PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 106, Number 3, July 1989

ON THE SECOND DERIVATIVES OF CONVEX FUNCTIONS ON HILBERT SPACES

NOBUYUKI KATO

(Communicated by William J. Davis)

ABSTRACT. Let ϕ be a proper l.s.c. convex function on a real Hilbert space H. We show that if H is separable, then ϕ is twice differentiable in some sense on a dense subset of the graph of $\partial\phi$.

1. INTRODUCTION

Let H be a (real) Hilbert space. The inner product and associated norm will be denoted by $\langle\cdot,\cdot\rangle$ and $|\cdot|$, respectively. Let $\phi: H \to (-\infty, +\infty]$ be a proper lower semicontinuous (l.s.c.) convex function. Then subdifferential $\partial\phi$ is defined by

$$\partial\phi(x) := \{y \in H | \phi(z) - \phi(x) \geq \langle y, z - x\rangle \quad \text{for all } z \in H\}.$$

We may regard $\partial\phi(x)$ as a generalized first derivative of ϕ at x. Actually, if ϕ is Gâteaux differentiable at x and has a continuous Gâteaux derivative $\nabla\phi(x)$, then $\partial\phi(x) = \{\nabla\phi(x)\}$.

Our purpose is to show that ϕ is twice differentiable in some sense on a dense subset of the graph of $\partial\phi$. Indeed, we show the existence of an operator which has nice properties and may be considered as a generalized second derivative of ϕ. Our main result is an infinite-dimensional version of an interesting work of Rockafellar [9].

2. PRELIMINARIES

To formulate our result, we will give some preparation in this section. Given a multi-valued operator T from H to 2^H, we define $D(T) := \{x \in H \mid Tx \neq \varnothing\}$, $R(T) := \bigcup_{x\in D(T)} Tx$. The graph of T is denoted by $G(T)$.

An operator T is said to be linear if its graph $G(T)$ is a linear subspace of $H \times H$. (T may be multi-valued.) For any linear operator T, we define its

Received by the editors January 25, 1988 and, in revised form, September 12, 1988.

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 47H05; Secondary 49A27, 49A50.

Key words and phrases. Convex function, subdifferential, second derivative, convergence in the sense of Mosco.

adjoint T^* by

$$G(T^*) := \{(v, g) \in H \times H \mid \langle f, v\rangle = \langle u, g\rangle \quad \text{for all } (u, f) \in G(T)\}$$

(cf. [1]). We say that T is self-adjoint if $T = T^*$, and T is positive if $\langle x, y\rangle \geq 0$ for all $(x, y) \in G(T)$.

We say that an operator $T: H \to 2^H$ is monotone if its graph $G(T)$ is a monotone set in $H \times H$, i.e.,

$$\langle x - y, x' - y'\rangle \geq 0 \quad \text{for all } (x, x'), (y, y') \in G(T).$$

T is said to be maximal monotone if it is monotone and there is no monotone extension of T. It is well known that T is maximal monotone iff T is monotone and the range condition $R(I + T) = H$ is satisfied. The subdifferential $\partial\phi$ is an important example of maximal monotone operator. For a maximal monotone operator T, we define its minimal section T^0 by

$$T^0 x := \{y \in Tx \mid |y| = \inf\{|z| : z \in Tx\}\}.$$

In Hilbert space setting, it is known that T^0 is a well defined single-valued monotone operator in H, with domain $D(T^0) = D(T)$. (See e.g. [5,6].)

Let $\{S_n\}$ be a sequence of sets in a Banach space X. Then we define the notion of set convergence as follows. (See e.g. [4,9].)

$$s\text{-}(w\text{-}) \overline{\lim_{n\to\infty}} S_n := \{v \in X \mid \text{there exist } \{n_k\} \subset \{n\} \text{ and } v_{n_k} \in S_{n_k} \text{ such that } s\text{-}(w\text{-}) \lim_{k\to\infty} v_{n_k} = v\},$$

$$s\text{-}(w\text{-}) \underline{\lim_{n\to\infty}} S_n := \{v \in X \mid \text{there exist } v_n \in S_n \text{ such that } s\text{-}(w\text{-}) \lim_{n\to\infty} v_n = v\},$$

$$s\text{-}(w\text{-}) \lim_{n\to\infty} S_n = S \quad \text{if } s\text{-}(w\text{-}) \overline{\lim_{n\to\infty}} S_n = s\text{-}(w\text{-}) \underline{\lim_{n\to\infty}} S_n = S,$$

where s-$(w$-$)$ means the strong (or weak) topology, respectively.

Let ϕ^n $(n = 0, 1, \dots)$ be proper l.s.c. convex functions on H. Then we say that ϕ^n converges to ϕ^0 in the sense of Mosco if

$$w\text{-} \overline{\lim_{n\to\infty}} \operatorname{Epi} \phi^n \subset \operatorname{Epi} \phi^0 \subset s\text{-} \underline{\lim_{n\to\infty}} \operatorname{Epi} \phi^n,$$

where $\operatorname{Epi} \phi^n$ is the epi-graph of ϕ^n, which is defined by

$$\operatorname{Epi} \phi^n = \{(x, \lambda) \in H \times \mathbf{R} \mid \phi^n(x) \leq \lambda\}.$$

According to Attouch [4, Theorem 1.2], the following conditions are equivalent:

(a) ϕ^n converges to ϕ^0 in the sense of Mosco.

(b) For any $\lambda > 0$ and $x \in H$,

$$(I + \lambda\partial\phi^n)^{-1} x \to (I + \lambda\partial\phi^0)^{-1} x \quad \text{as } n \to \infty,$$

and there exist $(u_n, v_n) \in G(\partial\phi^n)$ such that

$$u_n \to u_0, v_n \to v_0 \quad \text{and} \quad \phi^n(u_n) \to \phi^0(u_0).$$

(c) For any $\lambda > 0$ and $x \in H$,

$$\phi_\lambda^n(x) \to \phi_\lambda^0(x) \quad \text{as } n \to \infty,$$

where $\phi_\lambda^n(x)$ is the Yosida regularization of ϕ^n defined by

$$\phi_\lambda^n(x) = \min_{y \in H} \left\{ \phi^n(y) + \frac{1}{2\lambda}|x-y|^2 \right\}.$$

3. Main result

Let H be a Hilbert space and $\phi: H \to (-\infty, +\infty]$ be a proper l.s.c. convex function. For each $(x, y) \in G(\partial\phi)$ and $t > 0$, consider the second-order difference quotients

$$\Delta_{x,y,t}(h) := t^{-2}(\phi(x+th) - \phi(x) - t\langle y, h\rangle).$$

It is easily seen that $\Delta_{x,y,t}: H \to (-\infty, +\infty]$ is proper l.s.c. and convex. Our main theorem is now stated as follows:

Theorem. *Let H be a separable Hilbert space and ϕ be as above. Then there exists a dense subset E of $G(\partial\phi)$ such that ϕ is twice differentiable on E in the following sense:*

For any $(\overline{x}, \overline{y}) \in E$, there exists a proper l.s.c. convex function $q_{\overline{x},\overline{y}}: H \to (-\infty, +\infty]$ such that

(i) *$Q_{\overline{x},\overline{y}} \equiv \partial q_{\overline{x},\overline{y}}$ is a (multi-valued) linear positive self-adjoint operator in H (in the sense of §2).*

(ii) *The minimal section $Q^0_{\overline{x},\overline{y}}$ of $Q_{\overline{x},\overline{y}}$ is a single-valued linear positive self-adjoint operator in the closed linear subspace $\overline{D(Q_{\overline{x},\overline{y}})}$ of H, and*

$$q_{\overline{x},\overline{y}}(h) = \begin{cases} \frac{1}{2}|(Q^0_{\overline{x},\overline{y}})^{1/2}h|^2 & \text{if } h \in D((Q^0_{\overline{x},\overline{y}})^{1/2}) \\ +\infty & \text{otherwise}, \end{cases}$$

(iii) *$\Delta_{\overline{x},\overline{y},t}$ converges to $q_{\overline{x},\overline{y}}$ as $t \downarrow 0$ in the sense of Mosco.*

It should be noted that $Q_{\overline{x},\overline{y}}$ may be called the generalized second derivative of ϕ at $(\overline{x}, \overline{y})$ as shown below.

Remark. We observe that if $\phi: U \subset H \to \mathbf{R}$, U open convex, is of class C^2 and convex, then $Q_{\overline{x},\overline{y}} = d^2\phi(\overline{x})$ for $\overline{x} \in U$ satisfying $(\overline{x}, \overline{y}) \in E$ with $\overline{y} = d\phi(\overline{x})$, where $d\phi$ and $d^2\phi$ denote the first and second derivatives of ϕ.

Let $h \in H$ be fixed. By Taylor's theorem,

$$\phi(\overline{x} + th) = \phi(\overline{x}) + t\langle d\phi(\overline{x}), h\rangle + (t^2/2)\langle d^2\phi(\overline{x})h, h\rangle + o(t^2)$$

as $t \to 0$. Hence

$$\begin{aligned}\Delta_t(h) &:= t^{-2}(\phi(\overline{x} + th)\phi(\overline{x}) - t\langle d\phi(\overline{x}), h\rangle) \\ &\to \frac{1}{2}\langle d^2\phi(\overline{x})h, h\rangle \quad \text{as } t \to 0.\end{aligned}$$

Let $\tilde{q}(h) = \frac{1}{2}\langle d^2\phi(\overline{x})h, h\rangle$. Noting that $d^2\phi(\overline{x}): H \to H$ is self-adjoint, $\tilde{q}: H \to \mathbf{R}$ is continuous, convex, and $\partial\tilde{q} = d^2\phi(\overline{x})$. Since $\Delta_t: H \to \mathbf{R}$ is also continuous, convex, and $D(\Delta_t) = D(\tilde{q}) = H$, we have

$$\Delta_t \to \tilde{q} \quad \text{in the sense of Mosco}$$

by virtue of Salinetti and Wets [10, Corollary 2E]. On the other hand, by our theorem,

$$\Delta_t \to q_{\overline{x},\overline{y}}, \qquad \overline{y} = d\phi(\overline{x}), \quad \text{in the sense of Mosco}.$$

Then we have $(\tilde{q})_\lambda = (q_{\overline{x},\overline{y}})_\lambda$ by [4, Theorem 1.2]. See the condition (c) of §2. As a consequence, we obtain $\partial\tilde{q} = \partial q_{\overline{x},\overline{y}}$, i.e. $d^2\phi(\overline{x}) = Q_{\overline{x},\overline{y}}$.

4. Proof of Theorem

We begin with the following lemma.

Lemma 4.1. *Let $F: H \to H$ be a Lipschitz continuous mapping and suppose that F is Gâteaux differentiable at $\overline{u}$ and has a continuous linear Gâteaux derivative $\nabla F(\overline{u})$, i.e.,*

$$\nabla F(\overline{u})h = \lim_{t\downarrow 0} t^{-1}[F(\overline{u} + th) - F(\overline{u})] \quad \textit{for all } h \in H. \tag{4.1}$$

Then for any bijective continuous linear mapping $\Psi: H \times H \to H \times H$,

$$\lim_{t\downarrow 0} t^{-1}[\Psi(G(F)) - \Psi(\overline{u}, F(\overline{u}))] = \Psi(G(\nabla F(\overline{u}))),$$

where $\lim_{t\downarrow 0}$ is taken in the sense of set convergence with respect to the strong topology of $H \times H$ (see §2).

Proof. Let $(u, v) \in \Psi(G(\nabla F(\overline{u}))$ and $t_n \downarrow 0$. Setting $(x, y) = \Psi^{-1}(u, v)$, then

$$y = \nabla F(\overline{u})x = \lim_{n\to\infty} t_n^{-1}[F(\overline{u} + t_n x) - F(\overline{u})].$$

Puttting $x_n = \overline{u} + t_n x$, $y_n = F(\overline{u} + t_n x)$, then the linearity and continuity of Ψ imply that

$$\begin{aligned}\Psi(x, t_n^{-1}[F(\overline{u} + t_n x) - F(\overline{u})]) &= t_n^{-1}[\Psi(x_n, y_n) - \Psi(\overline{u}, F(\overline{u}))] \\ &\in t_n^{-1}[\Psi(G(F)) - \Psi(\overline{u}, F(\overline{u}))]\end{aligned}$$

and

$$\Psi(x, t_n^{-1}[F(\overline{u} + t_n x) - F(\overline{u})]) \to \Psi(x, y) = (u, v).$$

This shows that $\underline{\lim}_{t\downarrow 0} t^{-1}[\Psi(G(F)) - \Psi(\overline{u}, F(\overline{u}))] \supset \Psi(G(\nabla F(\overline{u})))$.

Conversely, let $(u, v) \in \overline{\lim}_{t\downarrow 0} t^{-1}[\Psi(G(F)) - \Psi(\overline{u}, F(\overline{u}))]$. Then there exist $t_n \downarrow 0$ and $(u_n, v_n) \to (u, v)$ such that $\Psi(\overline{u}, F(\overline{u})) + t_n(u_n, v_n) \in \Psi(G(F))$. Put $(x_n, y_n) = \Psi^{-1}(u_n, v_n)$ and $(x, y) = \Psi^{-1}(u, v)$. Then we have $(x_n, y_n) \to (x, y)$ since Ψ^{-1} is also continuous by the open mapping theorem. Thus, noting that $(\overline{u} + t_n x_n, F(\overline{u}) + t_n y_n) \in G(F)$, we have

$$t_n^{-1}[F(\overline{u} + t_n x_n) - F(\overline{u})] = y_n \to y.$$

But since F is Lipschitzian, we get

$$t_n^{-1}[F(\overline{u} + t_n x) - F(\overline{u})] \to y.$$

By the hypothesis (4.1), we have $y = \nabla F(\overline{u})x$ i.e. $(x, y) \in G(\nabla F(\overline{u}))$. Thus $(u, v) \in \Psi(G(\nabla F(\overline{u})))$ and the relation $\overline{\lim}_{t\downarrow 0} t^{-1}[\Psi(G(F)) - \Psi(\overline{u}, F(\overline{u}))] \subset \Psi(G(\nabla F(\overline{u})))$ holds. □

Next, we investigate the properties of multi-valued linear maximal monotone operators in H.

Lemma 4.2. *Let Q be a multi-valued linear maximal monotone operator in H. Suppose that Q is the subdifferential of some proper l.s.c. convex function $q: H \to (-\infty, +\infty]$. Then:*

(i) *Q is positive self-adjoint in the sense of* §2;

(ii) *the minimal section Q^0 of Q is a single-valued positive self-adjoint operator in the closed linear subspace $Y := \overline{D(Q)}$ of H, and*

$$q(h) = \begin{cases} \frac{1}{2}|(Q^0)^{1/2}h|^2 + const & \text{if } h \in D((Q^0)^{1/2})\ (\subset Y) \\ +\infty & \text{otherwise}. \end{cases} \tag{4.2}$$

Proof. (i) It follows from [7, Proposition 2.15] that $Q_\lambda = Q_\lambda^*$. Hence we see that $J_\lambda = J_\lambda^*$. Here $Q_\lambda = (1/\lambda)(I - J_\lambda)$ and $J_\lambda = (I + \lambda Q)^{-1}$ for $\lambda > 0$. Let $(v, g) \in G(Q)$. Then for any $(u, f) \in G(Q)$, we have

$$\langle f + u, v\rangle = \langle f + u, J_1(g + v)\rangle = \langle J_1(f + u), g + v\rangle = \langle u, g + v\rangle.$$

Thus $\langle f, v\rangle = \langle u, g\rangle$. This shows that $Q \subset Q^*$.

On the other hand, Q^* is monotone by [8, Theorem 2]. Then the maximality of Q implies that $Q = Q^*$. Positivity is obvious.

(ii) It is known [6, IV, Theorem 1.2; 7, Theorem 4.1] that there exists a unique contraction semigroup $\{S(t)\}$ on $Y = \overline{D(Q)}$ such that the minimal section $-Q^0$ of $-Q$ is the infinitesimal generator of $\{S(t)\}$. Notice that $\{S(t)\}$ is a linear (C_0)-contraction semigroup on the closed linear subspace Y.

Then Q^0 is a single-valued linear maximal monotone operator in Y (densely defined in Y). Since $Q^0 \subset Q = \partial q$, Q^0 is cyclically monotone by [7, Theorem

2.5]. Then it follows from [7, Proposition 2.15] that Q^0 is positive self-adjoint in Y and $Q^0 = \partial_Y \psi$ (the subdifferential considered in Y) with

$$\psi(h) = \begin{cases} \frac{1}{2}|(Q^0)^{1/2}h|^2 & \text{if } h \in D((Q^0)^{1/2}) \ (\subset Y) \\ +\infty & \text{otherwise}. \end{cases}$$

ψ is also proper l.s.c. convex in H, and it is easily checked that $Q^0 \subset \partial\psi$ and $D(Q)(= D(Q^0)) \subset D(\partial\psi) \subset \overline{D(Q)}\ (= Y)$. Then [7, Corollary 2.2] yields that $Q = \partial\psi$, and hence we conclude that $q = \psi + \text{const}$ by [7, Corollary 2.10]. □

Remark 4.1. It is shown that $Q = Q^0 + \partial I_Y$, where I_Y is the indicator function of Y (i.e. $I_Y(u) = 0$ if $u \in Y$, $= +\infty$ if $u \notin Y$). In fact, let $f \in Qu$, then f is written as $f = f_1 + f_2$ with $f_1 \in Y$ and $f_2 \in Y^\perp$ since $H = Y \oplus Y^\perp$ ($Y^\perp$ is the orthogonal complement of Y). As shown in Lemma 4.2, $Q = \partial\psi$ and so

$$\begin{aligned} \psi(z) - \psi(u) \geq \langle f, z-u\rangle &= \langle f_1 + f_2, z-u\rangle \\ &= \langle f_1, z-u\rangle \quad \text{for all } z \in Y. \end{aligned}$$

Therefore, $f_1 \in \partial_Y\psi(u) = Q^0u$. Noting that $\partial I_Y(u) = Y^\perp$ for $u \in Y$, we obtain $Q \subset Q^0 + \partial I_Y$. Conversely, $Q^0 + \partial I_Y \subset \partial\psi + \partial I_Y \subset \partial(\psi + I_Y) = \partial\psi = Q$.

Proof of Theorem. Put $F := 2(I + \partial\phi)^{-1} - I$, $\Phi(x, y) := (x + y, x - y)$. It is obvious that $F: H \to H$ is Lipschitz continuous, $\Phi: H \times H \to H \times H$ is bijective, continuous linear, and $\Phi^{-1}(u, v) = ((u+v)/2, (u-v)/2)$ is also continuous linear. Furthermore, we have $\Phi(G(\partial\phi)) = G(F)$.

Since H is separable, by the result of Mignot [10, Theorem 1.2], there exists a dense subset U of H on which F is Gâteaux differentiable and has a continuous linear Gâteaux derivative $\nabla F(\overline{u})$ at $\overline{u} \in U$. Thus for $\overline{u} \in U$, (4.1) holds.

Let $E := \Phi^{-1}(G(F|_U)) = \{((u + F(u))/2, (u - F(u))/2) | u \in U\}$, where $F|_U$ is the restriction of F to U. Then $E \subset \Phi^{-1}(G(F)) = G(\partial\phi)$ and E is dense in $G(\partial\phi)$ with respect to the product topology of $H \times H$. In fact, noting that F is Lipschitzian, we have $G(F|_U)$ is dense in $G(F)$. Thus by the continuity of Φ^{-1}, it is shown that E is dense in $\Phi^{-1}(G(F)) = G(\partial\phi)$.

Let $(\overline{x}, \overline{y}) \in E$. By the definition, there exists $\overline{u} \in U$ such that $(\overline{x}, \overline{y}) = \Phi^{-1}(\overline{u}, F(\overline{u}))$. Then it follows from Lemma 4.1 with $\Psi = \Phi^{-1}$ that

$$\lim_{t \downarrow 0} t^{-1}[G(\partial\phi) - (\overline{x}, \overline{y})] = \Phi^{-1}(G(\nabla F(\overline{u}))). \tag{4.3}$$

Define $Q_{\overline{x},\overline{y}}$ by $G(Q_{\overline{x},\overline{y}}) = \Phi^{-1}(G(\nabla F(\overline{u})))$. Then $Q_{\overline{x},\overline{y}}$ is a linear operator from H to 2^H with domain

$$D(Q_{\overline{x},\overline{y}}) = \{(x + \nabla F(\overline{u})x)/2 \,|\, x \in H\} (\overline{u} \text{ depends on } (\overline{x}, \overline{y}) \in E).$$

Moreover, $Q_{\overline{x},\overline{y}}$ is a maximal monotone operator in H. In fact, the monotonicity is a consequence of (4.3) since $t^{-1}[G(\partial\phi) - (\overline{x},\overline{y})]$ is monotone. Let $x \in H$. Then

$$\frac{x + \nabla F(\overline{u})x}{2} + Q_{\overline{x},\overline{y}}\left(\frac{x + \nabla F(\overline{u})x}{2}\right) \ni x$$

because $((x+\nabla F(\overline{u})x)/2, (x-\nabla F(\overline{u})x)/2) \in G(Q_{\overline{x},\overline{y}})$. Thus $R(I+Q_{\overline{x},\overline{y}}) = H$.

Now, consider the second-order difference quotients

$$\Delta_{x,y,t}(h) := t^{-2}(\phi(x+th) - \phi(x) - t\langle y, h\rangle)$$

for $(x,y) \in G(\partial\phi)$ and $t > 0$. It is easily verified that $\Delta_{x,y,t}: H \to (-\infty, +\infty]$ is a proper l.s.c. convex function with $\Delta_{x,y,t}(0) = 0$, and

$$G(\partial\Delta_{x,y,t}) = t^{-1}[G(\partial\phi) - (x,y)].$$

Therefore, we obtain from (4.3) that

$$\lim_{t\downarrow 0} G(\partial\Delta_{\overline{x},\overline{y},t}) = G(Q_{\overline{x},\overline{y}}). \tag{4.4}$$

Let $t_n \downarrow 0$. By (4.4), for any $(x,y) \in G(Q_{\overline{x},\overline{y}})$, there exists $(x_n, y_n) \in G(\partial\Delta_{\overline{x},\overline{y},t_n})$ such that $x_n \to x$ and $y_n \to y$ This is equivalent to: for any $\lambda > 0$ and $x \in H$,

$$(I + \lambda\partial\Delta_{\overline{x},\overline{y},t_n})^{-1}x \to (I + \lambda Q_{\overline{x},\overline{y}})^{-1}x \quad \text{as } n \to \infty$$

by [4, Corollary 1.1]. Then it follows from [4, Proposition 1.3] that

$$Q_{\overline{x},\overline{y}} = \partial\tilde{q}_{\overline{x},\overline{y}} \text{ for some proper l.s.c. convex function } \tilde{q}_{\overline{x},\overline{y}}.$$

Since $(0,0) \in G(\partial\tilde{q}_{\overline{x},\overline{y}})$, $\tilde{q}_{\overline{x},\overline{y}}(y) \geq \tilde{q}_{\overline{x},\overline{y}}(0)$ for every $y \in H$. Then setting

$$q_{\overline{x},\overline{y}}(x) := \tilde{q}_{\overline{x},\overline{y}}(x) - \tilde{q}_{\overline{x},\overline{y}}(0),$$

we have $\partial q_{\overline{x},\overline{y}} = \partial\tilde{q}_{\overline{x},\overline{y}} = Q_{\overline{x},\overline{y}}$ and $q_{\overline{x},\overline{y}}(0) = 0$. After all, we obtain:

$$\begin{gathered}\text{for any } \lambda > 0 \text{ and } x \in H,\ (I + \lambda\partial\ \Delta_{\overline{x},\overline{y},t_n})^{-1}x \to\ (I + \lambda\partial q_{\overline{x},\overline{y}})^{-1}x,\\ (0,0) \in G(\partial\overline{q}_{\overline{x},\overline{y}}),\ (0,0) \in G(\partial\Delta_{\overline{x},\overline{y},t_n}),\ \Delta_{\overline{x},\overline{y},t_n}(0) = q_{\overline{x},\overline{y}}(0) = 0.\end{gathered}$$

Hence by [4, Theorem 1.2] (see §2(b)), we conclude that $\Delta_{\overline{x},\overline{y},t_n}$ converges to $q_{\overline{x},\overline{y}}$ in the sense of Mosco. Thus (iii) is established.

The assertions (i) and (ii) are the consequences of Lemma 4.2. Note that the constant appeared in (4.2) becomes zero since $q_{\overline{x},\overline{y}}(0) = 0$. □

Remark 4.2. (1) The above pair of F and Φ was introduced previously in [12] and played an important role. But it is not essential. We may take, for example, $F := (I + \partial\phi)^{-1}$ and $\Phi(x,y) := (x+y, x)$ with appropriate changes of proof above.

(2) If $n := \dim H < +\infty$, E can be taken as $G(\partial\phi)\backslash E$ is of measure zero with respect to n-dimensional Hausdorff measure. Indeed, U can be taken as $H\backslash U$ is of measure zero with respect to n-dimensional Lebesgue measure by the

classical result of Rademacher. See e.g. [14, VIII, 3]. Since F is Lipschitzian, we have $\mathscr{H}^n(G(F|_{H\setminus U})) = 0$, where $\mathscr{H}^n$ is the n-dimensional Hausdorff outer measure. Since $\Phi^{-1} : H \times H \to H \times H$ is continuous linear, it is Lipschitz continuous. Denoting by L the Lipschitz constant of Φ^{-1}, $\mathscr{H}^n(G(\partial\phi)\setminus E) = \mathscr{H}^n(\Phi^{-1}(G(F|_{H\setminus U}))) \leq L^n \mathscr{H}^n(G(F|_{H\setminus U})) = 0$.

Remark 4.3. (1) Concerning (ii) of Theorem, Asakawa [3] previously suggested the expression where the restriction operator $Q|_Y$ (see [2]) is used instead of the minimal section Q^0. In our setting, it is shown that both of them coincide. Another expression is possible by using a square root $Q_{1/2}$ defined by the boundary value problem on half-axis of (elliptic) second order differential equations. See [6, V].

(2) After finishing the previous version of this paper, the author was informed by the referee of the J. L. Ndoutoume's work [11], which includes similar results to ours. $G(\partial\phi)$ is said to be smooth at $(\overline{x},\overline{y}) \in G(\partial\phi)$ if $\lim_{t\downarrow 0} t^{-1}[G(\partial\phi) - (\overline{x},\overline{y})] = S$ for some linear subspace S of $H \times H$ ([12]). [11, Proposition 2.2, Theorem 3.3] shows that if $G(\partial\phi)$ is smooth at $(\overline{x},\overline{y})$, then there exists a proper l.s.c. convex function $q_{\overline{x},\overline{y}}$ such that (i) the minimal section $Q^0_{\overline{x},\overline{y}}$ of $Q_{\overline{x},\overline{y}} \equiv \partial q_{\overline{x},\overline{y}}$ is a positive symmetric single-valued linear operator, and $q_{\overline{x},\overline{y}}(h) = \frac{1}{2}\langle Q^0_{\overline{x},\overline{y}}h, h\rangle$ for $h \in D(Q_{\overline{x},\overline{y}})$; (ii) $\Delta_{\overline{x},\overline{y},t}$ converges to $q_{\overline{x},\overline{y}}$ as $t \downarrow 0$ in the sense of Mosco.

The proof of our theorem shows that $G(\partial\phi)$ is smooth at each point in E, so that the points at which $G(\partial\phi)$ is smooth are dense in $G(\partial\phi)$ provided H is separable. Our result tells more about $Q^0_{\overline{x},\overline{y}}$, the self-adjointness in Y and the complete expression of $q_{\overline{x},\overline{y}}$ by the square root of $Q^0_{\overline{x},\overline{y}}$. It would be interesting that our proof of the self-adjointness of $Q^0_{\overline{x},\overline{y}}$ is based on the nonlinear as well as linear Hille-Yosida theorem.

The author was also informed by the referee that J.-B. Hiriart-Urruty [9] had defined the second differential for convex functions. Roughly speaking, it corresponds to a convex set C of which support function satisfies $\psi^*_C(h) = |(Q^0_{\overline{x},\overline{y}})^{1/2}h|$. See [9,11] for further details.

Finally, we note that, by (4.3), our second derivatives are related to the tangent cones (cf. [5]).

Acknowledgments

The author would like to thank Professor Y. Kōmura for giving a motivation of this work, and Professor I. Miyadera for his kind advice and constant encouragements. Also, the author would like to express his hearty thanks to Dr. H. Asakawa for pointing out the mistake of a previous preprint of this paper and for having a valuable discussion. In addition, the author is grateful to the referee for showing the references [9, 11] and for giving several comments that help to make the work more complete.

REFERENCES

1. R. Aren, *Operational calculus of linear relations*, Pacific J. Math. **11** (1961), 9–23.
2. H. Asakawa, *Restriction of maximal monotone operator to closed linear subspace*, TRU Math. **23** (1987), 97–116.
3. H. Asakawa, private communication.
4. H. Attouch, *Familles d'operateurs maximaux monotones et measurabilite*, Ann. Mat. Pura Appl. **120** (1979), 35–111.
5. J. P. Aubin and I. Ekeland, *Applied nonlinear analysis*, Wiley-Interscience, New York, 1984.
6. V. Barbu, *Nonlinear semigroups and differential equations in Banach spaces*, Noordhoff, Groningen, 1976.
7. H. Brezis, *Operateurs maximaux monotones*, North-Holland, Amsterdam, 1973.
8. H. Brezis and F. E. Browder, *Linear maximal monotone operators and singular nonlinear integral equations of Hammerstein type*, Nonlinear Analysis (L. Cesari, R. Kannan and H. F. Weinberger, eds.), Academic Press, New York, 1978.
9. J.-B. Hiriart-Urruty, *A new set-valued second order derivatives for convex functions*, Mathematics for Optimization, Mathematical Studies Series 129, North-Holland, Amsterdam, 1986.
10. F. Mignot, *Contrôle dans les inequations variationelles elliptiques*, J. Funct. Anal. **22** (1976), 130–185.
11. J. L. Ndoutoume, *Calcul differential generalise du second ordre*, Publ. AVAMAC Université du Perpignan, vol. 2, 1986.
12. R. T. Rockafellar, *Maximal monotone relations and the second derivatives of nonsmooth functions*, Ann. Int. Poincare Analyse Non Linéaire **2** (1985), 167–184.
13. G. Salinetti and R. J.-B. Wets, *On the relations between two types of convergence for functions*, J. Math. Anal. Appl. **60** (1977), 211–226.
14. E. M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Univ. Press, 1970.

DEPARTMENT OF MATHEMATICS, SCHOOL OF EDUCATION, WASEDA UNIVERSITY, TOKYO 160, JAPAN

Current address: Department of Mathematics, Faculty of Science, Shimane University, Matsue, Shimane 690, Japan