Journal of Convex Analysis
Volume 27 (2020), No. 1, 53–78

Echelons of Sets on Dispersion Gaps: Tools for Cluster Analysis

Jean-Pierre Aubin*
VIMADES, 14 rue Domat, 75005 Paris, France
aubin.jp@gmail.com

Dedicated to Alexander Ioffe on the occasion of his 80th birthday.

Received: October 25, 2018
Revised manuscript received: November 18, 2018
Accepted: November 29, 2018

Instead of analyzing time series of vectors or the problem of an allocation of vectors to clusters in vectors spaces, we shall investigate the same issues for time series and clusters of subsets K ranging over the hyperset $\mathcal{P}(X)$ of subsets of a set X of a plain set X deprived of any mathematical structure, let it be vectorial or topological. The arithmetic operations on vector spaces will be replaced by Boolean operations on hyperspaces.

For that purpose, we rely on the ideas going back to Abraham de Moivre based on

1. *dispersion gaps* $[[A_1, A_2]](K)$ between two disjoint subsets A_1 and A_2 (called *gists* of the dispersion gap) of subsets K such that $A_1 \subset K \subset \complement A_2$ (instead of dispersion intervals $[v_1, v_2] \subset \mathbb{R}$);
2. *magnitudes* which are increasing hyperfunctions $\mu\colon K \in \mathcal{P}(X) \mapsto \mu(K) \in \mathbb{R}_+$ vanishing at the empty set (encompassing measure, capacities, etc.) of all denominations.

The main instrument of measure of a set $K \in [[A_1, A_2]]$ between its two *gists* is its *echelon*

$$\mathbb{A}_\mu[[A_1, A_2]](K) \ := \ \frac{\mu(K \cap \complement A_1) - \mu(\complement(K) \cap \complement A_2)}{\mu(\complement A_1 \cap \complement A_2)} \ \in \ [-1, +1]$$

Its inverse associating with any echelon

$$e \in [-1, +1] \rightsquigarrow \mathbb{A}_\mu[[A_1, A_2]]^{-1}(e) \ \in \ [[A_1, A_2]]$$

the subsets sharing the same echelon. It plays the same role as the quantiles in statistics: it assigns the minimum value -1 to A_1 (instead of quantile 0) and $+1$ at $\complement A_2$, as the quantile 1. The subset $\mathbb{A}_\mu[[A_1, A_2]]^{-1}(0)$ plays the role of the median (instead of quantile 1/2). Actually, since we shall use the lattice operations sup and inf instead of the usual Kolmogorov measures $\int$, we were lead to use this new renormalization rule to compare all kind of magnitudes (Maslov measures, for instance).

We shall use magnitudes and echelons of sets to study time series of sets and clustering issues.

*The author thanks warmly Lionel Thibault for his corrections and suggestions, and Olivier Dordan and Pierre-Cyril Aubin-Frankowski for discussions on this topic. This work was supported by the US Air Force Grant FA9550-18-1-0254, "From vector spaces to hyperspaces, hypermaps and relations".

ISSN 0944-6532 / $ 2.50

Keywords: Abraham de Moivre, Boolean structure, Choquet capacity, cluster, dispersion gap, echelon, extremal box, extremal envelope, extremal quantile, Galois filtration, hyperset, Kolmogorov measure, magnitude, Maslov measure, set filtration, Tukey.

2010 Mathematics Subject Classification: 34G25, 34A60, 49J27, 49J53, 93B03, 37B55, 28B20, 45N05.

1. Introduction

Time series of sets and cluster analysis are investigated with the help of "echelons" built from 'dispersion gaps'" and 'magnitudes" we have to define first.

1.1. Magnitudes and Echelons

We introduce

1. *Magnitudes*: They are hyper-functions $\mu\colon K \in \mathcal{P}(X) \mapsto \mu(K) \in \mathbb{R}_+$ to evaluate subsets. There are many examples of such maps, for instance Kolmogorov and Maslov measures, Choquet capacities[1], etc., obeying different types of axioms, so that they became too polysemous.
   Fortunately, they all satisfy at least these two requirements :
   (a) $\mu(\emptyset) = 0$;
   (b) μ is increasing : if $K \subset L$, then $\mu(K) \leq \mu(L)$;
   so that we call *magnitudes* these increasing non negative set-defined functions.
2. *Dispersion intervals*: Given two disjoint subsets A_1 and A_2 of X, we define the *dispersion gap* $[[A_1, A_2]]$ as the family of subsets K satisfying $A_1 \subset K \subset \complement A_2$.
3. *Echelons*: The situation is a little complicated since, in many examples, we are lead to introduce three different magnitudes[2] $\mu_\downarrow \leq \mu \leq \mu_\uparrow$ for defining three kinds of echelon, the lower, central and upper *set echelon functions* of K in the dispersion A_1 and $\complement A_2$ covering different scenarii[3]

$$\mathbb{A}_{(\mu_\downarrow,\mu_\uparrow)}[[A_1, A_2]](K) := \frac{\mu_\downarrow(K \setminus A_1) - \mu_\uparrow(\complement(K) \setminus A_2)}{\mu_\uparrow(\complement(A_1 \cup A_2))} \in \left[-1, \frac{\mu_\downarrow(\complement(A_1) \setminus A_2)}{\mu_\uparrow(\complement(A_1 \cup A_2))}\right]$$

[1] The term *capacities* is quite polysemous. Following *Florin Vasilesco* [43] in 1937 and *Henri Cartan* [20] in 1945, *Gustave Choquet* provided in 1955 a comprehensive study [22] on capacities for solving problems in potential theory, and nowadays, to *quantile regression* (see the book *Quantile Regression* by *Roger Koenke* [28]). For a recent survey, see [24, Grabisch]. The Choquet capacity is defined on topological spaces as increasing and continuous from the right set function. Capacities are required to satisfy one of several properties for such and such purpose, such as monotonicity, supremum morphism, normalization, modularity, etc. *Gustave Choquet* also introduced exponentials, functions $\psi\colon K \in \mathcal{P}(X) \mapsto [0, 1]$ satisfying $\psi(K \cup L) = \psi(K)\psi(L)$. Choquet conjugates are also used as a way of measuring the expected utility of uncertain events, a topic of an abundant literature. The links with non smooth analysis were investigated in [41].

[2] For instance, magnitudes associated with lower Maslov measures $\mu_\downarrow$, Kolmogorov measures μ (the central ones used in standard statistics) and upper Maslov measures $\mu_\uparrow$.

[3] Recall that the set-difference is defined by $A \setminus B := A \cap \complement B$ so that, if $A \subset B$, then $A \setminus B = \emptyset$ and $K \setminus B \subset K \setminus A$, and that whenever $K_1 \subset K_2$, $K_1 \setminus A \subset K_2 \setminus A$ and $B \setminus K_2 \subset B \setminus K_1$.

$$\mathbb{A}_\mu[[A_1, A_2]](K) := \frac{\mu(K \setminus A_1) - \mu(\complement(K) \setminus A_2)}{\mu(\complement(A_1 \cup A_2))} \in [-1, +1]$$

$$\mathbb{A}_{(\mu_\uparrow, \mu_\downarrow)}[[A_1, A_2]](K) := \frac{\mu_\uparrow(K \setminus A_1) - \mu_\downarrow(\complement(K) \setminus A_2)}{\mu_\downarrow(\complement(A_1 \cup A_2))} \in \left[-1, \frac{\mu_\uparrow(\complement(A_1) \setminus A_2)}{\mu_\downarrow(\complement(A_1 \cup A_2))}\right]$$

Their values increase from -1 at A_1.

The inverses of each echelon maps assign to each (value of) echelons the family of subsets $K \in [[A_1, A_2]]$ evaluated by the same echelon and play the roles of quantiles in standard statistics for density probabilities. For instance, for the central echelon, the inverse $e \in [-1, +1] \rightsquigarrow \mathbb{A}_\mu[[A_1, A_2]]^{-1}(e)$ assigns to each (value of) echelons the family of subsets $K \in [[A_1, A_2]]$ evaluated by the same echelon $e \in [-1, +1]$, and play the role of *quantiles*[4] in statistics. By analogy, we suggest to call them also *quantile maps* to avoid an inflation of terminology. Hence, we shall obtain *upper median* and *lower median* surrounding the *central median* associated with μ.

They provide *measures of proximity* of a subset $K \in [[A_1, A_2]]$ to the gist of the dispersion gap: the subsets K associated with negative echelons are closer to A_1 and the ones with positive echelons closer to $\complement A_2$.

1.2. Clustering time series of sets

We consider a finite[5] *dispersion interval* $\{0, N\}$ of $N + 1$ scalars between 0 and N of scalars $n \in \mathbb{N}$. They are usually regarded as *dates* indexing a *series* of experiments described by subsets $K_n \subset X$. This justifies the term "time series" to qualify these sequences, even though the integers $n \in \mathbb{R}$ can describe other items than dates. For instance, the scalar can stand for incomes in economics (for studying social inequalities[6]), energy in physics, temperature in thermodynamics, velocities of vehicles in traffic studies[7], etc.

We can associate with each date $p \in \{0, N\}$ its *"sliding" downstream dispersion interval* $\{0, p-1\}$ *between the origin and the current date and downstream dispersion interval* $\{p, N\}$ *between the next date and the end* to take into account and compare what happened before and will happen this date (and not only what happened before).

We assume that lower and upper *magnitudes* $\mu_\downarrow \leq \mu_\uparrow$ are available to evaluate these subsets by $\mu_\downarrow(K_n)$ (thick horizontal) and $\mu_\uparrow(K_n)$ (dotted horizontam). For simplicity, we assume that the indexes $n \in \{0, N\}$ are integers. We shall address a

[4] They are the inverses of the density probabilities defined on "quantiles" $e \in [0, 1]$ instead of echelons $e \in [-1, +1]$. The median $1/2$ is replaced by 0 and the quartiles $1/4$ and $3/4$ are replaced by $-1/2$ and $+1/2$ respectively.

[5] For simplicity of the exposition, we consider here a finite numbers of scalars $n \in \mathbb{N}$, not necessarily integers, even though this is the most frequent cases.

[6] See for instance *Le Capital au* XXI^{e} *Siècle* by *Piketty* [37].

[7] See for instance *Traffic Networks as Information Systems. A Viability Approach* [13].

clustering issue to allocate them into *extremal boxes* displayed in Figure 1.1 which are constructed only with infimum and supremum operations:

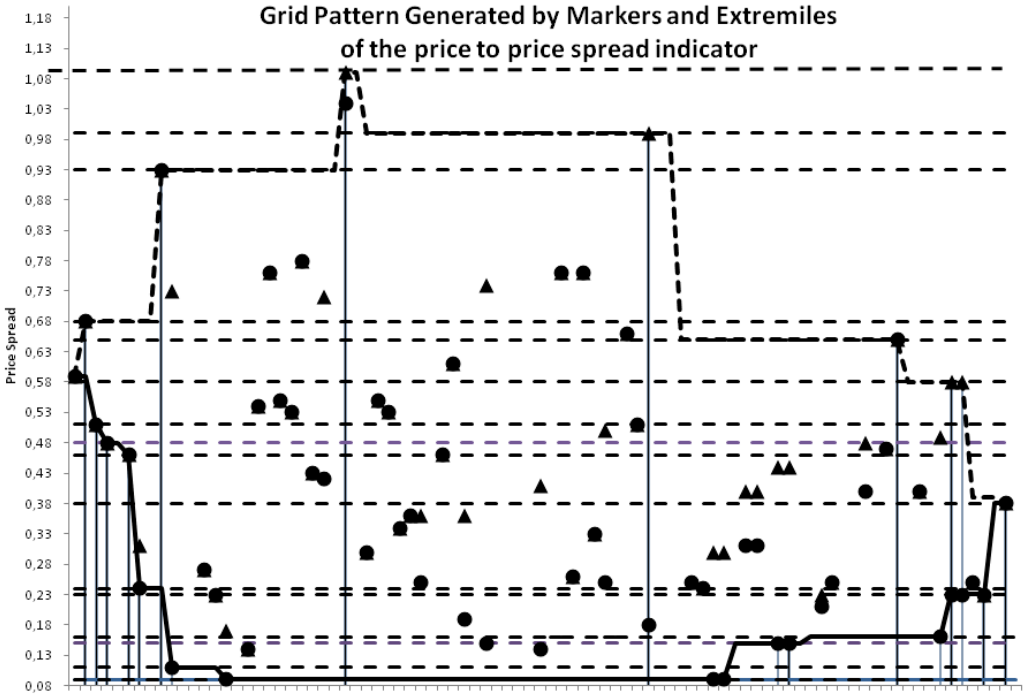


Figure 1.1: Extremal Envelope and Clustering of the Magnitudes

In Figure 1.1 the lower and upper observations $\mu_\downarrow(K_p)$ and $\mu_\uparrow(K_p)$ are in thick dotted horizontal and thick horizontal respectively, the vertical and horizontal bars delineating the extremal boxes are indicate the markers (signals) and extremiles (observations) defined in Definition 4.1, p. 69. The upper indicators are depicted by triangles and the lower ones by bullet (they may coincide). For that purpose, we shall distinguish four situations in formula (1), p. 69, by using:

1. *upper* magnitudes $\mu_\uparrow(K_p)$ during *upstream* dispersion intervals $\{0, n-1\}$;
2. *upper* magnitudes $\mu_\uparrow(K_p)$ during *downstream* dispersion intervals $\{n, N\}$;
3. *lower* magnitudes $\mu_\downarrow(K_p)$ during *upstream* dispersion intervals $\{0, n-1\}$;
4. *lower* magnitudes $\mu_\downarrow(K_p)$ during *downstream* dispersion intervals $\{n, N\}$.

1.3. Cluster analysis

Cluster analysis[8] covers many algorithms, among them (statistical) distribution models, graph-based models, neural network models, centroid models (e.g., the *k-means* algorithm and its numerous variants), density models, group models, etc.

It is a currently used as a nickname of *automatic classification*, or, in more pedantic terminology, *botryology*, from the Greek *botros* (grape). Other synonyms are

[8] Cluster analysis started in anthropology in 1932 by Driver and Kroeber and in psychology by Zubin in 1938, Tryon in 1939 and by Cattell in 1943.

numerical taxonomy and *typological analysis*. Together with pattern recognition, segmentation and many other techniques, clustering is used to detect alarms, anomalies or signals: see for instance *Diday*'s *dynamical clustering*[9], *Vapnik*'s *support vector machines and networks*, Pernot's *Choix d'un classifieur en discrimination*, [36], *neural networks*, etc.

These techniques are part of contemporary exploratory data analysis.

We begin with an initial *cluster* of ℓ given pairwise disjoint subsets $A_i \subset X$ of the set X. We shall call these subsets the (initial) *gists* of the cluster allocation mechanism. When the *gists* defining a cluster boil down to vectors of a vector space, we recover the standard cases of clustering allocation mechanisms.

The problem is to *allocate* (or *assign*) any subset K to the "closest" gist A_i of an initial cluster of ℓ given pairwise disjoint subsets A_i to form a new gist $A_i^2 := K \cup A_i$ defined a new cluster. At each step of the process, these sorting and clustering mechanisms are done *automatically* for providing new gists A_i^n associated with a subset K^{n-1}. Iterating this process, we can classify new sets, and thus, build the gists of new clusters.

This iterative process is not really an algorithm, since, at each step, the new subset K^n is not automatically computed, but provided from other sources. In this sense, this process is partly automatic for computing at each step a new gist A_i^n as soon as a new subset is provided. It can be regarded as an *adaptive algorithm*, in the sense that the construction of new clusters depends of new subsets, whenever no other algorithm is available to furnish the subsets to allocate.

This being said, we have to design a tool to choose a subset closer to one of the gists of a cluster. For that purpose, we suggest to use the concepts of *echelons* of sets to assign any set to one of the two gists of a dispersion gap.

2. The De Moivre heritage

We have to justify the introduction of statistics of subsets of a set X instead of its elements. To provide just one example, in our times greedy for money, at each day, the stock markets associate with each stock the *interval between the low and high prices* during this day, and, also, a selection of this interval, the closing price. Intervals are subsets and provide much more information than just the closing prices. Intervals embody also a description of *uncertainty* or chance, in a different way than probability laws do. Stock markets by the way are unable to provide their density distributions, leaving this task to financial analysts[10]. Moreover, closing prices, for instance, are highly oscillating between local suprema (modes) and infima, and thus "megamodal" rather than unimodal or multimodal : they often do not resemble any of the two hundred families[11] of probability densities

[9] See the bibliography in *Symbolic Data Analysis: Conceptual Statistics and Data Mining*, [18].

[10] See, for instance, *Tychastic Measure of Viability Risk*, [12].

[11] In reference to the political myth of the "two hundred families" denounced by *Georges Duchêne*, collaborator of *Pierre-Joseph Proudhon*, to stigmatize the financial "feudalism" in which "the 20 billion securities are at the discretion of 200 nabobs"", who have not committed

catalogued to date. Digital computers offer nowadays techniques enabling us to compensate such losses by analyzing subsets with a manifold of techniques (mathematic morphology, wavelets, neural networks, etc.).

For instance, once a set is given as a datum to observe, the information provided by its boundary is as relevant and important than the one provided by just its *center of gravity*[12], often too succinct to be an efficient summary. Even *Adolphe Quetelet*, author of *Treatise on Man & the Development of His Faculties*[13], a strong advocate of prevalence of the "average", nevertheless constantly returned to the terms of minima, maxima and gaps as soon as he mentions the average.

It may thus be useful to go back to *Abraham de Moivre*[14] and understand how he set the framework of data analysis:

Abraham de Moivre focused his attention in his book *The Doctrine of Chances* [33] of 1718 laying down the foundations of what was not yet known as statistics, even less "big data analysis", to both

1. data ranging over a *dispersion interval*;
2. numerical measure of the observations of these data, described by an indicator $\varphi\colon x \in X \mapsto \varphi(x) \in \mathbb{R}$.

At his time, experimental observations lacking, he used the Newtonian binomial coefficients as an example of observations, and, passing to the limit, forged the famous "bell curve", nowadays attributed to *Gauss* under the name of normal probability density. He paved the way to an exogenous approach, the "parametric way" followed by many mathematicians ever since.

Nowadays, it is time to study the observations of intervals $\Phi : x \in \mathbb{R} \rightsquigarrow \Phi(x) := [\Phi^\flat(x), \Phi^\sharp(x)] \subset \mathbb{R}$ instead of elements and go from and exogenous approach to an endogenous one.

200 million. He was jailed several times for his writings. Later, the term referred to the two hundred largest shareholders (out of nearly 40,000) who constituted the General Assembly of the Banque de France when it was created. Here too, we advocate to avoid one of this family of analytical densities, and thus, to "estimate" them.

[12] As well as the *Steiner point* which is also a selection of the convex hull of a subset.

[13] See [38, Quetelet], in which he claims that *The determination of the average man is not merely a matter of speculative curiosity; it may be of the most important service to the science of man and the social system. It ought necessarily to precede every other inquiry into social physics, since it is, as it were, the basis. The average man, indeed, is in a nation what the centre of gravity is in a body; it is by having that central point in view that we arrive at the apprehension of all the phenomena of equilibrium and motion.*

[14] Born in France two years after the death of Fermat, exiled in England because the criminal Edict of Fontainbleau (1685) persecuting Huguenots for their religion, he took up the book *Essai d'analyze sur les jeux de hasard*, [34], by *Pierre Rémond de Montmort* to write *The Doctrine of Chances* en 1718. Later, in 1733, he introduced the binomial indicator, and observed that it converges to a "bell curve", the *normal indicator*, as well as a special case of the central limit theorem, later called the de Moivre-Laplace theorem, among many other achievements. Newton told to his visitors: "Go to Mr. De Moivre; he knows these things better than I do". He noted that he was sleeping an extra 15 minutes each night and correctly calculated the date of his death as the day when the sleep time reached 24 hours.

2.1 An exogenous approach

It consists in *inventing* probability densities (depending on several parameters) by

(a) *building explicit analytical functions* motivated by typical "mathematical stories" making up a catalogue of about two hundred families of "parameter dependent" prototypes produced by an efficient "bell curve foundry" for furnishing means, modes, variance, to which was added the standard deviation[15], quantiles , and so on, as well as a myriad of results obtained so far by following this venerable tradition;

(b) *deriving* as many mathematical properties of these densities as possible;

(c) *approximating* observed probability densities by choosing (and estimating) one family of probability densities and computing its parameters for obtaining a reasonable estimation or approximation of the observed indicators of data and benefit of its properties.

This is the task of statistics[16], nowadays known as *data analysis*, "big" as it should since the apparition of digital computers.

2.2 Is the exogenous approach indispensable?

Not always, since we have quite often under our eyes bounded dispersion intervals, whilst the support of the most famous density functions is not bounded, out of reach of computers. We may avoid factitious and cumbersome debate on more or less thicker tails.

The exogenous approach strategy building up *a priori* models of probability densities may induce *undue losses of information*. Furthermore, many highly oscillating experimental indicators or provability densities are far away from the densities of the catalogue of nearly two hundred available parameterisable analytical densities: they are "megamodal" rather than multimodal, oscillating between many local maxima (modes) and minima, for instance.

However, whenever real experimental measures are not available, the immense reservoir of extensively studied examples of non experimental dispersion functions defined on dispersion intervals. or indicator maps relevant to such or such problem

[15] The term "standard deviation" was used for the first time by *Karl Pearson* in 1893 (who introduced the notation σ) and coined the word *histogram*, invented by *Oresme* in 1370. The concept of "variance" was introduced in 1918 by *Ronald Fisher* (not to be confused with *Irving Fisher*) in *The Correlation between Relatives on the Supposition of Mendelian Inheritance*, [23].

[16] Naturally, the origin of "statistics" activity goes back to the origins of History, since the use of numbers to provide the size of a population of soldiers to their rulers, whatever their regal power purposes. Demography, insurance and biology were other main sources of motivations. The formalization of statistical concepts have their origins in several books in the seventeen century, among which, *Natural and Political Observations upon the Bills of Mortality*, by *John Graunt* in 1663 and *political arithmetic* by *Sir William Petty* in 1676, without forgetting *Richard Cantillon* and *Condillac*. The sovereign and regal power using these observations lead *Gottfried Achenwall*, the "father of statistics", according to German economists, to suggest in 1752 the word *Statistik* in his *Staatsverfassung der Europäischen Reiche im Grundrisse*, in reference to the Latin *statisticum collegium* ("state council") and the Italian, *statista* "(statesman"). Up to the last letter "k", it became the universal concept superseded nowadays by "(big) data sciences".

can be "invented" arbitrarily to supplement the absence of measures. They tell us mathematical stories justifying their choice. At least, the mathematical ingenuity developed for more than two centuries is the honor of the human mind[17].

2.3 An endogenous approach

John Tukey[18] advocated in 1977 in his book *Exploratory Data Analysis*[19] an endogenous approach we suggest to adopt and adapt: *observe* one indicator map defined on a dispersion interval provided by actual experiments or enquiries and derive *as endogenously as possible few characteristic* features.

Algorithms processed by more and more efficient digital computers allow us nowadays to bypass the approximation processes by analytical probability densities and still derive directly some statistical information, such as the concept of "quantiles" to divide the dispersion gap into contiguous sub dispersion gaps and design Tukey whisker boxes, for instance.

We thus suggest to return to *de Moivre* to reestablish the preeminent role of the *dispersion intervals* and, nowadays, *dispersion gaps* between sets, which are the new data we shall deal with:

The dispersion gaps between sets and set-valued indicators are the parameters of our study.

Therefore,

(a) their choice is no longer an issue we deal with, since we assume that any indicator map is defined on dispersion gaps to extract available raw information, which should not be muddled or distorted by too many preliminary operations performed on them;

(b) the results depend upon this dispersion gap as well as the indicator map on which it is defined.

2.4. The rise of sets as data for analysis

Nowadays, digital computers offer techniques allowing us to derive information from the original knowledge of subsets taking into account the uncertainty they embody.

Instead of density functions defined on dispersion intervals, we can study an interval valued density map or indicator $\Phi : x \in X \rightsquigarrow [\Phi^\flat(x), \Phi^\sharp(x)]$. Boolean operations as well as arithmetic ones, such as *interval analysis*[20], can be used.

[17] As *Carl Jacobi* wrote: *"It is true that Fourier had the opinion that the principal aim of mathematics was public utility and explanation of natural phenomena; but a philosopher like him should have known that the sole end of science is the honor of the human mind, and that under this title a question about numbers is worth as much as a question about the system of the world".*

[18] He also coined in 1948 the word *bit* as a contraction of "binary digit" in [42, Tukey], and next popularized by *Claude E. Shannon*, as well as the word *software* in 1958 (also claimed by *Paul Niquette* in 1953 when he was only 19 years old and writing programs on the SWAC hardware at UCLA, then regarded as a prank by colleagues).

[19] See *Exploratory Data Analysis*, [42, Tukey]

[20] "Interval Analysis" was first advocated by *Ramon E. Moore* easier to deal with than Boolean

When a finite number ℓ of indicators $\Phi_i : x \in X_i \rightsquigarrow [\Phi_i^\flat(x), \Phi_i^\sharp(x)]$ were involved, independent observations provided hypercubes $\Phi(x) := \prod_{i=1}^{\ell} [\Phi_i^\flat(x), \Phi_i^\sharp(x)]$ were adequate and justified, so that the observations became the hypercubes $\Phi(x) := \prod_{i=1}^{\ell} \Phi_i(x) \subset \mathbb{R}^\ell$. The observations were no longer numbers or intervals, but vectors and subsets $K \in \mathcal{P}(\mathbb{R}^\ell)$.

Next, the independence of observations left way to the analysis of correlated observations, so that they became subsets of hypercubes. When these correlations were not observed, but guessed mathematically, the first class of subsets investigated which were *ellipsoids*[21], described by a positively definite symmetric matrix $K \in \mathcal{L}(\mathbb{R}^l, \mathbb{R}^\ell)$ (interpreted as a covariance matrix). They define a scalar product $((x, y)) := \langle Kx, y\rangle$, allowing us to enjoy the properties of Hilbert spaces.

Nowadays, these Hilbert spaces (finite dimensional or not) can be derived from observations $h \in H$. Whenever H is any set supplied with a *reproducing kernel*[22] ersatz of scalar products (deprived of "bilinearity"). They allow us to "embed" their elements into a "reproducing kernel Hilbert space", for a scalar product and thus, a symmetric positively operator K on this Hilbert space, playing the role of covariance matrices.

Other relevant linear operators can be constructed, for instance tensor products of vectors, or of functions of these vectors (they are Hebbian matrices in neural networks[23], tensor products of velocities, accelerations[24], etc.).

In summary, these examples provide motivations to study not only elements of a set X (such as vectors), but also and over all subsets ranging over the hyperspace $\mathcal{P}(X)$ of subsets of X. Even when X is deprived of any mathematical structure, the hyperspace $\mathcal{P}(X)$ enjoys at least Boolean operations. This allows us to devise statistical inferences of time series of sets and/or clustering issues, which deal with subsets by purpose, without using arithmetic operations.

3. Set echelons of sets on dispersion gaps under magnitudes

Here we provide more details on magnitudes, dispersion gaps and echelons.

operations. This approach became the topic of an abundant literature (see for instance [25]). However, these arithmetic operations, extended to hypercubes, did not replace Boolean operations in many set-valued problems (since the union of hypercubes is not easily described by arithmetic operations).

[21] See for instance [29, Kurzhanski & Valyi] *Ellipsoidal Calculus for Estimation and Control*.

[22] Reproducing kernels are symmetric positive definite functions $k\colon (h_1, h_2) \in H \times H \mapsto \mathbb{R}$ introduced in 1950 by Aronszajn in [5], later used when they invented the Support Vector Machines.

[23] See for instance *Neural Networks and Qualitative Physics: a Viability Approach*, [7, Aubin].

[24] See for instance [11].

3.1. Magnitudes

The first element we need for defining echelons of sets in a dispersion gap is the concept of magnitude:

Definition 3.1. (Magnitudes) Any function $\mu\colon K \in \mathcal{A} \mapsto \mu(K) \in \mathbb{R}$ satisfying

(1) $\mu(\emptyset) = 0$;

(2) μ is increasing : if $K \subset L$, then $\mu(K) \leq \mu(L)$;

is called a *magnitude*. Recall that a *Boolean algebra* $\mathcal{A} \subset \mathcal{P}(X)$ is a family of subsets stable under unions and complementarity (and thus, under intersection). A magnitude $\mu\colon \mathcal{A} \mapsto \mathbb{R}_+$ is said to be a

(1) a *sup morphism* if

$$\forall\, K_1, K_2,\ \ \mu(K_1 \cup K_2) = \sup(\mu(K_1), \mu(K_2))$$

(2) an *inf morphism* if

$$\forall\, K_1, K_2,\ \ \mu(K_1 \cap K_2) = \inf(\mu(K_1), \mu(K_2))$$

(3) a *Boolean morphism* if it both a sup and an inf morphism,

(4) is *additive* if $\mu(\emptyset) = 0$ and for all disjoint subsets $A, B \in \mathcal{A}$,

$$\mu(A \cup B) = \mu(A) + \mu(B).$$

There are as many echelons as magnitudes, among which the ones provided by measures.

Maslov and Kolmogorov magnitudes associated with indicators

Let $X := \mathbb{R}^\ell$ and $\Phi : X \rightsquigarrow \mathbb{R}$ be a set-valued map regarded as an *indicator*[25]. Historically, indicators were single-valued and associated with each data the number

[25] In other words, indicators are set-valued maps from $\mathbb{R}^\ell$ to $\mathbb{R}_+$. At the dawn of XXth century, the founders of set theory after *Georg Cantor*, *Felix Hausdorff*, *Maurice Fréchet*, *Kazimierz Kuratowski*, *René Baire*, to name a few, used the concept of set-valued maps and relations. *Nicolas Bourbaki* introduced in section 3.13 of *Theory of Sets* (*Théorie des ensembles*), [19, Bourbaki], in which he wrote *"[...] we have considered only the case of three sets merely to fix the ideas; analogous considerations hold for any finite number of sets"* (On n'a envisagé que le cas de trois ensembles, uniquement pour fixer les idées; des considérations analogues valent pour plusieurs ensembles en nombre fini quelconque). "He" next abandoned general relations which we could call *multinary* to concentrate on binary relations as the *Yamonamö*, still stone age tribes in the 1060's (see [21, *Noble Savages*] by *Napoleon Chagnon*) *had no word for numbers beyond* 2 *and had to sit down to estimate a long distance by using their toes*! Above all, "He" restricted "its" attention on input-output single-valued maps under the pretext that set-valued maps are single-valued maps from spaces to *hyperspaces* (families of subsets of a space, called also "power spaces"). Even category theory concentrates with input-output morphisms and functors. Unfortunately, hyperspaces do not faithfully inherit structures of the underlying spaces. This was not a problem as long as motivations for taking into consideration the set-valued character of maps popping up were ignored at that time and for such a long time, up to the 1960'. Since 1981, set-valued maps and relations could be differentiated both graphically (see [6, Aubin]) and pointwise (mutational analysis, see [8, Aubin]). These tools are not yet used at this stage of the study. See among many monographs *Set-valued analysis*, [15, Aubin & Frankowska],

of their occurrences, called *densities*[26]. When the indicator $\Phi := \{\varphi\}$ boils down to a single-valued and integrable non negative function φ, we recognize *density functions* (and, among them, whenever we assume further that their total measure is equal to 1, *density probabilities*).

The lower and upper *Maslov magnitudes* are the maps $\mu_\downarrow(\Phi)$ and $\mu_\uparrow(\Phi)$ defined by

$$\forall\ K \subset X,\ \begin{cases} \mu_\downarrow^\Phi(K) := \inf\limits_{x\in K} \inf\limits_{y\in\Phi(x)} y = \inf[\,\mathrm{Graph}(\Phi) \cap (K \times \mathbb{R})], \\ \mu_\uparrow^\Phi(K) := \sup\limits_{x\in K} \sup\limits_{y\in\Phi(K)} = \sup[\,\mathrm{Graph}(\Phi) \cap (K \times \mathbb{R})]. \end{cases}$$

Recall that we can define the Aumann integral of a set-valued map $\Phi\colon \mathcal{P}(X) \mapsto \mathbb{R}_+$ under a Kolmogorov measure $d\mu$ as the union

$$\int_K \Phi(x)d\mu(x) \ := \bigcup_\varphi \int_K \varphi(x)d\mu(x)$$

over the set of integrable selections $\varphi(x) \in \Phi(x)$. We shall set

$$\mu^\Phi(K) \ := \ \frac{\int_K \Phi(x)d\mu(x)}{\int_K d\mu(x)}$$

which is an additive magnitude, that we shall call *Kolmogorov magnitude*. We observe that

$$\forall\ K \in \mathcal{P}(X),\ \ \mu_\downarrow^\Phi(K) \ \le \mu^\Phi(K) \ \le \ \mu_\uparrow^\Phi(K).$$

3.2. Dispersion gaps between two disjoint sets

Back to *Abraham de Moivre*, who focused its attention to data ranging over a dispersion interval, which, nowadays, can be extended to dispersion gaps of subsets $K \in \mathcal{P}(X)$ in the following way:

For that purpose, we begin to introduce other operations on the hyperspace $\mathcal{P}(X)$:

Definition 3.2. (Demarcation between subsets) The *demarcation* operation "$\mho$" between two subsets is defined by

$$A_1 \mho A_2 \ := \ \complement A_1 \cap \complement A_2 \ = \ \complement(A_1 \cup A_2) \ = \ \complement A_2 \setminus A_1 \ = \ \complement A_1 \setminus A_2.$$

Variational Analysis, [39, Rockafellar & Wets], *Viability Theory. New Directions*, [10, Aubin, Bayen & Saint-Pierre], *Mutational and Morphological Analysis*, [8, Aubin], *Mutational Analysis. A Joint Framework for Cauchy Problems in and Beyond Vector Spaces*, [30, Lorenz], *derivatives of sets in measure* based on the Steiner formula for convex bodies by *Estáte Khmaladze*, [26, Khmaladze], [27, Khmaladze & Veil], etc.

[26] They gave their name to *density functions*, even when the values of the indicators were no longer the number of occurrences, leading sometimes to confusion. Since this is not always the case, we use the generic term of indicator, above all in the cases when indicators are intervals and not only singletons.

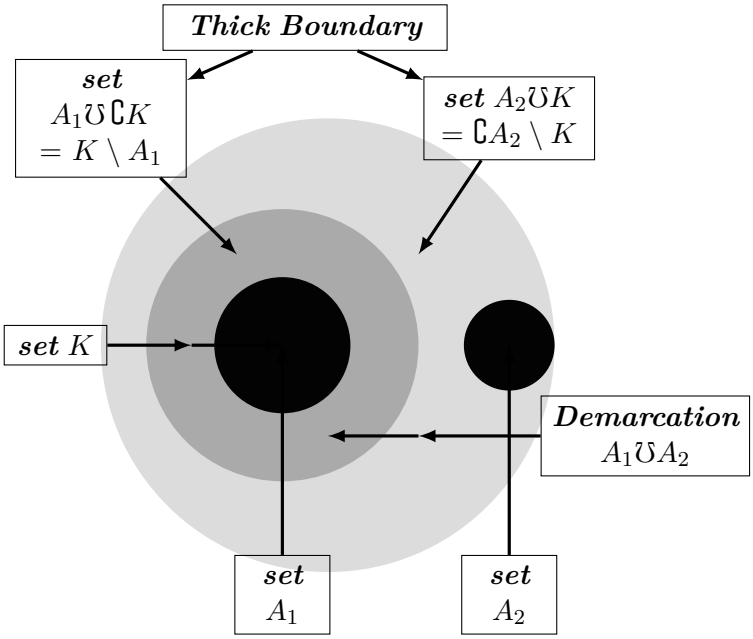
We observe that $A \,\mho\, \complement A = \emptyset$, and $A \,\mho\, A = \complement A$

Therefore, $X = [(A_1 \setminus A_2) \cup (A_2 \setminus A_1)] \cup (A_1 \cap A_2) \cup (A_1 \,\mho\, A_2)$

is a quadripartition of the space X. When A_1 and A_2 are disjoint, this boils down to the tripartition:

$$A_1 \cap A_2 = \emptyset \text{ implies } X = A_1 \cup A_2 \cup (A_1 \,\mho\, A_2)$$

Recall that a subset K of a topological space is not *connected* it there exists at least a *partition* $K = \overset{\circ}{A}_1 \cup \overset{\circ}{A}_2$ by two *nonempty open subsets* such that the demarcation $\overset{\circ}{A}_1 \,\mho\, \overset{\circ}{A}_2 \neq \emptyset$ is a *non empty* closed subset[27]. The union of two disjoint subsets is not connected.



Definition 3.3. (Dispersion gaps between two disjoint subsets) The *dispersion gap* $[[A_1, A_2]]$ between two disjoint subsets $A_1 \in \mathcal{A}$ and $A_2 \in \mathcal{A}$ is the family of subsets $K \in \mathcal{A}$ satisfying $A_1 \subset K \subset \complement A_2$.

Observe that $K \in [[A_1, A_2]]$ if and only if $\complement K \in [[A_2, A_1]]$.

Example 3.4. (Echelons of intervals) When $X := \mathbb{R}$ boils down to the real line $\mathbb{R}$, we consider the family of intervals $K :=]-\infty, v]$ indexed by $v \in \mathbb{R}$. We introduce two disjoint gists $A_1 :=]-\infty, v^\flat]$ and $A_2 := [v^\sharp, +\infty[$ whenever $v^\flat < v^\sharp$, and we observe that

(1) the dispersion interval $[[A_1, A_2]] = [v^\flat, v^\sharp]$

(2) $K \setminus A_1 = [v^\flat, v]$

(3) $\complement K \setminus A_2 = [v, v^\sharp]$ □

[27] Note that when X is a topological vector space, then whenever $y \in \overset{\circ}{K}$ and $z \in \complement \overset{\circ}{K}$, then there exists $\lambda \in [0, 1]$ such that $\lambda y + (1 - \lambda)z$ belongs to the boundary ∂K of K.

3.3. Echelons of sets of a dispersion gap under magnitudes

Definition 3.5. (Upper and lower echelons) Let us consider

(1) Two disjoint subsets $A_1 \in \mathcal{A}$ and $A_2 \in \mathcal{A}$ and their dispersion gap $[[A_1, A_2]]$;

(2) Three magnitudes $\mu_\downarrow \leq \mu \leq \mu_\uparrow$.

We associate with any subset $K \in [[A_1, A_2]]$ in the dispersion gap the lower, central and upper set echelons:

(1) the *lower set echelon function* defined by

$$\mathbb{A}_{(\mu_\downarrow,\mu_\uparrow)}[[A_1, A_2]](K) := \frac{\mu_\downarrow(A_1 \mho \complement K) - \mu_\uparrow(A_2 \mho K)}{\mu_\uparrow((A_1 \mho A_2))} \in \left[-1, \frac{\mu_\downarrow((A_1 \mho A_2))}{\mu_\uparrow(A_1 \mho A_2)}\right],$$

(2) the *central set echelon function* defined by

$$\mathbb{A}_{\mu}[[A_1, A_2]](K) := \frac{\mu(A_1 \mho \complement K) - \mu(A_2 \mho K)}{\mu(A_1 \mho A_2)} \in [-1, +1]\,,$$

(3) the *upper set echelon function* defined by

$$\mathbb{A}_{(\mu_\uparrow,\mu_\downarrow)}[[A_1, A_2]](K) := \frac{\mu_\uparrow(A_1 \mho \complement K) - \mu_\downarrow(A_2 \mho K)}{\mu_\downarrow(A_1 \mho A_2)} \in \left[-1, \frac{\mu_\uparrow(A_1 \mho A_2)}{\mu_\downarrow(A_1 \mho A_2)}\right].$$

The *inverses of set echelons* play an important role, since they associate with echelons the subsets having the same echelon, playing the role of quantiles in usual statistics which share the value of a probability distribution instead of an echelon measuring the position of a subset of a dispersion gap $[[A_1, A_2]]$. Pairs of echelons $e_1 \leq e_2$ are use as to divide the dispersion gap in subsdispersion gaps of subsets.

Definition 3.6. (Quantile Maps) The inverses of lower, central and upper set echelons

$$\begin{cases} e \in \left[-1, \frac{\mu_\downarrow((A_1 \mho A_2))}{\mu_\uparrow(A_1 \mho A_2)}\right] & \rightsquigarrow \mathbb{A}_{(\mu_\downarrow,\mu_\uparrow)}[[A_1, A_2]]^{-1}(e) \\ e \in [-1, +1] & \rightsquigarrow \mathbb{A}_{\mu}[[A_1, A_2]]^{-1}(e) \\ e \in \left[-1, \frac{\mu_\uparrow(A_1 \mho A_2)}{\mu_\downarrow(A_1)}\right] & \rightsquigarrow \mathbb{A}_{(\mu_\uparrow,\mu_\downarrow)}[[A_1, A_2]]^{-1}(e) \end{cases}$$

are called *lower, central and upper quantile maps*, respectively.

Each of them measure proximity relations, since subsets K are closer to A_1 whenever their echelon is small.
Subsets K with negative echelons are closer to A_1, with positive echelons are closer to A_2 and the ones with zero echelons constitute the *lower, central and upper medians* respectively.

Whiskers boxes. As for quantiles, we can associate with each of these echelons the corresponding *Tukey whiskers boxes* designed in 1969 by *John Wilder Tukey* to classify the data and their indicators in few boxes.

Each of these boxes summarizes a partition of a dispersion gap by five numerical data – the minimum, the median, the first and third quartiles (corresponding to the echelons -1, 0, $-\frac{1}{2}$ and $+\frac{1}{2}$) and the maximum. They delineate two boxes separating the quartiles by the median[28] the median, extended by *whiskers*, which are the echelons of the gists A_1 and $\complement A_2$.

The *choice of the median and of the quartiles is arbitrary* for constructing these boxes, even though they are quite reasonable to design non parametric methods.

Naturally, the lower, central and upper echelons of sets of a dispersion gap share the same properties. For avoiding cumbersome duplication of results and definitions, we shall present them in the framework of lower echelons only whenever there is no risk of muddle.

First, we observe that

$$\mathbb{A}_{(\mu_\downarrow,\mu_\uparrow)}[[A_1,A_2]](K) \;:=\; -\frac{\mu_\downarrow(A_1 \mho A_2)}{\mu_\uparrow(A_1 \mho A_2)}\, \mathbb{A}_{(\mu_\uparrow,\mu_\downarrow)}[[A_2,A_1]](\complement K)$$

The echelon defines a measure of proximity of a subset K to either A_1 of A_2: We thus observe the following properties:

Proposition 3.7. (Properties of Set Echelons) *Let us consider the three magnitudes $\mu_\downarrow \le \mu \le \mu_\uparrow$ and $A_1 \subset \complement A_2$. Then the three associated set echelons satisfy*

$$\mathbb{A}_{(\mu_\downarrow,\mu_\uparrow)}[[A_1,A_2]](A_1) \;=\; \mathbb{A}_{\mu}[[A_1,A_2]](A_1) \;=\; \mathbb{A}_{(\mu_\uparrow,\mu_\downarrow)}[[A_1,A_2]](A_1) \;=\; -1$$

and

$$\mathbb{A}_{(\mu_\downarrow,\mu_\uparrow)}[[A_1,A_2]](\complement A_2) \le \mathbb{A}_{\mu}[[A_1,A_2]](\complement A_2) = +1 \le \mathbb{A}_{(\mu_\uparrow,\mu_\downarrow)}[[A_1,A_2]](\complement A_2)$$

The three set echelons of sets under an magnitude are magnitudes*:*

$$\text{if } K_1 \subset K_2, \text{ then } \begin{cases} \mathbb{A}_{(\mu_\downarrow,\mu_\uparrow)}[[A_1,A_2]](K_1) & \le \mathbb{A}_{(\mu_\downarrow,\mu_\uparrow)}[[A_1,A_2]](K_2) \\ \mathbb{A}_{\mu}[[A_1,A_2]](K_1) & \le \mathbb{A}_{\mu}[[A_1,A_2]](K_2) \\ \mathbb{A}_{(\mu_\uparrow,\mu_\downarrow)}[[A_1,A_2]](K_1) & \le \mathbb{A}_{(\mu_\uparrow,\mu_\downarrow)}[[A_1,A_2]](K_2) \end{cases}$$

Assume that the three magnitudes $\mu_\downarrow \le \mu \le \mu_\uparrow$ are Boolean morphisms*. So are the lower set echelon $\mathbb{A}_{(\mu_\downarrow,\mu_\uparrow)}[[A_1,A_2]]$, the central set echelon $\mathbb{A}_{\mu}[[A_1,A_2]]$ and the upper set echelons $\mathbb{A}_{(\mu_\uparrow,\mu_\downarrow)}[[A_1,A_2]]$.*

Remark 3.8. (Echelons between disjoint open subsets) When the sets $A_1 := \overset{\circ}{A}_1$ and $A_2 := \overset{\circ}{A}_2$ are disjoint open subsets satisfying $A_1 \subset \complement A_2$, then, for any subset

[28] *John Wilder Tukey* justified the use of the median and quartiles which can be defined for all probability densities, unlike the mean and standard deviation. Moreover, the quantiles are more robust to skewed or heavy-tailed distributions than the traditional mean and standard deviation

$K \in [[A_1, A_2]]$, $A_1 \subset \overset{\circ}{K} \subset K \subset \overline{K} \subset \complement A_2$. We thus can define the *lower topological echelons* by

$$\mathbb{A}_{(\mu_\downarrow,\mu_\uparrow)}[\overset{\circ}{A}_1, \overset{\circ}{A}_2](K) := \frac{\mu_\downarrow(\overset{\circ}{A}_1 \cup \complement K) - \mu_\uparrow(\overset{\circ}{A}_2 \cup \overset{\circ}{K})}{\mu_\uparrow(\overset{\circ}{A}_1 \cup \overset{\circ}{A}_2)}$$

as well as the central topological echelon $\mathbb{A}_\mu[\overset{\circ}{A}_1, \overset{\circ}{A}_2](K)$ and upper topological echelon $\mathbb{A}_{(\mu_\uparrow,\mu_\downarrow)}[\overset{\circ}{A}_1, \overset{\circ}{A}_2](K)$.

Proposition 3.9. (Echelon of sets of an additive magnitude) *Let us consider a (finite positive) additive magnitude $\mu\colon K \in \mathcal{A} \mapsto \mu(K) \in \mathbb{R}_+$ defined on a set algebra $\mathcal{A} \subset \mathcal{P}(X)$ and two disjoint subsets $A_1 \in \mathcal{A}_1$ and $A_1 \in \mathcal{A}$ such that $A_1 \subset \complement A_2$. The* Kolmogorov echelon $\mathbb{A}_\mu[[A_1, A_2]](K)$ *of the $K \in \mathcal{A}$ of the additive magnitude μ on the set algebra $\mathcal{A}$ is, for all $K \in [[A_1, A_2]]$, equal to*

$$\mathbb{A}_\mu[[A_1, A_2]](K) := \frac{2\mu(K) - (\mu(\complement A_2) + \mu(A_1))}{\mu(\complement A_2) - \mu(A_1)} \in [-1, +1].$$

Remark 3.10. (Choquet conjugate) The *Choquet conjugate* of a set function $\mu\colon \mathcal{P}(X) \mapsto \overline{\mathbb{R}}$ is defined by

$$\forall\, K \in \mathcal{P}(X),\ \ \mu^\blacklozenge(K) := -\mu(\complement K)$$

It is a bijective transform since $\mu^{\blacklozenge\blacklozenge} = \mu$. If μ is increasing, so is its conjugate. The set function satisfies $\mu(K \cup L) \leq \mu(K) \leq \mu(K) + \mu(L)$ if and only if $\mu^\blacklozenge(K \cup L) \geq \mu^\blacklozenge(K) + \mu^\blacklozenge(L)$. Using the Choquet conjugate, the echelon between disjoint subsets A_1 and A_2 can be written in the form

$$\mathbb{A}_\mu[[A_1, A_2]](K) = \frac{\mu^\blacklozenge(K \cup A_2) - \mu^\blacklozenge(\complement K \cup A_1)}{\mu^\blacklozenge(A_1 \cup A_2)}.$$

Remark 3.11. Let us consider a closed subset $S \subset X$ such that $0 \in S$ and associate with any K the erosion

$$A_1 := K \ominus S := \bigcap_{s \in S}(K - s) \ \text{ and } \ A_2 := \complement(K - S) := (\complement K) \ominus S$$

the complement of the dilation $K - S := \bigcup_{s \in S}(K - s)$. Then K naturally belongs to the dispersion gap $[[K \ominus S, (\complement K) - S]]$. Since $K \setminus (K \ominus S) = ((\complement K) - S) \cap K$, we infer that the echelon of K in this dispersion gap is equal to

$$\mathbb{A}(\mu_\downarrow, \mu_\uparrow)[[K \ominus S, (\complement K) - S]](K) = \frac{(\mu_\downarrow\big(((\complement K) - S) \cap K\big) - \mu_\uparrow\big((K - S) \cap \complement K\big))}{\mu_\uparrow\big((K - S) \cup ((\complement K) - S)\big)}.$$

If μ is additive, we infer that

$$\mathbb{A}(\mu)[[K \ominus S, (\complement K) - S]](K) \;=\; \frac{(2\mu(K) - (\mu(K \ominus S) - \mu(K - S)))}{\mu(K - S) - \mu(K \ominus S)}\,.$$

By replacing the S by hS with $h > 0$, and taking the lower and upper limits of $\mathbb{A}(\mu)[[K \ominus hS, (\complement K) - hS]]$, we could obtain concepts of lower and upper morphological derivatives of the measure μ at K in the "direction" S. □

Example 3.12. (Maslov and Kolmogorov echelons) The main examples are echelons associated with indicators:

(1) the *lower Maslov echelon* defined by

$$\mathbb{A}_{(\Phi^\flat,\Phi^\sharp)}[[A_1, A_2]](K) := \frac{\Phi^\flat(A_1 \cup \complement K) - \Phi^\sharp(A_2 \cup K)}{\Phi^\sharp((A_1 \cup A_2))} \in \left[-1, \frac{\Phi^\flat((A_1 \cup A_2))}{\Phi^\sharp(A_1 \cup A_2)}\right];$$

(2) the *Kolmogorov echelon* defined by

$$\mathbb{A}_{\Phi^\natural}[[A_1, A_2]](K) \;:=\; \frac{\Phi^\natural(A_1 \cup \complement K) - \Phi^\natural(A_2 \cup K)}{\Phi^\natural(A_1 \cup A_2)} \;\in\; [-1, +1]\,;$$

(3) the *upper Maslov echelon* defined by

$$\mathbb{A}_{(\Phi^\sharp,\Phi^\flat)}[[A_1, A_2]](K) := \frac{\Phi^\sharp(A_1 \cup \complement K) - \Phi^\flat(A_2 \cup K)}{\Phi^\flat(A_1 \cup A_2)} \in \left[-1, \frac{\Phi^\sharp(A_1 \cup A_2)}{\Phi^\flat(A_1 \cup A_2)}\right].$$

Example 3.13. (Another measure of difference quotients) Let us consider an indicator $x \in X \rightsquigarrow \Phi(x) \subset \mathbb{R}$, a direction $u \in X$ and a duration Δ. We introduce the difference-quotient

$$\nabla_\Delta \Phi(x)(u) \;:=\; \frac{\Phi(x) - \Phi(x - \Delta u)}{\Delta} \;=\; \left\{\frac{a-b}{\Delta}\right\}_{a \in \Phi(x),\, b \in \Phi(x - hu)}$$

of x in the direction u. Instead of taking Painlevé-Kuratowski upper or lower limits of these difference-quotients when $\Delta \to 0+$, we may take the measures

$$\frac{\inf \Phi(x) - \sup \Phi(x - \Delta u)}{\Delta} \;\leq\; \frac{\sup \Phi(x) - \inf \Phi(x - \Delta u)}{\Delta}$$

and their difference

$$\frac{(\sup \Phi(x) - \inf \Phi(x)) + (\sup \Phi(x - \Delta u) - \inf \Phi(x - \Delta u))}{\Delta}$$

This may open a new way to explore. □

4. Extremal clustering of time series of sets

4.1. Extremal analysis of time series of sets

We refer to Figure 1.1, p. 56, to explain how were constructed the extremal boxes in which are allocated the upper and lower magnitudes of the time series of subsets K_n in order to classify them in a relevant way. For that purpose, we shall distinguish four situations described by four scenarios

1. *upper* magnitudes $\mu_\uparrow(K_p)$ during *upstream* dispersion intervals $\{0, n-1\}$;
2. *upper* magnitudes $\mu_\uparrow(K_p)$ during *downstream* dispersion intervals $\{n, N\}$;
3. *lower* magnitudes $\mu_\downarrow(K_p)$ during *upstream* dispersion intervals $\{0, n-1\}$;
4. *lower* magnitudes $\mu_\downarrow(K_p)$ during *downstream* dispersion intervals $\{n, N\}$.

We define the upper and lower envelopes by the formulas

$$\left\{\begin{array}{l} \overrightarrow{\beta}^\sharp(n) \in \sup_{p\in\{0,n-1\}} \mu_\uparrow(K_p), \quad \overleftarrow{\beta}^\sharp(n) \ := \ \sup_{p\in\{n,N\}} \mu_\uparrow(K_p) \\ \text{and the } \textit{upper extremal envelope } \ \beta^\sharp(n) := \min(\overrightarrow{\beta}^\sharp(n), \overleftarrow{\beta}^\sharp(n)); \\ \overrightarrow{\beta}^\flat(n) \in \inf_{p\in\{0,n-1\}} \mu_\downarrow(K_p), \quad \overleftarrow{\beta}^\flat(n) := \inf_{p\in\{n,N\}} \mu_\downarrow(K_p) \\ \text{and the } \textit{lower extremal envelope } \ \beta^\flat(n) \ := \ \max(\overrightarrow{\beta}^\flat(n), \overleftarrow{\beta}^\flat(n)) \end{array}\right. \tag{1}$$

Definition 4.1. (Extremal envelopes and their extremal boxes) The graph of the extremal envelope $p \rightsquigarrow [\beta^\sharp(n), \beta^\flat(n)]$ delineated by the upper and lower envelopes is called the *extremal envelope* enclosing the values of the magnitudes and the differences $\beta^\sharp(n) - \beta^\flat(n)$ are called the *amplitudes* at each date.

1. We define
   (a) the *upper jumps* $\overrightarrow{\beta}^\sharp(n+1) - \overrightarrow{\beta}^\sharp(n)$ which are positive for upper extremals on upstream dispersion intervals and negative on lower extremal on downstream dispersion intervals;
   (b) the *lower jumps* $\overrightarrow{\beta}^\flat(n+1) - \overrightarrow{\beta}^\flat(n)$ which are negative for upper extremals on upstream dispersion intervals and positive on lower extremal on downstream dispersion intervals.
2. We extract the subset of *markers* $j \in \{1, J\}$, which are the dates n_j such that the jump $\overrightarrow{\beta}^\sharp(n_j + 1) - \overrightarrow{\beta}^\sharp(n_j) \neq 0$ are not equal to 0.
3. We call *extremiles* the values of the extremal upper and lower envelopes at markers;
4. and build the extremal boxes

$$\{n_j, n_{j+1}\} \times [\overrightarrow{\beta}^\sharp(n_j), \overrightarrow{\beta}^\sharp(n_j + 1)]$$

which encompasses the magnitudes between two consecutive marker dates and between the two corresponding extremiles.

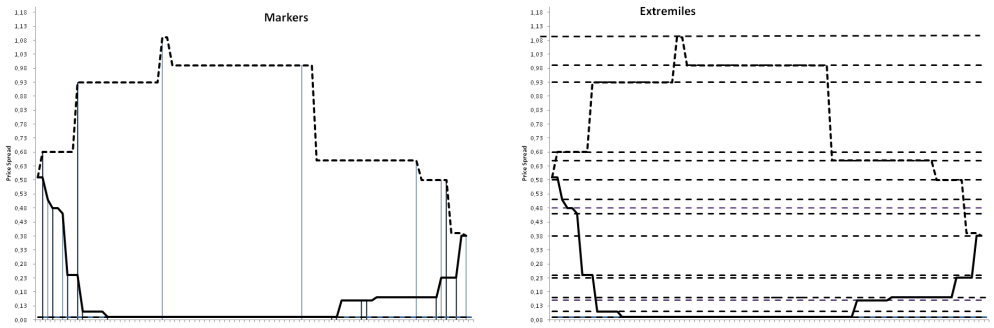


Figure 4.1: Markers and extremiles

Figure 4.1 (left) displays the graphs of the markers: the upstream maximal data and their magnitudes ones (thin vertical bars), the downstream ones (thick dotted horizonal), the upstream minimal data and their magnitudes ones (thick horizontal), the downstream ones (thick vertical bars).

Figure 4.1 (right) displays the extremiles which are described by the horizontal bars. They partition the magnitude interval between extremiles.

The extremal boxes are delineated by the intersection of markers vertical bars and extremile horizontal lines.

The extremal markers and their extremiles provide a classification or clustering of the data and their magnitudes.

4.2. Anchors of a time series

We wish to select the *local extrema*[29] of the upper and lower magnitudes which achieve the *extremiles*.

For this purpose, we need to proceed to detect and measure the local extrema.

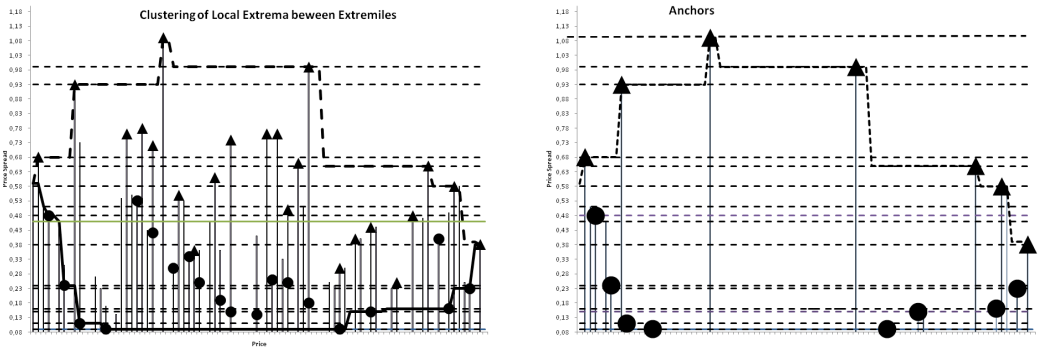


Figure 4.2: Local Extrema and anchors

In Figure 4.2 the local minima are displayed (with bullets) and the maximal ones (with triangles) are classified by the extremile partition from which are selected the

[29] This can be done by using the trendometer, adapted not only for single-valued magnitude maps (see [11]) and *Tychastic Measure of Viability Risk*, [12].

upper and *lower anchors*, the signal-magnitude pairs which are the local extrema of the upper and lower magnitudes achieving the upper and lower extremiles (in large bullets and triangles). They are the relevant signal/magnitude pairs summarizing the uppers and lower magnitudes of the observed sets.

In a nutshell, anchors are generated by the successive selection methods. First, by using the maximal and minimal selections of the magnitude, second, by allocating them in the clusters of the extremal grid defined by the markers and extremiles, and third, by selecting among them the local extrema of the maximal and minimal selections in this grid and isolating the anchors which are those local extrema lying at this extremal grid (and not the ones lying between the lines of the grid).

5. Clustering issues

We regard the two disjoint subsets A_1 and A_2 as *gists* around which subsets $K \in [[A_1, A_2]]$ could be compared. Indeed, knowing two capacities $(\mu_\downarrow, \mu_\uparrow)$, the associated echelon of K allows us to decide whether A_1 or A_2 is closer to K whenever K does not belong to the median of the dispersion gap $[[A_1, A_2]]$.

Namely, we associate with $K^1 := K$ the gists (A_1^2, A_2^2) of order 2 defined by

$$\begin{cases} A_1^2 = A_1 \cup K \text{ and } A_2^2 = A_2 & \text{if } \mathbb{A}_{(\mu_\downarrow,\mu_\uparrow)}[[A_1, A_2]](K) < 0, \\ A_1^2 = A_1 \text{ and } A_2^2 = A_2 \cup K & \text{if } \mathbb{A}_{(\mu_\downarrow,\mu_\uparrow)}[[A_1, A_2]](K) > 0, \end{cases}$$

and thus, a *second-order dispersion gap* $[[A_1^2, A_2^2]]$.

This is the first step of a *clustering process* associated with the pair $(\mu_\downarrow, \mu_\uparrow)$ of capacities. Indeed, we can assign to any new subset $K^2 \in [[A_1^2, A_2^2]]$ one of the two gists and cluster K to it, keeping the other one unchanged. This leads to a new pair of (A_1^3, A_2^3) of disjoint gists and a third smaller dispersion gap $[[A_1^3, A_2^3]]$ of order 3.

The clustering process stops at order n when $A_1^n = A_2^n$.

We turn our attention to the case of clusters of a finite number of subsets:

Definition 5.1. (Echelon of subsets of pairwise disjoint subsets) Let us denote by $\mathbb{L}$ a subset of ℓ indices and consider a finite family of *pairwise disjoint subsets* $\{A_k\}_{k\in\mathbb{L}}$ regarded as the *gists* of a sorting or clustering process. The *demarcation* between the pairwise disjoint subsets A_k defined by $\mho_{k\in\mathbb{L}} A_k := \complement\left(\bigcup_{k\in\mathbb{L}} A_k\right)$.
The *dispersion gap* $[[A_k]]_{k\in\mathbb{L}}$ is the family of subsets K such that no subset A_k satisfies both $A_k \cap K \neq \emptyset$ and $A_k \cap \complement K \neq \emptyset$. We associate with

1. any subset $K \in [[A_k]]_{k\in\mathbb{L}}$ the two subsets
   (a) $\mathbb{L}(K)$ of indices i such that $A_i \subset K$;
   (b) $\mathbb{L}(\complement K)$ of indices j such that $A_j \subset \complement K$.
2. two capacities $\mu_\downarrow \leq \mu_\uparrow$ the lower echelon of $K \in [[A_k]]_{k\in\mathbb{L}}$ with respect to the pairwise disjoint family:

$$\mathbb{A}_{(\mu_\downarrow,\mu_\uparrow)}\left[[A_k]\right]_{k\in\mathbb{L}}(K) := \frac{\mu_\downarrow\left(\left(\mho_{i\in\mathbb{L}(K)}A_i\right)\mho(\complement K)\right) - \mu_\uparrow\left(\left(\mho_{j\in\mathbb{L}(\complement K)}A_j\right)\mho K\right)}{\mu_\uparrow\left(\mho_{k\in\mathbb{L}}A_k\right)}$$

Intermediate echelon $\mathbb{A}^\flat_\mu\left[[A_k]\right]_{k\in\mathbb{L}}(K)$ and upper echelon $\mathbb{A}^\sharp_{(\mu_\uparrow,\mu_\downarrow)}\left[[A_k]\right]_{k\in\mathbb{L}}(K)$ are defined in the same way than in Definition 3.5, p. 65, for $\ell = 2$.

Remark 5.2. The subsets belonging to the *a priori* dispersion gap $\left[[A_k]\right]_{k\in\mathbb{L}}(K)$ actually belong to the *a posteriori* "bilateral" dispersion gap

$$\left[\left[\Big(\bigcup_{i\in\mathbb{L}(K)}A_i\Big),\ \Big(\bigcup_{j\in\mathbb{L}(\complement K)}A_j\Big)\right]\right],$$

since $A_1 := \bigcup_{i\in\mathbb{L}(K)}A_i$ and $A_2 := \bigcup_{j\in\in\mathbb{L}(\complement K)}A_j$ are disjoint, but depend on K. □

Let us consider ℓ *pairwise disjoint subsets* $\{A^1_k\}_{k\in\mathbb{L}^1}$ and a subset K^1 in the dispersion gap $[[A_{k^1}]]_{k\in\mathbb{L}}$. As in the case of two disjoint subsets, we can define a *clustering process* by associating with any subset $K^1 \in [[A^1_k]]_{k\in\mathbb{L}^1}$ of the first order dispersion gap the family $\{A^2_k\}_{k\in\mathbb{L}^2}$ of second order gists $(A^2_k)_{k\in\mathbb{L}^2}$ of order 2 defined by

$$\begin{cases} \text{if } \mathbb{A}_{(\mu_\downarrow,\mu_\uparrow)}\left[[A^1_k]\right]_{k\in\mathbb{L}}(K^1) < 0, \text{ then } \mathbb{L}^2 := \{k_2\}\cup\mathbb{L}^1(\complement K^1) \text{ and} \\ A^2_{k_2} = K^1\cup\bigcup_{i\in\mathbb{L}^\nVdash}(K^1)A^1_i \text{ and } A^2_j = A^1_j \text{ if } j\in\mathbb{L}^1(\complement K^1); \\ \text{if } \mathbb{A}_{(\mu_\downarrow,\mu_\uparrow)}\left[[A^1_k]\right]_{k\in\mathbb{L}}(K^1) > 0, \text{ then } \mathbb{L}^2 := \{k_2\}\cup\mathbb{L}^1(K^1) \text{ and} \\ A^2_i = A^1_i \text{if } i\in\mathbb{L}^1(K^1) \text{ and } A^2_{k_2} = K^1\cup\bigcup_{j\in\mathbb{L}^\nVdash(\complement K^1)}A^1_i, \end{cases}$$

and thus, a *second-order dispersion gap* $[[A^2_k]]_{k\in\mathbb{L}^2}$.

This is the first step of a *clustering process* by adding at each step a subset of the dispersion gap for building a new dispersion gap with fewer gists.

6. Echelons under Set Filtration

6.1. Set Filtrations

This section provides an example of construction of pair of subsets $A_{\gamma^\flat}(K) \subset A_{\gamma^\sharp}(K)$ associated with a *set calendar* $\mathcal{A}$, defined as an increasing sequence $n\in\mathbb{Z}\mapsto A_n\in\mathcal{A}$ of subsets A_n *pivoting* around the *pivot set* A_0 in the sense that

$$\begin{cases} A_{-\infty} := \bigcap_{n\in\mathbb{Z}}A_{-n} \subset \ldots A_{n+1}\subset A_n\ldots\subset A_1 \\ \subset A_0 \subset \\ A_{-1}\subset\ldots\subset A_{-n-1}\subset A_{-n}\subset\ldots A_{+\infty} := \bigcup_{n\in\mathbb{Z}}A_n \end{cases}$$

We associate with the set calendar its *slices*

$$\forall\, n \geq 1,\ \ S_n := A_n \setminus A_{n-1},\ \ S_{-n} := A_{-n+1} \setminus A_{-n} \ \text{ and } \ S_0 := \emptyset.$$

Then $S_1 \setminus S_{-1} = (A_1 \setminus A_0) \cup (A_0 \setminus A_{-1})$ is the *thick boundary* of the pivot set A_0 of the set calendar. The slices of the set calendar form a numerable partition, called a *stratification* of the space X:

$$X \ := \ A_{-\infty} \cup \bigcup_{n \in \mathbb{Z}} S_n \cup A_{+\infty}$$

The *spotting dates*[30] associated with a set calendar is the function $\gamma_{\mathcal{A}}$ are defined on the slices of the set calendar by

$$\gamma(x) \ := \ \gamma_{\mathcal{A}}(x) \ := \ n \text{ if and only if } x \ \in \ Sn$$

Let us associate with any subset K its *lower* and *upper spotting dates*

$$\gamma^\flat(K) \ := \ \inf_{K \subset A_n} n \text{ and } \gamma^\sharp(K) \ := \ \sup_{A_n \subset K} n$$

with which we associate the subsets $A_{\gamma^\flat}(K) \ \subset \ A_{\gamma^\sharp}(K)$ defining an dispersion gap $[[A_{\gamma^\flat}(K), A_{\gamma^\sharp}(K)]]$.

We then regard the difference $\delta(K) := \gamma^\sharp(K) - \gamma^\flat(K)$ as the *spotting duration of the temporal window* $[\gamma^\flat(K), \gamma^\sharp(K)]$ associated with the subset K relative to the set calendar $\mathcal{A}$.

Observing that if $K \setminus A_n = \emptyset$, then, $\forall\, m \geq n$, $K \setminus A_m = \emptyset$ and that if $A_p \setminus K = \emptyset$, then, $\forall\, q \leq p$, $A_q \setminus = \emptyset$, we infer that

$$\begin{cases} K \setminus A_{-\infty} \ = \ \bigcup_n K \setminus A_n \ = \ \bigcup_{n \geq \gamma^\flat(K)} K \setminus A_n \\ A_{+\infty} \setminus K \ = \ \bigcup_n A_n \setminus K \ = \ \bigcup_{n \leq \gamma^\sharp(K)} A_n \setminus K \end{cases}$$

Consequently, for any magnitude μ, we define the echelon of K of μ on the set calendar $\mathcal{A}$ by

$$\mathbb{A}_\mu[\mathcal{A}](K) \ := \ \mathbb{A}_\mu[A_{\gamma^\flat}(K), A_{\gamma^\sharp}(K)](K)$$

In this case, we identify the echelons of subsets K with the echelon $\mathbb{A}_\mu[A_n, A_m](K)$ having the same temporal window $[n = \gamma^\flat(K), m = \gamma^\sharp(K)]$.

6.2. Echelons under Galois measures

The next question which arises is: How can a set calendar be produced? An answer is provided by using any extensive hypermap from $\mathcal{P}(X)$ to $\mathcal{P}(X)$ and taking their powers for positive integers, and the powers of its reflection for negative integers. In other words, extensive hypermaps provide their own echelon measure. We first have to define these terms.

[30] This is justified by the fact that the index n of the set calendar is regarded as a discrete time.

Definition 6.1. (Hypermaps) Let $\mathcal{P}(X)$ denote the family of subsets of X and $\Lambda\colon \mathcal{P}(X) \rightsquigarrow \mathcal{P}(X)$ a map from $\mathcal{P}(X)$ to itself, called a *hypermap*. The *reflection* $\Lambda^\ddagger\colon \mathcal{P}(X) \rightsquigarrow \mathcal{P}(X)$ of $\Lambda\colon \mathcal{P}(X) \mapsto \mathcal{P}(X)$ is the hypermap defined by

$$\Lambda^\ddagger(K) := \complement\Lambda(\complement K)$$

A hypermap is said to be *extensive* if for all subsets $K \subset X$ we have $K \subset \Lambda(K)$, and *intensive* if $\Lambda(K) \subset K$. The extension of a hypermap Λ associates to any K the union $\Lambda(K) \cup K$ and intensification subset $\Lambda(K) \cap K$. A hypermap Λ is a *dilation* if it is a morphism for the union:

$$\Lambda\Big(\bigcup_{i\in\mathbb{I}} K_i\Big) := \bigcup_{i\in\mathbb{I}} \Lambda\left(K_i\right)$$

and an *erosion* if it stable for the intersection:

$$\Lambda\Big(\bigcap_{i\in\mathbb{I}} K_i\Big) := \bigcap_{i\in\mathbb{I}} \Lambda\left(K_i\right)$$

A union $\Lambda := \bigcup_{i\in\mathbb{I}} \Lambda_i$ of extensive dilations Λ_i is called a *shrub*. □

Observe that the *reflection*[31] $\Lambda \mapsto \Lambda^\ddagger$ is a bijection, the reflection of an extensive hypermap is an *intension*, the refection of an extension is the intensification of its reflection and the reflection of a dilation is an erosion.

Remark 6.2. (Echelons between disjoint extension of subsets) When the subsets $A_1 := \Lambda^\ddagger(A_1)$ and $A_2 := \Lambda^\ddagger(A_2)$ are disjoint open subsets satisfying $A_1 \subset \complement A_2$, then, for any subset $K \in [[A_1, A_2]]$, $A_1 \subset \Lambda^\ddagger(K) \subset K \subset \overline{K} \subset \complement A_2$. We thus can define the *topological echelons*:

1. the *lower set echelon function*

$$\mathbb{A}^\flat_{(\mu_\downarrow,\mu_\uparrow)}[\Lambda^\ddagger(A_1), \Lambda^\ddagger(A_2)](K) := \frac{\mu_\downarrow(\Lambda^\ddagger(A_1))\mho\Lambda^\ddagger(\complement K) - \mu_\uparrow(\Lambda^\ddagger(A_2)\mho\Lambda^\ddagger(K))}{\mu_\uparrow(\Lambda^\ddagger(A_1)\ \mho\ \Lambda^\ddagger(A_2))},$$

2. the *central set echelon function*

$$\mathbb{A}_\mu[\Lambda^\ddagger(A_1), \Lambda^\ddagger(A_2)](K) := \frac{\mu(\Lambda^\ddagger(A_1)\mho\Lambda^\ddagger(\complement K)) - \mu(\Lambda^\ddagger(A_2)\mho\Lambda^\ddagger(K))}{\mu(\Lambda^\ddagger(A_1)\ \mho\ \Lambda^\ddagger(A_2))},$$

3. the *upper set echelon function*

$$\mathbb{A}^\sharp_{(\mu_\uparrow,\mu_\downarrow)}[\Lambda^\ddagger(A_1), \Lambda^\ddagger(A_2)](K) := \frac{\mu_\uparrow(\Lambda^\ddagger(A_1))\mho\Lambda^\ddagger(\complement K)\mu_\downarrow(\Lambda^\ddagger(A_2)\mho\Lambda^\ddagger(K))}{\mu_\downarrow(\Lambda^\ddagger(A_1)\ \mho\ \Lambda^\ddagger(A_2))}.$$

[31] It plays the same role than Choquet conjugates for capacities instead of hypermaps (see remark, p. 67).

Powers Λ^n of extensive hypermaps are more and more extensive since, for any $K \subset X$, $K \subset \Lambda(K) \subset \ldots \subset \Lambda^n(K) \subset \Lambda^{n+1}(K) \subset \ldots$. In the same way, the powers $\Lambda^{\ddagger^n}(K)$ of its intensive reflection $\Lambda^\ddagger$ are more and more intensive since $\ldots \subset \Lambda^{\ddagger^{n+1}}(K) \subset \Lambda^{\ddagger^n}(K) \ldots \Lambda^\ddagger(K) \subset K$.

This suggests to associate to the extension Λ the set calendar $\mathcal{A}$ defined by for any $n \geq 0$ the subsets $A_n := \Lambda^n(K)$, $A^0 = K$ and $A_{-n} := \Lambda^{\ddagger^n}(K)$. In this case, we associate with the extension Λ and with any subset K the *Galois calendar*

$$\begin{cases} \partial_\Lambda^{-\infty}(K) := \bigcap_{n\geq 0} \Lambda^{\ddagger^n}(K) \cdots \subset \Lambda^{\ddagger^{n+1}}(K) \subset \Lambda^{\ddagger^n}(K) \cdots \subset \Lambda^\ddagger(K) \\ \subset K \subset \Lambda(K) \subset \ldots \Lambda^n(K) \subset \Lambda^{n+1}(K) \subset \ldots \\ \subset \bigcup_{n\geq 0} \Lambda^n(K) := \partial\Lambda^{+\infty}(K) \end{cases}$$

We observe that $\Lambda(\partial_\Lambda^{+\infty}(K)) = \partial_\Lambda^{+\infty}(K)$ is a fixed set of Λ and that $\Lambda^\ddagger(\partial_\Lambda^{-\infty}(K)) = \partial_\Lambda^{-\infty}(K)$ is a fixed set of is reflection $\Lambda^\ddagger$. We can regard the computations of the powers $\Lambda^n(K)$ as an *ascending algorithm* for computing the fixed set of the extension Λ containing K whereas the powers $\Lambda^{\ddagger^n}(K)$ describe a *descending algorithm* for computing the fixed of the intension $\Lambda^{\ddagger^n}$ contained in K.

We refer to [14] for examples of extensive hypermaps, among them, our favorite[32] ones, the hypermaps providing the *capture* and *absorption basins* of a subset under a set-valued map through an ascending algorithm, their reflection leading furnishing the descending algorithms converging to their *viability* and *invariant kernels* respectively. Dilation and erosions lead also to the concepts of openings and closing in *mathematical morphology*[33].

The slices of the Galois set calendar are called *marches*:

1. n-th order external marches

$$S_n := \partial_\Lambda^n(K) := \Lambda^n(K) \setminus \Lambda^{n-1}(K) = \Lambda^n(K) \cap \Lambda^{\ddagger^{n-1}}(\complement K),$$

2. n-th order internal marches

$$S_{-n} := \partial_\Lambda^{-n}(K) = \Lambda^{\ddagger^n}(K) \setminus \Lambda^{\ddagger^{n+1}}(K) = \Lambda^{\ddagger^n}(K) \cap \Lambda^{n+1}(K)(\complement K).$$

We obtain the *Galois stratifications* of X generated by K:

$$X = \partial_\Lambda^{-\infty}(K) \cup \Big(\bigcup_{n\geq 1} \partial_{\Lambda^\ddagger}^{-n}(K)\Big) \cup \Big(\bigcup_{n\geq 1} \partial_\Lambda^n(K)\Big) \cup \partial_\Lambda^{+\infty}(K).$$

[32] See [10] and [14].

[33] See for instance [1], [2], [8], [40], among a well-stocked literature.

Let us consider the case of extensions which are *shrubs*[34] $\Lambda := \bigcup_{i\in\mathbb{I}} \Lambda_i$ when the Λ_i are dilations. Their powers satisfy

$$\Lambda^2(K) \ = \ \bigcup_{(i_1,i_2)\in\mathbb{I}^2} (\Lambda_{i_1} \circ \Lambda_{i_2})(K)$$

and, more generally, $\Lambda^n(K) \ = \ \bigcup_{(i_1,i_2,...,i_n)\in\mathbb{I}^n} (\Lambda_{i_1} \circ \Lambda_{i_2} \ldots \Lambda_{i_n})(K).$

It is a dilation: $\Lambda^n\Big(\bigcup_{j\in\mathbb{J}} K_j\Big) \ = \ \bigcup_{j\in\mathbb{J}} \bigcup_{(i_1,i_2,...,i_n)\in\mathbb{I}^n} (\Lambda_{i_1} \circ \Lambda_{i_2} \ldots \Lambda_{i_n})(K_j).$

If Λ is a dilation, we associate the following maps of two variables $L \subset X$ and $M \subset X$ defined by

$$\Phi^{[1]}_\Lambda(L, M) \ := \ L \cap \Lambda(M)$$

We observe that

$$\Phi^{[1]}_\Lambda\left(\bigcup_{i_1\in\mathbb{I}_1} K_{i_1}, \bigcup_{i_2\in\mathbb{I}_2} K_{i_2}\right) \ = \ \bigcup_{(i_1,i_2)\in\mathbb{I}_1\times\mathbb{I}_2} \Phi^{[1]}_\Lambda(K_{i_2}, K_{i_2})$$

and, when $\Lambda := \bigcup_{j\in\mathbb{J}} \Lambda_j$ is a shrub, we obtain

$$\Phi^{[1]}_\Lambda\left(\bigcup_{i_1\in\mathbb{I}_1} K_{i_1}, \bigcup_{i_2\in\mathbb{I}_2} K_{i_2}\right) = \bigcup_{j\in\mathbb{J}} \bigcup_{(i_1,i_2)\in\mathbb{I}_1\times\mathbb{I}_2} \Phi^{[1]}_{\Lambda_j}(K_{i_2}, K_{i_2}).$$

Algorithms generated by the power of the intensive maps $\Phi^{[1]}_\Lambda$ associated with shrubs are parallel or distributed, since, *at each iteration, the cells are unions of subcells independently computed at the preceding iteration.*

Remark 6.3. Let us consider an extension $\Lambda\colon \mathcal{P}(X) \mapsto \mathcal{P}(X)$. Then the echelon of the subsets $\Lambda^n(K)$ between A and B are equal to

$$\mathbb{A}_\mu[A, B](\Lambda^n(K)) \ := \ \frac{\mu(\Lambda^n(K) \setminus A) - \mu(\Lambda^{\ddagger^n}(K) \setminus B)}{\mu(A \,\mho\, B)}$$

We associate with A and B the dates

$$\begin{cases} \gamma^\flat(K) \ := \ \inf_{A\subset\Lambda^n(K)} n, \\ \gamma^\sharp(K) \ = \ \sup_{B\subset(\Lambda^{\ddagger^n}(K))} m. \end{cases}$$

Taking $B = \Lambda(B)$ is a fixed set of Λ, then $\Lambda^n(B) = B$ so that

$$\mathbb{A}_\mu[A, B](\Lambda^n(B)) \ := \ \frac{\mu((B \setminus A)) - \mu(\Lambda^{\ddagger^n}(B) \setminus B)}{\mu(A \,\mho\, B)}.$$

[34] The reflections of shrubs are used in the construction of Cantor sets and among them, fractals. See for instance Section 2.9.4, p. 79, of [10]. Shrubs can be used for building *agglomerative hierarchical clustering* algorithms for approximating $\partial^{+\infty}_\Lambda(K)$ by the using the powers $\Lambda^n(K)$. Their reflection provides *divisive hierarchical clusterings.*

References

[1] M. Akian, R. Bapat, S. Gaubert: *Max-plus algebras*, in: *Handbook of Linear Algebra*, Discrete Mathematics and Its Applications 39, chapter 25, L. Hogben (ed.), Chapman & Hall, Boca Raton (2006).

[2] M. Akian, B. David, S. Gaubert: *Un théorème de représentation des solutions de viscosité d'une équation d'Hamilton-Jacobi-Bellman ergodique dégénérée sur le tore*, C. R. Acad. Sci. Paris, Ser I 346 (2008) 1149–1154.

[3] M. Akian, S. Gaubert, V. Kolokoltsov: *Invertibility of functional Galois connections I*, C. R. Acad. Sci. Paris, Série I 335 (2002) 883–888.

[4] M. Akian, J.-P. Quadrat, M. Viot: *Duality between probability and optimization*, in: *Idempotency*, J. Gunarwadernar (ed.), Cambridge University Press, Cambridge (1997) 331–353.

[5] N. Aronszajn: *Theory of reproducing kernels*, Trans. Amer. Math. Soc. 68 (1950) 337–404.

[6] J.-P. Aubin: *Contingent derivatives of set-valued maps and existence of solutions to nonlinear inclusions and differential inclusions*, in: Advances in Mathematics 7a, Mathematical Analysis and Applications, L. Nachbin (ed.) (1981) 159–229.

[7] J.-P. Aubin: *Neural Networks and Qualitative Physics: a Viability Approach*, Cambridge University Press, Cambridge (1996).

[8] J.-P. Aubin: *Mutational and Morphological Analysis*, Birkhäuser, Basel (1999).

[9] J.-P. Aubin: *Time and Money*, Lecture Notes Econ. Math. Systems 670, Springer, Cham (2014).

[10] J.-P. Aubin, A. M. Bayen, P. Saint-Pierre: *Viability Theory. New Directions*, Springer, Berlin (2011).

[11] J.-P. Aubin, Chen Luxi, O. Dordan: *Retro-prospective differential inclusions and their control by the differential connection tensors of their evolutions: the trendometer*, Complex Systems 23(2) (2014) 117–148.

[12] J.-P. Aubin, Chen Luxi, O. Dordan: *Tychastic Measure of Viability Risk*, Springer, Cham (2014).

[13] J.-P. Aubin, A. Désilles: *Traffic Networks as Information Systems. A Viability Approach*, Mathematical Engineering, Springer, Berlin (2017).

[14] J.-P. Aubin, O. Dordan: *A survey on Galois stratifications and measures of viability risk*, J. Convex Analysis 23(1) (2016) 181–225.

[15] J.-P. Aubin, H. Frankowska: *Set-Valued Analysis*, Birkhäuser, Basel (1990).

[16] J.-P. Aubin, R. B. Vinter (eds.): *Convex Analysis and Optimization*, Proceedings of a Conference, Research Notes in Mathematics 57, Pitman, Boston (1982).

[17] G. W. Bassett, R. Koenker, G. Kordas: *Pessimistic portfolio allocation and Choquet expected utility*, J. Financial Econometrics 2(4) (2004) 477–492.

[18] L. Billard, E. Diday: *Symbolic Data Analysis: Conceptual Statistics and Data Mining*, Wiley, Hoboken (2007).

[19] N. Bourbaki: *Théorie des Ensembles*, Hermann, Paris (1957).

[20] H. Cartan: *Théorie du potentiel newtonien: énergie, capacitée, suites de potentiels*, Bull. Soc. Math. France 73 (1945) 74–106.

[21] N. Chagnon: *Noble Savages*, Simon & Schuster, New York (2013)

[22] G. Choquet: *Theory of capacities*, Annales de l'Institut Fourier 5 (1953) 31–295.

[23] R. Fisher: *The correlation between relatives on the supposition of Mendelian inheritance*, Trans. Royal Soc. Edinburgh 52 (1918) 399–433.

[24] M. Grabisch: *Capacities and games on lattices: a survey of result*, Int. J. Uncertainty, Fuzziness and Knowledge-Based Systems 4(4) (2006) 371–392.

[25] L. Jaulin, M. Kieffer, O. Didrit, E. Walter: *Applied Interval Analysis*, Springer, London (2001).

[26] E. V. Khmaladze: *Differentiation of sets in measure*, J. Math. Anal. Appl. 334 (2007) 1055–1072.

[27] E. V. Khmaladze, W. Weil: *Differentiation of sets: the general case*, J. Math. Anal. Appl. 413 (2014) 291–310.

[28] R. Koenker: *Quantile Regression*, Econometric Society Monograph Series, Cambridge University Press, Cambridge (2005).

[29] A. B. Kurzhanski, I. Valyi: *Ellipsoidal Calculus for Estimation and Control*, Birkhäuser, Basel (1994).

[30] T. Lorenz: *Mutational Analysis. A Joint Framework for Cauchy Problems in and Beyond Vector Spaces*, Lecture Notes Math. 1996, Springer, Berlin et al. (2010).

[31] V.-P. Maslov: *Méthodes Opératorielles*, Éditions MIR, Moscow (1987).

[32] V.-P. Maslov, S. N. Samborski: *Idempotent Analysis*, American Mathematical Society, Providence (1993).

[33] A. de Moivre: *The Doctrine of Chances*, see http://www.ime.usp.br/∼walterfm/cursos/mac5796/DoctrineOfChances.pdf (1718).

[34] P. de Montmort: *Essay d'analyze sur les jeux de hazard*, see http://gallica.bnf.fr/ark:/12148/bpt6k110519q/f1.image (1708).

[35] R. E. Moore: *Interval Analysis*, Prentice-Hall, Englewood Cliffs (1966).

[36] E. Pernot: *Choix d'un classifieur en discrimination*, Strasbourg (1994).

[37] T. Piketty: *Le Capital au* XXI*ième Siècle*, Édition de Seuil, Paris (2013).

[38] A. Quetelet: *Sur l'homme et le développement de ses facultés, ou Essai de physique sociale*, 2 volumes, see http://gallica.bnf.fr/ark:/12148/bpt6k81570d.pdf (1835).

[39] R. T. Rockafellar, R. J. B. Wets: *Variational Analysis*, Grundlehren der mathematischen Wissenschaften 317, Springer, Berlin et al. (1998)

[40] M. Schmitt, J. Mattioli: *Éléments de Morphologie Mathématique*, Masson, Paris (1994).

[41] S. Shi: *Choquet theorem and nonsmooth analysis*, J. Math. Pures Appl. 67 (1988) 441–432.

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

[43] F. Vasilescu: *Problema lui Dirichlet Generalizata* (1937) .