### 0.1 Subdifferentials

In this section $X$ will denote a Banach space, $X^{*}$ its dual and $\langle\cdot, \cdot\rangle$ the pairing between the elements of $X^{*}$ and those of $X$. Monotone operators play the role of increasing or dicreasing functions from $\mathbb{R}$ to $\mathbb{R}$.

Definition 1 An opreator $T: X \rightarrow X^{*}$ is called monotone if

$$
\langle T x-T y, x-y\rangle \geq 0
$$

strictly monotone if

$$
\langle T x-T y, x-y\rangle>0
$$

and strongly monotone if

$$
\langle T x-T y, x-y\rangle \geq \alpha\|x-y\|^{2}
$$

for some $\alpha>0$ and all $x, y \in X$.
Definition 2 The graph of a multivalued operator $T: X \rightarrow 2^{X^{*}}$ is defined to be

$$
G(T)=\{(x, \xi): x \in X, \xi \in T x\}
$$

Definition 3 A multivalued operator $T: X \rightarrow 2^{X^{*}}$ is called monotone if

$$
\langle\xi-\eta, x-y\rangle \geq 0
$$

strictly monotone if

$$
\langle\xi-\eta, x-y\rangle>0
$$

and strongly monotone if

$$
\langle\xi-\eta, x-y\rangle \geq \alpha\|x-y\|^{2}
$$

for some $\alpha>0$, all $x, y \in X$ and all $\xi \in T x, \eta \in T y$.
Definition $4 T: X \rightarrow 2^{X^{*}}$ is called maximal monotone if it is monotone and has no proper extension to a monotone operator. In other words, if $S: X \rightarrow 2^{X^{*}}$ is a monotone operator such that $G(S) \supseteq G(T)$ then $G(S)=G(T)$.

Definition 5 An operator $F: X \rightarrow \mathbb{R}$ is called subdifferentiable at $x \in X$ if it has a linear minorant at $F(x)$. In other words, if there exists a $\xi \in X^{*}$ such that

$$
\langle\xi, y-x\rangle \leq F(y)-F(x) \forall y \in X
$$

The set of all such $\xi$ is calle the subdifferential of $F$ at $x$ and is denoted by $\partial F(x)$. If $F$ does not have a subdifferential at $x$ then $\partial F(x)=\phi$.

Roughly speaking, the subdifferential is the set of all "slopes" of lines that can be drawn supporting the epigraph of the operator $F$ at the point $(x, F(x))$.

Proposition $6 \partial F(x)$ is convex and $w^{*}$-closed.
Proof. Let $\xi, \eta \in \partial F(x), \alpha \in[0,1]$, then

$$
\begin{aligned}
\langle\alpha \xi+(1-\alpha) \eta, y-x\rangle & =\alpha\langle\xi, y-x\rangle+(1-\alpha)\langle\eta, y-x\rangle \\
& \leq \alpha(F(y)-F(x))+(1-\alpha)(F(y)-F(x)) \\
& =F(y)-F(x)
\end{aligned}
$$

hence, $\alpha \xi+(1-\alpha) \eta \in \partial F(x)$. To show the $w^{*}$-closedness, let $\xi_{n} \in \partial F(x)$ such that $\xi_{n} \xrightarrow{w^{*}} \xi$, then, for all $y \in X$,

$$
\left\langle\xi_{n}, y-x\right\rangle \leq F(y)-F(x)
$$

and hence,

$$
\langle\xi, y-x\rangle \leq F(y)-F(x),
$$

i.e., $\xi \in \partial F(x)$.

Proposition $7 \partial F(\cdot)$ is a monotone operator
Proof. Suppose $\xi \in \partial F(x)$ and $\eta \in \partial F(y)$, then

$$
\begin{aligned}
\langle\xi, y-x\rangle+F(x) & \leq F(y) \\
\langle\eta, x-y\rangle+F(y) & \leq F(x)
\end{aligned}
$$

The second inequality can be rearranged as

$$
\langle-\eta, y-x\rangle-F(x) \leq-F(y)
$$

which, when added to the first inequality yeilds

$$
\langle\xi-\eta, y-x\rangle \leq 0
$$

or

$$
\langle\xi-\eta, x-y\rangle \geq 0
$$

In the proof of the following proposition we will need the following version of the Hahn-Banach Theorem.

Theorem 8 Let $Y$ be a real locally convex space, $A, B$ be two nonempty convex sets in $Y$ such that $A$ is open and $A \cap B=\phi$. Then there exists a closed hyperplane $M$ that separates $A$ and $B$. In other words, there exist $\xi \in Y^{*}$ and $\beta \in \mathbb{R}$ such that $\langle\xi, y\rangle+\beta>0$ for all $y \in A,\langle\xi, y\rangle+\beta=0$ for all $y \in M$ and $\langle\xi, y\rangle+\beta \leq 0$ for all $y \in B$.

Proposition 9 Suppose $F: X \rightarrow \overline{\mathbb{R}}$ is convex, $x \in \operatorname{dom}(F)$ and $F$ is continuous at $x$ then $\partial F(x) \neq \phi$.

Proof. The idea is to obtain a hyperplane $M$ that supports the epigraph of $F$. Since $F$ is continuous at $x$, there exists an open set $\mathcal{O}$ such that $x \in \mathcal{O}$ and $F$ is bounded above, say by $c<\infty$ on $\mathcal{O}$. Notice that $\mathcal{O} \subset \operatorname{dom}(F)$ and $\mathcal{O} \times(c, \infty) \subset \operatorname{epi}(F)^{\circ}$ and thus epi $(F)^{\circ} \neq \phi$. epi $(F)^{\circ}$ is convex since $F$ is convex. Also, $(x, F(x)) \notin$ epi $(F)^{\circ}$. Hence, by the Hahn-Banach Theorem, there exists a closed hyperplane $M$ in $X \times \mathbb{R}$ that supports epi $(F)^{\circ}$ and contains $(x, F(x))$. The equation of $M$ can be written as $\langle\xi, y\rangle+m r+\beta=0$ for some $\xi \in X^{*}, m, \beta \in \mathbb{R}$ (not all zeros) and all $(y, r) \in M$. Furthermore,

$$
\begin{equation*}
\langle\xi, y\rangle+m r+\beta>0 \tag{1}
\end{equation*}
$$

for all $(y, r) \in \operatorname{epi}(F)^{\circ}$ and

$$
\langle\xi, x\rangle+m F(x)+\beta=0
$$

Hence, $\beta=-\langle\xi, x\rangle-m F(x)$ and (1) becomes

$$
\begin{equation*}
\langle\xi, y-x\rangle+m(r-F(x))>0 \forall(y, r) \in \operatorname{epi}(F)^{\circ} . \tag{2}
\end{equation*}
$$

If $m=0$ we get $\langle\xi, y-x\rangle>0$ for all $y \in \mathcal{O}$ (inparticular). Since $\mathcal{O}$ is open, then $\xi=0$ and $\beta=0$, which is a contradiction. Hence, we may assume withour loss of generality that $m>0$. In this case, (2) may be rearranged as

$$
\left\langle-\frac{\xi}{m}, y-x\right\rangle+F(x)<r \forall(y, r) \in \operatorname{epi}(F)^{\circ} .
$$

Therefore,

$$
\left\langle-\frac{\xi}{m}, y-x\right\rangle+F(x) \leq r \forall(y, r) \in \operatorname{epi}(F)
$$

and, in particular

$$
\left\langle-\frac{\xi}{m}, y-x\right\rangle+F(x) \leq F(y)
$$

Thus, $-\frac{\xi}{m} \in \partial F(x)$.
Exercise 10 If $F: X \rightarrow \overline{\mathbb{R}}$ is convex then $F(x+t h)-F(x) \leq t(F(x+h)-F(x))$ for sufficiently small $t>0$.

Exercise 11 If $F: X \rightarrow \overline{\mathbb{R}}$ is convex then

$$
\frac{F(x+t h)-F(x)}{t}
$$

is increasing as a function of $t$. Hence, it has a limit as $t \rightarrow 0^{+}$(which may be $\pm \infty)$.

Exercise 12 Denote the limit in the previous exercise by $d(x ; h)$. Then for any $h \in X$

$$
F(x)+\lambda d(x ; h) \leq F(x+\lambda h) \forall \lambda \in \mathbb{R} .
$$

Exercise 13 Let $x, y \in X$ and

$$
\mathcal{L}=\{x+\lambda y: \lambda \in \mathbb{R}\} .
$$

If $\Pi:\langle\xi, u\rangle+\beta=0$ is a hyperplane such that $\langle\xi, u\rangle+\beta \geq 0$ for all $u \in \mathcal{L}$, then $\langle\xi, u\rangle+\beta=0$ for all $u \in \mathcal{L}$. In other words, if a hyperplane contains a straight line on one side then it must actually contain the line in it.

### 0.2 Relationship to the Gateaux Derivative

Notice that from the definitions of the Gateaux derivative and the subdifferentials, the $F$ is assumed to be finite at the point of evaluation $x$. This will also be the case with convex functions.

Proposition 14 Suppose $F: X \rightarrow \overline{\mathbb{R}}$ is convex. If $F$ has a Gateaux derivative at $x \in \operatorname{dom} F$, then $\partial F(x)=\left\{F^{\prime}(x)\right\}$. Conversely, if $F$ is continuous at $x \in X$ and has a unique subdifferential $\{\xi\}$ at $x$, then $F$ is Gateaux differentiable at $x$ and $F^{\prime}(x)=\xi$.

Proof. Suppose $F$ has a Gateaux derivative at $x \in X$. we will show that $F^{\prime}(x) \in \partial F(x)$.

$$
\left\langle F^{\prime}(x), t h\right\rangle=F(x+t h)-F(x)+o(t) .
$$

Thus, for sufficiently small $t>0$, we have, using Exercise 10,

$$
\left\langle F^{\prime}(x), h\right\rangle \leq F(x+h)-F(x)+\frac{o(t)}{t}
$$

Taking the limit as $t \rightarrow 0^{+}$we get

$$
\left\langle F^{\prime}(x), h\right\rangle \leq F(x+h)-F(x),
$$

i.e., $F^{\prime}(x) \in \partial F(x)$. On the other hand, suppose $\xi \in \partial F(x)$, then

$$
\begin{aligned}
\langle\xi, t h\rangle+F(x) & \leq F(x+t h) \\
& =\left\langle F^{\prime}(x), t h\right\rangle+o(t)
\end{aligned}
$$

For $t>0$ we get

$$
\langle\xi, h\rangle \leq\left\langle F^{\prime}(x), h\right\rangle+\frac{o(t)}{t}
$$

which gives

$$
\langle\xi, h\rangle \leq\left\langle F^{\prime}(x), h\right\rangle .
$$

For $t<0$ we get

$$
\langle\xi, h\rangle \geq\left\langle F^{\prime}(x), h\right\rangle+\frac{o(t)}{t}
$$

which gives

$$
\langle\xi, h\rangle \geq\left\langle F^{\prime}(x), h\right\rangle .
$$

Combining the two inequalities gives

$$
\langle\xi, h\rangle=\left\langle F^{\prime}(x), h\right\rangle .
$$

Hence $\xi=F^{\prime}(x)$.
To show the converse, observe first that, for sufficiently small $t>0$,

$$
\langle\xi, t h\rangle \leq F(x+t h)-F(x) \leq t(F(x+h)-F(x))
$$

Therefore,

$$
\langle\xi, h\rangle \leq \frac{F(x+t h)-F(x)}{t} \leq F(x+h)-F(x)
$$

and, because $F$ is continuous at $x$, it follows that $d(x ; h)$ is finite. Next observe that from Exercise 12, for sufficiently small $h \in X$

$$
F(x)+\lambda d(x ; h) \leq F(x+\lambda h) \forall \lambda \in \mathbb{R} .
$$

Geometrically, this means that the line

$$
L=\{(x, F(x))+\lambda(h, d(x ; h)): \lambda \in \mathbb{R}\}
$$

in $X \times \mathbb{R}$ does not intersect (epi $F)^{\circ}$. Since the latter set is not empty we get by the Hahn-Banach therorem, the existemce of a hyperplane $\Pi:\langle\eta, y\rangle+m r+\beta=0$ that seperates $(\mathrm{epi} F)^{\circ}$ and $L$. By Exercise 13, $\Pi$ contains $L$.i.e.,

$$
\langle\eta, x+\lambda h\rangle+m(F(x)+\lambda d(x ; h))+\beta=0 \forall \lambda \in \mathbb{R} .
$$

It follows that $m \neq 0$ and $d(x ; h)=\left\langle\frac{\eta}{-m}, h\right\rangle$. i.e.,

$$
d(x ; h)=\lim _{t \rightarrow 0^{+}} \frac{F(x+t h)-F(x)}{t}
$$

is a continuous linear functional on $X$. Hence, $d(x ; h)=F^{\prime}(x)$. From the uniqueness of the subdifferential of $F$, we get $F^{\prime}(x)=\xi$.

Exercise 15 Give a direct proof that the operator $d(x ; h)$ is linear and continuous.

In terms of the Gateaux derivative, a convex function is charachterized as follows

Proposition 16 Suppose $F$ is defined on the convex set $\mathcal{A} \subset X$ inot $\mathbb{R}$. If $F$ is Gateaux differentiable on $\mathcal{A}$ then the following are equivalent

1. $F$ is convex
2. $F(y) \geq F(x)+\left\langle F^{\prime}(x), y-x\right\rangle$ for all $x, y \in \mathcal{A}$.

Proof. 1. $\Longrightarrow 2$. is clear from the previous proposition since the Gateaux derivative is a subdifferentiable. To show the converse, rewrite the inequality with $x$ replaced with $\alpha x+(1-\alpha) y$, then

$$
F(y) \geq F(\alpha x+(1-\alpha) y)+\left\langle F^{\prime}(\alpha x+(1-\alpha) y), \alpha(y-x)\right\rangle
$$

and similarly

$$
F(x) \geq F(\alpha x+(1-\alpha) y)+\left\langle F^{\prime}(\alpha x+(1-\alpha) y),(1-\alpha)(x-y)\right\rangle
$$

Multiplying the first inequality by $(1-\alpha)$ and the second by $\alpha$ and adding we get

$$
(1-\alpha) F(y)+\alpha F(x) \geq F(\alpha x+(1-\alpha) y)
$$

i.e., $F$ is convex.

In terms of the monotonicity of the Gateaux derivative, the convexity of a function $F$ is characterized as follows

Proposition 17 Suppose $F$ is defined on the convex set $\mathcal{A} \subset X$ inot $\mathbb{R}$. If $F$ is Gateaux differentiable on $\mathcal{A}$ then the following are equivalent

1. $F$ is convex
2. $F^{\prime}(x)$ is monotone on $\mathcal{A}$.

Proof. If $F$ is convex then its Gateaux derivative is a subdifferential and we saw before (Porposition 7) that the subdifferential is a monotone operator. On the other hand, suppose $F^{\prime}(x)$ is monotone on $\mathcal{A}$.The function

$$
\varphi(\lambda)=F(x+\lambda(y-x)), \lambda \in[0,1] .
$$

is differentiable and

$$
\varphi^{\prime}(\lambda)=\left\langle F^{\prime}(x+\lambda(y-x)), y-x\right\rangle
$$

If $\lambda_{1} \leq \lambda_{2}$ then

$$
\left\langle F^{\prime}\left(x+\lambda_{2}(y-x)\right)-F^{\prime}\left(x+\lambda_{1}(y-x)\right),\left(\lambda_{2}-\lambda_{1}\right) y-x\right\rangle \geq 0
$$

which implies that

$$
\left\langle F^{\prime}\left(x+\lambda_{2}(y-x)\right)-F^{\prime}\left(x+\lambda_{1}(y-x)\right), y-x\right\rangle \geq 0
$$

i.e.,

$$
\varphi^{\prime}\left(\lambda_{2}\right)-\varphi^{\prime}\left(\lambda_{1}\right) \geq 0
$$

Hence, $\varphi^{\prime}(\cdot)$ is increasing. Therefore, $\varphi(\cdot)$ is convex (see exercise below) and

$$
\varphi(\lambda) \leq(1-\lambda) \varphi(0)+\lambda \varphi(1)
$$

i.e.,

$$
F(x+\lambda(y-x)) \leq(1-\lambda) F(x)+\lambda F(y) .
$$

Exercise 18 If $\varphi^{\prime}: I \rightarrow \mathbb{R}$ is increasing, then $\varphi$ is convex.

