Quantitative translations for viscosity approximation methods

Pedro Pinto
(j.w.w. Ulrich Kohlenbach)

Technische Universität Darmstadt
Department of Mathematics

Computability and Complexity in Analysis
26 July, 2021

Outline

Basic notions on mappings

Let X be a Banach space.

- A map $\phi : X \to X$ is a **strict contraction** if

$$\exists r \in [0,1) \, \forall x, y \in X \; (\|\phi(x) - \phi(y)\| \leqslant r\|x - y\|)$$

Theorem (Banach's fixed point theorem)

Let X be a Banach space and ϕ a strict contraction.
Then ϕ has an unique fixed point $z (= \lim \phi^n(z_0))$, for any $z_0 \in X)$.

Basic notions on mappings

Let X be a Banach space.

- A map $\phi : X \to X$ is a **strict contraction** if

$$\exists r \in [0,1)\, \forall x, y \in X \ (\|\phi(x) - \phi(y)\| \leqslant r\|x - y\|)$$

Theorem (Banach's fixed point theorem)

Let X be a Banach space and ϕ a strict contraction.
Then ϕ has an unique fixed point $z(= \lim \phi^n(z_0)$, for any $z_0 \in X)$.

- A map $T : X \to X$ is **nonexpansive** if

$$\forall x, y \in X \ (\|T(x) - T(y)\| \leqslant \|x - y\|)$$

Viscosity approximation method

For C closed and convex, $u, x_0 \in C$, $T : C \to C$ nonexpansive and $(\alpha_n) \subset [0,1]$, consider the **Halpern iteration**

H: $x_0 \in C$ and $x_{n+1} = \alpha_n u + (1-\alpha_n)T(x_n)$, $n \in \mathbb{N}$.

Viscosity approximation method

For C closed and convex, $u, x_0 \in C$, $T : C \to C$ nonexpansive and $(\alpha_n) \subset [0, 1]$, consider the **Halpern iteration**

$$\text{H:} \quad x_0 \in C \ \text{ and } \ x_{n+1} = \alpha_n u + (1 - \alpha_n) T(x_n), \ n \in \mathbb{N}.$$

In connection with concrete problems in convex optimization, Yamada introduced in 2001 the *hybrid steepest descent method* (HSDM). Independently, Moudafi introduced the **viscosity method**

$$\text{vH:} \quad x_0 \in C \ \text{ and } \ x_{n+1} = \alpha_n \phi(x_n) + (1 - \alpha_n) T(x_n), \ n \in \mathbb{N},$$

where ϕ is a strict contraction, and has HSDM as a particular case.

The study of these iterations is highly relevant in many practical situations in optimization.

Iterations with sequences of n.e. maps

Let ϕ be a strict contraction, $(\alpha_n) \subset [0,1]$ a sequence of real numbers, and u an element of C.

For (S_n) a sequence of nonexpansive maps on C, we consider

- the (S_n)-Halpern iteration:

(S_n)-H: $w_0(u) = u$ and $w_{n+1}(u) = \alpha_n u + (1-\alpha_n) S_n(w_n(u))$, $n \in \mathbb{N}$

- the (S_n)-viscosity-Halpern iteration:

(S_n)-vH: $x_0 \in C$, and $x_{n+1} = \alpha_n \phi(x_n) + (1-\alpha_n) S_n(x_n), n \in \mathbb{N}$

Suzuki's theorem

- $(X, C, (S_n), (\alpha_n))$ has the **Halpern's property** if:

For any $u \in C$, the (S_n)-H iteration $(w_n(u))$ is strongly convergent.

Suzuki's theorem

- $(X, C, (S_n), (\alpha_n))$ has the **Halpern's property** if:

For any $u \in C$, the (S_n)-H iteration $(w_n(u))$ is strongly convergent.

E.g., if X is a Hilbert space, C is bounded, and (S_n) is constant, then $(X, C, (S_n), (\frac{1}{n+1}))$ has the Halpern's property.

Suzuki's theorem

- $(X, C, (S_n), (\alpha_n))$ has the **Halpern's property** if:

For any $u \in C$, the (S_n)-H iteration $(w_n(u))$ is strongly convergent.

E.g., if X is a Hilbert space, C is bounded, and (S_n) is constant, then $(X, C, (S_n), (\frac{1}{n+1}))$ has the Halpern's property.

Theorem (Suzuki (2007))

Let $(X, C, (S_n), (\alpha_n))$ have Halpern's property. Let ϕ be a strict contraction, and for each $u \in C$, define $Pu := \lim_n w_n(u)$. If $\lim_n \alpha_n = 0$ and $\sum \alpha_n = \infty$, then the (S_n)-vH iteration strongly converges to the unique point z satisfying $P\phi z = z$.

Rates and metastability

- **Rate of convergence** (to x) is a function $\rho : (0, \infty) \to \mathbb{N}$:

$$\forall \varepsilon > 0 \forall n \geqslant \rho(\varepsilon) \ (\|x_n - x\| \leqslant \varepsilon).$$

- **Cauchy rate** is a function $\rho : (0, \infty) \to \mathbb{N}$:

$$\forall \varepsilon > 0 \forall i, j \geqslant \rho(\varepsilon) \ (\|x_i - x_j\| \leqslant \varepsilon).$$

- **Rate of metastability** is a function $\varphi : (0, \infty) \times \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$:

$$\forall \varepsilon > 0 \, \forall f \in \mathbb{N}^{\mathbb{N}} \, \exists n \leqslant \varphi(\varepsilon, f) \, \forall i, j \in [n, f(n)] \ (\|x_i - x_j\| \leqslant \varepsilon) \, .$$

Rates and metastability

- **Rate of convergence** (to x) is a function $\rho : (0,\infty) \to \mathbb{N}$:

$$\forall \varepsilon > 0 \forall n \geqslant \rho(\varepsilon) \; (\|x_n - x\| \leqslant \varepsilon).$$

- **Cauchy rate** is a function $\rho : (0,\infty) \to \mathbb{N}$:

$$\forall \varepsilon > 0 \forall i,j \geqslant \rho(\varepsilon) \; (\|x_i - x_j\| \leqslant \varepsilon).$$

- **Rate of metastability** is a function $\varphi : (0,\infty) \times \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$:

$$\forall \varepsilon > 0 \, \forall f \in \mathbb{N}^{\mathbb{N}} \, \exists n \leqslant \varphi(\varepsilon, f) \, \forall i,j \in [n, f(n)] \; (\|x_i - x_j\| \leqslant \varepsilon) .$$

- For any $b \in \mathbb{N}^*$, $\theta_b : (0,\infty) \times \mathbb{N}^{\mathbb{N}} \times \mathbb{N} \to \mathbb{N}$ satisfies $(H[S_n])$:

For any $u \in C$ with $\|u\| \leqslant b$,
θ_b is a rate of metastability (with lower bound) for $(w_n(u))$.

Suzuki's proofs

(1) Consider z such that $P\phi z = z$.

(2) Then $z = P\phi z = \lim_n w_n(\phi z)$.

(3) Finally $\lim_n x_n = z$.

Suzuki's proofs

(1) Consider z such that $P\phi z = z$.

(2) Then $z = P\phi z = \lim_n w_n(\phi z)$.

(3) Finally $\lim_n x_n = z$.

From a proof mining perspective, the difficulty in analysing Suzuki's proof is in step (1), the existence of the point z.

The proof-theoretic interpretation of $\Pi^0_3 \to \Pi^0_3$, as in "Cauchy→Cauchy", translates classically (via a negative translation) into "Metastability rate→Metastability rate" and constructively into "Cauchy rate→Cauchy rate".

Why $z = P\phi z$ exists?

$$\begin{aligned} P \text{ nonexpansive } + \phi \text{ strict contraction} &\implies P\phi \text{ strict contraction} \\ &\implies \exists ! z \in C \ (z = P\phi(z)) \end{aligned}$$

Why $z = P\phi z$ exists?

$$P \text{ nonexpansive } + \phi \text{ strict contraction} \implies P\phi \text{ strict contraction}$$
$$\implies \exists! z \in C \; (z = P\phi(z))$$

Furthermore, the fixed point is the limit of the Picard iteration:

$$z = \lim (P\phi)^n(z_0), \quad \text{for any } z_0 \in C.$$

$$z_0 \in C \longrightarrow \phi z_0 \longrightarrow \lim w_n(\phi z_0) = P\phi z_0 = z_1$$

$$z_1 \longrightarrow \phi z_1 \longrightarrow \lim w_n(\phi z_1) = P\phi z_1 = (P\phi)^2 z_0 = z_2$$

$$\vdots$$

$$z = \lim z_n = \lim (P\phi)^n z_0$$

$$z_0 \in C \longrightarrow \phi z_0 \longrightarrow \lim w_n(\phi z_0) = P\phi z_0 = z_1$$
$$z_1 \longrightarrow \phi z_1 \longrightarrow \lim w_n(\phi z_1) = P\phi z_1 = (P\phi)^2 z_0 = z_2$$
$$\vdots$$
$$z = \lim z_n = \lim (P\phi)^n z_0$$

- Instead of using the limit point of $w_n(\phi z_i)$, we just take a 'good enough' Cauchy point.
- We make this construction long enough to ensure stability for some point z_M, i.e. such that $\|z_{M+1} - z_M\| \leqslant \varepsilon$.
- We then adapt the arguments to use such z_M instead of the fixed point z

Quantitative results

Theorem

Let $b \geqslant diam(C)$, ϕ *an r-contraction with* $\delta \in (0,1]$ *s.t.* $r \leqslant 1-\delta$, θ_b *satisfy* $(H[S_n])$, *and* A *be a rate of divergence for* $(\sum \alpha_n)$. *Then* (x_n) *satisfying* (S_n)*-vH is a Cauchy sequence with metastability rate*

$$\Psi(\varepsilon, f) = \Psi[b, \delta, A, \theta_b](\varepsilon, f) = \widetilde{A}\left(\psi_{M+1} + \left\lceil \log \frac{6b}{\varepsilon} \right\rceil\right) + 1$$

where $\Psi_0 = A(1) + 1$ *and* $\Psi_{m+1} = \theta_b(\varepsilon_0, f_{M-m}, \Psi_m)$, *with*

$$\widetilde{A}(k) := A\left(\left\lceil \frac{k}{\delta} \right\rceil\right), \ \varepsilon_0 = \frac{\varepsilon\delta^2}{30}, \ M = \left\lceil \log_{1-\frac{\delta}{2}}\left(\frac{\varepsilon\delta}{15b}\right) \right\rceil \ \textit{and}$$

$$\textit{for all } p \in \mathbb{N}, \begin{cases} f_0(p) = f\left(\widetilde{A}(p + \lceil \log \frac{6b}{\varepsilon} \rceil) + 1\right), \\ f_{m+1}(p) = \max\{f_0(p), \theta_b(\varepsilon_0, f_m, p)\} \quad \textit{for } m < M. \end{cases}$$

Corollary

Let now ρ be a common Cauchy rate for the T-Halpern iterations $(w_n(u))$, for all $u \in C$. Then

$$\Psi(\varepsilon, f) = \Psi[b, \delta, A, \rho](\varepsilon, f) = \sigma\left(\varepsilon, \max\{\rho(\varepsilon_0), A(1) + 1\}\right)$$

is a Cauchy rate for (x_n), where $\sigma(\varepsilon, p) := \widetilde{A}(p + \lceil \log \frac{6b}{\varepsilon} \rceil) + 1$ with $\widetilde{A}$, b, and ε_0 are as before,

Corollary

Let now ρ be a common Cauchy rate for the T-Halpern iterations $(w_n(u))$, for all $u \in C$. Then

$$\Psi(\varepsilon, f) = \Psi[b, \delta, A, \rho](\varepsilon, f) = \sigma\left(\varepsilon, \max\{\rho(\varepsilon_0), A(1) + 1\}\right)$$

is a Cauchy rate for (x_n), where $\sigma(\varepsilon, p) := \widetilde{A}(p + \lceil \log \frac{6b}{\varepsilon} \rceil) + 1$ with $\widetilde{A}$, b, and ε_0 are as before,

The quantitative result "Cauchy rate→Cauchy rate" shows that Suzuki's arguments are essentially constructive, which was not obvious.

Corollary

Let now ρ be a common Cauchy rate for the T-Halpern iterations $(w_n(u))$, for all $u \in C$. Then

$$\Psi(\varepsilon, f) = \Psi[b, \delta, A, \rho](\varepsilon, f) = \sigma\left(\varepsilon, \max\{\rho(\varepsilon_0), A(1) + 1\}\right)$$

is a Cauchy rate for (x_n), where $\sigma(\varepsilon, p) := \widetilde{A}(p + \lceil \log \frac{6b}{\varepsilon} \rceil) + 1$ with $\widetilde{A}$, b, and ε_0 are as before,

The quantitative result "Cauchy rate→Cauchy rate" shows that Suzuki's arguments are essentially constructive, which was not obvious.

We have shown that this analysis largely holds in a more general geodesic setting (‘*generalization of proofs*’)

Some Applications

- [Kohlenbach, Sipoș (2020)], using the results from [Kohlenbach, Leuștean (2012)], obtained a rate of metastability for the Halpern iteration in Banach spaces which are both uniformly smooth and uniformly convex.
- [Kohlenbach, Leuștean (2012)] extracted a rate of metastability in $\mathrm{CAT}(0)$.

Hence, one obtains rates of metastability for the viscosity version.

In Hilbert spaces, Bauschke generalized Halpern's iteration in a different way.
Let $T_1, \cdots, T_N$ be nonexpansive maps on C. Bauschke proved the convergence of (S_n)-H, for $S_n = T_{[n+1]}$, where $[n] = n \mod N$.

- In [Ferreira, Leuștean, P. (2019)], an analysis of Bauschke's proof allowed for the extraction of a rate of metastability for Bauschke's iteration (S_n)-H.

Hence, one obtains a rate of metastability for the viscosity version.

Körnlein gave a quantitative analysis of viscosity-Bauschke. Although, the rates are obtained through different methods, they are of similar complexity.

The Halpern-type Proximal Point Algorithm, a strongly convergent version of the well-known proximal point algorithm, is defined by

$$\text{HPPA: } x_{n+1} = \alpha_n u + (1-\alpha_n) J_{\gamma_n}(x_n),\ n \in \mathbb{N}$$

where J_{γ_n} are resolvent functions of a monotone operator.

- In [P. (2021)], the convergence of HPPA under conditions that allow for the natural choice of parameters was studied, and rates of metastability were obtained.

Hence, one obtains a rate of metastability for the viscosity-HPPA

$$\text{vHPPA: } x_{n+1} = \alpha_n \phi(x_n) + (1-\alpha_n) J_{\gamma_n}(x_n),\ n \in \mathbb{N}$$

In [Kohlenbach (2020)], metastability for HPPA in Banach spaces which are both uniformly smooth and uniformly convex under very general conditions.

Modulus of uniqueness

Suppose that $T : C \to C$ has at most one fixed point, i.e.

$$(1)\, \forall p_1, p_2 \in C\, (p_1 = Tp_1 \wedge p_2 = Tp_2 \to p_1 = p_2).$$

We say that T has uniformly at most one fixed point with **modulus of uniqueness** $\omega : (0, \infty) \to (0, \infty)$ if

$$(2)\, \forall \varepsilon > 0\, \forall p_1, p_2 \in C$$
$$(\|p_1 - Tp_1\|, \|p_2 - Tp_2\| \leqslant \omega(\varepsilon) \to \|p_1 - p_2\| \leqslant \varepsilon).$$

Modulus of uniqueness

Suppose that $T : C \to C$ has at most one fixed point, i.e.

$$(1)\, \forall p_1, p_2 \in C\, (p_1 = Tp_1 \wedge p_2 = Tp_2 \to p_1 = p_2).$$

We say that T has uniformly at most one fixed point with **modulus of uniqueness** $\omega : (0, \infty) \to (0, \infty)$ if

$$(2)\, \forall \varepsilon > 0\, \forall p_1, p_2 \in C$$
$$(\|p_1 - Tp_1\|, \|p_2 - Tp_2\| \leqslant \omega(\varepsilon) \to \|p_1 - p_2\| \leqslant \varepsilon).$$

If T is continuous and C is compact, (1) implies the existence of a modulus ω such that (2), but in general (2) is stronger than (1).

From uniqueness to Cauchy rates

- (x_n) is *asymptotically regular* (with respect to a map T) if
$$\lim \|x_n - T(x_n)\| = 0.$$
- In many cases, it is possible to obtain a **rate of asymptotic regularity**, i.e. a function $\mu : (0, \infty) \to (0, \infty)$ satisfying

$$\forall \varepsilon > 0 \forall n \geqslant \mu(\varepsilon) \ (\|x_n - T(x_n)\| \leqslant \varepsilon) \,.$$

From uniqueness to Cauchy rates

- (x_n) is *asymptotically regular* (with respect to a map T) if
$$\lim \|x_n - T(x_n)\| = 0.$$
- In many cases, it is possible to obtain a **rate of asymptotic regularity**, i.e. a function $\mu : (0, \infty) \to (0, \infty)$ satisfying

$$\forall \varepsilon > 0 \forall n \geqslant \mu(\varepsilon) \left(\|x_n - T(x_n)\| \leqslant \varepsilon \right).$$

- If ω is a modulus of uniqueness and μ is a rate of asymptotic regularity, then we have a Cauchy rate:

$$\forall \varepsilon > 0 \forall i, j \geqslant \mu(\omega(\varepsilon)) \left(\|x_i - x_j\| \leqslant \varepsilon \right).$$

Let X be a uniformly convex normed space.
Let T be a nonexpansive map on C such that $A := \mathrm{Id} - T$ is uniformly accretive operator in the sense of [Gwinner (1978)]:
there is a strictly increasing function $\varphi : [0, \infty) \to \mathbb{R}$ such that $\lim_{t\to\infty} \varphi(t) = \infty$ and

$$\forall x, y \in C \, \exists j \in J(x - y) \\ (\langle A(x) - A(y), j\rangle \geqslant (\varphi(\|x\|) - \varphi(\|y\|)) \cdot (\|x\| - \|y\|)),$$

where J is the normalized duality map of X, i.e.

$$J(x) := \{j \in X^* \, : \, \langle x, j\rangle = \|j\|^2 = \|x\|^2\}.$$

Let X be a uniformly convex normed space.
Let T be a nonexpansive map on C such that $A := \mathrm{Id} - T$ is uniformly accretive operator in the sense of [Gwinner (1978)]:
there is a strictly increasing function $\varphi : [0, \infty) \to \mathbb{R}$ such that $\lim_{t\to\infty} \varphi(t) = \infty$ and

$$\forall x, y \in C\, \exists j \in J(x-y) \\ \left(\langle A(x) - A(y), j\rangle \geqslant (\varphi(\|x\|) - \varphi(\|y\|)) \cdot (\|x\| - \|y\|)\right),$$

where J is the normalized duality map of X, i.e.

$$J(x) := \{j \in X^* \; : \langle x, j\rangle = \|j\|^2 = \|x\|^2\}.$$

There is a function $\Omega : (0, \infty)^2 \to (0, \infty)$ satisfying

$$\forall x, y \in C\, \forall K, \varepsilon > 0\, \exists j \in J(x-y) \\ \Big(\|x\|, \|y\| \leqslant K \wedge \big|\|x\| - \|y\|\big| \geqslant \varepsilon \to \langle A(x) - A(y), j\rangle \geqslant \Omega(\varepsilon, K)\Big),$$

Proposition

Let $b \in \mathbb{N}^*$ be such that $\forall x \in C\ (\|x\| \leqslant b)$, and write $\Omega(\varepsilon, b) = \Omega(\varepsilon)$. Let η be a modulus of uniform convexity. The function

$$\omega_b(\varepsilon) := \frac{1}{16b}\Omega\left(\frac{\varepsilon}{2} \cdot \eta\left(\frac{\varepsilon}{b}\right)\right) \cdot \eta\left(\frac{1}{8b^2}\Omega\left(\frac{\varepsilon}{2} \cdot \eta\left(\frac{\varepsilon}{b}\right)\right)\right)$$

is a modulus of uniqueness for being a fixed point of T.

Proposition

Let $b \in \mathbb{N}^*$ be such that $\forall x \in C$ $(\|x\| \leqslant b)$, and write $\Omega(\varepsilon, b) = \Omega(\varepsilon)$. Let η be a modulus of uniform convexity. The function

$$\omega_b(\varepsilon) := \frac{1}{16b}\Omega\left(\frac{\varepsilon}{2}\cdot\eta\left(\frac{\varepsilon}{b}\right)\right)\cdot\eta\left(\frac{1}{8b^2}\Omega\left(\frac{\varepsilon}{2}\cdot\eta\left(\frac{\varepsilon}{b}\right)\right)\right)$$

is a modulus of uniqueness for being a fixed point of T.

Using a rate from [Kohlenbach, Leuștean (2012)],

Theorem (Cauchy rate for Halpern)

For $x_0, u \in C$, consider $x_{n+1} = \frac{1}{n+1}u + (1 - \frac{1}{n+1})T(x_n)$. Then (x_n) is a Cauchy sequence with Cauchy rate $\mu(\omega_b(\varepsilon))$, where

$$\mu(\varepsilon) := \left\lceil \frac{4b}{\varepsilon} + \frac{32b^2}{\varepsilon^2} \right\rceil.$$

Final remarks

- For X complete and C closed, we have rate of convergence.

Final remarks

- For X complete and C closed, we have rate of convergence.
- The 'viscosity' results were proven in the general setting of W-hyperbolic spaces, for the more general Rakotch maps instead of strict contractions, and also for Browder-type sequences.
- This allows for additional applications using previous proof mining studies (e.g. in CAT(0) spaces for Browder-type iterations).

Final remarks

- For X complete and C closed, we have rate of convergence.
- The 'viscosity' results were proven in the general setting of W-hyperbolic spaces, for the more general Rakotch maps instead of strict contractions, and also for Browder-type sequences.
- This allows for additional applications using previous proof mining studies (e.g. in CAT(0) spaces for Browder-type iterations).
- We also considered algorithms with error terms.

Final remarks

- For X complete and C closed, we have rate of convergence.
- The 'viscosity' results were proven in the general setting of W-hyperbolic spaces, for the more general Rakotch maps instead of strict contractions, and also for Browder-type sequences.
- This allows for additional applications using previous proof mining studies (e.g. in CAT(0) spaces for Browder-type iterations).
- We also considered algorithms with error terms.
- Our proofs correspond to quantitative versions of the original arguments, and actually establish the convergence to the desirable limit point.

References

T. Suzuki. *Moudafi's viscosity approximations with Meir-Keeler contractions*, ***Journal of Mathematical Analysis and Applications***, 325(1):342–352, 2007.

U. Kohlenbach and A. Sipoș. *The finitary content of sunny nonexpansive retractions*, To appear in: ***Communications in Contemporary Mathematics***, arXiv:1812.04940, 2020.

U. Kohlenbach and L. Leuștean. *On the computational content of convergence proofs via Banach limits*, ***Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences***, 370.1971: 3449-3463, 2012.

U. Kohlenbach and L. Leuștean. *Effective metastability of Halpern iterates in CAT(0) spaces*, ***Advances in Mathematics***, 231.5: 2526-2556, 2012.

F. Ferreira, L. Leuștean and P. Pinto, *On the removal of weak compactness arguments in proof mining*, ***Advances in Mathematics***, vol. 354, 2019.

P. Pinto, *A rate of metastability for the Halpern type Proximal Point Algorithm*, ***Numerical Functional Analysis and Optimization***, 42(3): 320-343, 2021.

U. Kohlenbach, *Quantitative analysis of a Halpern-type Proximal Point Algorithm for accretive operators in Banach spaces*, ***Journal of Nonlinear and Convex Analysis***, 21(9):2125–2138, 2020.

J. Gwinner, *On the convergence of some iteration processes in uniformly convex Banach spaces*, ***Proceedings of the American Mathematical Society***, 71(1), 29-35, 1978.

Thank you