ESI **The Erwin Schrödinger International Institute for Mathematical Physics**
Boltzmanngasse 9
A-1090 Wien, Austria

A Generic Convergence Theorem for Continuous Descent Methods in Banach Spaces

Sergiu Aizicovici
Simeon Reich
Alexander J. Zaslavski

Vienna, Preprint ESI 1772 (2006)

February 2, 2006

Supported by the Austrian Federal Ministry of Education, Science and Culture
Available via http://www.esi.ac.at

A GENERIC CONVERGENCE THEOREM FOR CONTINUOUS DESCENT METHODS IN BANACH SPACES

Sergiu Aizicovici, Simeon Reich and Alexander J. Zaslavski

Abstract. We study continuous descent methods for minimizing convex functions defined on general Banach spaces and prove that most of them (in the sense of Baire category) converge.

1. Introduction

In this paper we continue our studies of descent methods. This is an important topic in optimization theory and in dynamical systems; see, for example, [1-3, 6-11]. Given a continuous convex function f on a Banach space X, we associate with f a complete metric space of vector fields $V : X \to X$ such that $f^0(x, Vx) \le 0$ for all $x \in X$. Here $f^0(x, u)$ is the right-hand derivative of f at x in the direction of $u \in X$. In [2] we studied the convergence of the values of the function f to its minimum along the trajectories of continuous dynamical systems governed by such vector fields and established a convergence result for most of them. Here by "most" we mean an everywhere dense G_δ subset of the space of vector fields (cf., for instance, [4, 5, 8-10, 12]). In [2] we considered a class of vector fields which are Lipschitz on bounded subsets. We assumed there that the convex function f has a unique point of minimum, and moreover, that the minimization problem for the function f on X is well-posed. In the present paper, we obtain a generic convergence result for a class of vector fields which are only locally Lipschitz and bounded on bounded subsets of X, and for a convex function f which has a (not necessarily unique) point of minimum. We equip the space of these vector fields with a natural complete metric and show that the function f tends to its minimum along any trajectory of the dynamical system determined by a generic pair consisting of an initial condition and a vector field.

More precisely, let $(X, ||\cdot||)$ be a Banach space and let $f : X \to R^1$ be a convex continuous function which satisfies the following conditions:

C(i)

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

C(ii) there is $\bar{x} \in X$ such that $f(\bar{x}) \le f(x)$ for all $x \in X$.

Note that if X is reflexive, then C(i) implies C(ii).

1991 *Mathematics Subject Classification.* 37L99, 47J35, 49M99, 54E35, 54E50, 54E52, 90C25.

Key words and phrases. Complete uniform space, convex function, descent method, generic property, initial value problem, minimization problem.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

For each $x \in X$, let

$$f^0(x,u) = \lim_{t \to 0^+} [f(x+tu) - f(x)]/t, \ u \in X. \tag{1.1}$$

For each $x \in X$ and $r > 0$, set

$$B(x,r) = \{z \in X : \ ||z-x|| \le r\} \text{ and } B(r) = B(0,r). \tag{1.2}$$

For each mapping $A : X \to X$ and each $r > 0$, put

$$\mathrm{Lip}(A,r) = \sup\{||Ax - Ay||/||x-y|| : \ x, y \in B(r) \text{ and } x \neq y\}. \tag{1.3}$$

Denote by $\mathcal{A}$ the set of all locally Lipschitz mappings $V : X \to X$ which are bounded on bounded subsets of X and satisfy the inequality $f^0(x, Vx) \le 0$ for all $x \in X$.

For the set $\mathcal{A}$ we consider the uniformity determined by the base

$$E_s(n,\epsilon) = \{(V_1,V_2) \in \mathcal{A} \times \mathcal{A} : \ \mathrm{Lip}(V_1 - V_2, n) \le \epsilon$$

$$\text{and } ||V_1x - V_2x|| \le \epsilon \text{ for all } x \in B(n)\}. \tag{1.4}$$

Clearly, this uniform space $\mathcal{A}$ is metrizable and complete. The topology induced by this uniformity in $\mathcal{A}$ will be called the strong topology.

We also equip the space $\mathcal{A}$ with the uniformity determined by the base

$$E_w(n,\epsilon) = \{(V_1,V_2) \in \mathcal{A} \times \mathcal{A} : \ ||V_1x - V_2x|| \le \epsilon$$

$$\text{for all } x \in B(n)\} \tag{1.5}$$

where $n, \epsilon > 0$. The topology induced by this uniformity will be called the weak topology.

The product space $X \times \mathcal{A}$ is equipped with a pair of topologies: a weak topology which is the product of the norm topology of X and the weak topology of $\mathcal{A}$, and a strong topology which is the product of the norm topology of X and the strong topology of $\mathcal{A}$.

Our paper is organized as follows. The main result, Theorem 2.1, is stated in the next section. An auxiliary result, Proposition 3.1, is presented in Section 3. The proof of our main result is carried out in Section 4.

2. The main result

In the proof of our main result, we are going to use the following existence result, the proof of which is almost identical with the proof of Proposition 1.1 in [2].

Proposition 2.1. *Let $x_0 \in X$ and $V \in \mathcal{A}$. Then there exists a unique continuously differentiable mapping $x : [0,\infty) \to X$ such that*

$$x'(t) = Vx(t), \ t \in [0,\infty),$$

$$x(0) = x_0.$$

Now we are ready to state the main result of our paper.

Theorem 2.1. *There exists a set $\mathcal{F} \subset X \times \mathcal{A}$ which is a countable intersection of open (in the weak topology) everywhere dense (in the strong topology) subsets of $X \times \mathcal{A}$ such that for each pair $(z, V) \in \mathcal{F}$ and each $\epsilon > 0$, the following property holds:*

There are a neighborhood $\mathcal{U}$ of (z, V) in $X \times \mathcal{A}$ with the weak topology and $T_0 > 0$ such that for each $(\xi, W) \in \mathcal{U}$ and each continuously differentiable mapping $y : [0, \infty) \to X$ satisfying

$$y(0) = \xi, \ y'(t) = Wy(t) \text{ for all } t \geq 0,$$

the inequality $f(y(t)) \leq f(\bar{x}) + \epsilon$ holds for all $t \geq T_0$.

3. An auxiliary result

We let $x \in W^{1,1}(0, T; X)$, i.e.,

$$x(t) = x_0 + \int_0^t u(s)ds, \ t \in [0, T],$$

where $T > 0$, $x_0 \in X$, and $u \in L^1(0, T; X)$. Then $x : [0, T] \to X$ is absolutely continuous and $x'(t) = u(t)$ for a.e. $t \in [0, T]$. Recall that the function $f : X \to R^1$ is assumed to be convex and continuous and therefore it is, in fact, locally Lipschitz. It follows that its restriction to the set $\{x(t) : \ t \in [0, T]\}$ is Lipschitz. Indeed, since the set $\{x(t) : \ t \in [0, T]\}$ is compact, the closure of its convex hull C is both compact and convex, and so the restriction of f to C is Lipschitz.

Hence the function $(f \circ x)(t) := f(x(t))$, $t \in [0, T]$, is absolutely continuous. It follows that for almost every $t \in [0, T]$, both the derivatives $x'(t)$ and $(f \circ x)'(t)$ exist:

$$x'(t) = \lim_{h \to 0} h^{-1}[x(t + h) - x(t)],$$

$$(f \circ x)'(t) = \lim_{h \to 0} h^{-1}[f(x(t + h)) - f(x(t))].$$

We now recall Proposition 3.1 in [11].

Proposition 3.1. *Assume that $t \in [0, T]$ and that both the derivatives $x'(t)$ and $(f \circ x)'(t)$ exist. Then*

$$(f \circ x)'(t) = \lim_{h \to 0} h^{-1}[f(x(t) + hx'(t)) - f(x(t))].$$

4. Proof of Theorem 2.1

For each $V \in \mathcal{A}$ and each $\gamma \in (0, 1)$, set

$$V_\gamma x = Vx + \gamma(\bar{x} - x), \ x \in X. \tag{4.1}$$

We first state four lemmata that can be proved by essentially the same arguments as those in [2].

Lemma 4.1. *Let $V \in \mathcal{A}$ and $\gamma \in (0, 1)$. Then $V_\gamma \in \mathcal{A}$.*

Lemma 4.2. *Let* $V \in \mathcal{A}$. *Then* $\lim_{\gamma \to 0^+} V_\gamma = V$ *in the strong topology.*

Lemma 4.3. *Let* $V \in \mathcal{A}$, $\gamma \in (0,1)$, $\epsilon > 0$, *and let* $x \in X$ *satisfy* $f(x) \geq f(\bar{x}) + \epsilon$. *Then* $f^0(x, V_\gamma x) \leq -\gamma\epsilon$.

Lemma 4.4. *Let* $V \in \mathcal{A}$, $\gamma \in (0,1)$, *and let* $x \in C^1([0,\infty); X)$ *satisfy*

$$x'(t) = V_\gamma x(t), \ t \in [0, \infty). \tag{4.2}$$

Assume that $T_0, \epsilon > 0$ *are such that*

$$T_0 > (f(x(0)) - f(\bar{x}))(\gamma\epsilon)^{-1}. \tag{4.3}$$

Then for each $t \geq T_0$,

$$f(x(t)) \leq f(\bar{x}) + \epsilon. \tag{4.4}$$

The next two lemmata play a key role in the proof of our theorem.

Lemma 4.5. *Let* $V \in \mathcal{A}$, $x_0 \in X$, $T > 0$, $\epsilon > 0$, *and let*

$$x \in C^1([0,T]; X) \tag{4.5}$$

satisfy

$$x'(t) = Vx(t), \ t \in [0, \infty), \tag{4.6}$$

and

$$x(0) = x_0. \tag{4.7}$$

Then there exists a neighborhood $\mathcal{U}$ *of* (x_0, V) *in* $X \times \mathcal{A}$ *with the weak topology such that for each* $(\xi, W) \in \mathcal{U}$ *and each* $y \in C^1([0,T]; X)$ *satisfying*

$$y(0) = \xi, \ y'(t) = Wy(t), \ t \in [0, \infty), \tag{4.8}$$

the following inequality holds:

$$||y(t) - x(t)|| \leq \epsilon, \ t \in [0, T]. \tag{4.9}$$

Proof. Denote by K the closure of the convex hull of $\{x(t) : \ t \in [0,T]\}$. It is clear that K is convex and compact. It is also not difficult to see that there is $r > 0$ such that V is Lipschitz on $B(z, r)$ for all $z \in K$.

We may assume that

$$f(z) \leq f(x_0) + 1 \text{ for all } z \in B(x_0, r). \tag{4.10}$$

Since K is compact, there are points $z_1, \ldots, z_n \in K$ such that

$$K \subset \cup_{i=1}^n B(z_i, r/2).$$

Then

$$E := \cup\{B(z, r/4) : z \in K\} \subset \cup_{i=1}^{n} B(z_i, 3r/4). \tag{4.11}$$

Clearly, E is a closed and convex set. By the choice of r, there is $L_1 > 0$ such that for each $i = 1, \ldots, n$,

$$||Vy_1 - Vy_2|| \le L_1||y_1 - y_2|| \text{ for all } y_1, y_2 \in B(z_i, r). \tag{4.12}$$

We claim that for each $y_1, y_2 \in E$,

$$||Vy_1 - Vy_2|| \le L_1||y_1 - y_2||. \tag{4.13}$$

Indeed, let $y_1, y_2 \in E$ with $y_1 \ne y_2$. To prove (4.13), put

$$\Omega = \{S \in [0, 1] : ||Vy_1 - V(ty_2 + (1 - t)y_1)|| \le tL_1||y_1 - y_2|| \text{ for all } t \in [0, S]\}. \tag{4.14}$$

First, we show that $\Omega \ne \emptyset$. By (4.11), there is $j \in \{1, \ldots, n\}$ such that

$$y_1 \in B(z_j, (3/4)r). \tag{4.15}$$

Clearly, there is $S_0 \in (0, 1)$ such that

$$ty_2 + (1 - t)y_1 \in B(z_j, r) \text{ for all } t \in [0, S_0].$$

When combined with (4.12) and (4.14), this relation implies that $S_0 \in \Omega$. Set $S_1 = \sup \Omega$. It is clear that $S_1 \in \Omega$. Next, we claim that $S_1 = 1$. Assume that $S_1 < 1$. Since the set E is convex, it follows from (4.11) that there is $p \in \{1, \ldots, n\}$ such that

$$S_1y_2 + (1 - S_1)y_1 \in B(z_p, (3/4)r). \tag{4.16}$$

Hence there is $S_2 \in (S_1, 1)$ such that

$$ty_2 + (1 - t)y_1 \in B(z_p, r) \text{ for all } t \in [S_1, S_2]. \tag{4.17}$$

By (4.17) and (4.12), we have for all $t \in [S_1, S_2]$,

$$||V(S_1y_2 + (1 - S_1)y_1) - V(ty_2 + (1 - t)y_1)|| \le (t - S_1)L_1||y_2 - y_1||.$$

In conjunction with the inclusion $S_1 \in \Omega$, this inequality implies that

$$||Vy_1 - V(ty_2 + (1 - t)y_1)|| \le tL_1||y_1 - y_2|| \text{ for all } t \in [0, S_2].$$

Thus $S_2 \in \Omega$, a contradiction. The contradiction we have reached proves that $S_1 = 1$, as claimed, and thus (4.13) is true.

By C(i), there is $n_0 > 0$ such that

$$\text{if } z \in X \text{ and } f(z) \le |f(x_0)| + 2, \text{ then } ||z|| \le n_0. \tag{4.18}$$

Choose a positive number δ such that

$$\delta(T+1)e^{L_1 T} < \min\{r/4,\ \epsilon/2\} \tag{4.19}$$

and set

$$\mathcal{U} = B(x_0, \delta) \times \{W \in \mathcal{A} :\ ||Wz - Vz|| \leq \delta \ \text{ for all } z \in B(n_0)\}. \tag{4.20}$$

Assume that

$$(y_0, W) \in \mathcal{U},\ y \in C^1([0,T]; X), \tag{4.21}$$

and

$$y(0) = y_0,\ y'(t) = Wy(t),\ t \in [0,T]. \tag{4.22}$$

By Proposition 3.1, (4.22), (4.6), and the inclusions $V, W \in \mathcal{A}$, the composite functions $(f \circ y)(t)$, $t \in [0,T]$, and $(f \circ x)(t)$, $t \in [0,T]$, are decreasing, so that

$$f(y(t)) \leq f(y(0)) = f(y_0),\ t \in [0,T],\ f(x(t)) \leq f(x(0)) = f(x_0),\ t \in [0,T]. \tag{4.23}$$

Relations (4.23) and (4.18) imply that

$$||x(t)|| \leq n_0,\ t \in [0,T]. \tag{4.24}$$

By (4.23), (4.21), (4.20), (4.19), and (4.10), we have for all $t \in [0,T]$,

$$f(y(t)) \leq f(y(0)) = f(y_0) \leq f(x_0) + 1. \tag{4.25}$$

Relations (4.25) and (4.18) imply that

$$||y(t)|| \leq n_0,\ t \in [0,T]. \tag{4.26}$$

Put

$$\Omega_0 = \{S \in [0,T] :\ ||y(t) - x(t)|| \leq \delta(T+1)e^{L_1 T},\ t \in [0,S]\}. \tag{4.27}$$

By (4.27), (4.22), (4.21), (4.20), and (4.6), $\Omega_0 \neq \emptyset$. Set

$$\tau_1 = \sup \Omega_0. \tag{4.28}$$

Clearly,

$$\tau_1 \in \Omega_0. \tag{4.29}$$

We now show that $\tau_1 = T$. Assume by way of contradiction that $\tau_1 < T$. In view of (4.29), (4.27) and (4.19), there holds for all $t \in [0, \tau_1]$,

$$||y(t) - x(t)|| \leq \delta(T+1)e^{L_1 T} < r/4. \tag{4.30}$$

Since $\tau_1 < T$, there is $\tau_2 \in (\tau_1, T)$ such that

$$||x(t) - y(t)|| < r/4, \ t \in [\tau_1, \tau_2]. \tag{4.31}$$

Relations (4.31) and (4.30) imply that

$$||x(t) - y(t)|| < r/4, \ t \in [0, \tau_2]. \tag{4.32}$$

Together with (4.11), this inequality implies that

$$y(t) \in E, \ t \in [0, \tau_2]. \tag{4.33}$$

By (4.22), (4.6), (4.7), (4.21), (4.20), (4.24), and (4.26), we have for $s \in [0, \tau_2]$,

$$||y(s) - x(s)|| = ||y_0 + \int_0^s y'(t)dt - x_0 - \int_0^s x'(t)dt|| \le ||x_0 - y_0||$$

$$+ \int_0^s ||y'(t) - x'(t)||dt \le \delta + \int_0^s ||Wy(t) - Vx(t)||dt$$

$$\le \delta + \int_0^s ||Vx(t) - Vy(t)||dt + \int_0^s ||Vy(t) - Wy(t)||dt$$

$$\le \delta + \int_0^s ||Vx(t) - Vy(t)||dt + \delta s \le \delta(1 + T) + \int_0^s L_1||x(t) - y(t)||dt.$$

It follows from this inequality and Gronwall's inequality that for all $s \in [0, \tau_2]$,

$$||y(s) - x(s)|| \le \delta(T + 1)e^{L_1 T}.$$

Hence $\tau_2 \in \Omega_0$, a contradiction. Therefore $\tau_1 = T$ and

$$||y(t) - x(t)|| < \delta(T + 1)e^{L_1 T}$$

for all $t \in [0, T]$. Together with (4.19), this latter inequality implies that $||y(t) - x(t)|| < \epsilon/2$ for all $t \in [0, T]$. Lemma 4.5 is proved.

Lemma 4.6. *Let $V \in \mathcal{A}$, $\gamma \in (0, 1)$, $\epsilon > 0$, and $x_0 \in X$. Then there exist a neighborhood $\mathcal{U}$ of (x_0, V_γ) in $X \times \mathcal{A}$ with the weak topology and $T_0 > 0$ such that for each $(\xi, W) \in \mathcal{U}$ and each continuously differentiable mapping $y : [0, \infty) \to X$ satisfying*

$$y'(t) = Wy(t), \ t \in [0, \infty), \ y(0) = \xi, \tag{4.34}$$

the following inequality holds: $f(x(T_0)) \le f(\bar{x}) + \epsilon$.

Proof. Let the mapping $x \in C^1([0, \infty); X)$ satisfy

$$x(0) = x_0, \ x'(t) = V_\gamma x(t), \ t \in [0, \infty). \tag{4.35}$$

(The existence of this mapping follows from Proposition 2.1.) By Lemma 4.4, there is $T_0 > 0$ such that

$$f(x(T_0)) \le f(\bar{x}) + \epsilon/2. \tag{4.36}$$

There also is $\delta > 0$ such that

$$|f(\xi) - f(x(T_0))| \leq \epsilon/4 \text{ for all } \xi \in B(x(T_0), \delta). \tag{4.37}$$

By Lemma 4.5, there exists a neighborhood $\mathcal{U}$ of (x_0, V_γ) in $X \times \mathcal{A}$ with the weak topology such that for each $(\xi, W) \in \mathcal{U}$ and each $y \in C^1([0,\infty); X)$ satisfying (4.34),

$$||y(T_0) - x(T_0)|| \leq \delta. \tag{4.38}$$

Assume that $(\xi, W) \in \mathcal{U}$ and that $y \in C^1([0,\infty); X)$ satisfies (4.34). Then (4.38) is true. Relations (4.38), (4.36) and (4.37) now imply that

$$f(y(T_0)) \leq f(x(T_0)) + \epsilon/4 \leq f(\bar{x}) + \epsilon.$$

Thus Lemma 4.6 is proved.

Completion of the proof of Theorem 2.1. The set

$$\{(z, V_\gamma) :\ z \in X,\ V \in \mathcal{A},\ \gamma \in (0,1)\}$$

is an everywhere dense subset of $X \times \mathcal{A}$ with the strong topology. Let $z \in X$, $V \in \mathcal{A}$, $\gamma \in (0,1)$, and let n be a natural number. By Lemma 4.6, there are $T(z, V, \gamma, n) > 0$, and an open neighborhood $\mathcal{U}(z, V, \gamma, n)$ of (z, V_γ) in $X \times \mathcal{A}$ with the weak topology, such that the following property holds:

(P) For each $(\xi, W) \in \mathcal{U}(z, V, \gamma, n)$ and each $y \in C^1([0,\infty); X)$ satisfying (4.34),

$$f(y(T(z, V, \gamma, n))) \leq f(\bar{x}) + 1/n.$$

Set

$$\mathcal{F} = \cap_{q=1}^{\infty} \cup \{\mathcal{U}(z, V, \gamma, q) :\ z \in X,\ V \in \mathcal{A},\ \gamma \in (0,1)\}.$$

It is obvious that $\mathcal{F}$ is a countable intersection of open (in the weak topology) everywhere dense (in the strong topology) subsets of $X \times \mathcal{A}$.

Let $(x_0, U) \in \mathcal{F}$ and $\epsilon > 0$. Choose $q > 8(\min\{1, \epsilon\}^{-1})$. There exist $z \in X$, $V \in \mathcal{A}$ and $\gamma \in (0,1)$ such that

$$(x_0, U) \in \mathcal{U}(z, V, \gamma, q).$$

Let $(\xi, W) \in \mathcal{U}(z, V, \gamma, q)$ and let $y \in C^1([0,\infty); X)$ satisfy (4.34). By property (P) and the choice of q,

$$f(y(T(z, V, \gamma, q))) \leq f(\bar{x}) + 1/q < f(\bar{x}) + \epsilon.$$

Since the function $f(y(t))$, $t \geq 0$, is decreasing, we conclude that

$$f(y(t)) < f(\bar{x}) + \epsilon \text{ for all } t \geq T(z, V, \gamma, q).$$

This completes the proof of Theorem 2.1.

Acknowledgments. The second author was partially supported by the Fund for the Promotion of Research at the Technion and by the Technion VPR Fund. He also thanks the Erwin Schrödinger International Institute for Mathematical Physics in Vienna for its support.

References

1. S. Aizicovici, S. Reich and A. J. Zaslavski, *Convergence theorems for continuous descent methods*, J. Evol. Equ. **4** (2004), 139-156.
2. S. Aizicovici, S. Reich and A. J. Zaslavski, *Most continuous descent methods converge*, Arch. Math. (Basel) **85** (2005), 268-277.
3. H. B. Curry, *The method of steepest descent for nonlinear minimization problems*, Quart. Appl. Math. **2** (1944), 258-261.
4. F. S. De Blasi and J. Myjak, *Sur la convergence des approximations successives pour les contractions non linéaires dans un espace de Banach*, C.R. Acad. Sci. Paris **283** (1976), 185-187.
5. F. S. De Blasi and J. Myjak, *Generic flows generated by continuous vector fields in Banach spaces*, Adv. Math. **50** (1983), 266-280.
6. J.-B. Hiriart-Urruty and C. Lemaréchal, *Convex Analysis and Minimization Algorithms*, Springer, Berlin, 1993.
7. J. W. Neuberger, *Sobolev Gradients and Differential Equations, Lecture Notes in Math.* **1670**, Springer, Berlin, 1997.
8. S. Reich and A. J. Zaslavski, *Generic convergence of descent methods in Banach spaces*, Math. Oper. Research **25** (2000), 231-242.
9. S. Reich and A. J. Zaslavski, *The set of divergent descent methods in a Banach space is σ-porous*, SIAM J. Optim. **11** (2001), 1003-1018.
10. S. Reich and A. J. Zaslavski, *Porosity of the set of divergent descent methods*, Nonlinear Anal. **47** (2001), 3247-3258.
11. S. Reich and A. J. Zaslavski, *Two convergence results for continuous descent methods*, Electron. J. Differential Equations **2003** (2003), 1-11.
12. G. Vidossich, *Most of the successive approximations do converge*, J. Math. Anal. Appl. **45** (1974), 127-131.

Department of Mathematics, Ohio University, Athens, OH 45701-2979, USA
E-mail address: aizicovi@math.ohiou.edu

Department of Mathematics, The Technion-Israel Institute of Technology, 32000 Haifa, Israel
E-mail address: sreich@tx.technion.ac.il; ajzasl@tx.technion.ac.il