Fixed Point Theory, 22(2021), No. 2, 511-526
DOI: 10.24193/fpt-ro.2021.2.34
http://www.math.ubbcluj.ro/~nodeacj/sfptcj.html

EXTENSION OF λ-PIR FOR WEAKLY CONTRACTIVE OPERATORS VIA FIXED POINT THEORY

A. BELHENNICHE*, S. BENAHMED* AND F.L. PEREIRA**

*École Nationale Polytechnique d'Oran Maurice Audin,
B.P 1523 El M'naouer Oran 31000, Algeria
E-mail: belhennicheabdelkader@yahoo.com / a_belhenniche@esc-alger.dz
sfyabenahmed@yahoo.fr /sfya.ghamnia@enp-oran.dz

**SYSTEC -Research Center for Systems and Technologies, Faculty of Electrical Engineering,
Porto University, Institute for Systems and Robotics, 4200-465, Porto, Portugal
E-mail: flp@fe.up.pt

Abstract. In this article, we apply methods of fixed point theory to investigate a Lambda policy iteration with a randomization algorithm for mappings that are merely weak contractions. As simple examples show, this class of mappings provide a much wider scope than the one afforded by strong contractions usually considered in the literature. More specifically, we investigate the properties of reinforcement learning procedures which have been developed for feedback control, in the framework of fixed point theory. Under fairly general assumptions, we determine sufficient conditions for the convergence with probability one in infinite dimensional policy spaces.
Key Words and Phrases: Fixed point theory, weakly contractive map, Lambda policy iteration with randomization.
2020 Mathematics Subject Classification: 47H10, 90C30, 49L20.

1. Introduction

Methods of fixed point theory play a key role in the investigation of a wide range of problems in Nonlinear Analysis. This article provides a powerful showcase of the use of methods of Fixed Point Theory to show how iterative schemes solve the optimal control synthesis by approximating the solution to the Hamilton-Jacobi equation via Dynamic Programming (DP) schemes under the weakest assumptions considered so far in the literature.

Dynamic programming became known in the 50's with the works of Richard Bellman, [6], being the key idea centered in solving a large decision problem by organizing it into simpler nested sub-problems which are solved recursively over time. Thus, if the sub-problems can be nested recursively within larger problems, so that dynamic programming method is applicable, then a relation - the so-called Bellman equation - between the value of the larger problem and the values of the sub-problems can be established. The impact of dynamic programming techniques in optimization based feedback control has been extremely significant, and, ever since, being expanded to a wide variety of control problems, notably, impulsive control (see [18], [4], and references therein) which can be regarded as an extension of the conventional optimal control of problems with absolutely continuous trajectories, [20].

These results are essential in many branches of science, economics, computer science, engineering, and, thus, its increasing use in control and optimization is not surprising, [7, 29, 22]. However, the huge computational complexity associated with solving the Bellman equation - the so called "curse of dimensionality" - is well known and, to circumvent it, iterative control policy or value function schemes approximating the optimal value function have been developed in recent years. An important class of these iterative schemes are designated by Reinforcement Learning (RL) algorithms. Bertsekas and Tsitsikilis in [12] present a broad class of RL algorithms in the context of Value Iteration (VI) and Policy Iteration (PI) methods. In [11], Bertsekas and Ioffe provided the analysis of Temporal Differences policy iteration ($TD(\lambda)$) scheme in the context of Neuro-Dynamic Programming framework in which they show that $TD(\lambda)$ scheme can be embedded into a PI scheme, designated by λ-PIR discussed in detail in [8]. In [10], Bertsekas investigates the connection between $TD(\lambda)$ and proximal algorithms which are more amenable for solving convex and optimization problems.

In [23], Yachao and Johanson and Jonas Martensson use abstract DP models, and extend the λ-PIR scheme for finite policy problems to contractive models with infinite policies. Moreover, they establish the well-posedness of the compact operator that plays a central in the algorithm, and determine the conditions for convergence with probability one of the λ-PIR scheme for problems with infinite dimensional policy spaces. The convergence results for the λ-PIR scheme are very important to provide guarantees for the convergence and other properties of deep learning techniques.

Fixed point theory is a powerful tool in topology, nonlinear analysis, optimal control, and machine learning. The well-known Banach contraction principle, commonly designated by **strong contraction**, states that if (X, d) is a complete metric space, and $T : X \to X$ is a mapping satisfying

$$d(T(x), T(y)) \leq Kd(x, y), \tag{1.1}$$

for some $K \in (0, 1)$ and for all $x, y \in X$, then T has a unique fixed point x^*, and the sequence $\{x_n\}$ generated by the iterative process $x_{n+1} = Tx_n$ converges to x^* for some $x \in X$. The generalization of the Banach contraction has been a heavily studied in several settings. In [13], Boyd and Wong replaced the constant K in (1.1) by an upper semicontinous function. In [3], Guerre-Delabriere introduced the notion of weakly contractive maps. In [5], a generalized Banach contraction conjecture was established. In [30], Suzuki has proved a generalization of the same nature to metric spaces. Many researchers have also obtained fixed point results for set-valued mappings satisfying generalized Banach contractions, notably, Nadler, [26], Durmaz, [15], Kikkawa and Suzuki, [21], M. Abbas, H. Iqbal, and Adrian Petruşel, [1].

By using methods of fixed point theory, in this article, we improve the results in [23] by showing that the properties of the proposed iterative procedure still hold for weakly contractive maps, a class of systems fundamentally much wider than the one in [23], as well as, the ones considered in previous works.

A mapping $f : X \to Y$ is said to be a weakly contractive or ψ-weak contraction, if there exists a continous function $\psi : [0,\infty) \to [0,\infty)$, with $\psi(0) = 0$, such that, for all $x, y \in X$, $\|f(x) - f(y)\| \leq \|x - y\| - \psi(\|x - y\|)$.

To see the relevance of this extension, just consider the following very simple example of weak contractive mapping that it is not a contraction. Let $f : [0,1] \to [0,1]$, defined by $f(x) = \sin(x)$.

First, let us show with a simple contradiction argument that f is not a contraction. Assume that sin is contraction. Then, for all $x, y \in \mathbb{R}$, there exists $K \in (0,1)$, such that $|\sin(x) - \sin(y)| \leq K|x - y|$. Now, let $x \to y$, and we conclude that, for any $y \in [0,1]$,

$$\lim_{x\to y} \frac{|\sin(x) - \sin(y)|}{|x - y|} = |\cos(y)| \leq K < 1.$$

The contradiction is obvious from the fact that $\cos(y)$ takes on the value 1 for $y = 0$.

Now, let us find a function ψ with the properties required to ensure that f is a ψ-weak contraction. Without any loss of generality, take $0 < y < x < 1$. First, notice the following inequality:

$$\begin{aligned} 5(x^3 - y^3) &= 5(x-y)\sum_{k=0}^{2} x^k y^{2-k} \geq 5(x-y)x^2 \\ &\geq 5(x-y)x^4 \geq (x-y)\sum_{k=0}^{4} x^k y^{4-k} = x^5 - y^5 \end{aligned} \tag{1.2}$$

From the Taylor's series expansion of f,

$$\sin(z) = \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k+1}}{(2k+1)!},$$

and, by considering Leibnitz's inequality

$$\frac{x^{2n+1} - y^{2n+1}}{(2n+1)!} \geq \frac{x^{2n+3} - y^{2n+3}}{(2n+3)!},$$

we have that

$$\begin{aligned} \sin(x) - \sin(y) &\leq x - y - \left(\frac{x^3 - y^3}{3!} - \frac{x^5 - y^5}{5!}\right) \\ &\leq x - y - \left(\frac{x^3 - y^3}{3!} - \frac{x^5 - y^5}{5!}\right) \\ &\qquad + \left(\frac{x^3 - y^3}{4!} - \frac{x^5 - y^5}{5!}\right) \\ &= (x-y) - \frac{1}{8}(x^3 - y^3) \\ &\leq (x-y) - \frac{1}{8}(x-y)^3, \end{aligned}$$

where the second inequality follows from (1.2). By reversing the roles of x and y, and by taking $\psi(t) = \dfrac{1}{8}t^3$ we may write

$$|\sin(x) - \sin(y)| \leq |x - y| - \psi(|x - y|).$$

We adopt the notion of weak contractivity defined by Guerre-Delabrier in [3] but in complete metric space as Rhoades showed in [28]. This is needed in order to guarantee the existence and uniqueness asserted by a fixed point theorem, as well as the convergence of the used operators. This challenge required the investigation of the convergence in norm, and of the existence and uniqueness of fixed point to the iterative process, the determination of error estimates, and the iteration of a Massa process, [24].

This article is organized as follows. In the next section, we formulate the iterative feedback control problem that we are going to investigate. Here, we provide the basic definitions, as well as the assumptions to be satisfied by its data. In the ensuing section, section 3, we present a number of fixed point results that are fundamental for the development of our contributions, notably, existence, monotonicity, and attainability. Also pertinent to our results is the Massa operator, [24], which will be also introduced in this section. The λ-policy iteration with randomization scheme is formulated in 4 in the context of the weakly contractive operators, where it is shown to inherit the properties already proved for the Massa operator. In section 5, we show that the convergence in norm with probability one of the iterative procedure defined for our problem under the stated assumptions, as well as its stability by proving appropriate error estimates. Finally, some conclusions, and prospective future work are briefly addressed in 6.

2. Problem formulation

We consider a set X of states, a set U of controls, and, for each $x \in X$, a nonempty control constraint $U(x) \subset U$. We denote by $\mathcal{M}$ the set of all functions $\mu : X \to U$ with $\mu(x) \in U(x)$ for all $x \in X$ which will be referred to as policies.
We denote by $\mathcal{J}(X)$ the set of functions $J : X \to \mathrm{R}$ and by $\bar{\mathcal{J}}(X)$ the set of functions $J : X \to \overline{\mathrm{R}}$ where $\overline{\mathrm{R}} = \mathrm{R} \cup \{-\infty, \infty\}$. We study the operator of the form

$$H : X \times U \times \mathcal{J}(X) \to \mathrm{R},$$

and, for each policy $\mu \in \mathcal{M}$, we consider the mapping $T_\mu : \mathcal{J}(X) \to \mathcal{J}(X)$ defined by

$$T_\mu J(x) := H(x, \mu(x), J), \quad \forall x \in X,$$

and the mapping $T : \mathcal{J}(X) \to \bar{\mathcal{J}}(X)$ defined by

$$TJ(x) := \inf_{\mu \in \mathcal{J}(X)} \{T_\mu J(x)\}, \quad \forall x \in X.$$

Example 2.1. Discounted semi Markov problem.

Let $x, y \in X$, $u \in U(x)$, and a mapping of the form

$$H(x, u, J) = G(x, u) + \sum_{y \in X} m_{xy}(u) J(y),$$

where G is some function representing the expected cost per stage, and, $\forall x \in X$, and $u \in U(x)$, $m_{xy}(u)$ are nonegative scalars satisfying

$$\sum_{y \in X} m_{xy}(u) < 1.$$

In view of the definition of $\mathcal{M}$, T, and T_μ, we have the following relations:

$$TJ(x) = \inf_{\mu \in \mathcal{M}} \{H(x, \mu(x), J)\} = \inf_{u \in U} \{H(x, u, J)\}.$$

Given some positive function $\nu : X \to \mathrm{R}$, we denote by $B(X)$ the set of functions J such that $\|J\| < \infty$, where the norm $\|\cdot\|$ on $B(X)$ is defined by

$$\|J\| = \sup_{x \in X} \left\{ \frac{|J(x)|}{\nu(x)} \right\}.$$

Lemma 2.1. *$B(X)$ is complete with respect to the topology induced by $\|\cdot\|$.*

It is not difficult to observe that $B(X)$ is closed, and convex.
Thus, given $\{J_k\}_{k=1}^{\infty} \subset B(X)$ and $J \in B(X)$, if $J_k \to J$ in the sense that

$$\lim_{k \to \infty} \|J_k - J\| = 0,$$

then $\lim_{k \to \infty} J_k(x) = J(x)$ for all $x \in X$.

Now, we introduce the following standard assumptions:

Assumption 2.1. *(Well posedness) For all $J \in B(X)$, and $\forall \mu \in \mathcal{M}$, we have that $T_\mu J \in B(X)$ and $TJ \in B(X)$.*

Definition 2.1. (ψ-weak contraction) A self map T of $B(X)$ is called a ψ-weak contraction if

$$\|TJ - TJ'\| \leq \|J - J'\| - \psi(\|J - J'\|), \;\; \forall\; J,\; J' \in B(X)$$

where $\psi : [0, \infty) \to [0, \infty)$ is a continuous and nondecreasing function satisfying $\psi(t) > 0$ if $t \in (0, \infty)$, and $\psi(0) = 0$.

Assumption 2.2. *The self map T_μ is a ψ-weak contraction.*

From definition 2, we conclude that every strong contraction T is also a ψ-weak contraction.

Every strong contraction map on $B(X)$ with contraction constant $k < 1$ is also a weak contraction with the map ψ given by $\psi(t) = (1 - k)t$ for all $t > 0$.

Obviously, T_μ is a continuous operator since, from the definition of the uniform continuity, it follows that

$$\forall\, \epsilon > 0, \exists\, \sigma > 0 : \|J - J'\| < \sigma \Longrightarrow \|T_\mu J - T_\mu J'\| < \epsilon$$

On other hand for all, for all $J, J' \in B(X)$ with $\|J - J'\| < \epsilon$ we have:

$$\|T_\mu J - T_\mu J'\| \leq \|J - J'\| - \psi(\|J - J'\|) \leq \|J - J'\| < \epsilon$$

Hence, T_μ is uniformly continous, and, thus, it is continous.

Example 2.2. Let $X = [0, \infty)$[1] be endowed with $d(x, y) = |x-y|$ and let $Tx = \dfrac{x}{x+1}$ for each $x \in X$. Define $\psi : [0, \infty) \to [0, \infty)$ by $\psi(t) = \dfrac{t^2}{1+t}$, then

$$\begin{aligned} d(Tx, Ty) &= \left| \frac{x}{1+x} - \frac{y}{1+y} \right| \\ &= \frac{|x-y|}{(1+x)(1+y)} \\ &\leq \frac{|x-y|}{1+|x-y|} \\ &= |x-y| - \frac{|x-y|^2}{1+|x-y|}. \end{aligned}$$

Thus,

$$d(Tx, Ty) \leq d(x, y) - \psi(d(x, y))$$

for all x, $y \in X$, and so T is a ψ-weak contraction. However, it is easily seen that it fails to be a contraction.

Example 2.3. Let $G = \{J \in B(X) : J(x) \geq 0\}$ be endowed by

$$\|J\| = \sup_{x \in X} \left\{ \frac{|J(x)|}{\nu(x)} \right\},$$

for a given nonnegative function ν defined on X, and let

$$TJ(x) := \frac{J(x)}{J(x)+1}$$

for each $J \in B(X)$, and $x \in X$.

Define $\psi : [0, \infty) \to [0, \infty)$ by $\psi(t) = \dfrac{t^2}{1+t}$. Let $\nu(x) = 1$. Then,

$$\begin{aligned} \|TJ - TJ'\| &= \left\| \frac{J}{1+J} - \frac{J'}{1+J'} \right\| \\ &= \frac{\|J-J'\|}{\|(1+J)(1+J')\|} \\ &\leq \frac{\|J-J'\|}{1+\|J-J'\|} \\ &= \|J-J'\| - \frac{\|J-J'\|^2}{1+\|J-J'\|}. \end{aligned}$$

Thus,

$$\|TJ - TJ'\| \leq \|J - J'\| - \psi(\|J - J'\|)$$

holds for all J, $J' \in G$, and so T is ψ-weak contraction. However, it is easily seen that it fails to be a contraction.

---

[1]We consider $[0, +\infty)$, and not $B(X)$ just to distinguish between contraction and weakly contractive map.

3. Auxiliary results

In this section, a number of results that will be instrumental in the proof of the main result of this article are presented.

Theorem 3.1. *(Existence) Let the operators $T, T_\mu : B(X) \to B(X)$ be ψ-weak contractions, then T and T_μ have, respectively, J^* and J_μ as fixed points.*

Note that Theorem 3.1 is special case of Theorem 1 in [13]. Remark that, although these results have been proved in the context of Hilbert spaces, they also hold for uniformly convex Banach spaces. From Theorem 3.1, it follows:

Lemma 3.1. *The following holds:*

i) *For an arbitrary $J_0 \in B(X)$, the sequence $\{J_k\}$ defined by $J_{k+1} = T_\mu J_k$ converges in norm to J_μ.*
ii) *For an arbitrary $J_0 \in B(X)$, the sequence $\{J_k\}$ defined by $J_{k+1} = TJ_k$ converges in norm to J^*.*

In fact, since $B(X)$ is convex and closed, we have the following theorem.

Theorem 3.2. (*Theorem 2 of* [28]) *Let T, and T_μ be ψ-weak contractions self-maps in the closed convex $B(X)$, then the iterative process $J_{k+1} = TJ_k$, and $J_{k+1} = T_\mu J_k$ converge in norm, respectively, to the fixed points J^*, and J_μ, with the following error estimates*

$$\|J_k - J^*\| \leq \Phi^{-1}(\Phi(\|J_1 - J^*\|) - (k-1)),$$

and

$$\|J_k - J_\mu\| \leq \Phi^{-1}(\Phi(\|J_1 - J_\mu\|) - (k-1)).$$

Here, Φ is the antiderivative defined by

$$\Phi(t) = \int \frac{dt}{\psi(t)}$$

and Φ^{-1} denotes its inverse.

We will require the following properties to hold.

Assumption 3.1. *(Monotonicity) $\forall\ J,\ J' \in B(X)$, we have that $J \leq J'$ implies*

$$H(x,u,J) \leq H(x,u,J'), \quad \forall x \in X,\ u \in U(x)$$

where $\leq$ is defined in a pointwise sense in X.

Assumption 3.2. *(Attainability) For all $J \in B(X)$, there exists $\mu \in \mathcal{M}$, such that $T_\mu J = TJ$.*

Now, we consider the convergence of the iteration process of Massa, see [28], in the context of weak contractions. The Massa operator is defined as follows.

$$S = \sum_{l=0}^{\infty} \alpha_l T_\mu^l, \tag{3.1}$$

where $\sum_{l=0}^{\infty} \alpha_l = 1$, and $\alpha_k \alpha_{k+1} \neq 0$ for at least one integer k. Note that this condition does not entail any loss of generality.

In the above article, the following result was proved.

Theorem 3.3. *(Theorem 5 of* [28]*) Let T be a weakly contractive self map in $B(X)$. Then, for any $J_0 \in B(X)$, the iteration scheme $S^k J_0$ converges to the unique fixed point of T_μ.*

This theorem plays a major role in supporting the development of our main contributions presented and discussed in the next two sections.

4. λ-policy iteration with randomization

The λ-PIR algorithm is a policy iteration iterative procedure with randomization that has been investigated by several authors [8, 9, 23], being infinite policies considered in the later.

However, all these results require the strong contractiveness of the operators being iterated. As mentioned in the introduction, this is where the added value to state-of-the-art of this article resides. We prove the most significant properties of the most recent publications under assumptions weaker than those considered so far. More precisely, we just require the weak contractiveness of the iterative operator.

Our developments rely on the following result showing that the function

$$\Theta(t) := \sum_{l=1}^{\infty} \alpha_l \psi(t),$$

where $\psi(\cdot)$ is the function appearing in the weak contractiveness property, inherits the properties established in Theorem 5 of [28].

Theorem 4.1. *Let T be a ψ-weakly contractive self map in $B(X)$. Then, the Massa operator S defined by (3.1) is a well defined and a well posed Θ-weak contraction with the function $\Theta(t) := \sum_{l=1}^{\infty} \alpha_l \psi(t)$.*

Proof. Let us consider the mapping S on the domain $B(X)$ defined pointwisely by

$$SJ(x) = \sum_{l=0}^{\infty} \alpha_l T_\mu^l J(x),$$

for all $x \in X$, and $J \in B(X)$. Now, we show that

1. S is well defined. The proof of this fact is analogous to the one of Lemma 3.1 in [23]. However, we include it here for the sake of completeness, and also because it is useful to follow the developments presented later.

Let us consider the sequence $\{\Im_l\}_{k=0}^{\infty}$ defined by $\Im_l = \sum_{l=0}^{k} \alpha_l (T_\mu^l J)(x)$.

From the above lemma, we have that $(T_\mu^l J)(x) \to J_\mu(x) \in \mathbb{R}$, and, thus, $\{(T_\mu^l J)(x)\}_{l=0}^{\infty}$ is bounded.

Denote this bound by $K_\mu(x) \in \mathbb{R}$. Then, $\forall\ k$, we have

$$\begin{aligned}\left\|\sum_{l=0}^{k} \alpha_l (T_\mu^l J(x)\right\| &\leq \sum_{l=0}^{n} \alpha_l \|(T_\mu^l J)(x)\| \\ &\leq \sum_{l=0}^{k} \alpha_l K_\mu(x) \leq K_\mu(x).\end{aligned}$$

This entails the boundedness of the sequence $\{\Im_l\}_{k=0}^{\infty}$.
If $J_\mu(x) > 0$, then $\exists\ N$ such that $(T_\mu^l J)(x) > 0$, $\forall l > N$.
Therefore, $\left\{\sum_{l=0}^{k} \alpha_l (T_\mu^l J)(x)\right\}_{k=N}^{\infty}$ is monotonically nondecreasing and bounded above by $K_\mu(x)$. Thus, the sequence $\{\Im_l\}$ converges to the limit $\sum_{l=0}^{\infty} \alpha_l (T_\mu^l J)(x) \in \mathbb{R}$.
A similar argument applies for the case $J_\mu(x) < 0$.
If $J_\mu(x) = 0$, then $\forall\ \epsilon$, $\exists\ N$ such that, $\forall l > N$, $\|T_\mu^l J(x)\| < \epsilon$. Therefore, $\forall\ k$, we have that

$$\begin{aligned}\left\|\sum_{l=0}^{N} \alpha_l (T_\mu^l J)(x) - \sum_{l=0}^{N+k} \alpha_l (T_\mu^l J)(x)\right\| &= \left\|\sum_{l=N}^{N+k} \alpha_l (T_\mu^l J)(x)\right\| \\ &\leq \sum_{l=N}^{N+k} \alpha_l \|(T_\mu^l J)(x)\| \\ &\leq \sum_{l=N}^{N+k} \alpha_l \leq \epsilon.\end{aligned}$$

This implies that the sequence $\{\Im_l\}$ is Cauchy, and, thus, it converges in $\mathbb{R}$.
Therefore, $\forall\ J \in B(X)$, and $x \in X$, the sequence $\{\Im_l\}$ converges in $\mathbb{R}$.

2. S is well posed. Remark that we have $SJ_\mu = J_\mu$, otherwise, if $J \neq J_\mu$, we would have

$$\begin{aligned}|SJ(x) - J_\mu(x)| &= \left|\sum_{l=0}^{\infty} \alpha_l T_\mu^l J(x) - J_\mu(x)\right| \\ &= \left|\sum_{l=0}^{\infty} \alpha_l T_\mu^l J(x)) - \sum_{l=0}^{\infty} \alpha_l T_\mu^l J_\mu(x)\right| \\ &\leq \sum_{l=0}^{\infty} \alpha_l |T_\mu^l J(x) - T_\mu^l J_\mu(x)|.\end{aligned}$$

Due to the fact that, $\forall\ l$, T_μ is weak contraction map, we have that

$$\begin{aligned}\|T_\mu^{l+1} J - T_\mu^{l+1} J_\mu\| &\leq \|T_\mu^l J - T_\mu^l J_\mu\| - \psi(\|T_\mu^l J - T_\mu^l J_\mu\|) \\ &\leq \|T_\mu^l J - T_\mu^l J_\mu\| \\ &\leq \|J - J_\mu\| - \psi(\|J - J_\mu\|),\end{aligned}$$

and, thus,

$$|T_\mu^l J(x) - T_\mu^l J_\mu(x)| \leq (\|J - J_\mu\| - \psi(\|J - J_\mu\|))\nu(x).$$

Therefore, we have

$$|SJ(x) - J_\mu(x)| \leq \sum_{l=0}^{\infty} \alpha_l(\|J - J_\mu\| - \psi(\|J - J_\mu\|))\nu(x),$$

and this entails that

$$\sup_{x\in X} \left\{ \frac{|SJ(x) - J_\mu(x)|}{\nu(x)} \right\} \leq \|J - J_\mu\| - \psi(\|J - J_\mu\|).$$

Thus,

$$\|SJ\| \leq (\|J - J_\mu\| - \psi(\|J - J_\mu\|)) + \|J_\mu\|.$$

The boundedness of $B(X)$ implies that $\psi(\|J - J_\mu\|)$ is finite, and, this together with $J_\mu \in B(X)$, leads to $SJ \in B(X)$.

3. S is a Θ-weak contraction with the function $\Theta(\cdot) = \sum_{l=1}^{\infty} \alpha_l \psi(\cdot)$.

$$\begin{aligned}
\|SJ - SJ'\| &\leq \left\| \sum_{l=0}^{\infty} \alpha_l (T^l J - T^l J') \right\| \\
&\leq \sum_{l=0}^{\infty} \alpha_l \|(T^l J - T^l J')\| \\
&= \alpha_0 \|T_\mu J - T_\mu J'\| + \sum_{l=1}^{\infty} \alpha_l \|(T^l J - T^l J')\| \\
&\leq \alpha_0 \|J - J'\| + \sum_{l=1}^{\infty} \alpha_l [\|J - J'\| - \psi(\|J - J'\|)] \\
&\leq \|J - J'\| - \sum_{l=1}^{\infty} \alpha_l \psi(\|J - J'\|)
\end{aligned}$$

Corollary 4.1. *The Massa iteration process defined by $J_{k+1} := SJ_k$ converges in norm to the fixed point J_μ with the error estimate*

$$\|J_k - J_\mu\| \leq \Phi^{-1}(\Phi(\|J_1 - J_\mu\|) - (k - 1)).$$

Proof. The convergence in norm to the fixed point J_μ is guarenteed by Theorem 3.6. Now, we just proof the second part. For all $k \in \mathbb{N}$, we may write

$$\|J_{k+1} - J_\mu\| = \|SJ_k - SJ_\mu\| \leq \|J_k - J_\mu\| - \psi(\|J_k - J_\mu\|).$$

The sequence of positive numbers $\{\lambda_k\}$ defined by $\lambda_k = \|J_k - J_\mu\|$ satisfies the following inequality

$$\lambda_{k+1} \leq \lambda_k - \psi(\lambda_k).$$

This implies that the sequence $\{\lambda_k\}$ is nonincreasing. Thus, it converges to some λ such that $\psi(\lambda) \le 0$. By the definition of ψ, this implies that $\lambda = 0$.
Let us show that, $\forall\, k \ge 1$, we have

$$\lambda_k \le \Phi^{-1}(\Phi(\lambda_1) - (k-1)).$$

Indeed, by definition of Φ, $\forall\, k \ge 1$ there exist $c_k \in [\lambda_{k+1}, \lambda_k]$ such that:

$$\begin{aligned} \Phi(\lambda_k) - \Phi(\lambda_{k+1}) &= \Phi'(c_k)(\lambda_k - \lambda_{k+1}) \\ &= \frac{\lambda_k - \lambda_{k+1}}{\psi(c_k)} \\ &\ge \frac{\lambda_k - \lambda_{k+1}}{\psi(\lambda_k)} \ge 1. \end{aligned}$$

This implies that, for all $k \ge 1$,

$$\Phi(\lambda_{k+1}) \le \Phi(\lambda_k) - 1 \le \cdots \le \Phi(\lambda_1) - k$$

and, thus,

$$\Phi(\lambda_k) \le \Phi(\lambda_1) - (k-1),$$

i.e.,

$$\|J_k - J_\mu\| \le \Phi^{-1}(\Phi(\|J_1 - J_\mu\|) - (k-1)).$$

Now, we are ready to put these general results at work in order to solve the problem stated in section 2.

Let us adopt the notation of [23]. Given some $\lambda \in [0,1)$, consider the mappings $T_\mu^\lambda \in B(X)$ defined pointwisely by

$$T_\mu^\lambda J(x) = (1-\lambda)\sum_{l=0}^{\infty} \lambda^l T_\mu^{l+1} J(x).$$

In what follows, we refer to this operator by "λ-operator". For given $J_k \in B(X)$, and $p_k \in (0,1)$, the current policy μ, and the next cost approximation J_{k+1} are computed as follows:

$$T_\mu J_k = T J_k \tag{4.1}$$

$$J_{k+1} = \begin{cases} T_\mu J_k & \text{w.p. } p_k \\ T_\mu^\lambda J_k & \text{w.p. } 1 - p_k \end{cases} \tag{4.2}$$

Here, "w.p." stands for "with probability". By putting $S = T_\mu^\lambda$, and replacing α_l by $(1-\lambda)\lambda^l$, and by noting that $\sum_{l=0}^{\infty}(1-\lambda)\lambda^l = 1$, we, immediately conclude, from Theorem 4.1, that:

- T_μ^λ is well defined.
- T_μ^λ is well posed.
- T_μ^λ is a weak contraction mapping with $\Theta(\cdot) = \sum_{l=1}^{\infty}(1-\lambda)\lambda^l \psi(t)$.

From these considerations, the next Lemma follows immediately.

Lemma 4.1. *Let $T_\mu : B(X) \to B(X)$ satisfy assumptions 2.2 and 2.3, then the mappings T_μ^λ are monotonic in the sense that*

$$J \leq J' \to T_\mu^\lambda J \leq T_\mu^\lambda J' \;\; \forall\, x \in X,\; \mu \in \mathcal{M}.$$

5. Convergence of the λ-PIR algorithm

In this section, we establish the convergence of the λ-PIR algorithm, and provide error estimates.

Now, we state the main result of this section.

Theorem 5.1. *Let assumptions 2.2, 3.1, and 3.2 hold. For a given $J_0 \in B(X)$ satisfying $TJ_0 < J_0$, the sequence $\{J_k\}_{k=0}^{\infty}$ generated by the algorithm represented by* (4.1) *and* (4.2) *converges in norm with probability one.*

Proof. Since $TJ_0 < J_0$, we have $T_{\mu^0}J_0 = TJ_0 < J_0$.

By monotonicity of T_{μ^0}, we have

$$T_{\mu^0}^l J_0 \leq T_{\mu^0}^{l-1} J_0, \quad T_{\mu^0}^l J_0 \leq TJ_0, \quad \forall l$$

which implies that

$$T_{\mu^0}^\lambda J_0 \leq T_{\mu^0} J_0 \leq J_0.$$

This means that J_1 bounded from above with probability one by TJ_0. In addition, we also have $J_{\mu^0} \geq J^*$ where J_{μ^0} is the fixed point of both $T_{\mu^0}^\lambda$, and T_{μ^0}, and, thus, we have that

$$J^* \leq J_{\mu^0} \leq T_{\mu^0}^\lambda J_0 \leq T_{\mu^0} J_0.$$

This implies that J_1 is bounded from below by J^* with probability one.

Now, due to the ψ-weak contraction assumption we have that

$$T^2 J_0 = T(T_{\mu^0} J_0) \leq T_{\mu^0} J_0,$$

and, due to the monotonicity of $T_{\mu^0}^\lambda$, and since $T_{\mu^0}^\lambda$ and T_μ can commute, we have that

$$T(T_{\mu^0}^\lambda J_0) \leq T_{\mu^0}(T_{\mu^0}^\lambda J_0) \leq T_{\mu^0}^\lambda J_0.$$

By proceeding with the iterations, we get that

$$J^* \leq J_k \leq T^k J_0$$

which means that $\{J_k\}$ converges in norm to J^* with probability one.

The following result provides an estimate of the approximation error.

Theorem 5.2. *Let the assumptions and conditions of theorem 5.1 hold, then the randomized iterative mix between VI and TD defined by the algorithm represented by the algorithm* (4.1), *and* (4.2) *generates values satisfying the following error estimate*

$$\|J_k - J^*\| \leq \Phi^{-1}(\Phi(\|J_1 - J^*\| - (k-1)).$$

Proof. Since we have

$$\|J_k - J^*\| \leq \|T^k J_0 - J^*\|,$$

and from corollary 4.1,

$$\|T^k J_0 - J^*\| \leq \Phi^{-1}(\Phi(\|J_1 - J^*\|) - (k-1))$$

Thus, we get that

$$\|J_k - J^*\| \leq \Phi^{-1}(\Phi(\|J_1 - J^*\|) - (k-1)).$$

Now we present an example illustrating our results. Let

$$H(x,u,J) = \int_X \left(g(x,u,y) + \frac{J}{J+1}(y) \right) d\mathbb{P}(y \mid x,u)$$

where $g : X \times U \times X \to \mathbb{R}$, and $P(\cdot \mid x,u)$ is the probability measure conditioned on (x,u) for a certain MDP.

Let $\nu(x) = 1$ for all $x \in X$, and let assumptions of theorem 5.1 hold. Given an arbitrary $J_0 \in B(X)$, and since $H(\cdot,\cdot,\cdot)$ is a weakly contractive map, the sequence $\{J_k\}_{k=0}^{\infty}$ generated by the algorithm represented by (4.1) and (4.2) converges in norm with probability one.

6. Conclusions

In this article, we extended the reinforcement learning procedure designated by Lambda policy iteration with randomization previously developed strongly contractive operators to the much wider class of weakly contractive operators by applying methods of fixed point theory. By resorting to the Massa operator, we prove the well definition, well-posedeness, and the weakly contractiveness of Lambda policy iteration scheme in Banach spaces. In future, we will seek to develop similar results to generalized metric spaces being the motivation fuelled by the investigation of convergence properties of general domain computational algorithms.

Acknowledgment. This work is supported by grant of the ministry of Higher Education and Scientific Research of Algeria, project number: 685/PNE/Enseignant/Chercheur/Portugal/2019-2020, awarded to Abdelkader Belhenniche. This work is supported by the R&D Unit SYSTEC, ref. POCI-010145-FEDER-006933 funded by ERDF - COMPETE2020 - FCT (Foundation for Science and Technology)/MEC - PT2020, Project STRIDE, ref. NORTE-01-0145FEDER-000033, funded by ERDF - NORTE 2020, and Project HARMONY, ref.AAC n2/SAICT/2017 — 031411, funded by FCT/MEC - ERDF - COMPETE2020

References

[1] M. Abbas, H. Iqbal, A. Petruşel, *Fixed points for multivalued Suzuki type $(\theta, \mathcal{R})$-contraction mapping with applications*, Journal of Function Spaces, **2019**(2019).
[2] Y.I. Alber, *Recurrence relations and variational inequalities*, Doklady Akademii Nauk. Russian Academy of Sciences, **270**(1983), no. 1, 11-17.
[3] Y.I. Alber, S. Guerre-Delabriere, *Principle of weakly contractive maps in Hilbert spaces. New Results in Operator Theory and its Applications*, Birkhäuser, Basel, 1997, 7-22.

[4] A. Arutyunov, V. Jakimovic, F. Pereira, *Second order necessary conditions for optimal impulsive control problems*, Journal of Dynamical and Control Systems, **9**(2003), no. 1, 131-153.
[5] A. Arvanitakis, *A proof of the generalized Banach contraction conjecture*, Proceedings of the American Mathematical Society, **131**(2003), no. 12, 3647-3656.
[6] R. Bellman, *The Theory of Dynamic Programming*, Rand Corporation Report P-550, Santa Monica, CA, 1954.
[7] D. Bertsekas, *Dynamic Programming and Optimal Control*, Athena Scientific, Belmont, MA, 1995.
[8] D.P. Bertsekas, *Abstract Dynamic Programming*, Athena Scientific, Belmont, Massachusetts, 2013.
[9] D.P. Bertsekas, *Abstract Dynamic Programming*, Athena Scientific, 2018.
[10] D.P. Bertsekas, *Proximal algorithms and temporal difference methods for solving fixed point problems*, Computational Optimization and Applications, **70**(2018), no. 3, 709-736.
[11] D.P. Bertsekas, S. Ioffe, *Temporal differences-based policy iteration and applications in neuro-dynamic programming*, Lab. for Info. and Decision Systems Report LIDS-P-2349, 1996.
[12] D.P. Bertsekas, J.N. Tsitsiklis, *Neuro-Dynamic Programming*, Athena Scientific, 1996.
[13] D.W. Boyd, J.S. Wong, *On nonlinear contractions*, Proceedings of the American Mathematical Society, **20**(1969), no. 2, 458-464.
[14] B.S. Choudhury, N. Metiya, M. Postolache, *A generalized weak contraction principle with applications to coupled coincidence point problems*, Fixed Point Theory and Applications, **2013**(2013), no. 1, 152.
[15] G. Durmaz, *Some theorems for a new type of multivalued contractive maps on metric space*, Turkish Journal of Mathematics, **41**(2017), no. 4, 1092-1100.
[16] G. Durmaz, I. Altun, *A new perspective for multivalued weakly picard operators*, Publications de l'Institut Mathematique, **101**(2017), no. 115, 197-204.
[17] P.N. Dutta, B. Choudhury, S. Binayak, *A generalization of the contraction principle in metric spaces*, Fixed Point Theory and Applications, **2008**(2008), no. 1, 406368.
[18] S. Fraga, F. Pereira, *Hamilton-Jacobi-Bellman equation and feedback synthesis for impulsive control*, IEEE Transactions on Automatic Control, **57**(2012), no. 1, 244-249.
[19] H.A. Hancer, A. Hatice, G. Mina, I. Altun, *On a broad category of multivalued weakly Picard operators*, Fixed Point Theory, **18**(2017), no. 1, 229-236.
[20] D. Karamzin, V. De Oliveira, F. Pereira, G. Silva, *On some extension of optimal control theory*, European Journal of Control, **20**(2014), no. 6, 284-291.
[21] M. Kikkawa, T. Suzuki, *Three fixed point theorems for generalized contractions with constants in complete metric spaces*, Nonlinear Analysis: Theory, Methods and Applications, **59**(2008), no. 9, 2942-2949.
[22] D.E. Kirk, *Optimal Control Theory: An Introduction, 1970*, Prentice-Hall, Englewood Cliffs, NJ, **3**(1998), 417-442.
[23] Y. Li, K.H. Johansson, J. Martensson, *Lambda-Policy Iteration with Randomization for Contractive Models with Infinite Policies: Well Posedness and Convergence* (Extended Version), arXiv preprint:1912.08504, 2019.
[24] S. Massa, *Convergence of an iterative process for a class of quasi-nonexpansive mappings*, Boll. U.M.I., **51-A**(1978), 154-158.
[25] J. Merryfield, B. Rothschild, JR.J. Stein, *An application of Ramsey's theorem to the Banach contraction principle*, Proceedings of the American Mathematical Society, **130**(2002), no. 4, 927-933.
[26] S.B. Nadler, *Multi-valued contraction mappings*, Pacific Journal of Mathematics, **30**(1969), no. 2, 475-488.
[27] B.E. Rhoades, *A comparison of various definitions of contractive mappings*, Transactions of the American Mathematical Society, **226**(1977), 257-290.
[28] B.E. Rhoades, *Some theorems on weakly contractive maps*, Nonlinear Analysis: Theory, Methods and Applications, **47**(2001), no. 4, 2683-2693.
[29] M. Sniedovich, *Dynamic Programming: Foundations and Principles*, CRC Press, 2010.

[30] T. Suzuki, *A generalized Banach contraction principle that characterizes metric completeness*, Proceedings of the American Mathematical Society, **136**(2008), no. 5, 1861-1869.

Received: May 16, 2020; Accepted: November 29, 2020.