Theory of Cost Measures: Convergence of Decision Variables

Marianne Akian

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Theory of cost measures : convergence of decision variables

Marianne Akian

N 2611

Juillet 1995

PROGRAMME 5

Rapport de recherche

ISSN 0249-6399

Theory of cost measures : convergence of decision variables

Marianne Akian *

Programme 5 — Traitement du signal, automatique et productique
Projet Méta2

Rapport de recherche n˚ 2611 — Juillet 1995 — 39 pages

Abstract: Considering probability theory in which the semifield of positive real numbers is replaced by the idempotent semifield of real numbers (union infinity) endowed with the min and plus laws leads to a new formalism for optimization. Probability measures correspond to minimums of functions that we call cost measures, whereas random variables correspond to constraints on these optimization problems that we call decision variables. We review in this context basic notions of probability theory – random variables, convergence of random variables, characteristic functions, L^p norms. Whenever it is possible, results and definitions are stated in a general idempotent semiring.

Key-words: Max-plus algebra, Dioid, Idempotent semiring, Idempotent measure, Decision variable, Fenchel transform, Optimization, Probability.

(Résumé : tsvp)

*e-mail : `Marianne.Akian@inria.fr`

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
Téléphone : (33 1) 39 63 55 11 – Télécopie : (33 1) 39 63 53 30

Théorie des mesures de coût : sur la convergence des variables de décision

Résumé : Si l'on considére la théorie des probabilités dans laquelle le demi-corps des réels positifs est remplacé par le demi-corps idempotent des réels (union l'infini) munis des lois min et plus, on obtient un nouveau formalisme pour l'optimisation. Les mesures de probabilité deviennent des minimums de fonctions que nous appellerons mesures de coût tandis que les variables aléatoires correspondent à des contraintes sur ces problèmes d'optimisation que nous appellerons variables de décision. Nous considérons dans ce contexte les notions de base de la théorie des probabilités – variables aléatoires, convergence de variables aléatoires, fonctions caracteristiques, normes L^p – et dès que cela est possible, nous établissons les théorèmes et les définitions dans un demi-anneau général.

Mots-clé : Algebra max-plus, Dioïde, Demi-anneau idempotent, Mesure idempotente, Variable de decision, Transformée de Fenchel, Optimisation, Probabilité.

Introduction

A probability or a positive measure is in some loose sense a continuous morphism from a Boolean σ-algebra $(\mathcal{A}, \cup, \cap)$ of subsets of some set Ω , into the semifield $(\mathbb{R}^+, +, \times)$. The $+$ law is used in the definition of probabilities and the $\times$ law in the definition of independence for instance. The notions of random variable, expectation, convergence,... are then introduced. If we replace $(\mathbb{R}^+, +, \times)$ by an idempotent semiring (or dioid) $(\mathbb{D}, \oplus, \otimes)$ [7], we can introduce the same notions, and then ask if the same theorems occur. Although this construction appears as a game, it is useful. Indeed, considering the semiring $\mathbb{R}_{\min} = (\mathbb{R} \cup \{+\infty\}, \min, +)$, addition becomes minimization and measures or integrals become infinite infimum. This type of measure has been first introduced by Maslov in [19]. The simplest measures over $\mathbb{R}_{\min}$ are those with density,

$$\mathbb{K}(A) = \inf_{x \in A} c(x) \quad \forall A \in \mathcal{A} ,$$

and in [1] we have proven that in a various class of algebras, any idempotent measure has a density. Then, the "measure" of a set A corresponds to the minimum of a function under the constraint to be in A and random variables, convergence of random variables,... correspond to constraints on an optimization problem, convergence of sequences of optimization problems,... This analogy then leads to a "new" formalism for optimization theory.

More generally, any idempotent structure defines a partial order. Then, measure or probability theory over this structure still corresponds to an optimization theory, but this time for the corresponding partial order. Multicriteria optimization problems can be treated along these lines, as in Samborski and Tarashchan [25] or Kolokoltzov and Maslov [18].

Another mapping from probability to optimization is the Cramer transform introduced large deviations theory (see for instance Azencott, Guivarc'h and Gundy [6]). It provides a morphism between probability laws and convex cost functions [2], thus between Wiener processes (resp. linear second order elliptic equations) and Bellman processes (see [2] for the definition) (resp. particular Bellman equations). This morphism, constructed by using Laplace and Fenchel transforms, directly connects characteristic functions of classical probabilities with characteristic functions or Fenchel transform of probabilities over $\mathbb{R}_{\min}$. Thus, it can leads to other analogies such as the L^p norms that we introduce in section 7.

The approach of optimization via the analogy with measure theory has first been investigated by Maslov in [19], where he constructed idempotent measures and integrals. Other results have followed by Maslov and others (see Maslov and Samborski [21], Maslov and Kolokoltzov [20],...). By using a "probabilistic" instead of "measure theory" approach, Quadrat proved in [23] the law of large numbers and the central limit theorems for decision variables with independent cost laws. These results are related with deterministic optimal control problems. Other studies motivated by this last approach have been done by Viot and Bellalouna [8], Del Moral, Thuillet, Rigal and Salut [12, 11]. In particular, the results of section 5 concerning the relations between almost sure and cost convergence together with a weakened version of the weak convergence (corresponding to vague convergence in classical probability theory) are also presented in [8] or [11], but only in $(\max, +)$ algebra and locally compact state spaces. In a survey paper [2] we have already presented some of the results of the present paper (without proof) together with the notions of Bellman processes. Properties of the weak convergence are more extensively studied in [3]. Applications of this approach to optimal control and game theory can be found in Whittle [27], Bernhard [10] and [2].

In this paper, we review the basic notions of probability theory after replacing the semifield of positive real numbers by a general idempotent semiring $(\mathbb{D}, \oplus, \otimes)$. The results are proven in a general idempotent semiring whose characteristics are defined in section 1. In section 2 and 3, we recall the definitions of idempotent or cost measures and integrals [19, 1]. Then in section 4, we introduce the notion of decision variable and convergence notions equivalent to the probabilistic ones that we compare in section 5. The main difference is that cost (probability) convergence is stronger than almost sure convergence. This is due to the non continuity of idempotent measures along nonincreasing sequences first noticed by Maslov [19] on the one hand, and the idempotency property on the other hand. Other properties are identical to the classical case and use almost the same proofs. As commutativity is not necessary, this theory and the classical probability theory can be constructed for matrices, operators,... Indeed, a global theory may be formulated in a general (non commutative) semiring $(\mathbb{D}, \oplus, \otimes)$ ordered by a relation compatible with the algebraic structure [19], but this would complicate the presentation.

We conclude this presentation by the notions of characteristic functions and L^p norms. Although characteristic functions may be generalized to a large family of semirings, we present them in $\mathbb{R}_{\max} = (\mathbb{R} \cup \{-\infty\}, \max, +)$, together with results of convergence (section 6). Note that the semiring $\mathbb{R}_{\max}$ has been preferred to $\mathbb{R}_{\min}$, to which it is isomorphic, for its corresponding order relation is the classical order relation $\leq$ of $\mathbb{R}$. However, L^p norms are introduced by analogy with those of random variables by means of the Cramer transform, and thus may not be defined for more general semirings than $\mathbb{R}_{\min}$ up to an isomorphism. We present them in $\mathbb{R}_{\max}$ (in section 7) together with relations between L^p convergence and other convergence notions. They lead to an easy proof of the law of large numbers of Quadrat and also to compactness criteria as in classical probabilities [3].

Let us note that although the convergence notions are introduced here following the analogy with probability theory, they may correspond to notions already introduced in optimization theory. For instance, the weak or in law convergence in $\mathbb{R}_{\max}$ is similar but not exactly equivalent [3] to the epigraph convergence introduced in convex analysis (see for instance Attouch [4], Attouch and Wets [5], Joly [16]). Finally, the main application of these convergence notions is relative to deterministic optimal control problems. For instance, the law of a Bellman chain X_n (the equivalent of a Markov chain) over $\mathbb{R}_{\min}$ is the value function v_n of some deterministic optimal control problem with finite horizon n. The weak convergence of X_n towards a decision variable X of law v is thus equivalent to the weak convergence of the function v_n towards v. This weak convergence is equivalent to weak convergence in Banach spaces but for the $(\min, +)$ scalar product and it is useful to study solutions of Hamilton-Jacobi equations (see Maslov and Samborski [22]). In addition, the L^p convergence of the Bellman chain X_n towards a constant decision variable x is equivalent to some coercity property of the function v_n around point x.

1 General assumptions on the dioid $(\mathbb{D}, \oplus, \otimes)$

Let us consider an idempotent semiring $(\mathbb{D}, \oplus, \otimes)$, that is a semiring with the idempotency property : $a \oplus a = a$ and consequently without the existence of opposites for the $\oplus$ law. The product law may be non commutative. The neutral elements (for the $\oplus$ and $\otimes$ laws respectively) are denoted by $\mathbb{0}$ and $\mathbb{1}$. We denote by $\preceq$ the partial order relation induced by the $\oplus$ law : $a \preceq b \quad \Leftrightarrow \quad a \oplus b = b$. Then $(\mathbb{D}, \preceq)$ is sup-semilattice and $a \oplus b$ is the supremum of a and b. We suppose that $(\mathbb{D}, \preceq)$ is a lattice such that any upper bounded set A has a least upper bound or supremum denoted $\oplus A$ or equivalently any nonempty set A has a greatest lower bound or infimum denoted $\wedge A$. Such a lattice will be called a locally complete lattice [1]. This property is equivalent to the existence of a complete lattice $\overline{\mathbb{D}}$ (that is a lattice such that any set admits a supremum and an infimum [15]) with top element (supremum) $\top$, such that $\mathbb{D}$ is a sublattice of $\overline{\mathbb{D}}$, $\overline{\mathbb{D}} = \mathbb{D} \cup \{\top\}$ and $\oplus \mathbb{D} = \top$. Then, the supremum $\oplus \mathbb{D}$ of $\mathbb{D}$ plays the role of $+\infty$ in $\mathbb{R}$. Note that, if $\top = \oplus \mathbb{D}$ does not belong to $\mathbb{D}$, the law $\otimes$ may be extended to $\overline{\mathbb{D}}$ so that $(\overline{\mathbb{D}}, \oplus, \otimes)$ becomes a semiring ($\top \otimes a = a \otimes \top = \top$ if $a \neq \mathbb{0}$ and $\top \otimes \mathbb{0} = \mathbb{0} \otimes \top = \mathbb{0}$). We suppose that the infinite distributivity of the $\otimes$ law with respect to the $\oplus$ law holds in $\mathbb{D}$ (but not necessarily in $\overline{\mathbb{D}}$) : $a \otimes (\oplus_{i \in I} b_i) = \oplus_{i \in I} (a \otimes b_i)$ for any set I such that $\oplus_{i \in I} b_i \in \mathbb{D}$ (together with the symmetrical relation in the non commutative case).

Example 1. $\mathbb{R}_{\max} = (\mathbb{R} \cup \{-\infty\}, \max, +)$ is an idempotent semifield (semiring with the existence of inverses for the $\otimes$ law) with neutral elements $\mathbb{0} = -\infty$ and $\mathbb{1} = 0$, associated order relation $\leq$ and supremum $\top = +\infty$. It is isomorphic to $\mathbb{R}_{\min} = (\mathbb{R} \cup \{+\infty\}, \min, +)$ (resp. $(\mathbb{R}^+, \max, \times)$) by the morphism $x \mapsto -x$ (resp. $x \mapsto e^x$). $\mathbb{R}_{\max}$ and $\mathbb{R}_{\min}$ are of particular interest in optimization (maximization and minimization) theory. Although vocabulary is suggested by minimization theory, the semiring $\mathbb{R}_{\max}$ is more practical since the order relation $\preceq$ coincides with the classical order relation of $\mathbb{R}$.

Example 2. $(\overline{\mathbb{R}} = [-\infty, +\infty], \max, \min)$ is an idempotent semiring with neutral elements $\mathbb{0} = -\infty$ and $\mathbb{1} = +\infty$, order relation $\leq$ and supremum $\top = +\infty \in \overline{\mathbb{R}}$. It is isomorphic to the semiring $([0, 1], \max, \min)$ with neutral elements $\mathbb{0} = 0$ and $\mathbb{1} = 1$ by the morphism $x \mapsto \phi(x)$ where ϕ is any bijective strictly increasing function from $\overline{\mathbb{R}}$ to $[0, 1]$ (ex : $\phi = (1 + \tanh)/2$).

Given an idempotent $\mathbb{D}$-measure (see section 2), an integral may be constructed if $\mathbb{D}$ is a metric space with a metric compatible with the idempotent semiring structure of $\mathbb{D}$ [19] and more generally, if $\mathbb{D}$ is a locally continuous lattice [1] (the notion of locally continuous lattice is an extension of that of continuous lattice [15] to locally complete lattices). However, the definition of cost (i.e. probability) convergence and the proof of some classical results on integrals require a distance on $\mathbb{D}$ compatible with the idempotent semiring structure of $\mathbb{D}$. In particular the topologies defined by δ and $\preceq$ (see below) have to be equivalent and the $\oplus$, $\wedge$ and $\otimes$ operations have to be continuous. We indeed impose to δ conditions that are similar to the properties of distances induced by (classical) algebra norms with $+$ replaced by $\max$. This led us to use the term "idempotent algebra distance" for δ . These conditions are equivalent to those of [19] and [20] for the $\oplus$ law and stronger for the $\otimes$ law. They imply the local continuity of $\mathbb{D}$ and the dual continuity of $[\mathbb{0}, \mathbb{1}]$ (see below).

Definition 3. We say that δ is an *idempotent algebra distance* on $(\mathbb{D}, \oplus, \otimes)$ if δ is a distance and if it satisfies :

1. $\delta(x \oplus x', y \oplus y') \leq \max(\delta(x, y), \delta(x', y'))$ for any x, x', y, y' in $\mathbb{D}$ (this property corresponds to $||x + x'|| \leq ||x|| + ||x'||$).

2. $\delta(\lambda\otimes x,\lambda\otimes y)\leq\delta_0(\lambda)\delta(x,y)$ and $\delta(x\otimes\lambda,y\otimes\lambda)\leq\delta_0(\lambda)\delta(x,y)$ for any x,y,λ in $\mathbb{D}$, where δ_0 is a morphism of monoids from $(\mathbb{D},\oplus)$ to $(\mathbb{R}^+,\max)$ such that $\delta_0(x\otimes y)\leq\delta_0(x)\delta_0(y)$ and $\delta_0(\mathbb{1})=1$ (this corresponds to the property of algebra norms : $\|xy\|\leq\|x\|\,\|y\|$). δ_0 is supposed to be bounded on bounded sets of $(\mathbb{D},\delta)$ (but not necessarily continuous).

3. For any $a\preceq b\preceq c$ and x in $\mathbb{D}$, we have $\delta(b,x)\leq\max(\delta(a,x),\delta(c,x))$.

4. $\delta(x\wedge x',y\wedge y')\leq\max(\delta(x,y),\delta(x',y'))$ for any x,x',y,y' in $\mathbb{D}$ (dual version of property 1).

5. δ is continuous for the lattice structure of $\mathbb{D}$: if x_i and y_i are $\mathbb{D}$-valued nets[1] such that $\lim_i x_i=x$ (that is $\liminf_i x_i\stackrel{\text{def}}{=}\oplus_i\wedge_{i\leq j}x_j=\limsup_i x_i\stackrel{\text{def}}{=}\wedge_i\oplus_{i\leq j}x_j=x$) and $\lim_i y_i=y$, then $\lim_i\delta(x_i,y_i)=\delta(x,y)$.

Using property 3, property 5 is equivalent to : x_i monotone and $\lim_i x_i=x$ implies $\lim_i\delta(x_i,x)=0$.

Proposition 4. *If δ is an idempotent algebra distance on $\mathbb{D}$, then for any (even non countable) set I and any subsets $\{a_i,i\in I\}$ and $\{b_i,i\in I\}$ of $\mathbb{D}$ such that $\oplus_{i\in I}a_i$ and $\oplus_{i\in I}b_i\in\mathbb{D}$, we have*

$$\delta(\bigoplus_{i\in I}a_i,\bigoplus_{i\in I}b_i)\leq\sup_{i\in I}\delta(a_i,b_i). \tag{1}$$

The same property holds for the $\wedge$ law instead of $\oplus$. The topology induced by δ coincides with the convergence notion associated with the lattice structure of $\mathbb{D}$, that is $x=\lim_i x_i$ in $\mathbb{D}$ if and only if $\lim_i\delta(x_i,x)=0$. It is also compatible with the semiring structure of $\mathbb{D}$, i.e. the functions $(x,y)\mapsto x\oplus y$ and $(x,y)\mapsto x\otimes y$ from $\mathbb{D}\times\mathbb{D}$ to $\mathbb{D}$ are continuous.

Proof. The compatibility with the lattice structure and inequality (1) are consequences of the continuity property 5 and properties 1 or 4 of an idempotent algebra distance. Let $\mathcal{P}$ denotes the set of finite subsets of I and $\alpha_p=\oplus_{i\in p}a_i$, $\beta_p=\oplus_{i\in p}b_i$ for $p\in\mathcal{P}$. Then $\mathcal{P}$ is a directed set (by the $\subset$ order relation) and α_p and β_p are nondecreasing with respect to p. By properties 5 and 1, we have

$$\delta(\bigoplus_{i\in I}a_i,\bigoplus_{i\in I}b_i)=\delta(\lim_p\alpha_p,\lim_p\beta_p)=\lim_p\delta(\alpha_p,\beta_p)\leq\sup_p\delta(\alpha_p,\beta_p)\leq\sup_{i\in I}\delta(a_i,b_i).$$

For the compatibility property, we only have to prove that $\lim_i\delta(x_i,x)=0$ implies $\lim_i x_i=x$. From $\limsup_i x_i=\lim_i\sup_{i\leq j}x_j$, we obtain

$$\delta(\limsup_i x_i,x)=\lim_i\delta(\sup_{i\leq j}x_j,x)\leq\lim_i\sup_{i\leq j}\delta(x_j,x)=\limsup_i\delta(x_i,x)=0.$$

Then, $\limsup_i x_i=x$. Replacing $\oplus$ by $\wedge$ and using property 4 instead of 1, we obtain inequality (1) for $\wedge$ law and $\lim_i\delta(x_i,x)=0$ implies $\liminf_i x_i=x$.

Property 1 implies that $(x,y)\mapsto x\oplus y$ is Lipschitz continuous. From property 2, we obtain

$$\delta(x\otimes y,x'\otimes y')\leq\max(\delta_0(x)\delta(y,y'),\delta_0(y')\delta(x,x')),$$

then $(x,y)\mapsto x\otimes y$ is Lipschitz continuous in every bounded set of $\mathbb{D}$, thus is continuous in $\mathbb{D}$. ☐

[1] A net is an application $i\in I\mapsto x_i\in\mathbb{D}$ where I is a directed set, that is a set endowed with an order relation such that any finite set as an upper bound.

Example 5. The exponential distance [19] $\delta(x,y) = |e^x - e^y|$ is an idempotent algebra distance on $\mathbb{R}_{\max}$ with $\delta_0(x) = e^x = \delta(x, \mathbb{0})$. In this example, the supremum $\top = +\infty \notin \mathbb{R}_{\max}$ and $\delta(\mathbb{0}, \top) = +\infty$, thus $\mathbb{R}_{\max}$ is not bounded.

For $\mathbb{D} = (\overline{\mathbb{R}}, \max, \min)$, $\delta(x,y) = |\phi(x) - \phi(y)|$ is an idempotent algebra distance (with $\delta_0(x) = 0$ if $x = \mathbb{0}$ and $\delta_0(x) = 1$ otherwise) if ϕ is a bijective strictly increasing function from $\overline{\mathbb{R}}$ into some compact interval of $\mathbb{R}$ (ex : $\phi = (1 + \tanh)/2$). In this case $\overline{\mathbb{D}} = \mathbb{D}$ then $\mathbb{D}$ is bounded.

Remark 6. In a normed vector space, property 3 is a consequence of properties 1 and 2 : if $b \in [a, c]$ i.e. $b = ta + (1-t)c$ with $t \in [0,1]$, then $||b|| \leq t||a|| + (1-t)||c|| \leq \max(||a||, ||b||)$. In an idempotent semiring $\mathbb{D}$, if $b = a \oplus t \otimes c$ with $t \preceq \mathbb{1}$, we obtain also from the first properties $\delta(b, x) = \delta(a \oplus t \otimes c, x \oplus t \otimes x) \leq \max(\delta(a, x), \delta_0(t)\delta(c, x)) \leq \max(\delta(a, x), \delta(c, x))$. Indeed, since δ_0 is a morphism from $(\mathbb{D}, \oplus)$ to $(\mathbb{R}^+, \max)$, δ_0 is nondecreasing and $0 \leq \delta_0(t) \leq \delta_0(\mathbb{1}) = 1$. Now, if $\mathbb{D}$ is an idempotent semifield, $a \preceq b \preceq c$ implies $b = a \oplus t \otimes c$ with $t = b \otimes c^{-1} \preceq \mathbb{1}$. Thus, in a semifield (likewise in a vector space), for instance in $\mathbb{R}_{\max}$ or $\mathbb{R}_{\min}$, property 3 is a consequence of the structure properties.

Remark 7. Let $(\mathbb{D}, \oplus, \otimes)$ be a semiring endowed with an idempotent algebra distance δ. The set $\mathbb{D}^n$, endowed with the laws $\oplus$ and $\otimes$ acting coordinate by coordinate, is an idempotent semiring with supremum $\oplus \mathbb{D}^n = (\oplus \mathbb{D}, \ldots, \oplus \mathbb{D})$. Moreover, the infinity distance $\delta_\infty(x, y) = \max_{i=1,\ldots,n} \delta(x_i, y_i)$ defines an idempotent algebra distance on it (with $(\delta_\infty)_0(x) = \max_{i=1,\ldots,n} \delta_0(x_i)$). We can also consider the non commutative idempotent semiring $\mathcal{M}_n(\mathbb{D})$ of $n \times n$ matrices with coefficients in $\mathbb{D}$, corresponding matrix $\oplus$ and $\otimes$ laws and supremum $\oplus \mathcal{M}_n(\mathbb{D}) = (\oplus \mathbb{D})_{i,j=1\ldots n}$, for which the infinity distance is also an idempotent algebra distance. The space $\mathbb{D}[X_1, \ldots, X_n]$ of polynomials with n indeterminates and coefficients in $\mathbb{D}$ is also an idempotent semiring, for which the infinity distance is also an idempotent algebra distance.

The infinite dimensional space $\mathcal{B}(X, \mathbb{D})$ of bounded functions from some (topological) space X to $\mathbb{D}$, endowed with the laws $\oplus$ and $\otimes$ acting coordinate by coordinate, is an idempotent semiring with supremum $\oplus \mathcal{B}(X, \mathbb{D}) = (\oplus \mathbb{D})^X$. However, the infinity distance $\delta_\infty(f, g) = \sup_{x \in X} \delta(f(x), g(x))$ satisfies (with $(\delta_\infty)_0(f) = \sup_{x \in X} \delta_0(f(x))$) all properties of an idempotent algebra distance (together with property (1) of Proposition 4) except the continuity property 5. Indeed, the topology associated to the lattice structure of $\mathcal{B}(X, \mathbb{D})$ is the pointwise convergence topology, whereas the topology induced by δ_∞ is the uniform convergence topology.

Remark 8. When the top element $\top$ of $\overline{\mathbb{D}}$ does not belong to $\mathbb{D}$, we have seen how to extend the $\otimes$ law such that $(\overline{\mathbb{D}}, \oplus, \otimes)$ becomes an idempotent semiring. Let us extend now the distance function δ to $\overline{\mathbb{D}}$ by : $\delta(x, \top) = \delta(\top, x) = +\infty$ if $x \in \mathbb{D}$, $\delta(\top, \top) = 0$ and $\delta_0(\top) = +\infty$. Then, properties 1-4 of an idempotent algebra distance are valid in $\overline{\mathbb{D}}$ and δ_0 becomes a morphism between the complete lattices (and monoids) $(\overline{\mathbb{D}}, \oplus)$ and $([0, +\infty], \max)$.

Definition 9. We say that δ is a *complete idempotent algebra distance* if δ is an idempotent algebra distance such that its extension to $\overline{\mathbb{D}}$ satisfies the following continuity property : if $\lim_i x_i = x \in \overline{\mathbb{D}}$ and $\lim_i y_i = y \in \mathbb{D}$ then $\lim_i \delta(x_i, y_i) = \delta(x, y)$.

An idempotent algebra distance is indeed complete iff $\lim_i \delta(x_i, y) = +\infty$ for any nondecreasing net x_i such that $\lim_i x_i = \top$ and any $y \in \mathbb{D}$. Note that the continuity property does not hold if $x = y = \top$ since the two nets x_i and y_i may converge with different rates. For instance in $\mathbb{R}_{\max}$, the two sequences $x_i = i$ and $y_i = i+1$ converge towards $\top = +\infty$, but for the exponential distance we have $\delta(x_i, y_i) = e^i(e-1) \longrightarrow_{i\to+\infty} +\infty \neq \delta(\top,\top)$.

Example 10. The distances defined in Example 5 are complete idempotent algebra distances. If δ is a complete idempotent algebra distance in $\mathbb{D}$, then the distances δ_∞ defined in Remark 7 are complete idempotent algebra distances on $\mathbb{D}^n$, $\mathcal{M}_n(\mathbb{D})$ and $\mathbb{D}[X_1,\dots,X_n]$.

Proposition 11. *If δ is a complete idempotent algebra distance on $\mathbb{D}$, then property (1) of Proposition 4 is valid for the $\oplus$ and $\wedge$ laws and for any a_i and b_i in $\overline{\mathbb{D}}$. As a consequence, the following "numbers" are elements of $\mathbb{D}$*

$$d_k = \bigoplus_{\delta(x,\mathbb{0})\le k} x \in \mathbb{D}, \tag{2}$$

and $\delta(d_k,\mathbb{0}) \le k$.

Proof. The generalization of property (1) is evident for the $\wedge$ law. For the $\oplus$ law, the proof of Proposition 4 can be applied. For (2), $d_k \in \overline{\mathbb{D}}$ exists ($\overline{\mathbb{D}}$ is complete) and $\delta(d_k,\mathbb{0}) \le \sup_{x,\delta(x,\mathbb{0})\le k} \delta(x,\mathbb{0}) \le k < +\infty$, thus $d_k \in \mathbb{D}$ ($d_k \notin \mathbb{D}$ would imply $\delta(d_k,\mathbb{0}) = +\infty$). $\square$

In order to use either the results of [19] and [20] or those of [1] on idempotent integrals, we relate the existence of an idempotent algebra distance with the continuity and dual continuity properties of the lattice $\mathbb{D}$. Recall that a subset D of $\mathbb{D}$ is said directed (resp. filtered) if any finite subset of D has an upper bound (resp. a lower bound) in $\mathbb{D}$.

Proposition 12. *Suppose that there exists an idempotent algebra distance δ on $\mathbb{D}$. The following properties hold :*

- *Let $\{D(j),\ j\in J\}$ be a family of upper bounded directed sets of $\mathbb{D}$. Let M be the set of all functions $f : J \to \cup_{j\in J} D(j)$ with $f(j) \in D(j)$ for all $j \in J$. Then the following identity holds :*

$$\bigwedge_{j\in J} \oplus D(j) = \bigoplus_{f\in M} \bigwedge_{j\in J} f(j). \tag{3}$$

- *Let $B \in \mathbb{D}$ and $\{F(j),\ j\in J\}$ be a family of filtered sets of the complete sublattice $[\mathbb{0}, B]$ of $\mathbb{D}$. Let M be the set of all functions $f : J \to \cup_{j\in J} F(j)$ with $f(j) \in F(j)$ for all $j \in J$. Then the following identity holds :*

$$\bigoplus_{j\in J} \wedge F(j) = \bigwedge_{f\in M} \bigoplus_{j\in J} f(j). \tag{4}$$

Proof. From the assumptions, all the numbers appearing in (3) are elements of $\mathbb{D}$. Moreover, the inequality $\wedge_{j\in J} \oplus D(j) \succeq \oplus_{f\in M} \wedge_{j\in J} f(j)$ is always true. The set M is ordered by the pointwise $\preceq$ relation and for this relation it is a directed set and $f \in M \mapsto \wedge_{j\in J} f(j)$ is nondecreasing. Then $\oplus_{f\in M} \wedge_{j\in J} f(j) = \lim_{f\in M} \wedge_{j\in J} f(j)$ and

$$\delta(\bigwedge_{j\in J} \oplus D(j), \bigoplus_{f\in M} \bigwedge_{j\in J} f(j)) = \lim_{f\in M} \delta(\bigwedge_{j\in J} \oplus D(j), \bigwedge_{j\in J} f(j)) \le \lim_{f\in M} \sup_{j\in J} \delta(\oplus D(j), f(j)).$$

Since $D(j)$ is a directed set and $x \in D(j) \mapsto \delta(\oplus D(j), x)$ is nonincreasing (by property 3 of δ),

$$\inf_{x \in D(j)} \delta(\oplus D(j), x) = \lim_{x \in D(j)} \delta(\oplus D(j), x) = \delta(\oplus D(j), \lim_{x \in D(j)} x) = 0.$$

Moreover, since $(\mathbb{R}, \leq)$ satisfies property (3) and the sets $\{\delta(\oplus D(j), x), x \in D(j)\}$ are directed for any $j \in J$, we obtain

$$\lim_{f \in M} \sup_{j \in J} \delta(\oplus D(j), f(j)) = \inf_{f \in M} \sup_{j \in J} \delta(\oplus D(j), f(j)) = \sup_{j \in J} \inf_{x \in D(j)} \delta(\oplus D(j), x) = 0$$

which implies (3). The identity (4) is obtained by replacing $\wedge$ by $\oplus$ in previous reasoning. $\square$

Property (3) corresponds to the local continuity of $\mathbb{D}$ and property (4) to the dual continuity of any sublattice $[\mathbb{0}, B]$ of $\mathbb{D}$ (see [15, 1]). In particular, let us define as in [15, 1] the "way below" $\ll$ relation on $\mathbb{D}$: $a \ll b$ if and only if for all upper bounded directed sets D of $\mathbb{D}$, such that $b \preceq \oplus D$, there exists $x \in D$ such that $a \preceq x$. Then, by definition of the local continuity, we have

$$x = \oplus\{y \in \mathbb{D},\ y \ll x\} \quad \text{for any} \quad x \in \mathbb{D}. \tag{5}$$

The "way below" relation corresponds to $<$ in $\mathbb{R}_{\max}$, the pointwise $<$ relation in $(\mathbb{R}_{\max})^n$ and in general it plays the same role as $<$ relation.

2 Decision spaces

In the following (section 2 to 5), we suppose given an idempotent semiring $(\mathbb{D}, \oplus, \otimes)$ endowed with an idempotent algebra distance δ .

Idempotent probabilities. Let U be a set and $\mathcal{U}$ a Boolean semi-σ-algebra of subsets of U, that is $\mathcal{U}$ contains $\emptyset$ and U and is stable by any countable union and finite intersection operation (ex : the set of open sets or the Borel sets σ-algebra of a topological space U). An *idempotent* $\mathbb{D}$*-measure* $\mathbb{K}$ on $(U, \mathcal{U})$ is by definition a map from $\mathcal{U}$ to $\overline{\mathbb{D}}$ such that $\mathbb{K}(\emptyset) = \mathbb{0}$, $\mathbb{K}(A \cup B) = \mathbb{K}(A) \oplus \mathbb{K}(B)$ and $\mathbb{K}(A_n) \nearrow_{n \to +\infty} \mathbb{K}(A)$ if $A_n \nearrow_{n \to +\infty} A$ with $A_n, A, B \in \mathcal{U}$. $\mathbb{K}$ is said *finite* if $\mathbb{K}(U) \in \mathbb{D}$. It is called an *idempotent probability* or a *cost measure* if $\mathbb{K}(U) = \mathbb{1}$ and in this case $(U, \mathcal{U}, \mathbb{K})$ is called a decision space. We consider only finite idempotent measures so that $[\mathbb{0}, \mathbb{K}(U)]$ is a dually continuous lattice and we denote by $\mathbb{K}^*$ the maximal extension of $\mathbb{K}$ to the set of all subsets of U : $\mathbb{K}^*(A) = \wedge_{B \in \mathcal{U},\ B \supset A} \mathbb{K}(B)$.

If $\mathbb{K}$ has a *density* i.e. if $\mathbb{K}(A) = \oplus_{u \in A}\, c(u)$ for any $A \in \mathcal{U}$ with $c : U \to \overline{\mathbb{D}}$ (a density always exists if $\mathcal{U}$ admits a countable basis, for instance if U is a separable metric space and $\mathcal{U}$ is the set of open sets [17, 20, 1]), then $c^*(u) = \mathbb{K}^*(\{u\})$ is the maximal density of $\mathbb{K}$ and, if $\mathcal{U}$ is a topology, the upper semicontinuous (u.s.c.) envelope of c (in $\mathbb{D}$). In this case we denote by $\mathrm{supp}(\mathbb{K})$ the domain of c^* i.e. $\{u \in U, c^*(u) \neq \mathbb{0}\}$; it is the complementary of the maximal "negligible" set. Let us note that in practice $\mathcal{U}$ will be a topology so that only u.s.c. (that is lower semi continuous (l.s.c.) if the dioid is $\mathbb{R}_{\min}$) densities will be considered.

Note also that an idempotent measure or probability is in general not continuous over nonincreasing sequences. However, this continuity holds in particular cases. We suppose given now a topological space U and denote by $\mathcal{U}$ the set of its open sets, $\mathcal{F}$ the set of its closed sets and $\mathcal{K}$ the set of its compact sets.

Proposition 13. *For any finite idempotent measure $\mathbb{K}$ on $(U,\mathcal{U})$, the following property holds for its maximal extension :*

$$\lim_n \mathbb{K}^*(C_n) = \mathbb{K}^*(C) \quad \forall C_n,\ C \in \mathcal{K},\ C_n \searrow_n C.$$

Proof. If C_n is a nonincreasing sequence converging to C then $\mathbb{K}^*(C_n)$ is nonincreasing and greater than $\mathbb{K}^*(C)$. If G is an open set containing the compact set C, then the sequence of compact sets $C_n \setminus G$ is nonincreasing and converges towards $\emptyset$. Therefore, $C_n \setminus G = \emptyset$ and $\mathbb{K}^*(C_n) \preceq \mathbb{K}(G)$ for n large enough. By taking the infimum over all open sets G containing C, we obtain $\lim_n \mathbb{K}^*(C_n) \preceq \mathbb{K}^*(C)$. □

In order to extend this result to closed sets, we have to impose a tightness condition on the measure $\mathbb{K}$. This condition is equivalent to that defined in probability theory. Individual classical probabilities are frequently tight (in Polish spaces for instance). However, the idempotent equivalent of tightness means that, in a finite dimensional vector space U, the density c^* tends to $\mathbb{0}$ (that is $+\infty$ in $\mathbb{R}_{\min}$) at infinity. This property is often used in optimization in order to ensure the existence of an optimum.

Definition 14. A finite idempotent $\mathbb{D}$-measure $\mathbb{K}$ on $(U,\mathcal{U})$ is said *tight* if

$$\bigwedge_{C\in\mathcal{K}} \mathbb{K}(C^c) = \mathbb{0}.$$

A set Γ of finite idempotent $\mathbb{D}$-measures is said *tight* if

$$\bigwedge_{C\in\mathcal{K}} \bigoplus_{\mathbb{K}\in\Gamma} \mathbb{K}(C^c) = \mathbb{0}.$$

Proposition 15. *For any tight idempotent measure $\mathbb{K}$ on $(U,\mathcal{U})$, the following property holds for its maximal extension :*

$$\lim_n \mathbb{K}^*(F_n) = \mathbb{K}^*(F) \quad \forall F_n,\ F \in \mathcal{F},\ F_n \searrow_n F.$$

Proof. Let C be a compact set. From previous proposition, we have $\lim_n \mathbb{K}^*(F_n \cap C) = \mathbb{K}^*(F \cap C)$, then $\lim_n \mathbb{K}^*(F_n) \preceq \mathbb{K}^*(F \cap C) \oplus \mathbb{K}^*(C^c) \preceq \mathbb{K}^*(F) \oplus \mathbb{K}(C^c)$. By taking the infimum over C and using the tightness condition, we obtain $\lim_n \mathbb{K}^*(F_n) \preceq \mathbb{K}^*(F)$ and the result is proved. □

Independence. Let $\mathbb{K}$ be an idempotent $\mathbb{D}$-probability. As in classical probability theory, we may define independence of events and conditional probabilities. For instance A and B are said *independent* if $\mathbb{K}(A \cap B) = \mathbb{K}(A) \otimes \mathbb{K}(B)$. For conditional probabilities, let us first suppose that $\mathbb{D}$ is a semifield. For any subsets A and B of U, the *conditional cost excess* (the name is chosen in reference to minimization problems) of being in A knowing that we are in B, denoted $\mathbb{K}(A|B)$, is by definition the unique solution of $\mathbb{K}(A|B) \otimes \mathbb{K}(B) = \mathbb{K}(A \cap B)$. Then $\mathbb{K}(.|B)$ is an idempotent probability on B. If $\mathbb{D}$ is only a semiring, but is such that $\otimes$ is distributive with respect to the $\wedge$ law, we may define $\mathbb{K}(A|B)$ as the minimal solution of $\mathbb{K}(A|B) \otimes \mathbb{K}(B) \succeq \mathbb{K}(A \cap B)$. Then, $\mathbb{K}(.|B)$ is a finite idempotent measure on B ($\mathbb{K}(.|B) \preceq \mathbb{1}$) but may not be a probability ($\mathbb{K}(B|B) \neq \mathbb{1}$). For instance, in $(\overline{\mathbb{R}}, \max, \min)$, $\mathbb{K}(A|B) = \mathbb{K}(A \cap B)$ and $\mathbb{K}(B|B) = \mathbb{K}(B)$ which is in general different from $\mathbb{1} = +\infty$. Moreover, the independence of A and B is not equivalent to $\mathbb{K}(A|B) = \mathbb{K}(A)$.

3 Idempotent integrals.

In classical probability theory, real random variables and expectations are just other words for measurable functions and integrals. We then recall here some results on the idempotent integral introduced by Maslov [19]. The notations are taken from [1].

We denote by $\mathcal{I}(U,\mathcal{U})$ the set of functions from U to $\mathbb{D}$ which are countable $\mathbb{D}$-linear combinations of characteristic functions $\mathbb{1}_A$ of sets $A \in \mathcal{U}$ ($\mathbb{1}_A(x) = \mathbb{1}$ in A and $\mathbb{1}_A(x) = \mathbb{0}$ outside A). Then, $\mathcal{I}(U,\mathcal{U})$ is a $\mathbb{D}$-semi-algebra and a lattice, stable by countable upper bounded (by any function) supremum. Moreover, there exists a unique $\mathbb{D}$-linear form $\mathbb{V}$ (with values in $\overline{\mathbb{D}}$) on $\mathcal{I}(U,\mathcal{U})$, continuous on converging nondecreasing sequences (i.e. such that $\mathbb{V}(f_n) \nearrow_{n\to+\infty} \mathbb{V}(f)$ if $f_n \nearrow_{n\to+\infty} f$) and extending $\mathbb{K}$, that is such that $\mathbb{V}(\mathbb{1}_A) = \mathbb{K}(A)$ for any $A \in \mathcal{U}$. We refer to it as the *Maslov integral* with respect to the idempotent probability $\mathbb{K}$. The integral $\mathbb{V}(f)$ will be sometimes denoted $\mathbb{V}_{\mathbb{K}}(f)$ or even $\mathbb{K}(f)$ in order to make the idempotent probability $\mathbb{K}$ explicit. We say that a function $f \in \mathcal{I}(U,\mathcal{U})$ is integrable if $\mathbb{V}(f) \in \mathbb{D}$. Let us note that, since the $\otimes$ law is not necessarily distributive with respect to countable $\oplus$ in $\overline{\mathbb{D}}$, the extension of $\mathbb{V}$ to the set of $\overline{\mathbb{D}}$-valued functions may not exist.

We now recall the characterization (similar to that of measurable functions) of the elements of $\mathcal{I}(U,\mathcal{U})$ by their level sets. We denote by $\mathcal{S}(U,\mathcal{U})$ the set of semi-measurable functions with respect to $\mathcal{U}$: $\mathcal{S}(U,\mathcal{U}) = \{f : U \mapsto \mathbb{D},\ U_f(a) \in \mathcal{U}\ \forall a \in \mathbb{D}\}$ where $U_f(a) \stackrel{\text{def}}{=} \{u \in U,\ a \ll f(u)\}$. For any semi-measurable function f, we also denote by $\mathcal{U}(f)$ the semi-σ-algebra generated by the sets $U_f(a)$ for $a \in \mathbb{D}$. Then, $\mathcal{I}(U,\mathcal{U})$ is exactly the set of functions f of $\mathcal{S}(U,\mathcal{U})$ such that $\mathcal{U}(f)$ has a countable basis, that is a countable subset $\mathcal{B}$ of $\mathcal{U}$ (not necessarily included in $\mathcal{U}(f)$) stable by finite intersection and such that the elements of $\mathcal{U}(f)$ are unions of elements of $\mathcal{B}$.

If $\mathbb{D}$ has a countable basis (as a locally continuous lattice [15, 1], this is the case for $\mathbb{R}_{\max}$) or $\mathcal{U}$ has a countable basis, then $\mathcal{I}(U,\mathcal{U}) = \mathcal{S}(U,\mathcal{U})$.

In any case we have the following general expression for $\mathbb{V}$:

$$\mathbb{V}(f) = \bigoplus_{a\in\mathbb{D}} a \otimes \mathbb{K}(U_f(a)).$$

Let $f \in \mathcal{I}(U,\mathcal{U})$, and let $\mathcal{B}$ denote any countable basis of $\mathcal{U}(f)$. Then, $\mathbb{K}$ has a density $c_{\mathcal{B}}$ on the semi-σ-algebra generated by $\mathcal{B}$ [1] and

$$\mathbb{V}(f) = \bigoplus_{u\in U} f(u) \otimes c_{\mathcal{B}}(u). \tag{6}$$

If $\mathcal{U}$ has a countable basis, then $\mathbb{K}$ has as density c^* on the entire algebra $\mathcal{U}$ and

$$\mathbb{V}(f) = \bigoplus_{u \in U} f(u) \otimes c^*(u).$$

Moreover, when considering a countable subset $\mathcal{F}$ of $\mathcal{I}(U, \mathcal{U})$, we can do as if $\mathbb{K}$ has a density, even it has no density on $\mathcal{U}$. Indeed, the expression (6) is valid with $\mathcal{B}$ equal to the countable basis generated by the basis of all elements of $\mathcal{F}$. As a consequence we have the following result.

Proposition 16. *Let $(U, \mathcal{U}, \mathbb{K})$ be a decision space and δ an idempotent algebra distance (resp. a complete idempotent algebra distance) on $\mathbb{D}$, then for any integrable (resp. not necessarily integrable) elements f and g of $\mathcal{I}(U, \mathcal{U})$, we have :*

$$\delta(\mathbb{V}(f), \mathbb{V}(g)) \leq \delta_\infty(f, g) \stackrel{\text{def}}{=} \sup_{u \in U} \delta(f(u), g(u)).$$

If $\mathbb{K}$ is only a finite idempotent $\mathbb{D}$-measure, we have

$$\delta(\mathbb{V}(f), \mathbb{V}(g)) \leq \delta_0(\mathbb{K}(U)) \delta_\infty(f, g).$$

Proof. Let $\mathcal{B}$ be a countable basis generated by the basis of $\mathcal{U}(f)$ and $\mathcal{U}(g)$. Then, from properties of δ, we have

$$\begin{aligned}
\delta(\mathbb{V}(f), \mathbb{V}(g)) &= \delta(\bigoplus_{u \in U} f(u) \otimes c_{\mathcal{B}}(u), \bigoplus_{u \in U} g(u) \otimes c_{\mathcal{B}}(u)) \\
&\leq \sup_{u \in U} \delta(f(u) \otimes c_{\mathcal{B}}(u), g(u) \otimes c_{\mathcal{B}}(u)) \\
&\leq \sup_{u \in U} \delta_0(c_{\mathcal{B}}(u)) \delta(f(u), g(u)) \\
&\leq \delta_\infty(f, g) \sup_{u \in U} \delta_0(c_{\mathcal{B}}(u)).
\end{aligned}$$

Since $c_{\mathcal{B}}(u) \preceq \mathbb{K}(U)$ for any $u \in U$, δ_0 is nondecreasing and $\delta_0(\mathbb{1}) = 1$, we get the proposition. $\square$

Let us finally note that since an idempotent probability may be extended to the set $\mathcal{P}(U)$ of all subsets of U, we may consider the idempotent integral, with respect to this extended probability, of any function of $\mathcal{I}(U, \mathcal{P}(U))$, that is of any function $f : U \to \mathbb{D}$ if $\mathcal{U}$ or $\mathbb{D}$ has a countable basis. This integral coincides with the previous one on the set of semi-measurable functions with respect to $\mathcal{U}$. In the sequel, we consider the maximal extension $\mathbb{K}^*$ of $\mathbb{K}$ and the associated integral, which is equal to the maximal extended integral. We may then denote it (without confusion) $\mathbb{V}^*$, $\mathbb{V}^*_{\mathbb{K}}$, $\mathbb{V}_{\mathbb{K}^*}$ or $\mathbb{K}^*$.

4 Decision variables

4.1 Basic definitions

We are now able to introduce the equivalent of random variables, expectations, laws. Here, all names are given in reference to minimization problems ($\mathbb{R}_{\min}$ case). The term "decision variables" (for general random variables) is chosen for the equivalent of random variables, for its value will determinate the decision to take in order to be optimal. They will act as a change of variables in the decision space. The equivalent of real random variables will be called "cost variables", for in $\mathbb{R}_{\min}$ they can be considered as additional costs (to add to c^*). The idempotent integral (expectation) of a cost variable is then the optimal cost, and for this reason will be called "value" as in optimal control problems.

We suppose given a decision space $(U, \mathcal{U}, \mathbb{K})$ and suppose that either $\mathcal{U}$ or $\mathbb{D}$ has a countable basis so that the maximal idempotent integral is defined for any function from $\mathcal{U}$ to $\mathbb{D}$.

Definition 17.

- A *decision variable* X on $(U, \mathcal{U}, \mathbb{K})$ with values in a topological space E is simply a map from U to E.
- The decision variables X and Y are *equal almost surely*, denoted $X \overset{\text{a.s.}}{=} Y$, if $\mathbb{K}^*(\{u, X(u) \neq Y(u)\}) = \mathbb{0}$. If $\mathbb{K}$ has a density, this means that X and Y coincide in $\operatorname{supp}(\mathbb{K})$.
- Denote by $\mathcal{V}$ the set of open sets of E. Any decision variable X with values in E induces on $(E, \mathcal{V})$ an idempotent measure $\mathbb{K}_X$ defined by : $\mathbb{K}_X(V) = \mathbb{K}^*(X^{-1}(V))$ for any $V \in \mathcal{V}$. The maximal density of $\mathbb{K}_X$ if it exists will be denoted by c^*_X and will be called the *cost or law density* of X.
- A decision variable X (resp. a set of decision variables) is *tight* if its law $\mathbb{K}_X$ (resp. the set of the laws of its elements) is tight (see Definition 14).
- Two decision variables X and Y are *independent* when $\mathbb{K}_{X,Y}(A \times B) = \mathbb{K}_X(A) \otimes \mathbb{K}_Y(B)$ for any open sets A and B. A set $(X_i)_{i \in I}$ of decision variables is *independent* when for any finite subset J of I and any open sets A_j, $\mathbb{K}_{(X_j)_{j\in J}}(\times_{j \in J} A_j) = \otimes_{j\in J} \mathbb{K}_X(A_j)$.
- The decision variable X is *measurable* if $X^{-1}(V) \in \mathcal{U}$ for any $V \in \mathcal{V}$ (if $\mathcal{U}$ is the set of open sets of U, this means continuous, if $\mathcal{U}$ is the Borel sets algebra, this means measurable in the usual sense).
- A *cost variable* on $(U, \mathcal{U}, \mathbb{K})$ is a decision variable with values in $\mathbb{D}$.
- The *value* of a cost variable X is the maximal Maslov integral $\mathbb{V}^*(X) = \mathbb{K}^*(X)$ with respect to the idempotent measure $\mathbb{K}$. The cost variable X is said *integrable* if $\mathbb{V}^*(X) \in \mathbb{D}$.

Remark 18. Suppose that $\mathbb{K}$ has a density and $\otimes$ is distributive with respect to $\wedge$. Then, the independence of the decision variables X and Y is equivalent to $c^*_{X,Y}(x, y) = c^*_X(x) \otimes c^*_Y(y)$ for any x and y.

A sequence of independent decision variables is a decision variable $X = (X_n)_{n\in\mathbb{N}}$ with values in $E = \times_{n\in\mathbb{N}} E_n$ such that $\mathbb{K}(X_0 \in A_0, \ldots, X_n \in A_n) = \mathbb{K}_{X_0}(A_0) \otimes \cdots \otimes \mathbb{K}_{X_n}(A_n)$ for any $n \in \mathbb{N}$. This implies $c^*_X(x) = \otimes_{n\in\mathbb{N}} c_n(x_n) \overset{\text{def}}{=} \lim{\downarrow}_N \otimes_{n \le N} c_n(x_n)$ for $x = (x_n) \in E^{\mathbb{N}}$ where c_n is the cost density of X_n. This last condition is equivalent to the independence of X_n if for any $a \ll \mathbb{1}$ there exists a sequence $a_n \ll \mathbb{1}$ such that $a \preceq \otimes_{n\in\mathbb{N}} a_n$. This is true in $\mathbb{R}_{\min}$, where the independence condition is then equivalent to $c^*_X(x) = \sum_{n\in\mathbb{N}} c_n(x_n)$. Hence, discrete optimal control problems correspond to $\mathbb{R}_{\min}$-Markov chains that we call Bellman chains.

Remark 19. If $\mathbb{K}$ has a density, the density c^*_X of a decision variable X is the upper semi-continuous envelope of the function $c_X(x) = \mathbb{K}^*(X^{-1}(\{x\})) = \oplus_{u,X(u)=x} c^*(u)$ ($= \inf_{X(u)=x} c^*(u)$ if $\mathbb{D} = \mathbb{R}_{\min}$). If c_X and c^*_X are not equal, the usual formula of change of variables in integrals may be false when considering maximal integrals. Indeed,

$$\mathbb{V}^*_{\mathbb{K}}(f(X)) \stackrel{\text{def}}{=} \bigoplus_{u\in U} f(X(u)) \otimes c^*(u) = \bigoplus_{x\in E} f(x) \otimes c_X(x)$$

and

$$\mathbb{V}^*_{\mathbb{K}_X}(f) \stackrel{\text{def}}{=} \bigoplus_{x\in E} f(x) \otimes c^*_X(x),$$

then $\mathbb{V}^*_{\mathbb{K}}(f(X))$ may be different from $\mathbb{V}^*_{\mathbb{K}_X}(f)$. However, for any lower semi continuous (l.s.c.) function f from E to $\mathbb{D}$ we have

$$\mathbb{V}^*_{\mathbb{K}}(f(X)) = \mathbb{V}^*_{\mathbb{K}_X}(f).$$

Indeed, l.s.c. functions coincide with semi-measurable functions with respect to $\mathcal{V}$ (we suppose that $\mathcal{V}$ or $\mathbb{D}$ has a countable basis).

Therefore, in our formalism we are allowed to consider non u.s.c. (resp. l.s.c. if $\mathbb{D} = \mathbb{R}_{\min}$) cost densities but only when calculating the supremum of this function multiplied by (resp. plus) a l.s.c. (resp. u.s.c.) function.

Remark 20. Suppose that $\mathcal{U}$ is a topology and that X is a measurable (i.e. continuous) decision variable. One can prove that c^*_X and c_X are equal, under the condition that $\mathbb{K}$ is tight. Indeed,

$$c^*_X(x) = \inf_{V\in\mathcal{V}, V\ni x} \mathbb{K}^*(X^{-1}(V)) = \lim_n \mathbb{K}^*(X^{-1}(B(x,\frac{1}{n}))) = \lim_n \mathbb{K}^*(X^{-1}(\overline{B}(x,\frac{1}{n}))),$$

where $B(x,\varepsilon)$ and $\overline{B}(x,\varepsilon)$ denote the open and closed ball of center x and radius ε in (E,d). Moreover, $X^{-1}(\overline{B}(x,\frac{1}{n}))$ is a nonincreasing sequence of closed sets of U converging to $X^{-1}(\{x\})$ and $c_X(x) = \mathbb{K}^*(X^{-1}(\{x\}))$. Then, if $\mathbb{K}^*$ is continuous under nonincreasing sequences of closed sets, c^*_X is equal to c_X. This property holds if $\mathbb{K}$ is tight (see section 2).

Otherwise, let us give a counter example. We consider on the dioid $\mathbb{R}_{\min}$, the usual topology $\mathcal{U}$ of $U = \mathbb{R}$, the cost measure $\mathbb{K}$ with density $c^*(u) = 1/|u|$ if $u \neq 0$, $c^*(0) = +\infty$ and the measurable decision variable $X(u) = |u|/(|u|+1)$. We obtain $c_X(x) = (1/x) - 1$ if $x \in (0,1)$ and $c_X(x) = +\infty$ otherwise. Then c_X is not l.s.c. and thus not equal to c^*_X.

Even if we have pointed out some difficulties due to maximal extensions, we will now omit to put the star on $\mathbb{K}$ or $\mathbb{V}$. We can think without ambiguity that the initial decision space is $(U, \mathcal{P}(U), \mathbb{K}^*)$ instead of $(U, \mathcal{U}, \mathbb{K})$. Moreover, since we will only consider continuous functions of decision variables on the same decision space, c_X and c^*_X play the same role and the formula of change of variables in integrals holds.

Proposition 21. *If X and Y are cost variables such that $X \stackrel{\text{a.s.}}{=} Y$, then $\mathbb{V}(X) = \mathbb{V}(Y)$ and $c^*_X = c^*_Y$. Moreover, for any countable families $(X_i)_{i\in I}$ and $(Y_i)_{i\in I}$ (if $\mathbb{K}$ has a density, I may be non countable), $X_i \stackrel{\text{a.s.}}{=} Y_i$ implies $\oplus_{i\in I} X_i \stackrel{\text{a.s.}}{=} \oplus_{i\in I} Y_i$.*

Then, $\mathbb{V}$ is a $\mathbb{D}$-linear form on the $\mathbb{D}$-linear space of classes of (even non integrable) cost variables. As a consequence $\mathbb{V}$ is nondecreasing and the set of classes of integrable cost variables is a $\mathbb{D}$-linear space.

Proof. The first assertion is obvious. For the second one, we have

$$\mathbb{K}(\{u, \bigoplus_{i\in I} X_i(u) \neq \bigoplus_{i\in I} Y_i(u)\}) \preceq \mathbb{K}(\bigcup_{i\in I} \{u, X_i(u) \neq Y_i(u)\}) = \bigoplus_{i\in I} \mathbb{K}(\{u, X_i(u) \neq Y_i(u)\}) = \mathbb{0},$$

and if $\mathbb{K}$ has a density, $\mathbb{K}$ is additive over any infinite union and the last equality holds for any set I. The third assertion is then a consequence of previous assertions and of the definition of the Maslov integral. □

Then, as in classical integration theory we may identify decision variables with their class for the $\overset{\text{a.s.}}{=}$ equivalence relation.

Since we do not use them in this paper, we do not give the definitions of conditional value (expectation), martingales,.... We refer to [2] and [11] for these points in particular spaces.

The optimum, that is the value of the decision variable in the point where the cost is optimal, is another interesting quantity, which in addition is the image of the expectation by the Cramer transform [2].

Definition 22. Suppose that $\mathbb{K}$ has a density and let X be a decision variable, then the optimum of X is defined as the element $\mathbb{O}(X) \overset{\text{def}}{=} \mathrm{Argmax}_{x\in E}\, c^*_X(x)$, when it exists and is unique.

Since it depends only on the cost density c^*_X, the optimum of X depends only on the class of X for the $\overset{\text{a.s.}}{=}$ equivalence relation. When considering multicriteria problems (for instance if $\mathbb{D} = (\mathbb{R}_{\max})^n$), the Argmax may be replaced by the Pareto set $\mathbb{O}(X) = \{x, (y \in E$ and $c^*_X(x) \preceq c^*_X(y)) \Rightarrow x = y\}$. In $\mathbb{R}_{\max}$, it may also be more interesting to consider the Argmax of the u.s.c. concave hull of c^*_X, since it is equal to the differential of the characteristic function in point 0 (see section 6).

4.2 Convergence

Definition 23. Consider a sequence X_n of decision variables and a decision variable X on $(U, \mathcal{U}, \mathbb{K})$ with values in the same metric space (E, d). We say that

- X_n *converges almost surely* towards X, denoted $X_n \xrightarrow{\text{a.s.}} X$, if and only if

$$\mathbb{K}(\{u, X_n(u) \underset{n\to+\infty}{\not\longrightarrow} X(u)\}) = \mathbb{0},$$

- X_n *converges in cost* towards X, denoted $X_n \xrightarrow{\mathbb{K}} X$, if and only if

$$\mathbb{K}(\{u, d(X_n(u), X(u)) \geq \varepsilon\}) \underset{n\to+\infty}{\longrightarrow} \mathbb{0} \quad \text{for any } \varepsilon > 0.$$

- X_n *converges in law or weakly* towards X, denoted $X_n \xrightarrow{\mathrm{w}} X$, if and only if

$$\mathbb{V}(f(X_n)) \underset{n\to+\infty}{\longrightarrow} \mathbb{V}(f(X))$$

  for any function $f \in \mathcal{C}_b(E,\mathbb{D})$, the set of bounded continuous functions from (E,d) to $(\mathbb{D},\delta)$.

- If $\mathcal{C}_b(E,\mathbb{D})$ is replaced by the set $\mathcal{C}'_b(E,\mathbb{D})$ of bounded and uniformly continuous functions from (E,d) to $(\mathbb{D},\delta)$, we say that X_n *converges in weak-law* towards X and denote it by $X_n \xrightarrow{\mathrm{w'}} X$.

The following definition generalizes the classical definition of equi-integrability to partially ordered idempotent semirings $\mathbb{D}$.

Definition 24. A sequence X_n of cost variables is said *equi-integrable* if and only if

$$\bigoplus_{n,m\in\mathbb{N}} \mathbb{V}(X_n \otimes \mathbb{1}_{\delta(X_m,\mathbb{0})\geq k}) \underset{k\to+\infty}{\longrightarrow} \mathbb{0}.$$

If $\mathbb{D}$ is totally ordered, dividing the integral into the component where $X_n \preceq X_m$ and its complementary, we obtain $\oplus_{n,m\in\mathbb{N}} \mathbb{V}(X_n \otimes \mathbb{1}_{\delta(X_m,\mathbb{0})\geq k}) = \oplus_{n\in\mathbb{N}} \mathbb{V}(X_n \otimes \mathbb{1}_{\delta(X_n,\mathbb{0})\geq k})$ and our definition of equi-integrability corresponds to the classical one.

In the following we will need the notion of strong integrability which has no equivalent in probability or integration. It is introduced in order to overcome the difficulty due to the non continuity of idempotent measures or idempotent integrals over nonincreasing sequences. Indeed, if we consider an integrable cost variable X, the sequence $(X \otimes \mathbb{1}_{\delta(X,\mathbb{0})\geq k})_k$ is nonincreasing and converges to $X \otimes \mathbb{1}_{\delta(X,\mathbb{0})=+\infty} = \mathbb{0}$ (since $X \in \mathbb{D}$), but $\mathbb{V}(X \otimes \mathbb{1}_{\delta(X,\mathbb{0})\geq k})$ may not converge to $\mathbb{V}(\mathbb{0}) = \mathbb{0}$. Therefore, the constant sequence $X_n = X$ may not be equi-integrable (contrary to the classical probability theory). Note also that the equi-integrability of the sequence X_n is equivalent to the condition $\mathbb{V}(X \otimes \mathbb{1}_{\delta(X,\mathbb{0})\geq k}) \longrightarrow_{k\to+\infty} \mathbb{0}$ for $X = \oplus_n X_n$ (indeed, $\mathbb{V}$ is additive and $\delta(\oplus_n X_n, \mathbb{0}) = \sup_n \delta(X_n, \mathbb{0})$ by properties 1 and 3 of δ).

Definition 25. The cost variable X is said *strongly integrable* if

$$\mathbb{V}(X \otimes \mathbb{1}_{\delta(X,\mathbb{0})\geq k}) \underset{k\to+\infty}{\longrightarrow} \mathbb{0}.$$

An upper bounded cost variable is strongly integrable. If $X \preceq Y$ and Y is strongly integrable, so does X (since $x \otimes \delta(x,\mathbb{0})$ is monotone). A sequence X_n is equi-integrable iff $\oplus_n X_n$ is strongly integrable. Then a sequence X_n upper bounded by a strongly integrable cost variable is equi-integrable. In probability theory, we often use the condition "X_n upper and lower bounded by integrable random variables", in place of the equi-integrability. Here, the lower bound is not required since all numbers are $\succeq \mathbb{0}$ and the upper bound has to be strongly integrable.

Example 26. In $\mathbb{R}_{\max}$, a cost variable X with cost density c^*_X is integrable if and only if $x + c^*_X(x)$ is upper bounded. It is strongly integrable if and only if $x + c^*_X(x) \longrightarrow_{x\to+\infty} -\infty$. If X is integrable, then $(1-\varepsilon)X$ is strongly integrable for any $\varepsilon > 0$.

Remark 27. In general (except if $\mathbb{D}$ is totally ordered) the set of strongly integrable cost variables is not stable by the $\oplus$ operation and then it is not a $\mathbb{D}$-linear space. Indeed, consider the semiring $(\mathbb{R}_{\max})^2$ and the space $U = \mathbb{R}^2$ with usual topology $\mathcal{U}$ and the idempotent probability $\mathbb{K}$ with density $c^*(u) = (-u_1^2, -u_2^2)$ for $u = (u_1, u_2)$. Then, the cost variables $X(u) = (\frac{1}{2}u_1^2, -u_1^2)$ and $Y(u) = (-u_2^2, \frac{1}{2}u_2^2)$ are strongly integrable whereas $Z(u) = X(u) \oplus Y(u) = (\frac{1}{2}u_1^2, \frac{1}{2}u_2^2)$ is not strongly integrable. Indeed, $\delta(X(u), \mathbb{0}) = e^{\frac{1}{2}u_1^2}$, then $\mathbb{V}(X \otimes \mathbb{1}_{\delta(X,\mathbb{0}) \geq k}) = (-\log k, -2\log k) \longrightarrow_{k \to +\infty} \mathbb{0}$ whereas $\mathbb{V}(Z \otimes \mathbb{1}_{\delta(Z,\mathbb{0}) \geq k}) = (0, 0) \neq \mathbb{0}$.

5 Relations between the different notions of convergence

The essential difference between idempotent probability and classical probability theories is that the cost convergence is stronger than the almost sure convergence. Almost all other differences will be consequences of this fact.

Proposition 28. *Let X_n and X denote decision variables with values in the same metric space (E, d).* $X_n \xrightarrow{\mathbb{K}} X$ *implies* $X_n \xrightarrow{\text{a.s.}} X$.

Proof. We use $\{u, X_n(u) \underset{n \to +\infty}{\not\longrightarrow} X(u)\} = \underset{k>0}{\cup} \lim\downarrow_N \underset{n \geq N}{\cup} \{u, d(X_n(u), X(u)) \geq \frac{1}{k}\}$. Using the properties of an idempotent measure, we have

$$\mathbb{K}(X_n \underset{n \to +\infty}{\not\longrightarrow} X) = \underset{k>0}{\oplus} \mathbb{K}(\lim\downarrow_N \underset{n \geq N}{\cup} \{d(X_n, X) \geq \frac{1}{k}\}) \tag{7}$$

$$\preceq \underset{k>0}{\oplus} \lim\downarrow_N \mathbb{K}(\underset{n \geq N}{\cup} \{d(X_n, X) \geq \frac{1}{k}\}) \tag{8}$$

$$= \underset{k>0}{\oplus} \lim\downarrow_N \underset{n \geq N}{\oplus} \mathbb{K}(d(X_n, X) \geq \frac{1}{k}). \tag{9}$$

Then, if $\mathbb{K}(d(X_n, X) \geq \frac{1}{k}) \underset{n \to +\infty}{\longrightarrow} \mathbb{0}$, $\lim\downarrow_N \underset{n \geq N}{\oplus} \mathbb{K}(d(X_n, X) \geq \frac{1}{k}) = \limsup_{n \to +\infty} \mathbb{K}(d(X_n, X) \geq \frac{1}{k}) = \mathbb{0}$ and $\mathbb{K}(X_n \underset{n \to +\infty}{\not\longrightarrow} X) = \mathbb{0}$. $\square$

Remark 29. The differences encountered with the classical probability theory were : a) (7) and (9) are equalities instead of inequalities ($\succeq$) because $\oplus$ is idempotent, b) (8) is in general an inequality instead of an equality because of the non continuity of idempotent measures over nonincreasing sequences. Let $X = (X_n)_{n \in \mathbb{N}}$ be a sequence of independent real-valued decision variables on $\mathbb{R}_{\min}$ with same cost density $c(x) = x^2$. Then $c_X^*(x = (x_n)) = \sum_{n=0}^{+\infty} x_n^2$ and $c_X^*(x) \neq \mathbb{0} = +\infty$ implies $x_n \longrightarrow_{n \to +\infty} 0$. Therefore $X_n \xrightarrow{\text{a.s.}} 0$, whereas $X_n \overset{\mathbb{K}}{\not\longrightarrow} 0$. However, an i.i.d. sequence of random variables converges almost surely to 0 iff each variable is almost surely equal to 0. This shows how much the idempotent a.s. convergence is poor compared to its classical equivalent.

Lemma 30 ("Fatou's lemma"). *Let X_n and X denote cost variables.*

1. *If* $X_n \underset{n\to+\infty}{\nearrow} X$ *almost surely, then* $\mathbb{V}(X_n) \underset{n\to+\infty}{\nearrow} \mathbb{V}(X)$.
2. *If* X_0 *is strongly integrable and* $X_n \underset{n\to+\infty}{\searrow} X$ *in cost, then* $\mathbb{V}(X_n) \underset{n\to+\infty}{\searrow} \mathbb{V}(X)$ *and* X *is strongly integrable.*

Proof. Property 1 is a consequence of Proposition 21. For property 2, $\mathbb{V}(X_n)$ is nonincreasing and $\lim_n \mathbb{V}(X_n) \succeq \mathbb{V}(X)$. Consider ε and $k > 0$. We have

$$\begin{aligned}\mathbb{V}(X_n) &= \mathbb{V}(X_n \otimes 1\!\!1_{\delta(X_n,X)\leq\varepsilon}) \oplus \mathbb{V}(X_n \otimes 1\!\!1_{\delta(X_n,\mathbb{0})\leq k} \otimes 1\!\!1_{\delta(X_n,X)\geq\varepsilon}) \oplus \mathbb{V}(X_n \otimes 1\!\!1_{\delta(X_n,\mathbb{0})\geq k}) \\ &\preceq \mathbb{V}(X_n \otimes 1\!\!1_{\delta(X_n,X)\leq\varepsilon}) \oplus d_k \otimes \mathbb{K}(\delta(X_n, X) \geq \varepsilon) \oplus \mathbb{V}(X_0 \otimes 1\!\!1_{\delta(X_0,\mathbb{0})\geq k}),\end{aligned}$$

where d_k is defined in (2).

Using $\mathbb{V}(X_n) \succeq \mathbb{V}(X) \succeq \mathbb{V}(X \otimes 1\!\!1_{\delta(X_n,X)\leq\varepsilon})$ and the properties of δ and $\mathbb{V}$, we obtain

$$\begin{aligned}\delta(\mathbb{V}(X_n), \mathbb{V}(X)) &\leq \max(\delta(\mathbb{V}(X_n \otimes 1\!\!1_{\delta(X_n,X)\leq\varepsilon}), \mathbb{V}(X \otimes 1\!\!1_{\delta(X_n,X)\leq\varepsilon})), \\ &\qquad \delta(d_k \otimes \mathbb{K}(\delta(X_n, X) \geq \varepsilon), \mathbb{0}), \delta(\mathbb{V}(X_0 \otimes 1\!\!1_{\delta(X_0,\mathbb{0})\geq k}), \mathbb{0})) \\ &\leq \max(\varepsilon, \delta_0(d_k)\delta(\mathbb{K}(\delta(X_n, X) \geq \varepsilon), \mathbb{0}), \delta(\mathbb{V}(X_0 \otimes 1\!\!1_{\delta(X_0,\mathbb{0})\geq k}), \mathbb{0}))\end{aligned} \tag{10}$$

For k large enough, the third term is less than ε. Then, $\limsup_{n\to+\infty} \delta(\mathbb{V}(X_n), \mathbb{V}(X)) \leq \varepsilon$ when $X_n \xrightarrow{\mathbb{K}} X$. This implies the second assertion of the lemma. □

Theorem 31 ("Lebesgue's dominated convergence"). *Let* X_n *and* X *denote cost variables.*

1. *If* $X_n \xrightarrow{a.s.} X$, *then* $\liminf_{n\to+\infty} \mathbb{V}(X_n) \succeq \mathbb{V}(X)$.
2. *If* X_n *is equi-integrable and* $X_n \xrightarrow{\mathbb{K}} X$, *then* $\lim_{n\to+\infty} \mathbb{V}(X_n) = \mathbb{V}(X)$ *and* X *is strongly integrable.*

Proof. 1) From the monotony of $\mathbb{V}$, we have $\liminf_n \mathbb{V}(X_n) \succeq \lim_n \mathbb{V}(Y_n)$, where $Y_n = \wedge_{m\geq n} X_m$. Moreover, Y_n is nondecreasing and converges almost surely towards X, then the first point of previous lemma implies $\lim_n \mathbb{V}(Y_n) = \mathbb{V}(X)$ and the first point of the theorem is proved.

2) Since the convergence in cost implies almost sure convergence, $\mathbb{V}(X) \preceq \liminf_n \mathbb{V}(X_n)$. Let us prove $\limsup_n \mathbb{V}(X_n) \preceq \mathbb{V}(X)$. We have

$$\limsup_n \mathbb{V}(X_n) = \lim_N\!\downarrow \bigoplus_{n\geq N} \mathbb{V}(X_n) = \lim_N\!\downarrow \mathbb{V}(Y_N),$$

where $Y_N = \oplus_{n\geq N} X_n$ is a nonincreasing sequence converging (at least almost surely) towards X. If now Y_N converges in cost towards X and Y_0 is strongly integrable, point 2 of the theorem will be a consequence of Fatou's lemma (Lemma 30).

But $\delta(Y_N, X) \leq \sup_{n\geq N} \delta(X_n, X)$, then using the properties of δ and $\mathbb{K}$, we have

$$\mathbb{K}(\delta(Y_N, X) \geq \varepsilon) \preceq \mathbb{K}(\bigcup_{n\geq N} \{\delta(X_n, X) \geq \varepsilon/2\}) = \bigoplus_{n\geq N} \mathbb{K}(\delta(X_n, X) \geq \varepsilon/2) \underset{N\to+\infty}{\longrightarrow} \mathbb{0},$$

then $Y_N \xrightarrow{\mathbb{K}} X$. Moreover, the equi-integrability of X_n is equivalent to the strong integrability of $Y_0 = \oplus_{n\geq 0} X_n$. □

In probability theory there is an equivalence between the equi-integrability of a sequence X_n together with its convergence in probability towards some variable X and the L^1 convergence of X_n towards X. Moreover, the mean convergence (L^1) is equivalent to the convergence of the expectation of X_n restricted to some subset A towards the expectation of X restricted to A uniformly in A. In "idempotent probability theory", the mean convergence (L^1) cannot be clearly defined, but the second equivalent definition may be considered, and the equivalence proved in probability theory holds for a certain class of idempotent semirings $\mathbb{D}$ including $\mathbb{R}_{\max}$.

Theorem 32. *Let X_n and X denote cost variables.*

1. *If X_n is equi-integrable and $X_n \xrightarrow{\mathbb{K}} X$, then X is strongly integrable and*

$$\mathbb{V}(X_n \otimes \mathbb{1}_A) \underset{n\to+\infty}{\longrightarrow} \mathbb{V}(X \otimes \mathbb{1}_A) \quad \textit{uniformly in } \; A \subset U \tag{11}$$

2. *Suppose that $\mathbb{K}$ has a density and that the distance δ of $\mathbb{D}$ satisfies :*

$$\delta(x \otimes \lambda, y \otimes \lambda) = \delta_0(\lambda)\delta(x, y) \quad \forall \lambda, x, y \in \mathbb{D}.$$

   Consider strongly integrable cost variables X and X_n such that $\oplus_{n\le N} X_n$ is strongly integrable for any $N \in \mathbb{N}$. Then the following propositions are equivalent :

   (a) X_n is equi-integrable and $X_n \xrightarrow{\mathbb{K}} X$.

   (b) (11) holds,

   (c) $\sup_{u\in U} \delta(X_n(u), X(u))\delta_0(c^*(u)) \underset{n\to+\infty}{\longrightarrow} 0$.

Proof. 1) For the assertion 1, "Lebesgue's dominated convergence" (Theorem 31) already implies that X is strongly integrable and $\mathbb{V}(X_n \otimes \mathbb{1}_A) \longrightarrow_{n\to+\infty} \mathbb{V}(X \otimes \mathbb{1}_A)$ for any $A \subset U$. Indeed, it is easy to show that $X_n \otimes \mathbb{1}_A \xrightarrow{\mathbb{K}} X \otimes \mathbb{1}_A$ and that $X_n \otimes \mathbb{1}_A$ is equi-integrable. In order to get the uniformity with respect to A, we review the proof of Theorem 31.

Let us consider $Z_N = \wedge_{n\ge N} X_n$ and $Y_N = \oplus_{n\ge N} X_n$, we have $Z_n \preceq X_n \preceq Y_n$, Y_0 is strongly integrable, Y_n decreases and converges in cost towards X (see the proof of Theorem 31) and Z_n increases and converges in cost towards X (by the same proof). Since

$$\delta(\mathbb{V}(X_n \otimes \mathbb{1}_A), \mathbb{V}(X \otimes \mathbb{1}_A)) \le \max(\delta(\mathbb{V}(Y_n \otimes \mathbb{1}_A), \mathbb{V}(X \otimes \mathbb{1}_A)), \delta(\mathbb{V}(Z_n \otimes \mathbb{1}_A), \mathbb{V}(X \otimes \mathbb{1}_A))),$$

it is sufficient to prove (11) for Z_n and Y_n that is for monotone sequences. For this, we use the same decomposition as in the proof of "Fatou's lemma". For Y_n we use (10), $(\delta(Y_n \otimes \mathbb{1}_A, X \otimes \mathbb{1}_A) \ge \varepsilon \Leftrightarrow \delta(Y_n, X) \ge \varepsilon)$ and the obvious fact that $Y_0 \otimes \mathbb{1}_A \preceq Y_0$. Then

$$\begin{aligned} &\sup_{A\subset U} \delta(\mathbb{V}(Y_n \otimes \mathbb{1}_A), \mathbb{V}(X \otimes \mathbb{1}_A)) \\ &\quad \le \quad \max(\varepsilon, \delta_0(d_k)\delta(\mathbb{K}(\delta(Y_n, X) \ge \varepsilon), \mathbb{0}), \delta(\mathbb{V}(Y_0 \otimes \mathbb{1}_{\delta(Y_0,\mathbb{0})\ge k}), \mathbb{0})), \end{aligned} \tag{12}$$

which implies the uniform convergence of $\mathbb{V}(Y_n \otimes \mathbb{1}_A)$ towards $\mathbb{V}(X \otimes \mathbb{1}_A)$.

Since the previous inequality (12) only used $X \preceq Y_n \preceq Y_0$, replacing X by Z_n and Y_n and Y_0 by X we obtain

$$\begin{aligned}&\sup_{A\subset U} \delta(\mathbb{V}(Z_n\otimes \mathbb{1}_A),\mathbb{V}(X\otimes \mathbb{1}_A))\\ &\quad\leq\quad \max(\varepsilon,\delta_0(d_k)\delta(\mathbb{K}(\delta(Z_n,X)\geq\varepsilon),\mathbb{0}),\delta(\mathbb{V}(X\otimes \mathbb{1}_{\delta(X,\mathbb{0})\geq k}),\mathbb{0})),\end{aligned}$$

and $\mathbb{V}(Z_n\otimes \mathbb{1}_A)$ tends towards $\mathbb{V}(X\otimes \mathbb{1}_A)$ uniformly in $A\subset U$.

2) For the assertion 2, we have already proved $2a \Rightarrow 2b$. Let us prove the converse implication. Suppose that (11) holds and let us prove the cost convergence of X_n towards X. Let c be a density of $\mathbb{K}$, we have

$$\delta(\mathbb{K}(\delta(X_n,X)\geq\varepsilon),\mathbb{0})=\delta(\bigoplus_{u,\delta(X_n(u),X(u))\geq\varepsilon} c(u),\mathbb{0})\leq \sup_{u,\delta(X_n(u),X(u))\geq\varepsilon}\delta(c(u),\mathbb{0}),$$

and using the additional property of δ supposed in the theorem, we have

$$\delta(c(u),\mathbb{0})=\delta(\mathbb{1},\mathbb{0})\delta_0(c(u))=\delta(\mathbb{1},\mathbb{0})\delta(X_n(u)\otimes c(u),X(u)\otimes c(u))/\delta(X_n(u),X(u)).$$

Then

$$\begin{aligned}\delta(\mathbb{K}(\delta(X_n,X)\geq\varepsilon),\mathbb{0}) &\leq \frac{\delta(\mathbb{1},\mathbb{0})}{\varepsilon}\sup_{u,\delta(X_n(u),X(u))\geq\varepsilon}\delta(\mathbb{V}(X_n\otimes \mathbb{1}_u),\mathbb{V}(X\otimes \mathbb{1}_u))\\ &\leq \frac{\delta(\mathbb{1},\mathbb{0})}{\varepsilon}\sup_{A\subset U}\delta(\mathbb{V}(X_n\otimes \mathbb{1}_A),\mathbb{V}(X\otimes \mathbb{1}_A)).\end{aligned}$$

Then, (11) implies $X_n \xrightarrow{\mathbb{K}} X$.

Let us prove now the equi-integrability of X_n. We have

$$\begin{aligned}&\delta(\mathbb{V}(X_n\otimes \mathbb{1}_{\delta(X_m,\mathbb{0})\geq k}),\mathbb{0})\\ &\quad\leq\quad \sup_A \delta(\mathbb{V}(X_n\otimes \mathbb{1}_A),\mathbb{V}(X\otimes \mathbb{1}_A))+\delta(\mathbb{V}(X\otimes \mathbb{1}_{\delta(X_m,\mathbb{0})\geq k}),\mathbb{0}).\end{aligned}\tag{13}$$

Moreover, $\delta(X_m,\mathbb{0})\leq\delta(X,\mathbb{0})+\delta(X_m,X)$, then for $k'\leq k-1$, we have

$$\mathbb{1}_{\delta(X_m,\mathbb{0})\geq k}\preceq \mathbb{1}_{\delta(X,\mathbb{0})\geq k'}\oplus \mathbb{1}_{\delta(X,\mathbb{0})\leq k'}\otimes \mathbb{1}_{\delta(X_m,X)\geq 1}$$

and

$$\delta(\mathbb{V}(X\otimes \mathbb{1}_{\delta(X_m,\mathbb{0})\geq k}),\mathbb{0})\leq \max(\delta(\mathbb{V}(X\otimes \mathbb{1}_{\delta(X,\mathbb{0})\geq k'}),\mathbb{0}),\delta_0(d_{k'})\delta(\mathbb{K}(\delta(X_m,X)\geq 1),\mathbb{0})).$$

Let us fix $\varepsilon>0$. From the strong integrability of X, we can fix k' such that the first term of the right hand side is less than ε. Since $X_n \xrightarrow{\mathbb{K}} X$, the second term is less than ε for m large enough and the second term of (13) is less than ε. Then, from (11)

$$\delta(\bigoplus_{n,m\geq n_0}\mathbb{V}(X_n\otimes \mathbb{1}_{\delta(X_m,\mathbb{0})\geq k}),\mathbb{0})\leq \sup_{n,m\geq n_0}\delta(\mathbb{V}(X_n\otimes \mathbb{1}_{\delta(X_m,\mathbb{0})\geq k}),\mathbb{0})\leq 2\varepsilon$$

for n_0 large enough. Now, the strong integrability of $\oplus_{n\leq n_0} X_n$ implies that the previous inequality holds with $n_0 = 0$. This implies the equi-integrability of X_n .

Let us prove now $2b \Leftrightarrow 2c$. We have

$$\delta(\mathbb{V}(X_n \otimes \mathbb{1}_A), \mathbb{V}(X \otimes \mathbb{1}_A)) \leq \sup_{u\in A} \delta(\mathbb{V}(X_n \otimes \mathbb{1}_{\{u\}}), \mathbb{V}(X \otimes \mathbb{1}_{\{u\}}))$$

with equality for $A = \{u\}$. Then

$$\begin{aligned} \sup_{A\subset U} \delta(\mathbb{V}(X_n \otimes \mathbb{1}_A), \mathbb{V}(X \otimes \mathbb{1}_A)) &= \sup_{u\in U} \delta(X_n(u) \otimes c^*(u), X(u) \otimes c^*(u)) \\ &= \sup_{u\in U} \delta(X_n(u), X(u))\delta_0(c^*(u)) \end{aligned}$$

and $2b \Leftrightarrow 2c$. $\square$

Remark 33. If (11) is satisfied, X is strongly integrable and A_n and A are subsets of U such that $\mathbb{K}(A_n \triangle A) \longrightarrow_{n\to+\infty} \mathbb{0}$ (where $A \triangle B = A\backslash B \cup B\backslash A$), then $\mathbb{V}(X_n \otimes \mathbb{1}_{A_n}) \longrightarrow_{n\to+\infty} \mathbb{V}(X \otimes \mathbb{1}_A)$. Indeed, by (11) this is equivalent to $\mathbb{V}(X \otimes \mathbb{1}_{A_n}) \longrightarrow_{n\to+\infty} \mathbb{V}(X \otimes \mathbb{1}_A)$, then we only need to prove that $X \otimes \mathbb{1}_{A_n}$ is equi-integrable and tends towards $X \otimes \mathbb{1}_A$ in cost. But $X \otimes \mathbb{1}_{A_n} \preceq X$ which is strongly integrable, then $X \otimes \mathbb{1}_{A_n}$ is equi-integrable. In addition, $\delta(X(u) \otimes \mathbb{1}_{A_n}(u), X(u) \otimes \mathbb{1}_A(u)) \geq \varepsilon > 0$ implies $\delta(\mathbb{1}_{A_n}(u), \mathbb{1}_A(u)) \geq \varepsilon/\delta_0(X(u)) > 0$ and $u \in A_n \triangle A$. Hence,

$$\mathbb{K}(\delta(X \otimes \mathbb{1}_{A_n}, X \otimes \mathbb{1}_A) \geq \varepsilon) \preceq \mathbb{K}(A_n \triangle A) \underset{n\to+\infty}{\longrightarrow} \mathbb{0}$$

and $X \otimes \mathbb{1}_{A_n} \xrightarrow{\mathbb{K}} X \otimes \mathbb{1}_A$ for any cost variable X .

Remark 34. The formula of point 2c may define a “mean convergence”. Indeed, the quantity

$$d(X,Y) = \sup_{u\in U} \delta(X(u), Y(u))\delta_0(c^*(u))$$

corresponds to the exponential of the L^1 “idempotent distance” defined in [11] for $\mathbb{D} = \mathbb{R}_{\max}$. It can be seen as the Maslov integral of $\delta(X,Y)$ with respect to the $(\mathbb{R}^+, \max, \times)$-measure $\delta_0(\mathbb{K}(.))$ with density $\delta_0(c^*(.))$. This distance can be generalized to the case of decision variables with values in (E,d) by taking

$$d(X,Y) = \sup_{u\in U} d(X(u), Y(u))\delta_0(c^*(u)),$$

but this does not lead to properties equivalent to those of Proposition 58 for independent decision variables.

Theorem 35. *Let X_n and X denote decision variables with values in the same metric space (E,d).*

$$X_n \xrightarrow{\mathbb{K}} X \quad \textit{implies} \quad X_n \xrightarrow{w'} X.$$

If X is a constant decision variable $X \equiv a$ and $\mathbb{D}$ is connected by arcs, then the converse proposition is true.

Proof. Consider a bounded and uniformly continuous function f from (E,d) to $(\mathbb{D},\delta)$. The sequence $f(X_n)$ of cost variables is bounded, thus equi-integrable. In addition, using the uniform continuity of f, for any $\varepsilon>0$ there exists $\varepsilon'>0$ such that $x,y\in E$ and $d(x,y)<\varepsilon'$ implies $\delta(f(x),f(y))<\varepsilon$. Therefore, $\mathbb{K}(\delta(f(X_n),f(X))\geq\varepsilon)\preceq\mathbb{K}(d(X_n,X)\geq\varepsilon')\longrightarrow_{n\to+\infty}\mathbb{0}$. Then, $f(X_n)\xrightarrow{\mathbb{K}}f(X)$ and $f(X_n)$ is equi-integrable, which by the previous theorem implies $\mathbb{V}(f(X_n))\longrightarrow_n\mathbb{V}(f(X))$.

For the converse proposition, we suppose that $\mathbb{D}$ is connected by arcs, then there exists a (uniformly) continuous function from $[0,1]$ to $[\mathbb{0},\mathbb{1}]$ such that $f(0)=\mathbb{0}$ and $f(1)=\mathbb{1}$ (if f is continuous and takes its values in $\mathbb{D}$ then $f\wedge\mathbb{1}$ is continuous and takes its values in $[\mathbb{0},\mathbb{1}]$). Then $g:E\to\mathbb{D},\ x\mapsto f(\min(d(x,a)/\varepsilon,1))$ is bounded and uniformly continuous and such that $\mathbb{1}_{d(x,a)\geq\varepsilon}\preceq g(x)$. Then

$$\mathbb{K}(d(X_n,a)\geq\varepsilon)=\mathbb{V}(\mathbb{1}_{d(X_n,a)\geq\varepsilon})\preceq\mathbb{V}(g(X_n))\underset{n\to+\infty}{\longrightarrow}\mathbb{V}(g(a))=\mathbb{0}.\quad\square$$

In classical probability theory, the almost sure convergence implies the weak convergence. Here, even if the cost convergence is stronger than the almost sure convergence, it does not imply the weak convergence, but only a weakened version of this convergence. Indeed, if $X_n\xrightarrow{\text{a.s.}}X$ and f is continuous then $f(X_n)\xrightarrow{\text{a.s.}}f(X)$ but this is false for the cost convergence. In order to overcome this difficulty, we have to impose a tightness condition on the limit.

Theorem 36. *Let X_n and X denote decision variables with values in the same metric space (E,d) and suppose that X is tight. Then*

$$X_n\xrightarrow{\mathbb{K}}X\ \textit{ implies }\ X_n\xrightarrow{w}X.$$

If X is a constant decision variable $X\equiv a$ and $\mathbb{D}$ is connected by arcs, then $X_n\xrightarrow{\mathbb{K}}X\Leftrightarrow X_n\xrightarrow{w}X$.

Proof. Let $f:E\to\mathbb{D}$ be a bounded continuous function, then f is uniformly continuous on every compact subset of E and more generally, for any compact set C of E, and $\varepsilon>0$, there exists $\varepsilon'>0$ such that if x or $y\in C$ and $d(x,y)<\varepsilon'$ then $\delta(f(x),f(y))<\varepsilon$. Then,

$$\{\delta(f(X_n),f(X))\geq\varepsilon\}\subset\{d(X_n,X)\geq\varepsilon'\}\cup\{X\in C^c\}$$

which implies when $X_n\xrightarrow{\mathbb{K}}X$

$$\limsup_n\mathbb{K}(\delta(f(X_n),f(X))\geq\varepsilon)\preceq\limsup_n\mathbb{K}(d(X_n,X)\geq\varepsilon')\oplus\mathbb{K}(X\in C^c)=\mathbb{K}(X\in C^c).$$

Taking the infimum over all compact sets leads to the conclusion, if X is tight. The last assertion is a consequence of the previous theorem and the fact that a constant decision variable is always tight. $\square$

Remark 37. If the assumption "X tight" is replaced by "the sequence X_n is tight", the conclusion holds again and X is tight. Indeed, using

$$\{\delta(f(X_n),f(X))\geq\varepsilon\}\subset\{d(X_n,X)\geq\varepsilon'\}\cup\{X_n\in C^c\},$$

we obtain the conclusion of Theorem 36. If C is compact, $X \in C^c$ is equivalent to $d(X, C) > 0$, then

$$\{X \in C^c\} \subset \cup_n \{X_n \in C^c\} \cup \cup_{p \in \mathbb{N}^*} \cap_n \{d(X_n, X) > \frac{1}{p}\}$$

and

$$\mathbb{K}(X \in C^c) \preceq \bigoplus_n \mathbb{K}(X_n \in C^c) \oplus \bigoplus_{p \in \mathbb{N}^*} \bigwedge_n \mathbb{K}(d(X_n, X) > \frac{1}{p}).$$

Since X_n tends to X in cost, the last term is equal to $\mathbb{0}$. Then, the tightness of X_n implies the tightness of X .

Remark 38. Theorem 36 may be proved using Theorem 35 if the weak convergence is equivalent to weak-law convergence when the limit is tight. This may be proved in $(\mathbb{R} \cup \{-\infty\}, \max, +)$ but not in a general semiring using the techniques of [3], that we do not develop here.

Proposition 39. *1. In general, the almost sure convergence does not imply the convergence in weak-law (thus does not imply the weak or the cost convergence).*

2. In general, the weak convergence does not imply the almost sure convergence (thus does not imply the convergence in cost).

Proof. 1) Consider the semiring $\mathbb{R}_{\min}$, the idempotent probability $\mathbb{K}$ on $U = \mathbb{R}$ with density $c^*(u) = u^2$ and the sequence of cost variables $X_n(u) = -1$ in $(0, 1/n]$ and $X_n(u) = 0$ otherwise (we may even use a smooth function of the same type). $X_n(u)$ tends to 0 for any u in $\mathbb{R}$ thus $X_n \xrightarrow{\text{a.s.}} 0$. Consider the bounded (in $\mathbb{R}_{\min}$) uniformly continuous function $f(x) = \max(x, -1)$, we have $\mathbb{V}(f(X_n)) = \mathbb{V}(X_n) = \min(\inf_{0<u\leq 1/n} u^2 - 1, \inf_{u>1/n \text{ or } u\leq 0} u^2) = -1 \not\longrightarrow_{n\to+\infty} \mathbb{V}(f(0)) = 0$. Thus, X_n does not tends to 0 in weak-law.

2) For the second point, the proof is almost identical to the classical probabilistic case. Consider $U = \{0, 1\}^{\mathbb{N}}$, the idempotent probability $\mathbb{K}$ with density $c^* \equiv \mathbb{1}$ and $X_n : u \mapsto u_n$. Then X_n are independent with identical laws and thus X_n weakly converges towards $X = X_0$, whereas $X_n \overset{\text{a.s.}}{\not\longrightarrow} X$. Indeed, for $u = (0, 1, \ldots, 1, \ldots)$, we have $c^*(u) = \mathbb{1}$, then $u \in \operatorname{supp}(\mathbb{K}) = U$ and $X_n(u) = 1$ and $X(u) = 0$. $\square$

The last proposition illustrates again the fact that a.s. convergence is poor. The role played by the a.s. convergence in probability is played here (in optimization) by the cost convergence. However, as we will see in section 7 the same techniques may be used in order to prove cost convergence or classical a.s. convergence. The weak convergence notions defined in classical and idempotent probability theory are similar and are related (in the $\mathbb{R}_{\max}$ case) by large deviations. Let us finally note that the weak convergence is in some cases equivalent to the epigraph convergence [3] for which a wide literature exists [4, 5, 16].

The following equivalence table compares the results of this section with their classical probabilistic equivalent. Implications are denoted by simple arrows.

decision variables	**random variables**
$\begin{array}{ccc} & \nearrow & \text{a.s.} \\ & & \not\Updownarrow \\ \text{cost} & \longrightarrow & \text{weak-law} \\ & \underset{\substack{+\text{tight}\\ \text{limit}}}{\searrow} & \uparrow \\ & & \text{weak} \end{array}$	$\begin{array}{ccccc} \text{a.s.} & \overset{\not\leftarrow}{\longrightarrow} & \text{proba} & \overset{\not\leftarrow}{\longrightarrow} & \text{weak-law} \\ & & \updownarrow & & \\ & & \text{weak} & & \end{array}$

and if the limit is constant :

cost $\leftrightarrow$ weak-law $\leftrightarrow$ weak	proba $\leftrightarrow$ weak-law $\leftrightarrow$ weak

cost variables	**real random variables**
equi-int. $+$ cost	equi-int. $+$ proba
$\downarrow \quad \uparrow +$ assumptions on $\mathbb{D}$	$\updownarrow$
$\lim_n \mathbb{V}(X_n \otimes \mathbb{1}_A) = \mathbb{V}(X \otimes \mathbb{1}_A)$ unif. in A	$\lim_n \mathbb{E}(X_n 1_A) = \mathbb{E}(X 1_A)$ unif. in A
$\uparrow \quad \downarrow +$ assumptions on $\mathbb{D}$	$\updownarrow$
"$L^1/\delta_0(\mathbb{K})$"	L^1

6 Characteristic functions of decision variables ($\mathbb{D} = \mathbb{R}_{\max}$)

In this section and the following, we consider an $\mathbb{R}_{\max}$-idempotent probability $\mathbb{K}$ on $(U, \mathcal{U})$ with density c^*, and decision variables with values in a reflexive Banach space $(E, \|.\|)$. δ will denote the exponential distance on $\mathbb{R}_{\max}$ and E' the dual space of E.

6.1 Characteristic functions and the Cramer transform

For any decision variable X with values in E, we denote by $\mathbb{F}(X)$ the Fenchel transform of the opposite of its cost density : $\mathbb{F}(X) = \mathcal{F}(-c^*_X)$. Then, for $\theta \in E'$,

$$\mathbb{F}(X)(\theta) = \sup_{x \in E} \langle \theta, x \rangle + c^*_X(x) = \sup_{u \in U} \langle \theta, X(u) \rangle + c^*(u) = \mathbb{V}(\langle \theta, X \rangle),$$

where $\langle \theta, X \rangle$ denotes the cost variable (with values in $\mathbb{R}$) $\langle \theta, X \rangle : u \mapsto \langle \theta, X(u) \rangle$. Since $x \mapsto \langle \theta, x \rangle$ is a morphism from $(E, +)$ to the multiplicative group $(\mathbb{R}, +)$ of $\mathbb{R}_{\max}$, $\mathbb{F}(X)$ corresponds to the Laplace transform of X (or characteristic function) in classical probability theory. It will be called the characteristic function of the decision variable X. Nevertheless, there is no bijective correspondence between laws and characteristic functions, except when c^*_X is supposed to be concave.

Beyond the analogy between probability and optimization, the Cramer transform introduced in the large deviation theory [6] yields to a morphism. Let us recall its definition. If $\mathcal{B}$ is the set of Borel sets of E and P is a (classical) probability measure on $(E, \mathcal{B})$, the Cramer transform of P is defined by

$$\mathcal{C}(P) \stackrel{\text{def}}{=} \mathcal{F}'(\log(\mathcal{L}(P))),$$

where $\mathcal{L}(P)$ is the Laplace transform of P i.e. $\mathcal{L}(P)(\theta) = \int e^{\langle \theta, x \rangle} P(dx)$ for $\theta \in E'$ and $\mathcal{F}'$ is the Fenchel transform from E' to E which is the "inverse" of $\mathcal{F}$: $\mathcal{F}'(c)(x) = \sup_{\theta \in E'} \langle \theta, x \rangle - c(\theta)$. $\mathcal{C}$ can also be applied to positive measures which it transforms into l.s.c. convex functions on E. As the logarithm of the Laplace transform of any positive measure is a l.s.c. convex function, $\mathcal{C}$ is injective in the subset of positive measures such that the Laplace transform has a nonempty interior domain. Finite measures are transformed into finite idempotent $\mathbb{R}_{\min}$-measures (lower bounded l.s.c. functions) and probabilities into $\mathbb{R}_{\min}$-probabilities (functions with an infimum equal to 0).

Consider a random variable X with values in E and denote by P_X its probability law and by $\mathcal{C}(X)$ the set of decision variables X' on $(U, \mathcal{U})$ such that $\mathcal{C}(P_X) = -c^*_{X'}$, then $\log \mathcal{L}(P_X) = \mathbb{F}(X')$ for any $X' \in \mathcal{C}(X)$. Therefore, there is a correspondence between the characteristic functions of random and decision variables. In addition, any linear continuous transformation of X corresponds to the same transformation of $X' \in \mathcal{C}(X)$ ($A\mathcal{C}(X) \subset \mathcal{C}(AX)$ if A is a linear continuous operator). This is however not the case for nonlinear transformations, so that theorems cannot proceed so easily (via the Cramer transform) from probability to optimization.

Finally, if X and Y are independent random variables with values in E, then $X' \in \mathcal{C}(X)$ and $Y' \in \mathcal{C}(Y)$ are independent decision variables. Indeed, if $P_{X,Y} = P_X P_Y$ then $\mathcal{C}(P_{X,Y}) = \mathcal{C}(P_X) + \mathcal{C}(P_Y)$. Consequently $X' + Y' \in \mathcal{C}(X+Y)$. In other terms, we have for any probabilities P and P', $\mathcal{C}(P * P') = \mathcal{C}(P) \,\square\, \mathcal{C}(P')$, where $*$ denotes the convolution and $\square$ the inf-convolution operator : $c \,\square\, c'(x) = \inf_{y \in E}(c(x-y) + c'(y))$. Then, if the Cramer transform were continuous, the "idempotent law of large numbers" proved in section 7 would have been a consequence of the classical law of large numbers, at least for cost densities that are images of the Cramer transform.

6.2 Convergence of characteristic functions

In this section, we exhibit some relations between convergence in law, in weak-law ... and convergence of characteristic functions.

Proposition 40. *If X_n is uniformly bounded in E and $X_n \xrightarrow{w'} X$, then*

$$\mathbb{F}(X_n)(\theta) \underset{n \to +\infty}{\longrightarrow} \mathbb{F}(X)(\theta) \quad \forall \theta \in E'.$$

Proof. Suppose that $X_n \xrightarrow{w'} X$ and $||X_n(u)|| \leq R$ for any $u \in U$ and $n \in \mathbb{N}$ and let us fix $\theta \in E'$. Consider a continuous function ψ from $\mathbb{R}$ to $[0, 1]$ such that $\psi(x) = 1$ in $(-\infty, 1]$ and $\psi(x) = 0$ on $[2, +\infty)$ and denote $\phi = \log \psi$. Since $f : x \mapsto \langle \theta, x \rangle + \phi(||x||/R)$ is a bounded uniformly continuous function from E to $\mathbb{R}_{\max}$, $\mathbb{V}(\langle\theta, X_n\rangle) = \mathbb{V}(f(X_n)) \longrightarrow_{n\to+\infty} \mathbb{V}(f(X))$. But this is also true if f is replaced by any function f_k defined with $k \geq R$ instead of R so that $\mathbb{V}(f_k(X))$ is independent of $k \geq R$. Since $\langle\theta, x\rangle = \sup_k f_k(x)$, $\mathbb{V}(f(X)) = \sup_{k\geq R} \mathbb{V}(f_k(X)) = \mathbb{V}(\langle\theta, X\rangle)$ and the conclusion follows. □

Proposition 41. *Under the assumptions of the previous proposition, we have*

$$\mathbb{F}(X_n) \underset{n\to+\infty}{\longrightarrow} \mathbb{F}(X)$$

uniformly in any compact set of E' and

$$\mathbb{F}(X_n)(\theta_n) \underset{n\to+\infty}{\longrightarrow} \mathbb{F}(X)(\theta) \quad \forall\theta_n \underset{n\to+\infty}{\longrightarrow} \theta \in E'.$$

Proof. Let us consider the function $f_\theta(x) = \langle\theta, x\rangle + \phi(\|x\|/R)$. From the previous proposition, we have $\mathbb{F}(X_n)(\theta) = \mathbb{V}(f_\theta(X_n))$ for any n, $\mathbb{F}(X)(\theta) = \mathbb{V}(f_\theta(X))$ and $\mathbb{V}(f_\theta(X_n)) \longrightarrow_n \mathbb{V}(f_\theta(X))$ for any $\theta \in E'$. Moreover,

$$\begin{aligned}
\delta(\mathbb{V}(f_\theta(X)), \mathbb{V}(f_{\theta'}(X))) &\leq \delta_\infty(f_\theta, f_{\theta'}) \\
&\leq \sup_{\|x\|\leq 2R} |e^{\langle\theta, x\rangle} - e^{\langle\theta', x\rangle}| \\
&\leq 2R\|\theta - \theta'\| e^{2R\max(\|\theta\|, \|\theta'\|)} \qquad (14)
\end{aligned}$$

implies that $\theta \mapsto \mathbb{V}(f_\theta(X))$ is Liptschitz continuous on every bounded set of E', uniformly in X. Thus, for any $\theta \in E'$, $\delta(\mathbb{F}(X)(\theta), \mathbb{F}(X_n)(\theta')) < \varepsilon$ if n is large enough and $\|\theta - \theta'\| \leq \varepsilon'$ small enough. This immediately implies the second assertion of the proposition. The first one is obtained by extraction of a finite subcovering by balls of radius ε' of any compact set. □

Using the same technique as above, we obtain the general result :

Lemma 42. *If $X_n \xrightarrow{w'} X$ and $\theta_n \longrightarrow_{n\to+\infty} \theta$ in E' then*

$$\liminf_{n\to+\infty} \mathbb{F}(X_n)(\theta_n) \geq \mathbb{F}(X)(\theta)$$

Proof. Indeed, by the previous proof we obtain $\mathbb{V}(f_{\theta_n}(X_n)) \longrightarrow_{n\to+\infty} \mathbb{V}(f_\theta(X))$ for any R defining the functions f_θ. Since $f_\theta(x) \leq \langle\theta, x\rangle$, we have

$$\liminf_{n\to+\infty} \mathbb{F}(X_n)(\theta_n) \geq \liminf_{n\to+\infty} \mathbb{V}(f_{\theta_n}(X_n)) = \mathbb{V}(f_\theta(X))$$

and by taking the supremum over R we obtain the Lemma. □

Remark 43. This property is one of the two conditions defining the epigraph convergence of l.s.c. functions [5, 4, 16]. In [3], we have shown that the weak (or weak-law) convergence implies the epigraph convergence and is equivalent to it under the tightness condition. In addition, the Fenchel transform is continuous for the epigraph convergence (at least if E has a finite dimension) only on the set of l.s.c. proper ($\not\equiv +\infty$) convex functions. Thus, if X_n has a concave cost density and $X_n \xrightarrow{w'} X$, $\mathbb{F}(X_n)$ converges in epigraph towards $\mathbb{F}(X)$. However, this is not the case for general decision variables.

As a generalization of Proposition 40, we have

Proposition 44. *If $X_n \xrightarrow{w'} X$ and if $\theta \in E'$ is such that the sequence $\mathbb{F}(X_n)(t\theta)$ is bounded in $\mathbb{R}_{\max}$ (i.e. upper bounded in $\mathbb{R}$) for some $t > 1$, then*

$$\mathbb{F}(X_n)(\theta) \underset{n\to+\infty}{\longrightarrow} \mathbb{F}(X)(\theta).$$

Proof. The proposition will follows from Proposition 45 below if $\langle\theta, X_n\rangle$ is equi-integrable.

Let $t > 1$ be such that the sequence $\mathbb{F}(X_n)(t\theta)$ is upper bounded. Then

$$\begin{aligned}\mathbb{V}(\langle\theta, X_n\rangle + \mathbb{1}_{\langle\theta,X_n\rangle\geq R}) &= \sup_{\langle\theta,x\rangle\geq R} \langle\theta, x\rangle + c^*_{X_n}(x) \\ &\leq \mathbb{F}(X_n)(t\theta) + \sup_{\langle\theta,x\rangle\geq R}(1-t)\langle\theta, x\rangle \\ &\leq \mathbb{F}(X_n)(t\theta) + (1-t)R.\end{aligned}$$

Since $\mathbb{F}(X_n)(t\theta)$ is upper bounded and the second term tends to $-\infty$ when R goes to $+\infty$, the equi-integrability of $\langle\theta, X_n\rangle$ holds. □

Proposition 45. *Consider an application f uniformly continuous from E to $\mathbb{R}$ (not necessarily bounded or uniformly continuous into $\mathbb{R}_{\max}$). If $X_n \xrightarrow{w'} X$ and $f(X_n)$ is equi-integrable, then*

$$\mathbb{V}(f(X_n)) \underset{n\to+\infty}{\longrightarrow} \mathbb{V}(f(X)).$$

Proof. The function $\psi_R(x) : \mathbb{R} \to \mathbb{R}_{\max},\ x \mapsto \min(x, R)$ is bounded and uniformly continuous. Then $f_R = \psi_R \circ f$ is bounded and uniformly continuous and $\mathbb{V}(f_R(X_n)) \longrightarrow_{n\to+\infty} \mathbb{V}(f_R(X))$. From $f = \sup_{R>0} f_R$, we obtain $\liminf_n \mathbb{V}(f(X_n)) \geq \lim_n \mathbb{V}(f_R(X_n)) = \mathbb{V}(f_R(X))$ for any $R > 0$, then $\liminf_n \mathbb{V}(f(X_n)) \geq \mathbb{V}(f(X))$.

For the other inequality, we use $x = \max(\psi_R(x), x + \mathbb{1}_{x\geq R}(x))$, which leads to

$$\mathbb{V}(f(X_n)) = \max(\mathbb{V}(f_R(X_n)), \mathbb{V}(f(X_n) + \mathbb{1}_{f(X_n)\geq R}))$$

and

$$\limsup_n \mathbb{V}(f(X_n)) \leq \max(\mathbb{V}(f_R(X)), \sup_n \mathbb{V}(f(X_n) + \mathbb{1}_{f(X_n)\geq R})).$$

The first term is less than $\mathbb{V}(f(X))$ and the second term tends to $\mathbb{0} = -\infty$ when R goes to $+\infty$ since $f(X_n)$ is equi-integrable. Thus $\limsup_n \mathbb{V}(f(X_n)) \leq \mathbb{V}(f(X))$ and the proposition follows. □

The set of θ such that $\mathbb{F}(X_n)(\theta)$ is upper bounded is convex and contains 0. If 0 is in the interior $\mathcal{O}$ of this domain and $X_n \xrightarrow{w'} X$, then Proposition 44 implies $\mathbb{F}(X_n)(\theta) \longrightarrow_{n\to+\infty} \mathbb{F}(X)(\theta)$ for any $\theta \in \mathcal{O}$.

Proposition 46. *If X_n and X have concave law densities, $\mathbb{O}(X_n)$ and $\mathbb{O}(X)$ exist, and if there exists an open neighborhood $\mathcal{O}$ of 0 in E' such that*

$$\mathbb{F}(X_n)(\theta) \underset{n\to+\infty}{\longrightarrow} \mathbb{F}(X)(\theta) \quad \forall\theta\in\mathcal{O},$$

then

$$\mathbb{O}(X_n) \underset{n\to+\infty}{\rightharpoonup} \mathbb{O}(X) \quad \textit{weakly in } E\,.$$

Proof. By definition $\mathbb{O}(X) = \mathrm{Argmax}_{x\in E}\, c_X^*(x)$ (if it is unique), then if $\mathbb{O}(X)$ exists and c_X^* is concave, $\langle\theta, \mathbb{O}(X)\rangle$ is the derivative of $\mathbb{F}(X)$ in the direction θ at point 0. More precisely,

$$\langle\theta, \mathbb{O}(X)\rangle = \lim_{t\searrow 0^+}\!\downarrow \mathbb{F}(X)(t\theta)/t = \lim_{t\nearrow 0^-}\!\uparrow \mathbb{F}(X)(t\theta)/t.$$

Then, if $t>0$ and $\theta\in E'$ are such that $t\theta\in\mathcal{O}$, $\limsup_n \langle\theta, \mathbb{O}(X_n)\rangle \le \limsup_n \mathbb{F}(X_n)(t\theta)/t = \mathbb{F}(X)(t\theta)/t$. By taking the infimum over $t>0$ we obtain $\limsup_n \langle\theta,\mathbb{O}(X_n)\rangle \le \langle\theta,\mathbb{O}(X)\rangle$. Similarly $\liminf_n \langle\theta,\mathbb{O}(X_n)\rangle \ge \langle\theta,\mathbb{O}(X)\rangle$. Since $\mathcal{O}$ is an open neighborhood of 0, it contains a open ball centered in 0, then for any $\theta\in E'$ there exists $t>0$ such that $t'\theta$ and $-t'\theta\in\mathcal{O}$ for $0\le t'\le t$, which from the previous assertions implies $\langle\theta,\mathbb{O}(X_n)\rangle \longrightarrow_{n\to+\infty} \langle\theta,\mathbb{O}(X)\rangle$ for any $\theta\in E'$. ☐

From Propositions 46 and 44, we obtain

Corollary 47. *Let X_n and X be decision variables with concave law densities such that $\mathbb{O}(X_n)$ and $\mathbb{O}(X)$ exist. If $X_n \xrightarrow{w'} X$ and $\mathbb{F}(X_n)(\theta)$ is upper bounded for any $\theta\in\mathcal{O}$, where $\mathcal{O}$ is a open neighborhood of 0 in E', then $\mathbb{O}(X_n) \underset{n\to+\infty}{\rightharpoonup} \mathbb{O}(X)$ weakly in E'.*

Until now, we have proved under some conditions that weak-law convergence implies the convergence of characteristic functions. Let us prove a converse result in a particular case.

Proposition 48. *If $\mathbb{F}(X_n)(\theta) \underset{n\to+\infty}{\longrightarrow} 0$ for any $\theta\in E'$, uniformly in θ bounded, then $X_n \xrightarrow{\mathbb{K}} 0$.*

Proof. Since for any $x\in E$, $\|x\| = \sup_{\theta\in E',\|\theta\|\le 1}\langle\theta, x\rangle$, we have

$$\mathbb{K}(\|X_n\|\ge\varepsilon) = \sup_{\theta\in E',\|\theta\|\le 1} \mathbb{K}(\langle\theta, X_n\rangle\ge\varepsilon) \le -t\varepsilon + \sup_{\theta\in E',\|\theta\|\le t} \mathbb{F}(X_n)(\theta).$$

Then $\limsup_n \mathbb{K}(\|X_n\|\ge\varepsilon) \le -t\varepsilon$ for any $t>0$ and by taking the infimum over $t>0$ we get the proposition. ☐

Consequently, if X_n is bounded with values in a finite dimensional space E, the cost convergence of X_n towards 0 which is equivalent to the weak convergence (Theorem 36), is also equivalent to the uniform convergence in all bounded sets of E', of the characteristic function of X_n towards $0 = \mathbb{F}(0)$.

Remark 49. If X_n is a cost variable with values in $\mathbb{R}$, then the cost convergence of X_n towards 0 in $(\mathbb{R}, |.|)$ is equivalent to that in $(\mathbb{R}_{\max}, \delta)$. Indeed, $\delta(x,0)\ge\varepsilon \Leftrightarrow x\ge\ln(1+\varepsilon)$ or $x\le\ln(1-\varepsilon)$. Then $|x|\ge -\ln(1-\varepsilon) \Rightarrow \delta(x,0)\ge\varepsilon \Rightarrow |x|\ge\ln(1+\varepsilon)$.

7 Linear spaces $\mathbb{L}^p(U, \mathcal{U}, \mathbb{K}, E)$ ($\mathbb{D} = \mathbb{R}_{\max}$)

We use the notations and assumptions of the previous section and introduce "L^p-norms". The analogy of the structures $(\mathbb{R}^+, +, \times)$ and $\mathbb{R}_{\max}$ on the one hand and the Cramer transform on the other hand allow to define two different types of "L^p-norms".

Since a cost variable X is the equivalent of a real random variable and is positive, we may introduce the "L^1-norm" of X (that is the distance between X and the cost variable $\mathbb{0}$) as the integral of X : $||X||_1 = \mathbb{V}(X)$ and then the "L^p-norm" will be : $||X||_p = \mathbb{V}(X^{\otimes p})^{\otimes \frac{1}{p}}$ where $a^{\otimes p}$ denotes the power p of a in the $\mathbb{R}_{\max}$ semiring, that is $a^{\otimes p} = p \times a$. Then, $||X||_p = \frac{1}{p} \times \mathbb{V}(p \times X) = \frac{1}{p} \sup_{x \in \mathbb{R}} px + c^*_X(x) = \sup_{x \in \mathbb{R}} x + \frac{1}{p} c^*_X(x)$. A "$L^p$ distance" related to this L^p-norm has been introduced in [11]. It yields to an idempotent distance (up to the logarithm) on the $\mathbb{R}_{\max}$-semimodule of cost variables. Relations between this L^p convergence and the other types of convergence are proved in [11]. Moreover, the L^2-norm defined in this way is in relation with the "scalar product" of Maslov [19] or Samborski [22]. However, it does not lead to a good notion to prove the law of large numbers [23]. For this purpose, we introduce another "L^p-norm" related to the classical one by the Cramer transform.

Consider a real random variable $X \in L^2(\Omega, \mathcal{A}, P)$. Its L^2-norm may be calculated by differentiating the logarithm of its Laplace transform $\mathcal{L}(X)$:

$$\begin{aligned} ||X||_2^2 &= \operatorname{Var}(X) + (\mathbb{E}(X))^2 \\ \operatorname{Var}(X) &= \left.\frac{\partial^2 \log \mathcal{L}(X)}{\partial \theta^2}\right|_{\theta=0} \\ \mathbb{E}(X) &= \left.\frac{\partial \log \mathcal{L}(X)}{\partial \theta}\right|_{\theta=0} \end{aligned}$$

Using the relations between random and decision variables described in section 6.1, we may state for a decision variable X with values in $\mathbb{R}$:

$$||X||_2^2 = \left.\frac{\partial^2 \mathbb{F}(X)}{\partial \theta^2}\right|_{\theta=0} + \left(\left.\frac{\partial \mathbb{F}(X)}{\partial \theta}\right|_{\theta=0} \right)^2 \tag{15}$$

Unfortunately, (15) does not proceed to a norm on the linear space (in the usual sense) of decision variables. For instance, if $c^*_X(x) = -\frac{3}{4}|x|^{\frac{4}{3}}$ then $X \overset{\text{a.s.}}{\neq} 0$ but $\mathbb{F}(X)(\theta) = \frac{1}{4}|\theta|^4$ and $||X||_2 = 0$. One of the reasons for which this could not occur in Probability is that even if $\frac{3}{4}|x|^{\frac{4}{3}}$ is a l.s.c. convex function, it is not the image of a probability by the Cramer transform. The space of l.s.c. convex functions is much larger than the space of probabilities and the L^2 norm we have to construct has to be more selective. Let us finally note that (15) is equivalent to $||X||_2^2 = \sigma^2 + (\mathbb{O}(X))^2$ where, if c^*_X is an u.s.c. concave function, $\mathbb{O}(X) = \operatorname{Argmax}_{x \in \mathbb{R}} c^*_X$ and

$$\mathbb{F}(X)(\theta) = \langle \theta, \mathbb{O}(X) \rangle + \frac{1}{2}(\sigma\theta)^2 + o(\theta^2)$$

which is equivalent to

$$c_X^*(x) = -\frac{1}{2}(\frac{x-\mathbb{O}(X)}{\sigma})^2 + o((x-\mathbb{O}(X))^2). \tag{16}$$

Then, $||X||_2 = 0$ implies $\mathbb{O}(X) = 0$ and $\sigma = 0$, but it may just mean that c_X^* is not in the order of x^2 in 0 but only of x^p with $1 < p < 2$. However, if we impose the condition

$$c_X^*(x) \leq -\frac{1}{2}(\frac{x-\mathbb{O}(X)}{\sigma})^2 \tag{17}$$

on the entire space $\mathbb{R}$, then $\mathbb{O}(X) = 0$ and $\sigma = 0$ imply $c_X^*(x) = -\infty$ except in 0, that is $X \overset{\text{a.s.}}{=} 0$. The following result, already proved in [3] for $E = \mathbb{R}$, shows that this leads to a norm on the linear space L^2 of decision variables such that (17) holds.

Proposition 50. *Let $(U, \mathcal{U}, \mathbb{K})$ be a decision space over $\mathbb{R}_{\max}$ (or by a change of signs $\mathbb{R}_{\min}$) such that $\mathbb{K}$ has a density and let $(E, ||.||)$ be a Banach space. Then, the numbers*

$$\begin{aligned} |X|_p &\overset{def}{=} \inf\{\sigma, c_X^*(x) \leq -\frac{1}{p}(\frac{||x-\mathbb{O}(X)||}{\sigma})^p\}, \qquad (18)\\ ||X||_p &\overset{def}{=} |X|_p + ||\mathbb{O}(X)|| \end{aligned}$$

define, for $p > 0$, respectively a seminorm and a norm in the linear space $\mathbb{L}^p(U, \mathcal{U}, \mathbb{K}, E)$ of classes (for the $\overset{\text{a.s.}}{=}$ relation) of decision variables with values in E such that the optimum $\mathbb{O}(X)$ is unique and $|X|_p$ is finite.

Moreover, $\mathbb{O}$ is a linear continuous operator from $\mathbb{L}^p(U, \mathcal{U}, \mathbb{K}, E)$ to E .

Between the local condition (16) and the global condition (17) an intermediary condition may be useful. For instance if (17) is satisfied only in the ball $B(\mathbb{O}(X), \varepsilon)$ and c_X^* is concave then $c_X^* \leq -\phi(\frac{1}{\sigma}(x - \mathbb{O}(X)))$ where $\phi(x) = \frac{x^2}{2}$ for $|x| \leq \varepsilon' = \frac{\varepsilon}{\sigma}$ and $\phi(x) = \varepsilon'|x| - \frac{\varepsilon'^2}{2}$ otherwise. If ε is fixed, this condition does not lead to a norm for the same reason as (16). However if ε' is fixed, that is ϕ is fixed, this leads to a norm.

Proposition 51. *Consider a symmetric quasiconvex function $\phi : E \to \mathbb{R}^+$, that is $\phi(-x) = \phi(x)$ and*

$$\phi((1-t)x + ty) \leq \max(\phi(x), \phi(y)) \quad \forall x, y \in E,\ t \in [0, 1], \tag{19}$$

such that :

$$\phi(x) \to +\infty \quad \textit{when} \quad ||x|| \to +\infty, \tag{20}$$

$$\phi(x) \to 0 \quad \Leftrightarrow \quad x \to 0. \tag{21}$$

Then, the numbers

$$\begin{aligned} |X|_\phi &\overset{def}{=} \inf\{\sigma, c_X^*(x) \leq -\phi(\frac{1}{\sigma}(x - \mathbb{O}(X)))\}\\ ||X||_\phi &\overset{def}{=} |X|_\phi + ||\mathbb{O}(X)||, \end{aligned}$$

define respectively a seminorm and a norm in the linear space $\mathbb{L}^{\phi}(U,\mathcal{U},\mathbb{K},E)$ of decision variables with values in E such that the optimum $\mathbb{O}(X)$ is unique and $|X|_{\phi}$ is finite.

Moreover, $\mathbb{O}$ is a linear continuous operator from $\mathbb{L}^{\phi}(U,\mathcal{U},\mathbb{K},E)$ to E.

Proof. Let us first note that from (19) and (21), $\phi(0) = 0$ and $t \in \mathbb{R}^{+} \mapsto \phi(tx)$ is nondecreasing. Then

$$\begin{aligned}
\sigma > |X|_{\phi} \quad &\Rightarrow \quad c_X^*(x) \leq -\phi(\frac{1}{\sigma}(x - \mathbb{O}(X))) \quad \forall x \in E \quad (\Rightarrow \sigma \geq |X|_{\phi}) \\
&\Leftrightarrow \quad \mathbb{V}(\phi(\frac{1}{\sigma}(X - \mathbb{O}(X))) \leq 0 = \mathbb{1} \\
&\Leftrightarrow \quad \mathbb{V}(\phi(\frac{1}{\sigma}(X - \mathbb{O}(X))) = 0
\end{aligned} \tag{22}$$

If there exists $\sigma > 0$ and $\mathbb{O}(X) \in E$ such that (22) holds, then by (21), $c_X^*(x) < 0$ for any $x \neq \mathbb{O}(X)$ and $c_X^*(x)$ is far from 0 if x is far from $\mathbb{O}(X)$, hence $\mathbb{O}(X)$ is the unique optimum of X. This implies that $X \in \mathbb{L}^{\phi}(U,\mathcal{U},\mathbb{K},E)$ if and only if there exists $\sigma > 0$ and $\mathbb{O}(X) \in E$ such that (22) hold. Moreover, $|X|_{\phi}$ is the smallest σ such that (22) holds.

If $X \in \mathbb{L}^{\phi}(U,\mathcal{U},\mathbb{K},E)$, $\lambda \in \mathbb{R}$ and $\sigma > |X|_{\phi}$, then by symmetry of ϕ

$$\mathbb{V}(\phi(\frac{1}{|\lambda|\sigma}(\lambda X - \lambda\mathbb{O}(X))) = \mathbb{V}(\phi(\frac{1}{\sigma}(X - \mathbb{O}(X))) \leq 0,$$

therefore $\lambda X \in \mathbb{L}^{\phi}(U,\mathcal{U},\mathbb{K},E)$, $\mathbb{O}(\lambda X) = \lambda\mathbb{O}(X)$ and $|\lambda X|_{\phi} = |\lambda||X|_{\phi}$.

If X and $Y \in \mathbb{L}^{\phi}(U,\mathcal{U},\mathbb{K},E)$, $\sigma > |X|_{\phi}$ and $\sigma' > |Y|_{\phi}$, $X' = \frac{1}{\sigma}(X - \mathbb{O}(X))$, $Y' = \frac{1}{\sigma'}(Y - \mathbb{O}(Y))$, then

$$\mathbb{V}(\max(\phi(X'), \phi(Y')) = \max(\mathbb{V}(\phi(X')), \mathbb{V}(\phi(Y'))) \leq 0$$

and by (19) (with $t = \frac{\sigma'}{\sigma+\sigma'}$)

$$\phi(\frac{1}{\sigma+\sigma'}((X+Y)(u) - \mathbb{O}(X) - \mathbb{O}(Y))) \ \leq \ \max(\phi(X'(u)), \phi(Y'(u))),$$

hence

$$\mathbb{V}(\phi(\frac{1}{\sigma+\sigma'}(X + Y - \mathbb{O}(X) - \mathbb{O}(Y)))) \leq 0$$

and $X + Y \in \mathbb{L}^{\phi}(U,\mathcal{U},\mathbb{K},E)$, $\mathbb{O}(X+Y) = \mathbb{O}(X) + \mathbb{O}(Y)$ and $|X+Y|_{\phi} \leq |X|_{\phi} + |Y|_{\phi}$.

Therefore $\mathbb{L}^{\phi}(U,\mathcal{U},\mathbb{K},E)$ is a linear space, $|.|_{\phi}$ and $||.||_{\phi}$ are seminorms and $\mathbb{O}$ is a linear operator from $\mathbb{L}^{\phi}$ to E. Moreover, $||X||_{\phi} = 0$ and (20) imply $c_X^* = \mathbb{1}_0$ thus $X \overset{\text{a.s.}}{=} 0$. Then $||.||_{\phi}$ is a norm and $\mathbb{O}$ is trivially continuous. $\square$

Remark 52. The previous result may be generalized to a semifield $\mathbb{D}$ satisfying the conditions of section 1. Let $\phi : E \to [\mathbb{1}, \oplus\mathbb{D}]$ be a symmetrical function such that

$$\begin{aligned}
&\phi((1-t)x + ty) \preceq \phi(x) \oplus \phi(y) \quad \forall x, y \in E,\ t \in [0,1], \\
&\forall a \in \mathbb{D}\ \exists A > 0,\ \text{s.t. } \phi(x) \succeq a \quad \text{when} \quad ||x|| \geq A\,, \\
&\phi(x) = \mathbb{1} \quad \Leftrightarrow \quad x = 0.
\end{aligned}$$

If there exist $O \in E$ and $\sigma > 0$ such that $\mathbb{V}(\phi(\frac{1}{\sigma}(X - O))) \preceq \mathbb{1}$, then O is unique and independent of σ. Indeed if O' and σ' also satisfy $\mathbb{V}(\phi(\frac{1}{\sigma'}(X - O'))) \preceq \mathbb{1}$, then using the same technique as in the proof of the subadditivity of $|.|_\phi$ and using the symmetry of ϕ, we obtain $\phi(\frac{O-O'}{\sigma+\sigma'}) \preceq \mathbb{1}$ and then $O = O'$. Note that in the proof of Proposition 51, the bicontinuity of ϕ in 0 was only required in order to prove that the element O is equal to the optimum $\mathbb{O}(X)$. Let us still denote by $\mathbb{O}(X)$ this unique element if it exists. The numbers

$$|X|_\phi \stackrel{\text{def}}{=} \inf\{\sigma, \mathbb{V}(\phi(\frac{1}{\sigma}(X - \mathbb{O}(X)))) \preceq \mathbb{1}\}$$
$$||X||_\phi \stackrel{\text{def}}{=} |X|_\phi + ||\mathbb{O}(X)||,$$

define respectively a seminorm and a norm in the linear space $\mathbb{L}^\phi(U, \mathcal{U}, \mathbb{K}, E)$ of decision variables with values in E such that $\mathbb{O}(X)$ exists and $|X|_\phi$ is finite and $\mathbb{O}$ is a linear continuous operator from $\mathbb{L}^\phi(U, \mathcal{U}, \mathbb{K}, E)$ to E. The proof is identical to that of Proposition 51. The existence of inverses for the $\otimes$ law is required to prove that $||.||_\phi$ is a norm.

Example 53. The conditions of the previous proposition are satisfied by the composition $\phi = f \circ \psi$ of a nondecreasing function f from $\mathbb{R}^+$ to $\mathbb{R}^+$ and a positive convex function ψ such that both f and ψ satisfy conditions (20,21). For instance, if we take $\psi(x) = ||x||$ and $f(x) = \frac{1}{p}x^p$ this leads to the Proposition 50. If we take now $f(x) = \frac{1}{p}x^p$ for $x \leq \varepsilon$ and $f(x) = \frac{1}{p}\varepsilon^p + \varepsilon^{p-1}(x - \varepsilon)$ for $x \geq \varepsilon$, then a decision variable X with concave cost density c^*_X belongs to $\mathbb{L}^\phi$ iff a condition of type (17)is satisfied on the ball $B(\mathbb{O}(X), \varepsilon\sigma)$ only. These functions f define other $\mathbb{L}^p$ spaces that we denote $\mathbb{L}^p_\varepsilon$ and which contrary to preceding ones satisfy the inclusion properties $\mathbb{L}^p_\varepsilon \subset \mathbb{L}^q_\varepsilon$ when $p < q$ and $\varepsilon \leq 1$.

Remark 54. In [2], we introduced the sensitivity of order $p > 1$ of real decision variables ($E = \mathbb{R}$) which is, for $p = 2$, the equivalent of the standard deviation in Probability. As it was already pointed out, this cannot be used for the definition of L^p-norms. However, we may define in $E = \mathbb{R}^n$ the variance-covariance matrix as the second derivative of $\mathbb{F}(X)$: $\mathbb{F}(X)(\theta) = \langle \theta, \mathbb{O}(X) \rangle + \frac{1}{2}\langle \mathrm{Var}(X)\theta, \theta \rangle + o(||\theta||^2)$. If c^*_X is concave, it may also be defined as the non singular matrix $\mathrm{Var}(X)$ (if it exists) such that $c^*_X(x + \mathbb{O}(X)) = \frac{1}{2}\langle \mathrm{Var}(X)^{-1}x, x \rangle + o(||x||^2)$. However this may not be generalized for $p \neq 2$.

Remark 55. If we consider the limit of the condition (18) when p tends to infinity, we find $||X - \mathbb{O}(X)|| \leq \sigma$ in the support of $\mathbb{K}$, i.e. almost surely. Thus the "limit" of $\mathbb{L}^p$ space is the space of almost surely bounded decision variables such that $\mathbb{O}(X)$ exists. Unfortunately the condition "$\mathbb{O}(X)$ exists" is in general not linear. Then, the good $\mathbb{L}^\infty$ space is the linear space of almost surely bounded decision variables endowed with the norm $||X||_\infty = \sup_{u \in \mathrm{supp}(\mathbb{K})} ||X(u)||$, in which X may not have an optimum.

Let us give a dual characterization of the L^p norm.

Proposition 56. *Let $||.||$ denotes the dual norm on E', $||\theta|| = \sup_{||x|| \leq 1} \langle \theta, x \rangle$, $p > 1$ and $1/p + 1/p' = 1$. Then, the seminorm $|.|_{p,\varepsilon}$ of $\mathbb{L}^p_\varepsilon(U, \mathcal{U}, \mathbb{K}, E)$ satisfies*

$$|X|_{p,\varepsilon} = \inf\{\sigma, \mathbb{F}(X)(\theta) \leq \langle \theta, \mathbb{O}(X) \rangle + \frac{1}{p'}||\sigma\theta||^{p'} \text{ for } ||\sigma\theta|| \leq \varepsilon^{p-1}\}.$$

More generally, let ϕ be a l.s.c. convex function satisfying the condition of Proposition 51, with Fenchel transform $\mathcal{F}(\phi)$. Then, the seminorm of $\mathbb{L}^{\phi}(U,\mathcal{U},\mathbb{K},E)$ satisfies

$$|X|_{\phi} = \inf\{\sigma, \mathbb{F}(X)(\theta) \leq \langle\theta, \mathbb{O}(X)\rangle + \mathcal{F}(\phi)(\sigma\theta)\ \forall\theta\in E'\}.$$

Proof. The last characterization is evident since the function c_X^* is less than an u.s.c. concave function ψ iff its u.s.c. concave envelope is less than ψ . Now, if $\phi(x) = f(||x||)$ with $f : \mathbb{R}^+ \to \mathbb{R}^+$ a l.s.c. convex function, then $\mathcal{F}(\phi)(\theta) = \mathcal{F}(f)(||\theta||)$ and for the function f defining $\mathbb{L}^p_\varepsilon$: $f(x) = \frac{1}{p}x^p$ for $x \leq \varepsilon$, $f(x) = \frac{1}{p}\varepsilon^p + \varepsilon^{p-1}(x-\varepsilon)$ for $x \geq \varepsilon$, we have $\mathcal{F}(f)(\theta) = \frac{1}{p'}\theta^{p'} + \chi_{\theta\leq\varepsilon^{p-1}}$, where χ_A denotes the characteristic function of the set A in $\mathbb{R}_{\min}$. □

Remark 57. By this characterization, we see that L^p norms of decision variables (for $p > 1$) are related by the Cramer transform to $L^{p'}$ norms of random variables. However, for $p < 1$, the L^p norms of decision variables do not correspond to any norm of random variables. Indeed, if c_X^* is a concave function in L^p with $p < 1$ then X is constant. Let us first recall that the Cramer transform of the Gaussian law, on $\mathbb{R}$ for instance, of center m and standard deviation σ is the quadratic function $\mathcal{M}^2_{m,\sigma}(x) = \frac{(x-m)^2}{2\sigma^2}$, such that $\mathbb{O}(\mathcal{M}^2_{m,\sigma}) = m$ and $|\mathcal{M}^2_{m,\sigma}|_2 = \sigma$. Moreover, if X is a random variable on $\mathbb{R}$ with expectation m and standard deviation σ , then $\log\mathcal{L}(X)(\theta) = m\theta + \frac{1}{2}(\sigma\theta)^2 + o(\theta^2)$ and thus $|X'|_2 \geq \sigma$ for any $X' \in \mathcal{C}(X)$ (see section 6.1 for the definition of $\mathcal{C}(X)$). If X has no second moment but has a moment of order $1 < p' < 2$, then the fractional derivative of order p' of $\log\mathcal{L}(X)$ exists in 0 and is related to the $L^{p'}$ norm of $X - m$. Moreover, up to some constant factor, it is smaller than $|X'|_p^{p'}$ for any $X' \in \mathcal{C}(X)$.

Proposition 58. *If X and $Y \in \mathbb{L}^p(U,\mathcal{U},\mathbb{K},E)$ are independent decision variables, then*

$$\begin{aligned} |X+Y|_p &\leq (|X|_p^{p'} + |Y|_p^{p'})^{\frac{1}{p'}} \quad \textit{for } p>1 \textit{ and } \frac{1}{p}+\frac{1}{p'}=1 \\ |X+Y|_p &\leq \max(|X|_p, |Y|_p) \quad \textit{for } p\leq 1. \end{aligned} \tag{23}$$

The inequality (23) is also true for the seminorm $|.|_{p,\varepsilon}$ of $\mathbb{L}^p_\varepsilon(U,\mathcal{U},\mathbb{K},E)$ with $\varepsilon > 0$.

Proof. Consider X and Y in $\mathbb{L}^p(U,\mathcal{U},\mathbb{K},E)$, $\sigma > |X|_p$ and $\sigma' > |Y|_p$. Let $\sigma'' = (\sigma^{p'} + (\sigma')^{p'})^{\frac{1}{p'}}$ if $p > 1$ and $\sigma'' = \max(\sigma,\sigma')$ if $p \leq 1$. By the Hölder inequality for $p > 1$ and the inequality $(x+y)^p \leq x^p + y^p$ for $x, y \geq 0$ and $p \leq 1$, we have

$$\begin{aligned} \frac{||(X+Y)(u) - \mathbb{O}(X) - \mathbb{O}(Y)||}{\sigma''} &\leq \frac{\sigma}{\sigma''}\frac{||X(u)-\mathbb{O}(X)||}{\sigma} + \frac{\sigma'}{\sigma''}\frac{||Y(u)-\mathbb{O}(y)||}{\sigma'} \\ &\leq ((\frac{||X(u)-\mathbb{O}(X)||}{\sigma})^p + (\frac{||Y(u)-\mathbb{O}(y)||}{\sigma'})^p)^{\frac{1}{p}}, \end{aligned}$$

then by the independence of X and Y we obtain

$$\mathbb{V}(\frac{1}{p}(\frac{||(X+Y)-\mathbb{O}(X)-\mathbb{O}(Y)||}{\sigma''})^p) \leq 0.$$

Another proof of (23) when $p > 1$ consists in using the characterization of Proposition 56. If X and $Y \in \mathbb{L}^\phi$, $\sigma > |X|_\phi$, $\sigma' > |Y|_\phi$ and X and Y are independent, we have

$$\mathbb{F}(X+Y)(\theta) = \mathbb{F}(X)(\theta) + \mathbb{F}(Y)(\theta) \leq \langle \theta, \mathbb{O}(X) + \mathbb{O}(Y) \rangle + \mathcal{F}(\phi)(\sigma''\theta)$$

and $\sigma'' \geq |X+Y|_\phi$ if $\mathcal{F}(\phi)(\sigma\theta) + \mathcal{F}(\phi)(\sigma'\theta) \leq \mathcal{F}(\phi)(\sigma''\theta)$ for any $\theta \in E'$. This is the case if $\mathcal{F}(\phi)(\theta) = \frac{1}{p'}\|\theta\|^{p'} + \chi_{\|\theta\| \leq \varepsilon^{p-1}}$ and $\sigma^{p'} + \sigma'^{p'} = \sigma''^{p'}$ ($0 < \varepsilon \leq +\infty$). Then, (23) holds for the seminorm $|.|_{p,\varepsilon}$. □

Using the same proof, we can generalize the previous result as follows.

Proposition 59. *Let ϕ be a l.s.c. convex function satisfying the condition of Proposition 51 and $\diamond$ be a commutative and associative law on $\mathbb{R}^+$ such that $x \mapsto x \diamond y$ is nondecreasing and continuous and*

$$\mathcal{F}(\phi)(\sigma\theta) + \mathcal{F}(\phi)(\sigma'\theta) \leq \mathcal{F}(\phi)((\sigma \diamond \sigma')\theta) \quad \forall \sigma, \sigma' \in \mathbb{R}^+,\ \theta \in E'. \tag{24}$$

Then for any independent decision variables X and $Y \in \mathbb{L}^\phi(U, \mathcal{U}, \mathbb{K}, E)$, we have

$$|X+Y|_\phi \ \leq\ |X|_\phi \diamond |Y|_\phi.$$

If $\phi(x) = f(\|x\|)$ with $f : \mathbb{R}^+ \to \mathbb{R}^+$ a l.s.c. convex function, then (24) is satisfied if $\times$ is distributive with respect to $\diamond$ and

$$\mathcal{F}(f)(\sigma) + \mathcal{F}(f)(\sigma') \leq \mathcal{F}(f)(\sigma \diamond \sigma') \quad \forall \sigma, \sigma' \in \mathbb{R}^+.$$

Let us now compare the "L^p-convergence" with the convergence notions of the previous sections.

Definition 60. We say that X_n converges in L^ϕ-norm towards X, denoted $X_n \xrightarrow{\mathbb{L}^\phi} X$, if X_n and $X \in \mathbb{L}^\phi(U, \mathcal{U}, \mathbb{K}, E)$ and $\|X_n - X\|_\phi \underset{n\to+\infty}{\longrightarrow} 0$.

Theorem 61. *If X_n and $X \in \mathbb{L}^\phi(U, \mathcal{U}, \mathbb{K}, E)$ (resp. $\mathbb{L}^\infty$) and $X_n \xrightarrow{\mathbb{L}^\phi} X$ (resp. $X_n \xrightarrow{\mathbb{L}^\infty} X$), then $X_n \xrightarrow{\mathbb{K}} X$. Conversely, the cost convergence does not imply the convergence in $\mathbb{L}^p$ or $\mathbb{L}^p_\varepsilon$-norm.*

Proof. Since the cost convergence of X_n towards X is equivalent to that of $X_n - X$ towards 0, we restrict ourselves to the case $X = 0$. Suppose that $X_n \xrightarrow{\mathbb{L}^\phi} 0$. We have $\mathbb{O}(X_n) \longrightarrow_{n\to+\infty} 0$ (in E) and $|X_n|_\phi \longrightarrow_{n\to+\infty} 0$, which implies that for any $\sigma > 0$ $\mathbb{V}(\phi(\frac{1}{\sigma}(X_n - \mathbb{O}(X_n)))) \leq 0$ for n large enough. Let us fix $\varepsilon > 0$. For n large enough, $\|\mathbb{O}(X_n)\| \leq \varepsilon/2$, then

$$\begin{aligned}
\mathbb{K}(\|X_n\| \geq \varepsilon) \quad &\leq \quad \mathbb{K}(\|X_n - \mathbb{O}(X_n)\| \geq \varepsilon/2) \\
&\leq \quad -\inf_{\|x\| \geq \frac{\varepsilon}{2\sigma}} \phi(x) + \mathbb{V}(\phi(\frac{1}{\sigma}(X_n - \mathbb{O}(X_n)))) \\
&\leq \quad -\inf_{\|x\| \geq \frac{\varepsilon}{2\sigma}} \phi(x).
\end{aligned}$$

Then, by property (20) of ϕ, $\mathbb{K}(\|X_n\| \geq \varepsilon) \longrightarrow_{n\to+\infty} -\infty = \mathbb{0}$.

If now $X_n \xrightarrow{\mathbb{L}^\infty} 0$, then for any $\sigma > 0$, $\mathbb{K}(||X_n|| > \sigma) = \mathbb{0}$ for n large enough and $X_n \xrightarrow{\mathbb{K}} 0$.

We prove now that the converse proposition is false for $\mathbb{L}^p$ with $p < +\infty$. Let us consider independent real ($E = \mathbb{R}$) decision variables with cost densities

$$\begin{aligned} c^*_{X_n}(x) &= -n|x|^p + n^{1-p} - n^{-1-p} \quad \text{for } |x| \geq 1/n \\ &= -|x|^p/n \quad \text{otherwise}. \end{aligned}$$

We get $\mathbb{K}(|X_n| \geq \varepsilon) = -n\varepsilon^p + n^{1-p} - n^{-1-p}$ if $1/n \leq \varepsilon$, then $\mathbb{K}(|X_n| \geq \varepsilon) \longrightarrow_{n\to+\infty} -\infty$ and $X_n \xrightarrow{\mathbb{K}} 0$. In addition, $\mathbb{O}(X_n) = 0$ and $|X_n|_p = \inf\{\sigma, c^*_{X_n}(x) \leq -\frac{1}{p}(\frac{|x|}{\sigma})^p\}$ then

$$(|X_n|_p)^p = \sup_x \frac{|x|^p}{-pc^*_{X_n}(x)} = \max(\sup_{0\leq x\leq 1/n} \frac{x^p}{\frac{p}{n}x^p}, \sup_{x\geq 1/n} \frac{x^p}{p(nx^p - n^{1-p} + n^{-1-p})}) = \frac{n}{p},$$

then $X_n \in \mathbb{L}^p$ and $||X_n||_p \longrightarrow_{n\to+\infty} +\infty$. Moreover, since $||X_n||_{p,\varepsilon} = ||X_n||_p$, the result is the same for L^p_ε .

Now for $\mathbb{L}^\infty$, we can consider

$$\begin{aligned} c^*_{X_n}(x) &= -\log(n\frac{|x|}{1-|x|}) \quad \text{for } 1/n \leq |x| < 1 \\ &= +\infty \quad \text{for } |x| \geq 1 \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

Then $||X_n||_\infty = 1$, $\mathbb{K}(|X_n| \geq \varepsilon) = -\log(n\varepsilon)$ if $1 > \varepsilon \geq 1/n$ and $\mathbb{K}(|X_n| \geq \varepsilon) \longrightarrow_{n\to+\infty} -\infty = \mathbb{0}$. Therefore $X_n \xrightarrow{\mathbb{K}} 0$ but X_n does not tend to 0 in L^∞-norm. □

Remark 62. In order to prove the L^p convergence from the cost convergence, we may need, as in classical probability theory, a new equi-integrability condition associated to our L^p-norms, distinct from that of section 4. For instance, we may define

$$|X|_{\phi,A} = \inf\{\sigma, \mathbb{V}(\phi(\frac{1}{\sigma}(X - \mathbb{O}(X))) \otimes \mathbb{1}_A(X)) \leq \mathbb{1}\}$$

and $||X||_{\phi,A} = |X|_{\phi,A} + ||\mathbb{O}(X)||$. However, the conditions $||X_n||_{\phi,\{||X_n||\geq k\}} \longrightarrow_{k\to+\infty} 0$ uniformly in n (equi-integrability with respect to our L^ϕ norm) and $X_n \xrightarrow{\mathbb{K}} X$ do not imply $X_n \xrightarrow{\mathbb{L}^\phi} X$. Indeed, if in the previous counter example $c^*_{X_n}(x)$ is replaced by $+\infty$ on $|x| > 1$, then $||X_n||_\infty = 1$, X_n is equi-integrable and the other properties are identical.

Let us summarize the additional convergence relations proved in the last two sections. In the following table, implications are denoted by simple arrows.

decision variables	**random variables**
$$\begin{tikzcd} \mathbb{L}^p \arrow[d, shift left] \\ \text{cost} \arrow[u, shift left] \arrow[d] & \\ \text{weak-law} \arrow[r, "\text{+unif. and const. limit}", "\text{+bounded}"'] & \text{Fenchel cv.} \arrow[ul] \end{tikzcd}$$	$$\begin{tikzcd} L^p \arrow[d, shift left] \\ \text{proba} \arrow[u, shift left, "+\text{ equi-int.}"'] \arrow[d] & \\ \text{weak-law} \arrow[r, "\text{+bounded}"'] \arrow[dr] & \text{Laplace cv.} \\ & \text{Fourier cv.} \end{tikzcd}$$

As an application of the previous convergence relations, we obtain the law of large numbers. Thus, the L^p-convergence appears to be the good notion for proving the idempotent equivalent of the classical theorems of probability.

Theorem 63 (Law of Large numbers (Quadrat [23, 2])). *Let X_n denote a sequence of independent decision variables with values in E and with identical cost density c^*_X. Suppose that $X_n \in \mathbb{L}^p(U, \mathcal{U}, \mathbb{K}, E)$ with $0 < p < +\infty$ or $\mathbb{L}^p_\varepsilon(U, \mathcal{U}, \mathbb{K}, E)$ with $p > 1$ and $\varepsilon > 0$, then*

$$S_N = \frac{1}{N} \sum_{n=0}^{N-1} X_n \xrightarrow{\mathbb{L}^p} \mathbb{O}(X),$$

which implies the cost, almost sure and weak convergence.

If we only suppose that $\mathbb{F}(X) = \mathcal{F}(c^*_X)$ is differentiable in 0 with $\langle d\mathbb{F}(X), \theta \rangle = \langle \theta, \mathbb{O}(X) \rangle$, then

$$S_N = \frac{1}{N} \sum_{n=0}^{N-1} X_n \underset{n \to +\infty}{\longrightarrow} \mathbb{O}(X),$$

where the convergence hold in cost, almost surely and weakly.

Proof. For the first assertion, we have $\mathbb{O}(S_N - \mathbb{O}(X)) = 0$, so that we can consider the case $\mathbb{O}(X) = 0$ only. But by Proposition 58, we have $|S_N|_p \leq |X_0|_p / N^{1/p}$ for $p > 1$ and $|S_N|_p \leq |X_0|_p / N$ for $p \leq 1$. Then, $|S_N|_p \longrightarrow_{n \to +\infty} 0$ in any case and by Theorem 61, $S_N \xrightarrow{\mathbb{K}} 0$. For the same reasons, this result holds for the L^p_ε-norm with $p > 1$ and $\varepsilon > 0$.

For the second assertion, we use $\mathbb{F}(S_N - \mathbb{O}(X))(\theta) = N\mathbb{F}(X_0)(\frac{\theta}{N}) - \mathbb{O}(X) \longrightarrow_{N \to +\infty} 0$ uniformly in θ bounded, then using Proposition 48 we obtain $S_N \xrightarrow{\mathbb{K}} \mathbb{O}(X)$. Then, the results of section 5 imply that the convergence also holds almost surely and weakly. □

The previous result may be translated in terms of an optimal control problem as follows. Let us consider the problem in $\mathbb{R}_{\min}$ instead of $\mathbb{R}_{\max}$ and write c instead of c_X^*. Then,

$$\mathbb{K}(S_N = s) = \inf_{x_0+\cdots+x_{N-1}=Ns} \sum_{n=0}^{N-1} c(x_n) = \inf_{y(.),\, y(0)=0,\, y(N)=s} \sum_{n=0}^{N-1} c(N(y(n+1)-y(n))).$$

If we consider now the optimal control problem with final cost f :

$$v_N(x) = \inf_{y(.),\, y(0)=x} \sum_{n=0}^{N-1} c(N(y(n+1)-y(n))) + f(y(N)),$$

we obtain that $v_N(x) = \mathbb{V}(f(x+S_N))$ and the weak convergence of S_N towards $\mathbb{O}(X)$ means that for any lower bounded continuous function $f : E \to \mathbb{R}_{\min}$ we have $v_N(x) \longrightarrow_{N\to+\infty} f(x+\mathbb{O}(X))$. This result may have been proved directly. Indeed, consider $N(y(n+1)-y(n))$ as the discretization (with step $1/N$) of the derivative of some trajectory $x(t)$ at time $t = n/N \in [0,1]$. The first term of the infimum defining v_N is in the order of $N\int_0^1 c(\dot{x}(t))dt$. Then, the second term is negligible with respect to the first one and does not influence the optimal trajectory. If now c has a unique optimum $\mathbb{O}(X)$ (which is a consequence of the assumptions of the law of large numbers), the unique trajectory starting from x and minimizing the first term is $x(t) = x + t\mathbb{O}(X)$ which leads to the value $f(x(1)) = f(x+\mathbb{O}(X))$ for the limit of $v_N(x)$.

References

[1] M. Akian, *Densities of idempotent measures and large deviations.* INRIA Report No. 2534, 1995.

[2] M. Akian, J.-P. Quadrat, M. Viot, *Bellman processes.* In *11th International Conference on Analysis and Optimization of Systems : Discrete Event Systems.* Lecture notes in Control and Information Sciences, **199**, Springer Verlag, 1994.

[3] M. Akian, J. P. Quadrat and M. Viot, *Duality between Probability and Optimization.* In *Proceedings of the workshop "Idempotency", Bristol, Oct. 94.* Cambridge University Press, 1995. To appear.

[4] H. Attouch, *Variational convergence for functions and operators.* Pitman, 1984.

[5] H. Attouch and R.J.B. Wets, *Isometries of the Legendre-Fenchel transform.* Transactions of the American Mathematical Society, **296**, p. 33-60, 1986.

[6] R. Azencott and Y. Guivarc'h and R.F. Gundy, *Ecole d'été de Saint Flour 8.* Lecture Notes in Mathematics, Springer-Verlag, Berlin, 1978.

[7] F. Baccelli, G. Cohen, G.J. Olsder and J.P. Quadrat, *Synchronization and linearity : an algebra for discrete events systems.* John Wiley & Sons, New-York, 1992.

[8] F. Bellalouna, *Un point de vue linéaire sur la programmation dynamique. Détection de ruptures dans le cadre des problèmes de fiabilité.* Thesis dissertation, Université Paris-IX Dauphine, 1992.

[9] R. Bellman and W. Karush, *Mathematical programming and the maximum transform.* SIAM Journal of Applied Mathematics, **10**, 1962.

[10] P. Bernhard, *Discrete and continuous time partial information minimax control* submitted to Annales de l'ISDG, 1994

[11] P. Del Moral, *Résolution particulaire des problèmes d'estimation et d'optimisation non-linéaires.* Thesis dissertation, Université Paul Sabatier de Toulouse, 1994.

[12] P. Del Moral, T. Thuillet, G. Rigal and G. Salut, *Optimal versus random processes: the non-linear case.* LAAS Report, Toulouse, 1991.

[13] W. Fenchel, *On the conjugate convex functions.* Canadian Journal of Mathematics, **1**, p. 73-77, 1949.

[14] M.I. Freidlin and A.D. Wentzell, *Random perturbations of dynamical systems.* Springer-Verlag, Berlin, 1979.

[15] G. Gierz, K.H. Hoffman, K. Keimel, J.D. Lawson, M. Mislove, D.S. Scott, *A compendium of continuous lattices.* Springer Verlag, Berlin, 1980.

[16] J.L. Joly, *Une famille de topologies sur l'ensemble des fonctions convexes pour lesquelles la polarité est bicontinue.* Journal de Mathématiques Pures et Appliquées, **52**, p. 421-441, 1973.

[17] V. N. Kolokoltsov and V .P. Maslov, *The general form of the endomorphisms in the space of continuous functions with values in a numerical commutative semiring (with the operation $\oplus = \max$).* Soviet Math. Dokl., Vol. 36, No. 1, 1988, pp 55–59.

[18] V.N. Kolokoltzov, V.P. Maslov, *New differential equation for the dynamic of the Pareto sets..* In *Proceedings of the workshop "Idempotency", Bristol, Oct. 94.* Cambridge University Press, 1995. To appear.

[19] V. Maslov, *Méthodes Opératorielles.* Editions MIR, Moscou, 1987.

[20] V.P. Maslov, V.N. Kolokoltzov, *Idempotent analysis and its applications to optimal control theory.* Moskow, Nauka, 1994 (Russian).

[21] V.P. Maslov, S.N. Samborski, *Idempotent analysis.* Advances In Soviet Mathematics, **13**, Amer. Math. Soc., Providence, 1992.

[22] V.P. Maslov, S.N. Samborski, *Stationary Hamilton-Jacobi and Bellman equations (existence and uniqueness of solutions)* in [21], p. 119-133, 1992.

[23] J.P. Quadrat, *Théorèmes asymptotiques en programmation dynamique.* Note CRAS, **311**, p.745-748, Paris, 1990.

[24] R.T. Rockafellar, *Convex analysis.* Princeton University Press Princeton, N.J., 1970.

[25] S.N. Samborski, A.A. Tarashchan, *The Fourier transform and semirings of Pareto sets.* in [21], p. 139.

[26] S.R.S. Varadhan, *Large Deviations and Applications.* CBMS-NSF Regional Conference Series in Applied Mathematics, N. 46, SIAM Philadelphia, Penn., 1984.

[27] P. Whittle, *Risk sensitive optimal control.* John Wiley & Sons, New-York,1990.

Unité de recherche INRIA Lorraine, Technopôle de Nancy-Brabois, Campus scientifique,
615 rue du Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unité de recherche INRIA Rennes, Irisa, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unité de recherche INRIA Rhône-Alpes, 46 avenue Félix Viallet, 38031 GRENOBLE Cedex 1
Unité de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unité de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
ISSN 0249-6399