IV. Mathematical characterizations of order functions

In this Section we prove several mathematical characterizations of order functions (and sometimes also of preorder functions) and describe some of their consequences. As we will see in Subsection IV.1, order functions can be characterized as: increasing preorder functions, continuous selection functions, functions commuting with thresholding, and functions commuting with increasing functions in one variable. These results are commented in Subsection IV.2, and their consequences for order filters will be discussed in Section VI.

IV.1. Characterization theorems

We showed in Theorem 6 that for $B = \{0,1\}$, a function $f_B : B^n \to B$ is an order function iff it is a non-constant increasing function. We will see how it is possible to extend this result to functions on D^n. But we must first introduce one more definition.

Given two finite sets C, C' having the same size, there is a unique strictly increasing function $\psi : C \to C'$, which is in fact a bijection. Then, given two functions $g_C : C^n \to C$ and $h_{C'} : C'^n \to C'$, we will say that $h_{C'}$ is *order-isomorphic* to g_C if for any $y, x_1, \ldots, x_n \in C$, $y = g_C(x_1, \ldots, x_n)$ implies that $\psi(y) = h_{C'}(\psi(x_1), \ldots, \psi(x_n))$, in other words if for any $x_1, \ldots, x_n \in C$,

$$\psi(g_C(x_1, \ldots, x_n)) = h_{C'}(\psi(x_1), \ldots, \psi(x_n)).$$

This means that g_C and $h_{C'}$ have the same behavior w.r.t. ordering.

It is easy to see that the relation "is order-isomorphic to" is an *equivalence relation*, *i.e.*, it is reflexive, symmetric, and transitive. Moreover, if g_C is an order function and if $h_{C'}$ is order-isomorphic to g_C, then $h_{C'}$ is an order function having the same choice map χ as g_C, in other words $g_C = f_C$ and $h_{C'} = f_{C'}$ for some order function f.

Now we can state our result:

Theorem 13. *Let f_D be a function $D^n \to \mathbf{R}$. Then f_D is an order function iff f_D is increasing and satisfies the following two conditions:*

- (i) *Given a subset C of size 1 or 2 of D, the restriction f_C of f_D to C^n is a selection function $C^n \to C$.*
- (ii) *Given two subsets C, C' of size 2 of D, the restrictions f_C and $f_{C'}$ of f_D to C^n and C'^n are order-isomorphic.*

Moreover, conditions (i) and (ii) are satisfied when f_D is a preorder function. Thus f_D is an order function iff it is an increasing preorder function.

Proof. (a) If f_D is a preorder function, then it is a selection function, and so for any $C \subseteq D$, its restriction to C^n is a selection function $C^n \to C$. Thus f_D satisfies (i). Given two subsets C, C' of size 2 of D, let ψ be the unique strictly increasing bijection $C \to C'$. For any $x_1, \ldots, x_n \in C$ and $i, j \in I_n$, we have $x_i < x_j$ iff $\psi(x_i) < \psi(x_j)$, and so by Definition 2 we must have $f_D(\psi(x_1), \ldots, \psi(x_n)) = \psi(f_D(x_1, \ldots, x_n))$. Thus $f_{C'}$ is order-isomorphic to f_C, and so f_D satisfies (ii).

(b) Suppose that f_D is an order function. Then f_D satisfies (i) and (ii), since it is a preorder function. Now f_D is increasing thanks to Corollary 7 and to the fact that the functions min and max are increasing.

(c) Suppose now that f_D is an increasing function $D^n \to \mathbf{R}$ satisfying (i) and (ii). Take a subset C of size 2 of D, and let ϕ be the unique strictly increasing bijection $B \to C$. Take then the function $g_B : B^n \to B$ which is order-isomorphic to f_C; in other words, for any $x_1, \ldots, x_n \in B$ we have

$$g_B(x_1, \ldots, x_n) = \phi^{-1}(f_D(\phi(x_1), \ldots, \phi(x_n))).$$

As the restriction f_C of f_D to C^n is an increasing selection function, g_B must be an increasing selection function. By Theorem 6, g_B is an order function; let χ be its choice map. But then f_C is also an order function, and it has the same choice map χ. Now for any other subset C' of size 2 of D, $f_{C'}$ is an order function with the same choice map χ, because $f_{C'}$ is order-isomorphic to f_C.

Take now $x_1, \ldots, x_n \in D$ such that $x_{i_1} \leq \ldots \leq x_{i_n}$ for some $(i_1, \ldots, i_n) \in \mathcal{T}_n$. Let $i_t = \chi(i_1, \ldots, i_n)$. We define then $y_1, \ldots, y_n, z_1, \ldots, z_n \in D$ as follows:

$$\begin{aligned} y_i &\doteq \begin{cases} x_{i_1} & \text{if } i = i_j \text{ for } j < t; \\ x_{i_t} & \text{if } i = i_j \text{ for } j \geq t; \end{cases} \\ z_i &\doteq \begin{cases} x_{i_t} & \text{if } i = i_j \text{ for } j \leq t; \\ x_{i_n} & \text{if } i = i_j \text{ for } j > t. \end{cases} \end{aligned} \tag{17}$$

Then for any $i \in I_n$ we have $y_i \leq x_i \leq z_i$, and as f_D is increasing, we get

$$f_D(y_1, \ldots, y_n) \leq f_D(x_1, \ldots, x_n) \leq f_D(z_1, \ldots, z_n). \tag{18}$$

If $x_{i_1} = x_{i_t}$, then $y_1 = \ldots = y_n = u$ by (17), and so (i) implies that $f_D(y_1, \ldots, y_n) = f_D(u, \ldots, u) = u = y_{i_t}$. If $x_{i_1} < x_{i_t}$, then we set $C = \{x_{i_1}, x_{i_t}\}$, and as $y_1, \ldots, y_n \in C$ and $y_{i_1} \leq \ldots \leq y_{i_n}$, we have $f_D(y_1, \ldots, y_n) = f_C(y_1, \ldots, y_n) = y_{i_t}$, because f_C is an order function with choice map χ. A similar argument shows that $f_D(z_1, \ldots, z_n) = z_{i_t}$. As $y_{i_t} = z_{i_t} = x_{i_t}$ (by (17)), we get $f_D(y_1, \ldots, y_n) = f_D(z_1, \ldots, z_n) = x_{i_t}$. Combining this equality with (18), we get $f_D(x_1, \ldots, x_n) = x_{i_t}$, where $i_t = \chi(i_1, \ldots, i_n)$. Hence f_D is an order function whose choice map is χ. ∎

Continuity is an essential requirement in any practical method for processing data, because the quantization of data (necessary for their digital processing) always implies a quantization error, which is propagated throughout subsequent processing. In this respect the fact that order functions are increasing can be used to prove the following:

Theorem 14. *Let f be a selection function $\mathbf{R}^n \to \mathbf{R}$. Then the following three statements are equivalent:*

(*i*) *f is an order function.*

(*ii*) *f is continuous.*

(*iii*) *For any $x_1, \ldots, x_n, y_1, \ldots, y_n \in \mathbf{R}$,*

$$|f(x_1, \ldots, x_n) - f(y_1, \ldots, y_n)| \leq \max\{|x_1 - y_1|, \ldots, |x_n - y_n|\}.$$

Proof. (*a*) (*i*) *implies* (*iii*).
Take $x_1, \ldots, x_n, y_1, \ldots, y_n \in \mathbf{R}$ and let $\epsilon = \max\{|x_1 - y_1|, \ldots, |x_n - y_n|\}$. For each $i \in I_n$ we have $x_i \leq y_i + \epsilon$, and as f is increasing (by Theorem 13), we get

$$f(x_1, \ldots, x_n) \leq f(y_1 + \epsilon, \ldots, y_n + \epsilon).$$

It is clear that for any $(i_1, \ldots, i_n) \in \mathcal{T}_n$, $y_{i_1} \leq \ldots \leq y_{i_n}$ iff $y_{i_1} + \epsilon \leq \ldots \leq y_{i_n} + \epsilon$, and so that for $t \in I_n$, $f(i_1, \ldots, i_n) = y_t$ iff $f(y_1 + \epsilon, \ldots, y_n + \epsilon) = y_t + \epsilon$. In other words,

$$f(y_1 + \epsilon, \ldots, y_n + \epsilon) = f(y_1, \ldots, y_n) + \epsilon.$$

Combining both equalities we get

$$f(x_1, \ldots, x_n) \leq f(y_1, \ldots, y_n) + \epsilon.$$

Now we can invert $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ in the above argument, and so we have

$$f(y_1, \ldots, y_n) \leq f(x_1, \ldots, x_n) + \epsilon.$$

Combining the last two equalities, we get (*iii*).

(*b*) (*iii*) *implies* (*ii*). This is evident.

(*c*) (*ii*) *implies* (*i*).
Let $(i_1, \ldots, i_n) \in \mathcal{T}_n$. Choose $y_1, \ldots, y_n \in \mathbf{R}$ such that $y_{i_1} < \ldots < y_{i_n}$. Let $y_t = f(y_1, \ldots, y_n)$. We set then $t = \chi(i_1, \ldots, i_n)$.

Let $x_1, \ldots, x_n \in \mathbf{R}$ such that $x_{i_1} \leq \ldots \leq x_{i_n}$. We must show that $f(x_1, \ldots, x_n) = x_t$. For any $\lambda \in \mathbf{R}$ such that $0 \leq \lambda \leq 1$ and each $i \in I_n$, we set $z_i(\lambda) = \lambda y_i + (1 - \lambda)x_i$. Thus $z_i(0) = x_i$ and $z_i(1) = y_i$. As f is continuous, the function g defined by $g(\lambda) = f(z_1(\lambda), \ldots, z_n(\lambda)) - z_t(\lambda)$ (for $0 \leq \lambda \leq 1$) is also continuous. Let ϵ be the minimum of all $|y_i - y_j|$ for $i, j \in I_n$, $i \neq j$. Take $\lambda > 0$. Then for any $i, j \in I_n$ with $i \neq j$, $x_i - x_j$ and $y_i - y_j$ have the same sign, and so for any real number μ such that $\lambda \leq \mu \leq 1$, we have

$$|z_i(\mu) - z_j(\mu)| = |\mu(y_i - y_j) + (1 - \mu)(x_i - x_j)| = \mu\,|y_i - y_j| + (1 - \mu)\,|x_i - x_j| \geq \mu\epsilon \geq \lambda\epsilon.$$

Let $X = \{\mu \mid \lambda \leq \mu \leq 1\}$. Now for any $\mu \in X$ we have either

$f(z_1(\mu), \ldots, z_n(\mu)) = z_t(\mu)$ and so $g(\mu) = 0$; or

$f(z_1(\mu), \ldots, z_n(\mu)) = z_s(\mu)$, where $s \neq t$, and so $|g(\mu)| = |z_s(\mu) - z_t(\mu)| \geq \lambda\epsilon$.

Moreover, $g(1) = f(y_1, \ldots, y_n) - y_t = 0$, and so $g(X)$ contains 0. But X is connected, and so its image $g(X)$ by the continuous function g is a connected subset of $\{0\} \cup \{u \in \mathbf{R} \mid |u| \geq \lambda\epsilon\}$. Thus $g(X) = \{0\}$, and so $g(\lambda) = 0$ for $\lambda > 0$. The continuity of g implies that

$$g(0) = \lim_{\lambda \to 0} g(\lambda) = \lim_{\lambda \to 0} 0 = 0,$$

in other words $f(x_1, \ldots, x_n) = x_t$ for $t = \chi(i_1, \ldots, i_n)$. Hence f is an order function. ∎

Thus order functions are continuous, and they are the only continuous selection functions. In particular, property *(iii)* (a particular case of uniform continuity) implies that an order function will never increase the quantization error of its variables. (This is interesting in the case of a succession of compositions of order functions).

For preorder functions one can show the following related result, whose proof is left to the reader.

Proposition 15. *Let* $U = \{(x_1, \ldots, x_n) \in \mathbf{R}^n \mid x_i \neq x_j \text{ for } i \neq j\}$. *Let* f *be a selection function* $U \to \mathbf{R}$. *Then* f *is a preorder function iff* f *is continuous on* U.

It is known that a rank function on $\mathbf{R}^n$ commutes with any increasing function $\mathbf{R} \to \mathbf{R}$ [9,23]. This can be generalized to order functions, and even to preorder functions if one replaces "increasing" by "strictly increasing":

Proposition 16. *Let* f_D *be a function* $D^n \to D$. *Then:*

(i) If f_D *is a preorder function, then* f_D *commutes with any strictly increasing function* $D \to D$.

(ii) If f_D *is an order function, then* f_D *commutes with any increasing function* $D \to D$.

Proof. *(i)* If f_D is a preorder function, let $g_D : D \to D$ be a strictly increasing function. For any $x_1, \ldots, x_n \in D$ and $i, j \in I_n$, we have $x_i < x_j$ iff $g_D(x_i) < g_D(x_j)$. Hence (by Definition 2) if $f_D(x_1, \ldots, x_n) = x_t$, then $f_D(g_D(x_1), \ldots, g_D(x_n)) = g_D(x_t)$, and so g_D commutes with f_D.

(ii) If f_D is an order function, consider an increasing function $g_D : D \to D$. For any $x_1, \ldots, x_n \in D$ such that $x_{i_1} \leq \ldots \leq x_{i_n}$ (where $(i_1, \ldots, i_n) \in \mathcal{T}_n$), we have $g_D(x_{i_1}) \leq \ldots \leq g_D(x_{i_n})$. Thus for $t = \chi(i_1, \ldots, i_n)$ we have $f_D(x_1, \ldots, x_n) = x_t$ and $f_D(g_D(x_{i_1}), \ldots, g_D(x_{i_n})) = g_D(x_t)$, in other words f_D commutes with g_D. ∎

The converse holds also to a certain extent. We will show in Corollary 19 that the converse of *(ii)* is true when $|D| \geq 3$. On the other hand, the converse of *(i)* cannot be proved when D is finite, because in this case the only strictly increasing function $D \to D$ is the identity, and so we cannot deduce anything from the fact that a function commutes with it. We can however prove the following partial converse of Proposition 16 *(i)*:

Proposition 17. *Let f be a function $\mathbf{R}^n \to \mathbf{R}$. If f commutes with any strictly increasing function $\mathbf{R} \to \mathbf{R}$, then f is a preorder function.*

Proof. Suppose that f commutes with any strictly increasing function $\mathbf{R} \to \mathbf{R}$. Let us first show that f is a selection function. Indeed, given $x_1, \ldots, x_n \in \mathbf{R}$, it is easy to see that there exists a strictly increasing function $g : \mathbf{R} \to \mathbf{R}$ such that $g(x_i) = x_i$ for $i \in I_n$, but $g(x) \neq x$ for every $x \notin \{x_1, \ldots, x_n\}$. Then we have

$$g(f(x_1, \ldots, x_n)) = f(g(x_1), \ldots, g(x_n)) = f(x_1, \ldots, x_n),$$

which implies that $f(x_1, \ldots, x_n) \in \{x_1, \ldots, x_n\}$, in other words f is a selection function.

Take $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_n)$ such that for $i, j \in I_n$, $x_i < x_j$ iff $y_i < y_j$. There is thus a strictly increasing function $g : \mathbf{R} \to \mathbf{R}$ such that $g(x_i) = y_i$ for $i \in I_n$. If $f(x_1, \ldots, x_n) = x_t$, then we have

$$f(y_1, \ldots, y_n) = f(g(x_1), \ldots, g(x_n)) = g(f(x_1, \ldots, x_n)) = g(x_t) = y_t.$$

This means that f is a preorder function. ∎

Before proving the converse of Proposition 16 (ii) for $|D| \geq 3$, we will concentrate on a particular class of increasing functions $D \to D$:

For any $y \in \mathbf{R}$, the *thresholding function* θ_y is a map $\mathbf{R} \to B$ defined by

$$\theta_y(x) = \begin{cases} 0 & \text{if } x < y; \\ 1 & \text{if } x \geq y. \end{cases} \tag{19}$$

It is clear that such a function θ_y is increasing. Thus for an order function f, we can apply Proposition 16 (ii), and so for $y \in \mathbf{R}$ and $x_1, \ldots, x_n \in D$ we have

$$\theta_y(f(x_1, \ldots, x_n)) = f(\theta_y(x_1), \ldots, \theta_y(x_n)).$$

In other words, an order function commutes with thresholding. Considering the restrictions f_D and f_B of f to D^n and B^n respectively, we can thus write: given any $y \in \mathbf{R}$,

$$\theta_y(f_D(x_1, \ldots, x_n)) = f_B(\theta_y(x_1), \ldots, \theta_y(x_n)) \qquad \text{for} \quad x_1, \ldots, x_n \in D. \tag{20}$$

The possible converse of this property would be that a function $f_D : D^n \to D$ satisfying (20) for every $y \in \mathbf{R}$ (with $B \subseteq D$) is an order function. More generally, consider two functions $g_D : D^n \to D$ and $h_B : B^n \to B$ and a relation between them of the form:

$$\theta_y(g_D(x_1, \ldots, x_n)) = h_B(\theta_y(x_1), \ldots, \theta_y(x_n)) \qquad \text{for} \quad x_1, \ldots, x_n \in D. \tag{21}$$

If (21) holds for any $y \in \mathbf{R}$, then does it imply that there is an order function f such that $g_D = f_D$ and $h_B = f_B$?

When $|D| = 2$, nothing can be deduced from (21), because in this case thresholding is either a constant function or a bijection $D \to B$. Indeed, for any $y \in \mathbf{R}$, h_B and g_D satisfy (21) provided that g_D and h_B are order-isomorphic selection functions, and when $D = B$, f_B satisfies (20) if it is a selection function.

On the other hand, when $|D| > 2$, that converse is true:

Theorem 18. *Suppose that $|D| \geq 3$. Given two functions $g_D : D^n \to D$ and $h_B : B^n \to B$, g_D and h_B satisfy (21) for any $y \in \mathbf{R}$ iff there is an order function f such that $g_D = f_D$ and $h_B = f_B$.*

Proof. As an order function f satisfies (20), (21) holds for $g_D = f_D$ and $h_B = f_B$.

Suppose now that g_D and h_B satisfy (21) for any $y \in \mathbf{R}$. Let us show that h_B is increasing. Take $x_1, \ldots, x_n, y_1, \ldots, y_n \in B$ such that $x_i \leq y_i$ for each $i \in I_n$. We set:

$$U \doteq \{i \in I_n \mid x_i = y_i = 0\};$$
$$V \doteq \{i \in I_n \mid x_i = 0, y_i = 1\};$$
$$W \doteq \{i \in I_n \mid x_i = y_i = 1\}.$$

As $|D| \geq 3$, there exist $u, v, w \in D$ such that $u < v < w$. Now define $z_1, \ldots, z_n$ as follows:

$$z_i \doteq \begin{cases} u & \text{for } i \in U; \\ v & \text{for } i \in V; \\ w & \text{for } i \in W. \end{cases}$$

Then $\theta_v(z_i) = 1$ for $i \in V \cup W$, and $\theta_w(z_i) = 1$ for $i \in W$, in other words

$$\begin{aligned} \theta_v(z_i) &= y_i, \\ \theta_w(z_i) &= x_i, \end{aligned} \qquad \text{for} \quad i \in I_n. \tag{22}$$

As $w > v$, for any $h \in D$, $\theta_w(h) \leq \theta_v(h)$. Thus

$$\theta_w(g_D(z_1, \ldots, z_n)) \leq \theta_v(g_D(z_1, \ldots, z_n)). \tag{23}$$

Combining (21, 22, 23) we get:

$$\begin{aligned} h_B(x_1, \ldots, x_n) &= h_B(\theta_w(z_1), \ldots, \theta_w(z_n)) = \theta_w(g_D(z_1, \ldots, z_n)) \\ \leq & \theta_v(g_D(z_1, \ldots, z_n)) = h_B(\theta_v(z_1), \ldots, \theta_v(z_n)) = h_B(y_1, \ldots, y_n). \end{aligned}$$

Thus h_B is increasing. By Theorem 6 one of the following holds:

(*a*) h_B is an order function.

(*b*) $h_B(x_1, \ldots, x_n) = 0$ for any $x_1, \ldots, x_n \in B$.

(*c*) $h_B(x_1, \ldots, x_n) = 1$ for any $x_1, \ldots, x_n \in B$.

If (a) holds, then there is an order function f such that $f_B = h_B$. By (20, 21) we have for any $y \in \mathbf{R}$ and $z_1, \ldots, z_n \in D$:

$$\theta_y(f_D(z_1, \ldots, z_n)) = f_B(\theta_y(z_1), \ldots, \theta_y(z_n)) = \theta_y(g_D(z_1, \ldots, z_n)).$$

Thus for any $y \in \mathbf{R}$ and $z_1, \ldots, z_n \in D$, $f_D(z_1, \ldots, z_n) \geq y$ iff $g_D(z_1, \ldots, z_n) \geq y$, in other words $g_D = f_D$.

If (b) holds, then for any $y \in \mathbf{R}$ and $z_1, \ldots, z_n \in D$ we have

$$\theta_y(g_D(z_1, \ldots, z_n)) = h_B(\theta_y(z_1), \ldots, \theta_y(z_n)) = 0,$$

that is $g_D(z_1, \ldots, z_n) < y$ for any $y \in \mathbf{R}$, a contradiction.

If (c) holds, then for any $y \in \mathbf{R}$ and $z_1, \ldots, z_n \in D$ we have

$$\theta_y(g_D(z_1, \ldots, z_n)) = h_B(\theta_y(z_1), \ldots, \theta_y(z_n)) = 1,$$

that is $g_D(z_1, \ldots, z_n) \geq y$ for any $y \in \mathbf{R}$, a contradiction. ∎

We derive from this result the following partial converse of Proposition 16 (ii):

Corollary 19. *Suppose that $|D| \geq 3$. Let f_D be a function $D^n \to D$. Then f_D is an order function iff it commutes with any increasing function $D \to D$.*

Proof. We showed in Proposition 16 (ii) that if f_D is an order function, then it commutes with any increasing function $D \to D$.

Suppose now that f_D commutes with any increasing function $D \to D$. We show that f_D is an order function. Take $a_0, a_1 \in D$ such that $a_0 < a_1$, and consider the bijection $\phi: B \to \{a_0, a_1\}$ defined by $\phi(0) = a_0$ and $\phi(1) = a_1$. Let h_B be the function $B^n \to B$ defined by

$$h_B(x_1, \ldots, x_n) = \phi^{-1}(f_D(\phi(x_1), \ldots, \phi(x_n))) \tag{24}$$

for $x_1, \ldots, x_n \in B$. (In other words, h_B is order-isomorphic to $f_{\{a_0, a_1\}}$). For any $y \in \mathbf{R}$, the map $\psi_y : D \to \{a_0, a_1\} : z \mapsto \phi(\theta_y(z))$ is increasing. Therefore it commutes with f_D, and so for any $z_1, \ldots, z_n \in D$ we have

$$\psi_y(f_D(z_1, \ldots, z_n)) = f_D(\psi_y(z_1), \ldots, \psi_y(z_n)),$$

in other words

$$\phi(\theta_y(f_D(z_1, \ldots, z_n))) = f_D(\phi(\theta_y(z_1)), \ldots, \phi(\theta_y(z_n))).$$

Applying ϕ^{-1} to both sides, we get by (24):

$$\theta_y(f_D(z_1, \ldots, z_n)) = \phi^{-1}(f_D(\phi(\theta_y(z_1)), \ldots, \phi(\theta_y(z_n)))) = h_B(\theta_y(z_1), \ldots, \theta_y(z_n)).$$

Thus f_D and h_B satisfy (21), and by Theorem 18 f_D is an order function. ∎

When $|D| = 2$, it is easy to see that an increasing function $D \to D$ is the identity or a constant function; in fact, for $D = B$ such a function is equal to a thresholding. Thus a function $D^n \to D$ commutes with any increasing function $D \to D$ iff it commutes with thresholding, and this happens iff it is a selection function. Note that when $n = 2$, it is easy to show that a selection function $D^n \to D$ is an order function.

IV.2. Some consequences

Commutation with thresholding (see (20)) implies that the behavior of an order function f_D on D^n is determined by that of the corresponding order function f_B on B^n. Indeed, for any $z \in D$, z is determined by the set of all $\theta_y(z)$ for $y \in \mathbf{R}$; thus for $x_1, \ldots, x_n \in D$ the value of $f_D(x_1, \ldots, x_n)$ is determined by all thresholded values $\theta_y(f_D(x_1, \ldots, x_n)) = f_B(\theta_y(x_1), \ldots, \theta_y(x_n))$, which are determined by the behavior of f_B. A practical application of this fact to the computation of $f_D(x_1, \ldots, x_n)$, called *threshold decomposition*, will be given in Section V.

The last four results of the preceding subsection (*i.e.*, Propositions 16 and 17, Theorem 18 and Corollary 19) dealt with the problem of commutation with increasing (or strictly increasing) functions $D \to D$. What about decreasing (or strictly decreasing) functions? We will show that similar results can be obtained, but here we obtain, instead of the commutation of f_D with such a function g_D, an equality of the form

$$g_D(f_D(x_1, \ldots, x_n)) = f_D^*(g_D(x_1), \ldots, g_D(x_n)). \tag{25}$$

We say then that g_D *commutes the pair* (f_D, f_D^*). We give below the four statements corresponding to the four results mentioned above:

Proposition 16'. *Let f_D and g_D be two functions $D^n \to D$ and $D \to D$ respectively. If either*

(i) f_D is a preorder function and g_D is strictly decreasing, or

(ii) f_D is an order function and g_D is decreasing,

then g_D commutes with the pair (f_D, f_D^*), where f_D^* is the dual of f_D.

Proof. (*i*) If f_D is a preorder function and g_D is strictly increasing, then for any $x_1, \ldots, x_n \in D$ and $i, j \in I_n$, we have $x_i < x_j$ iff $g_D(x_i) > g_D(x_j)$. Hence (by Definition 4) $f_D(x_1, \ldots, x_n) = x_t$ implies that $f_D^*(g_D(x_1), \ldots, g_D(x_n)) = g_D(x_t)$, and so f_D and f_D^* satisfy (25).

(*ii*) If f_D is an order function and g_D is increasing, then for any $x_1, \ldots, x_n \in D$ such that $x_{i_1} \leq \ldots \leq x_{i_n}$ (where $(i_1, \ldots, i_n) \in \mathcal{T}_n$), we have $g_D(x_{i_n}) \leq \ldots \leq g_D(x_{i_1})$. Thus for $t = \chi(i_1, \ldots, i_n)$ we have $f_D(x_1, \ldots, x_n) = x_t$, and as $\chi^*(i_n, \ldots, i_1) = t$ (by (2)), $f_D^*(g_D(x_1), \ldots, g_D(x_n)) = g_D(x_t)$, in other words f_D and f_D^* satisfy (25). ∎

Proposition 17'. *Let f, f' be two functions $\mathbf{R}^n \to \mathbf{R}$. If any strictly increasing function $\mathbf{R} \to \mathbf{R}$ commutes with the pair (f, f'), then f is a preorder function and $f' = f^*$, the dual of f.*

Proof. The map $\mathbf{R} \to \mathbf{R} : x \mapsto -x$ is strictly decreasing, and so it commutes with the pair (f, f'), in other words for any $y_1, \ldots, y_n \in \mathbf{R}$ we have

$$-f(y_1, \ldots, y_n) = f'(-y_1, \ldots, -y_n). \tag{26}$$

For any strictly increasing map $g: \mathbf{R} \rightarrow \mathbf{R}$, the map $\mathbf{R} \rightarrow \mathbf{R}: x \mapsto -g(x)$ is strictly decreasing. We have then for any $x_1, \ldots, x_n \in \mathbf{R}$:

$$-g(f(x_1, \ldots, x_n)) = f'(-g(x_1), \ldots, -g(x_n)).$$

Applying (26) with $y_i = g(x_i)$ $(i \in I_n)$, we get:

$$-f(g(x_1), \ldots, g(x_n)) = f'(-g(x_1), \ldots, -g(x_n)).$$

Combining the last two inequalities, we obtain

$$g(f(x_1, \ldots, x_n)) = f(g(x_1), \ldots, g(x_n)),$$

in other words f commutes with g. By Proposition 17, f is a preorder function. But then (26) means that $f' = f^*$. ∎

Recall the complementation $x \mapsto \overline{x} = 1 - x$ in B. The decreasing correspondent of thresholding is the *complemented thresholding* $\overline{\theta}_y$ defined by $\overline{\theta}_y(x) = \overline{\theta_y(x)}$. We get then instead of (20): given any $y \in \mathbf{R}$,

$$\overline{\theta}_y(f_D(x_1, \ldots, x_n)) = f_B^*(\overline{\theta}_y(x_1), \ldots, \overline{\theta}_y(x_n)) \qquad \text{for} \quad x_1, \ldots, x_n \in D. \tag{27}$$

We have then the following:

Theorem 18'. *Suppose that* $|D| \geq 3$. *Given two functions* $g_D: D^n \rightarrow D$ *and* $h_B: B^n \rightarrow B$, g_D *and* h_B *satisfy*

$$\overline{\theta}_y(g_D(x_1, \ldots, x_n)) = h_B(\overline{\theta}_y(x_1), \ldots, \overline{\theta}_y(x_n)) \qquad \text{for} \quad x_1, \ldots, x_n \in D \tag{28}$$

for any $y \in \mathbf{R}$ *iff there is an order function* f *such that* $g_D = f_D$ *and* $h_B = f_B^*$.

Proof. As an order function f satisfies (27), (28) holds for $g_D = f_D$ and $h_B = f_B^*$.

Suppose now that (28) holds. We can rewrite it as

$$\theta_y(g_D(x_1, \ldots, x_n)) = \overline{h_B(\overline{\theta_y(x_1)}, \ldots, \overline{\theta_y(x_n)})},$$

and defining the function $h'_B: B^n \rightarrow B$ by

$$h'_B(z_1, \ldots, z_n) = \overline{h_B(\overline{z_1}, \ldots, \overline{z_n})}, \tag{29}$$

we get

$$\theta_y(g_D(x_1, \ldots, x_n)) = h'_B(\theta_y(x_1), \ldots, \theta_y(x_n)),$$

in other words g_D and h'_D satisfy (21) for any $y \in \mathbf{R}$. Thus by Theorem 18 there is an order function f such that $g_D = f_D$ and $h'_B = f_B$. But then (29) means that $h_B = h'^*_B = f_B^*$. ∎

Corollary 19'. *Suppose that $|D| \geq 3$. Let f_D, f'_D be two functions $D^n \to D$. Then f_D is an order function and $f'_D = f^*_D$ iff any increasing function $D \to D$ commutes with the pair (f_D, f'_D).*

The proof is similar to that of Corollary 19, and it is left to the reader.

Given $a, b \in \mathbf{R}$, the map $x \mapsto ax + b$ is strictly increasing when $a > 0$, and strictly decreasing when $a < 0$. When $a = 0$, it is a constant map, and so it commutes with any selection function. We derive thus from Propositions 16 and 16' the following:

Corollary 20. *Let $f : \mathbf{R}^n \to \mathbf{R}$ be a preorder function. Then for any $a, b \in \mathbf{R}$ and $x_1, \ldots, x_n \in \mathbf{R}$ we have*

$$f(ax_1 + b, \ldots, ax_n + b) = \begin{cases} af(x_1, \ldots, x_n) + b & \text{if } a \geq 0; \\ af^*(x_1, \ldots, x_n) + b & \text{if } a \leq 0. \end{cases} \tag{30}$$

In particular, when f is equal to its dual f^* (for example if f is the median), then f commutes with the map $x \mapsto ax + b$.

V. Threshold decomposition and its unicity

Recall the definition of the thresholding function θ_y in (19): for $y \in \mathbf{R}$ and $x \in D$ we have $\theta_y(x) = 1$ if $x \geq y$, and $\theta_y(x) = 0$ if $x < y$. The value y is called the *threshold*. Given a order function $f : \mathbf{R}^n \to \mathbf{R}$, f commutes with θ_y for any $y \in \mathbf{R}$; this means (see (20)) that for $x_1, \ldots, x_n \in D$ and $z = f_D(x_1, \ldots, x_n)$, $\theta_y(z) = f_B(\theta_y(x_1), \ldots, \theta_y(x_n))$, where f_B and f_D are the order functions $B^n \to B$ and $D^n \to D$ corresponding to f.

As we explained at the beginning of Subsection IV.2, a consequence of this fact is that the behavior of f_D is determined by the behavior of f_B, because z is determined by its thresholded values $\theta_y(z)$ for all $y \in D$. Thus many properties found in the binary case can be directly extended to the non-binary case. For example, if two order functions f and g satisfy $f_B \leq g_B$, then $f_D \leq g_D$, or if $f_B = g_B \circ [h_B^{(1)}, \ldots, h_B^{(n)}]$ for n order functions $h^{(1)}, \ldots, h^{(n)}$, then $f_D = g_D \circ [h_D^{(1)}, \ldots, h_D^{(n)}]$.

In fact, many results of Section III were also found by relating the properties of f_D to those of f_B, for example the min-max decompositions of f_D were found by an analysis of the α-heavy sets for f_B.

In practice, thresholding can be useful for the computation of the values of an order function f_D only if we have to apply it with a finite number of thresholds. In other words, the set D should be finite. In this case it is possible to obtain $f_D(x_1, \ldots, x_n)$ by a linear combination of the results of f_B on all thresholded values of $x_1, \ldots, x_n$, as we explain in Subsection V.1. We show moreover in Subsection V.2 that there is no other possible method for a linear decomposition of any order function. For the reader with only practical applications in mind, the latter subsection can be skipped.

V.1. Threshold decomposition

In practical situations, when D is finite one generally assumes that D is the set of all integers k such that $0 \leq k \leq m$ for some integer $m > 0$. For example, in signal processing D will often be the set of integers $0, \ldots, 255$. Now more can be said about thresholding in this case:

Proposition 21. *Suppose that $D = \{0, \ldots, m\}$ for some integer $m > 0$. Then:*

(a) *for any $k \in D$, $k = \sum_{j=1}^{m} \theta_j(k)$.*

(b) *Let $f_D : D^n \to D$ be an order function and let $y, x_1, \ldots, x_n \in D$ such that $y = f_D(x_1, \ldots, x_n)$. For any $j = 1, \ldots, m$, set $y^{(j)} = \theta_j(y)$ and $x_i^{(j)} = \theta_j(x_i)$ ($i \in I_n$). Then:*

  (i) *For any $i \in I_n$, $x_i = \sum_{j=1}^{m} x_i^{(j)}$.*

  (ii) *For any $j = 1, \ldots, m$, $y^{(j)} = f(x_1^{(j)}, \ldots, x_n^{(j)})$.*

  (iii) *$y = \sum_{j=1}^{m} y^{(j)}$.*

The proof of (a) is straightforward (and we omit it), while (b) is an immediate consequence of (a) and (20).

Proposition 21 shows the functioning of the method of *threshold decomposition*, which was devised by Fitch [5] for the median filter, but is still valid for order functions and, as we will see in Section VI, for order filters. We compute the m thresholded vectors $(\theta_k(x_1), \dots, \theta_k(x_n))$ for $k = 1, \dots, m$, then we apply f_D to each one, and finally we sum the m results.

Another form of threshold decomposition exists when D is finite, but not of the form $\{0, \dots, m\}$:

Proposition 22. *Suppose that* $D = \{d_0, \dots, d_m\}$ *for some integer* $m > 0$, *with* $d_0 < \dots < d_m$. *For* $j = 1, \dots, m$, *let* $c_j = d_j - d_{j-1}$. *Then:*

(a) *For any* $k \in D$, $k = d_0 + \sum_{j=1}^{m} c_j \cdot \theta_{d_j}(k)$.

(b) *Let* $f_D : D^n \to D$ *be an order function and let* $y, x_1, \dots, x_n \in D$ *such that* $y = f_D(x_1, \dots, x_n)$. *For any* $j = 1, \dots, m$, *set* $y^{(j)} = \theta_{d_j}(y)$ *and* $x_i^{(j)} = \theta_{d_j}(x_i)$ $(i \in I_n)$. *Then:*

(i) *For any* $i \in I_n$, $x_i = d_0 + \sum_{j=1}^{m} c_j \cdot x_i^{(j)}$.

(ii) *For any* $j = 1, \dots, m$, $y^{(j)} = f(x_1^{(j)}, \dots, x_n^{(j)})$.

(iii) $y = d_0 + \sum_{j=1}^{m} c_j \cdot y^{(j)}$.

Again the proof is left to the reader. Note that a similar threshold decomposition formula has been given for another type of functions, *multivalued multithreshold functions* (see Definition 3 and Subsection IV.A.2 of [1]).

Proposition 22 can also be used for the computation of $f_D(x_1, \dots, x_n)$ for $x_1, \dots, x_n \in D$, whatever the form taken by D. Indeed, suppose that the distinct values of $x_1, \dots, x_n \in D$ are $y_1, \dots, y_k$, with $y_1 < \dots < y_k$. Then Proposition 22 implies that

$$f_D(x_1, \dots, x_n) = y_1 + \sum_{j=2}^{k} (y_j - y_{j-1}) f(\theta_{y_j}(x_1), \dots, \theta_{y_j}(x_n)).$$

When k is much smaller than the size of D, this reduces by far the total number of computations necessary to obtain $f_D(x_1, \dots, x_n)$.

One computational consequence of threshold decomposition is that for $D = \{0, \dots, m\}$ the complexity of computing $f_D(x_1, \dots, x_n)$ for $x_1, \dots, x_n \in D$ is asymptotically at most m times that of computing $f_B(y_1, \dots, y_n)$ for $y_1, \dots, y_n \in B$, plus a term linear in n.

We can now consider the reverse problem: are there other decompositions of a variable x as a linear combination $\lambda_1 x^{(1)} + \dots + \lambda_m x^{(m)}$ such that for any order function f_D we have $f_D(x_1, \dots, x_n) = \lambda_1 f_D(x_1^{(1)}, \dots, x_n^{(1)}) + \dots + \lambda_m f_D(x_1^{(m)}, \dots, x_n^{(m)})$? The answer is negative. We will show in the next subsection that the only such decomposition is the threshold decomposition. Readers uninterested in matheatical details can skip it and resume the reading in Section VI.

V.2. The impossibility of other linear decompositions

We consider first a collection $\mathcal{X}$ of n-tuples in $\mathbf{R}^n$ and contemplate the possibility of a linear decomposition of an order function on $\mathcal{X}$. This requires the introduction of a new concept. We say that $\mathcal{X}$ is *order-coherent* if there exists some $(i_1, \ldots, i_n) \in \mathcal{T}_n$ such that for any $(x_1, \ldots, x_n) \in \mathcal{X}$, $x_{i_1} \leq \ldots \leq x_{i_n}$, in other words if the elements of $\mathcal{X}$ admit a common ordering.

We have the following characterization:

Proposition 23. *Let $\mathcal{X}$ be a set (of size larger than 1) of n-tuples in $\mathbf{R}^n$. Then the following four statements are equivalent:*

(i) $\mathcal{X}$ is order-coherent.

(ii) For every $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in \mathcal{X}$ and $a, b \in \mathcal{I}_n$ with $a \neq b$, we may not have both $x_a < x_b$ and $y_a > y_b$.

(iii) For any $(z_1^{(1)}, \ldots, z_n^{(1)}), \ldots, (z_1^{(t)}, \ldots, z_n^{(t)}) \in \mathcal{X}$ and any nonnegative $\lambda_1, \ldots, \lambda_t \in \mathbf{R}$, where $t \geq 2$, we have

$$f(\lambda_1 z_1^{(1)} + \cdots + \lambda_t z_1^{(t)}, \ldots, \lambda_1 z_n^{(1)} + \cdots + \lambda_t z_n^{(t)}) = \lambda_1 f(z_1^{(1)}, \ldots, z_n^{(1)}) + \cdots + \lambda_t f(z_1^{(t)}, \ldots, z_n^{(t)})$$

for every order function $f: \mathbf{R}^n \to \mathbf{R}$.

(iv) For any $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in \mathcal{X}$, there exist $(z_1^{(1)}, \ldots, z_n^{(1)}), \ldots, (z_1^{(t)}, \ldots, z_n^{(t)}) \in \mathbf{R}^n$ and nonnegative $\lambda_1, \ldots, \lambda_t \in \mathbf{R}$, where $t \geq 2$, $\lambda_1, \lambda_2 > 0$, $(z_1^{(1)}, \ldots, z_n^{(1)}) = (x_1, \ldots, x_n)$, and $(z_1^{(2)}, \ldots, z_n^{(2)}) = (y_1, \ldots, y_n)$, such that

$$f(\lambda_1 z_1^{(1)} + \cdots + \lambda_t z_1^{(t)}, \ldots, \lambda_1 z_n^{(1)} + \cdots + \lambda_t z_n^{(t)}) = \lambda_1 f(z_1^{(1)}, \ldots, z_n^{(1)}) + \cdots + \lambda_t f(z_1^{(t)}, \ldots, z_n^{(t)})$$

for every order function $f: \mathbf{R}^n \to \mathbf{R}$.

Proof. *(a) (ii) implies (i).*

Suppose that *(ii)* holds. Given $a, b \in \mathcal{I}_n$, we have one of the following two possibilities:

— for any $(x_1, \ldots, x_n) \in \mathcal{X}$, $x_a \leq x_b$;

— for any $(x_1, \ldots, x_n) \in \mathcal{X}$, $x_a \geq x_b$.

In the first case, we write $a \preceq b$, and in the second one we write $b \preceq a$. Then the relation $\preceq$ is transitive, and so it can be used to order the elements of $\mathcal{I}_n$. Thus there is some $(i_1, \ldots, i_n) \in \mathcal{T}_n$ such that $i_1 \preceq \ldots \preceq i_n$, in other words *(i)*.

(b) (i) *implies* (iii).

There is some $(i_1, \ldots, i_n) \in \mathcal{T}_n$ such that for any $(x_1, \ldots, x_n) \in \mathcal{X}$, $x_{i_1} \leq \ldots \leq x_{i_n}$. Given an order function f with choice map χ, let $t = \chi(i_1, \ldots, i_n)$. For $j = 1, \ldots, t$, we have $z_{i_1}^{(j)} \leq \ldots \leq z_{i_n}^{(j)}$, and so

$$\lambda_1 z_1^{(1)} + \cdots + \lambda_t z_1^{(t)} \leq \ldots \leq \lambda_1 z_n^{(1)} + \cdots + \lambda_t z_n^{(t)}.$$

As f is an order function, this means that $f(z_1^{(j)}, \ldots, z_n^{(j)}) = z_t^{(j)}$ and

$$f(\lambda_1 z_1^{(1)} + \cdots + \lambda_t z_1^{(t)}, \ldots, \lambda_1 z_n^{(1)} + \cdots + \lambda_t z_n^{(t)}) = \lambda_1 z_t^{(1)} + \cdots + \lambda_t z_t^{(t)}.$$

Now we have also

$$\lambda_1 f(z_1^{(1)}, \ldots, z_n^{(1)}) + \cdots + \lambda_t f(z_1^{(t)}, \ldots, z_n^{(t)}) = \lambda_1 z_t^{(1)} + \cdots + \lambda_t z_t^{(t)}.$$

Combining both equalities, we get (iii).

(c) (iii) *implies* (iv). This is evident.

(d) (iv) *implies* (ii).

Suppose that there exist $a, b \in \mathcal{I}_n$ such that $z_a^{(1)} = x_a < x_b = z_b^{(1)}$ and $z_a^{(2)} = y_a > y_b = z_b^{(1)}$. We take the order function f defined by $f(z_1, \ldots, z_n) = \max\{z_a, z_b\}$. We have $f(z_1^{(1)}, \ldots, z_n^{(1)}) = \max\{z_a^{(1)}, z_b^{(1)}\} > z_a^{(1)}$, while for $j = 2, \ldots, t$ we have $f(z_1^{(j)}, \ldots, z_n^{(j)}) = \max\{z_a^{(j)}, z_b^{(j)}\} \geq z_a^{(j)}$. By taking a linear combination of these inequalities with $\lambda_1, \lambda_2 > 0$ and $\lambda_j \geq 0$ for $j = 3, \ldots, t$, we get:

$$\lambda_1 f(z_1^{(1)}, \ldots, z_n^{(1)}) + \cdots + \lambda_t f(z_1^{(t)}, \ldots, z_n^{(t)}) > \lambda_1 z_a^{(1)} + \cdots + \lambda_t z_a^{(t)}.$$

We have similarly $f(z_1^{(2)}, \ldots, z_n^{(2)}) = \max\{z_a^{(2)}, z_b^{(2)}\} > z_b^{(2)}$, while $f(z_1^{(j)}, \ldots, z_n^{(j)}) = \max\{z_a^{(j)}, z_b^{(j)}\} \geq z_b^{(j)}$ for $j = 1, 3, \ldots, t$, and so:

$$\lambda_1 f(z_1^{(1)}, \ldots, z_n^{(1)}) + \cdots + \lambda_t f(z_1^{(t)}, \ldots, z_n^{(t)}) > \lambda_1 z_b^{(1)} + \cdots + \lambda_t z_b^{(t)}.$$

Combining both inequalities, we obtain

$$\begin{aligned}
&\lambda_1 f(z_1^{(1)}, \ldots, z_n^{(1)}) + \cdots + \lambda_t f(z_1^{(t)}, \ldots, z_n^{(t)}) \\
>& \max\{\lambda_1 z_a^{(1)} + \cdots + \lambda_t z_a^{(t)}, \lambda_1 z_b^{(1)} + \cdots + \lambda_t z_b^{(t)}\} \\
=& f(\lambda_1 z_1^{(1)} + \cdots + \lambda_t z_1^{(t)}, \ldots, \lambda_1 z_n^{(1)} + \cdots + \lambda_t z_n^{(t)}),
\end{aligned}$$

which contradicts (iv). ∎

Statements (iii) and (iv) mean that order-coherence is a neccessary (by (iv)) and sufficient (by (iii)) condition for the linearity of an order function on elements of $\mathcal{X}$, while statement (ii) is another expression of the order-coherence of $\mathcal{X}$. Note that by (ii), $\mathcal{X}$ is order-coherent iff every pair in $\mathcal{X}$ is order-coherent.

In the binary case, we have a simple characterization of order-coherence. Recall the set $B = \{0,1\}$. Given two vectors $(x_1,\ldots,x_n)$ and $(y_1,\ldots,y_n)$, we write $(x_1,\ldots,x_n) \leq (y_1,\ldots,y_n)$ if $x_i \leq y_i$ for each $i \in I_n$.

Proposition 24. *Let $\mathcal{X}$ be a set (of size larger than 1) of n-tuples in B^n. Then $\mathcal{X}$ is order-coherent iff for every $(x_1,\ldots,x_n),(y_1,\ldots,y_n) \in \mathcal{X}$, either $(x_1,\ldots,x_n) \leq (y_1,\ldots,y_n)$, or $(y_1,\ldots,y_n) \leq (x_1,\ldots,x_n)$.*

Proof. By Proposition 23 (ii) the set $\mathcal{X}$ is not order-coherent iff there exist two n-tuples $(x_1,\ldots,x_n),(y_1,\ldots,y_n) \in \mathcal{X}$ and $a,b \in I_n$ such that $x_a < x_b$ and $y_a > y_b$. As $x_a, x_b, y_a, y_b \in B$, the following set of inequalities are equivalent:

$x_a < x_b$ and $y_a > y_b$.

$x_a = 0,\ x_b = 1,\ y_a = 1$, and $y_b = 0$.

$x_a < y_a$ and $x_b > y_b$.

But precisely we have neither $(y_1,\ldots,y_n) \leq (x_1,\ldots,x_n)$ nor $(x_1,\ldots,x_n) \leq (y_1,\ldots,y_n)$ iff $x_a < y_a$ and $x_b > y_b$ for some $a,b \in I_n$. Thus $\mathcal{X}$ is not order-coherent iff there are $(x_1,\ldots,x_n),(y_1,\ldots,y_n) \in \mathcal{X}$ such that we have neither $(y_1,\ldots,y_n) \leq (x_1,\ldots,x_n)$ nor $(x_1,\ldots,x_n) \leq (y_1,\ldots,y_n)$. ∎

Given a vector $(x_1,\ldots,x_n) \in \mathbf{R}^n$ and a number $y \in \mathbf{R}$, we will write $\theta_y(x_1,\ldots,x_n)$ for $(\theta_y(x_1),\ldots,\theta_y(x_n))$. We derive from Propositions 23 and 24 the following characterization of threshold decomposition:

Corollary 25. *Let $m \geq 2$ and $D = \{0,\ldots,m\}$. Consider m vectors $(x_1^{(1)},\ldots,x_n^{(1)})$, $\ldots$, $(x_1^{(m)},\ldots,x_n^{(m)})$ in B^n, and let $(x_1,\ldots,x_n)$ be their sum. Then $(x_1,\ldots,x_n) \in D^n$ and the following two statements are equivalent:*

(i) *For every order function $f_D : D^n \to D$,*

$$f_D(x_1,\ldots,x_n) = f_D(x_1^{(1)},\ldots,x_n^{(1)}) + \cdots + f_D(x_1^{(m)},\ldots,x_n^{(m)}).$$

(ii) $\{(x_1^{(1)},\ldots,x_n^{(1)}),\ldots,(x_1^{(m)},\ldots,x_n^{(m)})\} = \{\theta_1(x_1,\ldots,x_n),\ldots,\theta_n(x_1,\ldots,x_n)\}$.

Proof. As the sum of m elements of B is in $\{0,\ldots,m\}$, $(x_1,\ldots,x_n) \in D^n$. If (ii) holds, then (i) follows from Proposition 21. Let us now show that (i) implies (ii):

If (i) holds, then $(x_1^{(1)},\ldots,x_n^{(1)}),\ldots,(x_1^{(m)},\ldots,x_n^{(m)})$ constitute an order-coherent set by Proposition 23 (iv). Now Proposition 24 implies that for $u,v \in \{1,\ldots,m\}$, either $(x_1^{(u)},\ldots,x_n^{(u)}) \geq (x_1^{(v)},\ldots,x_n^{(v)})$, or $(x_1^{(v)},\ldots,x_n^{(v)}) \geq (x_1^{(u)},\ldots,x_n^{(u)})$. Thus we have

$$(x_1^{(t_1)},\ldots,x_n^{(t_1)}) \geq \ldots \geq (x_1^{(t_m)},\ldots,x_n^{(t_m)})$$

for some $t_1,\ldots,t_m \in \{1,\ldots,m\}$ with $\{t_1,\ldots,t_m\} = \{1,\ldots,m\}$.

For any $i \in I_n$, $x_i^{(t_1)} + \cdots + x_i^{(t_m)} = x_i$; as $x_i^{(t_1)} \geq \ldots \geq x_i^{(t_m)}$ and $x_i^{(t_1)}, \ldots, x_i^{(t_m)} \in B$, this implies that given $j \in \{1, \ldots, m\}$,

$$x_i^{(t_j)} = \begin{cases} 1 & \text{for } j \leq x_i; \\ 0 & \text{for } j > x_i. \end{cases}$$

Now the right-hand side of this equality is nothing but $\theta_j(x_i)$. Thus $x_i^{(t_j)} = \theta_j(x_i)$, and so $(x_1^{(t_j)}, \ldots, x_n^{(t_j)}) = \theta_j(x_1, \ldots, x_n)$. Hence (ii) holds. ∎

This shows that for $D = \{0, \ldots, m\}$, threshold decomposition is the unique way to decompose a signal $S \to D$ into a sum of binary signals in such a way that any order filter respects that decomposition.

More generally, one can apply Propositions 23 and 24 to other problems of signal decompositions; in particular one can give criteria for the order-coherence of a set of two or more non-binary n-tuples.

VI. Order filters

In this section we will consider local filters for digital signals based on order functions, what we call *order filters.* In Subsection VI.1 we define preorder and order filters, and apply to them concepts of Section II, such as the dual and the composition. In Subsection VI.2 we apply results of Sections IV and V to order filters, and in Subsection VI.3 we give some consequences of these results.

VI.1. Definition and main features

Let $\mathbf{Z}$ be the set of rational integers. Usually a physical signal is a real function, and it must be digitized. This is done as follows. A d-dimensional signal $\mathbf{R}^d \to \mathbf{R}$ is sampled at fixed intervals, and the sampled signal can thus be considered as a map $\mathbf{Z}^d \to \mathbf{R}$. The signal values are quantized, and one obtains thus a digital signal $\mathbf{Z}^d \to \mathbf{Z}$. In practice one works with a finite digitization, and so the digitized signal will be a map $S \to D$, where S and D are finite subsets of $\mathbf{Z}^d$ and $\mathbf{Z}$ respectively.

We will consider here *signals* $S \to D$, where S is a discrete set and D is any subset of $\mathbf{R}$. The elements of S are called *points*, and those of D are called *signal values.* Given such a signal X, its value on a point $p \in S$ will be written $X(p)$. We write D^S for the set of all signals $S \to D$.

We assume that S is totally ordered by a precedence relation $\prec$. For example if $S \subseteq \mathbf{Z}^d$, that ordering can be chosen as the "lexicographic order", which is used for ordering decimal numbers or words in dictionnaries, and is defined by $(x_1, \ldots, x_d) \prec (y_1, \ldots, y_d)$ if there is some k with $1 \leq k \leq d$ such that $x_k < y_k$ and $x_j = y_j$ for $1 \leq j < k$ (for $d = 2$ it is also called the "raster-scan order", while for $d = 1$ it reduces to the usual ordering relation $<$). Then given a signal $X \in D^S$, a subset P of size n of S, and a function $f_D : D^n \to D$, we will write $f_D(X(q) \mid q \in P)$ for $f_D(X(p_1), \ldots, X(p_n))$, where $p_1, \ldots, p_n$ are the elements of P, with $p_1 \prec \ldots \prec p_n$.

Let us now define a *local filter* for signals in D^S. We suppose that there exists a *windowing function* φ which associates to each point $p \in S$ a *finite* subset $\varphi(p)$ of S, generally containing p, called the *window* corresponding to p. Let n_p be the size of $\varphi(p)$ $(p \in S)$. Assume that for each $p \in S$ there is a function $f_D[p] : D^{n_p} \to D$. Then φ and the functions $f_D[p]$ for $p \in S$ induce a *local filter* F_D which transforms a signal $X : S \to D$ into another signal $X^{F_D} : S \to D$ defined by:

$$X^{F_D}(p) = f_D[p](X(q) \mid q \in \varphi(p)) \qquad (p \in S). \tag{31}$$

Then we say that the local filter F_D is a *selection filter* if the functions $f_D[p]$ $(p \in S)$ are selection functions, that F_D is a *preorder filter* if the functions $f_D[p]$ $(p \in S)$ are preorder functions, and that F_D is an *order filter* if the functions $f_D[p]$ $(p \in S)$ are order functions.

In fact the choice of the ordering $\prec$ of the points on the point set S does not matter. Indeed, changing that ordering leads to a permutation of the signal values entered as variables of each $f_D[p]$, and a permutation of variables does not change the property of being an order function, a preorder function, or a selection function (see the end of Subsection II.4). Thus the fact that a filter is a selection filter, a preorder filter, or an order filter, does not depend on the choice of the ordering of S.

Similarly the size and shape of the windows $\varphi(p)$ is not fundamental: one can extend a window $\varphi(p)$ to a larger window $\varphi'(p)$ containing it, and this leads to a void expansion of the function $f_D[p]$. Again this operation does not change the property of being an order function, a preorder function, or a selection function (see the end of Subsection II.4). Thus the fact that a filter is a selection filter, a preorder filter, or an order filter, is not modified by the enlargement of windows.

As it was the case with order functions (see Proposition 2), order filters can be made independent of the set D. Thus for $C \subseteq D$ the restriction to signals in C^S of an order filter F_D on D^S will be an order filter written F_C, and for $E \supseteq D$, F_D admits a unique extension to an order filter for signals in E^S, written F_E. We can assume that the order filter F_D for signals in D^S is the restriction to D^S of an order filter F for signals in $\mathbf{R}^S$.

Local filters have been used extensively in digital signal processing. Let us give here a few examples of preorder and order filters.

For $D = B$ and $S \subseteq \mathbf{Z}^2$ (or $\mathbf{Z}^3$), local filters are local operators for binary (black and white) digital images, and they are well-known. Here the resulting *grey-level* (1 for black and 0 for white) of a *pixel* p is determined by the configuration of grey-levels in the window $\varphi(p)$. One often takes $\varphi(p)$ to be the 3×3 neighborhood of p. Such a filter is a preorder filter iff it is a selection filter, and this happens iff each $f_B[p]$ satisfies $f_B[p](0, \ldots, 0) = 0$ and $f_B[p](1, \ldots, 1) = 1$, in other words iff every pixel in a neighborhood of constant grey-level is not modified by the operator. Moreover, it is an order filter iff it is an increasing preorder filter (see Theorem 13). There are several known examples of preorder filters for binary images:

— Shrinking and expansion; these are in fact order filters.

— For a *finite* image, the digital convex hull; this is also an order filter, with each window $\varphi(p)$ equal to the whole point set S.

— Thinning operators, and more generally topology-preserving shrinking or expansion operators.

— Border marking operators, etc..

These filters can be extended to images with grey-levels in D for a larger set D, but this extension is unambiguous only for order filters (see Propositions 2 and 3 in Subsection II.3).

As mentioned in the Introduction, several order filters (for digital signals with values in an arbitrary set D) have been considered in the literature, mainly *rank filters* [9,10], the *weighted median filter* [4,12], or filters built by a composition of rank filters, for example the *separable median filter* [16], or Min-Max filters [7,15], which extend shrinking/expansion operators for binary images.

For non-binary images, preorder filters which are not order filters are less frequent. One example is the *mode filter*, which assigns to a point p a new signal value $Y(p)$ equal to the *mode* [13] of the signal values $X(q)$ for points q within the corresponding window $\varphi(p)$, that is the most frequent occurence among them. Another possible example is the following contrast-enhancing filter for signals on $\mathbf{R}^S$, defined by:

$$X^F(p) = \begin{cases} med(X(q) \mid q \in \varphi(p)) & \text{if } X(p) = med(X(q) \mid q \in \varphi(p)); \\ min(X(q) \mid q \in \varphi(p)) & \text{if } X(p) < med(X(q) \mid q \in \varphi(p)); \\ max(X(q) \mid q \in \varphi(p)) & \text{if } X(p) > med(X(q) \mid q \in \varphi(p)). \end{cases}$$

(This filter was already suggested in Subsection II.1).

Some customary assumptions are generally made about local filters. Given $D \subseteq \mathbf{Z}^d$, one postulates that a local filter is translation-invariant. In other words, the windows $\varphi(p)$ are translates of a fixed template φ^*, and the functions $f_D[p]$ are identical.

In fact, when the digital set S is finite, then for a point p near the border of S, the translate $\varphi(p)$ of φ^* by p will not be completely contained in S. The most frequent solution to this problem is to extend S to a larger digital set S' by adding to it one or more layers of points, and to extend a signal X on S to a signal X' on S' by assigning to each point $r \in S' - S$ a signal value $X'(r)$ equal to the signal value $X(q)$ of the closest point $q \in S$. For example, if S is a two-dimensional rectangular grid and if each $\varphi(p)$ is a 3×3 window centered about p, the window $\varphi(p)$ will not be contained in S when p is in the border of S; then we add a layer L of points around S, and each point in L gets as signal value the one of the neighboring point in S. This type of construction is compatible with our definition of a local filter. Indeed, for a point p in the border zone (in other words with $\varphi(p)$ extending outside S), one can consider that the corresponding window is in fact $\varphi(p) \cap S$, and that within this window, certain signal values $X(q)$ occur several times as arguments of the function $f_D[p]$. In other words one applies on $\varphi(p) \cap S$ a function $g_D[p]$ which is a weighted expansion (see the end of Subsection II.4) of $f_D[p]$. As we explained there, that expansion does not modify the property of being a selection, preorder, or order function, and so by taking the windows $\varphi(p) \cap S$ and the functions $g_D[p]$ instead of $\varphi(p)$ and $f_D[p]$, we do not change the fact that the local filter F_D is a selection filter, a preorder filter, or an order filter.

Let us stress that we will not make here such an assumption of translation-invariance, because we do not require it. We will allow local filters having windows $\varphi(p)$ of varying sizes and shape, and applying within them distinct functions $f_D[p]$. There are several justifications for this non-uniformity. We have seen above that the postulate of translation-invariance is not preserved along the border of a finite digital set. We will introduce later a variant of the local filter in the case of a finite digital set S, called a *recursive* filter, and we will show that it is equivalent to a non-recursive filter which is not translation-invariant. We will also see in [20] that one can design order filters for digital images having a heterogeneous behavior, varying between the distinct image portions.

Now that we have defined local, selection, preorder, and order filters, let us describe how the basic operations of composition and dual described in Subsection II.4 can be applied to them.

Given two filters F_D and G_D, the *composition* $F_D G_D$ *of* F_D *by* G_D is a local filter defined by $X^{F_D G_D} = (X^{F_D})^{G_D}$ for a signal $X \in D^S$. If F_D is determined by the windowing function φ and the functions $f_D[p]$ $(p \in S)$, and G_D by the windowing function ψ and the functions $g_D[p]$ $(p \in S)$, then $F_D G_D$ will have the windowing function τ defined by

$$\tau(p) = \bigcup_{q \in \psi(p)} \varphi(q) \qquad \text{for} \quad p \in S,$$

and it will apply within each window $\tau(p)$ the function

$$g_D[p] \circ \left[f_D[q_1], \ldots, f_D[q_r] \right], \qquad \text{where} \quad \{q_1, \ldots, q_r\} = \psi(p) \quad \text{and} \quad q_1 \prec \ldots \prec q_r,$$

in other words the composition of the functions $f_D[q]$ (for $q \in \psi(p)$) by $g_D[p]$. Thanks to Proposition 5, the composition of two order filters, preorder filters, or selection filters is an order filter, a preorder filter, or a selection filter respectively.

The composition of functions can occur in another way in the design of local filters. Suppose that we have n local filters $F_D^1, \ldots, F_D^n$ for signals on D^S, and a function $g_D : D^n \to D$. Then we define the *composition* $(F_D^1, \ldots, F_D^n) \cdot g_D$ *of* $F_D^1, \ldots, F_D^n$ *by* g_D as follows:

$$X^{(F_D^1, \ldots, F_D^n) \cdot g_D}(p) = g_D(X^{F_D^1}(p), \ldots, X^{F_D^n}(p)) \qquad (p \in S)$$

for any signal $X \in D^S$. Clearly if $F_D^1, \ldots, F_D^n$ are determined by the windowing functions $\varphi_1, \ldots, \varphi_n$ and the functions $f_D^1[p], \ldots, f_D^n[p]$ $(p \in S)$, then $(F_D^1, \ldots, F_D^n) \cdot g_D$ is a local filter with windowing function ψ defined by

$$\psi(p) = \varphi_1(p) \cup \cdots \cup \varphi_n(p) \qquad \text{for} \quad p \in S,$$

and it will apply within each window $\psi(p)$ the function

$$g_D[p] \circ \left[f_D^1[p], \ldots, f_D^n[p] \right],$$

in other words the composition of the functions $f_D^1[p], \ldots, f_D^n[p]$ by $g_D[p]$. Again Proposition 5 implies that if f_D is a selection function, a preorder function, or an order function, and $F_D^1, \ldots, F_D^n$ are selection filters, preorder filters, or order filters respectively, then $(F_D^1, \ldots, F_D^n) \cdot g_D$ is a selection filter, a preorder filter, or an order filter respectively. Examples of such a composition are given in [7] with f_D being the maximum or minimum.

Let us now extend to preorder filters the definition of the *dual* of a preorder function. Given the preorder filter F_D (on D^S) determined by the windowing function φ and the preorder functions $f_D[p]$ ($p \in S$), its dual F_D^* will be the preorder filter (on D^S) determined by the windowing function φ and the dual preorder functions $f_D[p]^*$ ($p \in S$). Proposition 4 implies that:

— The dual of an order filter is an order filter.

— A preorder filter is the dual of its dual: $F_D^{**} = F_D$.

— Taking the dual commutes with restriction: Given a preorder filter F_D for signals in D^S and a subset C of D, the dual F_C^* of the restriction F_C of F_D to signals in C^S is equal to the restriction to signals in C^S of the dual F_D^* of F_D.

By Proposition 5, the dual of a composition of preorder functions is the composition of their duals. Thus, considering the use of composition of functions in local filters, we have the following:

— The dual of the composition of two preorder filters is the composition of their duals: $(F_D G_D)^* = F_D^* G_D^*$.

— The dual of the composition of several preorder filters by a preorder function is the composition of their duals: $((F_D^1, \ldots, F_D^n) \cdot g_D)^* = ({F_D^1}^*, \ldots, {F_D^n}^*) \cdot g_D^*$.

What we said in Subsection II.4 about the dual of rank functions applies also to rank filters: the dual of the minimum filter is the maximum filter, the median filter is its own dual, etc..

We can also extend the equality (3). For any map $\eta : D \to D$ we derive the *signal value transformation* T_η, which is a map $D^S \to D^S$ defined by

$$X^{T_\eta}(p) = \eta(X(p)) \qquad \text{for} \quad p \in S \quad \text{and} \quad X \in D^S. \tag{32}$$

Then, given a strictly decreasing bijection $\omega : D \to D$, (3) implies that

$$F_D^* T_\omega = T_\omega F_D \tag{33}$$

for any preorder filter F_D on D^S. Note that when D is finite, such a map ω exists and is unique, it is the "natural" reversion of the elements of D.

Given a filter F_D and two signals X and Y in D^S, equation (33) says that for $X' = X^{T_\omega}$ and $Y' = Y^{T_\omega}$, $Y = X^{F_D^*}$ is equivalent to $Y' = X'^{F_D}$. This means that the behavior of the dual F_D^* corresponds to that of F_D by complementation. If F_D erodes ridges in an image, then F_D^* will fill valleys in that image, if F_D deletes positive impulse noise in a signal, then F_D^* will delete negative impulse noise in that signal, and so on.

We have defined local filters, selection, preorder, and order filters, and have given their major features. When S is finite, there exists a variant type of local filters: *recursive* filters.

They were introduced in [17] in the case of the median filter for one-dimensional signals, but they can be derived from any kind of local filter.

As defined above in (31), filters for digital signals are *non-recursive* (from the point of view of signal processing), or *parallel* (from the point of view of algorithms). This means that the new signal value $X^{F_D}(p)$ on a point p is obtained independently from the new signal values $X^{F_D}(q)$ obtained on the other points q. When S is *finite*, one can also derive from the windowing function φ and from the functions $f_D[p]$ another filter $\widehat{F}_D$, which is called a *recursive* filter (in the signal processing sense, see [17]), or a *sequential* filter (in the algorithmic sense). Recall the ordering of S by $\prec$; then $\widehat{F}_D$ works like F_D, except that for each $p \in S$, the new signal value $X^{\widehat{F}_D}(p)$ is computed *after* having replaced $X(q)$ by $X^{\widehat{F}_D}(q)$ for every $q \prec p$. Let us describe this more formally. We suppose that the elements of S are $s_1, \ldots, s_n$, where $s_1 \prec \ldots \prec s_n$. For any $j = 2, \ldots, n$, we set

$$\begin{aligned} \varphi^-(s_i) &= \{s_j \in \varphi(s_i) \mid j < i\}; \\ \varphi^+(s_i) &= \{s_j \in \varphi(s_i) \mid j \geq i\}. \end{aligned} \tag{34}$$

Then for a digital signal X on S, $X^{\widehat{F}_D}$ is built recursively as follows:

$$\begin{aligned} X^{\widehat{F}_D}(s_1) &= f_D[s_1](X(s_j) \mid s_j \in \varphi(s_1)); \\ X^{\widehat{F}_D}(s_i) &= f_D[s_i](X^{\widehat{F}_D}(s_{j^-}), X(s_{j^+}) \mid s_{j^-} \in \varphi^-(s_i), s_{j^+} \in \varphi^+(s_i)) \qquad (i = 2, \ldots, n). \end{aligned} \tag{35}$$

It is easy to see that such a filter can be achieved by the composition of n non-recursive local filters $F_D^1, \ldots, F_D^n$ defined as follows:

$$X^{F_D^i}(s_j) = \begin{cases} X(s_j) & \text{if } j \neq i; \\ f_D[s_i](X(q) \mid q \in \varphi(s_i)) & \text{if } j = i. \end{cases} \tag{36}$$

When the functions $f_D[s_i]$ $(i = 1, \ldots, n)$ are selection, preorder, or order functions, the filters F_D^i are selection, preorder, or order filters respectively. As the composition of local filters preserves these three properties, a recursive selection, preorder, or order filter is equivalent to a non-recursive one.

Note that in the non-recursive equivalent of a recursive local filter, the corresponding windows $\psi(s_i)$ increase with i; they satisfy the following recursive equality:

$$\begin{aligned} \psi(s_1) &= \varphi(s_1); \\ \psi(s_i) &= \varphi^+(s_i) \cup \bigcup_{q \in \varphi^-(s_i)} \psi(q) \qquad (i = 2, \ldots, n). \end{aligned} \tag{37}$$

This implies in particular that

$$\psi(s_i) \subseteq \bigcup_{1 \leq j \leq i} \varphi(s_j) \qquad (i = 1, \ldots, n).$$

For example, given $S = \{1, \ldots, n\}$, $s_i = i$ and $\varphi(s_i) = \{j \in S \mid i - k \leq j \leq i + k\}$, we have $\psi(s_i) = \{j \in S \mid 1 \leq j \leq i + k\}$.

Thus the windows $\psi(s_i)$ of the non-recursive equivalent of a recursive local filter will not have the same size. This is one of the reasons why we did not require translation-invariance, as it is generally done in the literature on rank filters.

Recursive filters can be useful in the case of one-dimensional digital signals (indeed many recursive linear filters have been applied in speech processing). It has been shown in [17] that one pass of a (translation-invariant) recursive median filter reduces a one-dimensional digital signal to a *root* signal, that is a signal invariant under further median filtering. On the other hand, the (translation-invariant) non-recursive median filter requires in general several passes to reduce a one-dimensional digital signal to a *root* [6].

In the case of multi-dimensional signals, the application of recursive filters is more problematic, because there is no natural ordering of a multi-dimensional digital space, as it is the case for a time sequence. The result of a non-recursive filter relies heavily on the choice of the ordering of the elements of the point set S, and introduces an anisotropy on it.

More details on the definition of order filters will be given in [20].

VI.2. Properties and characterizations

We will give below the interpretation in terms of local filters of the results of Sections IV and V. We will first introduce some general definitions (extending to filters those made in the Introduction).

For any $y \in D$, let C_y be the *constant y signal* defined by $C_y(p) = y$ for every $p \in S$. Given $X, Y \in D^S$, we write $X \leq Y$ if $X(p) \leq Y(p)$ for any $p \in S$. Given a local filter F_D, we say that F_D is *increasing* if for any $X, Y \in D^S$, $X \leq Y$ implies that $X^{F_D} \leq Y^{F_D}$, and that F_D is *decreasing* if $X \leq Y$ implies that $X^{F_D} \geq Y^{F_D}$; given another local filter G_D, we say that F_D commutes with G_D if $F_D G_D = G_D F_D$. We recall from (32) that every map $\eta : D \to D$ induces a signal value transformation T_η defined by $X^{T_\eta}(p) = \eta(X(p))$ for $X \in D^S$ and $p \in S$; clearly T_η is a local filter with trivial windows $\varphi(p) = \{p\}$, and applying within each of them the same function $f_D[p] = \eta$.

We can now state several properties of preorder and order filters. Let F_D and G_D be two local filters for signals in D^S. Theorems 13 and 14 have the following consequences:

Property 1. *F_D is an order filter iff it is an increasing preorder filter.*

Property 2. *Suppose that $D = \mathbf{R}$ and that F_D is a selection filter. Then the following three statements are equivalent:*

(i) F_D is an order filter.

(ii) F_D is continuous.

(*iii*) *For any* $X, Y \in D^S$ *and* $\epsilon \in \mathbf{R}$, $C_{-\epsilon} \leq X - Y \leq C_\epsilon$ *implies* $C_{-\epsilon} \leq X^{F_D} - Y^{F_D} \leq C_\epsilon$.

We now apply Propositions 16, 16', 17 and 17', and Corollaries 19 and 19':

Property 3. *If* F_D *is a preorder filter, then:*

(*i*) *For any strictly increasing function* $\eta : D \to D$, F_D *commutes with* T_η.

(*ii*) *For any strictly decreasing function* $\eta : D \to D$, $F_D T_\eta = T_\eta F_D^*$.

Property 4. *Suppose that* $D = \mathbf{R}$. *Then:*

(*i*) F_D *is a preorder filter iff for any strictly increasing function* $\eta : \mathbf{R} \to \mathbf{R}$, F_D *commutes with* T_η.

(*ii*) F_D *is a preorder filter and* $G_D = F_D^*$ *iff for any strictly decreasing function* $\eta : \mathbf{R} \to \mathbf{R}$, $F_D T_\eta = T_\eta G_D$.

Property 5. *If* F_D *is an order filter, then:*

(*i*) *For any increasing function* $\eta : D \to D$, F_D *commutes with* T_η.

(*ii*) *For any decreasing function* $\eta : D \to D$, $F_D T_\eta = T_\eta F_D^*$.

Property 6. *Suppose that* $|D| \geq 3$. *Then:*

(*i*) F_D *is an order filter iff for any increasing function* $\eta : D \to D$, F_D *commutes with* T_η.

(*ii*) F_D *is an order filter and* $G_D = F_D^*$ *iff for any decreasing function* $\eta : \mathbf{R} \to \mathbf{R}$, $F_D T_\eta = T_\eta G_D$.

We can define on signals the arithmetic operation of scalar multiplication: for any $X \in \mathbf{R}^S$ and $\epsilon \in \mathbf{R}$, we define ϵX by $(\epsilon X)(p) = \epsilon(X(p))$. Then Corollary 20 is translated in the following way:

Property 7. *If* F_D *is preorder filter, then for any* $a, b \in \mathbf{R}$ *and* $X \in \mathbf{R}^S$ *we have*

$$(aX + C_b)^{F_D} = \begin{cases} a(X^{F_D}) + C_b & \text{if } a \geq 0; \\ a(X^{F_D^*}) + C_b & \text{if } a \leq 0. \end{cases}$$

In particular if F_D *is equal to its dual, then* F_D *commutes with the map* $X \mapsto aX + C_b$.

For any $y \in \mathbf{R}$, we derive from the thresholding $\theta_y : \mathbf{R} \to B$ defined in (19) the digital signal thresholding $\Theta_y : \mathbf{R}^S \to B^S : X \mapsto \Theta_y(X)$ by setting for any signal $X \in \mathbf{R}^S$ and point $p \in S$:

$$\Theta_y(X)(p) = \theta_y(X(p)) = \begin{cases} 0 & \text{if } X(p) < y; \\ 1 & \text{if } X(p) \geq y. \end{cases} \tag{38}$$

We derive similarly from the complemented thresholding $\overline{\theta}_y : \mathbf{R} \to B$ defined by $\overline{\theta}_y(x) = \overline{\theta_y(x)}$ the digital signal thresholding $\overline{\Theta}_y : \mathbf{R}^S \to B^S$ defined by $\overline{\Theta}_y(X)(p) = \overline{\Theta_y(X)(p)}$. Then we derive the following from (20), (27), and from Theorems 18 and 18':

Property 8. *If F_D is an order filter, then:*

(*i*) *F_D commutes with thresholding: for any $y \in \mathbf{R}$ and $X \in D^S$,*

$$\Theta_y(X^{F_D}) = (\Theta_y(X))^{F_B}. \tag{39}$$

(*ii*) *For any $y \in \mathbf{R}$ and $X \in D^S$, $\overline{\Theta}_y(X^{F_D}) = (\overline{\Theta}_y(X))^{F_B^*}$.*

Property 9. *Suppose that $|D| \geq 3$. Let H_B be a local filter for signals in B^S. Then:*

(*i*) *F_D is an order filter and $H_B = F_B$ iff for any $y \in \mathbf{R}$ and $X \in D^S$, $\Theta_y(X^{F_D}) = (\Theta_y(X))^{H_B}$.*

(*ii*) *F_D is an order filter and $H_B = F_B^*$ iff for any $y \in \mathbf{R}$ and $X \in D^S$, $\overline{\Theta}_y(X^{F_D}) = (\overline{\Theta}_y(X))^{H_B}$.*

We can also define on signals the arithmetic operation of addition: for any $X, Y \in \mathbf{R}^S$, we define $X + Y$ by $(X + Y)(p) = X(p) + Y(p)$. Then Propositions 21 and 22 imply the following:

Property 10. *Suppose that $D = \{0, \ldots, m\}$ for some integer $m > 0$, and that F_D is an order filter. Take $Y, X \in D^S$ such that $Y = X^{F_D}$, and for any $j = 1, \ldots, m$, set $Y_j = \Theta_j(Y)$ and $X_j = \Theta_j(X)$. Then:*

(*i*) $X = \sum_{j=1}^{m} X_j$.

(*ii*) *For any $j = 1, \ldots, m$, $Y_j = X_j^{F_B}$.*

(*iii*) $Y = \sum_{j=1}^{m} Y_j$.

Property 11. *Suppose that $D = \{d_0, \ldots, d_m\}$ for some integer $m > 0$, with $d_0 < \ldots < d_m$, and that F_D is an order filter. For $j = 1, \ldots, m$, let $c_j = d_j - d_{j-1}$. Take $Y, X \in D^S$ such that $Y = X^{F_D}$, and for any $j = 1, \ldots, m$, set $Y_j = \Theta_{d_j}(Y)$ and $X_j = \Theta_{d_j}(X)$. Then:*

(*i*) $X = C_{d_0} + \sum_{j=1}^{m} c_j X_j$.

(*ii*) *For any $j = 1, \ldots, m$, $Y_j = X_j^{F_B}$.*

(*iii*) $Y = C_{d_0} + \sum_{j=1}^{m} c_j Y_j$.

VI.3. Consequences

As explained in Subsection VI.1, a recursive local, selection, preorder, or order filter is equivalent to a non-recursive one. Thus all the properties given above apply indifferently to both recursive and non-recursive local filters.

The fact that an order filter F_D is increasing (see Property 1) implies that if F_D deletes from a signal some feature, it will also delete any smaller features. Let us explain this with more details. We suppose that we have a nonnegative signal X (*i.e.*, $X \geq C_0$), which is added to a constant signal C_g, and the order filter F_D deletes X from $C_g + X$, in other words

$(C_g + X)^{F_D} = C_g$. Consider now a smaller nonnegative signal Y (*i.e.*, $X \geq Y \geq C_0$). Then $C_g \leq C_g + Y \leq C_g + X$ and as F_D is increasing, $C_g = C_g^{F_D} \leq (C_g + Y)^{F_D} \leq (C_g + X)^{F_D} = C_g$, that is $(C_g + Y)^{F_D} = C_g$, in other words F_D deletes Y from $C_g + Y$. The same thing happens if we suppose instead that $X \leq Y \leq C_0$.

In practice, the median filter for example is used to delete noisy peaks in a signal, and it will delete all peaks of size less than half the window size. Other order filters can be used to delete some particular features whose size is bounded from above.

On the other hand, a preorder filter which is not an order filter is not increasing. It can thus delete a feature and preserve a smaller one.

In our comments on Theorem 14 we stressed the importance of continuity for any practical method for data processing. The same can be said for Property 2. Order filters are the only continuous selection filters. Moreover item (*iii*) implies that:

— An order filter will not increase the quantization error of a digital signal (this is not the case, for example, with linear filters, because they require a new quantization of the result after each pass).

— Given a digital signal corrupted by an additive noise signal whose amplitude never exceeds some bound ϵ, the image of this noisy signal by an order filter is equal to the image of the original signal corrupted by an additive noise signal whose amplitude also never exceeds ϵ. In other words, for any $X, N \in D^S$ such that $|N(p)| \leq \epsilon$ for every $p \in S$, and for any order filter F_D on D^S, we have $(X + N)^{F_D} = X^{F_D} + N'$, where $|N'(p)| \leq \epsilon$ for every $p \in S$.

Properties 5 and 6 are important for digital image processing. Indeed one often enhances a digital picture by applying a (non-linear) increasing grey-level transformation to each pixel. The original picture and the enhanced one normally represent the same scene (and this is often the case for the human eye), and so a filter should extract the same features whether applied before or after the grey-level transformation. Now order filters commute with all increasing grey-level transformations; moreover, when there are more than two grey-levels, they are the only filters having that property. Thus an order filter can be applied either before or after an increasing grey-level transformation, giving the same result in each case.

By Properties 3 and 4, the same can be said with preorder filters and strictly increasing grey-level transformations in the case when we have a continuous range of grey-levels. The fact that human vision recognizes features in an image after some strictly increasing grey-level transformations indicates that preorder functions might well intervene in this process.

Commutation with thresholding (see Property 8) has an important consequence — both theoretical and practical — for the behavior of order filters. We explained at the beginning of Subsection IV.2 that it implies that the behavior of an order function f_D on D^n is determined by that of the corresponding order function f_B on B^n. Now the same can be said about order filters: the properties of an order filter can be determined by those of its restriction to binary signal, since a digital signal X is determined by all the binary signals $\Theta_y(X)$ obtained by thresholding it for every possible threshold value $y \in D$.

By Property 9 order filters are the only filters for non-binary digital signals which commute with thresholding.

Consider now threshold decomposition (Properties 10 and 11). Its functioning for order filters is the same as for order functions (see Subsection V.1). The input digital signal X is decomposed into a sum of binary signals X_j obtained by thresholding; each of these binary signals is processed separately by one or a composition of several order filters, and the resulting binary signals Y_j are added in order to give the output signal Y. In order to stress the importance of this method, we can simply quote the first paragraph of the conclusion of [5]:

> *"The threshold decomposition and the set of binary signals perform the same function for median filters that superposition and sinusoids perform for linear filters—they allow complex problems to be decomposed into simpler problems. This has very fortunate practical and theoretical consequences."*

Of course, in that statement one can replace "median filter" by "order filter". One of the consequences of threshold decomposition mentioned in the conclusion of [5] is the fact that the behavior of the median filter is determined by its behavior on binary images (this is just what we said above about order filters). In fact the results in [6,17] about the invariant signals and the convergence of recursive and non-recursive median filters for one-dimensional digital signals can be easily proved in the binary case, and then extended to the non-binary case by threshold decomposition.

In practice, threshold decomposition can be used for the computation of the image of a signal X by an order filter F_D when only the behavior of F_B is known. Thus one can design an order filter by specifying its behavior for binary signals (according to certain requirements), and then computing its behavior on non-binary signals by threshold decomposition. We will give some examples in [20].

Note that there is another method for the derivation of the behavior of an order filter on D^S from its behavior on B^S: in Subsection III.2 we showed how the min-max decompositions of an order function f_D on D^n can be determined from the behavior of f_B, and more precisely from the α-heavy sets for f_B ($\alpha = 0, 1$).

We proved in Subsection V.2 that threshold decomposition is the unique linear decomposition method for an order function. It is easy to extend the results proved there to the case of order filters.

Finally, let us remark that any order filter will have a behavior in some ways similar to that of a smoother (or a low-pass filter). We note indeed that in practice, order filters have been used for purposes of this kind: noise smoothing with median filters [19,22], erosion of narrow peak features by composition of Min and Max filters [7,15], etc..

Several arguments can be given to justify this comparison:

(a) As it is a selection filter, an order filter will leave invariant a non-zero constant signal. Now such a signal has frequency zero, and so it is preserved by low-pass linear filters, but not by band-pass or high-pass filters.

(b) With a linear filter, the output signal value $Y(p)$ on a point p is a linear combination of the input signal $X(q)$ for all points q within some neighborhood of p: $Y(p) = \sum_{q\in\varphi(p)} \lambda_{pq}X(q)$. Now:

- This filter is an increasing signal transformation (see Property 1) iff the coefficients λ_{pq} of these linear combinations are all non-negative.
- This filter preserves non-zero constant signals and satisfies item (*iii*) of Property 2 iff the coefficients λ_{pq} of these linear combinations are all non-negative and $\sum_{q\in\varphi(p)} \lambda_{pq} = 1$ for any $p \in S$.

In both cases the coefficients λ_{pq} are nonnegative, and then the linear filter is a smoother. Now an order filter is an increasing signal transformation, it preserves non-zero constant signals, and it satisfies item (*iii*) of Property 2; thus it looks like a smoother from this point of view.

There are nevertheless some differences between the smoothing properties of a linear low-pass filter and those of an order filter. Given a signal decomposed into a linear combination of sinusoid components of various frequencies, a low-pass filter suppresses all sinusoid components whose period is below a given threshold, and preserves the other ones. Thus features of the signal corresponding to high frequencies are deleted. On the other hand, given the threshold decomposition of that signal, an order filter will delete in each binary threshold layer small groups of ones or of zeroes (see our comment on Property 1 at the beginning of this Subsection); in other words it will erode small peaks or fill small holes.

An order filter will never enhance a feature, it will rather delete it. But one can use it for feature enhancement by subtracting it from the identity; in other words we can enhance certain features in a signal X by taking $X - X^{F_D}$ for an order filter. For example, in a two-dimensional image the composition of a Min filter followed by a Max filter will erode all narrow peaks and ridges [15]; thus these peaks and ridges in an image $X \in S^D$ can be displayed by taking $X - X^{Min_D Max_D}$.

Acknowledgement

The author is endebted to his colleague Ph. Delsarte for suggesting him the proofs of Theorem 6 and Corollary 7 given in this report (a more intricate proof was initially obtained with the help of Proposition 1).

References

[1] M.H. Abd-El Barr, S.G. Zaky, Z.G. Vranesic: Synthesis of Multivalued Multi-threshold Functions for CCD Implementation. *IEEE Trans. Computers*, Vol. C-35, no. 2, pp. 124–133, Feb. 1986.

[2] R.P. Agnew: Minimax Functions, Configurations, and Partitions. *J. Indian Math. Soc.*, Vol. 24, pp. 1–21, 1961.

[3] G. Birkhoff: *Lattice Theory.* American Mathematical Society Colloquium Publications, Vol. 25, 3rd edition (1983).

[4] D.R.K. Brownrigg: The Weighted Median Filter. *Comm. ACM*, Vol. 27, no. 8, pp. 807–818, Aug. 1984.

[5] J.P. Fitch, E.J. Coyle, N.C. Gallagher, Jr.: Median Filtering by Threshold Decomposition. *IEEE Trans. Acoustics, Speech, & Signal Processing*, Vol. ASSP-32, no. 6, pp. 1183–1188, Dec. 1984.

[6] N.C. Gallagher, Jr., G.L. Wise: A Theoretical Analysis of the Properties of Median Filters. *IEEE Trans. Acoustics, Speech, & Signal Processing*, Vol. ASSP-29, no. 6, pp. 1136–1141, Dec. 1981.

[7] V. Goetcherian: From Binary to Grey Tone Image Processing Using Fuzzy Logic Concepts. *Pattern Recognition*, Vol. 12, no. 1, pp. 7–15, 1980.

[8] M.A. Harrison: *Introduction to Switching and Automata Theory.* McGraw-Hill (1965).

[9] G. Heygster: Rank Filters in Digital Image Processing. *Computer Graphics & Image Processing*, Vol. 19, no. 2, pp. 148–164, Jun. 1982.

[10] R.M. Hodgson, D.G. Bailey, M.J. Naylor, A.L.M. Ng, S.J. McNeill: Properties, Implementations and Applications of Rank Filters. *Image and Vision Computing*, Vol. 3, no. 1, pp. 3–14, Feb. 1985.

[11] M. Hollander, D.A. Wolfe: *Nonparametric Statistical Methods.* Wiley (1973).

[12] B.I. Justusson: Median Filtering: Statistical Properties. *Two-Dimensional Digital Signal Processing II. Transforms and Median Filters* (T.S. Huang, ed.), Topics in Applied Physics Vol. 43, Springer-Verlag (1981), pp. 161–196.

[13] C. Leach: *Introduction to Statistics. A Nonparametric Approach for the Social Sciences.* Wiley (1979).

[14] R.W. Murphy: A Positive-Integer Arithmetic for Data Processing. *IBM J. Res. & Dev.*, Vol. 1, no. 2, pp. 158–170, Apr. 1957.

[15] Y. Nakagawa, A. Rosenfeld: A Note on the Use of Local Min and Max Operations in Digital Picture Processing. *IEEE Trans. Systems, Man, & Cybernetics*, Vol. SMC-8, no. 8, pp. 632–635, Aug. 1978.

[16] P.M. Narendra: A Separable Median Filter for Image Noise Smoothing. *IEEE Trans. Pattern Analysis & Machine Intelligence*, Vol. PAMI-3, no. 1, pp. 20–29, Jan. 1981.

[17] T.A. Nodes, N.C. Gallagher, Jr.: Median Filters: Some Modifications and their Properties. *IEEE Trans. Acoustics, Speech, & Signal Processing*, Vol. ASSP-30, no. 5, pp. 739–746, Oct. 1982.

[18] S.K. Pal, R.A. King: On Edge Detection of X-Ray Images Using Fuzzy Sets. *IEEE Trans. Pattern Analysis & Machine Intelligence*, Vol. PAMI-5, no. 1, pp. 69–77, Jan. 1983.

[19] L.R. Rabiner, M.R. Sambur, C.E. Schmidt: Applications of a Nonlinear Smoothing Algorithm to Speech Processing. *IEEE Trans. Acoustics, Speech, & Signal Processing*, Vol. ASSP-23, no. 6, pp. 552–557, Dec. 1975.

[20] C. Ronse: Order Filters. *Philips Research Laboratory Brussels Manuscript in preparation.*

[21] C. Ronse: Threshold-Commuting Fuzzy Set Operations. *Philips Research Laboratory Brussels Manuscript in preparation.*

[22] J.W. Tukey: *Exploratory Data Analysis.* Addison-Wesley (1977).

[23] S.G. Tyan: Median Filtering: Deterministic Properties. *Two-Dimensional Digital Signal Processing II. Transforms and Median Filters* (T.S. Huang, ed.), Topics in Applied Physics Vol. 43, Springer-Verlag (1981), pp. 197–217.