

# Representation and approximation of convex dynamic risk measures with respect to strong-weak topologies

Ramin Okhrati\* and Hirbod Assa†

## Abstract

We provide a representation for strong-weak continuous dynamic risk measures from  $L^p$  into  $L_t^p$  spaces where these spaces are equipped respectively with strong and weak topologies and  $p$  is a finite number strictly larger than one. Conversely, we show that any such representation that admits a compact (with respect to the product of weak topologies) sub-differential generates a dynamic risk measure that is strong-weak continuous. Furthermore, we investigate sufficient conditions on the sub-differential for which the essential supremum of the representation is attained. Finally, the main purpose is to show that any convex dynamic risk measure that is strong-weak continuous can be approximated by a sequence of convex dynamic risk measures which are strong-weak continuous and admit compact sub-differentials with respect to the product of weak topologies. Throughout the arguments, no conditional translation invariance or monotonicity assumptions are applied.

**Keywords:** Sub-differential, Dynamic risk measures, Representation theorem, Convexity, Weak and strong continuity

---

\*University of Southampton, Southampton, UK, Email: r.okhrati@soton.ac.uk

†University of Liverpool, Liverpool, UK, Email: h.assa@liverpool.ac.uk

# 1 Introduction

Since the pioneering work of Artzner et al. (1999), risk measures have been playing an important role in developing mathematical finance models. Most risk measures are either sub-additive, convex, or quasiconvex, and they are considered as maps on functional analysis spaces modeling possible sets of risks. Therefore, the financial models based on these risk measures are heavily dependent on functional and convex analysis techniques.

The properties of a convex function, defined on a general Banach space, are related to its sub-differential. In Rockafellar (1966), it is proved that a real valued lower semi-continuous proper convex function, defined on a Banach space, is uniquely (up to an additive constant) determined by its sub-differential. This highlights the importance of sub-differential sets in convex analysis. The problem gets more complicated when the image space of the convex function is rather a functional space as it is the case for dynamic risk measures. Nevertheless, in either the static or dynamic case, the determination of a convex risk measure (or a convex function in general) is normally through a representation which is closely related to the sub-differential and continuity properties of the function.

The “continuity property” of a convex risk measure, either static or dynamic, (or in general a convex map) could be defined in several ways. The most common definitions of this property are continuity from above which is equivalent to Fatou property and lower semicontinuity with respect to weak\*-topologies, for instance, see Theorem 4.33 of Föllmer and Schied (2011) where this is discussed on an  $L^\infty$ -space for static risk measures that are monotone, translation invariant, and convex. There is also continuity from below which is stronger than continuity from above and it is equivalent to the Lebesgue property, see Corollary 4.35 of Föllmer and Schied (2011). The same results are extended in Chapter 11 of Föllmer and Schied (2011) to dynamic risk measures from  $L^\infty$  to  $L_t^\infty$  (for natural numbers  $t$ ) which satisfy conditional translation invariance, monotonicity, and conditional convexity. Fatou and Lebesgue property are also discussed on bounded càdlàg processes in Assa (2011).

Theorem 3.10 of Brezis (2011) shows that strong continuity of linear operators is equivalent to weak continuity. However, we cannot make the same conclusion about general (nonlinear) maps such as dynamic risk measures. One of the major challenges in convex analysis, in particular in risk measure theory, is that nonlinear maps which are strongly continuous are not necessarily continuous with respect to weak topologies. More precisely, any weak-weak continuous<sup>1</sup> risk measure satisfies the strong-weak continuity assumption while the other way

---

<sup>1</sup>The terminology weak-weak means that domain and the range of the risk measure are equipped with weak topologies. The other terminologies such as strong-weak or strong-strong are defined in the same fashion.

around does not necessarily hold unless the risk measure is linear which is not of interest neither theoretically nor practically.

Furthermore, note that strong-weak continuity does not imply strong-strong continuity again except for linear maps. We believe that a study of dynamic risk measures from  $L^p$  into  $L_t^p$  spaces, for integer  $t$  and  $1 < p < \infty$ , which are convex and strong-weak continuous with respect to the norm and weak topologies, is missing in the existing literature. The main motivation of this paper is to study these risk measures and investigate their approximation using dynamic convex risk measures which are strong-weak continuous and admit weakly compact sub-differential sets. Although the representations and results of this work are provided for  $L^2$  and  $L_t^2$  spaces, the arguments can be easily extended to  $L^p$  spaces for  $1 < p < \infty$ . However to simplify the notation, we concentrate on the case  $p = 2$ . Overall, there are three main purposes to this work.

First, we provide a representation theorem for dynamic convex risk measures from the space  $L^2 = L^2(\Omega, \mathcal{F}, \mathbb{P})$  to  $L_t^2 = L^2(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ ,  $t = 0, 1, \dots, n$ , that are continuous with respect to the strong topology on  $L^2$  induced by the  $L^2$ -norm, and the weak topology  $\sigma(L_t^2, L_t^2)$  on  $L_t^2$ . Note that the representation also holds for strong-strong continuous dynamic convex risk measures because they are strong-weak continuous as well. However, a representation of strong-strong continuous dynamic risk measures cannot be applied here because a strong-weak continuous map is not necessarily strong-strong continuous. In addition, we do not assume any regularity assumptions (also called local properties) such as those in Frittelli and Maggis (2014) though in their work, they study quasiconvex functions rather than the convex ones. In our work, the structure of the aforementioned dynamic convex risk measures is completely characterized using a dual representation without these regularity assumptions.

Second, we suppose that such representation is priorly given, and we investigate a mild sufficient condition over the sub-differential under which the dynamic risk measure is strong-weak continuous. More precisely, we show that the compactness of sub-differential with respect to the product of weak topologies guarantees the strong-weak continuity of the dynamic risk measure. In other words, compactness with respect to strong topologies is not required in the implication of strong-weak continuity. Although the main goal is to approximate dynamic convex risk measures which are strong-weak continuous, to the best of our knowledge, neither this form of representation nor these sufficient conditions are yet studied. Most popular risk measures and some new ones (either dynamic or static) can arise from our representation using the sufficient condition. In this context, we also discuss when the essential supremum in the representation theorem is actually attained.

Finally, we prove that any such dynamic risk measure which is convex and strong-weak continuous can be approximated by dynamic convex risk measures which are strong-weak continuous and admit weakly compact sub-differential sets. In what follows, we briefly review some related past work.

In Detlefsen and Scandolo (2005), representations of conditional convex risk measures are explained where the risk measures are assumed to be continuous from above. They also explain time consistency of conditional convex risk measures. Note that time consistency is not the interest of our work, and so we refer to Cheridito et al. (2006), Cheridito and Kupper (2011), and Riedel (2004) for discussions on this concept.

In Frittelli and Maggis (2011), a dual representation is obtained when the function is quasiconvex and it is lower semicontinuous with respect to a weak topology which is equivalent to continuity from below. The results are further extended to more general spaces in Frittelli and Maggis (2014). In Cerreia-Vioglio et al. (2011a), a dual representation is discussed for monotone quasiconcave function on  $M$ -spaces.

The outline of the paper is as follows: Some well-known definitions and an elementary lemma are recalled and discussed in Section 2. The main results of this paper and their proofs are explained in Section 3. Finally, Section 4 is devoted to conclusions and future work.

## 2 Preliminaries

Consider a complete probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  equipped with a filtration  $\{\mathcal{F}_t\}_{t=0, \dots, n}$ , where  $\mathcal{F}_n = \mathcal{F}$  and  $\mathcal{F}_0 = \{\emptyset, \Omega\}$ . We consider the functional spaces  $L^2 = L^2(\Omega, \mathcal{F}, \mathbb{P})$  and  $L_t^2 = L^2(\Omega, \mathcal{F}_t, \mathbb{P})$ ,  $t = 0, 1, \dots, n$ . The strong topology is induced by the norm  $\|\cdot\|$  defined by  $\|X\| = \mathbb{E}[X^2]^{\frac{1}{2}}$  for random variables  $X$  on this probability space where  $\mathbb{E}$  is the expectation with respect to the probability  $\mathbb{P}$ . Any map  $\rho_t : L^2 \mapsto L_t^2$ ,  $t = 0, 1, \dots, n$ , is referred as a dynamic risk measure. From a financial point of view,  $\rho_t(X)$  for  $X \in L^2$ , is interpreted as the (conditional) capital requirement for the financial position  $X$ . For the rest of the paper, unless otherwise stated, all equalities and inequalities of random variables are understood to hold  $\mathbb{P}$ -almost surely. To simplify the notation, most of time we drop the terminology “almost surely (a.s. for short)” though we explicitly mention it for important formulas.

**Definition 2.1.** *A dynamic risk measure  $\rho_t : L^2 \mapsto L_t^2$  satisfies normalization, conditional translation invariance, monotonicity, and conditional convexity if the following conditions hold respectively:*

- *Normalization:  $\rho_t(0) = 0$ , for all  $t$ ,*

- *Conditional cash (or translation) invariance*: for all  $t$ ,  $X \in L^2$ , and  $Y \in L_t^2$ ,  $\rho_t(X + Y) = \rho_t(X) - Y$ ,
- *Monotonicity*: if  $X \leq Y$ , for  $X$  and  $Y$  in  $L^2$  then  $\rho_t(X) \geq \rho_t(Y)$  for all  $t$ ,
- *Conditional convexity*:  $\rho_t(\lambda X + (1 - \lambda)Y) \leq \lambda \rho_t(X) + (1 - \lambda)\rho_t(Y)$ , for all  $t$ ,  $X \in L^2$ ,  $Y \in L^2$ ,  $\lambda \in L_t^2$ , and  $0 \leq \lambda \leq 1$ .

**Remark 2.1.** *It is important to note that neither monotonicity nor conditional cash invariance are assumed in the results of this work, and therefore dynamic risk measures  $\rho_t$  are not to be necessarily regular, i.e. the equation  $\rho_t(1_A X + 1_{A^c} Y) = 1_A \rho_t(X) + 1_{A^c} \rho_t(Y)$  does not necessarily hold for  $X, Y$  in  $L^2$  and  $A$  in  $\mathcal{F}_t$ .*

*In the context of non dynamic risk measures, El Karoui and Ravanelli (2009) show that the cash invariance property does not hold in the presence of stochastic interest rates, and therefore, they study convex risk measures under cash sub-invariance axiom (cash sub-additivity) instead. However, cash sub-invariance would lead to another problem. Under this axiom, diversification principal is not satisfied unless quasiconvex risk measures are used, see Cerreia-Vioglio et al. (2011b).*

*The problem of diversification and stochastic interest rates can still occur in the dynamic case. In what follows, in order to avoid the above diversification and interest rate problems, we do not assume cash invariance axiom in any of the results. In addition, we do not use monotonicity as well. For instance, monotonicity does not hold for variance or standard deviation.*

Similar to non-dynamic risk measures, the structure of the dynamic risk measure  $\rho_t$  is closely related to the properties of its sub-differential set that is defined in the next elementary lemma which we mention for the sake of completeness.

**Lemma 2.1.** *The sub-differential set  $\Delta$  of the dynamic risk measure  $\rho_t$ , defined by*

$$\{(z, k) \in L^2 \times L_t^2, \rho_t(y) \geq \mathbb{E}[zy | \mathcal{F}_t] + k, \forall y \in L^2\},$$

*is convex<sup>2</sup> and closed with respect to product topologies  $\sigma(L^2, L^2) \times \sigma(L_t^2, L_t^2)$ , where  $\sigma(L^2, L^2)$  and  $\sigma(L_t^2, L_t^2)$  are respectively the weak topologies on  $L^2$  and  $L_t^2$ .*

*Proof.* The convexity of  $\Delta$  is obvious and due to this property, in order to prove the claim, it is enough to show that  $\Delta$  is strong  $\times$  strong closed.<sup>3</sup> Suppose that  $(z_n, k_n) \in \Delta$  and  $(z_n, k_n) \xrightarrow{s} (z, k)$  where  $\xrightarrow{s}$  represents the strong convergence and  $(z, k) \in L^2 \times L_t^2$ . We show that  $(z, k) \in \Delta$ .

<sup>2</sup>Here, the convexity means that if  $(z_1, k_1) \in \Delta$  and  $(z_2, k_2) \in \Delta$  then  $(\lambda(z_1, k_1) + (1 - \lambda)(z_2, k_2)) \in \Delta$  for all  $0 \leq \lambda \leq 1$ ,  $\lambda \in \mathbb{R}$ .

<sup>3</sup>This can be proved using Hahan-Banach separating hyperplanes theorem.

Since  $\|z_n - z\| \rightarrow 0$ ; by Theorem 4.9 of Brezis (2011), there exists a subsequence  $\{z_{n_k}\}_{k=1}^\infty \subset \{z_n\}$  and  $h \in L^2$  such that  $z_{n_k} \rightarrow z$ , a.s., and  $|z_{n_k}| \leq h$ , a.s.,  $\forall k \geq 1$ , on  $L^2$ . Now by using Lebesgue dominated convergence theorem for conditional expectation, we have

$$\mathbb{E}[yz_{n_k} | \mathcal{F}_t] \rightarrow \mathbb{E}[yz | \mathcal{F}_t], \text{ a.s., } \forall y \in L^2.$$

On the other hand,  $\forall y \in L^2$ ,  $\rho_t(y) \geq \mathbb{E}[yz_{n_k} | \mathcal{F}_t] + k_n$  a.s., and we also have that  $k_n \xrightarrow{s} k$  in  $L^2$ , hence  $\|k_{n_k} - k\| \rightarrow 0$ . Therefore, there is a subsequence  $\{k_{n_{k_j}}\} \subset \{k_{n_k}\}$  such that  $\{k_{n_{k_j}}\} \rightarrow k$  a.s. and  $\mathbb{E}[yz_{n_{k_j}} | \mathcal{F}_t] \rightarrow \mathbb{E}[yz | \mathcal{F}_t]$ , which leads to  $\rho_t(y) \geq \mathbb{E}[yz] + k$ , a.s.. This shows that  $\Delta$  is strong $\times$ strong closed, and so it is weak $\times$ weak closed as well.  $\square$

Representation of dynamic risk measures is expressed by the essential supremum of a family of random variables. Though it is not always the case, it is of interest that this family is directed upward. We call a family of random variables directed upward if for the maximum of any two members of this family, there is a third member of the family which bounds them from the above. Then by Theorem A.33 of Föllmer and Schied (2011) the essential supremum of this family can be approximated by an increasing sequence of the members of the family. This theorem has been used a few times in the next section.

### 3 The main results

In the next proposition, we provide a representation of a dynamic risk measure that is strong-weak continuous and convex. Note that the representation of the proposition is still valid for strong-strong continuous dynamic risk measures as this concludes strong-weak continuity of the measure.

**Proposition 3.1.** *Suppose that  $\rho_t : L^2 \mapsto L_t^2$  is a normalized dynamic risk measure that satisfies convexity and is strong-weak continuous with respect to the norm topology on  $L^2$  and weak topology  $\sigma(L_t^2, L_t^2)$  on  $L_t^2$ . Then  $\rho_t$  admits the following representation:*

$$\rho_t(y) = \text{ess sup} \{ \mathbb{E}[zy | \mathcal{F}_t] + k; (z, k) \in \Delta \}, \quad (1)$$

where  $\Delta = \{(z, k) \in L^2 \times L_t^2; \rho_t(y) \geq \mathbb{E}[zy | \mathcal{F}_t] + k, \forall y \in L^2\}$  is convex and closed with respect to product topology  $\sigma(L^2, L^2) \times \sigma(L_t^2, L_t^2)$ .

*Proof.* Define the function  $f : L^2 \mapsto \mathbb{R}$  by  $f(y) = \mathbb{E}[\rho_t(y)]$ . Since  $\rho_t$  is strong-weak continuous then it is easy to show that  $f$  is also continuous with respect to the norm topology on  $L^2$  and

strong topology on  $\mathbb{R}$ . In addition, due to the convexity of  $\rho_t$ , function  $f$  is also convex, and by Lemma A.1 in the appendix, we can write

$$f(y) = \sup \{ \mathbb{E}[zy] + k; (z, k) \in \Delta' \}, \quad (2)$$

where

$$\Delta' = \{ (z, k) \in L^2 \times \mathbb{R}, f(y) \geq \mathbb{E}[zy] + k, \forall y \in L^2 \}.$$

Suppose that  $(z, k) \in \Delta'$ , and for each  $t \geq 0$ , we define the random variable:

$$k_t = \operatorname{ess\,inf}_{y \in L^2} \{ \rho_t(y) - \mathbb{E}[zy | \mathcal{F}_t]; k \leq \rho_t(y) - \mathbb{E}[zy | \mathcal{F}_t] \}.$$

Note that  $k \leq 0$  therefore, the random variable  $y = 0$  satisfies the equation  $k \leq \rho_t(y) - \mathbb{E}[zy | \mathcal{F}_t]$ , and so the essential infimum is not taken over an empty set. By Theorem A.33 of Föllmer and Schied (2011),  $k_t$  is a well-defined random variable as it is  $\mathcal{F}_t$ -measurable.

In addition, for a fixed  $k$ , we have  $k \leq k_t \leq 0$ , almost surely. As a result, for each element  $(z, k)$  in  $\Delta'$ , we can construct  $k_t \in L_t^2$  such that  $\mathbb{E}[k_t] \geq k$  and  $(z, k_t) \in \Delta$ , so from equation (2), we have

$$\begin{aligned} f(y) &= \sup \{ \mathbb{E}[\mathbb{E}[zy | \mathcal{F}_t]] + k; (z, k) \in \Delta' \} \\ &= \sup \{ \mathbb{E}[\mathbb{E}[zy | \mathcal{F}_t] + k]; (z, k) \in \Delta \} \\ &\leq \mathbb{E}[\operatorname{ess\,sup} \{ \mathbb{E}[zy | \mathcal{F}_t] + k; (z, k) \in \Delta \}]. \end{aligned}$$

Therefore, we obtain the following inequality

$$\mathbb{E}[\rho_t(y)] \leq \mathbb{E}[\operatorname{ess\,sup} \{ \mathbb{E}[zy | \mathcal{F}_t] + k; (z, k) \in \Delta \}]. \quad (3)$$

Also from the definition of  $\Delta$ , it is obvious that  $\rho_t(y) \geq \operatorname{ess\,sup} \{ \mathbb{E}[zy | \mathcal{F}_t] + k; (z, k) \in \Delta \}$ , hence

$$\mathbb{E}[\rho_t(y) - \operatorname{ess\,sup} \{ \mathbb{E}[zy | \mathcal{F}_t] + k; (z, k) \in \Delta \}] = 0. \quad (4)$$

From equation (4), we obtain the required representation

$$\rho_t(y) = \operatorname{ess\,sup} \{ \mathbb{E}[zy | \mathcal{F}_t] + k; (z, k) \in \Delta \}, \text{ almost surely.}$$

By Lemma 2.1, the set  $\Delta$  is convex and closed with respect to  $\sigma(L^2, L^2) \times \sigma(L_t^2, L_t^2)$ -topology. This completes the proof. □

Proposition 3.1 is similar to Theorem 1 of Zowe (1975) where a general duality theorem is provided for convex maps that are defined on topological vector spaces. However, the proof of Proposition 3.1 is made independently, and the proposition is more general in the sense that it is valid for convex maps that are strong-weak continuous rather than just strong-strong continuous. In addition, Proposition 3.1 specifies the sub-differential of the convex map (here, the dynamic risk measure).

Similar static risk measures (i.e. for  $\rho : L^2 \mapsto \mathbb{R}$ ) to (1) are applied in Balbás et al. (2010) where they used these real valued risk measures to minimize risk functions using saddle point conditions. The next proposition is somehow the converse of Proposition 3.1 where we show that a weak $\times$ weak compactness of the sub-differential leads to the strong-weak continuity of the dynamic risk measure.

**Proposition 3.2.** *Assume that  $\rho_t : L^2 \mapsto L_t^2$  is a dynamic risk measure that admits the following representation*

$$\rho_t(y) = \text{ess sup} \{ \mathbb{E}[zy | \mathcal{F}_t] + k; (z, k) \in \Delta \},$$

where  $\Delta \subset \{(z, k) \in L^2 \times L_t^2; \rho_t(y) \geq \mathbb{E}[zy | \mathcal{F}_t] + k, \forall y \in L^2\}$  is a compact set under the product topology  $\sigma(L^2, L^2) \times \sigma(L_t^2, L_t^2)$ . Then the dynamic risk measure  $\rho_t$  is strong-weak continuous.

*Proof.* The function  $\rho_t$  is convex because, any  $\alpha \in L_t^2$  is  $\mathcal{F}_t$ -measurable, and for all  $y$  and  $w$  in  $L^2$ , we have:

$$\begin{aligned} \rho_t(\alpha y + (1 - \alpha)w) &= \text{ess sup} \{ \alpha \mathbb{E}[zy | \mathcal{F}_t] + (1 - \alpha) \mathbb{E}[zw | \mathcal{F}_t] + k, (z, k) \in \Delta \} \\ &\leq \alpha \rho_t(y) + (1 - \alpha) \rho_t(w). \end{aligned}$$

In order to show the strong-weak continuity of  $\rho_t$ , it is enough to show that  $\varphi \circ \rho_t : L^2 \mapsto \mathbb{R}$  is strong-strong continuous for all  $\varphi$  in the dual space of  $L_t^2$ , denoted by  $(L_t^2)^*$ . But  $(L_t^2)^* = L_t^2$ , hence any such  $\varphi$  corresponds to an element  $y^* \in L_t^2$  such that

$$\varphi(\rho_t(y_n)) = \int \rho_t(y_n) y^* dP = \mathbb{E}[\rho_t(y_n) y^*].$$

Therefore, if we define the function  $f : L^2 \mapsto \mathbb{R}$  by  $f(y) = \mathbb{E}[\rho_t(y) y^*]$ , for an arbitrary fixed  $y^* \in L_t^2$ , in order to prove our claim, it is enough to show the continuity of this function with respect to the norm topology and strong topology on  $\mathbb{R}$ . Furthermore, without any loss in generality, one can assume that  $y^* \geq 0$ , because for a general  $y^*$ , the decomposition  $y^* = (y^*)^+ - (y^*)^-$  can be applied where  $(y^*)^+ = \max(y^*, 0)$  and  $(y^*)^- = \max(-y^*, 0)$  are both non-negative. With this decomposition, then  $(y^*)^+ + (y^*)^- = |y^*|$ ,  $\|(y^*)^+\| \leq \|y^*\|$ ,  $\|(y^*)^-\| \leq \|y^*\|$ , and

$f(y) = g(y) - h(y)$ , where  $g$  and  $h$  are given by  $g(y) = \mathbb{E}[\rho_t(y)(y^*)^+]$ ,  $h(y) = \mathbb{E}[\rho_t(y)(y^*)^-]$ . If one proves that functions  $g$  and  $h$  are continuous, then so is  $f$ .

It is obvious that  $f$  is convex, so we only need to show that it is continuous at zero<sup>4</sup>. Therefore, we suppose that  $y_n \xrightarrow{s} 0$  and we show that  $f(y_n) \rightarrow f(0)$  in the strong topology of  $\mathbb{R}$ . In what follows, it is important to note that  $\Delta$  is a weak  $\times$  weak compact set, so its projection on  $L^2$  or  $L_t^2$ , such as  $\{z \in L^2; (z, k) \in \Delta \text{ for some } k \in L_t^2\}$ , is weakly compact and hence bounded.

Note that the set  $\{\mathbb{E}[zy | \mathcal{F}_t] + k; (z, k) \in \Delta\}$  is not necessarily directed upwards, and therefore, we cannot simply approximate  $\rho_t(y)$  by a monotone sequence of the elements of this set. We fix this problem through a different approximation of  $\rho_t(y)$  and its sub-differential. The rest of the proof is divided into four steps.

**Step 1:** For a fixed  $m$  and  $y \in L^2$ , define  $\psi(y) = \lim_{m \rightarrow \infty} \rho_t^{(m)}(y)$ , where the limit is taken pointwise on the probability space,

$$\rho_t^{(m)}(y) = \text{ess sup}_{(z,k) \in \Delta_0} \mathbb{E}[zy 1_{\{|z| \leq m\}} | \mathcal{F}_t],$$

$$\Delta_0 = \{(z, k) \in L^2 \times L_t^0; \rho_t(y) \geq \mathbb{E}[zy | \mathcal{F}_t] + k, \forall y \in L^2\},$$

and  $L_t^0$  denotes the space of  $\mathcal{F}_t$ -measurable random variables. Each  $\rho_t^{(m)}(y)$  is  $\mathcal{F}_t$ -measurable and so is  $\psi(y)$ . For a fixed  $m, y$ , we have  $\mathbb{E}[zy 1_{\{|z| \leq m\}} | \mathcal{F}_t] \leq \text{ess sup}_{(z,k) \in \Delta_0} \mathbb{E}[zy 1_{\{|z| \leq m\}} | \mathcal{F}_t]$ , for all  $(z, k) \in \Delta$ ; by taking the limit as  $m$  goes to infinity and using monotone convergence theorem, we get  $\mathbb{E}[zy | \mathcal{F}_t] \leq \psi(y)$ . Therefore, we obtain

$$\text{ess sup}_{(z,k) \in \Delta} \mathbb{E}[zy | \mathcal{F}_t] \leq \psi(y). \quad (5)$$

**Step 2:** We show that for a fixed  $m$ , the set

$$\mathcal{A} = \{\mathbb{E}[zy 1_{\{|z| \leq m\}} | \mathcal{F}_t]; (z, k) \in \Delta_0, \text{ for some } k \in L_t^0\},$$

is directed upward. Consider the event  $A = \{\mathbb{E}[z_1 y | \mathcal{F}_t] \geq \mathbb{E}[z_2 y | \mathcal{F}_t]\}$ , for some  $k_1$  and  $k_2$ , such that  $(z_1, k_1)$  and  $(z_2, k_2)$  are in  $\Delta_0$ . Define  $z = |z_1| 1_{|z_1| \leq m} 1_A + |z_2| 1_{|z_2| \leq m} 1_{A^c}$ . Then, we observe that  $A \in \mathcal{F}_t$  and

$$\mathbb{E}[zy 1_{\{|z| \leq m\}} | \mathcal{F}_t] \geq \max(\mathbb{E}[z_1 y 1_{\{|z_1| \leq m\}} | \mathcal{F}_t], \mathbb{E}[z_2 y 1_{\{|z_2| \leq m\}} | \mathcal{F}_t]).$$

Note that by the definition of  $z$ , we have  $z = z 1_{\{|z| \leq m\}}$ .

---

<sup>4</sup>It can be proved that the continuity of a proper convex function on normed spaces is equivalent to its continuity at origin.

On the other hand, for  $i = 1, 2$ , we have  $\rho_t(y) \geq \mathbb{E}[z_i y | \mathcal{F}_t] + k_i$  for all  $y \in L^2$ . Therefore, we obtain

$$\begin{aligned} 1_A \rho_t(y) &\geq \mathbb{E}[|z_1| y 1_{\{|z_1| \leq m\}} 1_A | \mathcal{F}_t] + (k'_1) 1_A, \\ 1_{A^c} \rho_t(y) &\geq \mathbb{E}[|z_2| y 1_{\{|z_2| \leq m\}} 1_{A^c} | \mathcal{F}_t] + (k'_2) 1_{A^c}, \end{aligned}$$

where  $k'_i = \text{ess inf}_{y \in L^2} \{\mathbb{E}[z_i y 1_A | \mathcal{F}_t] - \mathbb{E}[|z_i| y 1_{\{|z_i| \leq m\}} 1_A | \mathcal{F}_t] + k_i\}$  for  $i = 1, 2$ . Hence, it is possible to construct a  $K \in L_t^0$  such that  $\rho_t(y) \geq \mathbb{E}[z y | \mathcal{F}_t] + K$  which means that  $(z, K) \in \Delta_0$ . This proves that  $\mathcal{A}$  is directed upward.

**Step 3:** From definition of  $\rho_t^{(m)}$  in Step 1, and by Step 2, there exists a sequence  $\{Z_{m_i}\}_{i \geq 1} \subset L^2$  such that  $(Z_{m_i}, K_{m_i}) \in L^2 \times L_t^0$  and we have the pointwise convergence

$$\rho_t^{(m)}(y) = \lim_{i \rightarrow \infty} \mathbb{E}[|z_{m_i}| y 1_{\{|z_{m_i}| \leq m\}} | \mathcal{F}_t], \text{ i.e. } \psi(y) = \lim_{m \rightarrow \infty} \lim_{i \rightarrow \infty} \mathbb{E}[|z_{m_i}| y 1_{\{|z_{m_i}| \leq m\}} | \mathcal{F}_t].$$

Now, suppose that  $y^* \in L_t^2$  and  $y^* \geq 0$ , then  $\psi(y) y^* = \lim_{m \rightarrow \infty} \lim_{i \rightarrow \infty} \mathbb{E}[|z_{m_i}| y y^* 1_{\{|z_{m_i}| \leq m\}} | \mathcal{F}_t]$ , since  $y^*$  is  $\mathcal{F}_t$  measurable. By Fatou's lemma we obtain

$$\mathbb{E}[\psi(y) y^*] \leq \liminf \|Z_{m_i}\| \|y\| \|y^*\|.$$

Since  $\Delta$  is weakly compact then  $\liminf \|Z_{m_i}\| < \infty$  and therefore,  $\lim_{\|y\| \rightarrow 0} \mathbb{E}[\psi(y) y^*] \leq 0$ . On the other hand, by equation (5) and definition of  $\rho_t(y)$ , we obtain  $\rho_t(y) \leq \psi(y) + \rho_t(0)$ . Since  $y^* \geq 0$ , this leads to  $\lim_{\|y\| \rightarrow 0} \mathbb{E}[\rho_t(y) y^*] \leq f(0)$ , i.e.,

$$\lim_{\|y\| \rightarrow 0} f(y) \leq f(0). \quad (6)$$

**Step 4:** For simplicity let  $\Delta' = \{(z, k) \in L^2 \times L_t^2; \rho_t(y) \geq \mathbb{E}[z y | \mathcal{F}_t] + k, \forall y \in L^2\}$ . Suppose that  $(z, k) \in \Delta'$  is fixed then for all  $y \in L^2$ ,  $\rho_t(y) \geq \mathbb{E}[z y | \mathcal{F}_t] + k$ . Therefore for any  $y^* \geq 0$  in  $L^2$ , we obtain  $\mathbb{E}[\rho_t(y) y^*] \geq \mathbb{E}[z y y^*] + \mathbb{E}[k y^*]$ ;  $\lim_{\|y\| \rightarrow 0} |\mathbb{E}[z y y^*]| \leq \lim_{\|y\| \rightarrow 0} \|z\| \|y\| \|y^*\| = 0$ . Hence for all  $k \in L_t^2$  such that  $(z, k) \in \Delta'$ , we get

$$\lim_{\|y\| \rightarrow 0} \mathbb{E}[\rho_t(y) y^*] \geq \lim_{\|y\| \rightarrow 0} \mathbb{E}[k y^*] = \mathbb{E}[k y^*]. \quad (7)$$

Similar to Step 2 and in fact much easier, one can show that  $\{k; (z, k) \in \Delta'\}$  is directed upward. Hence, there is an increasing sequence  $\{k_n\}_{n \geq 1} \subset L_t^2$  such that  $(z_n, k_n) \in \Delta'$  for some  $\{z_n\}_{n \geq 1} \subset L^2$ , and  $k_n \rightarrow \text{ess sup}\{k; (z, k) \in \Delta'\}$ . Therefore, by monotone convergence theorem,  $\lim_{n \rightarrow \infty} \mathbb{E}[k_n y^*] = \mathbb{E}[\text{ess sup}\{k; (z, k) \in \Delta'\} y^*]$ , but  $\text{ess sup}\{k; (z, k) \in \Delta'\} \geq \text{ess sup}\{k; (z, k) \in \Delta\}$  (because  $\Delta \subset \Delta'$ ) so  $\lim_{n \rightarrow \infty} \mathbb{E}[k_n y^*] \geq f(0) = \mathbb{E}[\text{ess sup}\{k; (z, k) \in \Delta\} y^*]$ . Also, equation (7) holds for all  $k$  such that  $(z, k) \in \Delta'$  for some  $z$ . In particular, one can take the sequence  $\{k_n\}_{n \geq 1}$ . Hence,  $\lim_{\|y\| \rightarrow 0} \mathbb{E}[\rho_t(y) y^*] \geq \lim_{n \rightarrow \infty} \mathbb{E}[k_n y^*] \geq f(0)$ , or  $\lim_{\|y\| \rightarrow 0} f(y) \geq f(0)$ . This together with equation (6), leads to  $\lim_{\|y\| \rightarrow 0} f(y) = f(0)$  which shows that  $f$  is strongly continuous at origin with respect to the norm topology. This completes the proof.  $\square$

The representation of Proposition 3.2 can be used to create risk measures that are strong-weak continuous, and despite its simplicity, many interesting risk measures are still covered. For example, in the non-dynamic case, i.e., for  $t = 0$ , all the following risk measures are special case of this representation: down side semi-deviations, CVaR, Wang's distortion measure, the absolute deviation, and the standard deviation. Proposition 3.2 can be used to create the dynamic version of these risk measures. Now, we investigate when the essential supremum in Proposition 3.2 is actually attained.

**Lemma 3.1.** *Assume that  $\rho_t : L^2 \mapsto L_t^2$  admits the following representation*

$$\rho_t(y) = \text{ess sup} \{ \mathbb{E}[zy | \mathcal{F}_t] + k; (z, k) \in \Delta \}, \quad \rho_t(0) = 0,$$

where  $\Delta \subset L^2 \times L_t^2$  is a convex compact set under the topology  $\sigma(L^2, L^2) \times \sigma(L_t^2, L_t^2)$ . We further assume that the set  $\{ \mathbb{E}[zy | \mathcal{F}_t] + k; (z, k) \in \Delta \}$  is directed upwards. Then the essential supremum is attained, i.e., there is an element  $(z^*, k^*) \in \Delta$  such that  $\rho_t(y) = \mathbb{E}[z^*y | \mathcal{F}_t] + k^*$ , almost surely.<sup>5</sup>

*Proof.* Since the set  $\{ \mathbb{E}[zy | \mathcal{F}_t] + k; (z, k) \in \Delta \}$  is directed upwards, there exists a sequence  $\{(z_n, k_n)\}_{n=1}^\infty$  in  $\Delta$  such that  $\mathbb{E}[z_n y | \mathcal{F}_t] + k_n \leq \mathbb{E}[z_{n+1} y | \mathcal{F}_t] + k_{n+1}$  and  $\rho_t(y) = \lim_{n \rightarrow \infty} \mathbb{E}[z_n y | \mathcal{F}_t] + k_n$ .

Define the projection  $\phi_1 : L^2 \times L_t^2 \mapsto L^2$  by  $\phi_1(x, y) = x$ . Then  $\phi_1(\Delta)$  is compact with respect to the weak topology  $\sigma(L^2, L^2)$  and hence it is strongly bounded. Since  $L^2$  is a reflexive space, by Theorem 3.18 of Brezis (2011), there is a subsequence  $\{z_{n_j}\}$  that converges weakly to an element  $z$  in  $L^2$ . By considering this subsequence, we still have the convergence  $\rho_t(y) = \lim_{j \rightarrow \infty} \mathbb{E}[z_{n_j} y | \mathcal{F}_t] + k_{n_j}$ .

Similarly by projection on the space  $L_t^2$ , we can construct a subsequence  $\{k_{n_j}\}$  of  $\{k_{n_j}\}$  that converges weakly to an element  $k$  in  $L^2$  with respect to the weak topology  $\sigma(L_t^2, L_t^2)$ . Since  $\{n_j\}$  is a subsequence of  $\{n_j\}$ , we have both the weak  $z_{n_j} \rightharpoonup z$  and the strong convergence

$$\rho_t(y) = \lim_{i \rightarrow \infty} \mathbb{E}[z_{n_{j_i}} y | \mathcal{F}_t] + k_{n_{j_i}}. \quad (8)$$

Note that  $L^2 \times L_t^2$  is a Hausdorff space, and  $\Delta$  is compact with respect to the product topology  $\sigma(L^2, L^2) \times \sigma(L_t^2, L_t^2)$ , so  $\Delta$  is also closed with respect to this topology. Because of the convergence  $(z_{n_{j_i}}, k_{n_{j_i}}) \rightarrow (z, k)$  in topology  $\sigma(L^2, L^2) \times \sigma(L_t^2, L_t^2)$ , we obtain  $(z, k) \in \Delta$ . Now we show that  $\rho_t(y) = \mathbb{E}[zy | \mathcal{F}_t] + k$ .

---

<sup>5</sup>Note that if the image space of  $\rho_t$  is a subspace of real numbers then this is rather a trivial result from functional analysis facts.

For every  $i$ ,  $\mathbb{E}[z_{n_i}y|\mathcal{F}_t] + k_{n_i} \in L_t^2$ , therefore for a fixed  $y \in L^2$ , by equation (8) and monotone convergence theorem we obtain

$$\mathbb{E}[\rho_t(y)] = \lim_{i \rightarrow \infty} \mathbb{E}[z_{n_i}y] + \mathbb{E}[k_{n_i}].$$

However, both  $z_{n_i}$  and  $k_{n_i}$  approaches weakly to  $z$  and  $k$  respectively, hence  $\mathbb{E}[z_{n_i}y] \rightarrow \mathbb{E}[zy]$  and  $\mathbb{E}[k_{n_i}] \rightarrow \mathbb{E}[k]$  which leads to  $\mathbb{E}[\rho_t(y)] = \mathbb{E}[zy] + \mathbb{E}[k]$  or  $\mathbb{E}[\rho_t(y) - \mathbb{E}[zy|\mathcal{F}_t] - k] = 0$ . On the other hand, by the definition of essential supremum,  $\rho_t(y) \geq \mathbb{E}[zy|\mathcal{F}_t] + k$ , hence  $\rho_t(y) = \mathbb{E}[zy|\mathcal{F}_t] + k$ , almost surely.  $\square$

In view of the above lemma, the next interesting question is whether or not we can approximate a continuous dynamic risk measure with continuous risk measures that admit weak $\times$ weak compact sub-differentials as this with the additional assumption ‘‘directed upward’’ guarantees that the essential supremum can be approximated. The following theorem answers this question.

**Theorem 3.1.** *Suppose that  $\rho_t : L^2 \mapsto L_t^2$  is a dynamic risk measure which is convex and strong-weak continuous. Then, there is a sequence of dynamic risk measures  $\rho_t^{(n)} : L^2 \mapsto L_t^2$ , with  $\sigma(L^2, L^2) \times \sigma(L_t^2, L_t^2)$  compact sub-differential sets, which are convex and strong-weak continuous such that  $\rho_t^{(n)}(y)$  converges to  $\rho_t(y)$  almost surely for all  $y \in L^2$ .*

*Proof.* Without any loss in generality, we can assume that  $\rho_t(0) = 0$ , otherwise, one can consider  $\psi(y) = \rho_t(y) - \rho_t(0)$ . First, we construct the sequence of  $\rho_t^{(n)}(y)$  for  $y \in L^2$ . From Proposition 3.1, we have the following representation for  $\rho_t(y)$ :

$$\rho_t(y) = \text{ess sup} \{ \mathbb{E}[zy | \mathcal{F}_t] + k_t; (z, k_t) \in \Delta \},$$

where  $\Delta = \{(z, k) \in L^2 \times L_t^2, \rho_t(y) \geq \mathbb{E}[zy | \mathcal{F}_t] + k, \forall y \in L^2\}$  is convex and closed with respect to product topologies  $\sigma(L^2, L^2) \times \sigma(L_t^2, L_t^2)$ . For each  $n \geq 1$ , define

$$\rho_t^{(n)}(y) = \text{ess sup} \{ \mathbb{E}[zy | \mathcal{F}_t] + k_t; (z, k_t) \in \Delta \cap (\bar{B}_n(L^2) \times \bar{B}_n(L_t^2)) \}, \quad (9)$$

where  $\bar{B}_n(L^2)$  and  $\bar{B}_n(L_t^2)$  are respectively the closures of open balls of radius  $n > 0$  in  $L^2$  and  $L_t^2$  spaces. Since  $L^2$  and  $L_t^2$  are reflexive Banach spaces, then  $\Delta \cap (\bar{B}_n(L^2) \times \bar{B}_n(L_t^2))$  is compact under the product topology  $\sigma(L^2, L^2) \times \sigma(L_t^2, L_t^2)$ .

For any  $y \in L^2$ , we can assume that  $\rho_t^{(n)}(y) \in L_t^2$ , for all  $n \geq 1$ . This is because one can take  $n$  sufficiently large such that  $\mathbb{E}[z^*y | \mathcal{F}_t] + k_t^* \leq \rho_t^{(n)}(y) \leq \rho_t(y)$  for some  $(z^*, k_t^*) \in L^2 \times L_t^2$ . Now note that since

$$\Delta \cap (\bar{B}_n(L^2) \times \bar{B}_n(L_t^2)) \subset \{(z, k) \in L^2 \times L_t^2, \rho_t^{(n)}(y) \geq \mathbb{E}[zy | \mathcal{F}_t] + k, \forall y \in L^2\},$$

by Proposition 3.2,  $\rho_t^{(n)}$  is strong-weak continuous.

Finally, we show the convergence. There is a sequence  $\{\phi_i\}_{i \geq 1}$ , defined by  $\phi_i = \mathbb{E}[z_i y | \mathcal{F}_t] + k_i$ ,  $(z_i, k_i) \in \Delta$  such that  $\max_{1 \leq i \leq n}(\phi_i)$  converges to  $\rho_t(y)$  as  $n$  goes to infinity<sup>6</sup>. For the rest of the argument, we consider a fixed element  $\omega \in \Omega$ , since  $\{\rho_t^{(n)}(y)\}_{n \geq 1}$  is increasing, one can find a strictly increasing subsequence  $\{n_i\}_{i \geq 1}$  such that  $\max_{1 \leq j \leq n_i}(\phi_j) \leq \rho_t^{(n_i)}(y)$  for all  $i \geq 1$ . On the other hand,  $\rho_t^{(n)}(y) \leq \rho_t(y)$  for all  $y$  and  $n \geq 1$ . This proves that  $\rho_t^{(n)}(y) \rightarrow \rho_t(y)$  almost surely as  $n \rightarrow \infty$ .  $\square$

Suppose that we are analyzing a convex risk measure problem in which the risk measure is strong-weak continuous. Proposition 3.1 provides a representation of such a risk measure with a sub-differential set  $\Delta$  that is closed with respect to the product of weak topologies. Then by Theorem 3.1 one can assume that  $\Delta$  is compact with respect to the product of weak topologies in the sense that it is possible to approximate the risk measure with a sequence of risk measures that admit compact sub-differential sets.

## 4 Conclusion

A representation of dynamic risk measures that are convex and strong-strong or strong-weak continuous is provided. It is also proved that if a risk measure satisfies such representation with a weak $\times$ weak compact sub-differential then it is strong-weak continuous. We then further discuss a sufficient condition under which the optimal value in the representation theorem is actually attained. Finally, an approximation procedure is carried out for those convex risk measures that are strong-weak continuous, but they do not necessarily admit a weak $\times$ weak compact sub-differential set.

## 5 Acknowledgments

The first author is very thankful to Alejandro Balbás for helping with the proof of Lemma A.1. The authors are also grateful to an anonymous referee for his/her constructive comments.

## References

- Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath  
 1999. Coherent measures of risk. *Math. Finance*, 9(3):203–228.

---

<sup>6</sup>Note that this is true whether or not the set  $\{\mathbb{E}[z y | \mathcal{F}_t] + k; (z, k) \in \Delta\}$  is directed upward.

- Assa, H.  
 2011. Lebesgue property of convex risk measures for bounded càdlàg processes. *Methods Appl. Anal.*, 18(3):335–349.
- Balbás, A., B. Balbás, and R. Balbás  
 2010. Minimizing measures of risk by saddle point conditions. *J. Comput. Appl. Math.*, 234(10):2924–2931.
- Brezis, H.  
 2011. *Functional analysis, Sobolev spaces and partial differential equations*, Universitext. Springer, New York.
- Cerreia-Vioglio, S., F. Maccheroni, M. Marinacci, and L. Montrucchio  
 2011a. Complete monotone quasiconcave duality. *Math. Oper. Res.*, 36(2):321–339.
- Cerreia-Vioglio, S., F. Maccheroni, M. Marinacci, and L. Montrucchio  
 2011b. Risk measures: rationality and diversification. *Math. Finance*, 21(4):743–774.
- Cheridito, P., F. Delbaen, and M. Kupper  
 2006. Dynamic monetary risk measures for bounded discrete-time processes. *Electron. J. Probab.*, 11:no. 3, 57–106.
- Cheridito, P. and M. Kupper  
 2011. Composition of time-consistent dynamic monetary risk measures in discrete time. *Int. J. Theor. Appl. Finance*, 14(1):137–162.
- Detlefsen, K. and G. Scandolo  
 2005. Conditional and dynamic convex risk measures. *Finance Stoch.*, 9(4):539–561.
- El Karoui, N. and C. Ravanelli  
 2009. Cash subadditive risk measures and interest rate ambiguity. *Math. Finance*, 19(4):561–590.
- Föllmer, H. and A. Schied  
 2011. *Stochastic finance*, extended edition. Walter de Gruyter & Co., Berlin. An introduction in discrete time.
- Frittelli, M. and M. Maggis  
 2011. Dual representation of quasi-convex conditional maps. *SIAM J. Financial Math.*, 2(1):357–382.

Frittelli, M. and M. Maggis

2014. Complete duality for quasiconvex dynamic risk measures on modules of the  $L^p$ -type. *Stat. Risk Model.*, 31(1):103–128.

Riedel, F.

2004. Dynamic coherent risk measures. *Stochastic Process. Appl.*, 112(2):185–200.

Rockafellar, R.

1966. Characterization of the subdifferentials of convex functions. *Pacific Journal of Mathematics*, 17(3):497–510.

Zowe, J.

1975. A duality theorem for a convex programming problem in order complete vector lattices. *J. Math. Anal. Appl.*, 50:273–287.

## A Appendix

Consider a Banach space  $Y$  with its dual space  $Z$ . As usual, the scalar product of  $y \in Y$  and  $z \in Z$  is defined and represented by  $z(y)$  and  $\langle y, z \rangle$ , respectively. A real valued function  $\rho : Y \mapsto \mathbb{R}$  is said to be convex if

$$\rho(\lambda y_1 + (1 - \lambda)y_2) \leq \lambda \rho(y_1) + (1 - \lambda)\rho(y_2),$$

holds for every  $y_1, y_2 \in Y$  and every  $0 \leq \lambda \leq 1$ .

**Lemma A.1.** *Suppose that  $\rho, \rho(0) = 0$ , is convex and continuous with respect to the topology induced by the norm of the space. Consider the sub-differential set*

$$\partial = \{(z, k) \in Z \times \mathbb{R}; \langle y, z \rangle + k \leq \rho(y) \quad \forall y \in Y\},$$

*then  $\partial$  is convex and closed with respect to the product topology of  $\sigma(Z, Y)$  and the strong topology on  $\mathbb{R}$ . Furthermore, for every  $y \in Y$ , the following representation holds:*

$$\rho(y) = \sup \{\langle y, z \rangle + k; (z, k) \in \partial\}.$$

*Proof.* It is easy to see that  $\partial$  is convex and closed with respect to the aforementioned topology. The inequality  $\rho(y) \geq \sup \{\langle y, z \rangle + k; (z, k) \in \partial\}$  is obvious too, so let us prove the opposite inequality. Consider  $y_0 \in Y$  and  $\varepsilon > 0$ , it is sufficient to prove that  $\rho(y_0) - \varepsilon \leq \sup \{\langle y, z \rangle + k; (z, k) \in \partial\}$ . It is obvious that  $(y_0, \rho(y_0) - \varepsilon) \notin A$ , where

$$A = \{(y, r) \in Y \times \mathbb{R}; r \geq \rho(y)\}.$$

Since  $\rho$  is convex and continuous, then the set  $A$  is closed and convex. In addition  $A$  is non-empty as  $\rho(0) = 0$ , and therefore, the Hahn-Banach Separation Theorem proves the existence of a non-null element  $(z_0, k) \in Z \times \mathbb{R}$  such that  $\langle y, z_0 \rangle + rk \geq \langle y_0, z_0 \rangle + k(\rho(y_0) - \varepsilon)$ , for every  $(y, r) \in A$ .

Note that  $k \geq 0$ , otherwise the latter inequality could not hold if  $r \rightarrow \infty$ . Moreover,  $k = 0$  is not possible either, since  $\langle y, z_0 \rangle \geq \langle y_0, z_0 \rangle$  for every  $y \in Y$  would lead to  $z_0 = 0$  which is against the fact that  $(z_0, k) \neq (0, 0)$ . Thus, taking  $z_0/k$  instead of  $z_0$ , and denoting it by  $z_0$  again, we have that  $\langle y, z_0 \rangle + r \geq \langle y_0, z_0 \rangle + (\rho(y_0) - \varepsilon)$ , holds for every  $(y, r) \in A$ . In particular, for  $r = \rho(y)$ , the inequality  $\langle y, z_0 \rangle + \rho(y) \geq \langle y_0, z_0 \rangle + (\rho(y_0) - \varepsilon)$ , holds for every  $y \in Y$ , and therefore  $\rho(y) \geq -\langle y, z_0 \rangle + \langle y_0, z_0 \rangle + (\rho(y_0) - \varepsilon)$ , which shows that  $(-z_0, \langle y_0, z_0 \rangle + \rho(y_0) - \varepsilon)$  belongs to  $\partial$ . This implies that

$$\sup \{ \langle y_0, z \rangle + k; (z, k) \in \partial \} \geq -\langle y_0, z_0 \rangle + \langle y_0, z_0 \rangle + \rho(y_0) - \varepsilon = \rho(y_0) - \varepsilon.$$

□