussion by Ovchinnikov on the referees' comments concerning his paper Ovchinnikov (2002).

THEOREM 4.14. *Assume that $f: \mathbf{R}^m \to \mathbf{R}$ is a piecewise linear function with characteristic pairs $\{(S_1, f_1), \ldots, (S_p, f_p)\}$ and let $\{K_1, K_2, \ldots, K_J\}$ be the set of cells induced by $\{f_1, \ldots, f_p\}$.*

(1) *If for each h we pick $x_h \in K_h$ and let $E_h = \{i \in \{1, \ldots, p\}: f_i(x_h) \geq f(x_h)\}$, then E_h is nonempty and*

$$f = \bigvee_{h=1}^{J} \bigwedge_{i \in E_h} f_i.$$

In particular, $f \in \{f_1, f_2, \ldots, f_p\}^{\vee\wedge}$.

(2) *If J^* is the subset of $\{1, \ldots, J\}$ having the property that for each $1 \leq h \leq J$ there exists a $j \in J^*$ such that $E_j \subseteq E_h$, then we have*

$$f = \bigvee_{j \in J^*} \bigwedge_{i \in E_j} f_i.$$

Proof. (1) For each $1 \leq h \leq J$ fix some $x_h \in K_h$ and then use Theorem 4.12 to choose some $1 \leq j \leq p$ such that $K_h \subseteq \operatorname{Int}(S_j)$. Clearly, $f_j(x_h) = f(x_h)$. This implies that if for each $1 \leq h \leq J$ we let

$$E_h = \{i \in \{1, \ldots, p\}: f_i(x_h) \geq f(x_h)\},$$

then, on the one hand, $E_h \neq \emptyset$ and, on the other hand, a glance at Lemma 4.11 guarantees that for each $i \in E_h$ and each $y \in K_h$ we have $f_i(y) \geq f(y)$. Now for each $1 \leq h \leq J$ consider the function

$$F_h = \bigwedge_{i \in E_h} f_i, \tag{$\star$}$$

and note that $F_h(y) \geq f(y)$ for each $y \in K_h$. Because for some $j \in E_h$ we have $f_j(x_h) = f(x_h)$, it follows from Lemma 4.11 that $f_j(y) = f(y)$ for all $y \in K_h$. Thus, $F_h(y) = f(y)$ for all $y \in K_h$.

Next, fix $y \in \mathbf{R}^m$. For each $1 \leq h \leq J$ there exists (according to Lemma 4.13) some f_j satisfying $f_j(y) \leq f(y)$ and $f_j(x_h) \geq f(x_h)$. In particular, it follows that we have $j \in E_h$ and consequently $F_h(y) = [\bigwedge_{i \in E_h} f_i](y) \leq f(y)$ for all $1 \leq h \leq J$. This implies $[\bigvee_{h=1}^{J} F_h](y) \leq f(y)$ for all $y \in \mathbf{R}^m$.

On the other hand, because for each $x \in K_h$ we have $F_h(x) = f(x)$, it must be the case that

$$\bigvee_{h=1}^{J} F_h = \bigvee_{h=1}^{J} \bigwedge_{i \in E_h} f_i = f,$$

on $\bigcup_{h=1}^{J} K_h$. Because $\bigcup_{h=1}^{J} K_h$ is dense in $\mathbf{R}^m$ and $\bigvee_{h=1}^{J} \bigwedge_{i \in E_h} f_i$ and f are both continuous functions, it follows that $\bigvee_{h=1}^{J} \bigwedge_{i \in E_h} f_i = f$ holds true on $\mathbf{R}^m$.

(2) To establish this identity, note first that if $E_j \subseteq E_h$, then $\bigwedge_{i \in E_h} f_i \le \bigwedge_{i \in E_j} f_i \le f$. This implies $\bigwedge_{i \in E_h} f_i \le \bigvee_{j \in J^*} \bigwedge_{i \in E_j} f_i \le f$ for each $1 \le h \le J$, and consequently we have $f = \bigvee_{h=1}^{J} \bigwedge_{i \in E_h} f_i \le \bigvee_{j \in J^*} \bigwedge_{i \in E_j} f_i \le f$, and the proof is finished. ■

Combining Corollary 4.6 and Theorem 4.14 we are now ready to state the fundamental result for this work.

THEOREM 4.15. *The vector space* **PL** *of all piecewise linear functions is a vector sublattice of the Riesz space* $C(\mathbf{R}^m)$ *and coincides with* $\mathbf{Aff}^{\vee\wedge}$, *that is,* $\mathbf{PL} = \mathbf{Aff}^{\vee\wedge}$.

In other words, **PL** *is precisely the Riesz subspace of* $C(\mathbf{R}^m)$ *generated by the* $(m+1)$*-dimensional vector subspace* **Aff** *of all affine functions.*

The next example reported in Ovchinnikov (2002) shows that piecewise polynomial functions need not admit a sup-inf representation.

Example 4.16. Define the piecewise quadratic function $f: \mathbf{R} \to \mathbf{R}$ as follows:

$$f(x) = \begin{cases} 0 & \text{if } x \le 0, \\ x^2 & \text{if } x > 0. \end{cases}$$

Notice that $x^2 \vee 0 = x^2$ and $x^2 \wedge 0 = 0$. ■

Noting that the set $\{f_1, f_2, \ldots, f_p\}^{\vee\wedge}$ is finite, Theorem 4.15 yields also the following.

COROLLARY 4.17. *If* $F = \{f_1, f_2, \ldots, f_p\}$ *is a finite set of affine functions on* $\mathbf{R}^m$, *then a function* $f \in C(\mathbf{R}^m)$ *is piecewise linear with components in* F *if and only if* f *belongs to the finite set* $F^{\vee\wedge}$.

Theorem 4.14 also provides an algorithm for constructing the sup-inf representation of a piecewise linear function with components $f_1, f_2, \ldots, f_p$ and unknown regions. The next example is a rudimentary algorithm illustrating this.

Example 4.18 (From $f, f_1, f_2, \ldots, f_p$ to $\mathbf{Aff}^{\vee\wedge}$). Take $f \in \mathbf{PL}$ with components $f_1, f_2, \ldots, f_p$. Following Theorem 4.14 the function f can be reconstructed using the following step:

Step I: Determine $E = \{(i, j)\colon 1 \le i < j \le p \text{ and } [f_i = f_j] \ne \emptyset\}$ and then for each $e = (i, j) \in E$ pick $\alpha_e \in \mathbf{R}$ and $a_e \in \mathbf{R}^m$ such that

$$H_e = \{x \in \mathbf{R}^m\colon a_e \cdot x = \alpha_e\} = [f_i = f_j].$$

Step II: Using the hyperplane arrangement $(H_e)_{e \in E}$ determine the cells $K_1, \dots, K_J$.

Step III: For each $h = 1, 2, \dots, J$ choose some x_h from the cell K_h.

Step IV: For each $1 \le h \le J$ determine $E_h = \{i \in \{1, \dots, p\} f_i(x_h) \ge f(x_h)\}$.

Step V: Select a "minimal" set $J^* \subseteq \{1, 2, \dots, J\}$ so that it satisfies property (2) of Theorem 4.14. Then we have

$$f = \bigvee_{j \in J^*} \bigwedge_{i \in E_j} f_i.$$

This procedure gives a desired sup-inf representation of f. ■

The next example illustrates the preceding algorithm. It also shows how in applying this algorithm, we can restrict our attention to a closed convex domain with nonempty interior.

Example 4.19. Consider once again Example 4.3 but with the restricted domain shown in Figure 5. Take the four affine components of the function f:

$$f_1(x_1, x_2) = x_1 - 5,$$
$$f_2(x_1, x_2) = x_2 - 5,$$
$$f_3(x_1, x_2) = x_1 + x_2 - 5,$$
$$f_4(x_1, x_2) = -x_1 - x_2 + 12.$$

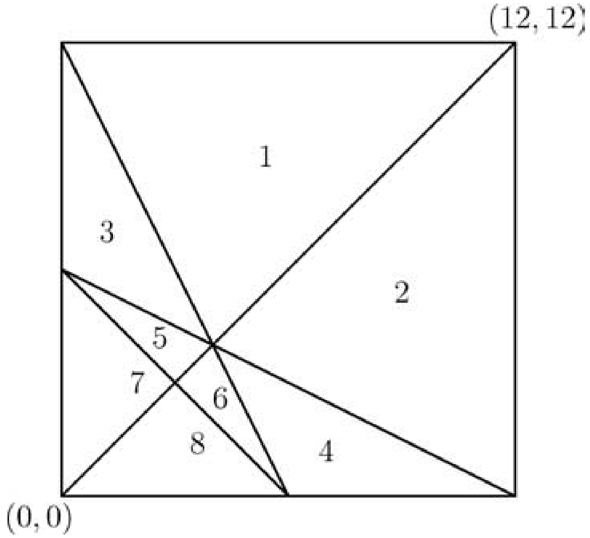

FIGURE 5. The eight regions of the oriented arrangement.

https://doi.org/10.1017/S1365100506050103 Published online by Cambridge University Press

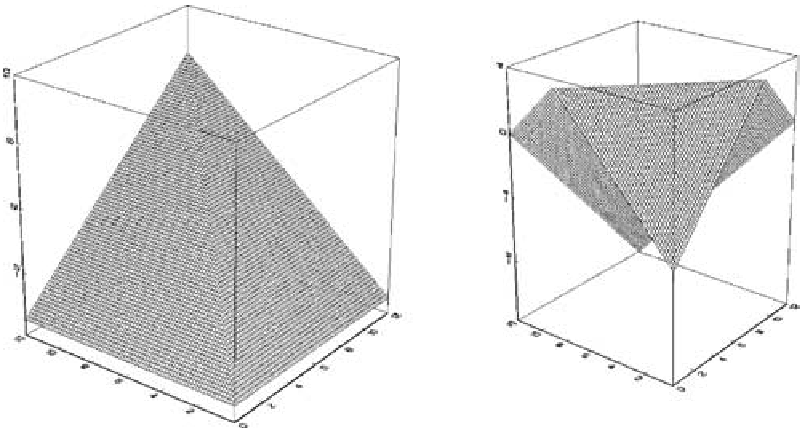

FIGURE 6. The graphs of $f_1 \wedge f_2$ and $f_3 \wedge f_4$.

These four affine functions induce eight cells. They are the eight regions of the oriented arrangement in Step I of the algorithm of Example 4.18 and they are depicted in Figure 5.

Notice that $E_1 = \{1, 2, 3\}$, $E_2 = \{1, 2, 3\}$, $E_3 = E_4 = E_5 = E_6 = E_7 = E_8 = \{3, 4\}$. Therefore, if we take $J^* = \{1, 3\}$, then we can write

$$f = (f_3 \wedge f_4) \vee (f_1 \wedge f_2 \wedge f_3).$$

Because we have restricted the domain, we can now write $f = (f_3 \wedge f_4) \vee (f_1 \wedge f_2)$; compare Figures 3 and 6. ■

A rudimentary algorithm for computing the regions of the functions in $\mathbf{Aff}^{\vee\wedge}$ by means of Theorem 4.12 is presented next.

Example 4.20 (From $\mathbf{Aff}^{\vee\wedge}$ to PL). Take $f \in \{f_1, f_2, \ldots, f_p\}^{\vee\wedge}$, where as usual $f_1, f_2, \ldots, f_p$ are affine functions on $\mathbf{R}^m$. Following Theorem 4.12 the regions of the function f can be obtained using the following steps:

Step I: Determine $E = \{(i, j)\colon 1 \le i < j \le p \text{ and } [f_i = f_j] \ne \emptyset\}$ and for each $e = (i, j) \in E$ and then pick $\alpha_e \in \mathbf{R}$ and $a_e \in \mathbf{R}^m$ such that $H_e = \{x \in \mathbf{R}^m\colon a_e \cdot x = \alpha_e\} = [f_i = f_j]$.

Step II: Use the hyperplane arrangement $(H_e)_{e \in E}$ to determine the cells $K_1, \ldots, K_J$.

Step III: For each $h = 1, 2, \ldots, J$ choose some $x_h \in K_h$ and then let

$$i_h = \min\{i \in \{1, \ldots, p\}\colon f_i(x_h) = f(x_h)\}.$$

Step IV: For each $h = 1, 2, \ldots, J$ determine the set $I_h = \{j \in \{1, \ldots, J\}\colon i_j = i_h\}$.

Step V: For each $h = 1, 2, \ldots, J$ let $S_{i_h} = \overline{\bigcup_{i \in I_h} K_j}$.

The characteristic pairs of the piecewise linear function f are distinct members of the family $\{(f_{i_h}, S_{i_h})\}_{h \in \{1, \ldots, J\}}$. ■

NOTES

1. Any norm on a Riesz space such that $|x| \leq |y|$ implies $\|x\| \leq \|y\|$ is called a **lattice** (or a **Riesz**) **norm**.

2. As a matter of fact, if we identify every vector $\mathbf{r} = (r_0, r_1, \ldots, r_m) \in \mathbf{R}^{m+1}$ with the affine function on $\mathbf{R}^m$ defined by $\mathbf{r}(x) = r_0 + r_1 x_1 + \cdots + r_m x_m$, then it is not difficult to see that we can identify **Aff** with the vector space $\mathbf{R}^{m+1}$.

3. If A is any subset of $\mathbf{R}^m$, then $\mathrm{Int}(A)$ denotes its interior and $\bar{A}$ its closure. We remark that the sets S_i are not assumed to be connected.

4. Keep in mind that this implies (by Theorem 4.5) that f is piecewise linear.

REFERENCES

Y.A. Abramovich and C.D. Aliprantis (2002a) An invitation to operator theory, vol. 50 of *Graduate Studies in Mathematics*. Providence, RI: American Mathematical Society.

Y.A. Abramovich and C.D. Aliprantis (2002b) Problems in operator theory, vol. 51 of *Graduate Studies in Mathematics*. Providence, RI: American Mathematical Society.

C.D. Aliprantis and K.C. Border (1999) *Infinite-dimensional analysis: a hitchhiker's guide*. Berlin: Springer-Verlag.

C.D. Aliprantis and O. Burkinshaw (2003) Locally solid Riesz spaces with applications to economics, vol. 105 of *Mathematical Surveys and Monographs*. Providence, RI: American Mathematical Society.

C.D. Aliprantis, D. Harris, and R. Tourky (2006) Riesz estimators. *Journal of Econometrics*, Forthcoming.

A. Björner, M. Las Vergnas, B. Sturmfels, N. White, and G.M. Ziegler (1999) Oriented matroids, vol. 46 of *Encyclopedia of Mathematics and its Applications*. Cambridge: Cambridge University Press.

D.J. Brown, C.B. Huijsmans, and B. de Pagter (1991) Approximating derivative securities in f-algebras, in C.D. Aliprantis, K.C. Border, and W.A.J. Luxemburg, eds., *Positive operators, Riesz Spaces, and Economics*, Studies in Economic Theory, Vol. 2, pp. 171–177. Springer-Verlag, New York and Heidelberg.

W.A.J. Luxemburg and A. C. Zaanen (1971) *Riesz spaces, Vol. I*. Amsterdam: North-Holland Publishing.

S. Ovchinnikov (2002) Max-min representation of piecewise linear functions. *Beirtrage zur Algebra und Geometrie* 43(1), 297–302.

H.H. Schaefer (1974) *Banach lattices and positive operators*, vol. 215 of Die Grundlehren der mathematischen Wissenschaften, New York: Springer-Verlag. Band 215.

https://doi.org/10.1017/S1365100506050103 Published online by Cambridge University Press