Philips Research Laboratory Brussels
Av. E. Van Becelaere 2, Box 8
B-1170 Brussels, Belgium

Report R500

Order Functions: Mathematical Properties and Applications to Digital Signal Processing

Christian Ronse

March 1986

Abstract: *A function f in n variables is called an* order function *if for any $x_1, \ldots, x_n$ such that $x_{i_1} \leq \ldots \leq x_{i_n}$ we have $f(x_1, \ldots, x_n) = x_t$, where t is determined by the n-tuple $(i_1, \ldots, i_n)$ corresponding to that ordering $x_{i_1} \leq \ldots \leq x_{i_n}$. Equivalently, it is a function built as a minimum of maxima, or a maximum of minima. Well-known examples are the minimum, the maximum, the median and more generally rank functions, or the composition of rank functions.*

In this report we study the mathematical properties of order functions and give several characterization theorems for them. We give then an interpretation of these properties in terms order filters, *that is, digital signal processing local filters based on order functions (such as the median or separable median filter, Min-Max filters, or rank filters).*

Keywords: order statistics, preorder functions, order functions, thresholding, threshold decomposition, digital signal processing.

ERRATA TO R500

— Page 31, line 9 (proof of Proposition 17):
"Take $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$" instead of "Take $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_n)$".

— Page 34, line 15 (after (25)):
"commutes with the pair" instead of "commutes the pair".

— Page 34, lines −14, −10, and −6, and Page 36, line 2 (proof of Proposition 16' and statement of Propositions 17' and 19'):
"decreasing" instead of "increasing".

— Page 38, line 6:
"we apply f_B" instead of "we apply f_D".

— Page 46, line −6:
"$g_D \circ [f_D^1[p], \ldots, f_D^n[p]]$" instead of "$g_D[p] \circ [f_D^1[p], \ldots, f_D^n[p]]$".

— Page 46, line −5:
"g_D" instead of "$g_D[p]$".

— Page 46, line −1:
"g_D" instead of "f_D".

— Page 50, line 14 (Property 6):
"$\eta : D \to D$" instead of "$\eta : \mathbf{R} \to \mathbf{R}$".

— Page 52, line −7:
"in the case where" instead of "in the case when".

— Page 53, line 2:
"restriction to binary signals" instead of "restriction to binary signal".

I. Introduction

Anyone working in non-parametric statistics or in digital signal processing has frequently met *rank functions*; such a function r_k in n variables selects the k-th smallest of its arguments as result:

$$r_k(x_1, \ldots, x_n) = x_{i_k} \qquad \text{if} \quad x_{i_1} \leq \ldots \leq x_{i_n}.$$

For $k = 1$ this is the minimum, while for $k = n$ this is the maximum. When n is odd and $k = (n+1)/2$, we have the well-known *median*.

In the same way as the median has been considered as an alternative to the mean in non-parametric statistics (see [11], p. 46, and [13], p. 11), in digital signal processing one has proposed to use smoothers based on running medians instead of running averages [22]. A signal filter using running medians is called a *median filter*. *Rank filters*, in other words filters based on rank functions, have also been applied in image processing (see in particular the recent survey [10]).

A rank function has the following feature: its result is not computed by an arithmetical function of its arguments; it is rather selected among these arguments, in function of their ordering. There are other functions possessing that same feature. We give here three examples of such functions.

Consider first the following simple generalization of rank functions, called *weighted rank functions*. Suppose that to each $i = 1, \ldots, n$ one associates a non-negative integer weight w_i. Then the k-th *weighted rank function* determined by the weights $w_1, \ldots, w_n$ is obtained by applying the ordinary k-th rank function r_k to a sample of $w_T = w_1 + \cdots + w_n$ variables containing w_i copies of x_i for each $i = 1, \ldots, n$. When w_T is odd and $k = (w_T + 1)/2$, we get the *weighted median* defined in [4,12].

The weighted median can be taken as an alternative to the weighted average in the design of smoothers for digital signals. Take for example a two-dimensional digital image. Then we can smooth it with a weighted median filter defined as follows. To each point we associate a 3×3 window centered about it; then in the image we replace the grey-level of each point by the weighted median of the grey-levels of the 9 points of the corresponding window, where the weights can be repartited within the window as follows:

1	2	1
2	3	2
1	2	1

Our second example is the following n^2 variable function:

$$med(med(x_{1,1}, \ldots, x_{1,n}), \ldots, med(x_{n,1}, \ldots, x_{n,n})).$$

It intervenes in the two-dimensional *separable median filter* [16], which consists in the succession of two one-dimensional median filterings, one on the rows and one on the columns of a two-dimensional image. It is not hard to see that this function is not a weighted rank function.

Our third example is given by the following operator for eroding narrow peaks and ridges in a two-dimensional digital image. It is obtained by the composition of a min filter followed by a max filter, and it is described in [7,15]. A point can be labelled with integer coordinates (i,j), and its grey-level is then written $x(i,j)$. Now the operator replaces each grey-level $x(i,j)$ by a new grey-level $y(i,j)$ obtained as follows: within each $d \times d$ square S containing (i,j), take the minimum $m(S)$ of the grey-levels $x(i',j')$ for all points (i',j') in S; then $y(i,j)$ is the maximum among all these minima $m(S)$. In other words, we have

$$y(i,j) = \max_{-d \leq u,v \leq 0} \left\{ \min_{0 \leq a,b \leq d} \{x(i+u+a, j+v+b)\} \right\}.$$

This means that for a grey-level value g, $y(i,j) \geq g$ if and only if there is a $d \times d$ square S containing (i,j), whose points (i',j') have all a grey-level $x(i',j') \geq g$. Thus peaks and ridges too narrow to contain a $d \times d$ square are eroded in the new image.

The functions described above have a common flavor, and they can be charaterized in two ways.

First, the result of such a function equals one of its arguments, and the choice of it depends only upon the ordering of these arguments. In other words, assuming that such function is in n variables, to every ordered n-tuple $(i_1, \ldots, i_n)$ obtained by a permutation of $(1, \ldots, n)$ one can associate an integer $t = \chi(i_1, \ldots, i_n)$ with $1 \leq t \leq n$, such that for any $x_1, \ldots, x_n$, $x_{i_1} \leq \ldots \leq x_{i_n}$ implies that $f(x_1, \ldots, x_n) = x_t$ for $t = \chi(i_1, \ldots, i_n)$. For example, $\chi(i_1, \ldots, i_n) = i_k$ for the k-th rank function.

Second, they can be obtained by a composition of the minimum and maximum functions only. For example, the k-th rank function is obtained by taking the minimum among all maxima of k-tuples of variables.

As we will show in Section III, these two properties are equivalent. We will call such functions *order functions* (because of their relations to the order statistics among their variables). The purpose of this report is the study of the fundamental mathematical properties of such functions. As can be seen from some of the examples given above, our motivation comes primarily from image processing, and we will indeed apply our results to digital signal processing operators built from order functions, what we call *order filters*.

There is a strong argument justifying a fundamental mathematical study of order functions. While many order filters have been built by heuristic methods, not much is known about their nature, their behavior, and even the reasons guiding the choice of a particular type of filter. The deepest theoretical results about them are mostly of a descriptive (see for example [10,23]) or statistical (see for example [12]) nature, they study convergence properties of certain types of rank filters (see for example [6,17]), or they relate order filters to their counterparts for binary digital signals (see for example [5,7,15]). Our aim is to give the first rudiments of a general theoretical understanding of order functions, and so to allow practitioners in digital signal processing to understand the nature of the order filters that they will use, and in particular to choose them knowingly.

The report is organized as follows:

In Section II we give an axiomatic definition of *order functions*, and of a wider class of functions that we call *preorder functions*. We characterize the possible values that order functions may take. We study also the *dual* of an order function, based on the duality between the two order relations $\leq$ and $\geq$, and the *composition* of order functions.

In Section III we show the equivalence between the two definitions of order functions given above and we study the possible min-max decompositions of an order function.

In Section IV we give several mathematical characterizations of order functions, relating to continuity, commutation with thresholding, etc..

In Section V we study the "threshold decomposition" method which was devised in [5] for median filters, but is still valid for order functions and order filters.

In Section VI we define the *order filters* corresponding to order functions. We give a practical interpretation in terms order filters of the properties of order functions stated in Sections IV and V.

A more detailed study of order filters, going beyond the mere application of properties of order functions, will be made in a second paper [20].

In a third paper [21], we will show that order functions relate also to fuzzy set theory. The link between fuzzy set theory and minimum/maximum filters was investigated in [7,15] and used in [18]; we will extend this link to order functions. We will explain why certain set operations, such as the union or the intersection of a finite number of sets, or the convex hull of a finite Euclidean set, etc., admit a unique fuzzy extension which commutes with thresholding. It is so because these operations are increasing.

General definitions and notation

We will recall here a few definitions from the theory of functions and introduce a particular notation that will be used throughout the report.

For a function f defined on a set X and a set $Y \supset X$, a function g defined on Y will be called an *extension of f to Y* if f is the restriction of g to X.

Given a set X and two functions $f: X^m \to X$ (where $m \geq 1$) and $g: X \to X$, we will say that f *commutes* with g if for every $x_1, \ldots, x_m \in X$ we have

$$g(f(x_1, \ldots, x_m)) = f(g(x_1), \ldots, g(x_m)).$$

Write $\mathbf{R}$ for the set of real numbers. Let E be a subset of $\mathbf{R}$. Let us recall the definition of *monotonous* functions. Given a function $f: E^m \to E$ (where $m \geq 1$), we will say that

— f is *increasing* if for any $x_1, \ldots, x_m, y_1, \ldots, y_m \in E$ such that $x_i \leq y_i$ for each $i = 1, \ldots, m$, we have $f(x_1, \ldots, x_m) \leq f(y_1, \ldots, y_m)$; and that

— f is *decreasing* if for any $x_1, \ldots, x_m, y_1, \ldots, y_m \in E$ such that $x_i \leq y_i$ for each $i = 1, \ldots, m$, we have $f(x_1, \ldots, x_m) \geq f(y_1, \ldots, y_m)$.

Now for a function $g: E \to E$, we will say that

— g is *strictly increasing* if for any $x, y \in E$, $x < y$ implies that $g(x) < g(y)$; and that

— g is *strictly decreasing* if for any $x, y \in E$, $x < y$ implies that $g(x) > g(y)$.

We consider a subset D of $\mathbf{R}$ of size > 1, and an integer $n > 1$. We will study order functions in n variables over D. In practice, D will be a set of possible signal values (for example $\{0, \ldots, 255\}$), or a set of values in a discrete statistical sample, while n will be the size of a window in an order filter for digital signals, or the size of a discrete statistical sample. Let B be the binary set $\{0, 1\}$. Two interesting particular cases will be $D = \mathbf{R}$ and $D = B$. However, in our mathematical analysis, we will make no particular restriction on D and n, except that $|D|, n > 1$.

Let $I_n = \{1, \ldots, n\}$ and let $\mathcal{T}_n$ be the set of ordered n-tuples $(i_1, \ldots, i_n)$ such that $\{i_1, \ldots, i_n\} = I_n$, in other words the set of ordered n-tuples $(i_1, \ldots, i_n)$ obtained by a permutation of $(1, \ldots, n)$. The elements $(i_1, \ldots, i_n)$ of $\mathcal{T}_n$ will correspond to the possible orderings $x_{i_1} \leq \ldots \leq x_{i_n}$ of the arguments of an order function.

II. Axiomatic definition of preorder and order functions

We will define *order functions*, and a generalization of them called *preorder functions*, as functions selecting one of their arguments as result, where that selection is determined by the ordering relations between these arguments. In Subsection II.1 we give an intuitive definition for them. In Subsection II.2 this is made more formal; we show also that the definition of an order function implies some constraints on the possible values that it may take, and we characterize them. This subsection contains mainly mathematical details, and so it can be skipped by readers interested only in practical applications of order functions. In Subsection II.3 we show that the domain of variables of an order function can be extended to $\mathbf{R}$. Then in Subsection II.4 we use the duality between the relations $\leq$ and $\geq$ in order to define the *dual* of a preorder or order function, and we consider the *composition* of preorder or order functions, and related constructions.

II.1. Selection, preorder, and order functions: an intuitive view

As explained at the end of the Introduction, we will consider a certain type of functions in n variables over a set D (where $n, |D| > 1$). As we will later consider certain particular cases for D (in particular, $D = \{0,1\}$), we will write these functions f_D, g_D, h_D, etc., in order to avoid any confusion.

The first property of such a function is that its result is always equal to one of its arguments. It is thus selected among these arguments by some specific rule. We call such a function a *selection function*. More precisely, we state the following:

Definition 1. Let f_D be a function $D^n \to D$. Then f_D is called a *selection function* if for any $x_1, \ldots, x_n \in D$, we have $f_D(x_1, \ldots, x_n) \in \{x_1, \ldots, x_n\}$.

This concept corresponds to the operation of selecting a value inside a sample of n values in D. The choice of that particular value must be made according to some criterion. Here we will consider that this criterion is based on the ordering of the samples. There are two ways to express this.

Recall the set $\mathcal{I}_n = \{1, \ldots, n\}$ and the set $\mathcal{T}_n$ of all n-tuples $(i_1, \ldots, i_n)$ obtained by a permutation of $(1, \ldots, n)$.

Consider a selection function $f_D : D^n \to D$. We require that for $x_1, \ldots, x_n \in D$, $f_D(x_1, \ldots, x_n)$ is chosen among $x_1, \ldots, x_n$ in function of their ordering. This ordering may be understood in two ways:

(1°) It is the set of all ordered pairs (i, j), where $i, j \in \mathcal{I}_n$, such that $x_i < x_j$.

(2°) It is a chain of the form $x_{i_1} \leq \ldots \leq x_{i_n}$, where $(i_1, \ldots, i_n) \in \mathcal{T}_n$.

In fact these two concepts are distinct, and (1°) leads to a wider class of functions than (2°). Indeed, given a chain $x_{i_1} \leq \ldots \leq x_{i_n}$, the set of ordered pairs (i, j) such that $x_i < x_j$ is completely determined if we specify for each $k = 1, \ldots, n-1$ whether we have $x_{i_k} = x_{i_{k+1}}$

or $x_{i_k} < x_{i_{k+1}}$. Now in (2°) the distinction between these two cases is irrelevant, and they both lead to the same result for $f_D(x_1,\ldots,x_n)$, while in (1°) this distinction matters, and the two cases can lead to distinct results for $f_D(x_1,\ldots,x_n)$.

As an example, consider the following function (with n odd):

$$f_D(x_1,\ldots,x_n) = \begin{cases} med(x_1,\ldots,x_n) & \text{if } x_1 = med(x_1,\ldots,x_n); \\ min(x_1,\ldots,x_n) & \text{if } x_1 < med(x_1,\ldots,x_n); \\ max(x_1,\ldots,x_n) & \text{if } x_1 > med(x_1,\ldots,x_n). \end{cases}$$

(It can be used for the design of a contrast-enhancing filter for digital images (with D being the set of greylevels): to each point p we associate a window $\varphi(p)$ of size n, and we apply f_D to the greylevels of the points of $\varphi(p)$, where x_1 corresponds to the reference point p.) Suppose that we have $x_{i_1} \leq \ldots \leq x_{i_n}$, where $i_{(n-1)/2} = 1$ (in other words x_1 is just before the median in the ordering of the variables x_i). Then we may have either $x_1 = x_{i_{(n+1)/2}} = med(x_1,\ldots,x_n)$ or $x_1 < med(x_1,\ldots,x_n)$. The two cases lead to two distinct results for $f_D(x_1,\ldots,x_n)$, namely $min(x_1,\ldots,x_n)$ and $med(x_1,\ldots,x_n)$. Thus here the result of f_D depends upon the ordering of the variables as in (1°).

On the other hand, the result of a rank function depends upon the ordering of the variables as in (2°).

As (2°) is more restrictive than (1°), (1°) gives rise to a wider class of functions. For the purpose of this report, we will call them *preorder functions*, while functions determined by (2°) will be called *order functions*. We have thus the following two definitions:

Definition 2. Let f_D be a function $D^n \to D$. Then f_D is called a *preorder function* if for any $x_1,\ldots,x_n \in D$, there is some $t \in I_n$ which is chosen as a function of the set of ordered pairs (i,j) for which $x_i < x_j$, such that we have $f_D(x_1,\ldots,x_n) = x_t$.

Definition 3. Let f_D be a function $D^n \to D$. Then f_D is called an *order function* if to every $(i_1,\ldots,i_n) \in \mathcal{T}_n$ one can associate an integer $t = \chi(i_1,\ldots,i_n) \in I_n$, such that for any $x_1,\ldots,x_n \in D$, $x_{i_1} \leq \ldots \leq x_{i_n}$ implies that $f_D(x_1,\ldots,x_n) = x_t$ for $t = \chi(i_1,\ldots,i_n)$.

It is clear that an order function is a preorder function (because the set of pairs (i,j) such that $x_i < x_j$ determines the set of all orderings of the form $x_{i_1} \leq \ldots \leq x_{i_n}$), and that order functions and preorder functions are selection functions.

The distinction between the two orderings (1°) and (2°) has been considered in [14]: a chain $x_{i_1} \leq \ldots \leq x_{i_n}$ (determining the ordering as in (2°)) was called a *configuration*, while a chain $x_{i_1} * \ldots * x_{i_n}$ (where each $*$ is either $=$ or $<$, determining thus the ordering as in (1°)) was called a *sub-configuration*. Preorder functions were called *configuration functions*, and order functions were identified with continuous configuration functions (it is indeed easy to see that a preorder function is an order function iff it is continuous).

For an order function f_D, the map χ associating to each $(i_1,\ldots,i_n) \in \mathcal{T}_n$ the value $\chi(i_1,\ldots,i_n) \in I_n$ will be called the *choice map* of f_D. It is clear that the behavior of f_D is determined by its choice map.

The functions described in the Introduction (*i.e.*, rank functions, weighted rank functions, compositions of rank functions, etc.) are order functions. For example, with the k-th rank function r_k we have $\chi(i_1, \ldots, i_n) = i_k$. With the k-th weighted rank function determined by the weights $w_1, \ldots, w_n$, we have $\chi(i_1, \ldots, i_n) = i_t$ if $\sum_{j<t} w_{i_j} < k \leq \sum_{j\leq t} w_{i_j}$.

In an order function f_D the choice of $t = \chi(i_1, \ldots, i_n)$ may not be arbitrary, because f_D must be well-defined. Take for example $n = 3$, and let $s = \chi(1, 2, 3)$ and $t = \chi(2, 1, 3)$. Given $x_1 = x_2 < x_3$, we have both $x_1 \leq x_2 \leq x_3$ and $x_2 \leq x_1 \leq x_3$; thus $f(x_1, x_2, x_3) = x_s = x_t$, and so we may not have for example $s = 3$ and $t = 2$, since $x_3 \neq x_2$. The restrictions on the possible choices $t = \chi(i_1, \ldots, i_n)$ will be characterized in the next subsection, where we will also give a more formal version of Definitions 2 and 3. The most important thing to remember is that the constraints to be satisfied by the choice map χ of an order function on D^n are independent of the set D. Readers uninterested in mathematical formalism may skip that subsection and resume the reading in Subsection II.3.

II.2. Formal characterization of preorder and order functions

We will give here a more formal definition of preorder and order functions. Then we will characterize the constraints that must be satisfied by the choices $t = \chi(i_1, \ldots, i_n)$ for $(i_1, \ldots, i_n) \in \mathcal{T}_n$.

Let us give some precisions on the selection of t in Definition 2. The set of pairs (i, j) such that $x_i < x_j$ is not arbitrary. Let $y_1, \ldots, y_m$ be the distinct values taken by $x_1, \ldots, x_n$, where $y_1 < \ldots < y_m$. The only constraints on the integer m are that $m \geq 1$, $m \leq n$ and $m \leq |D|$. For each $j = 1, \ldots, m$, let $P_j = \{i \in \mathcal{I}_n \mid x_i = y_j\}$. Then the sets $P_1, \ldots, P_m$ form a *partition* of $\mathcal{I}_n$, in other words they are non-void, pairwise disjoint, and their union is equal to $\mathcal{I}_n$. The ordered m-tuple $(P_1, \ldots, P_m)$ will be called an *ordered partition* of $\mathcal{I}_n$, and m will be its *length*; as it is determined by the values of $x_1, \ldots, x_n$, we will say that it is *induced by* $x_1, \ldots, x_n$.

Clearly, any ordered partition of $\mathcal{I}_n$ having length m, where $1 \leq m \leq \min\{n, |D|\}$, is induced by some $x_1, \ldots, x_n \in D$. Now it is obvious that the ordered partition $(P_1, \ldots, P_m)$ induced by $x_1, \ldots, x_n$ characterizes the set of ordered pairs (i, j) such that $x_i < x_j$, because for $i \in P_a$ and $j \in P_b$, we have $x_i < x_j$ iff $a < b$. Thus, given that ordered partition induced by $x_1, \ldots, x_n$, there is some $t \in \mathcal{I}_n$ such that $f_D(x_1, \ldots, x_n) = x_t$, and the choice of t is determined by that ordered partition; now if $t \in P_u$, then we have $f_D(x_1, \ldots, x_n) = x_s$ for any other $s \in P_u$. In other words, to each ordered partition $(P_1, \ldots, P_m)$ (with $1 \leq m \leq \min\{n, |D|\}$) corresponds one of its members P_u, such that if $x_1, \ldots, x_n$ induce that ordered partition, then $f_D(x_1, \ldots, x_n) = x_s$ for any $s \in P_u$. Hence we can give a new expression of Definition 2 as follows:

Definition 2'. Let f_D be a function $D^n \to D$. Let $s = \min\{n, |D|\}$. Then f_D is called a *preorder function* if there is a map σ associating to each ordered partition $(P_1, \ldots, P_m)$ of $\mathcal{I}_n$, where the length m satisfies $1 \leq m \leq s$, one of its members $P_u = \sigma(P_1, \ldots, P_m)$, and such that for any $x_1, \ldots, x_n \in D$, $f_D(x_1, \ldots, x_n) = x_t$ for $t \in \sigma(Q_1, \ldots, Q_h)$, where $(Q_1, \ldots, Q_h)$ is the ordered partition induced by $x_1, \ldots, x_n$.

It is clear that any such map σ corresponds to a preorder function, and that the correspondence between preorder functions $D^n \to D$ and such maps σ is one-to-one.

The formalization of Definition 3 is easier. We have only to remark that the choice $t = \chi(i_1, \ldots, i_n)$ for each $(i_1, \ldots, i_n) \in \mathcal{T}_n$ is given by a map $\chi: \mathcal{T}_n \to \mathcal{I}_n$. Hence we make the following:

Definition 3'. Let f_D be a function $D^n \to D$. Then f_D is called an *order function* if there is a map $\chi: \mathcal{T}_n \to \mathcal{I}_n$ such that for any $x_1, \ldots, x_n \in D$, $x_{i_1} \leq \ldots \leq x_{i_n}$ (with $(i_1, \ldots, i_n) \in \mathcal{T}_n$) implies that $f_D(x_1, \ldots, x_n) = x_t$ for $t = \chi(i_1, \ldots, i_n)$. The map χ is then called the *choice map* of f_D.

Now the correspondence between order functions $D^n \to D$ and their choice maps is also one-to-one, but it must be stressed that not every every map $\chi: \mathcal{T}_n \to \mathcal{I}_n$ is the choice map of an order function. We gave an example for $n = 3$ at the end of the previous subsection. More generally, when $x_a = x_b$ for $a \neq b$, there exist two distinct $(i_1, \ldots, i_n)$ and $(j_1, \ldots, j_n) \in \mathcal{T}_n$ such that $x_{i_1} \leq \ldots \leq x_{i_n}$ and $x_{j_1} \leq \ldots \leq x_{j_n}$; then for $c = \chi(i_1, \ldots, i_n)$ and $d = \chi(j_1, \ldots, j_n)$, we must have $x_c = x_d = f_D(x_1, \ldots, x_n)$ if χ is the choice map of an order function f_D, and so χ may not be arbitrary in this case.

Choice maps are determined by the following criterion:

Proposition 1. *Let χ be a map $\mathcal{T}_n \to \mathcal{I}_n$. Then the following three statements are equivalent:*

(i) χ is the choice map of an order function $f_D: D^n \to D$.

(ii) Given $(i_1, \ldots, i_n)$ and $(j_1, \ldots, j_n) \in \mathcal{T}_n$ such that for some $a, b \in \mathcal{I}_n$ with $a \leq b$ we have

$$\{i_k \mid k < a\} = \{j_k \mid k < a\} \quad \text{and}$$
$$\{i_k \mid a \leq k \leq b\} = \{j_k \mid a \leq k \leq b\} = M,$$

then $\chi(i_1, \ldots, i_n) \in M$ implies that $\chi(j_1, \ldots, j_n) \in M$.

(iii) Given $(u_1, \ldots, u_n)$ and $(v_1, \ldots, v_n) \in \mathcal{T}_n$ such that for some $k \in \{1, \ldots, n-1\}$ we have $v_k = u_{k+1}$, $v_{k+1} = u_k$, and $v_r = u_r$ for $r \neq k, k+1$, then $\chi(u_1, \ldots, u_n) = \chi(v_1, \ldots, v_n)$ or $\{\chi(u_1, \ldots, u_n), \chi(v_1, \ldots, v_n)\} = \{u_k, u_{k+1}\}$.

Proof. *(a) (i) implies (iii).*
Take $(u_1, \ldots, u_n)$ and $(v_1, \ldots, v_n)$ as in the statement of *(iii)*. We can write $\chi(u_1, \ldots, u_n) = u_t$ and $\chi(v_1, \ldots, v_n) = u_{t'}$, where $t, t' \in \mathcal{I}_n$. Suppose that $t \neq t'$; without loss of generality, we can assume that $t < t'$ (otherwise we invert t and t' in the following argument). As

$|D| \geq 2$, we can take $r, s \in D$ such that $r < s$. If $t < k$ we define $x_1, \ldots, x_n \in D$ as follows:

$$x_{u_j} = \begin{cases} r & \text{if } j \leq t; \\ s & \text{if } j > t. \end{cases}$$

As $t < k$ we have $x_{u_1} \leq \ldots \leq x_{u_k} = x_{u_{k+1}} \leq \ldots \leq x_{u_n}$, and so this means that $x_{u_1} \leq \ldots \leq x_{u_n}$ and $x_{v_1} \leq \ldots \leq x_{v_n}$. As χ is the choice map of f_D, we have thus $r = x_t = f_D(x_1, \ldots, x_n) = x_{t'} = s$, a contradiction. If $t' > k+1$ we define $x_1, \ldots, x_n \in D$ as follows:

$$x_{u_j} = \begin{cases} r & \text{if } j < t'; \\ s & \text{if } j \geq t'. \end{cases}$$

We get then the same contradiction $r = x_t = f_D(x_1, \ldots, x_n) = x_{t'} = s$. Thus $k \leq t < t' \leq k+1$, and so $\{u_t, u_{t'}\} = \{u_k, u_{k+1}\}$.

(b) (iii) implies (ii).
Take a, b and $(i_1, \ldots, i_n)$ as in *(ii)*. Let S be the set of all $(j_1, \ldots, j_n) \in \mathcal{T}_n$ such that

$$\{i_k \mid k < a\} = \{j_k \mid k < a\} \quad \text{and}$$
$$\{i_k \mid a \leq k \leq b\} = \{j_k \mid a \leq k \leq b\} = M.$$

We have then also

$$\{i_k \mid b < k\} = \{j_k \mid b < k\}.$$

We must show that if $\chi(i_1, \ldots, i_n) \in M$, then $\chi(j_1, \ldots, j_n) \in M$ for every $(j_1, \ldots, j_n) \in S$. It is clear that any element of S can be obtained from $(i_1, \ldots, i_n)$ by applying to it three permutations of the positions of its entries i_k: one of the entries i_k with $k < a$, one of those with $a \leq k \leq b$, and one of those with $b < k$. Each one of these 3 permutations can be decomposed as a succession of transpositions inverting neighboring entries, *i.e.*, transformations of S of the form

$$(u_1, \ldots, u_n) = (\ldots, u_k, u_{k+1}, \ldots) \;\mapsto\; (v_1, \ldots, v_n) = (\ldots, u_{k+1}, u_k, \ldots),$$

where either $k, k+1 < a$, or $a \leq k, k+1 \leq b$, or $b < k, k+1$. This is the transformation from $(u_1, \ldots, u_n)$ to $(v_1, \ldots, v_n)$ in *(ii)*. Thus either

$$\chi(v_1, \ldots, v_n) = \chi(u_1, \ldots, u_n), \qquad \text{or}$$
$$\{\chi(u_1, \ldots, u_n), \chi(v_1, \ldots, v_n)\} = \{u_k, u_{k+1}\}.$$

Now the restrictions on k and the definition of S imply that in the second case $u_k \in M$ iff $a \leq k \leq b$, iff $a \leq k+1 \leq b$, iff $u_{k+1} \in M$. Thus in both cases $\chi(u_1, \ldots, u_n) \in M$ implies that $\chi(v_1, \ldots, v_n) \in M$. Therefore, thanks to a succession of these transformations from $(u_1, \ldots, u_n)$ to $(v_1, \ldots, v_n)$, $\chi(j_1, \ldots, j_n) \in M$ for every $(j_1, \ldots, j_n) \in S$.

(c) (ii) implies (i).
We define the function $f_D : D^n \to D$ as follows: for any $x_1, \ldots, x_n \in D$, if $x_{i_1} \leq \ldots \leq x_{i_n}$

for some $(i_1,\ldots,i_n)\in\mathcal{T}_n$, then $f_D(x_1,\ldots,x_n)=x_t$ for $t=\chi(i_1,\ldots,i_n)$. We have only to show that f_D is uniquely defined, in other words that if we have $x_{j_1}\leq\ldots\leq x_{j_n}$ for some other $(j_1,\ldots,j_n)\in\mathcal{T}_n$, then $x_t=x_{t'}$ for $t'=\chi(j_1,\ldots,j_n)$.

Suppose that we have $x_1,\ldots,x_n\in D$ such that $x_{i_1}\leq\ldots\leq x_{i_n}$ and $x_{j_1}\leq\ldots\leq x_{j_n}$ for two distinct $(i_1,\ldots,i_n)$ and $(j_1,\ldots,j_n)\in\mathcal{T}_n$. Let $t=\chi(i_1,\ldots,i_n)$. Then there exist $a,b\in I_n$ such that $a\leq b$, and there are $a-1$ elements j of I_n such that $x_j<x_t$, and $b-a+1$ elements j of I_n such that $x_j=x_t$. Thus for every $k\in I_n$ we have:

$$x_{i_k}\begin{cases}<x_t & \text{if } k<a;\\ =x_t & \text{if } a\leq k\leq b;\\ >x_t & \text{if } b>k.\end{cases}$$

Now $j_1,\ldots,j_n$ satisfy the same property:

$$x_{j_k}\begin{cases}<x_t & \text{if } k<a;\\ =x_t & \text{if } a\leq k\leq b;\\ >x_t & \text{if } b>k.\end{cases}$$

Therefore

$$\{i_k\mid k<a\}=\{j_k\mid k<a\}\quad\text{and}$$
$$\{i_k\mid a\leq k\leq b\}=\{j_k\mid a\leq k\leq b\}=M.$$

Moreover $t=\chi(i_1,\ldots,i_n)\in M$ by definition of a and b. By (ii), $\chi(j_1,\ldots,j_n)\in M$, in other words $x_{t'}=x_t$ for $t'=\chi(j_1,\ldots,j_n)$. ∎

This result is very powerful for the determination of values of order functions, when some of them are specified. We show this on two simple examples.

(1°) Let χ be the choice map of an order function f_D. Assume that there exist $(u_1,\ldots,u_n)$ and $(v_1,\ldots,v_n)\in\mathcal{T}_n$ such that $u_1=v_n=t$ for some $t\in I_n$, and $\chi(u_1,\ldots,u_n)=\chi(v_1,\ldots,v_n)=t$. Then $\chi(r_1,\ldots,r_n)=t$ for any $(r_1,\ldots,r_n)\in\mathcal{T}_n$, in other words $f_D(x_1,\ldots,x_n)=x_t$ for any $x_1,\ldots,x_n\in D$.

Indeed, if $r_1=t$, this follows by applying (ii) to $(u_1,\ldots,u_n)$ and $(r_1,\ldots,r_n)$ with $a=b=1$. If $r_n=t$, it follows by applying (ii) to $(v_1,\ldots,v_n)$ and $(r_1,\ldots,r_n)$ with $a=b=n$. If $r_k=t$ for $1<k<n$, then $\chi(r_k,r_1,\ldots,r_{k-1},r_{k+1},\ldots,r_n)=r_k$, as shown above, and so (ii) applied to $(r_1,\ldots,r_n)$ and $(r_k,r_1,\ldots,r_{k-1},r_{k+1},\ldots,r_n)$ with $a=1$ and $b=k$ implies that $\chi(r_1,\ldots,r_n)\in\{r_1,\ldots,r_k\}$; but also $\chi(r_1,\ldots,r_{k-1},r_{k+1},\ldots,r_n,r_k)=r_k$, and so (ii) applied to $(r_1,\ldots,r_n)$ and $(r_1,\ldots,r_{k-1},r_{k+1},\ldots,r_n,r_k)$ with $a=k$ and $b=n$ implies that $\chi(r_1,\ldots,r_n)\in\{r_k,\ldots,r_n\}$; combining both results, we have $\chi(r_1,\ldots,r_n)=r_k=t$.

(2°) Consider now two integers $a,b\in I_n$ and an order function f_D such that for every $x_1,\ldots,x_n\in D$, $x_1=\min\{x_1,\ldots,x_n\}$ implies that $f_D(x_1,\ldots,x_n)=r_a(x_1,\ldots,x_n)$, and $x_1=\max\{x_1,\ldots,x_n\}$ implies that $f_D(x_1,\ldots,x_n)=r_b(x_1,\ldots,x_n)$. In other words, for any $(i_1,\ldots,i_n)\in\mathcal{T}_n$, $\chi(i_1,\ldots,i_n)=i_a$ if $i_1=1$ and $\chi(i_1,\ldots,i_n)=i_b$ if $i_n=1$. Then the method used in the preceding paragraph can be used to show that for any $(i_1,\ldots,i_n)\in\mathcal{T}_n$,

$\chi(i_1,\ldots,i_n) = i_a$ if $i_k = 1$ for $k < a$, and $\chi(i_1,\ldots,i_n) = i_b$ if $i_k = 1$ for $k > b$. We get thus a contradiction for $b < k < a$, and we show then that $f_D(x_1,\ldots,x_n) = x_1$ if $i_k = 1$ for $a \leq k \leq b$. Hence one of the following holds (where r_k is the k-th rank function):

(*i*) $a \leq b$, and for every $x_1,\ldots,x_n \in D$ we have

$$f_D(x_1,\ldots,x_n) = \begin{cases} r_a(x_1,\ldots,x_n) & \text{if } x_1 \leq r_a(x_1,\ldots,x_n); \\ x_1 & \text{if } r_a(x_1,\ldots,x_n) \leq x_1 \leq r_b(x_1,\ldots,x_n); \\ r_b(x_1,\ldots,x_n) & \text{if } r_b(x_1,\ldots,x_n) \leq x_1. \end{cases}$$

In fact, f_D is the b-th weighted rank function for the weights $w_1 = b-a+1$ and $w_i = 1$ for $i = 2,\ldots,n$. Two particular cases are: $a = b$, and then $f_D = r_a$; $a = 1$, $b = n$ and then $f_D(x_1,\ldots,x_n) = x_1$ as in the preceding example.

(*ii*) $a = b+1$ and for every $x_1,\ldots,x_n \in D$ we have $f_D(x_1,\ldots,x_n) = r_b(x_2,\ldots,x_n)$.

For $a = 2$ and $b = n-1$, the function described in (*i*) can be used for the design of a "no extreme" filter eliminating all isolated peaks in an image: if the greylevel of a point is minimum in the window around it, then it is replaced by the next to minimum greylevel; if it is maximum, then it is replaced by the next to maximum greylevel; otherwise it is preserved.

II.3. Extending the domain of preorder and order functions

The alert reader will have noted that the definition of order functions (see particularly Definition 3' in Subsection II.2) is independent of the structure of the domain D of its variables. It depends only upon the *choice map* χ associating to each n-tuple $(i_1,\ldots,i_n) \in \mathcal{T}_n$ an integer $\chi(i_1,\ldots,i_n) \in \mathcal{I}_n$. Moreover, the constraints that χ must satisfy (see Proposition 1 in Subsection II.2) have nothing to do with D. Hence, given another subset D' of size at least 2 of $\mathbf{R}$, to an order function f_D on D^n corresponds a similar order function $f_{D'}$ on D'^n, which has the same choice map χ as f_D. We can thus restrict f_D to C^n for $C \subseteq D$, or extend in a unique way f_D to E^n for $E \supseteq D$, and we get the corresponding order functions f_C and f_E, having the same choice map as f_D. In other words:

Proposition 2. *For $C \subset D$ (with $|C| \geq 2$), the restriction to C^n of an order function on D^n is an order function having the same choice map. Given $E \subseteq \mathbf{R}$ such that $E \supset D$, an order function on D^n can be uniquely extended to an order function on E^n, and that extension has moreover the same choice map.*

Thus, given an order function f_D on D^n, we will assume that f_D is the restriction to D^n of an order function f on $\mathbf{R}^n$. Conversely, the restriction to D^n of an order function f on $\mathbf{R}^n$ will be written f_D. The distinction between $f: \mathbf{R}^n \to \mathbf{R}$ and $f_D: D^n \to D$ is thus only a distinction on the domain of their variables, since both functions will have the same choice map.

For preorder functions, the situation is somewhat more complicated. Of course, for $C \subseteq D$, the restriction $f_C: C^n \to C$ of a preorder function $f_D: D^n \to D$ is still a preorder function. However, the extension of f_D to a larger set E^n is not always uniquely defined.

In fact, as can be seen from Definition 2' in Subsection II.2, one property of D intervenes in the definition of a preorder function on D^n: it is its size, which appears in the expression $\min\{n, |D|\}$. This number is the upper bound on the number of *distinct* values that n variables $x_1, \ldots, x_n \in D$ may take. Thus for $E \supset D$, if we wish to extend f_D to a preorder function f_E on E^n, then for $x_1, \ldots, x_n \in E$ the value of $f_E(x_1, \ldots, x_n)$ is uniquely determined when $x_1, \ldots, x_n$ take at most $\min\{n, |D|\}$ distinct values, because it is then possible to take $y_1, \ldots, y_n \in D$ such that for $i, j \in I_n$, $x_i < x_j$ iff $y_i < y_j$, and so we must take $f_E(x_1, \ldots, x_n) = x_t$ if $f_D(y_1, \ldots, y_n) = y_t$. On the other hand, when $x_1, \ldots, x_n$ take more than $\min\{n, |D|\}$ distinct values, $f_E(x_1, \ldots, x_n)$ cannot be determined from f_D. Let us explain this with a simple example:

We take $n = 3$, $D = \{0, 1\}$, and $E = \{0, 1, 2\}$. If we have $f_D(0, 0, 1) = 1$, then we will have $f_E(1, 1, 2) = 2$ and $f_E(0, 0, 2) = 2$, because in all three cases we have $x_1 = x_2 < x_3$. Here the extension works well, because we use no more than $2 = \min\{n, |D|\}$ values. But now $f_E(0, 1, 2)$ cannot be determined from f_D, because we have here $x_1 < x_2 < x_3$, something which cannot happen in D. Thus we can choose $f_E(0, 1, 2)$ equal to either 0, 1 or 2. The trouble happened because we had 3 distinct values, and $\min\{n, |D|\} < 3 \leq \min\{n, |E|\}$ Now if we extend E to a larger set F, then the extension from f_E to f_F is unique. Suppose for example that we have $f_E(0, 1, 2) = 1$. Then for any $a, b, c \in F$ such that $a < b < c$, we will have $f_F(a, b, c) = b$. Here everything is all right because we have $\min\{n, |E|\} = \min\{n, |F|\} = 3$.

Hence the extension from f_D to f_E for $E \supset D$ is unique iff $\min\{n, |D|\} = \min\{n, |E|\}$, in other words iff $|D| \geq n$. We have thus the following:

Proposition 3. *For $C \subset D$, the restriction to C^n of a preorder function on D^n is a preorder function. Given $E \subseteq \mathbf{R}$ such that $E \supset D$, a preorder function on D^n admits an extension to E^n which is a preorder function, and this extension is unique iff $|D| \geq n$.*

II.4. The dual and the composition of preorder and order functions

There is a natural duality between the two strict order relations $<$ and $>$ on $\mathbf{R}$. It induces a duality on preorder and order functions.

Consider a preorder function $f_D : D^n \to D$. Given $x_1, \ldots, x_n \in D$, $f_D(x_1, \ldots, x_n) = x_t$, where t is determined by the set $\mathcal{R}$ of all ordered pairs (i, j) such that $x_i < x_j$. In other words f_D is determined by a map τ associating to $\mathcal{R}$ some $t = \tau(\mathcal{R}) \in I_n$. The inversion of $<$ and $>$ leads to the dual set $\mathcal{R}^*$ of all ordered pairs (u, v) such that $x_u > x_v$, in other words $(v, u) \in \mathcal{R}$. This dual $\mathcal{R}^*$ of $\mathcal{R}$ induces a *dual* f_D^* of f_D. It is defined as follows: for any $x_1, \ldots, x_n \in D$, given the set $\mathcal{R}$ of ordered pairs (i, j) such that $x_i < x_j$, $f_D^*(x_1, \ldots, x_n) = x_t$, where $t = \tau(\mathcal{R}^*)$. Taking $x'_1, \ldots, x'_n \in D$ such that for any $i, j \in I_n$, $x_i < x_j$ iff $x'_i > x'_j$, then $\mathcal{R}^*$ is the set of all ordered pairs (i, j) such that $x'_i < x'_j$, and so we have $f(x'_1, \ldots, x'_n) = x'_t$ iff $f_D^*(x_1, \ldots, x_n) = x_t$. Thus we can state the following:

Definition 4. Given a preorder function $f_D : D^n \to D$, its dual f_D^* is a preorder function $D^n \to D$ built as follows: given $x_1, \ldots, x_n, x'_1, \ldots, x'_n \in D$ such that $x_i < x_j$ iff $x'_i > x'_j$ for each $i, j \in I_n$, $f_D^*(x_1, \ldots, x_n) = x_t$ when $f_D(x'_1, \ldots, x'_n) = x'_t$.

Let us explain why f_D^* is well-defined and a preorder function. Take $x_1, \ldots, x_n \in D$. First, it is always possible to choose $x'_1, \ldots, x'_n \in D$ such that for each $i, j \in I_n$, $x_i < x_j$ iff $x'_i > x'_j$. Indeed, let $y_1, \ldots, y_m$ be the distinct values taken by $x_1, \ldots, x_n$, with $y_1 < \ldots < y_m$. For each $i \in I_n$, there is some $u \in I_m$ such that $x_i = y_u$; we set then $x'_i = y_{m+1-u}$. Now for $i, j \in I_n$, if $x_i = y_u$ and $x_j = y_v$, then $x_i = y_u < y_v = x_j$ iff $u < v$, iff $m+1-u > m+1-v$, iff $x'_i = y_{m+1-u} > y_{m+1-v} = x'_j$.

Second, the value of $f_D^*(x_1, \ldots, x_n) = x_t$ does not depend on the choice of $x'_1, \ldots, x'_n$. Indeed, if we have $x'_1, \ldots, x'_n, x''_1, \ldots, x''_n$ such that for each $i, j \in I_n$, $x_i < x_j$ iff $x'_i > x'_j$ iff $x''_i > x''_j$, then $f_D(x'_1, \ldots, x'_n) = x'_t$ iff $f_D(x''_1, \ldots, x''_n) = x''_t$ (since f_D is a preorder function), and so $f_D^*(x_1, \ldots, x_n) = x_t$ is well-defined.

Finally, f_D^* is a preorder function. Indeed, as f_D is a preorder function, we have $f_D(x'_1, \ldots, x'_n) = x'_t$, where t is determined by the set of ordered pairs (j, i) such that $x'_j < x'_i$, in other words $x_i < x_j$. Thus $f_D(x_1, \ldots, x_n) = x_t$, where t is chosen as a function of the set of ordered pairs (i, j) such that $x_i < x_j$.

In Subsection II.2, we gave a formal definition of preorder functions. We showed there that the set $\mathcal{R}$ of all ordered pairs (i, j) such that $x_i < x_j$ is characterized by an ordered m-tuple $(P_1, \ldots, P_m)$ of subsets of I_n forming a partition of it (where $1 \leq m \leq \min\{n, |D|\}$). Thus a preorder function f_D is determined by a map σ associating to any such ordered m-tuple $(P_1, \ldots, P_m)$ some $t \in I_n$. Then it is easy to show that f_D^* is a preorder function determined by the dual map σ^* defined by

$$\sigma^*(P_1, \ldots, P_m) = \sigma(P_m, \ldots, P_1). \tag{1}$$

This equation can be taken as an alternative definition of the dual of a preorder function.

As can be seen from Definition 4 or (1), we have $f_D^{**} = f_D$: a preorder function is the dual of its dual.

It is also clear from the definition that the operation of taking the dual of a preorder function commutes with the restriction from D to a subset C of it.

When f_D is an order function, then f_D^* is also an order function. Indeed, given $x_1, \ldots, x_n, x'_1, \ldots, x'_n \in D$ such that $x_i < x_j$ iff $x'_i > x'_j$ for each $i, j \in I_n$, assume that $x'_{i_1} \leq \ldots \leq x'_{i_n}$. Then $x_{i_n} \leq \ldots \leq x_{i_1}$, $f_D(x_1, \ldots, x_n) = x_t$ for $t = \chi(i_n, \ldots, i_1)$, and we have also $f_D^*(x'_1, \ldots, x'_n) = x'_t$. Thus f_D^* is the order function determined by the choice map $\chi^* : \mathcal{T}_n \to I_n$ defined by

$$\chi^*(i_1, \ldots, i_n) = \chi(i_n, \ldots, i_1) \qquad \text{for} \quad (i_1, \ldots, i_n) \in \mathcal{T}_n. \tag{2}$$

We can summarize our results as follows:

Proposition 4. *The dual of a preorder function f_D is a preorder function f_D^* characterized by* (1), *with the following properties:*

— *If f_D is an order function with choice map χ, then its dual f_D^* is an order function with choice map χ^*, where χ^* is defined by* (2).

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

— *Given a subset C of D, the dual f_C^* of the restriction f_C of f_D to C^n is equal to the restriction to C^n of the dual f_D^* of f_D.*

Let us describe briefly the dual of some well-known order functions: the maximum is the dual of the minimum, the median is its own dual, the dual of the rank function r_k (selecting the kth smallest value of a sample) is the rank function r_{n+1-k} (selecting the kth largest value of a sample). The dual of the k-th weighted rank function determined by the weights $w_1, \ldots, w_n$ is the k'-th weighted rank function determined by the same weights $w_1, \ldots, w_n$, where $k' = w_1 + \cdots + w_n + 1 - k$. In particular the weighted median is its own dual.

There is often a "natural" bijection $\omega : D \to D$ which reverses the order of the elements of D (in other words, which is strictly decreasing). For example, if $D = \mathbf{R}$ or $\mathbf{Z}$ (the set of rational integers), ω is the map $D \to D : x \mapsto -x$, while if D is finite, there is in fact a unique strictly decreasing bijection $D \to D$. Then for any $x_1, \ldots, x_n \in D$, we can take $\omega(x_1), \ldots, \omega(x_n)$ for $x'_1, \ldots, x'_n$ in Definition 4, and so we have

$$\omega(f_D^*(x_1, \ldots, x_n)) = f_D(\omega(x_1), \ldots, \omega(x_n)), \tag{3}$$

or

$$f_D^*(x_1, \ldots, x_n) = \omega^{-1}(f_D(\omega(x_1), \ldots, \omega(x_n))). \tag{4}$$

Note that ω^2 is often the identity on D (in other words, $x = \omega(\omega(x))$ for any $x \in D$). This is the case for the two examples given above. Then, since $\omega = \omega^{-1}$, we get

$$f_D^*(x_1, \ldots, x_n) = \omega(f_D(\omega(x_1), \ldots, \omega(x_n))). \tag{5}$$

Of course we can then take this equation as an alternative definition of the dual of f_D in this case.

As we will see later, duality takes an important place in the theory of order functions and order filters. Another important operation for order functions is the composition. We mentioned in the Introduction the composition of rank functions or rank filters, and announced the main result of Section III, that an order function is a composition of min and max functions. We will end this section by giving its definition and main properties, and some constructions derived from it.

Definition 5. Given $m, n \geq 2$, m functions $g_D^1, \ldots, g_D^m : D^n \to D$ and a function $f_D : D^m \to D$, the *composition* $f_D \circ [g_D^1, \ldots, g_D^m]$ *of* $g_D^1, \ldots, g_D^m$ *by* f_D is the function

$D^n \to D$ defined by

$$f_D \circ [g_D^1, \ldots, g_D^m](x_1, \ldots, x_n) = f_D(g_D^1(x_1, \ldots, x_n), \ldots, g_D^m(x_1, \ldots, x_n))$$

for $x_1, \ldots, x_n \in D$.

The proof of the following result is left to the reader:

Proposition 5. *Composition preserves the set of order functions, preorder functions, and selection functions respectively: in other words, if $g_D^1, \ldots, g_D^m$ and f_D are order functions, preorder functions, or selection functions, then $f_D \circ [g_D^1, \ldots, g_D^m]$ is an order function, a preorder function, or a selection function respectively. The dual of a composition of preorder functions is the composition of their duals: $(f_D \circ [g_D^1, \ldots, g_D^m])^* = f_D^* \circ [(g_D^1)^*, \ldots, (g_D^m)^*]$.*

Thus a composition of weighted rank functions is an order function, as we stated in the Introduction. We will show in the next section that the reverse holds: an order function is the composition of a particular type of weighted rank functions, namely min and max functions applied to subsets of the set of variables of the function.

Let us now describe a simple method related to the composition, which allows the building of a selection function, a preorder function, or an order function from another one.

We introduce a function $D^m \to D$ (where $m \geq 2$), the k-th projection p_k defined by $p_k(x_1, \ldots, x_m) = x_k$ for any $x_1, \ldots, x_m \in D$ (with $1 \leq k \leq m$). It is an order function with a constant choice map: $\chi(i_1, \ldots, i_m) = k$ for any $(i_1, \ldots, i_m) \in \mathcal{T}_m$. It is moreover equal to its own dual. We can now build from a function $f_D : D^n \to D$ a function $g_D : D^m \to D$ of the form $f_D \circ [p_{a_1}, \ldots, p_{a_n}]$, where $a_1, \ldots, a_n \in \{1, \ldots, m\}$. In other words, for $x_1, \ldots, x_n \in D$ we have

$$g_D(x_1, \ldots, x_n) = f_D(x_{a_1}, \ldots, x_{a_n}).$$

As g_D is built by a composition of projections p_k by f_D, and as projections are order functions, it is clear from Proposition 5 that this construction preserves the set of order functions, preorder functions, and selection functions respectively. As the projection is its own dual, Proposition 5 again implies that this construction commute with the taking of the dual of a preorder function (or an order function), in other words that we have also

$$g_D^*(x_1, \ldots, x_n) = f_D^*(x_{a_1}, \ldots, x_{a_n})$$

when f_D is a preorder function. Note also that the set of such constructions is closed under repetition.

Let us give three particular cases of this construction. They are: the *permutation of variables*, the *weighted expansion*, and the *void expansion*.

(1°) We take $m = n$ and we choose $a_1, \ldots, a_n$ to be a permutation of $1, \ldots, n$ (in other words and $(a_1, \ldots, a_n) \in \mathcal{T}_n$). We call g_D a *permutation of variables of f_D*.

(2°) *Weighted expansion* is the method by which we defined weighted rank functions from rank functions in the Introduction. Suppose that we have m non-negative integer weights $w_1, \ldots, w_m$ such that $w_1 + \cdots + w_m = n$. Then we define g_D from f_D and $w_1, \ldots, w_m$ by letting $a_1, \ldots, a_n$ consist in w_i copies of i for $i = 1, \ldots, n$ (in increasing order). Thus for any $x_1, \ldots, x_m$ we set $g_D(x_1, \ldots, x_m) = f_D(y_1, \ldots, y_n)$, where $y_1, \ldots, y_n$ consist in w_i copies of x_i for $i = 1, \ldots, m$. In other words,

$$\begin{aligned} &g_D(x_1, \ldots, x_m) = f_D(y_1, \ldots, y_n), \qquad \text{where} \\ &y_j = x_i \qquad \text{for} \quad \sum_{t<i} w_t < j \leq \sum_{t \leq i} w_t, \quad j = 1, \ldots, n. \end{aligned}$$

We call g_D the *weighted expansion of* f_D *by* $w_1, \ldots, w_m$. Note that the weighted rank functions are the weighted expansions of rank functions.

(3°) A particular case of weighted expansion is when $w_i = 0$ or 1 for each i. As $w_1 + \cdots + w_m = n$, we have $m \geq n$, and $1 \leq a_1 \leq \ldots \leq a_n \leq m$. We call g_D the *void expansion of* f_D *to* D^m *by* $a_1, \ldots, a_n$.

As we explained above (see Proposition 2), it is possible to extend the domain D of the variables of an order function. Now by void expansion it is also possible to extend the number n of variables intervening in that function.

Note also that each one of these three subsets of constructions is closed under repetition. In other words, the succession of two permutations of variables is a permutation of variables, a weighted expansion of a weighted expansion of a function is again a weighted expansion of that function, and the void expansion of a void expansion of a function is again a void expansion of that function.

Weighted and void expansion will intervene in the definition of order filters for *finite* images, as we will explain in Section VI.

III. Order functions as compositions of min and max functions

As we said in the Introduction, order functions can be defined in two ways, either as we have done in Subsection II.1 (with Definition 3), or as a composition of the minimum and maximum functions. We give a simple proof of the equivalence between the two definitions in Subsection III.1. In Subsection III.2 we analyze the possible min-max decompositions of order functions. Then in Subsection III.3 we give a description of order functions as a generalization of weighted rank functions, that we call *set-weighted rank functions.* These two subsections are rather technical, and can be skipped in a first reading.

III.1. The main argument

For any set $\mathcal{H}$ of subsets of I_n, we define the minimum of maxima and maximum of minima functions $minmax[\mathcal{H}], maxmin[\mathcal{H}]: \mathbf{R}^n \to \mathbf{R}$ by

$$\begin{aligned} minmax[\mathcal{H}](x_1,\ldots,x_n) &= \min_{S\in\mathcal{H}}(\max_{j\in S}(x_j)), \\ maxmin[\mathcal{H}](x_1,\ldots,x_n) &= \max_{S\in\mathcal{H}}(\min_{j\in S}(x_j)). \end{aligned} \tag{6}$$

We will give here a simple proof of the fact that a function $D^n \to D$ is an order function iff it is equal to $minmax[\mathcal{H}]$ for some $\mathcal{H}$ (or $maxmin[\mathcal{H}']$ for some $\mathcal{H}'$). This result was first shown in [2] for $D = \mathbf{R}$. (In fact, following [14], order functions were then defined as continuous preorder functions).

As we explained at the beginning of Subsection II.3 (see Proposition 2), the definition of an order function f_D on D^n is independent of the domain D of its variables, and so the behavior of f_D is completely determined by that of the corresponding function f_B on B^n, where $B = \{0,1\}$. It suffices thus to show the result for $D = B$. We have the following characterization of order functions with boolean variables:

Theorem 6. *Recall the set $B = \{0,1\}$. Consider a function $f_B: B^n \to B$. Then the following four statements are equivalent:*

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

(*ii*) *f_B is a non-constant increasing function.*

(*iii*) *There is a set $\mathcal{H}$ of subsets of I_n such that $f_B = minmax[\mathcal{H}]_B$.*

(*iv*) *There is a set $\mathcal{H}'$ of subsets of I_n such that $f_B = maxmin[\mathcal{H}']_B$.*

Proof. (*a*) (*i*) *implies* (*ii*).
Clearly an order function f_B is non-constant. Let us show that it is increasing. Take $x_1,\ldots,x_n,y_1,\ldots,y_n \in B$ such that $x_i \le y_i$ for each $i \in I_n$. We set

$$U \doteq \{i \in I_n \mid x_i = y_i = 0\},$$

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

$$W \doteq \{i \in I_n \mid x_i = y_i = 1\},$$

and let $u = |U|$, $v = |V|$, and $w = |W|$. Take $(i_1, \ldots, i_n) \in \mathcal{T}_n$ such that

$$U = \{i_j \mid 1 \leq j \leq u\},$$
$$V = \{i_j \mid u+1 \leq j \leq u+v\},$$
$$W = \{i_j \mid u+v+1 \leq j \leq n\}.$$

Then $x_{i_1} \leq \ldots \leq x_{i_n}$, $y_{i_1} \leq \ldots \leq y_{i_n}$, and so for $t = \chi(i_1, \ldots, i_n)$ we have $f_B(x_1, \ldots, x_n) = x_t$ and $f_B(y_1, \ldots, y_n) = y_t$. As $x_t \leq y_t$, we get $f_B(x_1, \ldots, x_n) \leq f_B(y_1, \ldots, y_n)$.

(*b*) (*ii*) *implies* (*iii*).
This is a well-known result in the theory of boolean functions. A proof of it can be found in Theorem 5 on page 189 of [8]. We will also give some explicit decompositions of f_B as a minimum of maxima in Subsection III.2.

(*c*) (*iii*) *implies* (*i*).
This follows by the distributivity law for the minimum and maximum (a minimum of maxima can be decomposed as a maximum of minima, in the same way as a product of sums can be decomposed as a sum of products).

(*d*) (*iv*) *implies* (*i*).
As the maximum and the partial minima $\min_{j \in S}(x_j)$ (for $X \in \mathcal{H}$) are all order functions, their composition is an order function by Proposition 5. ∎

We deduce then the following characterization of order functions $D^n \to D$:

Corollary 7. *Consider a function $f_D : D^n \to D$. Then the following three statements are equivalent:*

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

(*ii*) *There is a set $\mathcal{H}$ of subsets of $\mathcal{I}_n$ such that $f_D = minmax[\mathcal{H}]_D$.*

(*iii*) *There is a set $\mathcal{H}'$ of subsets of $\mathcal{I}_n$ such that $f_D = maxmin[\mathcal{H}']_D$.*

Proof. (*a*) Each one of (*ii*) and (*iii*) *implies* (*i*).
The argument is the same as in point (*d*) of the proof of Theorem 6.

(*b*) (*i*) *implies* (*ii*) and (*iii*).
Let f be the (unique) extension of f_D to $\mathbf{R}^n$, and let f_B be the restriction of f to B^n; then f and f_B are order functions (see Proposition 2). By Theorem 6 there is a set $\mathcal{H}$ for which $f_B = minmax[\mathcal{H}]_B$. Now $minmax[\mathcal{H}]$ is an order function (by (*a*)), and by Proposition 2 we must then have $f = minmax[\mathcal{H}]$, and so $f_D = minmax[\mathcal{H}]_D$. Thus (*ii*) holds, and we prove similarly that (*iii*) holds. ∎

It follows that the set of order functions on D^n is equal to the set of functions built from n variables $x_1, \ldots, x_n$ in D by arbitrary combinations of min and max functions. This set is isomorphic to the free distributive lattice generated by n symbols (see [3], pages 59–63). In particular, the number of order functions in n variables is the size of that lattice. The determination of this number is known as the "Dedekind problem", and it is still unsolved for $n > 7$.

In the next subsection, we will describe the sets $\mathcal{H}$ and $\mathcal{H}'$ satisfying (ii) and (iii) in Corollary 7 in terms of the properties of f_B.

III.2. Characterization of the possible min-max decompositions

We consider an order function f_D on D^n and the corresponding order function f_B on B^n. We will derive min-max decompositions of f_D from the behavior of f_B. But we must first introduce some notation.

Given a subset S of I_n, we write S^c for its complement in I_n, in other words $S^c = I_n - S$. Following the classical use in Boolean algebra, for any $\alpha \in B$ we write $\overline{\alpha}$ for the other element of B, in other words $\overline{\alpha} = 1 - \alpha$. The map $\alpha \mapsto \overline{\alpha}$ is called the *complementation.* It reverses the order of B and is its own inverse. Thus (5) implies that given the order function f_B on B^n, its dual f_B^* satisfies the equality

$$f_B^*(x_1, \ldots, x_n) = \overline{f_B(\overline{x_1}, \ldots, \overline{x_n})} \tag{7}$$

for any $x_1, \ldots, x_n \in B$.

Now, given $i \in I_n$, $S \subseteq I_n$, $\alpha \in B$ and a function $g_B : B^n \to B$, we define the following two quantities:

$$\varepsilon(\alpha, i, S) \doteq \begin{cases} \alpha & \text{if } i \in S, \\ \overline{\alpha} & \text{if } i \notin S, \end{cases} \tag{8}$$

and

$$\delta(\alpha, g_B, S) \doteq g_B(\varepsilon(\alpha, 1, S), \ldots, \varepsilon(\alpha, n, S)). \tag{9}$$

Let us mention here some of their properties with respect to complementation. Clearly (8) implies that

$$\varepsilon(\alpha, i, S^c) = \varepsilon(\overline{\alpha}, i, S) = \overline{\varepsilon(\alpha, i, S)}. \tag{10}$$

Then by (9) and (10) we get:

$$\delta(\alpha, g_B, S^c) = \delta(\overline{\alpha}, g_B, S). \tag{11}$$

In the case of the order function f_B, we obtain by (7), (9), and (10):

$$\delta(\alpha, f_B^*, S) = \overline{\delta(\overline{\alpha}, f_B, S)} = \overline{\delta(\alpha, f_B, S^c)}. \tag{12}$$

We can now introduce the sets that may belong to the two sets $\mathcal{H}$ and $\mathcal{H}'$ mentioned in Corollary 7 for the min-max decomposition of an order function. Given $\alpha \in B$, $S \subseteq I_n$ and $g_B : B^n \to B$, we say that S is *α-heavy for g_B* if $\delta(\alpha, g_B, S) = \alpha$. Such a set satisfies the following properties:

Lemma 8. *Given $\alpha \in B$ and the order function $f_B : B^n \to B$, we have:*

(*i*) *For every $S, T \subseteq I_n$, if S is α-heavy for f_B and $S \subseteq T$, then T is also α-heavy for f_B.*

(*ii*) *I_n is α-heavy for f_B, while $\emptyset$ is not.*

(*iii*) *For every $S \subseteq I_n$, S is α-heavy for f_B iff it is $\overline{\alpha}$-heavy for f_B^*.*

(*iv*) *For every $S \subseteq I_n$, S is α-heavy for f_B iff S^c is not $\overline{\alpha}$-heavy for f_B.*

Proof. (*i*) Suppose first that $\alpha = 0$. Then (8) implies that $\varepsilon(0,i,S) \geq \varepsilon(0,i,T)$ for each $i \in I_n$. As f_B is increasing, we have by (9)

$$0 = \delta(0, f_B, S) = f_B(\varepsilon(0,1,S), \ldots, \varepsilon(0,n,S)) \geq f_B(\varepsilon(0,1,T), \ldots, \varepsilon(0,n,T)) = \delta(0, f_B, T),$$

in other words $\delta(0, f_B, T) = 0$ and so T is 0-heavy.

Suppose last that $\alpha = 1$. Then a similar argument shows that $\varepsilon(1,i,S) \leq \varepsilon(1,i,T)$ and $1 = \delta(1, f_B, S) \leq \delta(1, f_B, T)$, in other words $\delta(1, f_B, T) = 1$ and so T is 1-heavy.

(*ii*) follows from the fact that $f(0,\ldots,0) = 0$ and $f(1,\ldots,1) = 1$, in other words $\delta(\alpha, f_B, I_n) = \alpha$ and $\delta(\alpha, f_B, \emptyset) = \overline{\alpha}$.

(*iii*) Applying (12) with $\overline{\alpha}$ instead of α, we have $\delta(\overline{\alpha}, f_B^*, S) = \overline{\delta(\alpha, f_B, S)}$. Thus S is α-heavy for f_B iff $\delta(\alpha, f_B, S) = \alpha$, iff $\delta(\overline{\alpha}, f_B^*, S) = \overline{\alpha}$, iff S is $\overline{\alpha}$-heavy for f_B^*.

(*iv*) Applying (11) with $\overline{\alpha}$ instead of α, we have $\delta(\overline{\alpha}, f_B, S^c) = \delta(\alpha, f_B, S)$. Thus S is α-heavy for f_B iff $\delta(\alpha, f_B, S) = \alpha$, iff $\delta(\overline{\alpha}, f_B, S^c) = \alpha$, iff S^c is not $\overline{\alpha}$-heavy for f_B. ∎

The denomination "α-heavy" that we introduced above can be explained by property (*i*), since a set having a heavy subset is itself heavy. It will also be justified in the next subsection, where we will characterize order functions as a generalization of weighted rank functions, with a weight associated to each subset rather than to each element of I_n; of course, that weight will depend upon the heavy subsets of that set.

For $\alpha \in B$ and an order function f_B, write $\mathcal{H}_\alpha[f_B]$ for the set of all α-heavy sets for f_B. Let $\mathcal{M}_\alpha[f_B]$ be the set of all minimal elements of $\mathcal{H}_\alpha[f_B]$; its elements will be called *minimal α-heavy sets for f_B*. Then by Lemma 8 (*i*) we have

$$\mathcal{H}_\alpha[f_B] = \bigcup_{S \in \mathcal{M}_\alpha[f_B]} \{T \subseteq I_n \mid S \subseteq T\}. \tag{13}$$

We can now characterize the possible min-max decompositions of an order function f_D on D^n in terms of α-heavy sets for f_B.

Proposition 9. *Given the order function $f_D : D^n \to D$ and the corresponding order function $f_B : B^n \to B$, for any sets $\mathcal{H}, \mathcal{H}'$ of subsets of I_n, we have:*

(*i*) $\quad f_D = minmax[\mathcal{H}]_D \quad$ *iff* $\quad \mathcal{M}_0[f_B] \subseteq \mathcal{H} \subseteq \mathcal{H}_0[f_B]$.

(*ii*) $\quad f_D = maxmin[\mathcal{H}']_D \quad$ *iff* $\quad \mathcal{M}_1[f_B] \subseteq \mathcal{H}' \subseteq \mathcal{H}_1[f_B]$.

Proof. (*i*) As explained in the proof of Corollary 7, $f_D = minmax[\mathcal{H}]_D$ iff $f_B = minmax[\mathcal{H}]_B$. Now this equality is equivalent to the following statement:

— For any $x_1, \ldots, x_n \in B$, $f_B(x_1, \ldots, x_n) = 0$ iff $minmax[\mathcal{H}]_B(x_1, \ldots, x_n) = 0$.

Define $N(x_1, \ldots, x_n)$ as the set of all $j \in I_n$ such that $x_j = 0$. Then $f_B(x_1, \ldots, x_n) = 0$ means that $N(x_1, \ldots, x_n)$ is 0-heavy, while $minmax[\mathcal{H}]_B(x_1, \ldots, x_n) = 0$ means by (6) that there is some $S \in \mathcal{H}$ such that $\max_{j \in S}(x_j) = 0$, in other words $S \subseteq N(x_1, \ldots, x_n)$. Thus that statement can be rewritten as follows:

— For any $x_1, \ldots, x_n \in B$, $N(x_1, \ldots, x_n)$ is 0-heavy iff there is some $S \in \mathcal{H}$ such that $S \subseteq N(x_1, \ldots, x_n)$.

Now $N(x_1, \ldots, x_n)$ can be any subset P of I_n. Thus the statement is equivalent to the following one:

(∗) For any $P \subseteq I_n$, P is 0-heavy iff there is some $S \in \mathcal{H}$ such that $S \subseteq P$.

Now we have three possibilities:

(*a*) $\mathcal{M}_0[f_B] \subseteq \mathcal{H} \subseteq \mathcal{H}_0[f_B]$.

Take $P \subseteq I_n$. If P is 0-heavy, let S be a minimal 0-heavy subset of P; then $S \subseteq P$, and $S \in \mathcal{H}$, since $\mathcal{M}_0[f_B] \subseteq \mathcal{H}$. Conversely, if there is some $S \in \mathcal{H}$ such that $S \subseteq P$, then S is 0-heavy, since $\mathcal{H} \subseteq \mathcal{H}_0[f_B]$, and so P is 0-heavy by Lemma 8 (*i*). Hence the statement (∗) is satisfied in this case.

(*b*) $\mathcal{H} \subseteq \mathcal{H}_0[f_B]$, but $\mathcal{M}_0[f_B] \not\subseteq \mathcal{H}$.

Let P be an element of $\mathcal{M}_0[f_B]$ not contained in $\mathcal{H}$; then P is 0-heavy, and as it is minimal 0-heavy, it does not contain another 0-heavy set S. As every element of $\mathcal{H}$ is 0-heavy, there is no element S of $\mathcal{H}$ such that $S \subseteq P$. Hence the statement (∗) is contradicted in this case.

(*c*) $\mathcal{H} \not\subseteq \mathcal{H}_0[f_B]$.

Take $S \in \mathcal{H}$ such that S is not 0-heavy. Then for $P = S$, P is not 0-heavy and $S \subseteq P$ with $S \in \mathcal{H}$. Hence the statement (∗) is contradicted in this case.

Therefore (∗) is equivalent to (*a*), in other words $f_B = minmax[\mathcal{H}]_B$ iff $\mathcal{M}_0[f_B] \subseteq \mathcal{H} \subseteq \mathcal{H}_0[f_B]$.

(*ii*) As the minimum is the dual of the maximum, by Proposition 5 $maxmin[\mathcal{H}']_D$ is the dual of $minmax[\mathcal{H}']_D$. Thus $f_D = maxmin[\mathcal{H}']_D$ iff $f_D^* = minmax[\mathcal{H}']_D$, iff $\mathcal{M}_0[f_B^*] \subseteq \mathcal{H}' \subseteq \mathcal{H}_0[f_B^*]$. Now Lemma 8 (*iii*) states that for $S \subseteq I_n$, S is 1-heavy for f_B iff it is 0-heavy for f_B^*. In other words $\mathcal{H}_1[f_B] = \mathcal{H}_0[f_B^*]$. It follows then (by identifying the minimal elements in both sets) that $\mathcal{M}_1[f_B] = \mathcal{M}_0[f_B^*]$. Therefore $f_D = maxmin[\mathcal{H}']_D$ iff $\mathcal{M}_1[f_B] \subseteq \mathcal{H}' \subseteq \mathcal{H}_1[f_B]$. ∎

We have thus characterized the possible min-max decompositions of an order function f_D on D^n. Note that for $\alpha = 0, 1$, $\mathcal{H}_\alpha[f_B]$ is determined by $\mathcal{M}_\alpha[f_B]$ (see (13)). Thus the sets $\mathcal{H}$ and $\mathcal{H}'$ for which f_D can be decomposed as $minmax[\mathcal{H}]_D$ and as $maxmin[\mathcal{H}']_D$ are determined by $\mathcal{M}_0[f_B]$ and $\mathcal{M}_1[f_B]$ respectively, which are also the smallest sets giving these two decompositions. It is thus natural to consider $minmax[\mathcal{M}_0[f_B]]_D$ and $maxmin[\mathcal{M}_1[f_B]]_D$

as the two standard min-max decompositions of the order function f_D.

Let us now give simple characterization of α-heavy sets for f_B in terms of f_D:

Proposition 10. *Let $P \subseteq I_n$. Then:*

(*i*) *P is 0-heavy for f_B iff $f_D(x_1, \dots, x_n) \leq \max_{j \in P}(x_j)$ for any $x_1, \dots, x_n \in D$.*

(*ii*) *P is 1-heavy for f_B iff $f_D(x_1, \dots, x_n) \geq \min_{j \in P}(x_j)$ for any $x_1, \dots, x_n \in D$.*

Proof. (*i*) The set P is 0-heavy for f_B iff $\mathcal{H}_0[f_B] = \mathcal{H}_0[f_B] \cup \{P\}$, in other words iff $minmax[\mathcal{H}_0[f_B]]_D = minmax[\mathcal{H}_0[f_B] \cup \{P\}]_D$. Now Proposition 9 (*i*) implies that $minmax[\mathcal{H}_0[f_B]]_D = f_D$, while for any $x_1, \dots, x_n \in D$ we have (by (6))

$$minmax[\mathcal{H}_0[f_B] \cup \{P\}](x_1, \dots, x_n) = \min\big(\max_{j \in P}(x_j), minmax[\mathcal{H}_0[f_B]]_D(x_1, \dots, x_n)\big).$$

Thus P is 0-heavy iff for any $x_1, \dots, x_n \in D$ we have

$$f_D(x_1, \dots, x_n) = \min\big(\max_{j \in P}(x_j), f_D(x_1, \dots, x_n)\big).$$

But the latter equality is equivalent to $f_D(x_1, \dots, x_n) \leq \max_{j \in P}(x_j)$.

(*ii*) is proved in the same way as (*i*). ▮

III.3. Set-weighted rank functions

Recall the weighted rank functions mentioned in the Introduction. Let us give here a slightly different formulation of their definition. To each $i \in I_n$ we associate a non-negative integer weight w_i. Let $w_T = w_1 + \cdots + w_n$. Then for any integer k such that $0 < k \leq w_T$, the k-th *weighted rank function* $\hat{r}_{k;w_1,\dots,w_n}$ determined by the weights $w_1, \dots, w_n$ is built as follows: given $x_1, \dots, x_n \in D$ such that $x_{i_1} \leq \dots \leq x_{i_n}$ (with $(i_1, \dots, i_n) \in \mathcal{T}_n$), we have

$$\hat{r}_{k;w_1,\dots,w_n}(x_1, \dots, x_n) = x_{i_t}, \qquad \text{where} \quad \sum_{j<t} w_{i_j} < k \leq \sum_{j \leq t} w_{i_j}. \tag{14}$$

Thus to each subset P of I_n we associate a *cumulative weight* $W(P)$ equal to the sum of the weights of its elements; then for each $(i_1, \dots, i_n)$, we look at the successive weights

$$W(\{i_1\}), W(\{i_1, i_2\}), \dots, W(\{(i_1, \dots, i_t)\}), \dots, W(\{(i_1, \dots, i_n)\}),$$

and we take $\chi(i_1, \dots, i_n) = i_t$, where t is the smallest integer $j = 1, \dots, n$ such that $W(\{(i_1, \dots, i_j)\}) \geq k$. Note that the inequality $\sum_{j<t} w_{i_j} < k \leq \sum_{j \leq t} w_{i_j}$ in (14) can also be expressed in a dual way as $\sum_{j>t} w_{i_j} < w_T + 1 - k \leq \sum_{j \geq t} w_{i_j}$. Here we look at the successive weights

$$W(\{i_n\}), W(\{i_{n-1}, i_n\}), \dots, W(\{(i_t, \dots, i_n)\}), \dots, W(\{(i_1, \dots, i_n)\}),$$

and we take $\chi(i_1,\ldots,i_n) = i_t$, where t is the largest integer $j = 1,\ldots,n$ such that $W(\{(i_j,\ldots,i_n)\}) \geq w_T + 1 - k$.

One can generalize weighted rank functions by associating to each subset P of I_n a weight $W(P)$ which is not equal to the sum of the weights of its elements. Here W is a real-valued *increasing* function on the set of parts of I_n, in other words, for $P \subseteq Q$, $W(P) \leq W(Q)$. We call it a *weight function*. Set $W_0 = W(\emptyset)$ and $W_T = W(I_n)$. Take a threshold K such that $W_0 < K \leq W_T$. Then we can define from W and K two *set-weighted rank functions* $R^0_{K;W}$ and $R^1_{K;W}$ having respective choice maps $\chi^0_{K;W}$ and $\chi^1_{K;W}$ defined by setting for any $(i_1,\ldots,i_n) \in \mathcal{T}_n$:

$$\begin{aligned} \chi^0_{K;W} &= i_t, \qquad \text{where} \quad W(\{i_j \mid j < t\}) < K \leq W(\{i_j \mid j \leq t\}); \\ \chi^1_{K;W} &= i_t, \qquad \text{where} \quad W(\{i_j \mid j > t\}) < K \leq W(\{i_j \mid j \geq t\}). \end{aligned} \tag{15}$$

When W is a linear function, set-weighted rank functions reduce to ordinary weighted rank functions.

Before making a mathematical analysis of set-weighted rank functions, let us indicate here some possible uses for them. We will give below two examples.

(1°) Assume that we have n devices or observers $T_1,\ldots,T_n$ making n respective measurements $x_1,\ldots,x_n$ of a quantity X. If we make no further assumptions, then an estimation of X will be given by $med(x_1,\ldots,x_n)$, while for $0 < \gamma < 1$, a γ confidence interval will be obtained by rejecting from the sample the smallest and largest values in a proportion of $\eta = (1-\gamma)/2$; in other words it will be bounded by $r_a(x_1,\ldots,x_n)$ and $r_b(x_1,\ldots,x_n)$ for $a = 1 + \eta(n-1)$ and $b = 1 + (1-\eta)(n-1)$.

Suppose now that the devices T_i have distinct degrees of accuracy. Such a degree may be measured by a weight w_i associated to the measurement x_i. Then we will use weighted rank functions instead of rank functions. Thus the estimation of X will be the weighted median $\widehat{med}_{w_1,\ldots,w_n}(x_1,\ldots,x_n)$, and the confidence interval will be bounded by $\hat{r}_{\eta w_T;w_1,\ldots,w_n}(x_1,\ldots,x_n)$ and $\hat{r}_{(1-\eta)w_T;w_1,\ldots,w_n}(x_1,\ldots,x_n)$.

Suppose further that the agreement between certain particular devices has a particular weight; for example with two devices T_j and T_k, the fact that for an estimation x^* of X we have both $x_j \leq x^*$ and $x_k \leq x^*$ increases their weight by w_{jk}. We associate thus to each $P \subseteq I_n$ a weight $W(P)$ measuring the accuracy of the fact that for an estimation x^* of X we have $x_i \leq x^*$ for each $i \in P$; in our example we have $W(\{i\}) = w_i$, $W(\{j\}) = w_j$, and $W(\{i,j\}) = w_i + w_j + w_{ij}$. Then the lower bound of the confidence interval will be given by $R^0_{\eta W_T;W}(x_1,\ldots,x_n)$. Similarly the higher bound of that interval will be given by $R^1_{\eta W'_T;W'}(x_1,\ldots,x_n)$, where for $P \subseteq I_n$, the weight $W'(P)$ measures the accuracy of the fact that for an estimation x^* of X we have $x_i \geq x^*$ for each $i \in P$ (one can generally choose $W' = W$).

(2°) Suppose that we want to design a noise smoothing filter for two-dimensional digital images that replaces the grey-level of each point by the result of the application of an order function to the grey-levels of the 9 points of a 3×3 window centered about it. The usual choice for that order function is the median (and so we get a median filter). Suppose that the window around a given point p is as follows:

$$\begin{matrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{matrix}$$

Then the grey-level of p will be changed from 1 to 0. But here p may be the corner of a square of grey-level 1, and it should not necessarily be erased. In fact, as the three points around p having grey-level 1 are connected, the triple that they constitute should have a bigger weight than a triple of disconnected points, as in the following example:

$$\begin{matrix} 1 & 0 & 0 \\ 0 & * & 1 \\ 1 & 0 & 0 \end{matrix}$$

Thus it will be more convenient to apply within the window a set-weighted rank function such that connected sets have more weight than disconnected sets having the same size.

Let us now make a mathematical analysis of set-weighted rank functions. As n variables $x_1, \ldots, x_n \in D$ may satisfy $x_{i_1} \leq \ldots \leq x_{i_n}$ and $x_{j_1} \leq \ldots \leq x_{j_n}$ for two distinct $(i_1, \ldots, i_n)$ and $(j_1, \ldots, j_n) \in \mathcal{T}_n$, we must show that these two functions $R^0_{K;W}$ and $R^1_{K;W}$ are well-defined, in other words that their choice maps $\chi^0_{K;W}$ and $\chi^1_{K;W}$ satisfy the requirements stated in Subsection II.2 (see Proposition 1). This will be done in Proposition 11. They are then order functions, and then by (2) $R^1_{K;W}$ is the dual of $R^0_{K;W}$. We will then show in Proposition 12 that every order function can be expressed as a set-weighted rank function. Thus the two concepts of order function and set-weighted rank function are equivalent.

It is clear from (15) that the behavior of $\chi^0_{K;W}$ and $\chi^1_{K;W}$ does not depend on the weight $W(P)$ of each $P \subseteq I_n$, but only on the set $\mathcal{H}$ of all sets $P \subseteq I_n$ such that $W(P) \geq K$. These sets P are called *heavy*. The set $\mathcal{H}$ satisfies the following two conditions:

(i) For every $S, T \subseteq I_n$, if $S \in \mathcal{H}$ and $S \subseteq T$, then $T \in \mathcal{H}$.

(ii) $I_n \in \mathcal{H}$, while $\emptyset \in \mathcal{H}$.

Indeed (i) follows from the fact that W is increasing, and (ii) from the fact that $W_0 < K \leq W_T$. Conversely, given a set $\mathcal{H}$ of subsets of I_n satifying conditions (i) and (ii), we define the weight W by $W(P) = 1$ if $P \in \mathcal{H}$ and $W(P) = 0$ otherwise, and we take the threshold $K = 1$; then $\mathcal{H}$ will be the set of all P such that $W(P) \geq K$. A set $\mathcal{H}$ of subsets of I_n satisfying conditions (i) and (ii) above will be called a *heavy set collection in I_n*.

Therefore we can define set-weighted rank functions by their heavy set collection rather than by their weight function. We will thus write $R^0[\mathcal{H}]$, $R^1[\mathcal{H}]$, $\chi^0[\mathcal{H}]$, and $\chi^1[\mathcal{H}]$ for the set-weighted rank functions $R^0_{K;W}$ and $R^1_{K;W}$ and their choice maps $\chi^0_{K;W}$ and $\chi^1_{K;W}$

respectively. Here we have for any $(i_1, \ldots, i_n) \in \mathcal{T}_n$:

$$\begin{aligned} \chi^0[\mathcal{H}] &= i_t, \qquad \text{where} \quad \{i_j \mid j < t\} \notin \mathcal{H} \quad \text{and} \quad \{i_j \mid j \leq t\} \in \mathcal{H}; \\ \chi^1[\mathcal{H}] &= i_t, \qquad \text{where} \quad \{i_j \mid j > t\}) \notin \mathcal{H} \quad \text{and} \quad \{i_j \mid j \geq t\} \in \mathcal{H}. \end{aligned} \tag{16}$$

As announced above, set-weighted rank functions are well-defined:

Proposition 11. *Let $\mathcal{H}$ be a heavy set collection in $\mathcal{I}_n$. Then the two maps $\chi^0[\mathcal{H}]$ and $\chi^1[\mathcal{H}]$ defined in (16) are well-defined, and they are choice maps. Thus the two set-weighted functions $R^0[\mathcal{H}]$ and $R^1[\mathcal{H}]$ that they define are well-defined order functions.*

Proof. Given $(i_1, \ldots, i_n) \in \mathcal{T}_n$, we have $\{i_j \mid j \leq 0\} = \emptyset \notin \mathcal{H}$ and $\{i_j \mid j \leq n\} = \mathcal{I}_n \in \mathcal{H}$ (by (ii)). There is thus some $t \in \mathcal{I} - n$ which is the smallest $a \in \mathcal{I}_n$ such that $\{i_j \mid j \leq a\} \in \mathcal{H}$. Then for every $b \in \mathcal{I}_n$ such that $t < b$, we have $\{i_j \mid j \leq b\} \in \mathcal{H}$ by (i). Thus t is the unique $a \in \mathcal{I}_n$ such that $\{i_j \mid j \leq a\} \in \mathcal{H}$ and $\{i_j \mid j < a\} \in \mathcal{H}$. Therefore $\chi^0[\mathcal{H}]$ is well-defined by (16). We show similarly that $\chi^1[\mathcal{H}]$ is well-defined.

To prove that the two functions $R^0[\mathcal{H}]$ and $R^1[\mathcal{H}]$ are well-defined, we only have to show that the two maps $\chi^0[\mathcal{H}]$ and $\chi^1[\mathcal{H}]$ are choice maps. It is sufficient to show that they satisfy the condition (ii) in Proposition 1. Suppose that we have $(i_1, \ldots, i_n)$ and $(j_1, \ldots, j_n) \in \mathcal{T}_n$ and $a, b \in \mathcal{I}_n$ with $a \leq b$, such that

$$\{i_k \mid k < a\} = \{j_k \mid k < a\} = L$$

and

$$\{i_k \mid a \leq k \leq b\} = \{j_k \mid a \leq k \leq b\} = M.$$

We must prove that for $\alpha = 0, 1$, if $\chi^\alpha[\mathcal{H}](i_1, \ldots, i_n) \in M$, then $\chi^\alpha[\mathcal{H}](j_1, \ldots, j_n) \in M$.

By (16), if $\chi^0[\mathcal{H}](i_1, \ldots, i_n) \in M$, then there is some $t \in \mathcal{I}_n$ with $i_t \in M$, such that $\{i_k \mid k < t\} \notin \mathcal{H}$ and $\{i_k \mid k \leq t\} \in \mathcal{H}$. By property (i) above, this implies that $L \notin \mathcal{H}$ and $L \cup M \in \mathcal{H}$. Again by property (i), there is an integer s such that $a \leq s \leq b$, $\{j_k \mid k < s\} \notin \mathcal{H}$, and $\{j_k \mid k \leq s\} \in \mathcal{H}$. Thus $\chi^0[\mathcal{H}](j_1, \ldots, j_n) \in M$.

Let $N = \{i_k \mid b < k\} = \{j_k \mid b < k\}$. We show similarly that $\chi^1[\mathcal{H}](i_1, \ldots, i_n) \in M$ implies that $N \notin \mathcal{H}$ and $N \cup M \in \mathcal{H}$, which implies in turn that $\chi^1[\mathcal{H}](j_1, \ldots, j_n) \in M$.

Thus $\chi^0[\mathcal{H}]$ and $\chi^1[\mathcal{H}]$ are choice maps. ∎

Given the equivalence between the representations (15) and (16), it follows that set-weighted rank functions defined by (15) are well-defined order functions.

The reader may have noted that by Lemma 8 (i) and (ii), for any order function f_D and $\alpha \in B$, $\mathcal{H}_\alpha[f_B]$ is a heavy set collection in $\mathcal{I}_n$. This is no coincidence, and we will show in the next proposition that the two maps $f_D \mapsto \mathcal{H}_\alpha[f_B]$ and $\mathcal{H} \mapsto R^\alpha[\mathcal{H}]_D$ establish a one-to-one correspondence between order functions on D^n and heavy set collections in $\mathcal{I}_n$:

Proposition 12. *Let $\alpha = 0$ or 1. Then:*

(*i*) *Given an order function f_D with choice map χ, $f_D = R^\alpha[\mathcal{H}_\alpha[f_B]]_D$, that is $\chi = \chi^\alpha[\mathcal{H}_\alpha[f_B]]$. In particular, every order function is a set-weighted rank function.*

(*ii*) *Given a heavy set collection $\mathcal{H}$ in $\mathcal{I}_n$, $\mathcal{H} = \mathcal{H}_\alpha[R^\alpha[\mathcal{H}]_B]$.*

Proof. (*i*) For any $k \in \mathcal{I}_n$ and $(i_1, \ldots, i_n) \in \mathcal{T}_n$, let $i_t = \chi(i_1, \ldots, i_n)$, and take $x_1, \ldots, x_n \in B$ such that for any $j \in \mathcal{I}_n$,

$$x_{i_j} = \begin{cases} 0 & \text{if } j \leq k, \\ 1 & \text{if } j > k. \end{cases}$$

Then $x_{i_1} \leq \ldots \leq x_{i_n}$ and so we have

$$f_D(x_1, \ldots, x_n) = x_{i_t} = \begin{cases} 0 & \text{if } t \leq k, \\ 1 & \text{if } t > k. \end{cases}$$

Now $f_D(x_1, \ldots, x_n) = \delta(0, f_B, \{i_j \mid 1 \leq j \leq k\})$ (by (8) and (9)), and so $\{i_j \mid 1 \leq j \leq k\}$ is 0-heavy for f_B iff $k \geq t$. Thus $\{i_j \mid j < t\} \notin \mathcal{H}_0[f_B]$ and $\{i_j \mid j \leq t\} \in \mathcal{H}_0[f_B]$. By (16) this means that $\chi = \chi^0[\mathcal{H}_0[f_B]]$.

Now $\{i_j \mid j \geq t\} = \{i_j \mid j < t\}^c$ and $\{i_j \mid j > t\} = \{i_j \mid j \leq t\}^c$, and so Lemma 8 (*iv*) implies that $\{i_j \mid j \geq t\} \in \mathcal{H}_1[f_B]$ and $\{i_j \mid j > t\} \notin \mathcal{H}_1[f_B]$. By (16) this means that $\chi = \chi^1[\mathcal{H}_1[f_B]]$.

Thus $f_D = R^\alpha[\mathcal{H}_\alpha[f_B]]_D$ for $\alpha = 0, 1$.

(*ii*) Let $\mathcal{H}'$ be a heavy set collection such that $R^\alpha[\mathcal{H}] = R^\alpha[\mathcal{H}']$. Then $\chi^\alpha[\mathcal{H}] = \chi^\alpha[\mathcal{H}']$. Let S be a subset of size k of $\mathcal{I}_n$, where $0 \leq k \leq n$. There is then some $(i_1, \ldots, i_n) \in \mathcal{T}_n$ such that $S = \{i_j \mid 1 \leq j \leq k\}$. By (16), $S \in \mathcal{H}$ iff $\chi^\alpha[\mathcal{H}](i_1, \ldots, i_n) = \chi^\alpha[\mathcal{H}'](i_1, \ldots, i_n) = i_t$ for some $t \leq k$, in other words iff $S \in \mathcal{H}'$. Thus $\mathcal{H}' = \mathcal{H}$.

Applying (*i*) with $R^\alpha[\mathcal{H}]$ and $\chi^\alpha[\mathcal{H}]$ for f and χ, we obtain $\chi^\alpha[\mathcal{H}] = \chi^\alpha[\mathcal{H}_\alpha[R^\alpha[\mathcal{H}]_B]]$. By the preceding paragraph, this means that $\mathcal{H} = \mathcal{H}_\alpha[R^\alpha[\mathcal{H}]_B]$. ∎

Proposition 12 states that if $\mathcal{F}(n, D)$ is the set of order functions on D^n and $\mathcal{S}(n)$ is the set of heavy set collections in $\mathcal{I}_n$, then the two maps $f_D \mapsto \mathcal{H}_\alpha[f_B]$ and $\mathcal{H} \mapsto R^\alpha[\mathcal{H}]_D$ constitute a bijection $\mathcal{F}(n, D) \to \mathcal{S}(n)$ and its inverse. This explains why we called the elements of $\mathcal{H}_\alpha[f_B]$ "α-heavy sets", since they are the heavy sets in a set-weighted rank function $R^\alpha_{K;W}$ (in other words, the sets P having $W(P) \geq K$).