The continuous steepest descent method with convex-like potential

Simon K. Niederländer

Ingolstadt University of Applied Sciences, 85049 Ingolstadt, Germany (E-mail: simon.niederlaender@thi.de)

Abstract: In a real Hilbert space setting, we investigate the asymptotic properties of the solutions of the classical continuous steepest descent method with convex-like potential. Despite the absence of convexity, we show that the solutions preserve the remarkable minimizing properties typically associated with convex functions. In particular, we find that the values of the convex-like potential decay asymptotically at a sublinear rate. If, moreover, the potential function is weakly lower semi-continuous, we prove that the solutions weakly converge toward a minimizer. Under a quadratic growth condition on the convex-like potential, we further provide a strong convergence result on the solutions along with a linear decay rate of the function values. Numerical experiments illustrate our theoretical findings.

Keywords: Continuous steepest descent; weak quasi-convexity; Lyapunov analysis

1. INTRODUCTION

Let $\mathcal{H}$ be a real Hilbert space endowed with inner product $\langle \cdot\,, \cdot \rangle$ and associated norm $\|\cdot\|$. Consider the minimization problem

$$\inf \{f(x) : x \in \mathcal{H}\}, \tag{P}$$

where $f : \mathcal{H} \to \mathbb{R}$ is a continuously differentiable real-valued function. The classical continuous steepest descent method associated with (P) consists of the first-order evolution equation

$$\dot{x} + \nabla f(x) = 0 \tag{SD}$$

with $\nabla f : \mathcal{H} \to \mathcal{H}$ denoting the (Fréchet) gradient of f. We say that $x : [0, +\infty[\to \mathcal{H}$ is a (classical) solution of (SD) if $x \in \mathcal{C}^1([0, +\infty[; \mathcal{H})$ such that it verifies (SD) on $[0, +\infty[$.

The minimizing properties of (SD) relative to f are underlined by the fact that the solutions $x(t)$ of (SD) satisfy, for every $t \geq 0$, the "descent property"

$$\frac{\mathrm{d}}{\mathrm{d}t} f(x(t)) + \|\dot{x}(t)\|^2 = 0.$$

Hence, as long as the solutions of (SD) do not reach the critical points of f, that is, $\operatorname{crit} f := \{x \in \mathcal{H} : \nabla f(x) = 0\}$, the mapping

$$t \longmapsto f(x(t))$$

is decreasing. Assuming now that the function f is minorized, that is, $\inf_{\mathcal{H}} f > -\infty$, and that ∇f is Lipschitz continuous on the bounded subsets of $\mathcal{H}$, the Cauchy problem associated with (SD) is well posed and the solutions $x : [0, +\infty[\to \mathcal{H}$ of (SD) satisfy the "finite-energy property"

$$\int_0^\infty \|\dot{x}(\tau)\|^2 \,\mathrm{d}\tau < +\infty.$$

Moreover, the bounded solutions of (SD) are such that

$$\lim_{t \to +\infty} \nabla f(x(t)) = 0.$$

Consequently, if there exists $\bar{x} \in \mathcal{H}$ such that $x(t_n) \to \bar{x}$ strongly in $\mathcal{H}$ as $n \to +\infty$ for some sequence $t_n \to +\infty$, then $\bar{x} \in \operatorname{crit} f$. However, without any additional assumptions on f, the bounded solutions $x(t)$ of (SD) may fail to converge as $t \to +\infty$; see Palis and de Melo (1982) for a counterexample.

In contrast, the solutions of (SD) are known to inherit remarkable minimizing properties whenever the function f is convex. Indeed, assuming that f is convex, the solutions of (SD) are minimizing in the sense that (see Brézis (1971, 1973))

$$\lim_{t \to +\infty} f(x(t)) = \inf_{\mathcal{H}} f.$$

If, in addition, the set $\operatorname{argmin}_{\mathcal{H}} f$ is nonempty, then the solutions of (SD) obey the asymptotic estimate (see, again, Brézis (1971, 1973))

$$f(x(t)) - \min_{\mathcal{H}} f = \mathcal{O}\Big(\frac{1}{t}\Big) \text{ as } t \to +\infty.$$

In this case, the solutions $x(t)$ of (SD) are weakly convergent, as $t \to +\infty$, toward an element of $\operatorname{argmin}_{\mathcal{H}} f$; cf. Bruck (1975). On the other hand, if the solutions of (SD) are known to be strongly convergent, then the above estimate improves to (see Güler (2005))

$$f(x(t)) - \min_{\mathcal{H}} f = \mathcal{o}\Big(\frac{1}{t}\Big) \text{ as } t \to +\infty.$$

Note, however, that the weakly convergent solutions of (SD) need not be strongly convergent; see Baillon (1978) for a counterexample.

The above estimates have further been developed in the recent work of Attouch et al. (2024). Assuming again that f is convex such that $\operatorname{argmin}_{\mathcal{H}} f$ is nonempty, the solutions $x : [0, +\infty[\to \mathcal{H}$ of (SD) satisfy, in fact, the refined integral estimate

$$\int_0^\infty \tau \|\dot{x}(\tau)\|^2 \,\mathrm{d}\tau < +\infty.$$

Moreover, the asymptotic estimate

$$f(x(t)) - \min_{\mathcal{H}} f = \mathcal{o}\Big(\frac{1}{t}\Big) \text{ as } t \to +\infty$$

holds independent of the strong convergence assumption. In addition, the following estimate on the velocity $\dot{x}(t)$ of a (SD) solution is verified:

$$\|\dot{x}(t)\| = o\Big(\frac{1}{t}\Big) \text{ as } t \to +\infty;$$

we refer the reader to Niederländer (2021, 2023) (see also Battahi et al. (2024)) for the respective results on the "generalized steepest descent method" in the context of linearly constrained convex minimization.

In this work, we aim at deriving asymptotic estimates on the (SD) solutions based on a convexity-like assumption on the potential function f. In particular, we assume that the continuously differentiable function $f : \mathcal{H} \to \mathbb{R}$ is such that there exist $\bar{x} \in \operatorname{argmin}_{\mathcal{H}} f$ and $\theta > 0$ with

$$\forall x \in \mathcal{H}, \quad \langle \nabla f(x), x - \bar{x}\rangle \geq \theta\big(f(x) - \min\nolimits_{\mathcal{H}} f\big). \tag{C}$$

Condition (C) can be viewed as a generalization of the notion of convexity. Indeed, if f is continuously differentiable and convex, then the above condition is satisfied with $\theta = 1$ for every $\bar{x} \in \mathcal{H}$. On the other hand, condition (C) also generalizes Euler's identity for homogeneous functions. The above inequality has already been investigated by Cabot et al. (2009) (see also Su et al. (2016), Aujol et al. (2019), Sebbouh et al. (2020)) in the context of second-order evolution equations solving the minimization problem (P). In particular, it has been shown by Su et al. (2016) that every continuously differentiable function $f : \mathcal{H} \to \mathbb{R}$ with the property that

$$x \longmapsto \sqrt[\theta]{f(x) - \min\nolimits_{\mathcal{H}} f}$$

is convex for some $\theta > 0$ satisfies condition (C). More recently, Aujol et al. (2019) (see also Sebbouh et al. (2020)) provided a geometric interpretation of (a local version of) condition (C) in terms of a "flatness property" of the function f in the neighborhood of its minimizers; we also refer the reader to Cabot et al. (2009) for particular examples of functions f which satisfy condition (C).

The main contribution of this work is to extend the minimizing properties of the classical continuous steepest descent method (SD), established in the convex case, to convex-like functions f satisfying condition (C). In particular, we show that the solutions $x(t)$ of (SD) with convex-like potential f evolve according to the estimate

$$\limsup_{t \to +\infty} t\big(f(x(t)) - \min\nolimits_{\mathcal{H}} f\big) < +\infty,$$

suggesting that $f(x(t)) - \min_{\mathcal{H}} f$ behaves as $\mathcal{O}(1/t)$ as $t \to +\infty$. More precisely, we show that the solutions $x(t)$ of (SD) obey, as in the convex case, the estimate

$$f(x(t)) - \min\nolimits_{\mathcal{H}} f = o\Big(\frac{1}{t}\Big) \text{ as } t \to +\infty.$$

Moreover, the refined integral estimate

$$\int_0^\infty \tau \|\nabla f(x(\tau))\|^2 \, \mathrm{d}\tau < +\infty$$

remains valid. If, in addition, the convex-like function f is weakly lower semi-continuous, we show that there exists $\bar{x} \in \operatorname{argmin}_{\mathcal{H}} f$ such that

$$\operatorname{w-}\lim_{t \to +\infty} x(t) = \bar{x}.$$

Finally, under the additional assumption that the convex-like potential f satisfies a "quadratic growth condition", that is, there exist $\bar{x} \in \operatorname{argmin}_{\mathcal{H}} f$ and $\alpha > 0$ such that

$$\forall x \in \mathcal{H}, \quad f(x) - \min\nolimits_{\mathcal{H}} f \geq \alpha \|x - \bar{x}\|^2, \tag{G}$$

we show that the solutions $x(t)$ of (SD) strongly converge, as $t \to +\infty$, toward the unique element of $\operatorname{argmin}_{\mathcal{H}} f$. In this case, we further derive the estimates

$$\begin{aligned} f(x(t)) - \min\nolimits_{\mathcal{H}} f &= \mathcal{O}\big(\mathrm{e}^{-2\alpha\theta t}\big) \text{ as } t \to +\infty; \\ \|x(t) - \bar{x}\|^2 &= \mathcal{O}\big(\mathrm{e}^{-2\alpha\theta t}\big) \text{ as } t \to +\infty; \\ \|\dot{x}(t)\|^2 &= \mathcal{O}\big(\mathrm{e}^{-2\alpha\theta t}\big) \text{ as } t \to +\infty. \end{aligned}$$

A simple yet representative numerical example illustrates our findings.

2. PRELIMINARY FACTS

In the following, we presuppose that

- (A1) $f : \mathcal{H} \to \mathbb{R}$ is continuously differentiable and minorized, i.e., $\inf_{\mathcal{H}} f > -\infty$;
- (A2) $\nabla f : \mathcal{H} \to \mathcal{H}$ is Lipschitz continuous on bounded sets.

Consider again the first-order evolution equation

$$\dot{x} + \nabla f(x) = 0 \tag{SD}$$

with initial data $x_0 \in \mathcal{H}$. Following Haraux (1991), we say that $x : [0, +\infty[\, \to \mathcal{H}$ is a (classical) solution of the above Cauchy problem if $x \in \mathcal{C}^1([0, +\infty[; \mathcal{H})$ such that it verifies (SD) on $[0, +\infty[$ with $x(0) = x_0$. Equivalently, the mapping $x : [0, +\infty[\, \to \mathcal{H}$ is a solution of (SD) if it is continuous such that

$$x(t) + \int_0^t \nabla f(x(\tau)) \, \mathrm{d}\tau = x_0, \quad t \geq 0.$$

Let us begin our discussion by recalling some preliminary facts on the (SD) differential system. The following results are outlined, for instance, in Chill and Fašangová (2010).

Theorem 2.1. For every $x_0 \in \mathcal{H}$ there exists a unique solution $x : [0, +\infty[\, \to \mathcal{H}$ of (SD). Moreover,

- (i) $t \mapsto f(x(t))$ is non-increasing and
$$\frac{\mathrm{d}}{\mathrm{d}t} f(x(t)) + \|\dot{x}(t)\|^2 = 0, \quad t \geq 0;$$
- (ii) it holds that
$$\int_0^\infty \|\dot{x}(\tau)\|^2 \, \mathrm{d}\tau < +\infty.$$

The above result identifies the correspondence

$$x \longmapsto f(x)$$

as a Lyapunov function for the continuous steepest descent (SD) whose decay property will be essential for the analysis of the asymptotic behavior of its solutions. Indeed, assuming the boundedness of the (SD) solutions, we have the following asymptotic properties.

Proposition 2.2. Let $x : [0, +\infty[\, \to \mathcal{H}$ be a bounded solution of (SD). Then it holds that

$$\lim_{t \to +\infty} \dot{x}(t) = \lim_{t \to +\infty} \nabla f(x(t)) = 0.$$

Moreover, if there exists $\bar{x} \in \mathcal{H}$ such that $x(t_n) \to \bar{x}$ strongly in $\mathcal{H}$ as $n \to +\infty$ for some sequence $t_n \to +\infty$, then $\bar{x} \in \operatorname{crit} f$.

Remark 2.3. We note that the solutions of (SD) remain bounded, for instance, whenever f is coercive, that is,

$$\lim_{\|x\| \to +\infty} f(x) = +\infty.$$

Indeed, it suffices to observe from Theorem 2.1(i) that for every $t \geq 0$, it holds that

$$f(x(t)) \leq f(x(0)).$$

This majorization together with the fact that f is coercive clearly implies that $x(t)$ remains bounded on $[0, +\infty[$.

3. THE (SD) EVOLUTION EQUATION

Let us now investigate the (SD) evolution equation under the assumption that there exist $\bar{x} \in \operatorname{argmin}_{\mathcal{H}} f$ and $\theta > 0$ such that, for every $x \in \mathcal{H}$, it holds that

$$\langle \nabla f(x), x - \bar{x} \rangle \geq \theta\big(f(x) - \min_{\mathcal{H}} f\big). \qquad \text{(C)}$$

Condition (C) clearly implies that every critical point of f is a minimizer. Indeed, assuming that $x \in \operatorname{crit} f$, we readily obtain

$$0 \geq f(x) - \min_{\mathcal{H}} f,$$

implying that $x \in \operatorname{argmin}_{\mathcal{H}} f$. Note, however, that f may admit multiple (non-isolated) minima. Yet another consequence of condition (C) is that, for every $x \in \mathcal{H}$ and $\lambda \in [0, 1]$, it holds that

$$f(\bar{x} + \lambda(x - \bar{x})) - \min_{\mathcal{H}} f \leq \lambda^{\theta}\big(f(x) - \min_{\mathcal{H}} f\big),$$

suggesting that the function value contracts by at most λ^{θ} when contracting the line segment from x to $\bar{x}$ by a factor λ. To see this, let $x \in \mathcal{H}$ and take $\lambda \in [0, 1]$. Upon defining the auxiliary function $\phi : [0, +\infty[\to [0, +\infty[$ by

$$\phi(\lambda) = f(\bar{x} + \lambda(x - \bar{x})) - \min_{\mathcal{H}} f,$$

in view of condition (C), we obtain

$$\lambda\phi'(\lambda) \geq \theta\phi(\lambda), \quad \lambda \in]0, 1].$$

An immediate integration then shows

$$\lambda^{\theta}\phi(1) \geq \phi(\lambda).$$

Let us start our discussion by deriving some basic properties of the (SD) solutions under condition (C).

Theorem 3.1. Let $f : \mathcal{H} \to \mathbb{R}$ satisfy condition (C) with $\theta > 0$ and let $x : [0, +\infty[\to \mathcal{H}$ be a solution of (SD). Then, for every $\bar{x} \in \operatorname{argmin}_{\mathcal{H}} f$, the following assertions hold:

(i) $t \mapsto \|x(t) - \bar{x}\|^2$ is non-increasing and
$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|x(t) - \bar{x}\|^2 + \theta\big(f(x(t)) - \min_{\mathcal{H}} f\big) \leq 0, \quad t \geq 0;$$
(ii) $\lim_{t\to+\infty}\|x(t) - \bar{x}\|$ exists;
(iii) $x \in \mathcal{L}^{\infty}([0, +\infty[; \mathcal{H})$;
(iv) it holds that
$$\int_0^{\infty} f(x(\tau)) - \min_{\mathcal{H}} f \,\mathrm{d}\tau < +\infty.$$

Proof. Let $\bar{x} \in \operatorname{argmin}_{\mathcal{H}} f$. Taking the inner product with $x(t) - \bar{x}$ in (SD) and subsequently applying the chain rule, for every $t \geq 0$, we have

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|x(t) - \bar{x}\|^2 + \langle \nabla f(x(t)), x(t) - \bar{x} \rangle = 0.$$

In view of condition (C), we obtain

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|x(t) - \bar{x}\|^2 + \theta\big(f(x(t)) - \min_{\mathcal{H}} f\big) \leq 0.$$

Owing to the fact that $f(x(t)) - \min_{\mathcal{H}} f \geq 0$, it readily follows that

$$t \longmapsto \|x(t) - \bar{x}\|^2$$

is non-increasing. This quantity being bounded from below, we infer that

$$\lim_{t\to+\infty}\|x(t) - \bar{x}\|^2 \text{ exists},$$

implying that $x(t)$ remains bounded on $[0, +\infty[$. On the other hand, integrating the above inequality over $[0, t]$ entails

$$\frac{1}{2}\|x(t) - \bar{x}\|^2 + \theta\int_0^t f(x(\tau)) - \min_{\mathcal{H}} f \,\mathrm{d}\tau \leq \frac{1}{2}\|x(0) - \bar{x}\|^2.$$

Taking into account that $\|x(t) - \bar{x}\|^2 \geq 0$ and subsequently dividing by $\theta > 0$ gives

$$\int_0^t f(x(\tau)) - \min_{\mathcal{H}} f \,\mathrm{d}\tau \leq \frac{1}{2\theta}\|x(0) - \bar{x}\|^2.$$

This majorization being valid for every $t \geq 0$, passing to the limit as $t \to +\infty$ yields the desired conclusion.

As an immediate consequence of the above result together with the decay property of the mapping

$$t \longmapsto f(x(t)),$$

we have the following asymptotic estimate.

Corollary 3.2. Under the hypotheses of Theorem 3.1, the following assertion holds:

$$f(x(t)) - \min_{\mathcal{H}} f = o\Big(\frac{1}{t}\Big) \text{ as } t \to +\infty$$

Proof. In view of Theorem 2.1(i), we readily observe that the mapping

$$t \longmapsto f(x(t)) - \min_{\mathcal{H}} f$$

is non-increasing. Hence, for every $t \geq 0$, it holds that

$$\int_{t/2}^t f(x(\tau)) - \min_{\mathcal{H}} f \,\mathrm{d}\tau \geq \frac{t}{2}\big(f(x(t)) - \min_{\mathcal{H}} f\big).$$

Owing to the fact that $f(x) - \min_{\mathcal{H}} f \in \mathcal{L}^1([0, +\infty[; \mathbb{R})$, we classically deduce that the above integral vanishes as $t \to +\infty$ and thus,

$$\lim_{t\to+\infty} t\big(f(x(t)) - \min_{\mathcal{H}} f\big) = 0,$$

concluding the desired estimate.

Remark 3.3. We emphasize that the above asymptotic estimate clearly implies that

$$\lim_{t\to+\infty} f(x(t)) = \min_{\mathcal{H}} f.$$

We leave the details to the reader.

Let us now investigate the convergence properties of the solutions of (SD). To this end, we assume, in addition to condition (C), that the function f is weakly (sequentially) lower semi-continuous, that is, for every sequence $(x_n)_{n\in\mathbb{N}}$ and $x \in \mathcal{H}$ such that $x_n \rightharpoonup x$ weakly in $\mathcal{H}$ as $n \to +\infty$, it holds that

$$\liminf_{n\to+\infty} f(x_n) \geq f(x).$$

Given this additional assumption, we have the following weak convergence result.

Proposition 3.4. Let $f : \mathcal{H} \to \mathbb{R}$ be weakly lower semi-continuous such that condition (C) holds. Let $x : [0, +\infty[\to \mathcal{H}$ be a solution of (SD). Then there exists $\bar{x} \in \operatorname{argmin}_{\mathcal{H}} f$ such that

$$\mathrm{w}\text{-}\lim_{t\to+\infty} x(t) = \bar{x}.$$

Proof. Let $\bar{x} \in \operatorname{argmin}_{\mathcal{H}} f$ and recall from Theorem 3.1 (ii) that

$$\lim_{t\to+\infty} \|x(t)-\bar{x}\| \text{ exists.}$$

In view of Lemma A.1, it suffices to show that every weak sequential cluster point of $(x(t))_{t\geq 0}$ belongs to $\operatorname{argmin}_{\mathcal{H}} f$. Let $x \in \mathcal{H}$ and suppose that $x(t_n) \rightharpoonup x$ weakly in $\mathcal{H}$, as $n \to +\infty$, for a sequence $t_n \to +\infty$. By virtue of Corollary 3.2 and the weak lower semi-continuity of f, we have

$$\begin{aligned}\min_{\mathcal{H}} f &= \lim_{n\to+\infty} f(x(t_n)) \\ &= \liminf_{n\to+\infty} f(x(t_n)) \\ &\geq f(x),\end{aligned}$$

implying that $x \in \operatorname{argmin}_{\mathcal{H}} f$. We conclude by applying Lemma A.1 to the set $\operatorname{argmin}_{\mathcal{H}} f$.

Remark 3.5. We note that every convex function $f : \mathcal{H} \to \mathbb{R}$ is weakly lower semi-continuous; see, e.g., Ekeland and Témam (1999), Bauschke and Combettes (2017). On the other hand, when $\mathcal{H}$ is finite-dimensional, the notions of weak and (strong) lower semi-continuity coincide.

Let us conclude this section with a result on the strong convergence of the solutions of (SD) under the additional assumption that there exist $\bar{x} \in \operatorname{argmin}_{\mathcal{H}} f$ and $\alpha > 0$ such that

$$\forall x \in \mathcal{H}, \quad f(x) - \min_{\mathcal{H}} f \geq \alpha \|x-\bar{x}\|^2. \tag{G}$$

Condition (G) may be interpreted as a "quadratic growth condition" which clearly implies that $\bar{x}$ is the unique element of $\operatorname{argmin}_{\mathcal{H}} f$.

Proposition 3.6. Under the hypotheses of Theorem 3.1, suppose that $f : \mathcal{H} \to \mathbb{R}$ satisfies condition (G). Then, for $\bar{x} \in \operatorname{argmin}_{\mathcal{H}} f$, it holds that

$$\lim_{t\to+\infty} x(t) = \bar{x}.$$

Proof. Let $\bar{x} \in \operatorname{argmin}_{\mathcal{H}} f$ and recall from Corollary 3.2 that

$$\lim_{t\to+\infty} f(x(t)) = \min_{\mathcal{H}} f.$$

In view of condition (G), we readily obtain

$$\lim_{t\to+\infty} \|x(t)-\bar{x}\|^2 = 0,$$

concluding the result.

4. ASYMPTOTIC ESTIMATES

In this section, we aim at deriving explicit decay rate estimates for the (SD) evolution equation. In particular, we show that the solutions $x(t)$ of (SD) obey, under condition (C), the refined integral estimate

$$\int_0^\infty \tau\|\dot{x}(\tau)\|^2 \,\mathrm{d}\tau < +\infty.$$

The technique we use to prove this fact essentially relies on a combination of the previous arguments.

Theorem 4.1. Let $f : \mathcal{H} \to \mathbb{R}$ satisfy condition (C) with $\theta > 0$ and let $x : [0, +\infty[\to \mathcal{H}$ be a solution of (SD). Then, for every $\bar{x} \in \operatorname{argmin}_{\mathcal{H}} f$, the following assertions hold:

(i) $t \mapsto \|x(t)-\bar{x}\|^2/2\theta + t\big(f(x(t)) - \min_{\mathcal{H}} f\big)$ is non-increasing and

$$\frac{1}{2\theta}\frac{\mathrm{d}}{\mathrm{d}t}\|x(t)-\bar{x}\|^2 + \frac{\mathrm{d}}{\mathrm{d}t}\Big(t\big(f(x(t)) - \min_{\mathcal{H}} f\big)\Big) + t\|\dot{x}(t)\|^2 \leq 0, \quad t \geq 0;$$

(ii) it holds that

$$\limsup_{t\to+\infty} t\big(f(x(t)) - \min_{\mathcal{H}} f\big) < +\infty;$$

(iii) it holds that

$$\int_0^\infty \tau\|\dot{x}(\tau)\|^2 \,\mathrm{d}\tau < +\infty.$$

Proof. Let $\bar{x} \in \operatorname{argmin}_{\mathcal{H}} f$ and let $\rho : [0, +\infty[\to]0, +\infty[$ be some continuously differentiable function to be chosen. Taking the inner product with $\dot{x}(t) + \rho(t)(x(t)-\bar{x})$ in (SD), for every $t \geq 0$, we have

$$\frac{\rho(t)}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|x(t)-\bar{x}\|^2 + \langle \nabla f(x(t)), \dot{x}(t)\rangle + \rho(t)\langle \nabla f(x(t)), x(t)-\bar{x}\rangle + \|\dot{x}(t)\|^2 = 0.$$

Dividing by $\rho(t)$ and using the chain rule yields

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|x(t)-\bar{x}\|^2 + \frac{1}{\rho(t)}\frac{\mathrm{d}}{\mathrm{d}t}\big(f(x(t)) - \min_{\mathcal{H}} f\big) + \langle \nabla f(x(t)), x(t)-\bar{x}\rangle + \frac{1}{\rho(t)}\|\dot{x}(t)\|^2 = 0.$$

In view of the basic identity

$$\frac{1}{\rho(t)}\frac{\mathrm{d}}{\mathrm{d}t}\big(f(x(t)) - \min_{\mathcal{H}} f\big) = \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{1}{\rho(t)}\big(f(x(t)) - \min_{\mathcal{H}} f\big)\Big) - \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{1}{\rho(t)}\Big)\big(f(x(t)) - \min_{\mathcal{H}} f\big),$$

the above equality reads as

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|x(t)-\bar{x}\|^2 + \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{1}{\rho(t)}\big(f(x(t)) - \min_{\mathcal{H}} f\big)\Big) - \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{1}{\rho(t)}\Big)\big(f(x(t)) - \min_{\mathcal{H}} f\big) + \frac{1}{\rho(t)}\|\dot{x}(t)\|^2 + \langle \nabla f(x(t)), x(t)-\bar{x}\rangle = 0.$$

Upon applying condition (C), we obtain

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|x(t)-\bar{x}\|^2 + \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{1}{\rho(t)}\big(f(x(t)) - \min_{\mathcal{H}} f\big)\Big) + \Big(\theta - \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{1}{\rho(t)}\Big)\Big)\big(f(x(t)) - \min_{\mathcal{H}} f\big) + \frac{1}{\rho(t)}\|\dot{x}(t)\|^2 \leq 0.$$

Let us now choose $\rho(t) = 1/\theta t$ such that, for every $t > 0$, it holds that

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{1}{\rho(t)}\Big) = \theta.$$

Consequently, we have

$$\frac{1}{2\theta}\frac{\mathrm{d}}{\mathrm{d}t}\|x(t)-\bar{x}\|^2 + \frac{\mathrm{d}}{\mathrm{d}t}\Big(t\big(f(x(t)) - \min_{\mathcal{H}} f\big)\Big) + t\|\dot{x}(t)\|^2 \leq 0.$$

Taking into account that $t\|\dot{x}(t)\|^2 \geq 0$, we immediately obtain that

$$t \longmapsto \frac{1}{2\theta}\|x(t)-\bar{x}\|^2 + t\big(f(x(t)) - \min_{\mathcal{H}} f\big)$$

is non-increasing. Moreover, integrating the above inequality over $[0, t]$ entails

$$\frac{1}{2\theta}\|x(t)-\bar{x}\|^2 + t\big(f(x(t)) - \min_{\mathcal{H}} f\big) + \int_0^t \tau\|\dot{x}(\tau)\|^2 \,\mathrm{d}\tau \leq \frac{1}{2\theta}\|x(0)-\bar{x}\|^2.$$

Since we have both $\|x(t)-\bar{x}\|^2 \geq 0$ and $t\|\dot{x}(t)\|^2 \geq 0$, we deduce that

$$t\big(f(x(t)) - \min_{\mathcal{H}} f\big) \leq \frac{1}{2\theta}\|x(0)-\bar{x}\|^2.$$

Passing to the upper limit as $t \to +\infty$ ensures

$$\limsup_{t\to+\infty} t\big(f(x(t)) - \min_{\mathcal{H}} f\big) < +\infty.$$

On the other hand, using that $\|x(t)-\bar{x}\|^2/2\theta + t\big(f(x(t)) - \min_{\mathcal{H}} f\big) \geq 0$, we readily obtain

$$\int_0^t \tau\|\dot{x}(\tau)\|^2 \,\mathrm{d}\tau \leq \frac{1}{2\theta}\|x(0)-\bar{x}\|^2.$$

This majorization being valid for every $t \geq 0$, taking the supremum yields

$$\int_0^\infty \tau\|\dot{x}(\tau)\|^2 \,\mathrm{d}\tau < +\infty,$$

concluding the result.

Remark 4.2. Theorem 4.1(ii) asserts that the solutions $x(t)$ of (SD) obey the asymptotic estimate

$$f(x(t)) - \min_{\mathcal{H}} f = \mathcal{O}\Big(\frac{1}{t}\Big) \text{ as } t \to +\infty.$$

In fact, as previously shown in Corollary 3.2, the above estimate improves to $o(1/t)$ as $t \to +\infty$.

We conclude this section by deriving fast asymptotic estimates for the (SD) solutions under the assumption that the convex-like potential $f : \mathcal{H} \to \mathbb{R}$ satisfies again condition (G).

Proposition 4.3. Under the hypotheses of Theorem 4.1, suppose that $f : \mathcal{H} \to \mathbb{R}$ satisfies condition (G) with $\alpha > 0$. Then, for $\bar{x} \in \operatorname{argmin}_{\mathcal{H}} f$, it holds that

$$\begin{aligned} f(x(t)) - \min_{\mathcal{H}} f &= \mathcal{O}\big(\,\mathrm{e}^{-2\alpha\theta t}\big) \text{ as } t \to +\infty;\\ \|x(t)-\bar{x}\|^2 &= \mathcal{O}\big(\,\mathrm{e}^{-2\alpha\theta t}\big) \text{ as } t \to +\infty;\\ \|\dot{x}(t)\|^2 &= \mathcal{O}\big(\,\mathrm{e}^{-2\alpha\theta t}\big) \text{ as } t \to +\infty. \end{aligned}$$

Proof. Let $\bar{x} \in \operatorname{argmin}_{\mathcal{H}} f$ and recall from Theorem 3.1(i) that, for every $t \geq 0$, we have

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|x(t)-\bar{x}\|^2 + \theta\big(f(x(t)) - \min_{\mathcal{H}} f\big) \leq 0.$$

In view of condition (G), we obtain

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|x(t)-\bar{x}\|^2 + \alpha\theta\|x(t)-\bar{x}\|^2 \leq 0.$$

Multiplying the above inequality by $\mathrm{e}^{2\alpha\theta t}$ gives

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\big(\,\mathrm{e}^{2\alpha\theta t}\|x(t)-\bar{x}\|^2\big) \leq 0.$$

Successively integrating over $[0,t]$ and dividing by $\mathrm{e}^{2\alpha\theta t}/2$ entails

$$\|x(t)-\bar{x}\|^2 \leq \mathrm{e}^{-2\alpha\theta t}\|x(0)-\bar{x}\|^2.$$

On the other hand, by virtue of condition (C) and the Cauchy–Schwarz inequality, we have

$$f(x(t)) - \min_{\mathcal{H}} f \leq \frac{1}{\theta}\|\nabla f(x(t))\|\|x(t)-\bar{x}\|.$$

Observing that ∇f is Lipschitz continuous on the bounded subsets of $\mathcal{H}$, there exists $L \geq 0$ such that

$$\|\nabla f(x(t))\| \leq L\|x(t)-\bar{x}\|.$$

Combining the above inequalities yields

$$f(x(t)) - \min_{\mathcal{H}} f \leq \frac{L}{\theta}\|x(t)-\bar{x}\|^2$$

and thus,

$$f(x(t)) - \min_{\mathcal{H}} f \leq \frac{L}{\theta}\,\mathrm{e}^{-2\alpha\theta t}\|x(0)-\bar{x}\|^2.$$

Moreover, in view of (SD) and the above derivations, we readily obtain

$$\|\dot{x}(t)\| \leq L\,\mathrm{e}^{-\alpha\theta t}\|x(0)-\bar{x}\|,$$

concluding the desired estimates.

5. NUMERICAL EXPERIMENT

In this section, we provide a simple, yet representative, example that allows for a direct exposition of our main results.

Example 5.1. Let $\mathcal{H} = \mathbb{R}^2$ and consider the convex-like potential $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} \phi(x)\|x\|^2, & x \neq 0,\\ 0, & x = 0 \end{cases}$$

with $\phi \in \mathcal{C}^1(\mathbb{R}^2 \setminus \{0\};\mathbb{R})$ taking the form $\phi(x) = (1/2 + \varepsilon\cos(k\arg x))$ for some $\varepsilon \in \,]0,1/2[$ and $k \geq 3$. Here, $\arg x$ denotes the polar angle of $x \neq 0$, that is, $x = r(\cos\vartheta, \sin\vartheta)$ with $r = \|x\|$ and $\vartheta = \arg x \in \,]-\pi,\pi]$. Clearly, f admits the unique minimizer $\bar{x} = 0$ with $\min_{\mathbb{R}^2} f = 0$. Moreover, for every $x \neq 0$, it holds that $\nabla f(x) = 2\phi(x)x + \|x\|^2\nabla\phi(x)$ with

$$\nabla\phi(x) = \frac{1}{\|x\|}\partial_\vartheta\phi(x)(-\sin\vartheta, \cos\vartheta)$$

and $\partial_\vartheta\phi(x) = -\varepsilon k\sin(k\vartheta)$ for $x = r(\cos\vartheta,\sin\vartheta)$. Consequently, for every $x \neq 0$, we have

$$\begin{aligned} \|\nabla f(x)\| &= \|x\|\sqrt{\big(2\phi(x)\big)^2 + \big(\partial_\vartheta\phi(x)\big)^2}\\ &\leq \|x\|\sqrt{(1+2\varepsilon)^2 + (\varepsilon k)^2}, \end{aligned}$$

which clearly implies that $\nabla f(x)$ tends to zero as $x \to 0$. Upon defining $\nabla f(0) := 0$, we infer that ∇f is continuous (in fact, Lipschitz continuous on bounded sets) and thus $f \in \mathcal{C}^1(\mathbb{R}^2;\mathbb{R})$.

On the other hand, f is positively 2-homogeneous, that is, $f(\lambda x) = \lambda^2 f(x)$ holds for every $x \in \mathbb{R}^2$ and $\lambda > 0$. As f is continuously differentiable, Euler's identity gives, for every $x \in \mathbb{R}^2$,

$$\langle\nabla f(x), x-\bar{x}\rangle = 2f(x),$$

implying that condition (C) holds with equality and $\theta = 2$. Moreover, since $\phi(x) \geq 1/2 - \varepsilon$ for every $x \in \mathbb{R}^2$, we have

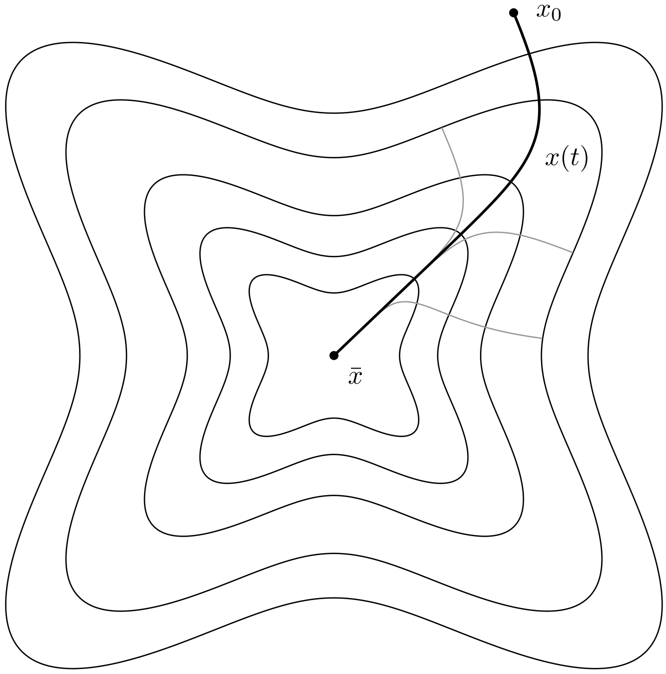
$$f(x) \geq \Big(\frac{1}{2} - \varepsilon\Big)\|x-\bar{x}\|^2,$$

so that condition (G) is verified with $\alpha = 1/2 - \varepsilon$. Figure 1 depicts the evolution of a solution $x(t)$ of the (SD) differential system for the convex-like potential f, with parameters $\varepsilon = 1/4$ and $k = 4$, and initial data $x_0 = (1,2)$.

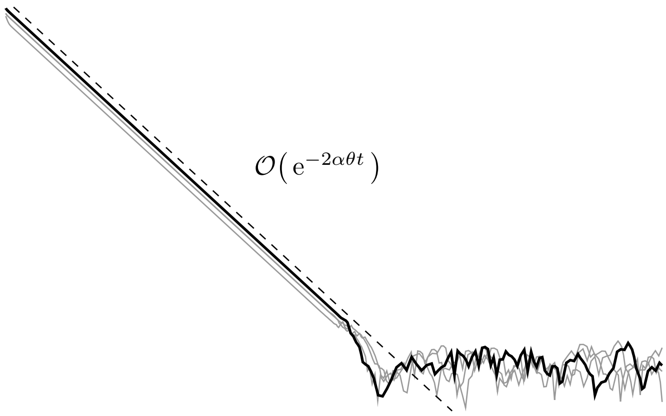
Analyzing Figure 1, we observe that the solutions $x(t)$ of (SD) converge, as $t \to +\infty$, to the unique minimizer $\bar{x}$ of f; cf. Proposition 3.4. Moreover, since f verifies both conditions (C) and (G), we find that $\|x(t)-\bar{x}\|^2$ evolves according to the estimate $\mathcal{O}\big(\,\mathrm{e}^{-2\alpha\theta t}\big)$ as $t \to +\infty$; cf. Proposition 4.3. A similar behavior is observed numerically for the quantities $f(x(t)) - \min_{\mathbb{R}^2} f$ and $\|\dot{x}(t)\|^2$.

APPENDIX

For the following classical result, named the Opial lemma, the reader is referred to Opial (1967).



(a) Trajectories $\{x(t) : t \geq 0\}$



(b) Evolution of $\|x(t) - \bar{x}\|^2/2$

Fig. 1. (a) Graphical view on the trajectories of solutions $x(t)$ of (SD) for different initial conditions, overlaid with contour lines of the sublevel sets of f.
(b) Semilogarithmic view on $\|x(t)-\bar{x}\|^2/2$ for the corresponding initial data together with the asymptotic estimate $\mathcal{O}\big(\mathrm{e}^{-2\alpha\theta t}\big)$ as $t \to +\infty$.

Lemma A.1. (Opial). Let $\mathcal{H}$ be a real Hilbert space and let $x : [0, +\infty[\to \mathcal{H}$ be such that there exists a nonempty subset S of $\mathcal{H}$ which verifies

(i) for every $\bar{x} \in S$, $\lim_{t\to+\infty}\|x(t) - \bar{x}\|$ exists;
(ii) $\forall t_n \to +\infty$ such that $x(t_n) \rightharpoonup \bar{x}$ weakly in $\mathcal{H}$ as $n \to +\infty$, it holds that $\bar{x} \in S$.

Then $x(t)$ converges weakly, as $t \to +\infty$, to some element $\bar{x} \in S$.

REFERENCES

Attouch, H., Boţ, R.I., and Nguyen, D.K. (2024). Fast convex optimization via time scale and averaging of the steepest descent. *Math. Oper. Res.* Available at `https://doi.org/10.1287/moor.2023.0186`.

Aujol, J.F., Dossal, C., and Rondepierre, A. (2019). Optimal convergence rates for Nesterov acceleration. *SIAM J. Optim.*, 29, 3131–3153.

Baillon, J.B. (1978). Un exemple concernant le comportement asymptotique de la solution du problème $\mathrm{d}u/\mathrm{d}t + \partial\phi(u) \ni 0$. *J. Funct. Anal.*, 28, 369–376.

Battahi, F., Chbani, Z., Niederländer, S.K., and Riahi, H. (2024). Asymptotic behavior of the Arrow–Hurwicz differential system with Tikhonov regularization. Available at `https://arxiv.org/pdf/2411.17656`.

Bauschke, H.H. and Combettes, P.L. (2017). *Convex analysis and monotone operator theory in Hilbert spaces.* CMS Books in Mathematics. Springer, New York.

Brézis, H. (1971). Monotonicity methods in Hilbert spaces and some applications to nonlinear partial differential equations. In E. Zarantonello (ed.), *Contributions to Nonlinear Functional Analysis*, 101–156. Acad. Press, Madison.

Brézis, H. (1973). *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert.* Mathematics Studies 5. North-Holland, Amsterdam.

Bruck, R.E. (1975). Asymptotic convergence of nonlinear contraction semi-groups in Hilbert spaces. *J. Funct. Anal.*, 18, 15–26.

Cabot, A., Engler, H., and Gadat, S. (2009). On the long time behavior of second order differential equations with asymptotically small dissipation. *Trans. Amer. Math. Soc.*, 361, 5983–6017.

Chill, R. and Fašangová, E. (2010). *Gradient systems.* Lecture Notes of the 13th International Internet Seminar. Matfyzpress, Prague.

Ekeland, I. and Témam, R. (1999). *Convex analysis and variational problems.* Classics in Applied Mathematics. SIAM, Philadelphia.

Güler, O. (2005). Convergence rate estimates for the gradient differential inclusion. *Optim. Methods Softw.*, 20, 729–735.

Haraux, A. (1991). *Systèmes dynamiques dissipatifs et applications.* Recherches en Mathématiques Appliquées. Masson, Paris.

Niederländer, S.K. (2021). Second-order dynamics with Hessian-driven damping for linearly constrained convex minimization. *SIAM J. Control Optim.*, 59, 3708–3736.

Niederländer, S.K. (2023). On the Arrow–Hurwicz differential system for linearly constrained convex minimization. *Optimization*, 73, 2313–2345.

Opial, Z. (1967). Weak convergence of the sequence of successive approximations for nonexpansive mappings. *Bull. Amer. Math. Soc.*, 73, 591–597.

Palis, J. and de Melo, W. (1982). *Geometric theory of dynamical systems.* Springer, New York.

Sebbouh, O., Dossal, C., and Rondepierre, A. (2020). Convergence rates of damped inertial dynamics under geometric conditions and perturbations. *SIAM J. Optim.*, 30, 1850–1877.

Su, W., Boyd, S., and Candès, E.J. (2016). A differential equation for modeling Nesterov's accelerated gradient method. *J. Mach. Learn. Res.*, 17, 5312–5354.