Chapter 5

Warped Yosida Approximation and its Properties

5.1 Introduction

Newton-like dynamical system governed by maximally monotone operator is as follows:

$$\dot{x}(t) + Tx(t) = 0,$$

where T is a maximally monotone operator. This dynamical system is ill-posed. One can regularize the monotone operators on Hilbert into single-valued Lipschitzian operators via a process known as the Yosida approximation. This approximation is very applicable due to ill-posedness of Newton-like dynamical system governed by maximally monotone operator.

The resolvent and Yosida approximation plays an important role in the convergence of dynamical system associated with maximal monotone operator [96, 97, 98]. In 2020, Bui et al. [99] have introduced warped resolvent and discussed its properties. In this sequel, in this manuscript, we have defined warped Yosida approximation and analyzed it properties.

5.2 Warped Yosida Approximation

In this section, we define Yosida approximation and provide some characterization and properties.

Definition 5.2.1. Let B be a reflexive Banach space with dual space B^*. Assume that $C(\neq \emptyset) \subseteq B$, $M : C \to B^*$ and $T : B \to 2^{B^*}$ are such that $\text{ran}(M) \subset \text{ran}(T + \gamma M)$ and $T + \gamma M$ is injective. For any $\gamma \in (0, \infty)$, warped Yosida approximation of T with kernel M is defined by $A_\gamma^M = \frac{1}{\gamma}\left(M - M \circ J_{\gamma T}^M\right)$, where $J_{\gamma T}^M$ is warped resolvent.

Example 5.2.1. Let $C(\neq \emptyset)$ be a subset of B and $\phi : B \to (-\infty, \infty]$ be a proper convex lower semicontinuous map. Let $\gamma > 0$. Assume that $M : C \to B^*$ is an operator with $\text{ran}(M) \subset \text{ran}(M + \gamma\partial\phi)$ and $M + \gamma\partial\phi$ is injective. Then warped Yosida approximation of $\partial\phi$ is $(\partial\phi)_\gamma^M = \frac{1}{\gamma}(M - M \circ Prox_{\gamma\partial\phi}^M)$, where $Prox_{\gamma\partial\phi}^M = (M + \gamma\partial\phi)^{-1} \circ M$.

Let M be an injective operator. The warped Yosida approximation of $\partial\phi$ is described by the following variational inequality:

$$z = (\partial\phi)_\gamma^M \Leftrightarrow (\forall y \in B)\ \langle y - x + \gamma M^{-1}z, \gamma z\rangle + \phi(x - \gamma M^{-1}z) \leq \phi(z)\ \forall (x, z) \in B \times B.$$

Example 5.2.2. Let $T : B \to 2^{B^*}$ be a maximal monotone operator such that $Zer(T) \neq \emptyset$. Suppose that $f : B \to (-\infty, \infty]$ is an admissible function such that $\mathcal{D}(T) \subset int\ \mathcal{D}(f)$. Set $M = \nabla f$. Then $T_\gamma^{\nabla f}$ is well defined warped Yosida approximation defined in [100].

Now, we provide an example of warped Yosida approximation with respect to different choices of admissible function f.

Example 5.2.3. Let $A : (0, \infty) \to \mathbb{R}$ be a monotone mapping. Define an admissible function (Boltzmann-Shannon entropy) $\mathcal{BS} : (0, \infty) \to (0, \infty)$ by $x \mapsto x \log x - x$.

Then warped resolvent of A is [101]

$$J_A^{\mathcal{BS}} x = (\log + A)^{-1} \circ Ax = xe^{Ax}.$$

Next, warped Yosida approximation

$$\begin{aligned} A^{\mathcal{BS}} x &= \nabla BSx - \nabla \mathcal{BS} \circ J_A^{\mathcal{BS}} x \\ &= \log x - \log(xe^{Ax}) \\ &= \log x - \log x - \log e^{Ax} \\ &= -Ax. \end{aligned}$$

In the next proposition, we give an explicit formula for warped Yosida approximation of a monotone operator, which is displacement mapping of a linear isometry of finite order.

Proposition 5.2.1. Let H be a real Hilbert space and $R : H \to H$ be a linear isometry of finite order m. Define $T := I - R$, referred as the displacement mapping of R, and hence T is a maximal monotone operator. Assume that $M : H \to H$ is an injective operator with $\text{ran}(M) \subset \text{ran}(T + \gamma M)$ and $T + \gamma M$ is injective, for $\gamma > 0$. We have the following estimates for warped resolvent and warped Yosida approximation:

(i) $J^M_{\gamma(I-S)} = \frac{I}{M+\gamma I} \sum_{n=0}^{\infty} \left(\frac{\gamma}{M+\gamma I}\right)^n S^n \circ M$, where $S : H \to H$ is a nonexpansive and linear operator.

(ii) $J^M_{\gamma T} = \frac{I}{(M+\gamma I)^m - (\gamma I)^m} \sum_{n=0}^{m-1} (M + \gamma I)^{m-1-n} \gamma^n R^n \circ M$.

(iii) $T^M_\gamma = \frac{1}{\gamma} \left(M - M \circ J^M_{\gamma T}\right)$.

Proof. (i) Let $S : H \to H$ be a nonexpansive and linear operator.

$$\begin{aligned} J^M_{\gamma(I-S)} &= (M + \gamma(I - S))^{-1} \circ M \\ &= M^{-1}(I + \gamma M^{-1}(I - S))^{-1} \circ M \\ &= M^{-1}\left[(I + \gamma M^{-1})\left(I - \frac{\gamma M^{-1}}{I + \gamma M^{-1}} S\right)\right]^{-1} \circ M \\ &= \frac{M^{-1}}{I + \gamma M^{-1}} \sum_{n=0}^{\infty} \left(\frac{\gamma M^{-1}}{I + \gamma M^{-1}}\right)^n S^n \circ M \\ &= \frac{I}{M + \gamma I} \sum_{n=0}^{\infty} \left(\frac{\gamma}{M + \gamma I}\right)^n S^n \circ M \end{aligned}$$

(ii) Let $T := I - R$, where R is an isometry of finite order m. So, R is surjective and $\|R\| = 1$, and hence R is nonexpansive. Therefore from (i), we have

$$\begin{aligned} J^M_{\gamma T} &= \frac{I}{M + \gamma I} \sum_{n=0}^{\infty} \left(\frac{\gamma}{M + \gamma I}\right)^n R^n \circ M \\ &= \left(\frac{I}{M + \gamma I}\right)\left(\sum_{n=0}^{m-1} \left(\frac{\gamma}{M + \gamma I}\right)^n R^n + \left(\frac{\gamma}{M + \gamma I}\right)^m \sum_{n=0}^{m-1} \left(\frac{\gamma}{M + \gamma I}\right)^n R^n + \cdots\right) \circ M \\ &= \left(\frac{I}{M + \gamma I}\right)\left(1 + \left(\frac{\gamma}{M + \gamma I}\right)^m + \left(\frac{\gamma}{M + \gamma I}\right)^{2m} + \cdots\right) \sum_{n=0}^{m-1} \left(\frac{\gamma}{M + \gamma I}\right)^n R^n \circ M \\ &= \left(\frac{I}{M + \gamma I}\right)\left(\frac{1}{1 - \left(\frac{\gamma I}{M + \gamma I}\right)^m}\right) \sum_{n=0}^{m-1} \left(\frac{\gamma}{M + \gamma I}\right)^n R^n \circ M \\ &= \left(\frac{(M + \gamma I)^{m-1}}{(M + \gamma I)^m - (\gamma I)^m}\right) \sum_{n=0}^{m-1} \left(\frac{\gamma}{M + \gamma I}\right)^n R^n \circ M \\ &= \frac{I}{(M + \gamma I)^m - (\gamma I)^m} \sum_{n=0}^{m-1} (M + \gamma I)^{m-1-n} \gamma^n R^n \circ M. \end{aligned}$$

□

Now, we explore some properties and characteristics of warped Yosida approximation.

Proposition 5.2.2. Let $T : B \to 2^{B^*}$ and $M : B \to B^*$ be the operators with $\text{ran}(M) \subset \text{ran}(T + \gamma M)$ and $T + \gamma M$ is injective, for $\gamma > 0$. Then we have the following:

(i) $\left(J^M_{\gamma T}(x), T^M_\gamma(x)\right) \in \mathcal{G}(T)$, $x \in B$.

(ii) $0 \in T(x)$ if and only if $0 \in T^M_\gamma(x)$, $x \in B$.

Proof. (i) For $x \in B$, we have

$$\begin{aligned}
&J^M_{\gamma T}(x) = (M + \gamma T)^{-1} \circ M(x) \\
\Leftrightarrow &M(x) \in (M + \gamma T) \circ J^M_{\gamma T}(x) \\
\Leftrightarrow &\frac{1}{\gamma}\left(M - M \circ J^M_{\gamma T}\right)(x) \in T\left(J^M_{\gamma T}(x)\right) \\
\Leftrightarrow &T^M_\gamma(x) \in T\left(J^M_{\gamma T}(x)\right).
\end{aligned}$$

(ii) For $x \in B$,

$$\begin{aligned}
&0 \in T(x) \Leftrightarrow 0 \in \gamma T(x) \\
\Leftrightarrow &M(x) \in (M + \gamma T)(x) \\
\Leftrightarrow &x \in (M + \gamma T)^{-1} \circ M(x) \\
\Leftrightarrow &M(x) \in M(J^M_{\gamma T}(x)) \\
\Leftrightarrow &0 \in \left(M - M \circ J^M_{\gamma T}\right)(x) \\
\Leftrightarrow &0 \in \gamma T^M_\gamma x \Leftrightarrow 0 \in T^M_\gamma x.
\end{aligned}$$

□

Proposition 5.2.3. Let $T : B \to 2^{B^*}$ be an operator and $M : B \to B^*$ be an injective operator with $\operatorname{ran}(M) \subset \operatorname{ran}(T+\gamma M)$ and $T+\gamma M$ is injective, for $\gamma > 0$. Let $\lambda > 0$. Then we have the following:

(i) $T_\gamma^M = (\gamma M^{-1} + T^{-1})^{-1}$.

(ii) $T_\gamma^M = J_{\gamma^{-1}T^{-1}}^{M^{-1}} \circ \gamma^{-1} M$.

(iii) $T_{\gamma+\lambda}^M = \left(T_\gamma^M\right)_\lambda^M$.

(iv) Let $x, y \in B$. Then $x = T_\gamma^M y \Leftrightarrow (y - \gamma M^{-1}x, x) \in \mathcal{G}(T)$.

(v) Assume that T is monotone and M^{-1} is β-strongly monotone. Then T_γ^M is $\beta\gamma$-cocoercive.

Proof. (i) Let $x, y \in B$. Indeed,

$$\begin{aligned}
x \in T_{\gamma M} y \Leftrightarrow x &\in \frac{1}{\gamma}\left(M - M \circ J_{\gamma T}^M\right)(y) \\
\Leftrightarrow &\gamma x \in \left(M - M \circ J_{\gamma T}^M\right)(y) \\
\Leftrightarrow &M(y) - \gamma x \in (M \circ J_{\gamma T}^M)(y) \\
\Leftrightarrow &M^{-1}(M(y) - \gamma x) \in J_{\gamma T}^M(y) \\
\Leftrightarrow &y - \gamma M^{-1}(x) \in (M + \gamma T)^{-1} \circ M(y) \\
\Leftrightarrow &(M + \gamma T)(y - \gamma M^{-1}(x)) \in M(y) \\
\Leftrightarrow &M(y) + \gamma T(y) - \gamma x - \gamma^2 T \circ M^{-1}(x) \in M(y) \\
\Leftrightarrow &x \in T(y - \gamma M^{-1}(x)) \qquad (5.1) \\
\Leftrightarrow &x \in (\gamma M^{-1} + T^{-1})^{-1}(y).
\end{aligned}$$

(ii) Let $x, y \in B$. From part (i),

$$x \in T_{\gamma M} y \Leftrightarrow x \in (\gamma M^{-1} + T^{-1})^{-1}(y)$$
$$\Leftrightarrow x \in (M^{-1} + \gamma^{-1} T^{-1})^{-1} \circ M^{-1}(\gamma^{-1} M(y))$$
$$\Leftrightarrow x \in J^{M^{-1}}_{\gamma^{-1} T^{-1}} \circ \gamma^{-1} M(y).$$

(iii) From part(i), for $x, y \in B$,

$$x \in T^M_{(\gamma+\lambda)}(y) \Leftrightarrow x \in ((\gamma+\lambda) M^{-1} + T^{-1})^{-1}(y)$$
$$\Leftrightarrow x \in T(y - (\gamma+\lambda) M^{-1}(x))$$
$$\Leftrightarrow x \in T(y - \gamma M^{-1} y - \lambda M^{-1} y)$$
$$\Leftrightarrow x \in T^M_\gamma (y - \lambda M^{-1}(y))$$
$$\Leftrightarrow x \in \left(T^M_\gamma\right)^M_\lambda y.$$

(iv) For $x, y \in B$ and from (5.1), we have

$$x \in T^M_\gamma \Leftrightarrow x \in T(y - \gamma M^{-1} x) \Leftrightarrow (y - \gamma M^{-1} x, x) \in \mathcal{G}(T). \tag{5.2}$$

(v) For $x_1 = T^M_\gamma y_1, x_2 \in T^M_\gamma y_2 \in B$, from (5.1) and using monotonicity of T, we obtain

$$\langle x_1 - x_2, y_1 - \gamma M^{-1} x - y_2 + \gamma M^{-1} x_2 \rangle \geq 0$$
$$\Rightarrow \langle x_1 - x_2, y_1 - y_2 \rangle \geq \gamma \langle x_1 - x_2, M^{-1} x_1, M^{-1} x_2 \rangle. \tag{5.3}$$

Since M^{-1} is β-strongly monotone, so from (5.3), we deduce that

$$\beta \|x_1 - x_2\|^2 \leq \langle x_1 - x_2, M^{-1} x_1 - M^{-1} x_2 \rangle$$

$$\leq \frac{1}{\gamma}\langle x_1 - x_2, y_1 - y_2\rangle,$$

which implies that

$$\beta\gamma\|x_1 - x_2\|^2 \leq \langle x_1 - x_2, y_1 - y_2\rangle. \tag{5.4}$$

Therefore, T_γ^M is $\beta\gamma$-cocoercive and hence $\frac{1}{\beta\gamma}$-Lipschitz continuous.

□

Proposition 5.2.4. Let $T : B \to 2^{B^*}$ be a maximal monotone operator and $M : B \to B^*$ be an injective operator such that $\operatorname{ran}(M) \subset \operatorname{ran}(T + \gamma M)$ and $T + \gamma M$ is injective for $\gamma > 0$. Then, we have the following:

(i) For $x, y,$ *and* $p \in B$

$$(y, p) = \left(J_{\gamma T}^M, T_\gamma^M\right) \Leftrightarrow \begin{cases} (y, p) \in \mathcal{G}(T), \\ x = y + \gamma M^{-1}p. \end{cases}$$

(ii) Assume that M is α-Lipschitz continuous and β-strongly monotone. Let $S : B \to \mathcal{G}(T)$ be a map defined by

$$x \mapsto (J_{\gamma T}^M x, T_\gamma^M x).$$

Then S is a Lipschitz continuous. Further, if M^{-1} is λ-Lipschitzian, then S^{-1} is also Lipschitz continuous.

Proof. (i) From the definition of warped resolvent and warped Yosida approximation, we have

$$\begin{cases} y = J^M_{\gamma T}, \\ p = T^M_\gamma, \end{cases} \Leftrightarrow \begin{cases} (y, \gamma^{-1}(Mx - My) \in \mathcal{G}(T), \\ p = \frac{1}{\gamma}(Mx - My), \end{cases} \Leftrightarrow \begin{cases} (y, p) \in \mathcal{G}(T), \\ x = y + \gamma M^{-1}p. \end{cases}$$

(ii) Let $x, y \in B$. Using Lipschitz continuity of $J^M_{\gamma T}$ and T^M_γ, we have

$$\begin{aligned} \|Sx - Sy\|^2 &= \|J^M_{\gamma T}x - J^M_{\gamma T}y\|^2 + \|T^M_\gamma x - T^M_\gamma y\|^2 \\ &\leq \frac{\alpha}{\beta}\|x - y\|^2 + \frac{1}{\beta}\|x - y\|^2. \end{aligned}$$

Thus, S is $\sqrt{\frac{\alpha+1}{\beta}}$-Lipschitz continuous.

Conversely, let $(x, p), (y, q) \in \mathcal{G}(T)$. By Cauchy-Schwarz inequality, we deduce that

$$\begin{aligned} \|S^{-1}(x, p) - S^{-1}(y, q)\|^2 &= \|(x - y) + \gamma(M^{-1}p - M^{-1}q)\|^2 \\ &\leq (\|x - y\| + \gamma\|M^{-1}p - M^{-1}q\|)^2 \\ &\leq (\|x - y\| + \gamma\lambda\|p - q\|)^2 \\ &\leq (1 + \gamma^2\lambda^2)(\|x - y\| + \|p - q\|)^2 \\ &\leq (1 + \gamma^2\lambda^2)(\|(x, p) - (y, q)\|^2). \end{aligned}$$

Hence, S^{-1} is $(1 + \gamma^2\lambda^2)$-Lipschitz continuous.

□

Proposition 5.2.5. Let $T : H \to 2^H$ be a maximal monotone operator and $M : H \to H$ be an injective operator such that $\text{ran}(M) \subset \text{ran}(T + \gamma M)$, $\text{ran}(M) \subset \text{ran}(T + \mu M)$, $T + \mu M$ and $T + \gamma M$ are injective for $\mu, \gamma > 0$. Then, for any $x \in H$, we have the following:

(i) $J^M_{\mu T}x\left(\frac{\mu}{\gamma}x+(1-\frac{\mu}{\gamma})J^M_{\gamma T}x\right)=J^M_{\gamma T}x.$

(ii) If M is β-Lipschitz and α-strongly monotone, then

$$\|J^M_{\gamma T}x-J^M_{\mu T}x\|\leq\frac{\beta}{\alpha}\left(1-\frac{\mu}{\gamma}\right)\|M^{-1}T^M_{\gamma}x\|.$$

Proof. (i) For $x\in H$ and $\mu=\lambda\gamma$, we have

$$\begin{aligned}
&x\in((M+\gamma T)^{-1}\circ M)^{-1}\circ(M+\gamma T)^{-1}\circ Mx\\
&\Leftrightarrow x\in M^{-1}\circ(M+\gamma T)\circ J^M_{\gamma T}x\\
&\Leftrightarrow (Mx-M\circ J^M_{\gamma T}x)\in\gamma T(J^M_{\gamma T}x)\\
&\Leftrightarrow \lambda Mx-(1-\lambda)M\circ J^M_{\gamma T}x\in(M+\mu T)J^M_{\gamma T}x\\
&\Leftrightarrow (M+\mu T)^{-1}\circ M(\lambda x+(1-\lambda)J^M_{\gamma T}x)=J^M_{\gamma T}x\\
&\Leftrightarrow J^M_{\mu T}(\lambda x+(1-\lambda)J^M_{\gamma T}x)=J^M_{\gamma T}x.
\end{aligned}$$

Putting $\lambda=\frac{\mu}{\gamma}$, we have

$$J^M_{\mu T}x\left(\frac{\mu}{\gamma}x+(1-\frac{\mu}{\gamma})J^M_{\gamma T}x\right)=J^M_{\gamma T}x.$$

(ii) From (i), we get

$$\begin{aligned}
\|J^M_{\gamma T}x-J^M_{\mu T}x\|&=\|J^M_{\mu T}x\left(\frac{\mu}{\gamma}x+(1-\frac{\mu}{\gamma})J^M_{\gamma T}x\right)-J^M_{\mu T}x\|\\
&\leq\frac{\beta}{\alpha}\left\|\left(1-\frac{\mu}{\gamma}\right)(x-J^M_{\gamma T}x)\right\|\\
&=\frac{\beta}{\alpha}\left(1-\frac{\mu}{\gamma}\right)\|M^{-1}T^M_{\gamma}x.\|
\end{aligned}$$

****

□