S. Kusuoka
T. Maruyama (Eds.)

Advances in MATHEMATICAL ECONOMICS

Volume 15

Springer

Advances in MATHEMATICAL ECONOMICS

Aims and Scope. The project is to publish *Advances in Mathematical Economics* once a year under the auspices of the Research Center for Mathematical Economics. It is designed to bring together those mathematicians who are seriously interested in obtaining new challenging stimuli from economic theories and those economists who are seeking effective mathematical tools for their research.

The scope of *Advances in Mathematical Economics* includes, but is not limited to, the following fields:

- Economic theories in various fields based on rigorous mathematical reasoning.
- Mathematical methods (e.g., analysis, algebra, geometry, probability) motivated by economic theories.
- Mathematical results of potential relevance to economic theory.
- Historical study of mathematical economics.

Authors are asked to develop their original results as fully as possible and also to give a clear-cut expository overview of the problem under discussion. Consequently, we will also invite articles which might be considered too long for publication in journals.

S. Kusuoka, T. Maruyama (Eds.)

Advances in Mathematical Economics

Volume 15

Springer

Shigeo Kusuoka
Professor
Graduate School of Mathematical Sciences
The University of Tokyo
3-8-1 Komaba, Meguro-ku
Tokyo 153-0041, Japan

Toru Maruyama
Professor
Department of Economics
Keio University
2-15-45 Mita, Minato-ku
Tokyo 108-8345, Japan

ISSN 1866-2226 e-ISSN 1866-2234
ISBN 978-4-431-53929-2 e-ISBN 978-4-431-53930-8
DOI 10.1007/978-4-431-53930-8
Springer Tokyo Dordrecht Heidelberg London New York

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Table of Contents

Research Articles

Historical Perspective

Note

Adv. Math. Econ. 15, 1–26 (2011)

Advances in
MATHEMATICAL
ECONOMICS

Some various convergence results for normal integrands

Charles Castaing

Département de Mathématiques, Faculté des Sciences et Techniques
Université Montpellier II, 34095 Montpellier Cedex, France
(e-mail: charles.castaing@gmail.com)

Received: October 4, 2010
Revised: October 25, 2010

JEL classification: C01, C02

Mathematics Subject Classification (2010): 28B20, 60G42, 46A17, 54A20

Abstract. We prove various results for the conditional expectation of multifunctions and normal integrands. Applications to the epiconvergence of integrand reversed martingales and lower semicontinuous superadditive random sequences are presented.

Key words: conditional expectation, epiconvergence, ergodic, martingale, normal integrands, superadditive

1. Introduction

In this paper, we present some convergence problems for normal integrands involving new tools in the conditional expectation of normal integrands and multifunctions and its applications to the epiconvergence of integrand reversed martingales and parametric Birkhoff–Kingman ergodic theorem. The paper is organized as follows. In Sect. 3 we state and summarize for references some results on the conditional expectation of closed convex valued integrable multifunctions in separable Banach spaces. Section 4 is devoted to the conditional expectation of closed convex valued integrable multifunctions in the dual of a separable Banach space, here the dual is no longer assumed to be separable. Main results on the conditional expectation of normal integrands and the epiconvergence results for integrand reversed martingales and lower semicontinuous superadditive random sequences are presented in Sect. 5.

DOI: 10.1007/978-4-431-53930-8_1,

2. Notations and preliminaries

Throughout this paper $(\Omega, \mathcal{F}, P)$ is a complete probability space, $(\mathcal{F}_n)_{n\in\mathbf{N}}$ is an increasing sequence of sub-σ-algebras of $\mathcal{F}$ such that $\mathcal{F}$ is the σ-algebra generated by $\cup_{n\in\mathbf{N}}\mathcal{F}_n$. E is a separable Banach space and E^* is its topological dual. Let $\overline{B}_E$ (resp. $\overline{B}_{E^*}$) be the closed unit ball of E (resp. E^*) and 2^E the collection of all subsets of E. Let $cc(E)$ (resp. $cwk(E)$) (resp. $\mathcal{R}wk(E)$) be the set of nonempty closed convex (resp. convex weakly compact) (resp. *ball-weakly compact* closed convex) subsets of E, here a closed convex subset in E is *ball-weakly compact* if its intersection with any closed ball in E is weakly compact. For $A \in cc(E)$, the distance and the support function associated with A is defined respectively by

$$d(x, A) = \inf\{\|x - y\| \,:\, y \in A\}, (x \in E)$$

$$\delta^*(x^*, A) = \sup\{\langle x^*, y\rangle \,:\, y \in A\}, (x^* \in E^*).$$

We also define

$$|A| = \sup\{||x|| \,:\, x \in A\}.$$

Given a sub-σ-algebra $\mathcal{B}$ in Ω, a mapping $X : \Omega \to 2^E$ is $\mathcal{B}$-measurable if for every open set U in E the set

$$X^- U := \{\omega \in \Omega : X(\omega) \cap U \neq \emptyset\}$$

is a member of $\mathcal{B}$. A function $f : \Omega \to E$ is a $\mathcal{B}$-measurable selection of X if $f(\omega) \in X(\omega)$ for all $\omega \in \Omega$. A Castaing representation of X is a sequence $(f_n)_{n\in\mathbf{N}}$ of $\mathcal{B}$-measurable selections of X such that

$$X(\omega) = cl\{f_n(\omega), n \in \mathbf{N}\} \quad \forall \omega \in \Omega$$

where the closure is taken with respect to the topology of associated with the norm in E. It is known that a nonempty closed-valued multifunction $X : \Omega \to c(E)$ is $\mathcal{B}$-measurable iff it admits a Castaing representation. If $\mathcal{B}$ is complete, the $\mathcal{B}$-measurability is equivalent to the measurability in the sense of graph, namely the graph of X is a member of $\mathcal{B} \otimes \mathcal{B}(E)$, here $\mathcal{B}(E)$ denotes the Borel tribe on E. A $cc(E)$-valued $\mathcal{B}$-measurable $X : \Omega \to cc(E)$ is integrable if the set $S^1_X(\mathcal{B})$ of all $\mathcal{B}$-measurable and integrable selections of X is nonempty. We denote by $L^1_E(\mathcal{B})$ the space of E-valued $\mathcal{B}$-measurable and Bochner-integrable functions defined on Ω and $\mathcal{L}^1_{cwk(E)}(\mathcal{B})$ the space of all $\mathcal{B}$-measurable multifunctions $X : \Omega \to cwk(E)$, such that $|X| \in L^1_{\mathbf{R}}(\mathcal{B})$. We refer to [5] for the theory of Measurable Multifunctions and Convex Analysis, and to [13, 16] for basic theory of martingales and adapted sequences.

3. Multivalued conditional expectation theorem

Given a sub-σ-algebra, $\mathcal{B}$ of $\mathcal{F}$ and an integrable $\mathcal{F}$-measurable $cc(E)$-valued multifunction $X : \Omega \rightrightarrows E$, Hiai and Umegaki [14] showed the existence of a $\mathcal{B}$-measurable $cc(E)$-valued integrable multifunction, denoted by $E^{\mathcal{B}}X$ such that

$$\mathcal{S}^1_{E^{\mathcal{B}}X}(\mathcal{B}) = cl\{E^{\mathcal{B}}f : f \in \mathcal{S}^1_X(\mathcal{F})\}$$

the closure being taken in $L^1_E(\Omega, \mathcal{A}, P)$; $E^{\mathcal{B}}X$ is the multivalued conditional expectation of X relative to $\mathcal{B}$. If $X \in \mathcal{L}^1_{cwk(E)}(\mathcal{F})$, and the strong dual E^*_b is separable, then $E^{\mathcal{B}}X \in \mathcal{L}^1_{cwk(E)}(\mathcal{B})$ with $\mathcal{S}^1_{E^{\mathcal{B}}X}(\mathcal{B}) = \{E^{\mathcal{B}}f : f \in \mathcal{S}^1_X(\mathcal{F})\}$. A unified approach for general conditional expectation of $cc(E)$-valued integrable multifunctions is given in [17] allowing to recover both the $cc(E)$-valued conditional expectation of $cc(E)$-valued integrable multifunctions in the sense of [14] and the $cwk(E)$-valued conditional expectation of $cwk(E)$-valued integrably bounded multifunctions given in [3]. For more information on multivalued conditional expectation and related subjects we refer to [1, 5, 14, 17]. In the context of this paper we summarize two specific versions of conditional expectation in a separable Banach space and its dual (see Sect. 4).

Proposition 3.1. *Assume that E^*_b is separable. Let $\mathcal{B}$ be a sub-σ-algebra of $\mathcal{F}$ and an integrable $\mathcal{F}$-measurable $cc(E)$-valued multifunction $X : \Omega \rightrightarrows E$. Assume further there is a $\mathcal{F}$-measurable ball-weakly compact $cc(E)$-valued multifunction $K : \Omega \rightrightarrows E$ such that $X(\omega) \subset K(\omega)$ for all $\omega \in \Omega$. Then there is a unique (for the equality a.s.) $\mathcal{B}$-measurable $cc(E)$-valued multifunction Y satisfying the property*

$$(*) \quad \forall v \in L^\infty_{E^*}(\mathcal{B}), \int_\Omega \delta^*(v(\omega), Y(\omega))dP(\omega) = \int_\Omega \delta^*(v(\omega), X(\omega))dP(\omega).$$

$E^{\mathcal{B}}X := Y$ is the conditional expectation of X.

Proof. The proof is an adaptation of the one of Theorem VIII.35 in [5]. Let u_0 be an integrable selection of X. For every $n \in \mathbf{N}$, let

$$X_n(\omega) = X(\omega) \cap (u_0(\omega) + n\overline{B}_E) \quad \forall n \in \mathbf{N} \quad \forall \omega \in \Omega.$$

As $X(\omega) \subset K(\omega)$ for all $\omega \in \Omega$, we get

$$X_n(\omega) = X(\omega) \cap (u_0(\omega) + n\overline{B}_E) \subset K(\omega) \cap (u_0(\omega) + n\overline{B}_E) \quad \forall n \in \mathbf{N}\, \forall \omega \in \Omega.$$

As $K(\omega)$ is ball-weakly compact, it is immediate that $X_n \in \mathcal{L}^1_{cwk(E)}(\mathcal{F})$. so that, by virtue of ([3] or [17], Remarks of Theorem 3), the conditional expectation $E^{\mathcal{B}}X_n \in \mathcal{L}^1_{cwk(E)}(\mathcal{B})$. It follows that

$$(**) \qquad \int_\Omega \delta^*(v(\omega), E^{\mathcal{B}}X_n(\omega))P(d\omega) = \int_\Omega \delta^*(v(\omega), X_n(\omega))P(d\omega).$$

$\forall n \in \mathbf{N}, \forall v \in L^\infty_{E^*}(\mathcal{B})$. Now let

$$Y(\omega) = cl(\cup_{n\in\mathbf{N}} E^{\mathcal{B}}X_n(\omega)) \quad \forall \omega \in \Omega.$$

Then Y is $\mathcal{B}$-measurable and a.s. convex. Now the required property (*) follows from (**) and the monotone convergence theorem. Indeed

$$\forall n \in \mathbf{N}, \forall v \in L^\infty_{E^*}(\mathcal{B}), \langle u_0, v\rangle \leq \delta^*(v, X_n) \uparrow \delta^*(v, X)$$
$$\langle v, E^{\mathcal{B}}u_0\rangle \leq \delta^*(v, E^{\mathcal{B}}X_n) \uparrow \delta^*(v, Y).$$

Now the uniqueness follows exactly as in the proof of Theorem VIII.35 in [5], via the measurable projection theorem ([5], Theorem III.32). ■

4. Conditional expectation in a dual space

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space, $(\mathcal{F}_n)_{n\in\mathbf{N}}$ an increasing sequence of sub σ-algebras of $\mathcal{F}$ such that $\mathcal{F}$ is the σ-algebra generated by $\cup_{n\geq 1}^\infty \mathcal{F}_n$. Let E be a separable Banach space, $D_1 = (x_p)_{p\in\mathbf{N}}$ is a dense sequence in the closed unit ball of E, E^* is the topological dual of E, $\overline{B}_E$ (resp. $\overline{B}_{E^*}$) is the closed unit ball of E (resp. E^*). We denote by E^*_s (resp. E^*_c) (resp. E^*_b) (resp. $E^*_{m^*}$) the topological dual E^* endowed with the topology $\sigma(E^*, E)$ of pointwise convergence, alias w^* topology (resp. the topology τ_c of compact convergence) (resp. the topology s^* associated with the dual norm $||.||_{E^*_b}$) (resp. the topology $m^* = \sigma(E^*, H)$, where H is the linear space of E generated by D, that is the Hausdorff locally convex topology defined by the sequence of semi-norms

$$p_k(x^*) = \max\{|\langle x^*, x_p\rangle| : p \leq k\}, \;\; x^* \in E^*, k \geq 1.$$

Recall that the topology m^* is metrizable by the metric

$$d_{E^*_{m^*}}(x_1^*, x_2^*) := \sum_{p=1}^{p=\infty} \frac{1}{2^p}|\langle x_p, x_1^*\rangle - \langle x_p, x_2^*\rangle|, \quad x_1^*, x_2^* \in E^*.$$

Further we have

$$d_{E^*_{m^*}}(x^*, y^*) \le ||x^* - y^*||_{E^*_b}, \quad \forall x^*, y^* \in E^* \times E^*.$$

We assume from now that $d_{E^*_{m^*}}$ is held fixed. Further, we have $m^* \subset w^* \subset \tau_c \subset s^*$. When E is infinite dimensional these inclusions are strict. On the other hand, the restrictions of m^*, w^* τ_c to any bounded subset of E^* coincide and the Borel tribes $\mathcal{B}(E^*_s)$, $\mathcal{B}(E^*_c)$ and $\mathcal{B}(E^*_{m^*})$ associated with E^*_s E^*_c and $E^*_{m^*}$ are equal. Noting that E^* is the countable union of closed balls, we deduce that the space E^*_s is Suslin, as well as the metrizable topological space $E^*_{m^*}$. A $2^{E^*_s}$-valued multifunction (alias mapping for short) $X : \Omega \rightrightarrows E^*_s$ is $\mathcal{F}$-measurable if its graph belongs to $\mathcal{F} \otimes \mathcal{B}(E^*_s)$. Given a $\mathcal{F}$-measurable mapping $X : \Omega \rightrightarrows E^*_s$ and a Borel set $G \in \mathcal{B}(E^*_s)$, the set

$$X^- G = \{\omega \in \Omega : X(\omega) \cap G \neq \emptyset\}$$

is $\mathcal{F}$-measurable, that is $X^- G \in \mathcal{F}$. In view of the completeness hypothesis on the probability space, this is a consequence of the Projection Theorem (see e.g. Theorem III.23 of [5]) and of the equality

$$X^- G = \text{proj}_\Omega \{Gr(X) \cap (\Omega \times G)\}.$$

Further if $u : \Omega \to E^*_s$ is a scalarly $\mathcal{F}$-measurable mapping, that is, for every $x \in E$, the scalar function $\omega \mapsto \langle x, u(\omega)\rangle$ is $\mathcal{F}$-measurable, then the function $f : (\omega, x^*) \mapsto ||x^* - u(\omega)||_{E^*_b}$ is $\mathcal{F} \otimes \mathcal{B}(E^*_s)$-measurable, and for every fixed $\omega \in \Omega$, $f(\omega, .)$ is lower semicontinuous on E^*_s, shortly, f is a normal integrand, indeed, we have

$$||x^* - u(\omega)||_{E^*_b} = \sup_{j \in \mathbf{N}} \langle e_j, x^* - u(\omega)\rangle$$

here $D_1 = (e_j)_{j \ge 1}$ is a dense sequence in the closed unit ball of E. As each function $(\omega, x^*) \mapsto \langle e_j, x^* - u(\omega)\rangle$ is $\mathcal{F} \otimes \mathcal{B}(E^*_s)$-measurable and continuous on E^*_s for each $\omega \in \Omega$, it follows that f is a normal integrand. Consequently, the graph of u belongs to $\mathcal{F} \otimes \mathcal{B}(E^*_s)$. Besides these facts, let us mention that the function distance $d_{E^*_b}(x^*, y^*) = ||x^* - y^*||_{E^*_b}$ is lower semicontinuous on $E^*_s \times E^*_s$, being the supremum of w^*-continuous functions. If X is a $\mathcal{F}$-measurable mapping, the distance function $\omega \mapsto d_{E^*_b}(x^*, X(\omega))$ is $\mathcal{F}$-measurable, by using the lower semicontinuity of the function $d_{E^*_b}(x^*, .)$ on E^*_s and measurable projection theorem ([5], Theorem III.23) and recalling that E^*_s is a Suslin space. A mapping $u : \Omega \rightrightarrows E^*_s$ is said to be scalarly integrable, if, for every $x \in E$, the scalar function $\omega \mapsto \langle x, u(\omega)\rangle$ is $\mathcal{F}$-measurable and integrable. We denote by $L^1_{E^*}[E](\mathcal{F})$ the subspace of all $\mathcal{F}$-measurable mappings u such that the function

$|u| : \omega \mapsto ||u(\omega)||_{E_b^*}$ is integrable. The measurability of $|u|$ follows easily from the above considerations. By $cwk(E_s^*)$ we denote the set of all nonempty convex $\sigma(E^*, E)$-compact subsets of E_s^*. A $cwk(E_s^*)$-valued mapping $X : \Omega \Rightarrow E_s^*$ is scalarly $\mathcal{F}$-measurable if the function $\omega \to \delta^*(x, X(\omega))$ is $\mathcal{F}$-measurable for every $x \in E$. Let us recall that any scalarly $\mathcal{F}$-measurable $cwk(E_s^*)$-valued mapping is $\mathcal{F}$-measurable. Indeed, let $(e_k)_{k\in\mathbf{N}}$ be a sequence in E which separates the points of E^*, then we have $x \in X(\omega)$ iff, $\langle e_k, x\rangle \le \delta^*(e_k, X(\omega))$ for all $k \in \mathbf{N}$. By $\mathcal{L}^1_{cwk(E_s^*)}(\Omega, \mathcal{F}, P)$ (shortly $\mathcal{L}^1_{cwk(E_s^*)}(\mathcal{F})$) we denote the of all scalarly integrable $cwk(E)$-valued multifunctions X such that the function $|X| : \omega \to |X(\omega)|$ is integrable, here $|X(\omega)| := \sup_{y^*\in X(\omega)} ||y^*||_{E_b^*}$, by the above consideration, it is easy to see that $|X|$ is $\mathcal{F}$-measurable. For the convenience of the reader we recall and summarize the existence and uniqueness of the conditional expectation in $\mathcal{L}^1_{cwk(E_s^*)}(\mathcal{F})$. See ([17], Theorem 3).

Theorem 4.1. *Given a* $\Gamma \in \mathcal{L}^1_{cwk(E_s^*)}(\mathcal{F})$ *and a sub-*σ*-algebra* $\mathcal{B}$ *of* $\mathcal{F}$*, there exists a unique (for equality a.s.) mapping* $\Sigma := E^{\mathcal{B}}\Gamma \in \mathcal{L}^1_{cwk(E_s^*)}(\mathcal{B})$*, that is the conditional expectation of* Γ *with respect to* $\mathcal{B}$*, which enjoys the following properties:*

a) $\int_\Omega \delta^*(v, \Sigma)dP = \int_\Omega \delta^*(v, \Gamma)dP$ *for all* $v \in L^\infty_E(\mathcal{B})$.
b) $\Sigma \subset E^{\mathcal{B}}|\Gamma|\overline{B}_{E^*}$ *a.s.*
c) $\mathcal{S}^1_\Sigma(\mathcal{B})$ *is sequentially* $\sigma(L^1_{E^*}[E](\mathcal{B}), L^\infty_E(\mathcal{B}))$ *compact (here* $\mathcal{S}^1_\Sigma(\mathcal{B})$ *denotes the set of all* $L^1_{E^*}[E](\mathcal{B})$ *selections of* Σ*) and satisfies the inclusion*

$$E^{\mathcal{B}}\mathcal{S}^1_\Gamma(\mathcal{F}) \subset \mathcal{S}^1_\Sigma(\mathcal{B}).$$

d) Furthermore one has

$$\delta^*(v, E^{\mathcal{B}}\mathcal{S}^1_\Gamma(\mathcal{F})) = \delta^*(v, \mathcal{S}^1_\Sigma(\mathcal{B}))$$

for all $v \in L^\infty_E(\mathcal{B})$.
e) $E^{\mathcal{B}}$ *is increasing:* $\Gamma_1 \subset \Gamma_2$ *a.s. implies* $E^{\mathcal{B}}\Gamma_1 \subset E^{\mathcal{B}}\Gamma_2$ *a.s.*

This result involves the existence of conditional expectation for $\sigma(E^*, E)$ closed convex integrable mapping in E^*, namely

Theorem 4.2. *Given a* $\mathcal{F}$*-measurable* $\sigma(E^*, E)$ *closed convex mapping* Γ *in* E^* *which admits a integrable selection* $u_0 \in L^1_{E^*}[E](\mathcal{F})$ *and a sub-*σ*-algebra* $\mathcal{B}$ *of* $\mathcal{F}$*, For every* $n \in \mathbf{N}$ *and for every* $\omega \in \Omega$ *let*

$$\Gamma_n(\omega) = \Gamma(\omega) \cap (u_0(\omega) + n\overline{B}_{E^*}).$$

$$\Sigma(\omega) = \sigma(E^*, E) - cl[\cup E^{\mathcal{B}}\Gamma_n(\omega)].$$

Then $\Sigma(\omega)$ *is a.s. convex and is a* $\mathcal{B}$*-measurable* $\sigma(E^*, E)$ *closed convex mapping which satisfies the properties:*

a) $\int_\Omega \delta^*(v, \Sigma)dP = \int_\Omega \delta^*(v, \Gamma)dP$ *for all* $v \in L_E^\infty(\mathcal{B})$.
b) Σ *is the unique (for = a.s.)*$\mathcal{B}$*-measurable* $\sigma(E^*, E)$ *closed convex mapping with property a).*
c) $\Sigma := E^{\mathcal{B}}\Gamma$ *is the conditional expectation of* Γ.
d) $E^{\mathcal{B}}$ *is increasing:* $\Gamma_1 \subset \Gamma_2$ *a.s. implies* $E^{\mathcal{B}}\Gamma_1 \subset E^{\mathcal{B}}\Gamma_2$ *a.s.*

Proof. Follows the same line of the proof of Theorem VIII-35 in [5] and is omitted. ■

For more information in the conditional expectation of multifunctions, we refer to [1, 14, 17]. In particular recent existence results for conditional expectation in Gelfand and Pettis integration as well as the multivalued Dunford–Pettis representation theorem are available [1]. These results involve several new convergence problems, for instance, the Mosco convergence of sub-super martingales, pramarts in Bochner, Pettis or Gelfand integration, (see [1, 8, 10]). In the context of this paper we will discuss in the next section some conditional expectation results for the normal integrands.

5. On various convergence problems for normal integrands

We present in this section some convergence results for integrand reversed martingales and superadditive sequences. For this purpose we need to develop some tools concerning the conditional expectation for normal integrands. Let S be a topological Suslin space and $\mathcal{B}(S)$ the Borel tribe of S. A mapping $\Psi : \Omega \times S \to \mathbf{R}$ is a $\mathcal{F}$-normal integrand, if it satisfies the two following properties:

a) $\Psi(\omega, .)$ is lower semicontinuous on S for all $\omega \in \Omega$,
b) Ψ is $\mathcal{F} \times \mathcal{B}(S)$-measurable.

Let us recall the following result ([4], Theorem 2.1).

Theorem 5.1. *Let* S *be a topological Suslin space. Let* $\Psi : \Omega \times S \to \mathbf{R}^+$ *be a* $\mathcal{F} \times \mathcal{B}(S)$*-measurable integrand and let* $\mathcal{G}$ *be a sub-*σ*-algebra of* $\mathcal{F}$*. Then there exists a* $\mathcal{G} \times \mathcal{B}(S)$*-measurable integrand,* $E^{\mathcal{G}}\Psi$*, satisfying: for every* $(\mathcal{G}, \mathcal{B}(S))$*-measurable mapping* $u : \Omega \to S$*, the following holds*

$$\int_A E^{\mathcal{G}}\Psi(\omega, u(\omega))dP(\omega) = \int_A \Psi(\omega, u(\omega))dP(\omega)$$

for all $A \in \mathcal{G}$*. The integrand* $E^{\mathcal{G}}\Psi$ *is unique modulo the sets of the form* $N \times S$*, where* N *is a* P*-negligible set in* $\mathcal{G}$*.* $E^{\mathcal{G}}\Psi$ *is the conditional expectation of* Ψ *relative to* $\mathcal{G}$*.*

Now we provide an existence result of conditional expectation for normal integrands on separable Banach spaces.

Theorem 5.2. *Let E be separable Banach space and $\Psi : \Omega \times E \to \mathbf{R}^+$ be a $\mathcal{F}$-normal integrand satisfying*

(i) There exist $a \in L^1_{\mathbf{R}^+}(\Omega, \mathcal{F}, P)$, $b \in \mathbf{R}^+$ and $u_0 \in L^1_E(\Omega, \mathcal{F}, P)$ such that

$$-a(\omega) - b||x - u_0(\omega)||_E \le \Psi(\omega, x) \quad \forall(\omega, x) \in \Omega \times E.$$

(ii) $\Psi(., .u(.))$ is integrable for all $u \in L^1_E(\Omega, \mathcal{G}, P)$.
Then there exists a $\mathcal{G}$-normal integrand $E^{\mathcal{G}}\Psi : \Omega \times E \to \mathbf{R}$ such that

$$\int_A E^{\mathcal{G}}\Psi(\omega, u(\omega))dP(\omega) = \int_A \Psi(\omega, u(\omega))dP(\omega)$$

for all $u \in L^1_E(\Omega, \mathcal{G}, P)$ and for all $A \in \mathcal{G}$. Further, the integrand $E^{\mathcal{G}}\Psi$ is unique modulo the sets of the form $N \times E$, where N is a P-negligible set in $\mathcal{G}$. $E^{\mathcal{G}}\Psi$ is the conditional expectation of Ψ relative to $\mathcal{G}$.

Proof. In order to construct the conditional expectation of the $\mathcal{F}$-normal integrand Ψ we will produce some arguments given in the proof of Theorem 2.3 in [4] and Theorem 7.4 in [15] with appropriate modifications. For notational convenience, we set

$$\theta_0(\omega, x) = a(\omega) + b||x - u_0(\omega))||_E \quad \forall(\omega, x) \in \Omega \times E.$$

For each $k \in \mathbf{N}$, let us set

$$\Psi^k(\omega, x) := \inf_{y \in E}[\Psi(\omega, y) + k||x - y||_E] \quad \forall(\omega, x) \in \Omega \times E.$$

Then for any $\omega \in \Omega$ and for any $x, y \in E$ and for any $k \in \mathbf{N}$ we have the estimates

$$|\Psi^k(\omega, x) - \Psi^k(\omega, y)| \le k||x - y||_E. \tag{5.2.1}$$

$$-\theta_0(\omega, x) \le \Psi^k(\omega, x) \le \Psi^{k+1}(\omega, x) \le \Psi(\omega, x). \tag{5.2.2}$$

$$\sup_{k \in \mathbf{N}} \Psi^k(\omega, x) = \Psi(\omega, x). \tag{5.2.3}$$

From the estimate (5.2.2) we deduce

$$|\Psi^k(\omega, x)| \le |\Psi(\omega, x)| + 2\theta_0(\omega, x). \tag{5.2.4}$$

From (ii) and the estimate (5.2.4) it is immediate that $\Psi^k(.,u(.))$ is integrable for all $u \in L^1_E(\Omega, \mathcal{G}, P)$.

Step 1 Conditional expectation of $E^{\mathcal{G}}\Psi^k$.
In this step $k \in \mathbf{N}$ is fixed. For each $x \in E$, let us denote by $E^{\mathcal{G}}\Psi^k_x(.)$ the conditional expectation of the integrable function $\Psi^k_x(.) := \Psi^k(.,x)$. Then using (5.2.1), it is obvious that

$$(5.2.5) \qquad |E^{\mathcal{G}}\Psi^k_x(\omega) - E^{\mathcal{G}}\Psi^k_y(\omega) \leq k||x-y||_E$$

outside a negligible set $N_{x,y}$. Now let $u \in L^1_E(\Omega, \mathcal{G}, P)$ and D be a countable dense subset in E. Then the following hold:

$$(5.2.6) \quad |E^{\mathcal{G}}\Psi^k_x(\omega) - E^{\mathcal{G}}[\Psi^k(\omega, u(\omega))]| \leq kE^{\mathcal{G}}[||x-u(\omega)||_E] \quad a.s.$$

$$(5.2.7) \qquad E^{\mathcal{G}}\Psi^k_x(.) \in L^1_{\mathbf{R}}(\Omega, \mathcal{G}, P).$$

$$(5.2.8) \quad |E^{\mathcal{G}}\Psi^k_x(\omega) - E^{\mathcal{G}}\Psi^k_y(\omega)| \leq k||x-y||_E \quad \forall (x,y) \in D \times D \quad a.s.$$

Whence there exist a negligible set N in $\mathcal{G}$ such that for all $\omega \in \Omega \setminus N$ the mapping $x \mapsto E^{\mathcal{G}}\Psi^k_x(\omega)$ is k-Lipschitzean on D. Let us set

(5.2.9)
$$E^{\mathcal{G}}\Psi^k(\omega, x) = \inf_{y \in D}[E^{\mathcal{G}}\Psi^k_y(\omega) + k||x-y||_E] \quad \forall (\omega, x) \in \Omega \setminus N \times E.$$

and $E^{\mathcal{G}}\Psi^k(\omega, x) = 0 \quad \forall (\omega, x) \in N \times E$. In others words, for each $\omega \in \Omega \setminus N$, $x \mapsto E^{\mathcal{G}}\Psi^k(\omega, x)$ is the k-Lipschitzean extension of the mapping $x \mapsto E^{\mathcal{G}}\Psi^k_x(\omega)$ defined on D. Furthermore, it is not difficult to check that the above extension formula (5.2.9) yields the $\mathcal{G}$-measurable dependence of $\omega \mapsto E^{\mathcal{G}}\Psi^k(\omega, x)$, shortly $E^{\mathcal{G}}\Psi^k(\omega, x)$ is a $\mathcal{G}$-normal since it is separately $\mathcal{G}$-measurable, and separately k-Lipschitzean. Now we have to show that the $\mathcal{G}$-normal integrand $E^{\mathcal{G}}\Psi^k$ satisfies

$$(5.2.10) \qquad \int_A E^{\mathcal{G}}\Psi^k(\omega, u(\omega))dP(\omega) = \int_A \Psi^k(\omega, u(\omega))dP(\omega) < \infty$$

for all $A \in \mathcal{G}$. It is easy to check that, $E^{\mathcal{G}}\Psi^k(., u(.)) \in L^1_E(\Omega, \mathcal{G}, P)$ when u is a $\mathcal{G}$-measurable step function taking values in D and (5.2.10) holds. In general case, if $u \in L^1_E(\Omega, \mathcal{G}, P)$, there is a sequence $(u_n)_{n\in\mathbf{N}}$ of $\mathcal{G}$-measurable step functions taking values in D with $||u_n(\omega)||_E \leq ||u(\omega)||_E + 1$, $\forall n \in \mathbf{N}$, $\forall \omega \in \Omega$ which pointwisely converges in norm to u. Now since Ψ^k is k-Lipschitzean, we have the estimate

$$\begin{aligned} |\Psi^k(\omega, u_n(\omega))| &\leq |\Psi^k(\omega, u(\omega))| + k||u_n(\omega) - u(\omega)||_E \\ &\leq |\Psi^k(\omega, u(\omega))| + 2k||u(\omega)||_E + 1 \end{aligned}$$

and similarly since $E^{\mathcal{G}}\Psi^k$ is k-Lipschitzean, namely

$$|E^{\mathcal{G}}\Psi^k(\omega,x) - E^{\mathcal{G}}\Psi^k(\omega,y)| \le k||x-y||_E \quad \forall(\omega,x,y)\in\Omega\times E\times E$$

so we have

$$\begin{aligned}|E^{\mathcal{G}}\Psi^k(\omega,u_n(\omega))| &\le |E^{\mathcal{G}}\Psi^k(\omega,u(\omega))| + k||u_n(\omega)-u(\omega)||_E\\ &\le |E^{\mathcal{G}}\Psi^k(\omega,u(\omega))| + 2k||u(\omega)||_E + 1\end{aligned}$$

Note that

$$\begin{aligned}|E^{\mathcal{G}}\Psi^k(\omega,u(\omega))| &\le |E^{\mathcal{G}}\Psi^k(\omega,u_n(\omega))| + k||u(\omega)-u_n(\omega)||_E\\ &\le |E^{\mathcal{G}}\Psi^k(\omega,u_n(\omega))| + 2k||u(\omega)||_E + 1\end{aligned}$$

so that $E^{\mathcal{G}}\Psi^k(\omega,u(\omega))$ is integrable, too. As $\Psi^k(.,u_n(.))$ pointwisely converges to $\Psi^k(.,u(.))$ and $E^{\mathcal{G}}\Psi^k(.,u_n(.))$ pointwisely converges to $E^{\mathcal{G}}\Psi^k(.,u(.))$, from the Lebesgue dominated convergence theorem we get

$$\begin{aligned}\int_A E^{\mathcal{G}}\Psi^k(\omega,u(\omega))dP(\omega) &= \lim_{n\to\infty}\int_A E^{\mathcal{G}}\Psi^k(\omega,u_n(\omega))dP(\omega)\\ &= \lim_{n\to\infty}\int_A \Psi^k(\omega,u_n(\omega))dP(\omega) = \int_A \Psi^k(\omega,u(\omega))dP(\omega).\end{aligned}$$

Step 2 Conditional expectation of $E^{\mathcal{G}}\Psi$ and conclusion.
From the results obtained in Step 1 we can provide the conditional expectation $E^{\mathcal{G}}\Psi$ of the $\mathcal{F}\otimes\mathcal{B}(E)$-measurable integrand Ψ relative to $\mathcal{G}$ satisfying the required property. Taking account of the above estimates and using the measurable projection theorem ([5], Theorem III-23) we provide a null set N' in $\mathcal{G}$ (which does not depend on $x\in E$) such that

$$E^{\mathcal{G}}\Psi^k(\omega,x) \le E^{\mathcal{G}}\Psi^{k+1}(\omega,x) \le E^{\mathcal{G}}\Psi(\omega,x) \quad \forall(\omega,x\in\Omega\setminus N'\times E$$

and

$$|E^{\mathcal{G}}\Psi^k(\omega,x)| \le E^{\mathcal{G}}|\Psi|^k|(\omega,x) \le E^{\mathcal{G}}|\Psi|(\omega,x) + 2E^{\mathcal{G}}\theta_0(\omega,x).$$

In particular we get the estimate

(5.2.11)
$$|E^{\mathcal{G}}\Psi^k(\omega,u(\omega))| \le E^{\mathcal{G}}|\Psi|(\omega,u(\omega)) + 2E^{\mathcal{G}}\theta_0(\omega,u(\omega))\forall\omega\in\Omega\setminus N'$$

here $E^{\mathcal{G}}\theta_0$ denotes the conditional expectation of $\theta_0(.,u(.))$. Using these facts, we may assume that the sequence $(E^{\mathcal{G}}\Psi^k(\omega,x))_{k\in\mathbf{N}}$ is increasing for each $(\omega,x)\in\Omega\times E$, clearly, this argument does not generate uncountable null sets in $\mathcal{G}$ depending on $x\in E$. Now the conditional expectation $E^{\mathcal{G}}\Psi$ can be defined as the supremum of the $\mathcal{G}$-normal integrands $E^{\mathcal{G}}\Psi^k$, so it is a $\mathcal{G}$

normal integrand on $\Omega \times E$ and is the conditional expectation of Ψ relative to $\mathcal{G}$, namely, set $E^{\mathcal{G}}\Psi(\omega, x) := \sup_{k\in\mathbf{N}} E^{\mathcal{G}}\Psi^k(\omega, x) \quad \forall(\omega, x) \in \Omega \setminus N' \times E$ and $E^{\mathcal{G}}\Psi(\omega, x) = 0 \quad \forall(\omega, x) \in N' \times E$. Indeed, by using the estimates (5.2.4)–(5.2.11) and by applying again the Lebesgue dominated convergence theorem we have

$$\begin{aligned}\int_A E^{\mathcal{G}}\Psi(\omega, u(\omega))dP(\omega) &= \uparrow \lim_{k\to\infty} \int_A E^{\mathcal{G}}\Psi^k(\omega, u(\omega))dP(\omega)\\ = \uparrow \lim_{k\to\infty} \int_A \Psi^k(\omega, u(\omega))))]dP(\omega) &= \int_A \Psi(\omega, u(\omega))dP(\omega)\end{aligned}$$

for all $A \in \mathcal{G}$. ■

A Suslin version of the preceding result is available ([15], Theorem 5.6). We only present a scalar version of this theorem.

Theorem 5.3. *Let (S, d) be a Suslin metric space with $d \le 1$, $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. Let $\Psi : \Omega \times S \to \mathbf{R}$ be a $\mathcal{F}$-normal integrand satisfying:*

(i) There is $a \in L^1_{\mathbf{R}^+}(\Omega, \mathcal{F}, P)$ and $b \in \mathbf{R}^+$ and a $\mathcal{F}$-measurable mapping $u_0 : \Omega \to S$ such that

$$-a(\omega) - d(x, u_0(\omega))b \le \Psi(\omega, x) \quad \forall(\omega, x) \in \Omega \times S.$$

(ii) For any $u \in L^0_S(\Omega, \mathcal{G}, P)$, $\Psi(., u(.)) \in L^1_{\mathbf{R}}(\Omega, \mathcal{F}, P)$.
Then there exists a $\mathcal{G}$-normal integrand $E^{\mathcal{G}}\Psi : \Omega \times S \to \mathbf{R}$ such that

$$\int_A E^{\mathcal{G}}\Psi(\omega, u(\omega))dP(\omega) = \int_A \Psi(\omega, u(\omega))dP(\omega)$$

for all $u \in L^0_S(\Omega, \mathcal{G}, P)$ and for all $A \in \mathcal{G}$. Further, the integrand $E^{\mathcal{G}}\Psi$ is unique modulo the sets of the form $N \times S$, where N is a P-negligible set in $\mathcal{G}$. $E^{\mathcal{G}}\Psi$ is the conditional expectation of Ψ relative to $\mathcal{G}$.

We provide an application of this result yielding an existence theorem of conditional expectation for normal integrands defined on a dual space.

Corollary 5.1. *Let E be a separable Banach space, $\overline{B}_{E^*}$ endowed with the weak star topology, $\mathcal{G}$ a sub-σ-algebra of $\mathcal{F}$ and $\Psi : \Omega \times \overline{B}_{E^*} \to \mathbf{R}$ a $\mathcal{F} \times \mathcal{B}(\overline{B}_{E^*})$ normal integrand satisfying:*

(i) There is $a \in L^1_{\mathbf{R}^+}(\Omega, \mathcal{G}, P)$ such that

$$-a(\omega) \le \Psi(\omega, x) \quad \forall(\omega, x) \in \Omega \times \overline{B}_{E^*}.$$

(ii) $\Psi(., u(.))$ *is integrable for all* $u \in L^\infty_{\overline{B}_{E^*}}(\Omega, \mathcal{G}, P)$.

Then there exists a $\mathcal{G}$*-normal integrand* $E^{\mathcal{G}}\Psi : \Omega \times \overline{B}_{E^*} \to \mathbf{R}$ *such that*

$$\int_A E^{\mathcal{G}}\Psi(\omega, u(\omega))dP(\omega) = \int_A \Psi(\omega, u(\omega))dP(\omega)$$

for all $u \in L^\infty_{\overline{B}_{E^*}}(\Omega, \mathcal{G}, P)$ *and for all* $A \in \mathcal{G}$. *Further, the integrand* $E^{\mathcal{G}}\Psi$ *is unique modulo the sets of the form* $N \times \overline{B}_{E^*}$, *where* N *is a* P*-negligible set in* $\mathcal{G}$. $E^{\mathcal{G}}\Psi$ *is the conditional expectation of* Ψ *relative to* $\mathcal{G}$.

Proof. Recall that $(E^*_{m^*}, d_{E^*_{m^*}})$ is a Suslin metric space and the following properties hold:

$$d_{E^*_{m^*}}(x^*, y^*) \le ||x^* - y^*||_{E^*_b}, \quad \forall (x^*, y^*) \in E^* \times E^*.$$

$$\mathcal{B}(E^*_c) = \mathcal{B}(E^*_s) = \mathcal{B}(E^*_m).$$

In addition, the topologies m^*, τ_c, w^* coincide on weak star compact sets of E^*. Now the proof follows by applying Theorem 5.3 with the Suslin metric space (S, d) replaced by the Suslin metric space $(\overline{B}_{E^*}, d_{E^*_{m^*}})$. Alternatively, one may apply the techniques of Theorem 5.2 with $(E, ||.||)$ replaced by $(\overline{B}_{E^*}, d_{E^*_{m^*}})$. ∎

As an example, let us consider $\Gamma \in \mathcal{L}^1_{cwk(E)}(\mathcal{F})$. Let us set

$$\Psi(\omega, x^*) := \delta^*(x^*, \Gamma(\omega)) \quad \forall (\omega, x^*) \in \Omega \times \overline{B}_{E^*}.$$

Let us mention another useful variant dealing with normal *convex* integrands. See ([5], Chap. 8) and [2] for related variants.

Theorem 5.4. *Let* E *be a separable Banach space,* $\mathcal{G}$ *be a sub-*σ*-algebra of* $\mathcal{F}$ *and* $f : \Omega \times E^*_s \to \overline{\mathbf{R}}$ *be a* $\mathcal{F} \times \mathcal{B}(E^*_s)$*-measurable normal convex integrand, such that* $f^+(\omega, u(\omega))$ *is integrable for some* $u \in L^1_{E^*}[E](\Omega, \mathcal{F}, P)$ *and let* $f^* : \Omega \times E \to \overline{\mathbf{R}}$ *be the polar of* f. *Then there is a* $\mathcal{G} \times \mathcal{B}(E^*_s)$*-measurable normal convex integrand* $g : \Omega \times E^*_s \to \overline{\mathbf{R}}$ *such that, for all* $v \in L^\infty_E(\Omega, \mathcal{G}, P)$, *the function* $f^*(., v(.))$ *and* $g^*(., v(.))$ *are quasi-integrable, and*

$$\int_A f^*(\omega, v(\omega))dP(\omega) = \int_A g^*(\omega, v(\omega))dP(\omega)$$

$\forall A \in \mathcal{G}$. *Moreover* g *is unique a.s. and the mapping* $f \to g$ *is increasing;* $g^* := E^{\mathcal{G}} f^*$ *can be called the conditional expectation of* f^* *relative to* $\mathcal{G}$.

Proof. Theorem 5.4 is a slight improvement of Theorem VIII.36 in [5] involving Theorem 4.2. So we only summarize some useful facts of the proof. Let us set

$$\Gamma(\omega) := epi f(\omega, .) \quad \forall \omega \in \Omega.$$

Then Γ is a closed convex integrable mapping in $E_s^* \times \mathbf{R}$ and admitting an integrable selection: $\omega \to (u_0(\omega), f^+(\omega, u_0(\omega)))$. Now apply Theorem 4.2 provides the conditional expectation

$$\Sigma(\omega) := E^{\mathcal{B}}\Gamma(\omega) \quad \forall \omega \in \Omega.$$

Using the arguments given in Theorem VIII. 36 in [5] we check that $\Sigma(\omega)$ is a.s. an epigraph. Put $r(\omega) := f^+(\omega, u_0(\omega))$. Then for every $n \in \mathbf{N}$, $(u_0(\omega), r(\omega) + n) \in \Gamma(\omega)$, and

$$E^{\mathcal{B}}(u_0, r + n) = E^{\mathcal{B}}(u_0, r) + (0, n) \in \mathcal{S}^1_\Sigma.$$

If N is a negligible set such that for all $\omega \in \Omega \setminus N$, for all $n \in \mathbf{N}$

$$E^{\mathcal{B}}(u_0, r)(\omega) + (0, n) \in \Sigma(\omega)$$

we have for all $\omega \in \Omega \setminus N$

$$E^{\mathcal{B}}(u_0, r)(\omega) + \{0\} \times [0, \infty[\subset \Sigma(\omega)$$

and so $\Sigma(\omega)$ is an epigraph. Let g be the normal integrand associated with Σ

$$g(\omega, x^*) = \begin{cases} \inf\{\lambda \in \mathbf{R} : (x^*, \lambda) \in \Sigma(\omega)\} & \text{if} \quad \omega \in \Omega \setminus N \\ 0 \quad \text{if} \quad \omega \in N \end{cases}$$

As

$$f^*(\omega, v(\omega)) := \delta^*((v(\omega), -1), \Gamma(\omega))$$

and

$$g^*(\omega, v(\omega)) := \delta^*((v(\omega), -1), \Sigma(\omega))$$

these functions are quasi integrable and have the required properties thanks to Theorem 4.2. The uniqueness and the increasing property can be proved as in Theorem VIII.35 in [5]. ■

Remarks. The integrand g is the conditional infimum convolution of the functions $f(\omega, .)$. It has been studied in [2], and in Chap. VIII of [5].

Now we can define the notion of integrand reversed martingales according to the Definition 2.4 in [4] and the existence Theorems 5.2, 5.3, and 5.4 of conditional expectations. This concept leads us to new variational convergence problems with further applications.

Definition 5.1. *Let E be separable Banach space , $(\mathcal{B}_n)_{n\in\mathbf{N}}$ be an decreasing sequence of sub-σ-algebras of $\mathcal{F}$ with $\mathcal{B}_\infty = \cap_{n=1}^\infty \mathcal{B}_n$ and $\Psi_n : \Omega \times E \to \mathbf{R}^+$ $(n \in \mathbf{N})$ be a $\mathcal{B}_n$-normal integrand such that $\Psi_n(., u(.))$ is integrable for any $u \in L^1_E(\Omega, \mathcal{B}_n, P)$.*

The sequence $(\Psi_n, \mathcal{B}_n)_{n\in\mathbf{N}}$ of $\mathcal{B}_n$-normal integrands is a lower semicontinuous integrand reversed martingale (resp. supermartingale) if

$$\Psi_{n+1}(\omega, x) = E^{\mathcal{B}_{n+1}}\Psi_n(\omega, x) \quad \forall(\omega, x) \in \Omega \times E \quad \forall n \in \mathbf{N}.$$

respectively

$$\Psi_{n+1}(\omega, x) \geq E^{\mathcal{B}_{n+1}}\Psi_n(\omega, x) \quad \forall(\omega, x) \in \Omega \times E \quad \forall n \in \mathbf{N}.$$

Using the ideas developed in [4] we present a variational convergence result for the *reversed integrand martingales* associated with a decreasing sequence of sub-σ-algebras $(\mathcal{B}_n)_{n\in\mathbf{N}}$ of $\mathcal{F}$ with $\mathcal{B}_\infty = \cap_{n=1}^\infty \mathcal{B}_n$. Let S be a Polish space. We denote by $\mathcal{M}^1_+(S)$ the set of all probability Borel measures on S is endowed with the narrow topology, so that $\mathcal{M}^1_+(S)$ is a Polish space, $\mathcal{Y}(\Omega, \mathcal{M}^1_+(S))$ the space of all $(\mathcal{F}, \mathcal{B}(\mathcal{M}^1_+(S)))$-measurable mappings (alias Young measures) $\lambda : \Omega \to \mathcal{M}^1_+(S)$ (see e.g. [6] for Young measures on topological spaces). A sequence $(\lambda^n)_{n\in\mathbf{N}}$ in $\mathcal{Y}(\Omega, \mathcal{M}^1_+(S))$ *stably* converges to a Young measure λ if

$$\lim_{n\to\infty} \int_\Omega [\int_S h(\omega, s)\lambda^n_\omega(ds)]P(d\omega) = \int_\Omega [\int_S h(\omega, s)\lambda_\omega(ds)]P(d\omega)$$

for all bounded Caratheodory integrand $h : \Omega \times S \to \mathbf{R}$. A sequence of lower semicontinuous function $(\varphi_n)_{n\in\mathbf{N}}$ defined on S *epilower* converges to a lower semicontinuous function φ defined on S, if, for any sequence $(x_n)_{n\in\mathbf{N}}$ in S converging to $x \in S$, we have

$$\liminf_{n\to\infty} \varphi_n(x_n) \geq \varphi(x)$$

Let us recall the following lemma.

Lemma 5.1. *Let S be a Polish space, and let $\varphi_n, \varphi_\infty$, $(n \in \mathbf{N})$ be a non-negative sequence of normal integrands defined on $\Omega \times S$ such that for each $\omega \in \Omega$, $\varphi_n(\omega, .)$ epilower converges to $\varphi_\infty(\omega, .)$. Let $\lambda^n, \lambda^\infty$ $(n \in \mathbf{N})$ be a sequence of Young measures in $\mathcal{Y}(\Omega, \mathcal{M}^1_+(S))$ which stably converges to λ^∞. Then we have*

$$\liminf_{n\to\infty} \int_\Omega [\int_S \varphi_n(\omega, s)\lambda^n_\omega(ds)]dP(\omega) \geq \int_\Omega [\int_S \varphi_\infty(\omega, s)\lambda^\infty_\omega(ds)]dP(\omega).$$

Proof. See ([7], Lemma 3.4). ■

Now we provide an epiconvergence results for integrand reversed martingales.

Theorem 5.5. *Let E be a separable Banach space, $(\mathcal{B}_n)_{n\in\mathbf{N}}$ be a decreasing sequence of sub-σ-algebras of $\mathcal{F}$ with $\mathcal{B}_\infty = \cap_{n=1}^\infty \mathcal{B}_n$, $\Psi : \Omega \times E \to \mathbf{R}^+$ a $\mathcal{F}$-normal integrand such that $\Psi(., .u(.))$ is integrable for all $u \in L^1_E(\Omega, \mathcal{B}_n, P)$ and for all $n \in \mathbf{N} \cup \{\infty\}$. Let $E^{\mathcal{B}_n}\Psi$ $(n \in \mathbf{N} \cup \{\infty\})$ be the conditional expectation of Ψ relative to $\mathcal{B}_n$ whose existence is given by Theorem 7.2. Then the following properties hold:*

(a) For any $u \in L^1_E(\Omega, \mathcal{B}_\infty, P)$, $(E^{\mathcal{B}_n}\Psi(., u(.)))_{n\in\mathbf{N}}$ is a reversed integrable martingale which converges a.s. to $E^{\mathcal{B}_\infty}\Psi_u$.

(b) If $(u^n)_{n\in\mathbf{N}}$ is a sequence in $L^1_E(\Omega, \mathcal{B}_\infty, P)$ which stably converges to a Young measure $\lambda \in \mathcal{Y}(\Omega, \mathcal{M}^1_+(E))$, then

$$\liminf_{n\to\infty} \int_\Omega E^{\mathcal{B}_n}\Psi(\omega, u^n(\omega))dP(\omega) \geq \int_\Omega [\int_E E^{\mathcal{B}_\infty}\Psi(\omega, s)\lambda_\omega(ds)]dP(\omega).$$

(c) The sequence $(E^{\mathcal{B}_n}\Psi)_{n\in\mathbf{N}}$ forms a lower semicontinuous integrand reversed martingale and for any $u \in L^1_E(\Omega, \mathcal{B}_\infty, P)$, the following epiconvergence result holds:

$$\sup_{k\in\mathbf{N}} \limsup_{n\to\infty} \inf_{y\in E} [E^{\mathcal{B}_n}\Psi(\omega, y) + k||u(\omega) - y||_E] \leq E^{\mathcal{B}_\infty}\Psi(\omega, u(\omega)) \quad a.s.$$

Proof. (a) Let $n \in \mathbf{N} \cup \{\infty\}$. According to our assumption and Theorem 5.2, the conditional expectation $E^{\mathcal{B}_n}\Psi$ of Ψ relative to $\mathcal{B}_n$ is a $\mathcal{B}_n$-normal integrand on $\Omega \times E$ satisfying

$$\int_A E^{\mathcal{B}_n}\Psi(\omega, u(\omega))dP(\omega) = \int_A \Psi(\omega, u(\omega))dP(\omega) < \infty$$

for all $u \in L^1_E(\Omega, \mathcal{B}_n, P)$ and for all $A \in \mathcal{B}_n$. In particular, for each $u \in L^1_E(\Omega, \mathcal{B}_\infty, P)$, we have $\Psi_u := \Psi(., u(.)) \in L^1_{\mathbf{R}}(\Omega, \mathcal{F}, P)$. Hence $(E^{\mathcal{B}_n}\Psi_u)_{n\in\mathbf{N}}$ is a reversed integrable martingale which converges a.s. to $E^{\mathcal{B}_\infty}\Psi_u$ by virtue of Corollary V-III-12 in [16].

(b) We show first that for a.s. $\omega \in \Omega$, $E^{\mathcal{B}_n}\Psi(\omega, .)$ epilower converges to $E^{\mathcal{B}_\infty}\Psi(\omega, .)$. Let $(x_n)_{n\in\mathbf{N}}$ in E converging to $x \in E$, then by the lower semicontinuity of $\Psi(\omega, .)$, for each $\omega \in \Omega$ we have

$$\liminf_{n\to\infty} \Psi(\omega, x_n) \geq \Psi(\omega, x).$$

Now using the Doob a.s. convergence for positive reversed martingales (see Corollary V-III-12 in [16]) and Lemma 3.3 in [4], we have for a.s. $\omega \in \Omega$

$$\liminf_{n\to\infty} E^{\mathcal{B}_n}[\Psi(., x_n)](\omega) \geq E^{\mathcal{B}_\infty}[\liminf_{n\to\infty} \Psi(., x_n)](\omega) \geq E^{\mathcal{B}_\infty}[\Psi(., x)](\omega)$$

showing the required convergence. Now (b) follows by applying this result and Lemma 5.1 to the normal integrands $E^{\mathcal{B}_n}\Psi$ and $E^{\mathcal{B}_\infty}\Psi$.

(c) Here we use some arguments developed in the proof of Theorem 3.5 in [4] and the notations and results obtained in the proof of Theorem 5.2. By virtue of Theorem 5.2, the conditional expectation $E^{\mathcal{B}_\infty}\Psi$ is a $\mathcal{B}_\infty \otimes \mathcal{B}(E)$-measurable normal integrand satisfying

$$\int_A E^{\mathcal{B}_\infty}\Psi(\omega, u(\omega))dP(\omega) = \int_A \Psi(\omega, u(\omega))dP(\omega) < \infty$$

for all $u \in L^1_E(\Omega, \mathcal{B}_\infty, P)$ and for all $A \in \mathcal{B}_\infty$. Let us set

$$\Phi(\omega, x) := E^{\mathcal{B}_\infty}\Psi(\omega, x) \quad \forall(\omega, x) \in \Omega \times E.$$

$$\Phi^k(\omega, x) = \inf_{y\in E}[\Phi(\omega, y) + k||x - y||_E] \quad \forall(\omega, x) \in \Omega \times E.$$

Then by the normality of $\Phi := E^{\mathcal{B}_\infty}\Psi$

$$0 \leq \Phi^k(\omega, x) \leq \Phi^{k+1}(\omega, x) \leq \Phi(\omega, x) \quad \forall k \in \mathbf{N} \quad \forall(\omega, x) \in \Omega \times E.$$

$$\sup_{k\in\mathbf{N}} \Phi^k(\omega, x) = \Phi(\omega, x) \quad \forall(\omega, x) \in \Omega \times E.$$

Let $u \in L^1_E(\Omega, \mathcal{B}_\infty, P)$, From the definition of Φ we have for each k

$$0 \leq \Phi^k(\omega, u(\omega)) \leq \Phi(\omega, u(\omega)) = E^{\mathcal{B}_\infty}\Psi(\omega, u(\omega))$$

so that $\Phi^k(., u(.)) \in L^1_{\mathbf{R}}(\Omega, \mathcal{F}, P)$. Let $p \in \mathbf{N}$. Since E is separable, applying measurable selection theorem ([5], Theorem III-22), it is not difficult to provide a $(\mathcal{B}_\infty, \mathcal{B}(E))$-measurable mapping $v_{k,p,u} : \Omega \to E$ such that

$$0 \leq \Phi(\omega, v_{k,p,u}(\omega)) + k||u(\omega) - v_{k,p,u}(\omega)||_E \leq \Phi^k(\omega, u(\omega)) + \frac{1}{p}$$

for all $\omega \in \Omega$, so that $\omega \mapsto \Phi(\omega, v_{k,p,u}(\omega))$ and $\omega \mapsto k||u(\omega) - v_{k,p,u}(\omega)||_E$ are integrable, and so $v_{k,p,u} \in L^1_E(\Omega, \mathcal{B}_\infty, P)$. By our assumption, $\omega \mapsto \Psi(\omega, v_{k,p,u}(\omega))$ is integrable, too. Applying ([16], Corollaire V-3-12) yields a negligible set $N_{k,p,u}$ such that for all $\omega \notin N_{k,p,u}$

$$\lim_{n\to\infty} E^{\mathcal{B}_n}[\Psi(\omega, v_{k,p,u}(\omega))] = E^{\mathcal{B}_\infty}[\Psi(\omega, v_{k,p,u}(\omega))] = \Phi(\omega, v_{k,p,u}(\omega)).$$

When we deduce

$$\begin{aligned}
&\limsup_{n\to\infty} \inf_{y\in E} [E^{\mathcal{B}_n}\Psi(\omega, y) + k||u(\omega) - y||_E] \\
&\quad \le \limsup_{n\to\infty} [E^{\mathcal{B}_n}\Psi(\omega, v_{k,p,u}(\omega)) + k||u(\omega) - v_{k,p,u}(\omega)||_E] \\
&\quad = \Phi(\omega, v_{k,p,u}(\omega)) + k||u(\omega) - v_{k,p,u}(\omega)||_E \le \Phi^k(\omega, u(\omega)) \\
&\qquad + \frac{1}{p} \quad \forall\omega \notin N_{k,p,u}.
\end{aligned}$$

Set $N_u := \cup_{k\in\mathbf{N},\, p\in\mathbf{N}} N_{k,p,u}$. Then N_u is negligible. Taking the supremum on $k \in \mathbf{N}$ in the extreme terms yields

$$\begin{aligned}
&\sup_{k\in\mathbf{N}} \limsup_{n\to\infty} \inf_{y\in E} [E^{\mathcal{B}_n}\Psi(\omega, y) + k||u(\omega) - y||_E] \\
&\quad \le \sup_{k\in\mathbf{N}} \Phi^k(\omega, u(\omega)) = \Phi(\omega, u(\omega)) = E^{\mathcal{B}_\infty}\Psi(\omega, u(\omega)) \qquad \forall\omega \notin N_u.
\end{aligned}$$

We finish the paper by providing some epiconvergence results related to the Birkhoff–Kingman ergodic theorem. We refer to ([11], Theorem 2.3), ([18], Theorem 6) for sharp variants dealing with general Suslin metric spaces. Here we also improve some related results in [4].

Theorem 5.6. *Let E be a separable Banach space, T a measurable transformation of Ω preserving P, $\mathcal{I}$ the σ algebra of invariants sets. Let $\Psi : \Omega \times E \to \mathbf{R}^+$ be a $\mathcal{F}$-normal integrand such that, for any $u \in L^1_E(\Omega, \mathcal{I}, P)$, $\Psi(., u(.)) \in L^1_{\mathbf{R}}(\Omega, \mathcal{F}, P)$. Then for any $u \in L^1_E(\Omega, \mathcal{I}, P)$ the following epiconvergence result holds:*

$$\begin{aligned}
&\sup_{k\in\mathbf{N}} \liminf_{n\to\infty} \inf_{y\in E} [\frac{1}{n}\sum_{j=0}^{n-1} \Psi(T^j\omega, y) + k||u(\omega) - y||_E] \\
&\quad \ge E^{\mathcal{I}}\Psi(\omega, u(\omega)) \quad a.s.
\end{aligned}$$

Proof. Here we use some arguments of Theorem 5.3 in [4] and the notations and results obtained in the proof of Theorem 5.2. For $k \in \mathbf{N}$, $n \in \mathbf{N}$ and for $(\omega, x) \in \Omega \times E$, we set

$$\Psi^k(\omega, x) = \inf_{y\in E} [\Psi(\omega, y) + k||x - y||_E].$$

$$\Psi_n(\omega, x) = \frac{1}{n}\sum_{j=0}^{n-1} \Psi(T^j\omega, x).$$

$$\Psi^k_n(\omega, x) = \inf_{y\in E} [\Psi_n(\omega, y) + k||x - y||_E].$$

Then the following hold

(5.6.1) $|\Psi^k(\omega, x) - \Psi^k(\omega, y)| \leq k||x - y||_E \quad \forall(\omega, x, y) \in \Omega \times E \times E.$

(5.6.2) $$0 \leq \Psi^k(\omega, x) \leq \Psi(\omega, x) \quad \forall(\omega, x) \in \Omega \times E.$$

(5.6.3) $$\sup_{k\in\mathbf{N}} \Psi^k(\omega, x) = \Psi(\omega, x) \quad \forall(\omega, x) \in \Omega \times E.$$

There is a negligible set N which does not depend on $x \in E$ such that

(5.6.4) $$E^{\mathcal{I}}\Psi(\omega, x) = \begin{cases} \sup_{k\in\mathbf{N}} E^{\mathcal{I}}\Psi_k(\omega, x) & \text{if} \quad \omega \in \Omega \setminus N \times E \\ 0 \quad \text{if} \quad (\omega, x) \in N \times E \end{cases}$$

where $E^{\mathcal{I}}\Psi^k$ and $E^{\mathcal{I}}\Psi$ denote the conditional expectation relative to $\mathcal{I}$ of Ψ^k and Ψ respectively.

(5.6.5) $$\Psi_n^k(\omega, x) \geq \frac{1}{n}\sum_{j=0}^{n-1} \Psi^k(T^j\omega, x) \quad \forall(\omega, x) \in \Omega \times E.$$

By virtue of classical Birkhoff ergodic theorem (see e.g. Lemma 5 in [18]) it follows that

(5.6.6) $$\lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} \Psi^k(T^j\omega, u(\omega)) = E^{\mathcal{I}}[\Psi^k(\omega, u(\omega))] \quad a.s.$$

Using (5.6.5) and (5.6.6) yield a negligible set $N_{k,u}$ such that

$$\liminf_{n\to\infty} \inf_{y\in E} [\frac{1}{n}\sum_{j=0}^{n-1} \Psi(T^j\omega, y) + k||u(\omega) - y||_E)]$$
$$\geq E^{\mathcal{I}}[\Psi^k(\omega, u(\omega))] \quad \forall\omega \notin N_{k,u}$$

Using (5.6.4) and the preceding limits, we produce a negligible set $N_u = \cup_{k\in\mathbf{N}} N_{k,u} \cup N$ such that

$$\sup_{k\in\mathbf{N}} \liminf_{n\to\infty} \inf_{y\in E} [\frac{1}{n}\sum_{j=0}^{n-1} \Psi(T^j\omega, y) + k||u(\omega) - y||_E] \geq E^{\mathcal{I}}\Psi(\omega, u(\omega)) \quad \forall\omega \notin N_u.$$

■

We finish the paper by providing some applications to the epiconvergence of superadditive normal integrands (alias lsc superadditive random sequences [4]).

Definition 5.2. *Let S be a topological space and let T be a measure preserving transformation of Ω into itself. A sequence $(\Psi_n)_{n\in\mathbf{N}}$ of $\mathcal{F}$-normal integrands defined on $\Omega \times S$ is superadditive if for all $m, n \in \mathbf{N}$, and for all $(\omega, x) \in \Omega \times S$*

$$\Psi_n(T^m\omega, x) \le \Psi_{n+m}(\omega, x) - \Psi_m(\omega, x).$$

Corollary 5.2. *Let E be a separable Banach space, T a measurable transformation of Ω preserving P, $\mathcal{I}$ the σ algebra of invariants sets. Let $(\Psi_n)_{n\in\mathbf{N}}$ be a nonnegative sequence of superadditive $\mathcal{F}$-normal integrands defined on $\Omega \times E$ such that for all $n \in \mathbf{N}$, for all $u \in L^1_E(\Omega, \mathcal{I}, P)$, $\Psi_n(., u(.)) \in L^1_{\mathbf{R}}(\Omega, \mathcal{F}, P)$. Then for all $u \in L^1_E(\Omega, \mathcal{I}, P)$, the following epiconvergence result holds:*

$$(^*) \quad \sup_{k\in\mathbf{N}} \liminf_{n\to\infty} \inf_{y\in E} [\frac{1}{k}\Psi_{n+k-1}(\omega, y) + k||u(\omega) - y||_E] \ge \sup_{k\in\mathbf{N}} \frac{1}{k} E^{\mathcal{I}}\Psi_k(\omega, u(\omega)).$$

Proof. By virtue of Theorem 5.2, the conditional expectation $E^{\mathcal{I}}\Psi_k$ $(k \in \mathbf{N})$ is a $\mathcal{I} \otimes \mathcal{B}(E)$-measurable normal integrand satisfying

$$\int_A E^{\mathcal{I}}\Psi_k(\omega, u(\omega))dP(\omega) = \int_A \Psi_k(\omega, u(\omega))dP(\omega) < \infty$$

for all $u \in L^1_E(\Omega, \mathcal{I}, P)$ and for all $A \in \mathcal{I}$. By superadditivity we have

$$\Psi_k(T^j\omega, y) \le \Psi_{k+j}(\omega, y) - \Psi_j(\omega, y)$$

for all $j, k \in \mathbf{N}$ and for all $(\omega, y) \in \Omega \times E$. This gives

$$\sum_{j=0}^{n-1} \Psi_k(T^j\omega, y) \le \sum_{j=0}^{n-1} [\Psi_{k+j}(\omega, y) - \Psi_j(\omega, y)] \le n\Psi_{n+k-1}(\omega, y).$$

Whence

$$\frac{1}{k}\Psi_{n+k-1}(\omega, y) \ge \frac{1}{k}\frac{1}{n}\sum_{j=0}^{n-1} \Psi_k(T^j\omega, y).$$

It follows that

$$\inf_{y\in E} [\frac{1}{k}\Psi_{n+k-1}(\omega, y) + k||u(\omega) - y||_E] \ge \frac{1}{k} \inf_{y\in E} [\frac{1}{n}\sum_{j=0}^{n-1} \Psi_k(T^j\omega, y) + k||u(\omega) - y||_E].$$

Using this estimate and (i)–(ii) and applying Theorem 5.6 with Ψ replaced by Ψ_k yields a negligible set $N_{k,u}$ such that for $\omega \notin N_{k,u}$

$$\liminf_{n\to\infty} \inf_{y\in E} [\frac{1}{k}\Psi_{n+k-1}(\omega, y) + k||u(\omega) - y||_E]$$
$$\geq \frac{1}{k} \liminf_{n\to\infty} \inf_{y\in E} [\frac{1}{n}\sum_{j=0}^{n-1} \Psi_k(T^j\omega, y) + k||u(\omega) - y||_E]$$
$$\geq \frac{1}{k} E^{\mathcal{I}}\Psi_k(\omega, u(\omega)).$$

By taking the supremum over k in extreme terms we get a negligible set $N_u := \cup_{k\in\mathbf{N}} N_{k,u}$ such that for $\omega \notin N_u$

$$\sup_{k\in\mathbf{N}} \liminf_{n\to\infty} \inf_{y\in E} [\frac{1}{k}\Psi_{n+k-1}(\omega, y) + k||u(\omega) - y||_E)]$$
$$\geq \sup_{k\in\mathbf{N}} \frac{1}{k} E^{\mathcal{I}}\Psi_k(\omega, u(\omega))$$

proving the epiliminf inequality (*). ■

We finish the paper with some epilimsup results.

Corollary 5.3. *Let E be a separable Banach space, T a measurable transformation of Ω preserving P, $\mathcal{I}$ the σ algebra of invariants sets. Let $\Psi : \Omega \times E \to \mathbf{R}^+$ be a $\mathcal{F}$-normal integrand such that $\Psi(., .u(.))$ is integrable for all $u \in L^1_E(\Omega, \mathcal{I}, P)$. Let $(f_n)_{n\in\mathbf{N}}$ be a positive superadditive uniformly bounded sequence in $L^\infty_{\mathbf{R}}(\Omega, \mathcal{F}, P)$. Then for any $u \in L^1_E(\Omega, \mathcal{I}, P)$, the following epiconvergence result holds:*

$$(^{**}) \quad \sup_{k\in\mathbf{N}} \limsup_{n\to\infty} \inf_{y\in E} [\frac{1}{n} f_n(\omega).\frac{1}{n}\sum_{j=0}^{n-1} \Psi(T^j\omega, y) + k||u(\omega) - y||_E]$$
$$\leq \sup_{n\in\mathbf{N}} \frac{1}{n} E^{\mathcal{I}} f_n(\omega).\ E^{\mathcal{I}}\Psi(\omega, u(\omega)) \quad a.s.$$

Proof. Here we use some arguments developed in the proof of Proposition 5.5 in [4]. By virtue of Theorem 5.2, the conditional expectation $E^{\mathcal{I}}\Psi$ is a $\mathcal{I} \otimes \mathcal{B}(E)$-measurable normal integrand satisfying

$$\int_A E^{\mathcal{I}}\Psi(\omega, u(\omega))dP(\omega) = \int_A \Psi(\omega, u(\omega))dP(\omega) < \infty$$

for all $u \in L^1_E(\Omega, \mathcal{I}, P)$ and for all $A \in \mathcal{I}$. Let us set

$$\gamma(\omega := \sup_{n\in\mathbf{N}} \frac{1}{n} E^{\mathcal{I}} f_n(\omega) \quad \forall \omega \in \Omega.$$

$$\Delta(\omega, x) := \gamma(\omega)\, E^{\mathcal{I}} \Psi(\omega, x) \quad \forall (\omega, x) \in \Omega \times E.$$

$$\Delta^k(\omega, x) = \inf_{y\in E} [\Delta(\omega, y) + k||x - y||_E] \quad \forall (\omega, x) \in \Omega \times E.$$

Then

$$\sup_{k\in\mathbf{N}} \Delta^k(\omega, x) = \Delta(\omega, x) \quad \forall (\omega, x) \in \Omega \times E.$$

By Kingman theorem for superadditive integrable sequences ([12], Theorem 10.7.1) we may assume that

$$\lim_{n\to\infty} \frac{1}{n} f_n(\omega) = \gamma(\omega) \quad \forall \omega \in \Omega.$$

Let $u \in L^1_E(\Omega, \mathcal{I}, P)$. Let $p \in \mathbf{N}$. Since E is separable, applying measurable selection theorem ([5], Theorem III-22), it is not difficult to provide a $(\mathcal{I}, \mathcal{B}(E))$-measurable mapping $v_{k,p,u} : \Omega \to E$ such that

$$0 \le \gamma(\omega) E^{\mathcal{I}} \Psi(\omega, v_{k,p,u}(\omega)) + k||u(\omega) - v_{k,p,u}(\omega)||_E \le \Delta^k(\omega, u(\omega)) + \frac{1}{p}$$

for all $\omega \in \Omega$, so that $\omega \mapsto \gamma(\omega) E^{\mathcal{I}} \Psi(\omega, v_{k,p,u}(\omega))$ and $\omega \mapsto k||u(\omega) - v_{k,p,u}(\omega)||_E$ are integrable, and so $v_{k,p,u} \in L^1_E(\Omega, \mathcal{I}, P)$. By our assumption, $\omega \mapsto \Psi(\omega, v_{k,p,u}(\omega))$ is integrable, too, so that

$$\int_A E^{\mathcal{I}} \Psi(\omega, v_{k,p,u}(\omega)) dP(\omega) = \int_A \Psi(\omega, v_{k,p,u}(\omega)) dP(\omega) < \infty$$

for all $A \in \mathcal{I}$. Hence $E^{\mathcal{I}} \Psi(\omega, v_{k,p,u}(\omega)) = E^{\mathcal{I}}[\Psi(\omega, v_{k,p,u}(\omega))]$ a.s. From the classical Birkhoff ergodic theorem (see e.g. [18]) we provide a negligible set $N_{k,p,u}$ such that for $\omega \notin N_{k,p,u}$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \Psi(T^j \omega, v_{k,p,u}(\omega)) = E^{\mathcal{I}}[\Psi(\omega, v_{k,p,u}(\omega))] = E^{\mathcal{I}} \Psi(\omega, v_{k,p,u}(\omega)).$$

Whence we provide a negligible set $N_{k,p,u}$ such that

$$\limsup_{n\to\infty} \inf_{y\in E}[\frac{1}{n} f_n(\omega).\frac{1}{n}\sum_{j=0}^{n-1}\Psi(T^j\omega, y) + k||u(\omega) - y||_E]$$

$$\leq \limsup_{n\to\infty}[\frac{1}{n} f_n(\omega).\frac{1}{n}\sum_{j=0}^{n-1}\Psi(T^j\omega, v_{k,p,u}(\omega)) + k||u(\omega) - v_{k,p,u}(\omega)||_E]$$

$$= \gamma(\omega)E^{\mathcal{I}}\Psi(\omega, v_{k,p,u}(\omega) + k||u(\omega) - v_{k,p,u}(\omega)||_E \leq \Delta^k(\omega, u(\omega)) + \frac{1}{p} \quad \forall\omega \notin N_{k,u}.$$

Taking the supremum on $k \in \mathbf{N}$ in this inequality yields a negligible set $N_u := \cup_{k\in\mathbf{N}, p\in\mathbf{N}} N_{k,p,u}$ such that

$$\sup_{k\in\mathbf{N}} \limsup_{n\to\infty} \inf_{y\in E}[\frac{1}{n} f_n(\omega).\frac{1}{n}\sum_{j=0}^{n-1}\Psi(T^j\omega, y) + k||u(\omega) - y||_E]$$

$$\leq \sup_{k\in\mathbf{N}} \Delta^k(\omega, u(\omega))$$

$$= \Delta(\omega, u(\omega)) = \sup_{n\in\mathbf{N}} \frac{1}{n}E^{\mathcal{I}} f_n(\omega).\, E^{\mathcal{I}}\Psi(\omega, u(\omega)) \qquad \forall\omega \notin N_u.$$

■

Here is a convex variant dealing with $L^\infty_{\overline{B}_E}(\Omega, \mathcal{I}, P)$ and involving the application of Theorem 5.4.

Corollary 5.4. *Let E be a separable Banach space, T a measurable transformation of Ω preserving P, $\mathcal{I}$ the σ algebra of invariants sets and $f : \Omega \times E^*_s \to \overline{\mathbf{R}}$ be a $\mathcal{F} \times \mathcal{B}(E^*_s)$-measurable normal convex integrand such that $f^+(\omega, 0) = 0$ for all $\omega \in \Omega$. Let*

$$f^*(\omega, x) := \delta^*((x, -1), epi f(\omega, .)) \quad \forall(\omega, x) \in \Omega \times E.$$

Let $(f_n)_{n\in\mathbf{N}}$ be a positive superadditive integrable sequence. Then for any $u \in L^\infty_{\overline{B}_E}(\Omega, \mathcal{I}, P)$ such that $f^*(., u(.)) \in L^1_{\mathbf{R}}(\Omega, \mathcal{F}, P)$,[1] the following epi-convergence result holds:

$$(^{**}) \quad \sup_{k\in\mathbf{N}} \limsup_{n\to\infty} \inf_{y\in\overline{B}_E} [\frac{1}{n} f_n(\omega).\frac{1}{n}\sum_{j=0}^{n-1} f^*(T^j\omega, y) + k||u(\omega) - y||_E]$$

$$\leq \sup_{n\in\mathbf{N}} \frac{1}{n}E^{\mathcal{I}} f_n(\omega).\, g^*(\omega, u(\omega))$$

$$= \sup_{n\in\mathbf{N}} \frac{1}{n}E^{\mathcal{I}} f_n(\omega).\, E^{\mathcal{I}} f^*(\omega, u(\omega)) \quad a.s.$$

[1] For more consideration on the finiteness or continuity assumption of $I_{f^*}(u) := \int_\Omega f^*(\omega, u(\omega))dP(\omega)$ on $L^\infty_E(\Omega, \mathcal{F}, P)$, see [5], Chap. VIII.

where

$$g^*(\omega, x) = \delta^*((x, -1), E^{\mathcal{I}} epi f(\omega, .)) \quad \forall(\omega, x) \in \Omega \times E.$$

Proof. By virtue of Theorem 4.2 or Theorem 5.4, the conditional expectation g^* is a $\mathcal{I} \otimes \mathcal{B}(E)$-measurable normal integrand satisfying

$$\int_A f^*(\omega, v(\omega))dP(\omega) = \int_A g^*(\omega, v(\omega))dP(\omega)$$

for all $v \in L^\infty_E(\Omega, \mathcal{I}, P)$, and for all $A \in \mathcal{I}$ with

$$g^*(\omega, v(\omega)) := \delta^*((v(\omega), -1), E^{\mathcal{I}} epi f(\omega, .)) \quad \forall \omega \in \Omega.$$

Let us set

$$\gamma(\omega) := \sup_{n \in \mathbf{N}} \frac{1}{n} E^{\mathcal{I}} f_n(\omega) \quad \forall \omega \in \Omega.$$

$$\Delta(\omega, x) := \gamma(\omega)\, g^*(\omega, x) \quad \forall(\omega, x) \in \Omega \times \overline{B}_E.$$

$$\Delta^k(\omega, x) = \inf_{y \in \overline{B}_E} [\Delta(\omega, y) + k||x - y||_E] \quad \forall(\omega, x) \in \Omega \times \overline{B}_E.$$

Then

$$\sup_{k \in \mathbf{N}} \Delta^k(\omega, x) = \Delta(\omega, x) \quad \forall(\omega, x) \in \Omega \times \overline{B}_E.$$

By Kingman theorem for superadditive integrable sequences ([12], Theorem 10.7.1) we may assume that

$$\lim_{n \to \infty} \frac{1}{n} f_n(\omega) = \gamma(\omega) \quad \forall \omega \in \Omega.$$

Let $u \in L^\infty_{\overline{B}_E}(\Omega, \mathcal{I}, P)$ such that $f^*(., u(.)) \in L^1_{\mathbf{R}}(\Omega, \mathcal{F}, P)$. We may assume that $\gamma(\omega) \in [0, \infty[$ for all $\omega \in \Omega$. Let $p \in \mathbf{N}$. Since $\overline{B}_E$ is borelian in the separable Banach space E, applying measurable selection theorem ([5], Theorem III-22) yields a $(\mathcal{I}, \mathcal{B}(\overline{B}_E))$-measurable mapping $v_{k,p,u} : \Omega \to \overline{B}_E$ such that

$$0 \leq \gamma(\omega) g^*(\omega, v_{k,p,u}(\omega)) + k||u(\omega) - v_{k,p,u}(\omega))||_E \leq \Delta^k(\omega, u(\omega)) + \frac{1}{p} \quad \forall \omega \in \Omega.$$

From the classical Birkhoff ergodic theorem (see e.g. [18]) we provide a negligible set $N_{k,p,u}$ such that for $\omega \notin N_{k,p,u}$

$$\lim_{n \to \infty} \frac{1}{n} \sum_{j=0}^{n-1} f^*(T^j \omega, v_{k,p,u}(\omega)) = E^{\mathcal{I}}[f^*(\omega, v_{k,p,u}(\omega))] = g^*(\omega, v_{k,p,u}(\omega)) \quad a.s.$$

Whence we provide a negligible set $N_{k,p,u}$ such that

$$\limsup_{n\to\infty} \inf_{y\in \overline{B}_E} [\frac{1}{n} f_n(\omega).\frac{1}{n}\sum_{j=0}^{n-1} f^*(T^j\omega, y) + k||u(\omega) - y||_E]$$

$$\leq \limsup_{n\to\infty} [\frac{1}{n} f_n(\omega).\frac{1}{n}\sum_{j=0}^{n-1} f^*(T^j\omega, v_{k,p,u}(\omega)) + k||u(\omega) - v_{k,p,u}(\omega)||_E]$$

$$= \gamma(\omega)g^*(\omega, v_{k,p,u}(\omega)) + k||u(\omega) - v_{k,p,u}(\omega)||_E \leq \Delta^k(\omega, u(\omega))$$

$$+\frac{1}{p} \quad \forall \omega \notin N_{k,p,u}.$$

Taking the supremum on $k \in \mathbf{N}$ in this inequality yields a negligible set $N_u := \cup_{k\in\mathbf{N},p\in\mathbf{N}} N_{k,p,u}$ such that for $\omega \notin N_u$

$$\sup_{k\in\mathbf{N}} \limsup_{n\to\infty} \inf_{y\in \overline{B}_E} [\frac{1}{n} f_n(\omega).\frac{1}{n}\sum_{j=0}^{n-1} f^*(T^j\omega, y) + k||u(\omega) - y)||_E]$$

$$\leq \sup_{k\in\mathbf{N}} \Delta^k(\omega, u(\omega)) = \Delta(\omega, u(\omega)) = \sup_{n\in\mathbf{N}} \frac{1}{n} E^{\mathcal{I}} f_n(\omega).\, g^*(\omega, u(\omega)).$$

■

Comments. Several open variational convergence problems appear with the lower semicontinuous (lsc) normal integrands, for instance, the epiconvergence problems for lsc superadditive sequences and lsc integrand reversed martingales $(\Psi_n, \mathcal{B}_n)_{n\in\mathbf{N}}$, partial results are given in Theorem 5.5 with the *regular* lsc integrand reversed martingale $(E^{\mathcal{B}_n}\Psi, \mathcal{B}_n)_{n\in\mathbf{N}}$ and in Theorem 5.6 and Corollaries 5.2, 5.3, and 5.4 with specific lsc superadditive random sequences. Further, in view of applications to the strong law of large numbers, it would be interesting to develop the lsc integrand pramarts associated with a decreasing sequence of sub-σ-algebras of $\mathcal{F}$. Further contributions to the epiconvergence for the normal integrands, in particular, the parametric Birkhoff ergodic theorem, are given in [4, 11, 15, 18, 19], at this point, partial results were obtained in [4] and some of them are improved here. A powerful tool allowing to prove the existence of conditional expectation for normal integrands is provided in Theorem 2 in [18]. Another tool [9] involving a specific Hausdorff–Baire approximation for separately additive and separately lower semicontinuous integrand $F(A, x)$ $(A \in \mathcal{F}, x \in E)$ provides an integral representation theorem for this object and also the existence of conditional expectation for normal integrands. There is an abundant bibliography on these subjects, see [4, 11, 18]. The above stated theorems are related to other results in convex analysis, economics, probability, variational analysis.

References

1. Akhiat, F., Castaing, C., Ezzaki, F.: Some various convergence results for multivalued martingales. Adv. Math. **13**, 1–33 (2010)
2. Bismut, J.M.: Intégrales convexes et probabilités. J. Math. Anal. Appl. **42**, 639–673 (1973)
3. Castaing, C.: Compacité et inf-equicontinuity dans certains espaces de Köthe-Orlicz. Sém. Anal. Convexe, Montpellier, Exposé No 6 (1979)
4. Castaing, C., Ezzaki, F.: Variational inequalities for integrand martingales and additive random sequences. Seminaire Analyse Convexe, Montpellier 1992, Exposé No 1. Acta Math. Vietnam. **18**(1), 137–171 (1993)
5. Castaing, C., Valadier, M.: Convex Analysis and Measurable Multifunctions. Lecture Notes in Mathematics, vol. 580. Springer, Berlin (1977)
6. Castaing, C., Raynaud de Fitte, P., Valadier, M.: Young Measures on Topological Spaces. With Applications in Control Theory and Probability Theory. Kluwer Academic, Dordrecht (2004)
7. Castaing, C., Raynaud de Fitte, P., Salvadori, A.: Variational convergence results with applications to evolution inclusions. Adv. Math. Econ. **8**, 33–73 (2006)
8. Castaing, C., Hess, Ch., Saadoune, M.: Tightness conditions and integrability of the sequential weak upper limit of a sequence of multifunctions. Adv. Math. Econ. **11**, 11–44 (2008)
9. Castaing, C., Raynaud de Fitte, P.: Variational convergence on $\mathcal{Y}([0, 1], H) \times \mathcal{M}^b_H([0, 1])$. In: Working Paper, 08/2005
10. Castaing, C., Ezzaki, F., Tahri, K.: Convergences of multivalued pramarts. J. Nonlinear Convex Anal. **11**(2), 243–266 (2010)
11. Choirat, C., Hess, C., Seri, R.: A functional version of the Birkhoff ergodic theorem for a normal integrand: a variational approach. Ann. Probab. **1**, 63–92 (2003)
12. Dudley, R.M.: Real Analysis and Probability. Mathematics Series. Chapman-Hall, Wadsworth Inc. (1989)
13. Egghe, L.: Stopping Time Techniques for Analysts and Probabilists. Cambridge University Press, London (1984)
14. Hiai, F., Umegaki, H.: Integrals, conditional expectations and martingales of multivalued functions. J. Multivar. Anal. **7**, 149–182 (1977)
15. Jalby, V.: Semi-continuité, convergence et approximation des applications vectorielles. Loi des grands nombres, Université Montpellier II, Laboratoire Analyse Convexe, 34095 Montpellier Cedex 05, France, Janvier 1992

16. Neveu, N.: Martingales a Temps Discret. Masson et Cie, Paris (1972)
17. Valadier, M.: On conditional expectation of random sets. Annali Math. Pura Appl. (iv) **CXXVI**, 81–91 (1980)
18. Valadier, M.: Conditional expectation and ergodic theorem for a positive integrand. J. Nonlinear Convex Anal. **1**, 233–244 (2002)
19. Valadier, M.: Corrigendum to: conditional expectation and ergodic theorem for a positive integrand. J. Nonlinear Convex Anal. **3**, 123 (2002)

Adv. Math. Econ. 15, 27–39 (2011)

Advances in
MATHEMATICAL
ECONOMICS

A note on robust representations of law-invariant quasiconvex functions

Samuel Drapeau[1], Michael Kupper[1], and Ranja Reda[2,*]

[1] Humboldt University Berlin, Unter den Linden 6, 10099 Berlin, Germany
(e-mail: drapeau@mathematik.hu-berlin.de, kupper@mathematik.hu-berlin.de)

[2] Vienna University of Technology, Wiedner Haupstrasse 8/105-1, 1040 Vienna, Austria
(e-mail: ranja.reda@fam.tuwien.ac.at)

Received: June 2, 2010
Revised: October 7, 2010

JEL classification: C02, D80

Mathematics Subject Classification (2010): 91B02, 91B06, 91B30

Abstract. We give robust representations of law-invariant monotone quasiconvex functions. The results are based on Jouini et al. (Adv Math Econ 9:49–71, 2006) and Svindland (Math Financ Econ, 2010), showing that law-invariant quasiconvex functions have the Fatou property.

Key words: Fatou property, law-invariance, risk measure, robust representation

1. Introduction

The theory of *monetary risk measures* dates back to the end of the twentieth century where Artzner et al. in [1] introduced the *coherent cash additive risk measures* which were further extended to the *convex cash additive risk measures* by Föllmer and Schied in [7] and Frittelli and Gianin in [9].

* Financial support from the DFG IRTG 1339 and the Berlin Mathematical School is gratefully acknowledged by Samuel Drapeau. Financial support from MATHEON project E.11 is gratefully acknowledged by both Samuel Drapeau and Michael Kupper. Financial support by the Christian Doppler Research Association (CDG) is gratefully acknowledged by Ranja Reda. The author would like to thank Christina Ziehaus for helpful comments and fruitful discussions.

S. Kusuoka, T. Maruyama (eds.), *Advances in Mathematical Economics Volume 15*,
DOI: 10.1007/978-4-431-53930-8_2,

Monetary risk measures aim at specifying the capital requirement that financial institutions have to reserve in order to cope with severe losses from their risky financial activities. Recently, motivated by the study of *risk orders* in a general framework, Drapeau and Kupper in [4] defined *risk measures* as quasiconvex monotone functions. Building upon the latter, the aim of this note is to specify the robust representation of risk measures in the law-invariant case.

Robust representation of law-invariant monetary risk measures for bounded random variables have first been studied by Kusuoka in [13], then Frittelli and Gianin in [10] and further Jouini et al. in [11]. In a recent paper [2], Cerreia-Voglio et al. provide a robust representation for law-invariant risk measures which are weakly[1] upper semicontinuous.

In this note, we provide a robust representation of law-invariant risk measures for bounded random variables which are norm lower semicontinuous. This is based on results by Jouini et al. in [11] and Svindland in [14] showing that law-invariant norm-closed convex sets of bounded random variables are Fatou closed. This robust representation takes the form

$$\rho(X) = \sup_{\psi} R\left(\psi, \int_0^1 q_{-X}(s)\,\psi(s)\,ds\right),$$

where R is a *maximal risk function* which is uniquely determined, ψ are some nondecreasing right-continuous functions whose integral is normalized to 1, and q_X is the quantile function of the random variable X. We further provide a representation in the special case of norm lower semicontinuous law-invariant *convex cash subadditive risk measures* introduced by El Karoui and Ravanelli in [5]. Finally, we give a representation of time-consistent law-invariant monotone quasiconcave functions in the spirit of Kupper and Schachermayer in [12]. We illustrate these results by a couple of explicit computations for examples of law-invariant risk measures given by certainty equivalents.

2. Notations, definitions and the Fatou property

Throughout, $(\Omega, \mathcal{F}, \mathbb{P})$ is a standard probability space. We identify random variables which are almost surely (a.s.) identical. All equalities and inequalities between random variables are understood in the a.s. sense. As usual, $\mathbb{L}^\infty := \mathbb{L}^\infty(\Omega, \mathcal{F}, \mathbb{P})$ is the space of bounded random variables with topological dual $(\mathbb{L}^\infty)^*$. Following [4], a risk measure is defined as follows.

[1] For the weak*-topology $\sigma\left(\mathbb{L}^\infty, \mathbb{L}^1\right)$.

Definition 1. A *risk measure* on $\mathbb{L}^\infty$ is a function $\rho : \mathbb{L}^\infty \to [-\infty, +\infty]$ satisfying for any $X, Y \in \mathbb{L}^\infty$ the axioms of

(i) Monotonicity:

$$\rho(X) \geq \rho(Y), \quad \text{whenever } X \leq Y,$$

(ii) Quasiconvexity:

$$\rho(\lambda X + (1-\lambda)Y) \leq \max\{\rho(X), \rho(Y)\}, \quad \text{for any } 0 \leq \lambda \leq 1.$$

Further particular risk measures used in this paper satisfy some of the following additional properties,

(i) Cash additivity if $\rho(X+m) = \rho(X) - m$ for any $m \in \mathbb{R}$.
(ii) Cash subadditivity if $\rho(X+m) \geq \rho(X) - m$ for any $m \geq 0$.
(iii) Convexity if $\rho(\lambda X + (1-\lambda)Y) \leq \lambda\rho(X) + (1-\lambda)\rho(Y)$, for any $0 \leq \lambda \leq 1$.
(iv) Law-invariance if $\rho(X) = \rho(Y)$ whenever X and Y have the same law.

A risk measure satisfies the *Fatou property* if

$$\rho(X) \leq \liminf_{n\to\infty} \rho(X_n) \quad \text{whenever} \quad \sup_n \|X_n\|_\infty < \infty \quad \text{and} \quad X_n \xrightarrow{\mathbb{P}} X, \tag{1}$$

where $\xrightarrow{\mathbb{P}}$ denotes convergence in probability. A reformulation of the results in [11] in the context of quasiconvex law-invariant functions yields the following result.

Proposition 1. *Let $f : \mathbb{L}^\infty \to [-\infty, +\infty]$ be a $||\cdot||_\infty$-lower semicontinuous, quasiconvex and law-invariant function. Then, f is $\sigma(\mathbb{L}^\infty, \mathbb{L}^1)$-lower semicontinuous and has the Fatou property.*

Proof. Let $C \subset \mathbb{L}^\infty$ be a $||\cdot||_\infty$-closed, convex, law-invariant set with polar C° in $(\mathbb{L}^\infty)^*$. In view of Proposition 4.1 in [11], it follows that $C^\circ \cap \mathbb{L}^1$ is $\sigma((\mathbb{L}^\infty)^*, \mathbb{L}^\infty)$-dense in C°. Hence, $C = (C^\circ \cap \mathbb{L}^1)^\circ$, showing that C is $\sigma(\mathbb{L}^\infty, \mathbb{L}^1)$-closed.

Consider now a law-invariant, quasiconvex and $||\cdot||_\infty$-lower semicontinuous function $f : \mathbb{L}^\infty \to [-\infty, +\infty]$. By assumption, the level sets $\mathcal{A}_m := \{X \in \mathbb{L}^\infty \mid f(X) \leq m\}$, $m \in \mathbb{R}$, are $||\cdot||_\infty$-closed, convex and law-invariant. Hence, $\mathcal{A}_m$ are $\sigma(\mathbb{L}^\infty, \mathbb{L}^1)$-closed, showing that f is $\sigma(\mathbb{L}^\infty, \mathbb{L}^1)$-lower semicontinuous. Finally a similar argumentation as in [3] yields the Fatou property.

Remark 1. Any law-invariant, proper convex function $f : \mathbb{L}^\infty \to [-\infty, \infty]$ is $\sigma(\mathbb{L}^\infty, \mathbb{L}^\infty)$-lower semicontinuous, see [6].

Recently, a similar result is shown in [14] in the more general setting of non-atomic probability spaces rather than standard probability spaces.

3. Representation results for law-invariant risk measures

Throughout,

$$q_X(t) := \inf\left\{s \in \mathbb{R} \,\middle|\, \mathbb{P}[X \le s] \ge t\right\}, \qquad t \in (0,1)$$

denotes the quantile function of a random variable $X \in \mathbb{L}^\infty$. Let Ψ be the set of integrable, nondecreasing, right-continuous functions $\psi : (0,1) \to [0,+\infty)$ and define the subsets

$$\Psi_1 := \left\{\psi \in \Psi \,\middle|\, \int_0^1 \psi(u)\,du = 1\right\},$$
$$\Psi_{1,s} := \left\{\psi \in \Psi \,\middle|\, \int_0^1 \psi(u)\,du \le 1\right\}.$$

Denote by Ψ_1^∞ and $\Psi_{1,s}^\infty$ the set of all bounded functions in Ψ_1 and $\Psi_{1,s}$, respectively. It is shown in [8], Theorem 4.54, that any law-invariant cash additive risk measure ρ on $\mathbb{L}^\infty$ that satisfies the Fatou property has the robust representation

$$\rho(X) = \sup_{\psi \in \Psi_1} \left(\int_0^1 \psi(t) q_{-X}(t)\,dt - \alpha_{min}(\psi)\right), \quad X \in \mathbb{L}^\infty, \tag{2}$$

where $\alpha_{\min}(\psi) = \sup_{X \in \mathcal{A}_\rho} \int_0^1 \psi(t) q_{-X}(t)\,dt$ is the minimal penalty function for the acceptance set $\mathcal{A}_\rho := \{X \in \mathbb{L}^\infty \mid \rho(X) \le 0\}$.

In a first step, we derive the following representation result for law-invariant cash sub-additive convex risk measures.

Proposition 2. *Let ρ be a law-invariant cash sub-additive convex risk measure on $\mathbb{L}^\infty$. Then ρ has the robust representation*

$$\rho(X) = \sup_{\psi \in \Psi_{1,s}^\infty} \left(\int_0^1 \psi(t) q_{-X}(t)\,dt - \alpha_{\min}(\psi)\right), \quad X \in \mathbb{L}^\infty,$$

for the minimal penalty function

$$\alpha_{\min}(\psi) = \sup_{X \in \mathbb{L}^\infty} \left(\int_0^1 \psi(t) q_{-X}(t)\,dt - \rho(X)\right), \quad \psi \in \Psi_{1,s}.$$

Proof. According to Theorem 4.3 in [5] it follows

$$\rho(X) = \sup_{\mathbb{Q} \in \mathcal{M}_{1,s}(\mathbb{P})} \left(\mathbb{E}_{\mathbb{Q}}[-X] - \tilde{\alpha}_{\min}(\mathbb{Q})\right),$$

where $\tilde{\alpha}_{\min}(\mathbb{Q}) = \sup_{X\in\mathbb{L}^\infty}(\mathbb{E}_{\mathbb{Q}}[-X] - \rho(X))$ and $\mathcal{M}_{1,s}(\mathbb{P})$ denotes the set of measures $\mathbb{Q}$ absolutely continuous with respect to $\mathbb{P}$ such that $\mathbb{E}[d\mathbb{Q}/d\mathbb{P}] \leq 1$. By Lemma 4.55 in [8] and the law-invariance of ρ we deduce

$$\begin{aligned}\tilde{\alpha}_{\min}(\mathbb{Q}) &= \sup_{X\in\mathbb{L}^\infty} \sup_{Y\sim X} \left(\mathbb{E}_{\mathbb{Q}}[-Y] - \rho(Y)\right) \\ &= \sup_{X\in\mathbb{L}^\infty} \left(\int_0^1 \psi(t) q_{-X}(t)\, dt - \rho(X)\right) = \alpha_{min}(\psi),\end{aligned}$$

for any $\psi \in \Psi_{1,s}$ and $\mathbb{Q} \in \mathcal{M}_{1,s}(\mathbb{P})$ with $\psi = q_{d\mathbb{Q}/d\mathbb{P}}$. Finally, under consideration of Remark 1 and Lemma 4.55 in [8], it follows

$$\begin{aligned}\rho(X) &= \sup_{\mathbb{Q}\in\mathcal{M}^\infty_{1,s}(\mathbb{P})} \left(\mathbb{E}_{\mathbb{Q}}[-X] - \tilde{\alpha}_{\min}(\mathbb{Q})\right) \\ &= \sup_{\mathbb{Q}\in\mathcal{M}^\infty_{1,s}(\mathbb{P})} \sup_{\tilde{\mathbb{Q}}\sim\mathbb{Q}} \left(\mathbb{E}_{\tilde{\mathbb{Q}}}[-X] - \tilde{\alpha}_{\min}(\tilde{\mathbb{Q}})\right) \\ &= \sup_{\psi\in\Psi^\infty_{1,s}} \left(\int_0^1 \psi(t) q_{-X}(t) dt - \alpha_{\min}(\psi)\right),\end{aligned}$$

where $\mathcal{M}^\infty_{1,s}(\mathbb{P})$ are those elements in $\mathcal{M}_{1,s}(\mathbb{P})$ with a bounded Radon–Nikodým derivative.

As a second step, we state our main result: a quantile representation for $||\cdot||_\infty$-lower semicontinuous law-invariant risk measures. Beforehand, as in [4], we define the class of *maximal risk functions* $\mathcal{R}^{\max}$ as the set of functions $R : \Psi_1 \times \mathbb{R} \to [-\infty, +\infty]$ which

(i) Are nondecreasing and left-continuous in the second argument,
(ii) Are jointly quasiconcave,
(iii) Have a uniform asymptotic minimum, that is,

$$\lim_{s\to-\infty} R(\psi_1, s) = \lim_{s\to-\infty} R(\psi_2, s)$$

for any $\psi_1, \psi_2 \in \Psi_1$,
(iv) Right-continuous version $R^+(\psi, s) := \inf_{s'>s} R(\psi, s')$, are $\sigma(\mathbb{L}^1, \mathbb{L}^\infty)$-upper semicontinous in the first argument.

Theorem 1. *Let $\rho : \mathbb{L}^\infty \to [-\infty, +\infty]$ be a law invariant $||\cdot||_\infty$-lower semicontinuous risk measure. Then, there exists a unique risk function $R \in \mathcal{R}^{\max}$ such that*

$$\rho(X) = \sup_{\psi\in\Psi_1} R\left(\psi, \int_0^1 q_{-X}(t)\psi(t)\, dt\right), \quad X \in \mathbb{L}^\infty$$

where

$$R(\psi, x) = \sup_{m\in\mathbb{R}} \left\{ m \;\middle|\; \alpha_{\min}(\psi, m) < x \right\}, \quad \psi \in \Psi_1$$

for

$$\alpha_{\min}(\psi, m) = \sup_{X\in\mathcal{A}^m} \int_0^1 q_{-X}(t)\psi(t)\,dt$$

and $\mathcal{A}^m = \{X \in \mathbb{L}^\infty \mid \rho(X) \leq m\}$.

The proof of the previous theorem is based on the following proposition.

Proposition 3. *Suppose that* $\mathcal{A} \subset \mathbb{L}^\infty$ *is law-invariant,* $\|\cdot\|_\infty$*-closed, convex and such that* $\mathcal{A} + \mathbb{L}^\infty_+ \subset \mathcal{A}$. *Then*

$$X \in \mathcal{A} \iff \int_0^1 q_{-X}(t)\psi(t)\,dt \leq \alpha_{\min}(\psi) \quad \textit{for all } \psi \in \Psi_1, \tag{3}$$

where

$$\alpha_{\min}(\psi) := \sup_{X\in\mathcal{A}} \int_0^1 q_{-X}(t)\psi(t)\,dt, \quad \psi \in \Psi_1.$$

Proof. Associated to the set $\mathcal{A}$ we define

$$\rho_{\mathcal{A}}(X) := \inf\{m \in \mathbb{R} \mid X + m \in \mathcal{A}\}, \quad X \in \mathbb{L}^\infty.$$

The function $\rho_{\mathcal{A}} : \mathbb{L}^\infty \to [-\infty, +\infty]$ is a law-invariant, convex risk measure. Since

$$\{X \in \mathbb{L}^\infty \mid \rho_{\mathcal{A}}(X) \leq m\} = \mathcal{A} - m, \tag{4}$$

which by Proposition 1 is $\sigma(\mathbb{L}^\infty, \mathbb{L}^1)$-closed, it follows that $\rho_{\mathcal{A}}$ is $\sigma(\mathbb{L}^\infty, \mathbb{L}^1)$-l.s.c.. Moreover, one of the following cases must be valid:

(i) $\mathcal{A} = \emptyset$, $\rho_{\mathcal{A}} \equiv +\infty$ and $\alpha_{\min} \equiv -\infty$;
(ii) $\mathcal{A} = \mathbb{L}^\infty$, $\rho_{\mathcal{A}} \equiv -\infty$ and $\alpha_{\min} \equiv +\infty$;
(iii) $\mathcal{A} \neq \emptyset$ and $\mathcal{A} \neq \mathbb{L}^\infty$, in which case $\rho_{\mathcal{A}}$ is real-valued. Indeed, if there is $X, Y \in \mathbb{L}^\infty$ such that $X \in \mathcal{A}$ and $Y \notin \mathcal{A}$, then there is $n \in \mathbb{R}$ such that $X + n \notin \mathcal{A}$ showing that $\rho_{\mathcal{A}}(X) \in \mathbb{R}$. By monotonicity and translation invariance of $\rho_{\mathcal{A}}$, it follows that $\rho_{\mathcal{A}}(Z) \in \mathbb{R}$ for all $Z \in \mathbb{L}^\infty$.

For the cases (i) and (ii), the equivalence (3) is obvious. As for the third case, it follows from (2) that

$$\rho_{\mathcal{A}}(X) = \sup_{\psi\in\Psi_1} \left(\int_0^1 q_{-X}(t)\psi(t)dt - \alpha_{\min}(\psi) \right),$$

which together with (4) implies (3).

We are now ready for the proof of Theorem 1.

Proof. The *risk acceptance family* $\mathcal{A} = (\mathcal{A}_m)_{m\in\mathbb{R}}$ defined as

$$\mathcal{A}^m := \left\{X \in \mathbb{L}^\infty \mid \rho(X) \leq m\right\},$$

is law-invariant, $||\cdot||_\infty$-closed, convex and such that $\mathcal{A} + \mathbb{L}^\infty_+ \subset \mathcal{A}$. Thus, Proposition 3 implies

$$X \in \mathcal{A}^m \iff \int_0^1 q_{-X}(t)\psi(t)dt - \alpha_{\min}(\psi, m) \leq 0 \quad \text{for all } \psi \in \Psi_1, \tag{5}$$

for the family of penalty functions

$$\alpha_{\min}(\psi, m) = \sup_{X\in\mathcal{A}^m} \int_0^1 q_{-X}(t)\psi(t)\,dt, \quad \psi \in \Psi_1.$$

Since for all $X \in \mathbb{L}^\infty$

$$\rho(X) = \inf\left\{m \in \mathbb{R} \;\middle|\; X \in \mathcal{A}^m\right\}, \tag{6}$$

it follows from (5) that

$$\rho(X) = \inf\left\{m \in \mathbb{R} \;\middle|\; \int_0^1 q_{-X}(t)\psi(t)\,dt \leq \alpha_{\min}(\psi, m) \text{ for all } \psi \in \Psi_1\right\}. \tag{7}$$

The goal is to show that

$$\rho(X) = \sup_{\psi\in\Psi_1} \inf_{m\in\mathbb{R}} \left\{m \;\middle|\; \int_0^1 q_{-X}(t)\psi(t)\,dt \leq \alpha_{\min}(\psi, m)\right\}. \tag{8}$$

To begin with, the equation (3) implies:

$$\rho(X) \geq \sup_{\psi\in\Psi_1} \inf_{m\in\mathbb{R}} \left\{m \;\middle|\; \int_0^1 q_{-X}(t)\psi(t)\,dt \leq \alpha_{\min}(\psi, m)\right\}.$$

As for the reverse inequality, suppose that $\rho(X) > -\infty$, otherwise (8) is trivial, and fix $m_0 < \rho(X)$. Define $C = \{Y \in \mathbb{L}^\infty \mid \rho(Y) \leq m_0\}$, which is law-invariant, $\|\cdot\|_\infty$-closed, convex, such that $C + \mathbb{L}^\infty \subset C$. Thus, Proposition 3 yields

$$Y \in C \iff \int_0^1 q_{-Y}(t)\psi(t)\,dt \leq \alpha_C(\psi) \quad \text{for all } \psi \in \Psi_1, \tag{9}$$

for the penalty function $\alpha_C(\psi) = \sup_{Y\in C} \int_0^1 q_{-Y}(t)\psi(t)\,dt$. Since $X \notin C$, it follows from (9) that there is $\psi^* \in \Psi_1$ such that

$$\int_0^1 q_{-X}(t)\psi^*(t)\,dt > \alpha_C(\psi^*) \geq \int_0^1 q_{-Y}(t)\psi^*(t)\,dt \quad \text{for all } Y \in C. \tag{10}$$

Since $\mathcal{A}^m \subset C$ for all $m \leq m_0$ and therefore $\alpha_{\min}(\psi^*, m) \leq \alpha_C(\psi^*)$, it follows

$$\begin{aligned}\int_0^1 q_{-X}(t)\psi^*(t)\,dt &- \alpha_{\min}(\psi^*, m)\\ &\geq \int_0^1 q_{-X}(t)\psi^*(t)\,dt - \sup_{Y\in C}\int_0^1 q_{-Y}(t)\psi^*(t)\,dt > 0.\end{aligned} \tag{11}$$

Hence,

$$m_0 \leq \sup_{\psi\in\Psi_1} \inf_{m\in\mathbb{R}} \left\{ m \;\middle|\; \int_0^1 q_{-X}(t)\psi(t)\,dt \leq \alpha_{\min}(\psi, m) \right\}. \tag{12}$$

Since (12) holds for all $m_0 < \rho(X)$ we deduce

$$\rho(X) \leq \sup_{\psi\in\Psi_1} \inf_{m\in\mathbb{R}} \left\{ m \;\middle|\; \int_0^1 q_{-X}(t)\psi(t)\,dt \leq \alpha_{\min}(\psi, m) \right\},$$

and (8) is established.

Let $R(\psi, x) := \sup_{m\in\mathbb{R}} \left\{ m \;\middle|\; \alpha_{\min}(\psi, m) < x \right\}$ be the left-inverse of $\alpha_{\min}$. Then

$$\rho(X) = \sup_{\psi\in\Psi_1} R\left(\psi, \int_0^1 q_{-X}(t)\psi(t)\,dt\right) \quad \text{for all } X \in \mathbb{L}^\infty. \tag{13}$$

The proof of the existence is completed. The uniqueness follows from a similar argumentation as in [4].

4. Time-consistent law-invariant quasiconcave functions

As an application of Proposition 1 we discuss an extension of the representation results for time-consistent law-invariant strictly monotone functions given in [12]. In this subsection, we work on a standard filtered probability space[2] $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{N}_0}, \mathbb{P}\right)$. We fix $-\infty \leq a < b \leq +\infty$ and denote by $\mathbb{L}_t^\infty(a,b)$ and $\mathbb{L}^\infty(a,b)$ the set of all random variables X such that $a < \operatorname{ess\,inf} X \leq \operatorname{ess\,sup} X < b$ and which are $\mathcal{F}_t$-measurable and $\mathcal{F}$-measurable, respectively. A function $c_0 : \mathbb{L}^\infty(a,b) \to \mathbb{R}$ is

(i) *Normalized on constants* if $c_0(m) = m$ for all $m \in (a,b)$;
(ii) *Strictly monotone* if $X \geq Y$ and $\mathbb{P}[X > Y] > 0$ imply $c_0(X) > c_0(Y)$;
(iii) *Time-consistent* if for any $t \in \mathbb{N}$ there exists a mapping $c_t : \mathbb{L}^\infty(a,b) \to \mathbb{L}_t^\infty(a,b)$ which satisfies the *local property*, that is, for any $X, Y \in \mathbb{L}^\infty(a,b)$

$$1_A X = 1_A Y \quad \text{implies} \quad 1_A c_t(X) = 1_A c_t(Y) \quad \text{for all } A \in \mathcal{F}_t, \tag{14}$$

and

$$c_0(X) = c_0(c_t(X)) \quad \text{for all } X \in \mathbb{L}^\infty(a,b). \tag{15}$$

Under the additional assumption of quasiconcavity we deduce as a corollary of Theorem 1.4 in [12]:

Theorem 2. *A function $c_0 : \mathbb{L}^\infty(a,b) \to \mathbb{R}$ is normalized on constants, strictly monotone, $||\cdot||_\infty$-continuous, law-invariant, time-consistent and quasiconcave if and only if*

$$c_0(X) = u^{-1} \circ \mathbb{E}[u(X)], \tag{16}$$

for an increasing, concave function $u : (a,b) \to \mathbb{R}$. In this case, the function u is uniquely defined up to positive affine transformations, and

$$c_t(X) = u^{-1} \circ \mathbb{E}[u(X) \mid \mathcal{F}_t] \quad \textit{for all } t \in \mathbb{N}. \tag{17}$$

Proof. Fix a compact interval $[A,B] \subset (a,b)$. Since $\mathbb{L}^\infty[A,B] := \{X \in \mathbb{L}^\infty \mid A \leq X \leq B\}$ is $||\cdot||_\infty$-closed in $\mathbb{L}^\infty$, it follows that

$$c_0^{A,B}(X) := \begin{cases} c_0(X) & \text{if } X \in \mathbb{L}^\infty[A,B] \\ -\infty & \text{else} \end{cases}, \quad X \in \mathbb{L}^\infty,$$

is law-invariant, quasiconcave and $||\cdot||_\infty$-upper semicontinuous. Due to Proposition 1, the function $c_0^{A,B}$ has the Fatou property and consequently the condition (C) in [12] is satisfied. Hence, by Theorem 1.4 in [12] there is $u_{A,B} : (A,B) \to \mathbb{R}$ such that

---

[2] Recall that a standard filtered probability space is isomorphic to $([0,1]^{\mathbb{N}_0}, \mathcal{B}([0,1]^{\mathbb{N}_0}), (\mathcal{F}_t)_{t\in\mathbb{N}_0}, \lambda^{\mathbb{N}_0})$ where $\mathcal{B}([0,1]^{\mathbb{N}_0})$ is the Borel sigma-algebra, $\lambda^{\mathbb{N}_0}$ is the product of Borel measures, and $(\mathcal{F}_t)_{t\in\mathbb{N}_0}$ is the filtration generated by the coordinate functions.

$$c_0^{A,B}(X) = u_{A,B}^{-1} \circ \mathbb{E}\left[u_{A,B}(X)\right], \quad X \in \mathbb{L}^\infty(A,B).$$

Exhausting (a, b) by increasing compact intervals, as in the proof of the "only if"-part of Theorem 1.4 in [12], yields $u : (a, b) \to \mathbb{R}$ such that $c_0(X) = u^{-1} \circ \mathbb{E}[u(X)]$ for all $X \in \mathbb{L}^\infty(a, b)$. Finally, it is shown in Lemma 2 in [2] that c_0 is quasiconcave if and only if u is concave.

5. Examples

The certainty equivalent of a random variable provides a typical example of a law-invariant risk measure which is not necessarily convex nor cash additive. Let $l : \mathbb{R} \to]-\infty, +\infty]$ be a loss function, that is, a lower semicontinuous proper convex nondecreasing function. By $l^{-1} : \mathbb{R} \to [-\infty, +\infty[$ we denote the left-inverse of l given by

$$l^{-1}(s) = \inf\left\{x \in \mathbb{R} \,\middle|\, l(x) \geq s\right\}, \quad s \in \mathbb{R}.$$

We further denote by $l(x+) := \lim_{t \searrow x} l(t)$ for $x \in \mathbb{R}$ the right-continuous version of l. By Proposition B.2 in [4], we have

$$l^{-1}(s) \leq x \iff s \leq l(x+). \tag{18}$$

We now define the risk measure

$$\rho(X) := l^{-1}\mathbb{E}[l(-X)], \qquad X \in \mathbb{L}^\infty, \tag{19}$$

with convention that $l^{-1}(+\infty) = \lim_{s\to+\infty} l^{-1}(s)$. In [2, 4] it is shown that ρ is a risk measure. Note that in [2], it is assumed that l is real-valued and increasing, and therefore does not include some of the examples below. In [4], a constructive method is given to compute the robust representation. To be self contained, we present this method in the law-invariant context hereafter. For simplicity we suppose that l is differentiable on the interior of its domain and first compute the minimal penalty function at any risk level m. From relation (18) follows

$$\begin{aligned}
\alpha_{\min}(\psi, m) &= \sup_{X\in\mathcal{A}^m} \int_0^1 q_{-X}(s)\psi(s)\,ds \\
&= \sup_{\left\{X \mid \int_0^1 l(q_{-X}(s))ds \leq l(m+)\right\}} \int_0^1 q_{-X}(s)\psi(s)\,ds \\
&= \sup_{X\in\mathbb{L}^\infty} \int_0^1 \left[q_{-X}(s)\psi(s) - \frac{1}{\beta}\Big(l(q_{-X}(s)) - l(m+)\Big)\right] ds,
\end{aligned} \tag{20}$$

for some Lagrange multiplier $\beta := \beta(\psi, m) > 0$. The first order condition implies

$$\psi - \frac{1}{\beta} l'\left(q_{-\hat{X}}\right) = 0.$$

Since l' is nondecreasing, denote by h its right-inverse. Assuming that $q_{-\hat{X}} = h(\beta\psi)$ fulfills the previous condition,[3] then, under integrability and positivity conditions, β is determined through the equation

$$\int_0^1 l\Big(h\big(\beta\psi(s)\big)\Big) ds = l(m+). \tag{21}$$

Plugging the optimizer $q_{-\hat{X}}$ in (20) yields

$$\alpha_{\min}(\psi, m) = \int_0^1 h(\beta\psi(s))\, \psi(s)\, ds. \tag{22}$$

We subsequently list closed form solutions for some specific loss functions.

– **Quadratic Function:** Suppose that $l(x) = x^2/2 + x$ for $x \geq -1$ and $l(x) = -1/2$ elsewhere. In this case, $E[l(-X)]$ corresponds to a monotone version of the mean-variance risk measure of Markowitz. Here, $l^{-1}(s) = \sqrt{2s+1} - 1$ if $s > -1/2$ and $-\infty$ elsewhere, therefore

$$\rho(X) := \begin{cases} \sqrt{2\mathbb{E}\left[-X + \dfrac{X^2}{2}\right] + 1} - 1 & \text{if } \mathbb{E}[-X] + \dfrac{1}{2}\mathbb{E}\left[X^2\right] > -\dfrac{1}{2} \\ -\infty & \text{else} \end{cases}. \tag{23}$$

For $m \leq -1$, since $1 \in \mathcal{A}^m$, it is clear that $\alpha_{\min}(\psi, m) = -\int_0^1 \psi(s)\, ds = -1$. Otherwise, the first order condition yields $q_{-\hat{X}} = \beta\psi - 1$ and therefore

$$\alpha_{\min}(\psi, m) = (1+m)\left(\int_0^1 \psi(s)^2\, ds\right)^{1/2} - 1.$$

By inversion follows

$$R(\psi, s) = (s+1) \Big/ \left(\int_0^1 \psi(s)^2 ds\right)^{1/2} - 1,$$

if $s > -1$, and $R(\psi, s) = -\infty$ elsewhere, and therefore

[3] This is often the case, in particular when l' is increasing.

$$\rho(X) = \sup_{\psi \in \Psi_1} \left\{ \frac{\int_0^1 q_{-X}(s)\,\psi(s)\,ds + 1}{\left(\int_0^1 \psi(s)^2 ds\right)^{1/2}} - 1 \,\middle|\, \int_0^1 q_{-X}(s)\,\psi(s)\,ds > -1 \right\}. \tag{24}$$

– **Exponential Function:** If $l(x) = e^x - 1$, then

$$\rho(X) := \ln\left(\mathbb{E}\left[e^{-X}\right]\right) = \sup_{\psi \in \Psi_1} \left\{ \int_0^1 \left(q_{-X}(s)\,\psi(s) - \psi(s) \log \psi(s) \right) ds \right\}. \tag{25}$$

– **Logarithm Function:** If $l(x) = -\ln(-x)$ for $x < 0$ and $l = +\infty$ elsewhere, then

$$\rho(X) := -\exp\left(\mathbb{E}\left[\ln(X)\right]\right) = \sup_{\psi \in \Psi_1} \left\{ \frac{\int_0^1 q_{-X}(s)\,\psi(s)\,ds}{\exp\left(\int_0^1 \ln \psi(s)\,ds\right)} \right\}. \tag{26}$$

– **Power Function:** If $l(x) = -(-x)^{1-\gamma}/(1-\gamma)$ for $x \leq 0$ and $l = +\infty$ elsewhere whereby $0 < \gamma < 1$, we obtain

$$\rho(X) = \sup_{\psi \in \Psi_1} \left\{ \left(\int_0^1 \psi(s)^{\frac{\gamma-1}{\gamma}} ds \right)^{\frac{\gamma}{1-\gamma}} \int_0^1 q_{-X}(s)\,\psi(s)\,ds \right\}. \tag{27}$$

References

1. Artzner, Ph., Delbaen, F., Eber, J.M., Heath, D.: Coherent risk measures. Math. Finance **9**(3), 203–228 (1999)
2. Cerreia-Vioglio, S., Maccheroni, F., Marinacci, M., Montrucchio, L.: Risk measures: rationality and diversification. Math. Finance (2010)
3. Delbaen, F.: Coherent measure of risk on general probability spaces. In: Sandmann, K., Schonbucher, P.J. (eds.) Advances in Finance and Stochastics, Essays in Honor of Dieter Sondermann, pp. 1–37. Springer, Berlin (2002)
4. Drapeau, S., Kupper, M.: Risk preferences and their robust representation. Preprint (SSRN) (2010)
5. El Karoui, N., Ravanelli, C.: Cash sub-additive risk measures and interest rate ambiguity. Math. Finance **19**(4), 561–590 (2008)

6. Filipović, D., Svindland, G.: The canonical model space for law-invariant convex risk measures is L^1. Math. Finance (2008)
7. Föllmer, H., Schied, A.: Convex measure of risk and trading constraints. Finance Stoch. **6**, 429–447 (2002)
8. Föllmer, H., Schied, A.: Stochastic finance, an introduction in discrete time. de Gruyter Studies in Mathematics **27** (2002)
9. Frittelli, M., Rosazza Gianin, E.: Putting order in risk measures. J. Bank. Finance **26**(7), 1473–1486 (2002)
10. Frittelli, M., Rosazza Gianin, E.: Law-invariant convex risk measures. Adv. Math. Econ. **7**, 33–46 (2005)
11. Jouini, E., Schachermayer, W., Touzi, N.: Law invariant risk measures have the Fatou property. Adv. Math. Econ. **9**, 49–71 (2006)
12. Kupper, M., Schachermayer, W.: Representation results for law invariant time consistent functions. Math. Financ. Econ. **2**(3), 189–210 (2009)
13. Kusuoka, S.: On law-invariant coherent risk measures. Adv. Math. Econ. **3**, 83–95 (2001)
14. Svindland, G.: Continuity properties of law-invariant (quasi-)convex risk functions. Math. Financ. Econ. **1**, 39–43 (2010)

Adv. Math. Econ. 15, 41–65 (2011)

Advances in
MATHEMATICAL
ECONOMICS

On the Fourier analysis approach to the Hopf bifurcation theorem

Toru Maruyama*

Department of Economics, Keio University, 2-15-45 Mita, Minato-ku,
Tokyo 108-8345, Japan
(e-mail: maruyama@econ.keio.ac.jp)

Received: July 12, 2010
Revised: October 19, 2010

JEL classification: E32

Mathematics Subject Classification (2010): 34C23, 34C25, 42A20

Abstract. Ambrosetti and Prodi (A primer of nonlinear analysis, Cambridge University Press, Cambridge, 1993) formulated an abstract version of the Hopf bifurcation theorem and tried to deduce the well-known classical result from it. In this paper, we examine the Hopf bifurcation phenomena in the framework of a Sobolev space (rather than C^r), having recourse to the Carleson–Hunt theory. Some more careful reasonings to evaluate the magnitudes of the Fourier coefficients seem to be required in order to implement the Ambrosetti–Prodi approach in their classical setting. We incidentally try to fortify their way of proof from the standpoint of classical Fourier analysis.

Key words: Hopf bifurcation, periodic solution for ordinary differential equation, Fourier series

1. Introduction

The Hopf bifurcation theorem provides an effective criterion for finding out periodic solutions for ordinary differential equations. Although various proofs of this classical theorem are known, there seems to be no easy way to arrive at the goal. Among them, the idea of Ambrosetti and Prodi [1] is particularly noteworthy.

* I would like to express my sincere thanks to Professors S. Kusuoka, P.H. Rabinowitz and Y. Takahashi for their invaluable comments on the earlier draft.

S. Kusuoka, T. Maruyama (eds.), *Advances in Mathematical Economics Volume 15*, 41
DOI: 10.1007/978-4-431-53930-8_3,

They start with formulating the Hopf theorem in an abstract fashion and then try to deduce the classical result from it. Let $F(\omega, \mu, \cdot)$ be a smooth function of a Banach space $\mathfrak{X}$ into another one $\mathfrak{Y}$ with a couple (ω, μ) of real parameters. In order to find out some bifurcation point (ω^*, μ^*) of the equation $F(\omega, \mu, x) = 0$, the derivative $D_x F(\omega^*, \mu^*, 0)$ of F with respect to $x \in \mathfrak{X}$ at $(\omega^*, \mu^*, 0)$ plays a crucial role. As is stated in Theorem 1 exactly, the condition that both of the dimension of the kernel of $D_x F(\omega^*, \mu^*, 0)$ and the codimension of the image of $D_x F(\omega^*, \mu^*, 0)$ are 2, together with other condition, assures that there occurs a bifurcation phenomenon at (ω^*, μ^*).

Being based upon this result, Ambrosetti and Prodi successfully pave the way to deduce the classical Hopf theorem and elucidate its mathematical structure.

Let $f(\mu, x)$ be a given function of $\mathbb{R} \times \mathbb{R}^n$ into $\mathbb{R}^n$. We consider the ordinary differential equation

$$F(\omega, \mu, x(\cdot)) \equiv \omega \frac{dx(\cdot)}{dt} - f(\mu, x(\cdot)) = 0 \tag{1}$$

with a couple (ω, μ) of parameters. Here Ambrosetti and Prodi adopt some suitable function space C^r (resp. C^{r-1}) as $\mathfrak{X}$ (resp. $\mathfrak{Y}$). In order to apply the abstract theorem mentioned above to this concrete problem, we have to calculate the dimension of the kernel of the operator

$$D_x F(\omega^*, \mu^*, 0) : x \mapsto \omega^* \frac{dx}{dt} - D_x f(\mu^*, 0)x \tag{2}$$

of $\mathfrak{X}$ into $\mathfrak{Y}$ and the codimension of its image. Ambrosetti and Prodi's idea to overcome this problem is very simple in a sense. They first expand the function $x(t)$ in the uniformly convergent Fourier series

$$x(t) = \sum_{k=-\infty}^{\infty} u_k e^{ikt}. \tag{3}$$

If the derivative $dx/dt = \dot{x}(\cdot)$ can be expanded in the form

$$\dot{x}(t) = \sum_{k=-\infty}^{\infty} iku_k e^{ikt}, \tag{4}$$

we obtain

$$D_x F(\omega^*, \mu^*, 0)x = \sum_{k=-\infty}^{\infty} [ik\omega^* I - D_x f(\mu^*, 0)]u_k e^{ikt} \tag{5}$$

by substituting (3) and (4) into (2), where I is the $(n \times n)$-identity matrix. Hence the kernel of $D_x F(\omega^*, \mu^*, 0)$ consists of all the $x(\cdot)$ such that

$[ik\omega^* I - D_x f(\mu^*, 0)]u_k = 0$ for all $k \in \mathbb{Z}$ (the set of all the integers). And when we wish to determine the image of $D_x F(\omega^*, \mu^*, 0)$, we have only to examine the equation

$$[ik\omega^* I - D_x f(\mu^*, 0)]u_k = v_k \quad \text{for all} \quad k \in \mathbb{Z}, \tag{6}$$

where v_k's are the Fourier coefficients of $y \in \mathfrak{Y}$. The image of $D_x F(\omega^*, \mu^*, 0)$ is the set of all the elements y of $\mathfrak{Y}$, for which the equation (6) is solvable.

The main purpose of the present paper is to establish the Hopf theorem in the framework of some Sobolev space instead of C^r. This approach seems to enable us to simplify the technical details in the course of the proof to some extent. The basic result due to Carleson [2] and Hunt [7] plays a crucial role in our theory.

Incidentally we have to note that we require the condition that $x(\cdot)$ is of the class C^r, $r \geqq 3$ and $y(\cdot)$ is of the class C^{r-1} in order to justify the expression (4) if we remain in the space C^r. As will be shown later, we have, for any $\varepsilon > 0$, that

$$\begin{aligned}\left\| \sum_{k=-\infty}^{\infty} iku_k e^{ikt} \right\| &\leqq \sum_{k=-\infty}^{\infty} \|ku_k\| \\ &\leqq \sum_{|k|<N} \|ku_k\| + \varepsilon \sum_{|k|\geqq N} \frac{1}{|k|^{r-1}}\end{aligned}$$

for sufficiently large N. Hence if we assume $r \geqq 3$, the left-hand side is uniformly convergent and the expression (4) is valid. Thus although Ambrosetti and Prodi assume only $r = 1$, we have to impose additional restrictions on the smoothness of $x(\cdot)$ and $y(\cdot)$, and some more subtle and careful account of the magnitudes of the Fourier coefficients seems to be required.

The another object of this paper is to fortify these analytical details and to complete the Ambrosetti and Prodi theory from the standpoint of the classical Fourier analysis.

2. Abstract Hopf bifurcation theorem

Let $\mathfrak{X}$ and $\mathfrak{Y}$ be a couple of real Banach spaces. And $F(\omega, \mu, x)$ is assumed to be a function of the class $C^2(\mathbb{R}^2 \times \mathfrak{X}, \mathfrak{Y})$ which satisfies

$$F(\omega, \mu, 0) = 0 \quad \text{for all} \quad (\omega, \mu) \in \mathbb{R}^2.$$

A point $(\omega^*, \mu^*) \in \mathbb{R}^2$ is called a *bifurcation point* of F if $(\omega^*, \mu^*, 0)$ is in the closure of the set

$$S = \{(\omega, \mu, x) \in \mathbb{R}^2 \times \mathfrak{X} \mid x \neq 0, F(\omega, \mu, x) = 0\}. \tag{1}$$

We shall use several notations for the sake of simplicity.

$$\begin{aligned} T &= D_x F(\omega^*, \mu^*, 0), \\ \mathfrak{V} &= \operatorname{Ker} T, \quad \mathfrak{R} = T(\mathfrak{X}), \\ M &= D^2_{x,\mu} F(\omega^*, \mu^*, 0), \\ N &= D^2_{x,\omega} F(\omega^*, \mu^*, 0). \end{aligned}$$

T is the derivative of F with respect to x at $(\omega^*, \mu^*, 0)$. It is a bounded linear operator of $\mathfrak{X}$ into $\mathfrak{Y}$. $\mathfrak{V}$ and $\mathfrak{R}$ are the kernel and the image of T, respectively. M (resp. N) is the second derivative of F with respect to (x, μ) (resp. (x, ω)) at $(\omega^*, \mu^*, 0)$. Since $D^2_{x,\mu} F$ is the bounded linear operator of $\mathbb{R}$ into $\mathfrak{L}(\mathfrak{X}, \mathfrak{Y})$ (the Banach space of all the bounded linear operators of $\mathfrak{X}$ into $\mathfrak{Y}$), it can be identified with an element of $\mathfrak{L}(\mathfrak{X}, \mathfrak{Y})$. The same is true for $D^2_{x,\omega} F$.

The following theorem is an abstract version of the Hopf bifurcation theorem due to Ambrosetti and Prodi [1] (pp. 136–139).

Theorem 1. *Let $\mathfrak{X}$ and $\mathfrak{Y}$ be real Banach spaces. Assume that $F(\omega, \mu, x)$ is a function of the class $C^2(\mathbb{R}^2 \times \mathfrak{X}, \mathfrak{Y})$ which satisfies the following two conditions.*

1. $\dim \mathfrak{V} = 2$. $\mathfrak{R}$ *is closed and* $\operatorname{codim} \mathfrak{R} = 2$.
   We represent $\mathfrak{X}$and $\mathfrak{Y}$ in the forms of topological direct sums:

$$\mathfrak{X} = \mathfrak{V} \oplus \mathfrak{W}, \quad \mathfrak{Y} = \mathfrak{Z} \oplus \mathfrak{R}.$$

   We denote by P the projection of $\mathfrak{Y}$ into $\mathfrak{Z}$, and by Q the projection of $\mathfrak{Y}$ into $\mathfrak{R}$.
2. *There exists some point $v* \in \mathfrak{V}$ such that PMv^* and PNv^* are linearly independent.*
   Then (ω^*, μ^*) is a bifurcation point of F.

The condition that $\dim \mathfrak{V} = 2$ and $\operatorname{codim} \mathfrak{R} = 2$ implies that $T : \mathfrak{X} \to \mathfrak{Y}$ is a Fredholm operator with index zero.

The proof of this theorem is based upon the Ljapunov–Schmidt reduction method, which is also neatly explained in Ambrosetti and Prodi [1] (pp. 89–91).

3. Classical Hopf bifurcation for ordinary differential equations

We now turn to the classical bifurcation phenomena of periodic solutions for some ordinary differential equation. Let $f(\mu, x)$ be a function of the class $C^2(\mathbb{R} \times \mathbb{R}^n, \mathbb{R}^n)$. And consider the differential equation

$$\frac{dx}{ds} = f(\mu, x). \tag{1}$$

Changing the time variable s by the relation

$$t = \omega s \quad (\omega \neq 0), \tag{2}$$

we rewrite the equation (1) as

$$\frac{dx}{dt} = \frac{1}{\omega} f(\mu, x), \tag{3}$$

$$i.e. \quad \omega \frac{dx}{dt} = f(\mu, x). \tag{3'}$$

This is an ordinary differential equation with two real parameters, ω and μ. For the sake of simplicity, we assume that the function $f(\mu, x)$ satisfies

$$f(\mu, 0) = 0 \quad \text{for all} \quad \mu \in \mathbb{R}. \tag{4}$$

We denote by $\mathfrak{W}^{1,2}_{2\pi}(\mathbb{R}, \mathbb{R}^n)$ the set of all the 2π-periodic and absolutely continuous functions $x : \mathbb{R} \to \mathbb{R}^n$ such that $\dot{x}|_{[0,2\pi]} \in \mathfrak{L}^2([0, 2\pi], \mathbb{R}^n)$, where $\dot{x}|_{[0,2\pi]}$ denotes the restriction of $\dot{x} = dx/dt$ to the interval $[0, 2\pi]$; *i.e.*

$$\mathfrak{W}^{1,2}_{2\pi} = \{x : \mathbb{R} \to \mathbb{R}^n | x \text{ is } 2\pi\text{-periodic, absolutely continuous and } \dot{x}(\cdot)|_{[0,2\pi]} \in \mathfrak{L}^2([0, 2\pi], \mathbb{R}^n)\}. \tag{5}$$

$\mathfrak{W}^{1,2}_{2\pi}$ is a Banach space under the norm

$$\|x\|_{\mathfrak{W}^{1,2}_{2\pi}} = \left(\int_0^{2\pi} \|x(t)\|^2 dt\right)^{1/2} + \left(\int_0^{2\pi} \|\dot{x}(t)\|^2 dt\right)^{1/2}. \tag{6}$$

We also denote by $\mathfrak{L}^2_{2\pi}(\mathbb{R}, \mathbb{R}^n)$ the set of all the 2π-periodic measurable functions $y : \mathbb{R} \to \mathbb{R}^n$ such that $y|_{[0,2\pi]} \in \mathfrak{L}^2([0, 2\pi], \mathbb{R}^n)$; *i.e.*

$$\mathfrak{L}^2_{2\pi} = \{y : \mathbb{R} \to \mathbb{R}^n | y \text{ is } 2\pi\text{-periodic and } y|_{[0,2\pi]} \in \mathfrak{L}^2([0, 2\pi], \mathbb{R}^n)\}. \tag{7}$$

$\mathfrak{L}^2_{2\pi}$ is a Banach space under the norm

$$\|y\|_{\mathfrak{L}^2_{2\pi}} = \left(\int_0^{2\pi} \|y(t)\|^2 dt\right)^{1/2}. \tag{8}$$

In this section, we adopt $\mathfrak{W}^{1,2}_{2\pi}$ as $\mathfrak{X}$ and $\mathfrak{L}^2_{2\pi}$ as $\mathfrak{Y}$, respectively; *i.e.*

$$\mathfrak{X} = \mathfrak{W}^{1,2}_{2\pi}, \quad \mathfrak{Y} = \mathfrak{L}^2_{2\pi}. \tag{9}$$

Define the function $F : \mathbb{R}^2 \times \mathfrak{X} \to \mathfrak{Y}$ by (cf. (9))

$$F(\omega, \mu, x) = \omega \frac{dx}{dt} - f(\mu, x). \tag{10}$$

Then we can prove that F is a function of the class $C^2(\mathbb{R}^2 \times \mathfrak{X}, \mathfrak{Y})$ provided that the following assumptions are satisfied.

Assumption 1. (i) There exists some constants α and $\beta \in \mathbb{R}$ such that

$$\|f(\mu, x)\| \leqq \alpha + \beta\|x\| \quad \text{for all} \quad x \in \mathbb{R}^n.$$

(ii) There exists some constant ρ such that

$$\|D_x f(\mu, x)\|, \quad \|D^2 f(\mu, x)\| \leqq \rho \quad \text{for all} \quad x \in \mathbb{R}^n.$$

The proof of the fact $F(\cdot) \in C^2(\mathbb{R}^2 \times \mathfrak{X}, \mathfrak{Y})$ will be given in the next section.

It is obvious that

$$F(\omega, \mu, 0) = 0 \quad \text{for all} \quad (\omega, \mu) \in \mathbb{R}^2. \tag{11}$$

$(\omega^*, \mu^*) \in \mathbb{R}^2$ is called a *bifurcation point* of F if there exists a sequence (ω_n, μ_n, x_n) in $\mathbb{R}^2 \times \mathfrak{X}$ such that

$$\begin{cases} F(\omega_n, \mu_n, x_n) = 0, \\ (\omega_n, \mu_n) \to (\omega^*, \mu^*) \quad \text{as} \quad n \to \infty, \quad \text{and} \\ x_n \neq 0, \quad x_n \to 0 \quad \text{as} \quad n \to \infty. \end{cases}$$

Each x_n is a non-trivial (not identically zero) periodic solution with period 2π of the equation

$$\omega_n \frac{dx}{dt} - f(\mu_n, x) = 0.$$

Hence, changing the time-variable to s again, we obtain a periodic solution

$$X_n(s) = x_n(\omega_n s)$$

with period $\tau_n = 2\pi/\omega_n$ for the original equation (1). Consequently we have that

$$\tau_n \to \tau^* = \frac{2\pi}{\omega^*}, \quad \| X_n \|_{\mathfrak{W}^{1,2}_{2\pi/\omega_n}} \to 0 \quad \text{as} \quad n \to \infty,$$

provided that $\omega^* \neq 0$. The target of our investigations is to find out a bifurcation point of F according to the principle of Theorem 1. We have to note that the derivative

$$D_x F(\omega, \mu, 0) : \; x \mapsto \omega \dot{x} - D_x f(\mu, 0)x \tag{12}$$

is to play the most important role in the course of our discussions. ($\dot{x}$ means dx/dt.) Of course, $D_x F(\omega, \mu, 0)$ is a bounded linear operator of $\mathfrak{X}$ into $\mathfrak{Y}$. If we denote

$$A_\mu = D_x f(\mu, 0),$$

A_μ is an $(n \times n)$-matrix and (12) can be rewritten in the form

$$D_x F(\omega, \mu, 0)x = \omega \dot{x} - A_\mu x. \tag{12'}$$

Here we need a couple of assumptions to be imposed upon the matrix A_μ at some (ω^*, μ^*).

Assumption 2. *A_{μ^*} is regular, and $\pm i\omega^* (\omega^* > 0)$ are simple eigenvalues of A_{μ^*}.*

Assumption 3. *None of $\pm ik\omega^*$ $(k \neq \pm 1)$ is an eigenvalue of A_{μ^*}.*

4. Smoothness of *F*

In this section, we examine the differentiability of the so called Nemyckii operator.[1]

Let us define the operator Φ on $\mathfrak{W}^{1,2}_{2\pi}$ by

$$\Phi(x(\cdot)) = f(\mu, x(\cdot)), \quad x \in \mathfrak{W}^{1,2}_{2\pi} \tag{1}$$

for any fixed μ. If Assumption 1(i) is satisfied, then $\Phi(x(\cdot))$ is in $\mathfrak{L}^2_{2\pi}$.

Lemma 1. *Under Assumption 1(i), the operator $\Phi : \mathfrak{W}^{1,2}_{2\pi} \to \mathfrak{L}^2_{2\pi}$ is continuous.*

[1] Related topics are discussed in Ambrosetti–Prodi [1] pp. 17–21.

Proof. Assume that $x_n(\cdot) \to x_0(\cdot)$ (as $n \to \infty$) in $\mathfrak{W}^{1,2}_{2\pi}$. It obviously implies that $x_n(\cdot) \to x_0(\cdot)$ (as $n \to \infty$) in $\mathfrak{L}^2_{2\pi}$.

Then there exists some subsequence $\{x_{n'}(\cdot)\}$ and a function $\varphi(\cdot) \in \mathfrak{L}^2([0, 2\pi], \mathbb{R})$ such that

$$x_{n'}(t) \to x_0(t) \quad \text{a.e. on} \quad [0, 2\pi], \text{ and} \tag{2}$$

$$\|x_{n'}(t)\| \leqq \varphi(t) \quad \text{a.e. on} \quad [0, 2\pi]. \tag{3}$$

Since f is C^2 by assumption, we must have

$$\Phi(x_{n'}(\cdot)) = f(\mu, x_{n'}(\cdot)) \to \Phi(x_0(\cdot)) = f(\mu, x_0(\cdot)) \quad \text{a.e.} \tag{4}$$

Taking account of Assumption 1(i), we obtain

$$\|f(\mu, x_{n'}(\cdot))\| \leqq \alpha + \beta\|x_{n'}(\cdot)\| \leqq \alpha + \beta\varphi(\cdot) \quad \text{a.e.} \tag{5}$$

Now (4) and (5) imply that

$$\begin{aligned}\|\Phi(x_{n'}(\cdot)) - \Phi(x_0(\cdot))\|^2_{\mathfrak{L}^2_{2\pi}} &= \int_0^{2\pi} \|f(\mu, x_{n'}(t)) - f(\mu, x_0(t))\|^2 dt \\ &\to 0 \quad \text{as} \quad n' \to \infty\end{aligned} \tag{6}$$

by the dominated convergence theorem.

If follows that $\Phi(x_n(\cdot)) \to \Phi(x_0(\cdot))$ in $\mathfrak{L}^2_{2\pi}$. (If not, a contradiction would occur to the above argument.) Q.E.D.

Lemma 2. *If Assumption 1 is satisfied, then the operator* $\Phi : \mathfrak{W}^{1,2}_{2\pi} \to \mathfrak{L}^2_{2\pi}$ *is twice continuously differentiable.*

Proof. Let us begin by evaluating the first variation of Φ. For any $x, z \in \mathfrak{W}^{1,2}_{2\pi}$, we have

$$\begin{aligned}&\frac{1}{\lambda}[\Phi(x(\cdot) + \lambda z(\cdot)) - \Phi(x(\cdot))](t) \\ &\quad = \frac{1}{\lambda}[f(\mu, x(t) + \lambda z(t)) - f(\mu, x(t))] \\ &\quad = \frac{1}{\lambda}[D_x f(\mu, x(t))\lambda z(t) + o(\lambda z(t))] \\ &\quad \to D_x f(\mu, x(t))z(t) \quad \text{as} \quad \lambda \to 0,\end{aligned} \tag{7}$$

where we must have

$$D_x f(\mu, x(\cdot))z(\cdot) \in \mathfrak{L}^2_{2\pi} \quad \text{for any} \quad x(\cdot), z(\cdot) \in \mathfrak{W}^{1,2}_{2\pi} \tag{8}$$

by Assumption 1 and $z(\cdot) \in \mathfrak{W}^{1,2}_{2\pi}$.

We can also confirm that, for any $\varepsilon > 0$,

$$\begin{aligned} &\|\frac{1}{\lambda}[\Phi(x(\cdot)+\lambda z(\cdot)) - \Phi(x(\cdot))](t) - D_x f(\mu, x(t))z(t)\|^2 \\ &= \|D_x f(\mu, x(t))z(t) + \frac{o(\lambda z(t))}{\lambda} - D_x f(\mu, x(t))z(t)\|^2 \\ &\leqq \varepsilon^2 \cdot \|z(t)\|^2 \qquad \text{for sufficiently small } \lambda. \end{aligned} \tag{9}$$

Hence we obtain, by (7) and the dominated convergence theorem, that

$$\frac{1}{\lambda}[\Phi(x(\cdot)+\lambda z(\cdot)) - \Phi(x(\cdot))] \to D_x f(\mu, x(\cdot))z(\cdot) \quad \text{in} \quad \mathfrak{L}^2_{2\pi} \quad \text{as} \quad \lambda \to 0. \tag{10}$$

Therefore Φ has the first variation of the form (8).

It is easy to check that the mapping

$$z(\cdot) \mapsto D_x f(\mu, x(\cdot))z(\cdot)$$

is a bounded linear operator of $\mathfrak{W}^{1,2}_{2\pi}$ into $\mathfrak{L}^2_{2\pi}$. Hence Φ is Gâteaux-differentiable.

Furthermore Φ turns out to be Fréchet-differentiable[2] since

$$x(\cdot) \mapsto D_x f(\mu, x(\cdot)) \tag{11}$$

is a continuous mapping of $\mathfrak{W}^{1,2}_{2\pi}$ into $\mathfrak{L}(\mathfrak{W}^{1,2}_{2\pi}, \mathfrak{L}^2_{2\pi})$ (the space of bounded linear operators of $\mathfrak{W}^{1,2}_{2\pi}$ into $\mathfrak{L}^2_{2\pi}$).[3]

Finally we can prove that Φ is twice continuously differentiable by a similar method as above. So we omit the details. Q.E.D.

Denoting $T = D_x F(\omega^*, \mu^*, 0)$, we have

$$Tx = 0 \quad \text{if and only if} \quad \omega^* \dot{x} - A_{\mu^*} x = 0. \tag{12}$$

In order to apply Theorem 1 to our classical problem in Sect. 3, we have to start with confirming that (a) the dimension of the kernel of T is 2, and

---

[2] Let $\mathfrak{V}$ and $\mathfrak{W}$ be a couple of Banach spaces. Assume that a function φ of an open subset U of $\mathfrak{V}$ into $\mathfrak{W}$ is Gâteaux-differentiable in a neighborhood V of $x \in U$. We denote by $\delta\varphi(v)$ the Gâteaux-derivative of φ at v. If the function $v \mapsto \delta\varphi(v)(V \to \mathfrak{L}(\mathfrak{V}, \mathfrak{W}))$ is continuous, then φ is Fréchet-differentiable.

[3] The continuity of the mapping (11) can be proved in the same manner as in the proof of Lemma 1. Assumption 1(i) is used again for the dominated convergence argument.

(b) the codimension of the image of T is also 2. For brevity, we have to show that

$$\dim \operatorname{Ker} T = 2, \quad \text{and}$$
$$\operatorname{codim} T(\mathfrak{X}) = 2.$$

Henceforth, we denote Ker T by $\mathfrak{V}$ and $T(\mathfrak{X})$ by $\mathfrak{R}$.

Remark. Actually we can confirm a priori that T is a Fredholm operator with index zero when $\omega^* = 0$. Hence we do not have to check both of $\dim\mathfrak{V} = 2$ and $\operatorname{codim}\mathfrak{R} = 2$, because either one follows from the other automatically.

The operator $T : x \mapsto \omega^* \dot{x} - A_{\mu^*} x \quad (\mathfrak{W}^{1,2}_{2\pi} \to \mathfrak{L}^2_{2\pi})$ can be rewritten as

$$\begin{aligned} T_x &= \omega^* \dot{x} - A_{\mu^*} x \\ &= \omega^* \dot{x} + x - x - A_{\mu^*} x \\ &= \omega^* \dot{x} + x - (I + A_{\mu^*}) x. \end{aligned}$$

Since the mapping $x \mapsto \omega^* \dot{x} + x$ $(\omega^* \neq 0)$ is an isomorphism between $\mathfrak{W}^{1,2}_{2\pi}$ and $\mathfrak{L}^2_{2\pi}$, it is a Fredholm operator with index zero. We also know that the inclusion mapping of $\mathfrak{W}^{1,2}_{2\pi}$ into $\mathfrak{L}^2_{2\pi}$ is a compact operator.[4] Consequently, the mapping $x \mapsto (I + A_{\mu^*})x$ is also a compact operator of $\mathfrak{W}^{1,2}_{2\pi}$ into $\mathfrak{L}^2_{2\pi}$. (Note that A_{μ^*} is a bounded operator of $\mathfrak{L}^2_{2\pi}$ into itself.) Thus we confirmed that the operator T can be expressed as a sum of a Fredholm operator with index zero and a compact operator. It follows from a well-known theorem[5] that T is also a Fredholm operator with index zero.

However we prefer an elementary way to prove both of $\dim\mathfrak{V} = 2$ and $\operatorname{codim}\mathfrak{R} = 2$ without having recourse to the Fredholm operator theory discussed above.

I appreciate Professor S. Kusuoka's suggestion on this point.

5. $\dim\mathfrak{V} = 2$

Expanding $x \in \mathfrak{X}$ in the Fourier series, we obtain

$$x(t) = \sum_{k=-\infty}^{\infty} u_k e^{ikt}, \quad u_k \in \mathbb{C}^n, \tag{1}$$

[4] This is a special case of the Rellich–Kondrachov compactness theorem. Evans [4] pp. 272–274.

[5] See, for instance, Zeidler [11] pp. 300–301.

where u_k is a vector, the j-th coordinate of which is given by

$$\frac{1}{\sqrt{2\pi}}\int_0^{2\pi} x_j(t)e^{ikt}dt, \quad j = 1, 2, \ldots, n.$$

We must have the relation

$$u_{-k} = \bar{u}_k \quad \text{(conjugate)}, \quad k \in \mathbb{Z}, \tag{2}$$

since $x(t)$ is a real vector. Furthermore we have to keep in mind that the series (1) is uniformly convergent.[6]

The k-th Fourier coefficient of $\dot{x}(\cdot)$ is given by iku_k. Since $x(\cdot) \in \mathfrak{W}^{1,2}_{2\pi}$, it is clear that $\dot{x}(\cdot) \in \mathfrak{L}^2_{2\pi}$. Hence we obtain

$$\dot{x}(t) = \sum_{k=-\infty}^{\infty} iku_k e^{ikt} \quad \text{a.e.}, \tag{3}$$

that is, the Fourier series of $\dot{x}(\cdot)$ given by the right-had side of (3) converges a.e. and equal to $\dot{x}(\cdot)$. This result is justified by the Carleson–Hunt theorem ([2], [7]). It follows that

$$\omega^*\dot{x} - A_{\mu^*}x = \sum_{k=-\infty}^{\infty} [ik\omega^* I - A_{\mu^*}]u_k e^{ikt} = 0 \quad \text{a.e.} \tag{4}$$

By the uniqueness of the Fourier coefficients (with respect to the complete orthonormal system $(1/\sqrt{2\pi})e^{ikt}$; $k = 0, \pm 1, \pm 2, \ldots$), we must have

$$[ik\omega^* I - A_{\mu^*}]u_k = 0 \quad \text{for all} \quad k \in \mathbb{Z}. \tag{5}$$

It is enough to find out all $x \in \mathfrak{X}$, the Fourier coefficients of which satisfy (5), in order to determine $\mathfrak{V}$.

By the regularity of A_{μ^*} in Assumptions 2 and 3,

$$ik\omega^* I - A_{\mu^*}, \quad k \neq \pm 1$$

are all invertible. Therefore we have simply

$$u_k = 0 \quad \text{for} \quad k \neq \pm 1. \tag{6}$$

---

[6] If a 2π-periodic function $\varphi : \mathbb{R} \to \mathbb{R}$ is absolutely continuous and its derivative φ' belongs to $\mathfrak{L}^2([0, 2\pi], \mathbb{R})$, then the Fourier series of φ uniformly converges to φ on $\mathbb{R}$. cf. Katznelson [8] Theorem 6.2, pp. 33–34. The k-th Fourier coefficient of φ' is given by $ik\hat{\varphi}(k)$, where $\hat{\varphi}(k)$ is the k-th Fourier coefficient of φ.

Thus all the coefficients in (1) besides $k = \pm 1$ must be zero. Consequently all we have to do is to find out functions whose Fourier coefficients $u_{\pm 1}$ satisfy

$$[\pm i\omega^* I - A_{\mu^*}]u_{\pm 1} = 0. \tag{7}$$

Since $\pm i\omega^*$ are simple eigenvalues of A_{μ^*} by Assumption 2, there exists $\xi \in \mathbb{C}^n$ $(\xi \neq 0)$ such that[7]

$$\text{Ker}\{i\omega^* I - A_{\mu^*}\} = \text{span}\{\xi\}. \tag{8}$$

On the other hand,

$$\text{Ker}\{-i\omega^* I - A_{\mu^*}\} = \text{span}\{\bar{\xi}\}. \tag{9}$$

The Fourier coefficients $u_{\pm 1}$ which satisfy (7) can be expressed as

$$u_{+1} = a\xi, \quad u_{-1} = b\bar{\xi} \quad (a, b \in \mathbb{C}). \tag{10}$$

Then any real solution $x(t)$ of (5) is of the form:

$$x(t) = a\xi e^{it} + \bar{a}\bar{\xi}e^{-it}. \tag{11}$$

If we put $a = \alpha + i\beta$, $\xi = \gamma + i\delta$ $(\alpha, \beta \in \mathbb{R};\ \gamma, \delta \in \mathbb{R}^n)$, it follows from simple calculations that

$$\begin{aligned} x(t) &= 2Re[a\xi e^{it}] \\ &= 2[\alpha(\gamma \cos t - \delta \sin t) - \beta(\gamma \sin t + \delta \cos t)]. \end{aligned}$$

Denoting $p(t) = \gamma \cos t - \delta \sin t$ and $q(t) = \gamma \sin t + \delta \cos t$, we have

$$x(t) = 2\alpha p(t) - 2\beta q(t), \tag{12}$$

where α and β are any real numbers. It can easily be checked that $p(t)$ and $q(t)$ are linearly independent.[8]

Thus we have shown that $p(\cdot)$ and $q(\cdot)$ form a basis of $\mathfrak{V}$. And so $\dim\mathfrak{V} = 2$.

[7] span$\{\xi\}$ denotes the subspace of $\mathbb{C}^n$ spanned by ξ.

[8] Put $\mu p(t) + \nu q(t) = \mu(\gamma \cos t - \delta \sin t) + \nu(\gamma \sin t + \delta \cos t) = (\mu\gamma + \nu\delta)\cos t + (\nu\gamma - \mu\delta)\sin t = 0$. Then we have

$$\begin{cases} \mu\gamma + \nu\delta = 0, \\ \nu\gamma - \mu\delta = 0. \end{cases}$$

It follows that

$$\begin{cases} \mu\nu\gamma + \nu^2\delta = 0, \\ \mu\nu\gamma - \mu^2\delta = 0. \end{cases}$$

Hence $(\nu^2 + \mu^2)\delta = 0$. If $\mu \neq 0$ or $\nu \neq 0$, δ must be zero. And so $\xi = \gamma$, that is $\xi = \bar{\xi}$ (real vector). Thus we get a contradiction.

6. codim$\mathfrak{R} = 2$

We now proceed to examining the dimension of the quotient space $\mathfrak{Y}/\mathfrak{R}$ of $\mathfrak{Y}$ modulo $\mathfrak{R}$.

Writing

$$Tx = y, \quad x \in \mathfrak{X}, \quad y \in \mathfrak{Y},$$

we again expand x and y in the Fourier series. Let u_k (resp. v_k) be the Fourier coefficients of x (resp. y). Then we must have

$$\sum_{k=-\infty}^{\infty} [ik\omega^* I - A_{\mu^*}] u_k e^{ikt} = \sum_{k=-\infty}^{\infty} v_k e^{ikt}. \tag{1}$$

Since $x \in \mathfrak{W}^{1,2}_{2\pi}$, $y \in \mathfrak{L}^2_{2\pi}$ and both of them are 2π-periodic, the right-hand side of (1) converges a.e. again by the Carleson–Hunt theorem. By the uniqueness of the Fourier coefficients, we must have

$$[ik\omega^* I - A_{\mu^*}] u_k = v_k \quad \text{for all} \quad k \in \mathbb{Z}. \tag{2}$$

By the regularity of A_{μ^*} in Assumptions 2 and 3, u_k $(k \neq \pm 1)$ in (2) can be solved uniquely in the form:

$$\begin{aligned} u_0 &= -A_{\mu^*}^{-1} v_0, \\ u_k &= [ik\omega^* I - A_{\mu^*}]^{-1} v_k, \quad k \neq 0, \pm 1. \end{aligned} \tag{3}$$

Since the norm of $(1/ik\omega^*)A_{\mu^*}$ is less than 1 for sufficiently large $|k|$'s (say $|k| \geqq k_0$), we have

$$\begin{aligned} [ik\omega^* I - A_{\mu^*}]^{-1} &= \frac{1}{ik\omega^*}\left[I - \frac{1}{ik\omega^*}A_{\mu^*}\right]^{-1} \\ &= \frac{1}{ik\omega^*}\left[I + \frac{1}{ik\omega^*}A_{\mu^*} + \frac{1}{(ik\omega^*)^2}A_{\mu^*}^2 + \dots\right] \\ &= \frac{1}{ik\omega^*}I + O\left(\frac{1}{k^2}\right) \quad \text{for} \quad |k| \geqq k_0. \end{aligned} \tag{4}$$

It follows, from (3) and (4), that

$$\begin{aligned} u_k &= [ik\omega^* I - A_{\mu^*}]^{-1} v_k \\ &= \frac{1}{ik\omega^*} v_k + O\Big(\frac{1}{k^2}\Big) v_k \quad \text{for} \quad |k| \geqq k_0. \end{aligned}$$

Consequently,

$$\sum_{k\neq 0,\pm 1} u_k e^{ikt} = \sum_{\substack{|k|<k_0 \\ k\neq 0,\pm 1}} [ik\omega^* I - A_{\mu^*}]^{-1} v_k e^{ikt} + \sum_{|k|\geqq k_0} \frac{v_k}{ik\omega^*} e^{ikt} + \sum_{|k|\geqq k_0} O\left(\frac{1}{k^2}\right) v_k e^{ikt}. \tag{5}$$

We claim that the second term of the right-hand side of (5) is of the class $\mathfrak{W}^{1,2}_{2\pi}$. In order to prove it, we define the function $\theta(t)$ by

$$\theta(t) = \sum_{|k|\geqq k_0} v_k e^{ikt}. \tag{6}$$

Then $\theta(t)$ is of the class $\mathfrak{L}^2_{2\pi}$. Let $\Theta(t)$ be the indefinite integral of $\theta(t)$, that is

$$\Theta(t) = \int_0^t \theta(\tau) d\tau. \tag{7}$$

Clearly $\Theta(t)$ is of the class $\mathfrak{W}^{1,2}_{2\pi}$ and $\dot{\Theta}(t) = \theta(t)$ a.e. Expanding $\theta(t)$ in the Fourier series, we confirm that the coefficient $\hat{\theta}(0)$ corresponding to $k = 0$ is zero. Hence we obtain by the classical result in Fourier Analysis[9] that

$$\hat{\Theta}(k) = \frac{\hat{\theta}(k)}{ik} \quad \text{for all} \quad k \neq 0,$$

$$i.e. \quad \frac{1}{\omega^*}\Theta(t) = \sum_{|k|\geqq k_0} \frac{v_k}{ik\omega^*} e^{ikt}$$

$$= \text{the second term of (5).}$$

This is of the class $\mathfrak{W}^{1,2}_{2\pi}$ by (7).

---

[9] Let $\varphi : \mathbb{R} \to \mathbb{R}$ (we may replace $\mathbb{R}$ by $\mathbb{R}^n$) be a 2π-periodic function which is integrable on $[-\pi.\pi]$. Furthermore we assume $\hat{f}(0) = 0$ ($\hat{f}(0)$ is the Fourier coefficient corresponding to $k = 0$). If we define

$$\Phi(t) = \int_0^t \varphi(\tau) d\tau,$$

Φ is a 2π-periodic continuous function and

$$\hat{\Phi}(k) = \frac{1}{ik}\hat{f}(k), \quad k \neq 0.$$

See Katznelson [8] Theorem 1.6, p. 4 or Zygmund [12] Vol. 1, p. 42.

We can also prove that the third term of the right-hand side of (5) is of the class $\mathfrak{W}^{1,2}_{2\pi}$ by a similar argument as above. In this case we define the function $\theta(t) \in \mathfrak{L}^2_{2\pi}$ by

$$\theta(t) = \sum_{|k| \geqq k_0} ikO\left(\frac{1}{k^2}\right) v_k \cdot e^{ikt} \tag{8}$$

instead of (6). $\Theta(t)$ is the indefinite integral of $\theta(t)$ as in (7).

The first term of the right-hand side of (5) is obviously of the class $\mathfrak{W}^{1,2}_{2\pi}$. This finishes the proof of the claim that the right-hand side of (5) is of the class $\mathfrak{W}^{1,2}_{2\pi}$.

Thus denoting the Fourier coefficients of any $y \in \mathfrak{Y}$ by v_k's, the function

$$\sum_{k \neq \pm 1} u_k e^{ikt} = -A^{-1}_{\mu^*} v_0 + \sum_{k \neq 0, \pm 1} [ik\omega^* I - A_{\mu^*}]^{-1} v_k e^{ikt} \tag{9}$$

is of the class $\mathfrak{W}^{1,2}_{2\pi}$ and u_k's defined here satisfy the relation (1).

We shall now go over to $k = \pm 1$. The equation

$$[\pm i\omega^* I - A_{\mu^*}] u_{\pm 1} = v_{\pm 1} \tag{10}$$

does or does not have a solution. Since

$$\operatorname{codim}[\pm i\omega^* I - A_{\mu^*}](\mathbb{C}^n) = 1$$

by Assumption 2, there must exist some $\varphi \in \mathbb{C}^n$ $(\varphi \neq 0)$ such that

$$\mathbb{C}^n / [i\omega^* I - A_{\mu^*}](\mathbb{C}^n) = \operatorname{span}\{\varphi + [i\omega^* I - A_{\mu^*}](\mathbb{C}^n)\}. \tag{11}$$

On the other hand, we also have

$$\mathbb{C}^n / [-i\omega^* I - A_{\mu^*}](\mathbb{C}^n) = \operatorname{span}\{\bar{\varphi} + [-i\omega^* I - A_{\mu^*}](\mathbb{C}^n)\}. \tag{12}$$

An element of $\mathfrak{Y}$ is not contained in $\mathfrak{R}$ if and only if its Fourier coefficients $v_{\pm 1}$ (corresponding to $k = \pm 1$) do not admit the existence of $u_{\pm 1}$ which satisfy (10). Such vector v_1 (resp. v_{-1}) is contained in the equivalence class $\varphi + [i\omega^* I - A_{\mu^*}](\mathbb{C}^n)$ (resp. $\bar{\varphi} + [-i\omega^* I - A_{\mu^*}](\mathbb{C}^n)$). Therefore any element of $\mathfrak{Y}/\mathfrak{R}$ can be expressed as

$$a\varphi e^{it} + b\bar{\varphi} e^{-it} + \mathfrak{R}; \quad a, b \in \mathbb{C}.$$

In order to find out a real solution, we should put $b = \bar{a}$. If we write $\varphi = \gamma + i\delta$, $a = \alpha + i\beta$ $(\alpha, \beta \in \mathbb{R};\ \gamma, \delta \in \mathbb{R}^n)$, we obtain (in the same manner as the calculations on page 52)

$$\begin{aligned} a\varphi e^{it} + b\bar{\varphi} e^{-it} &= 2[\alpha(\gamma \cos t - \delta \sin t) - \beta(\gamma \sin t + \delta \cos t)] \\ &= 2\alpha p(t) - 2\beta q(t), \end{aligned}$$

where $p(t) = \gamma \cos t - \delta \sin t$ and $q(t) = \gamma \sin t + \delta \cos t$. Thus we have confirmed that any element of $\mathfrak{Y}/\mathfrak{R}$ can be expressed as a linear combination of the two linearly independent elements, $p(t) + \mathfrak{R}$ and $q(t) + \mathfrak{R}$. Hence $\mathrm{codim}\mathfrak{R} = 2$.

Remark 1. We can specify $\varphi = \xi$ (see p. 52 for the definition of ξ). Hence we will henceforth choose ξ as φ.

Remark 2. We can represent $\mathfrak{X} = \mathfrak{W}^{1,2}_{2\pi}$ and $\mathfrak{Y} = \mathfrak{L}^2_{2\pi}$ in the form of topological direct sums:

$$\mathfrak{X} = \mathfrak{V} \oplus \mathfrak{W}, \quad \mathfrak{Y} = \mathfrak{Z} \oplus \mathfrak{R},$$

by choosing suitable subspaces $\mathfrak{W} \subset \mathfrak{X}$ and $\mathfrak{Z} \subset \mathfrak{Y}$, respectively. Of course, $\mathfrak{V} = \mathrm{Ker}T$ and $\mathfrak{R} = T(\mathfrak{X})$ as stated on page 50. We also denote by P the projection of $\mathfrak{Y}$ into $\mathfrak{Z}$ corresponding to the direct product defined above.

7. Linear independence of *PMv** and *PNv** (I)

According to the Assumption 2, $\pm i\omega^*$ are simple eigenvalues of A_{μ^*}. Hence $\mathbb{C}^n$ can be expressed as a direct sum

$$\mathbb{C}^n = \mathrm{Ker}[\pm i\omega^* I - A_{\mu^*}] \oplus [\pm i\omega^* I - A_{\mu^*}](\mathbb{C}^n). \tag{1}$$

We shall now concentrate on the case $+i\omega^*$. (The case $-i\omega^*$ can be discussed similarly.)

Let $\eta \in \mathbb{C}^n$ $(\eta \neq 0)$ be any vector which is orthogonal to $[i\omega^* I - A_{\mu^*}](\mathbb{C}^n)$. And define a function $g : \mathbb{R} \times \mathbb{C} \times \mathbb{C}^n \to \mathbb{C}^n \times \mathbb{C}$ by

$$g(\mu, \lambda, \theta) = \begin{pmatrix} (\lambda I - A_\mu)(\xi + \theta) \\ \langle \eta, \theta \rangle \end{pmatrix}. \tag{2}$$

($\langle \cdot, \cdot \rangle$ denotes the inner product, that is $\langle \eta, \theta \rangle = \sum_{j=1}^n \eta_j \bar{\theta}_j$.) Be sure again that the vector ξ is defined by (8) on page 52 (see also Fig. 1). Then the function g is of the class C^1 and satisfies

$$g(\mu^*, i\omega^*, 0) = 0. \tag{3}$$

We are now going to solve the equation $g(\mu, \lambda, \theta) = 0$ locally with respect to (λ, θ) in terms of μ in some neighborhood of $(\mu^*, i\omega^*, 0)$. The derivative of g with respect to (λ, θ) is given by

$$D_{(\lambda,\theta)} g(\mu^*, i\omega^*, 0)(\lambda, \theta) = \begin{pmatrix} \lambda\xi + (i\omega^* I - A_{\mu^*})\theta \\ \langle \eta, \theta \rangle \end{pmatrix}. \tag{4}$$

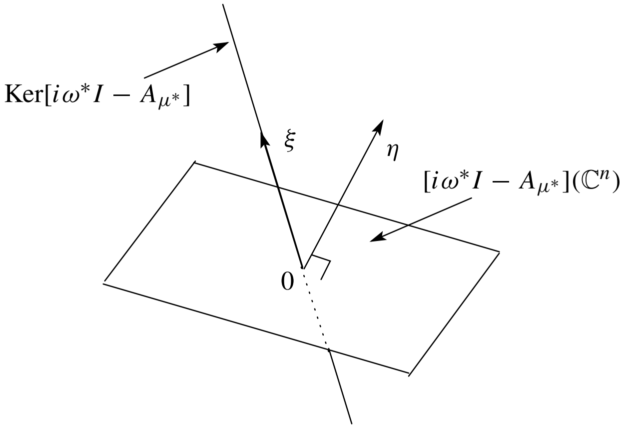

Fig. 1 Geometry of ξ and η

Here $D_{(\lambda,\theta)} g(\mu^*, i\omega^*, 0)$ ($(n+1)\times(n+1)$ matrix) is regular by (1).[10] Applying the implicit function theorem, we obtain the following lemma.

Lemma 3. *There exist a couple of functions, $\lambda(\mu)$ and $\theta(\mu)$ of the C^1-class which are defined in some neighborhood of μ^* and satisfy*

$$\begin{pmatrix} (\lambda(\mu)I - A_\mu)(\xi + \theta(\mu)) \\ \langle \eta, \theta(\mu) \rangle \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{5}$$

and

$$\lambda(\mu^*) = i\omega^*, \quad \theta(\mu^*) = 0. \tag{6}$$

Denoting $\xi + \theta(\mu)$ in (5) by $\xi(\mu)$, we can rewrite the relations (5) and (6) as follows:

$$A_\mu \xi(\mu) = \lambda(\mu)\xi(\mu), \tag{5'}$$

$$\lambda(\mu^*) = i\omega^*, \ \xi(\mu^*) = \xi. \tag{6'}$$

[10] For any $(\alpha_0, \beta_0) \in \mathbb{C}^n \times \mathbb{C}$, there exist some $\lambda_0 \in \mathbb{C}$ and $\gamma_0 \in (i\omega^* I - A_{\mu^*})(\mathbb{C}^n)$ such that $\alpha_0 = \lambda_0 \xi + \gamma_0$. And such λ_0 and γ_0 are unique. Let $(\alpha_0, \beta_0) = (0, 0)$. Then we must have $\lambda_0 = 0$ and $\gamma_0 = 0$. The equation

$$\begin{pmatrix} (i\omega^* I - A_{\mu^*})\theta \\ \langle \eta, \theta \rangle \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

has a unique solution $\theta = 0 (\in \mathbb{C}^n)$ because $\mathrm{Ker}[i\omega^* I - A_{\mu^*}] \cap \mathrm{Ker}\langle \eta, \cdot \rangle = \mathrm{Ker}[i\omega^* I - A_{\mu^*}] \cap [i\omega^* I - A_{\mu^*}](\mathbb{C}^n) = \{0\}$. Thus we conclude that $D_{(\lambda,\theta)} g(\mu^*, i\omega^*, 0)$ is injective.

8. Linear independence of *PMv** and *PNv** (II)

Since $\operatorname{codim}[i\omega^* I - A_{\mu^*}](\mathbb{C}^n) = 1$ by Assumption 2, there exists a nonzero vector $\eta \in \mathbb{C}^n$ such that

$$\langle \eta, \kappa \rangle = 0 \quad \text{for all} \quad \kappa \in [i\omega^* I - A_{\mu^*}](\mathbb{C}^n) \tag{1}$$

as we have seen above. Taking account of the fact[11] that $\xi \notin [i\omega^* I - A_{\mu^*}](\mathbb{C}^n)$, η can be chosen so that

$$\langle \eta, \xi \rangle = 1. \tag{2}$$

The function

$$\Pi : \kappa \mapsto \langle \eta, \kappa \rangle \xi, \ \kappa \in \mathbb{C}^n \tag{3}$$

is called the *spectral projection*[12] associated with ξ. We can similarly define the spectral projection $\overline{\Pi}$ associated with $\bar{\xi}$ by using $\bar{\eta}$ instead of η.

---

[11] $i\omega^*$ is a simple eigenvalue, again by Assumption 2.

[12] Look at the Fig. 2. For the sake of an intuitive exposition, the vectors η, ξ and κ are treated as real vectors. By $\langle \eta, \xi \rangle = \| \eta \| \cdot \| \xi \| \cos\theta = 1$, it follows that $\| \eta \| = 1/ \| \xi \| \cos\theta$. Hence

$$\Pi(\kappa) = \xi \langle \eta, \kappa \rangle = (\| \kappa \| \cos\zeta / \| \xi \| \cos\theta) \cdot \xi.$$

Since $\| \kappa \| \cos\zeta = OA$ and $\| \xi \| \cos\theta = OB$,

$$\Pi(\kappa) = \frac{OA}{OB} \cdot \xi.$$

Here ζ is the angle between η and κ, and θ is the one between ξ and η. κ can be represented uniquely as $\kappa = \alpha\eta + \beta z$ for some $\alpha, \beta \in \mathbb{C}$ and $z \in [i\omega^* I - A_{\mu^*}](\mathbb{C}^n)$. On the other hand, η can be represented uniquely in the form $\eta = a\xi + bz'$ for some $a, b \in \mathbb{C}$ and $z' \in [i\omega^* I - A_{\mu^*}](\mathbb{C}^n)$. Since $\| \eta \|^2 = \langle a\xi + bz', \eta \rangle = a\langle \xi, \eta \rangle + b\langle z', \eta \rangle = a$, it follows that $\eta = \| \eta \|^2 \xi + bz'$. Hence we have

$$\kappa = \alpha\eta + \beta z = \alpha \| \eta \|^2 \xi + (\alpha bz' + \beta z).$$

Furthermore $\Pi(\kappa) = \langle \eta, \xi \rangle \xi = \langle \eta, \alpha \| \eta \|^2 \xi + (\alpha bz'P + \beta z) \rangle \xi = \alpha \| \eta \|^2 \xi$.($\langle \eta, \alpha bz' + \beta z \rangle = 0$ because $\alpha bz' + \beta z \in [i\omega^* I - A_{\mu^*}](\mathbb{C}^n)$.) Thus we obtain

$$\kappa = \Pi(\kappa) + (\alpha bz' + \beta z).$$

This is the direct sum of $\mathbb{C}^n$ corresponding to $\operatorname{span}\{\xi\}$ and $[i\omega^* I - A_{\mu^*}](\mathbb{C}^n)$.

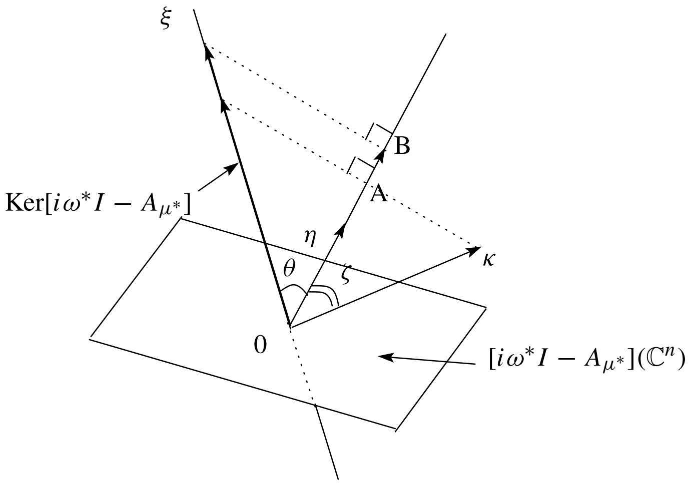

Fig. 2 Spectral projection

Denoting

$$A'_{\mu^*} = \frac{d}{d\mu} A_\mu \Big|_{\mu=\mu^*} \tag{4}$$

for the sake of simplicity, we get the following result.

Lemma 4. $\Pi A'_{\mu^*}\xi = \lambda'(\mu^*)\xi,\ \overline{\Pi} A'_{\mu^*}\bar{\xi} = \overline{\lambda'(\mu^*)}\bar{\xi}.$

Proof. It is enough to prove only the first part. It is straightforward that

$$\begin{aligned} A_\mu \xi &= A_\mu(\xi - \xi(\mu)) + A_\mu \xi(\mu) \\ &= A_\mu(\xi - \xi(\mu)) + \lambda(\mu)\xi(\mu). \end{aligned}$$

Denote

$$\xi' = \frac{d}{d\mu}\xi(\mu)\Big|_{\mu=\mu^*}.$$

Then it follows that

$$\begin{aligned} A'_{\mu^*}\xi &= A'_{\mu^*}(\xi - \xi(\mu^*)) - A_{\mu^*}\xi' + \lambda'(\mu^*)\xi(\mu^*) + \lambda(\mu^*)\xi' \\ &= \lambda'(\mu^*)\xi(\mu^*) + (\lambda(\mu^*)I - A_{\mu^*})\xi' \\ &= \lambda'(\mu^*)\xi + (i\omega^* I - A_{\mu^*})\xi'. \end{aligned}$$

Taking account of the fact $\Pi[(i\omega^* I - A_{\mu^*})\xi'] = 0$, we must have

$$\Pi A'_{\mu^*}\xi = \lambda'(\mu^*)\xi.$$ Q.E.D.

We have finished the preparation for the spectral projection. And we shall now go back to our task to evaluate PNv^* and PMv^*. Recall that $M = D^2_{x,\mu}F(\omega^*, \mu^*, 0)$ and $N = D^2_{x,\omega}F(\omega^*, \mu^*, 0)$.

If we specify $v^* \in \mathfrak{V}$ as

$$v^* = \xi e^{it} + \bar{\xi} e^{-it}, \tag{5}$$

it follows that

$$\begin{aligned} D_x F(\omega, \mu, 0)v^* &= \omega \dot{v}^* - A_\mu v^* \\ &= i\omega(\xi e^{it} - \bar{\xi} e^{-it}) - A_\mu(\xi e^{it} + \bar{\xi} e^{-it}). \end{aligned} \tag{6}$$

Evaluation of *PNv** We obtain, by (6), that

$$\underbrace{D^2_{x\omega} F(\omega^*, \mu^*, 0)}_{N} v^* = i\xi e^{it} - i\bar{\xi} e^{-it}. \tag{7}$$

In general, the projection Py of $y \in \mathfrak{Y}$ into $\mathfrak{Z}$ can be calculated as

$$Py = \Pi(v_1)e^{it} + \Pi(v_{-1})e^{-it},$$

where $v_{\pm 1}$ are the Fourier coefficients of y corresponding to $k = \pm 1$.[13] Hence, by (7), PNv^* is evaluated as

$$\begin{aligned} PNv^* &= i\langle \eta, \xi \rangle \xi e^{it} - i\langle \bar{\eta}, \bar{\xi} \rangle \bar{\xi} e^{-it} \\ &= i\xi e^{it} - i\bar{\xi} e^{-it}. \end{aligned} \tag{8}$$

Evaluation of *PMv** Dividing $\lambda(\mu)$ (obtained by Lemma 3) into real and imaginary parts, we write

$$\lambda(\mu) = \alpha(\mu) + i\beta(\mu).$$

And we also write

$$\lambda'(\mu) = \alpha'(\mu) + i\beta'(\mu).$$

It follows from (6) that

$$\underbrace{D^2_{x\mu} F(\omega, \mu, 0)}_{M} v^* = -A'_{\mu^*}(\xi e^{it} + \bar{\xi} e^{-it}). \tag{9}$$

---

13 Express each of the Fourier coefficients of y by the direct sum corresponding to $\text{span}\{\xi\}$ and $[i\omega^* I - A_{\mu^*}](\mathbb{C}^n)$. And delete all the terms which do not contribute to the former.

Therefore we have by Lemma 4 that

$$
\begin{aligned}
PMv^* &= -\Pi A'_{\mu^*}\xi e^{it} - \overline{\Pi} A_{\mu^*}\bar{\xi} e^{-it} \\
&= -\lambda'(\mu^*)\xi e^{it} - \overline{\lambda'(\mu^*)}\bar{\xi} e^{-it} \\
&= -\alpha'(\mu^*)(\xi e^{it} + \bar{\xi} e^{-it}) - \beta'(\mu^*)(\xi e^{it} - \bar{\xi} e^{-it}). \qquad (10)
\end{aligned}
$$

Comparing (8) and (10), we get a simple fact that PNv^* and PMv^* are linearly independent if and only if $\alpha'(\mu^*) \neq 0$.

Thus all the requirements in Theorem 1 are fulfilled if we make an additional assumption that $\alpha'(\mu^*) \neq 0$.

Theorem 2. *Let $f(\mu, x) : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ be a function of the class $C^2(\mathbb{R} \times \mathbb{R}^n, \mathbb{R}^n)$ which satisfies $f(\mu, 0) = 0$ for all $\mu \in \mathbb{R}$. Suppose that Assumptions 1–3 as well as the condition $\alpha'(\mu^*) \neq 0$ are satisfied. Then (ω^*, μ^*) is a bifurcation point of $F(\omega, \mu, x) = \omega dx/dt - f(\mu, x)$.*

Remark. (due to S. Kusuoka) The additional condition $\alpha'(\mu^*) \neq 0$ can be expressed in an alternative equivalent form.

We denote by $\mathfrak{V}_0$ (resp. $\mathfrak{W}_0$) the kernel (resp. the image) of ${\omega^*}^2 I + A_{\mu^*}^2$; i.e.

$$\mathfrak{V}_0 = \mathrm{Ker}\{{\omega^*}^2 I + A_{\mu^*}^2\}, \text{ and}$$

$$\mathfrak{W}_0 = [{\omega^*}^2 I + A_{\mu^*}^2](\mathbb{R}^n).$$

We have, of course, that

$$\mathbb{R}^n = \mathfrak{V}_0 \oplus \mathfrak{W}_0. \qquad (11)$$

Let $\pi_0 : \mathbb{R}^n \to \mathfrak{V}_0$ be the projection of $\mathbb{R}^n$ into $\mathfrak{V}_0$ corresponding to the above direct product (11). Then we can prove that

$$\alpha'(\mu^*) = 0 \quad \text{if and only if} \quad \mathrm{tr} A'_{\mu^*}\pi_0 \neq 0, \qquad (12)$$

where tr is the trace of a matrix.

Proof. Let $\mathfrak{V}_0^{\mathbb{C}} = \mathbb{C} \otimes \mathfrak{V}_0$ and $\pi_0^{\mathbb{C}} : \mathbb{C}^n \to \mathfrak{V}_0^{\mathbb{C}}$ be the complexifications of $\mathfrak{V}_0$ and π_0, respectively. Recall the formula (8) in Sect. 5; i.e.

$$\mathrm{Ker}\{i\omega^* I - A_{\mu^*}\} = \mathrm{span}\{\xi\}.$$

We denote

$$\xi = a + ib; \quad a, b \in \mathbb{R}^n.$$

It can easily be checked that

$$A_{\mu^*} = -\omega^* b \quad \text{and} \quad A_{\mu^*} b = \omega^* a.$$

Consequently we also have that a and b are contained in $\mathfrak{V}_0$. They must be linearly independent. Hence there exist some $\tilde{a}$ and $\tilde{b} \in \mathbb{R}^n$ which satisfy the following conditions:

$$\begin{aligned} &\langle \tilde{a}, a \rangle = 1, \quad \langle \tilde{a}, b \rangle = 0, \\ &\langle \tilde{b}, a \rangle = 0, \quad \langle \tilde{b}, b \rangle = 1, \text{ and} \\ &\langle \tilde{a}, w \rangle = \langle \tilde{b}, w \rangle = 0 \quad \text{for all } w \in \mathfrak{W}_0. \end{aligned}$$

We, then, define the vector η by

$$\eta = \frac{1}{2}(\tilde{a} + i\tilde{b}).$$

This η clearly satisfies the condition in Sect. 7 (p. 56) that $\eta \neq 0$ and η is orthogonal to $[i\omega^* I - A_{\mu^*}](\mathbb{C}^*)$.

Since $\Pi A'_{\mu^*}\xi = \lambda'(\mu^*)\xi$ by Lemma 4, it follows that

$$\lambda'(\mu^*) \equiv \langle \eta, A'_{\mu^*}\xi \rangle = \frac{1}{2}\langle \tilde{a} + i\tilde{b}, A'_{\mu^*}\xi \rangle.$$

Finally we obtain the desired result by

$$\alpha'(\mu^*) = Re\lambda'(\mu^*) = \frac{1}{2}\{\langle \tilde{a}, A_{\mu^*} a \rangle + \langle \tilde{b}, A'_{\mu^*} b \rangle\} = \frac{1}{2}\mathrm{tr} A'_{\mu^*}\pi_0.$$

This proves the desired result.

9. Hopf bifurcation in C^r

In the preceding sections, we examined the Hopf bifurcation phenomena in the framework of the Sobolev space $\mathfrak{X} = \mathfrak{W}^{1,2}_{2\pi}$. However we have to note that some technical modifications are required when we consider the same problem in the alternative space consisting of periodic smooth functions.

Here we specify a couple of function spaces, $\mathfrak{X}$ and $\mathfrak{Y}$, as

$$\mathfrak{X} = \{x \in C^r(\mathbb{R}, \mathbb{R}^n) | x(t + 2\pi) = x(t) \text{ for all } t\}, \text{ and}$$

$$\mathfrak{Y} = \{x \in C^{r-1}(\mathbb{R}, \mathbb{R}^n) | y(t + 2\pi) = y(t) \text{ for all } t\},$$

where $r \geqq 3$.

Furthermore the function $f : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ is assumed to be of the class $C^{r-1}(\mathbb{R} \times \mathbb{R}^n, \mathbb{R}^n)$.

1° The first point to be examined is the formula (3) in Sect. 5. In the alternative setting, we have to proceed somewhat more carefully as follows.

Since x is of class $C^r \quad (r \geqq 3)$, we must have

$$u_k = o\Big(\frac{1}{|k|^r}\Big) \quad \text{as} \quad |k| \to \infty.\text{[14]} \tag{1}$$

Therefore, for any $\varepsilon > 0$, there exists some $N \in \mathbb{N}$ such that $|u_k| \leqq \varepsilon \cdot (1/|k|^r) \quad (|k| \geqq N)$. By differentiating the right-hand side of (1) of Sect. 5 termwise, we obtain

$$\begin{aligned}\Big\| \sum_{k=-\infty}^{\infty} iku_k e^{ikt} \Big\| &\leqq \sum_{k=-\infty}^{\infty} \|ku_k\| \\ &\leqq \sum_{|k|<N} \|ku_k\| + \varepsilon \sum_{|k| \geqq N} |k| \cdot \frac{1}{|k|^r} \\ &= \sum_{|k|<N} \|ku_k\| + \varepsilon \sum_{|k| \geqq N} \frac{1}{|k|^{r-1}}.\end{aligned}$$

Thus, taking account of the condition $r \geqq 3$, the series obtained by the termwise differentiation of the right-hand side of (1) is uniformly convergent. Hence we obtain

$$\dot{x}(t) = \sum_{k=-\infty}^{\infty} iku_k e^{ikt}. \tag{2}$$

2° The second modification is required in Sect. 6. The argument which succeeds the formula (5) should be replaced by the following one.

We claim that the second term of the right-hand side of (5) in Sect. 6 is of the class C^r. In order to prove it, we define the function $\theta(t)$ by

$$\theta(t) = \sum_{|k| \geqq k_0} v_k e^{ikt}. \tag{3}$$

Then $\theta(t)$ is of the class C^{r-1}. Let $\Theta(t)$ be the indefinite integral of $\theta(t)$, that is

$$\Theta(t) = \int_0^t \theta(\tau) d\tau. \tag{4}$$

[14] See Katznelson [8] p. 26.

Clearly $\Theta(t)$ is continuously differentiable and $\dot{\Theta}(t) = \theta(t)$. Expanding $\theta(t)$ in the Fourier series, we confirm that the coefficient $\hat{\theta}(0)$ corresponding to $k = 0$ is zero. Hence we obtain by the classical result in Fourier Analysis[15] that

$$\hat{\Theta}(k) = \frac{\hat{\theta}(k)}{ik} \quad \text{for all} \quad k \neq 0,$$

$$i.e. \quad \frac{1}{\omega^*}\Theta(t) = \sum_{|k| \geqq k_0} \frac{v_k}{ik\omega^*} e^{ikt}$$

$$= \text{the second term of (5).}$$

This is of the class C^r by (4).

The third term of the right-hand side of (5) in Sect. 6 is of the class C^{r-1}, and converges at every t. Differentiating termwise formally, we get

$$\sum_{|k| \geqq k_0} O\left(\frac{1}{k^2}\right) v_k \cdot ike^{ikt}. \tag{5}$$

The series (5) is uniformly convergent since $v_k = o(1/k^{r-1})$ $(r \geqq 3)$. Hence the third term of (5) in Sect. 6 is differentiable and (5) deduced above is exactly its derivative. Thus the third term is of the class C^r.

The first term of the right-hand side of (5) is obviously of the class C^r. This finishes the proof of the claim that the right-hand side of (5) is of the class C^r.

No change is required at all in the remaining part of the proof.

Theorem 3. *Let $f(\mu, x) : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ be a function of the class $C^{r-1}(\mathbb{R} \times \mathbb{R}^n, \mathbb{R}^n)$ $(r \geqq 3)$ which satisfies $f(\mu, 0) = 0$ for all $\mu \in \mathbb{R}$. Suppose that Assumption 2, Assumption 3 and $\alpha'(\mu^*) \neq 0$ are satisfied. Then (ω^*, μ^*) is a bifurcation point of $F(\omega, \mu, x) = \omega dx/dt - f(\mu, x)$.*

In Maruyama [9], an application of Theorem 3 to a dynamic economic problem is illustrated. N. Kaldor's view on the mechanism causing business cycles is formulated in terms of a differential equation of Liénard type and an existence proof of periodic solutions (business cycles) is provided based upon the Hopf theorem.

[15] See note 9 on page 54.

References[16]

1. Ambrosetti, A., Prodi, G.: A Primer of Nonlinear Analysis. Cambridge University Press, Cambridge (1993)
2. Carleson, L.: On convergence and growth of partial sums of Fourier series. Acta Math. **116**, 135–157 (1966)
3. Crandall, M.G., Rabinowitz, P.H.: The Hopf Bifurcation Theorem. MRC Technical Summary Report, No.1604. University of Wisconsin Mathematics Research Center (1976)
4. Evans, L.C.: Partial Differential Equations. American Mathematical Society, Providence (1998)
5. Gallego, F.B., Fajardo, J.C.V.: Bifurcations under nondegenerated conditions of higher degree and a new simple proof of the Hopf–Neimark–Sacker bifurcation theorem. J. Math. Anal. Appl. **237**, 93–105 (1999)
6. Golubitsky, M., Langford, W.F.: Classification and unfoldings of degenerate Hopf bifurcations. J. Differ. Equ. **41**, 375–415 (1981)
7. Hunt, R.A.: On the convergence of Fourier series. In: Proceedings of the Conference on Orthogonal Expansions and Their Continuous Analogues, pp. 234–255 (1968)
8. Katznelson, Y.: An Introduction to Harmonic Analysis, 3rd edn. Cambridge University Press, Cambridge (2004)
9. Maruyama, T.: Existence of periodic solutions for Kaldorian business fluctuations. In: Mordukhovich, B.S., et al. (eds.) Nonlinear Analysis and Optimization II. Contemporary Mathematics, vol. 514, pp. 189–197. American Mathematical Society, Providence (2010)
10. Neĭmark, J.I.: On some cases of periodic motions depending on parameters. Dokl. Acad. Nauk SSSR **129**, 736–739 (1959)
11. Zeidler, E.: Applied Functional Analysis. Springer, New York (1995)
12. Zygmund, A.: Trigonometric Series, 2nd edn., vol. 1. Cambridge University Press, London (1959)

[16] The references [3, 5, 6, 10] are not cited in text. However these are closely related works which must attract readers' interest.

Adv. Math. Econ. 15, 67–88 (2011)

Advances in MATHEMATICAL ECONOMICS

Fixed point theorems and ergodic theorems for nonlinear mappings in Banach spaces

Pavel Kocourek[1], **Wataru Takahashi**[2,1], **and Jen-Chih Yao**[1,*]

[1] Department of Applied Mathematics, National Sun Yat-sen University, Kaohsiung 80424, Taiwan
(e-mail: pakocica@gmail.com; yaojc@math.nsysu.edu.tw)

[2] Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Tokyo 152-8552, Japan
(e-mail: wataru@is.titech.ac.jp)

Received: September 29, 2010
Revised: October 25, 2010

JEL classification: C62, C68

Mathematics Subject Classification (2010): 47H09, 47H10, 47H25

Abstract. In this paper, we first introduce a class of nonlinear mappings called generalized nonspreading which contains the class of nonspreading mappings in a Banach space and then prove a fixed point theorem, a nonlinear mean convergence theorem of Baillon's type and a weak convergence theorem of Mann's type for such nonlinear mappings in a Banach space. Using these theorems, we obtain some fixed point theorems, nonlinear mean convergence theorems and weak convergence theorems in a Banach space.

Key words: Banach space, fixed point, mean convergence, weak convergence, nonexpansive mapping, nonspreading mapping

1. Introduction

Let H be a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and norm $\|\cdot\|$ and let C be a nonempty subset of H. Let T be a mapping of C into H. Then we denote by $F(T)$ the set of fixed points of T. A mapping $T : C \to H$ is

* The second author and the third author were partially supported by Grant-in-Aid for Scientific Research No. 19540167 from Japan Society for the Promotion of Science and by the grant NSC, respectively.

DOI: 10.1007/978-4-431-53930-8_4,

called nonexpansive if $\|Tx - Ty\| \leq \|x - y\|$ for all $x, y \in C$. An important example of nonexpansive mappings in a Hilbert space is a firmly nonexpansive mapping. Let C be a nonempty subset of H. A mapping $F : C \to H$ is said to be firmly nonexpansive if

$$\|Fx - Fy\|^2 \leq \langle x - y, Fx - Fy \rangle$$

for all $x, y \in C$; see, for instance, Browder [6] and Goebel and Kirk [8]. It is known that a firmly nonexpansive mapping F can be deduced from an equilibrium problem in a Hilbert space; see, for instance, [4] and [7]. Recently, Kohsaka and Takahashi [26], and Takahashi [35] introduced the following nonlinear mappings which are deduced from a firmly nonexpansive mapping in a Hilbert space. A mapping $T : C \to H$ is called nonspreading [26] if

$$2\|Tx - Ty\|^2 \leq \|Tx - y\|^2 + \|Ty - x\|^2$$

for all $x, y \in C$. A mapping $T : C \to H$ is called hybrid [35] if

$$3\|Tx - Ty\|^2 \leq \|x - y\|^2 + \|Tx - y\|^2 + \|Ty - x\|^2$$

for all $x, y \in C$. They proved fixed point theorems for such mappings; see also Kohsaka and Takahashi [25] and Iemoto and Takahashi [17]. Recently, Kocourek, Takahashi and Yao [22] defined a broad class of nonlinear mappings containing the classes of nonexpansive mappings, nonspreading mappings and hybrid mappings in a Hilbert space: A mapping $T : C \to H$ is called generalized hybrid [22] if there are $\alpha, \beta \in \mathbb{R}$ such that

$$\alpha\|Tx - Ty\|^2 + (1-\alpha)\|x - Ty\|^2 \leq \beta\|Tx - y\|^2 + (1-\beta)\|x - y\|^2 \tag{1.1}$$

for all $x, y \in C$. We call such a mapping an (α, β)-generalized hybrid mapping. Then, Kocourek, Takahashi and Yao [22] proved a fixed point theorem for such mappings in a Hilbert space. Further, they proved a nonlinear mean convergence theorem of Baillon's type [3] in a Hilbert space.

In this paper, motivated by these results, we first introduce a class of nonlinear mappings called generalized nonspreading which contains the class of nonspreading mappings in a Banach space and then prove a fixed point theorem, a nonlinear mean convergence theorem of Baillon's type and a weak convergence theorem of Mann's type for such nonlinear mappings in a Banach space. Using these theorems, we obtain some fixed point theorems, nonlinear mean convergence theorems and weak convergence theorems in a Banach space.

2. Preliminaries

Let E be a real Banach space with norm $\|\cdot\|$ and let E^* be the topological dual space of E. We denote the value of $y^* \in E^*$ at $x \in E$ by $\langle x, y^* \rangle$. When $\{x_n\}$ is a sequence in E, we denote the strong convergence of $\{x_n\}$ to $x \in E$ by $x_n \to x$ and the weak convergence by $x_n \rightharpoonup x$. The modulus δ of convexity of E is defined by

$$\delta(\epsilon) = \inf\left\{1 - \frac{\|x+y\|}{2} : \|x\| \le 1, \|y\| \le 1, \|x-y\| \ge \epsilon\right\}$$

for every ϵ with $0 \le \epsilon \le 2$. A Banach space E is said to be uniformly convex if $\delta(\epsilon) > 0$ for every $\epsilon > 0$. A uniformly convex Banach space is strictly convex and reflexive. Let C be a nonempty subset of a Banach space E. A mapping $T : C \to E$ is nonexpansive [5, 9, 19] if $\|Tx - Ty\| \le \|x - y\|$ for all $x, y \in C$. A mapping $T : C \to E$ is quasi-nonexpansive if $F(T) \neq \emptyset$ and $\|Tx - y\| \le \|x - y\|$ for all $x \in C$ and $y \in F(T)$, where $F(T)$ is the set of fixed points of T. If C is a nonempty closed convex subset of a strictly convex Banach space E and $T : C \to C$ is quasi-nonexpansive, then $F(T)$ is closed and convex; see Itoh and Takahashi [18]. Let E be a Banach space. The duality mapping J from E into 2^{E^*} is defined by

$$Jx = \{x^* \in E^* : \langle x, x^* \rangle = \|x\|^2 = \|x^*\|^2\}$$

for every $x \in E$. Let $U = \{x \in E : \|x\| = 1\}$. The norm of E is said to be Gâteaux differentiable if for each $x, y \in U$, the limit

$$\lim_{t \to 0} \frac{\|x + ty\| - \|x\|}{t} \tag{2.1}$$

exists. In the case, E is called smooth. We know that E is smooth if and only if J is a single-valued mapping of E into E^*. We also know that E is reflexive if and only if J is surjective, and E is strictly convex if and only if J is one-to-one. Therefore, if E is a smooth, strictly convex and reflexive Banach space, then J is a single-valued bijection. The norm of E is said to be uniformly Gâteaux differentiable if for each $y \in U$, the limit (2.1) is attained uniformly for $x \in U$. It is also said to be Fréchet differentiable if for each $x \in U$, the limit (2.1) is attained uniformly for $y \in U$. A Banach space E is called uniformly smooth if the limit (2.1) is attained uniformly for $x, y \in U$. It is known that if the norm of E is uniformly Gâteaux differentiable, then J is uniformly norm-to-weak* continuous on each bounded subset of E, and if the norm of E is Fréchet differentiable, then J is norm-to-norm continuous. If E is uniformly smooth, J is uniformly norm-to-norm continuous on each bounded subset of E. For more details, see [31, 32]. The following result is also well known; see [32].

Theorem 2.1. *Let E be a smooth Banach space and let J be the duality mapping on E. Then, $\langle x-y, Jx-Jy\rangle \geq 0$ for all $x, y \in E$. Further, if E is strictly convex and $\langle x-y, Jx-Jy\rangle = 0$, then $x=y$.*

Let E be a smooth Banach space. The function $\phi\colon E\times E \to (-\infty,\infty)$ is defined by

$$\phi(x,y) = \|x\|^2 - 2\langle x, Jy\rangle + \|y\|^2$$

for $x, y \in E$, where J is the duality mapping of E; see [1] and [20]. We have from the definition of ϕ that

$$\phi(x,y) = \phi(x,z) + \phi(z,y) + 2\langle x-z, Jz-Jy\rangle \tag{2.2}$$

for all $x, y, z \in E$. From $(\|x\|-\|y\|)^2 \leq \phi(x,y)$ for all $x, y \in E$, we can see that $\phi(x,y) \geq 0$. Further, we can obtain the following equality:

$$2\langle x-y, Jz-Jw\rangle = \phi(x,w) + \phi(y,z) - \phi(x,z) - \phi(y,w) \tag{2.3}$$

for $x, y, z, w \in E$. If E is additionally assumed to be strictly convex, then

$$\phi(x,y) = 0 \Longleftrightarrow x = y. \tag{2.4}$$

The following theorems are in Xu [39] and Kamimura and Takahashi [20].

Theorem 2.2 (Xu [39]). *Let E be a uniformly convex Banach space and let $r>0$. Then there exists a strictly increasing, continuous and convex function $g:[0,\infty)\to[0,\infty)$ such that $g(0)=0$ and*

$$\|\lambda x + (1-\lambda)y\|^2 \leq \lambda\|x\|^2 + (1-\lambda)\|y\|^2 - \lambda(1-\lambda)g(\|x-y\|)$$

for all $x, y \in B_r$ and λ with $0\leq\lambda\leq 1$, where $B_r = \{z \in E : \|z\| \leq r\}$.

Theorem 2.3 (Kamimura and Takahashi [20]). *Let E be smooth and uniformly convex Banach space and let $r>0$. Then there exists a strictly increasing, continuous and convex function $g:[0,2r]\to\mathbb{R}$ such that $g(0)=0$ and*

$$g(\|x-y\|) \leq \phi(x,y)$$

for all $x, y \in B_r$, where $B_r = \{z \in E : \|z\| \leq r\}$.

Let E be a smooth Banach space and let C be a nonempty subset of E. Then a mapping $T: C\to E$ is called generalized nonexpansive [11, 13, 15] if $F(T)\neq\emptyset$ and

$$\phi(Tx,y) \leq \phi(x,y)$$

for all $x\in C$ and $y\in F(T)$. Let D be a nonempty subset of a Banach space E. A mapping $R: E\to D$ is said to be sunny if

$$R(Rx + t(x - Rx)) = Rx$$

for all $x \in E$ and $t \geq 0$. A mapping $R : E \to D$ is said to be a retraction or a projection if $Rx = x$ for all $x \in D$. A nonempty subset D of a smooth Banach space E is said to be a generalized nonexpansive retract (resp. sunny generalized nonexpansive retract) of E if there exists a generalized nonexpansive retraction (resp. sunny generalized nonexpansive retraction) R from E onto D; see [10–12] for more details. The following results are in Ibaraki and Takahashi [11].

Theorem 2.4 (Ibaraki and Takahashi [11]). *Let C be a nonempty closed sunny generalized nonexpansive retract of a smooth and strictly convex Banach space E. Then the sunny generalized nonexpansive retraction from E onto C is uniquely determined.*

Theorem 2.5 (Ibaraki and Takahashi [11]). *Let C be a nonempty closed subset of a smooth and strictly convex Banach space E such that there exists a sunny generalized nonexpansive retraction R from E onto C and let $(x, z) \in E \times C$. Then the following hold:*

(i) *$z = Rx$ if and only if $\langle x - z, Jy - Jz \rangle \leq 0$ for all $y \in C$;*
(ii) *$\phi(Rx, z) + \phi(x, Rx) \leq \phi(x, z)$.*

In 2007, Kohsaka and Takahashi [24] proved the following results:

Theorem 2.6 (Kohsaka and Takahashi [24]). *Let E be a smooth, strictly convex and reflexive Banach space and let C be a nonempty closed subset of E. Then the following are equivalent:*

(a) *C is a sunny generalized nonexpansive retract of E;*
(b) *C is a generalized nonexpansive retract of E;*
(c) *JC is closed and convex.*

Theorem 2.7 (Kohsaka and Takahashi [24]). *Let E be a smooth, strictly convex and reflexive Banach space and let C be a nonempty closed sunny generalized nonexpansive retract of E. Let R be the sunny generalized nonexpansive retraction from E onto C and let $(x, z) \in E \times C$. Then the following are equivalent:*

(i) *$z = Rx$;*
(ii) *$\phi(x, z) = min_{y \in C} \phi(x, y)$.*

Very recently, Ibaraki and Takahashi [16] also obtained the following result concerning the set of fixed points of a generalized nonexpansive mapping.

Theorem 2.8 (Ibaraki and Takahashi [16]). *Let E be a reflexive, strictly convex and smooth Banach space and let T be a generalized nonexpansive mapping from E into itself. Then, $F(T)$ is closed and $JF(T)$ is closed and convex.*

The following is a direct consequence of Theorems 2.6 and 2.8.

Theorem 2.9 (Ibaraki and Takahashi [16]). *Let E be a reflexive, strictly convex and smooth Banach space and let T be a generalized nonexpansive mapping from E into itself. Then, $F(T)$ is a sunny generalized nonexpansive retract of E.*

3. Fixed point theorems

In this section, we try to extend Kocourek, Takahashi and Yao's fixed point theorem [22] in a Hilbert space to that in a Banach space. Let E be a smooth Banach space, let C be a nonempty closed convex subset of E and let J be the duality mapping from E into E^*. Then, a mapping $T : C \to E$ is called generalized nonspreading if there are $\alpha, \beta, \gamma, \delta \in \mathbb{R}$ such that

$$\begin{aligned} &\alpha\phi(Tx, Ty) + (1-\alpha)\phi(x, Ty) + \gamma\{\phi(Ty, Tx) - \phi(Ty, x)\} \\ &\leq \beta\phi(Tx, y) + (1-\beta)\phi(x, y) + \delta\{\phi(y, Tx) - \phi(y, x)\} \end{aligned} \tag{3.1}$$

for all $x, y \in C$, where $\phi(x, y) = \|x\|^2 - 2\langle x, Jy\rangle + \|y\|^2$ for $x, y \in E$. We call such a mapping an $(\alpha, \beta, \gamma, \delta)$-generalized nonspreading mapping. Let T be an $(\alpha, \beta, \gamma, \delta)$-generalized nonspreading mapping. Observe that if $F(T) \neq \emptyset$, then $\phi(u, Ty) \leq \phi(u, y)$ for all $u \in F(T)$ and $y \in C$. Indeed, putting $x = u \in F(T)$ in (3.1), we obtain

$$\phi(u, Ty) + \gamma\{\phi(Ty, u) - \phi(Ty, u)\} \leq \phi(u, y) + \delta\{\phi(y, u) - \phi(y, u)\}.$$

So, we have that

$$\phi(u, Ty) \leq \phi(u, y) \tag{3.2}$$

for all $u \in F(T)$ and $y \in C$. Further, if E is a Hilbert space, then we have $\phi(x, y) = \|x - y\|^2$ for $x, y \in E$. So, from (3.1) we obtain the following:

$$\begin{aligned} &\alpha\|Tx - Ty\|^2 + (1-\alpha)\|x - Ty\|^2 + \gamma\{\|Tx - Ty\|^2 - \|x - Ty\|^2\} \\ &\leq \beta\|Tx - y\|^2 + (1-\beta)\|x - y\|^2 + \delta\{\|Tx - y\|^2 - \|x - y\|^2\} \end{aligned} \tag{3.3}$$

for all $x, y \in C$. This implies that

$$(\alpha+\gamma)\|Tx - Ty\|^2 + \{1-(\alpha+\gamma)\}\|x - Ty\|^2$$
$$\le (\beta+\delta)\|Tx - y\|^2 + \{1-(\beta+\delta)\}\|x - y\|^2 \tag{3.4}$$

for all $x, y \in C$. That is, T is a generalized hybrid mapping [22] in a Hilbert space. Now, using the technique developed by [30], we prove a fixed point theorem for generalized nonspreading mappings in a Banach space.

Theorem 3.1. *Let E be a smooth, strictly convex and reflexive Banach space and let C be a nonempty closed convex subset of E. Let T be a generalized nonspreading mapping of C into itself. Then, the following are equivalent:*

(a) *$F(T) \neq \emptyset$;*
(b) *$\{T^n x\}$ is bounded for some $x \in C$.*

Proof. Let T be a generalized nonspreading mapping of C into itselt. Then, there are $\alpha, \beta, \gamma, \delta \in \mathbb{R}$ such that

$$\alpha\phi(Tx, Ty) + (1-\alpha)\phi(x, Ty) + \gamma\{\phi(Ty, Tx) - \phi(Ty, x)\}$$
$$\le \beta\phi(Tx, y) + (1-\beta)\phi(x, y) + \delta\{\phi(y, Tx) - \phi(y, x)\}$$

for all $x, y \in C$. If $F(T) \neq \emptyset$, then $\phi(u, Ty) \le \phi(u, y)$ for all $u \in F(T)$ and $y \in C$. So, if u is a fixed point in C, then we have $\phi(u, T^n x) \le \phi(u, x)$ for all $n \in \mathbb{N}$ and $x \in C$. This implies (a) $\Longrightarrow$ (b). Let us show (b) $\Longrightarrow$ (a). Suppose that there exists $x \in C$ such that $\{T^n x\}$ is bounded. Then for any $y \in C$ and $k \in \mathbb{N} \cup \{0\}$, we have

$$\begin{aligned}
&\alpha\phi(T^{k+1}x, Ty) + (1-\alpha)\phi(T^k x, Ty) + \gamma\{\phi(Ty, T^{k+1}x) - \phi(Ty, T^k x)\} \\
&\quad \le \beta\phi(T^{k+1}x, y) + (1-\beta)\phi(T^k x, y) + \delta\{\phi(y, T^{k+1}x) - \phi(y, T^k x)\} \\
&\quad = \beta\{\phi(T^{k+1}x, Ty) + \phi(Ty, y) + 2\langle T^{k+1}x - Ty, JTy - Jy\rangle\} \\
&\qquad + (1-\beta)\{\phi(T^k x, Ty) + \phi(Ty, y) + 2\langle T^k x - Ty, JTy - Jy\rangle\} \\
&\qquad + \delta\{\phi(y, T^{k+1}x) - \phi(y, T^k x)\}.
\end{aligned} \tag{3.5}$$

This implies that

$$\begin{aligned}
0 &\le (\beta-\alpha)\{\phi(T^{k+1}x, Ty) - \phi(T^k x, Ty)\} + \phi(Ty, y) \\
&\quad + 2\langle \beta T^{k+1}x + (1-\beta)T^k x - Ty, JTy - Jy\rangle \\
&\quad - \gamma\{\phi(Ty, T^{k+1}x) - \phi(Ty, T^k x)\} + \delta\{\phi(y, T^{k+1}x) - \phi(y, T^k x)\}.
\end{aligned} \tag{3.6}$$

Summing up these inequalities (3.6) with respect to $k = 0, 1, \dots, n-1$, we have

$$\begin{aligned} 0 &\le (\beta - \alpha)\{\phi(T^n x, Ty) - \phi(x, Ty)\} + n\phi(Ty, y) \\ &+ 2\langle x + Tx + \cdots + T^{n-1}x + \beta(T^n x - x) - nTy, JTy - Jy\rangle \\ &- \gamma\{\phi(Ty, T^n x) - \phi(Ty, x)\} + \delta\{\phi(y, T^n x) - \phi(y, x)\}. \end{aligned} \tag{3.7}$$

Dividing by n in (3.7), we have

$$\begin{aligned} 0 &\le \frac{1}{n}(\beta - \alpha)\{\phi(T^n x, Ty) - \phi(x, Ty)\} + \phi(Ty, y) \\ &+ 2\langle S_n x + \frac{1}{n}\beta(T^n x - x) - Ty, JTy - Jy\rangle \\ &- \frac{1}{n}\gamma\{\phi(Ty, T^n x) - \phi(Ty, x)\} + \frac{1}{n}\delta\{\phi(y, T^n x) - \phi(y, x)\}, \end{aligned} \tag{3.8}$$

where $S_n x = \frac{1}{n}\sum_{k=0}^{n-1} T^k x$. Since $\{T^n x\}$ is bounded by assumption, $\{S_n x\}$ is bounded. Thus we have a subsequence $\{S_{n_i} x\}$ of $\{S_n x\}$ such that $\{S_{n_i} x\}$ converges weakly to a point $u \in C$. Letting $n_i \to \infty$ in (3.8), we obtain

$$0 \le \phi(Ty, y) + 2\langle u - Ty, JTy - Jy\rangle.$$

Putting $y = u$, we obtain

$$\begin{aligned} 0 &\le \phi(Tu, u) + 2\langle u - Tu, JTu - Ju\rangle \\ &= \phi(Tu, u) + \phi(u, u) + \phi(Tu, Tu) - \phi(u, Tu) - \phi(Tu, u) \\ &= -\phi(u, Tu). \end{aligned} \tag{3.9}$$

Hence we have $\phi(u, Tu) \le 0$ and then $\phi(u, Tu) = 0$. Since E is strictly convex, we obtain $u = Tu$. Therefore $F(T)$ is nonempty. This completes the proof. □

Using Theorem 3.1, we have the following fixed point theorems in a Banach space.

Theorem 3.2 (Kohsaka and Takahashi [26]). *Let E be a smooth, strictly convex and reflexive Banach space and let C be a nonempty closed convex subset of E. Let $T : C \to C$ be a nonspreading mapping, i.e.,*

$$\phi(Tx, Ty) + \phi(Ty, Tx) \le \phi(Tx, y) + \phi(Ty, x)$$

for all $x, y \in C$*. Then, the following are equivalent:*

(a) $F(T) \neq \emptyset$*;*
(b) $\{T^n x\}$ *is bounded for some* $x \in C$*.*

Proof. Putting $\alpha = \beta = \gamma = 1$ and $\delta = 0$ in (3.1), we obtain that

$$\phi(Tx, Ty) + \phi(Ty, Tx) \leq \phi(Tx, y) + \phi(Ty, x)$$

for all $x, y \in C$. So, we have the desired result from Theorem 3.1. □

See [2, 21, 23, 33, 38] for examples and convergence theorems for non-spreading mappings in a Banach space.

Theorem 3.3. *Let* E *be a smooth, strictly convex and reflexive Banach space and let* C *be a nonempty closed convex subset of* E*. Let* $T : C \to C$ *be a hybrid mapping [35], i.e.,*

$$2\phi(Tx, Ty) + \phi(Ty, Tx) \leq \phi(Tx, y) + \phi(Ty, x) + \phi(x, y)$$

for all $x, y \in C$*. Then, the following are equivalent:*

(a) $F(T) \neq \emptyset$*;*
(b) $\{T^n x\}$ *is bounded for some* $x \in C$*.*

Proof. Putting $\alpha = 1$, $\beta = \gamma = \frac{1}{2}$ and $\delta = 0$ in (3.1), we obtain that

$$2\phi(Tx, Ty) + \phi(Ty, Tx) \leq \phi(Tx, y) + \phi(Ty, x) + \phi(x, y)$$

for all $x, y \in C$. So, we have the desired result from Theorem 3.1. □

Theorem 3.4. *Let* E *be a smooth, strictly convex and reflexive Banach space and let* C *be a nonempty closed convex subset of* E*. Let* $T : C \to C$ *be a mapping such that*

$$\alpha\phi(Tx, Ty) + (1 - \alpha)\phi(x, Ty) \leq \beta\phi(Tx, y) + (1 - \beta)\phi(x, y)$$

for all $x, y \in C$*. Then, the following are equivalent:*

(a) $F(T) \neq \emptyset$*;*
(b) $\{T^n x\}$ *is bounded for some* $x \in C$*.*

Proof. Putting $\gamma = \delta = 0$ in (3.1), we obtain that

$$\alpha\phi(Tx, Ty) + (1 - \alpha)\phi(x, Ty) \leq \beta\phi(Tx, y) + (1 - \beta)\phi(x, y)$$

for all $x, y \in C$. So, we have the desired result from Theorem 3.1. □

As a direct consequence of Theorem 3.4, we have Kocourek, Takahashi and Yao's fixed point theorem [22] in a Hilbert space.

Theorem 3.5 (Kocourek, Takahashi and Yao [22]). *Let C be a nonempty closed convex subset of a Hilbert space H and let $T : C \to C$ be a generalized hybrid mapping, i.e., there are $\alpha, \beta \in \mathbb{R}$ such that*

$$\alpha\|Tx - Ty\|^2 + (1-\alpha)\|x - Ty\|^2 \leq \beta\|Tx - y\|^2 + (1-\beta)\|x - y\|^2 \tag{3.10}$$

for all $x, y \in C$. Then, the following are equivalent:

(a) $F(T) \neq \emptyset$;
(b) $\{T^n x\}$ *is bounded for some* $x \in C$.

Proof. We know that $\phi(x, y) = \|x - y\|^2$ for all $x, y \in C$ in Theorem 3.4. So, we have the desired result from Theorem 3.4. □

4. Some properties of generalized nonspreading mappings

In this section, we first discuss the demiclosedness property of generalized nonspreading mappings in a Banach space. Let E be a Banach space and let C be a nonempty subset of E. Let $T : C \to C$ be a mapping. Then, $p \in C$ be an asymptotic fixed point of T [29] if there exists $\{x_n\} \subset C$ such that $x_n \rightharpoonup p$ and $\lim_{n\to\infty} \|x_n - Tx_n\| = 0$. We denote by $\hat{F}(T)$ the set of asymptotic fixed points of T. A mapping T of C into itself is said to have the demiclosedness property on C if $\hat{F}(T) = F(T)$.

Proposition 4.1. *Let E be a strictly convex Banach space with a uniformly Gâteaux differentiable norm, let C be a nonempty closed convex subset of E and let T be a generalized nonspreading mapping of C into itself. Then $\hat{F}(T) = F(T)$.*

Proof. Since $T : C \to C$ is a generalized nonspreading mapping, there are $\alpha, \beta, \gamma, \delta \in \mathbb{R}$ such that

$$\begin{aligned} &\alpha\phi(Tx, Ty) + (1-\alpha)\phi(x, Ty) + \gamma\{\phi(Ty, Tx) - \phi(Ty, x)\} \\ &\leq \beta\phi(Tx, y) + (1-\beta)\phi(x, y) + \delta\{\phi(y, Tx) - \phi(y, x)\} \end{aligned} \tag{4.1}$$

for all $x, y \in C$. The inclusion $F(T) \subset \hat{F}(T)$ is obvious. Thus we show $\hat{F}(T) \subset F(T)$. Let $u \in \hat{F}(T)$ be given. Then we have a sequence $\{x_n\}$ of C such that $x_n \rightharpoonup u$ and $\lim_{n\to\infty} \|x_n - Tx_n\| = 0$. Since the norm of E is uniformly Gâteaux differentiable, the duality mapping J on E is uniformly norm-to-weak* continuous on each bounded subset of E; see Takahashi [31]. Thus

$$\lim_{n\to\infty} \langle w, JTx_n - Jx_n \rangle = 0$$

for all $w \in E$. On the other hand, since $T : C \to C$ is a generalized nonspreading mapping, we obtain that

$$
\begin{aligned}
&\alpha\phi(Tx_n, Tu) + (1-\alpha)\phi(x_n, Tu) + \gamma\{\phi(Tu, Tx_n) - \phi(Tu, x_n)\} \\
&\quad \le \beta\phi(Tx_n, u) + (1-\beta)\phi(x_n, u) + \delta\{\phi(u, Tx_n) - \phi(u, x_n)\} \\
&\quad = \beta\{\phi(Tx_n, Tu) + \phi(Tu, u) + 2\langle Tx_n - Tu, JTu - Ju\rangle\} \\
&\qquad + (1-\beta)\{\phi(x_n, Tu) + \phi(Tu, u) + 2\langle x_n - Tu, JTu - Ju\rangle\} \\
&\qquad + \delta\{\phi(u, Tx_n) - \phi(u, x_n)\}.
\end{aligned} \tag{4.2}
$$

This implies that

$$
\begin{aligned}
0 &\le (\beta - \alpha)\{\phi(Tx_n, Tu) - \phi(x_n, Tu)\} + \phi(Tu, u) \\
&\quad + 2\langle \beta Tx_n + (1-\beta)x_n - Tu, JTu - Ju\rangle \\
&\quad - \gamma\{\phi(Tu, Tx_n) - \phi(Tu, x_n)\} + \delta\{\phi(u, Tx_n) - \phi(u, x_n)\} \\
&= (\beta - \alpha)\{\|Tx_n\|^2 - \|x_n\|^2 - 2\langle Tx_n - x_n, JTu\rangle\} + \phi(Tu, u) \\
&\quad + 2\langle \beta(Tx_n - x_n) + x_n - Tu, JTu - Ju\rangle \\
&\quad - \gamma\{\|Tx_n\|^2 - \|x_n\|^2 - 2\langle Tu, JTx_n - Jx_n\rangle\} \\
&\quad + \delta\{\|Tx_n\|^2 - \|x_n\|^2 - 2\langle u, JTx_n - Jx_n\rangle\}.
\end{aligned} \tag{4.3}
$$

From

$$
\begin{aligned}
\left|\|Tx_n\|^2 - \|x_n\|^2\right| &= (\|Tx_n\| + \|x_n\|)\left|\|Tx_n\| - \|x_n\|\right| \\
&\le (\|Tx_n\| + \|x_n\|)\|Tx_n - x_n\|,
\end{aligned}
$$

we have $\|Tx_n\|^2 - \|x_n\|^2 \to 0$ as $n \to \infty$. So, letting $n \to \infty$ in (4.3), we have that

$$
\begin{aligned}
0 &\le \phi(Tu, u) + 2\langle u - Tu, JTu - Ju\rangle \\
&= \phi(Tu, u) + \phi(u, u) + \phi(Tu, Tu) - \phi(u, Tu) - \phi(Tu, u) \\
&= -\phi(u, Tu).
\end{aligned}
$$

Thus $\phi(u, Tu) \le 0$ and then $\phi(u, Tu) = 0$. Since E is strictly convex, we obtain $u = Tu$. This completes the proof. □

From Matsushita and Takahashi [28], we also know the following result.

Lemma 4.2 (Matsushita and Takahashi [28]). *Let E be a smooth and strictly convex Banach space, let C be a nonempty closed convex subset of E and let T be a mapping of C into itself such that $F(T)$ is nonempty. Assume that*

$$\phi(u, Ty) \leq \phi(u, y)$$

for all $u \in F(T)$ and $y \in C$. Then $F(T)$ is closed and convex.

Using this lemma (Lemma 4.2) and (3.2), we have the following result.

Proposition 4.3. *Let E be a smooth and strictly convex Banach space, let C be a nonempty closed convex subset of E and let T be a generalized nonspreading mapping of C into itself such that $F(T)$ is nonempty. Then $F(T)$ is closed and convex.*

Proof. It is sufficient to consider the case that $F(T)$ is nonempty. Then we have from (3.2) that $\phi(u, Ty) \leq \phi(u, y)$ for all $u \in F(T)$ and $y \in C$. From Lemma 4.2, $F(T)$ is closed and convex. □

Let E be a reflexive, smooth and strictly convex Banach space. Let C be a nonempty subset of E. Matsushita and Takahashi [28] also gave the following definition: A mapping $T : C \to C$ is relatively nonexpansive if $F(T) \neq \emptyset$, $\hat{F}(T) = F(T)$ and

$$\phi(y, Tx) \leq \phi(y, x)$$

for all $x \in C$ and $y \in F(T)$. Using Proposition 4.1, we prove the following theorem.

Theorem 4.4. *Let E be a strictly convex Banach space with a uniformly Gâteaux differentiable norm, let C be a nonempty closed convex subset of E and let T be a generalized nonspreading mapping of C into itself such that $F(T)$ is nonempty. Then, T is relatively nonexpansive.*

Proof. By assumption, $F(T)$ is nonempty. Since T is a generalized nonspreading mapping of C into itself, we have that

$$\phi(y, Tx) \leq \phi(y, x)$$

for all $x \in C$ and $y \in F(T)$. From Proposition 4.1, we also have $\hat{F}(T) = F(T)$. Thus T is relatively nonexpansive. □

As a direct consequence of Theorem 4.4, we have the following result.

Theorem 4.5 (Kohsaka and Takahashi [26]). *Let E be a strictly convex Banach space with a uniformly Gâteaux differentiable norm, let C be a nonempty closed convex subset of E and let T be a nonspreading mapping of C into itself such that $F(T)$ is nonempty. Then, T is relatively nonexpansive.* □

Proof. An $(\alpha, \beta, \gamma, \delta)$-generalized hybrid mapping T of C into itself such that $\alpha = \beta = \gamma = 1$ and $\delta = 0$ is a nonspreading mapping. From Theorem 4.4, we have the desired result.

5. Nonlinear ergodic theorems

Let E be a smooth Banach space and let C be a nonempty closed convex subset of E. Let $T : C \to E$ be a generalized nonspreading mapping; see (3.1). This mapping has the property that if $u \in F(T)$ and $x \in C$, then $\phi(u, Tx) \leq \phi(u, x)$. This property can be revealed by putting $x = u \in F(T)$ in (3.1). Similarly, putting $y = u \in F(T)$ in (3.1), we obtain that for $x \in C$,

$$\alpha\phi(Tx, u) + (1-\alpha)\phi(x, u) + \gamma\{\phi(u, Tx) - \phi(u, x)\}$$
$$\leq \beta\phi(Tx, u) + (1-\beta)\phi(x, u) + \delta\{\phi(u, Tx) - \phi(u, x)\} \tag{5.1}$$

and hence

$$(\alpha - \beta)\{\phi(Tx, u) - \phi(x, u)\} + (\gamma - \delta)\{\phi(u, Tx) - \phi(u, x)\} \leq 0. \tag{5.2}$$

Therefore, we have that $\alpha > \beta$ together with $\gamma \leq \delta$ implies that

$$\phi(Tx, u) \leq \phi(x, u).$$

Now, we can prove the following nonlinear ergodic theorem for generalized nonspreading mappings in a Banach space.

Theorem 5.1. *Let E be a uniformly convex Banach space with a Fréchet differentiable norm and let C be a nonempty closed convex sunny generalized nonexpansive retract of E. Let $T : C \to C$ be a generalized nonspreading mapping with $F(T) \neq \emptyset$ such that $\phi(Tx, u) \leq \phi(x, u)$ for all $x \in C$ and $u \in F(T)$. Let R be the sunny generalized nonexpansive retraction of E onto $F(T)$. Then, for any $x \in C$,*

$$S_n x = \frac{1}{n}\sum_{k=0}^{n-1} T^k x$$

converges weakly to an element q of $F(T)$, where $q = \lim_{n\to\infty} RT^n x$.

Proof. We know that since C is a sunny generalized nonexpansive retract of E, there exists the sunny generalized nonexpansive retraction P of E onto C. On the other hand, by assumption, $T : C \to C$ is a generalized nonexpansive mapping, i.e., $F(T) \neq \emptyset$ and

$$\phi(Tx, u) \leq \phi(x, u)$$

for all $x \in C$ and $u \in F(T)$. Putting $S = TP$, we have that S is a generalized nonexpansive mapping of E into itself such that $F(S) = F(T)$. Indeed, it is obvious that $F(S) = F(T)$. We also have that for any $x \in E$ and $u \in F(S) = F(T)$,

$$\phi(Sx, u) = \phi(TPx, u) \leq \phi(Px, u) \leq \phi(x, u).$$

So, S is a generalized nonexpansive mapping of E into itself such that $F(S) = F(T)$. From Theorems 2.9 and 2.4, there exists the sunny generalized nonexpansive retraction R of E onto $F(T)$. From Theorem 2.7, this retraction R is characterized by

$$Rx = \arg\min_{u \in F(T)} \phi(x, u).$$

We also know from Theorem 2.5 that

$$0 \leq \langle v - Rv, JRv - Ju \rangle, \quad \forall u \in F(T),\ v \in C.$$

Adding up $\phi(Rv, u)$ to both sides of this inequality, we have

$$\begin{aligned} \phi(Rv, u) &\leq \phi(Rv, u) + 2\langle v - Rv, JRv - Ju \rangle \\ &= \phi(Rv, u) + \phi(v, u) + \phi(Rv, Rv) \\ &\quad - \phi(v, Rv) - \phi(Rv, u) \\ &= \phi(v, u) - \phi(v, Rv). \end{aligned} \tag{5.3}$$

Since $\phi(Tz, u) \leq \phi(z, u)$ for any $u \in F(T)$ and $z \in C$, it follows that

$$\begin{aligned} \phi(T^n x, RT^n x) &\leq \phi(T^n x, RT^{n-1} x) \\ &\leq \phi(T^{n-1} x, RT^{n-1} x). \end{aligned}$$

Hence the sequence $\phi(T^n x, RT^n x)$ is nonincreasing. Putting $u = RT^n x$ and $v = T^m x$ with $n \leq m$ in (5.3), we have from Theorem 2.3 that

$$\begin{aligned} g(\|RT^m x - RT^n x\|) &\leq \phi(RT^m x, RT^n x) \\ &\leq \phi(T^m x, RT^n x) - \phi(T^m x, RT^m x) \\ &\leq \phi(T^n x, RT^n x) - \phi(T^m x, RT^m x), \end{aligned}$$

where g is a strictly increasing, continuous and convex real-valued function with $g(0) = 0$. From the properties of g, $\{RT^n x\}$ is a Cauchy sequence.

Therefore $\{RT^n x\}$ converges strongly to a point $q \in F(T)$ since $F(T)$ is closed from Theorem 2.8. Next, consider a fixed $x \in C$ and an arbitrary subsequence $\{S_{n_i} x\}$ of $\{S_n x\}$ convergent weakly to a point v. From the proof of the fixed point theorem (Theorem 3.1) we know that $v \in F(T)$. Rewriting the characterization of the retraction R, we have that

$$0 \leq \left\langle T^k x - RT^k x, JRT^k x - Ju \right\rangle$$

and hence

$$\begin{aligned}\left\langle T^k x - RT^k x, Ju - Jq \right\rangle &\leq \left\langle T^k x - RT^k x, JRT^k x - Jq \right\rangle \\ &\leq \|T^k x - RT^k x\| \cdot \|JRT^k x - Jq\| \\ &\leq K \|JRT^k x - Jq\|,\end{aligned}$$

where K is an upper bound for $\|T^k x - RT^k x\|$. Summing up these inequalities for $k = 0, 1, \ldots, n-1$, we arrive to

$$\left\langle S_n x - \frac{1}{n} \sum_{k=0}^{n-1} RT^k x, Ju - Jq \right\rangle \leq K \frac{1}{n} \sum_{k=0}^{n-1} \|JRT^k x - Jq\|,$$

where $S_n x = \frac{1}{n} \sum_{k=0}^{n-1} T^k x$. Letting $n_i \to \infty$ and remembering that J is continuous, we get

$$\langle v - q, Ju - Jq \rangle \leq 0.$$

This holds for any $u \in F(T)$. Therefore $Rv = q$. But because $v \in F(T)$, we have $v = q$. Thus the sequence $\{S_n x\}$ converges weakly to the point q. □

Using Theorem 5.1, we obtain the following theorems.

Theorem 5.2. *Let E be a uniformly convex Banach space with a Fréchet differentiable norm. Let $T : E \to E$ be an $(\alpha, \beta, \gamma, \delta)$-generalized nonspreading mapping such that $\alpha > \beta$ and $\gamma \leq \delta$. Assume that $F(T) \neq \emptyset$ and let R be the sunny generalized nonexpansive retraction of E onto $F(T)$. Then, for any $x \in E$,*

$$S_n x = \frac{1}{n} \sum_{k=0}^{n-1} T^k x$$

converges weakly to an element q of $F(T)$, where $q = \lim_{n\to\infty} RT^n x$.

Proof. Since the identity mapping I is a sunny generalized nonexpansive retraction of E onto E, E is a nonempty closed, convex sunny generalized nonexpansive retract of E. We also know that $\alpha > \beta$ together with $\gamma \leq \delta$ implies that

$$\phi(Tx, u) \leq \phi(x, u)$$

for all $x \in E$ and $u \in F(T)$. So, we have the desired result from Theorem 5.1. □

Theorem 5.3 (Kocourek, Takahashi and Yao [22]). *Let H be a Hilbert space and let C be a nonempty closed convex subset of H. Let $T : C \to C$ be a generalized hybrid mapping with $F(T) \neq \emptyset$ and let P be the metric projection of H onto $F(T)$. Then, for any $x \in C$,*

$$S_n x = \frac{1}{n} \sum_{k=0}^{n-1} T^k x$$

converges weakly to an element p of $F(T)$, where $p = \lim_{n\to\infty} PT^n x$.

Proof. Since C is a nonempty closed convex subset of H, there exists the metric projection of H onto C. In a Hilbert space, the metric projection of H onto C is equivalent to the sunny generalized nonexpansive retraction of E onto C. On the other hand, a generalized hybrid mapping $T : C \to C$ with $F(T) \neq \emptyset$ is quasi-nonexpansive, i.e.,

$$\phi(Tx, u) = \|Tx - u\|^2 \leq \|x - u\|^2 = \phi(x, u)$$

for all $x \in C$ and $u \in F(T)$. So, we have the desired result from Theorem 5.1. □

Remark. We do not know whether a nonlinear ergodic theorem of Baillon's type for nonspreading mappings holds or not.

6. Weak convergence theorems

In this section, we prove a weak convergence theorem of Mann's iteration [34] for generalized nonspreading mappings in a Banach space. For proving it, we need the following lemma obtained by Takahashi and Yao [37].

Lemma 6.1 (Takahashi and Yao [37]). *Let E be a smooth and uniformly convex Banach space and let C be a nonempty closed subset of E such that JC is closed and convex. Let $T : C \to C$ be a generalized nonexpansive mapping such that $F(T) \neq \emptyset$. Let $\{\alpha_n\}$ be a sequence of real numbers such that $0 \leq \alpha_n < 1$ and let $\{x_n\}$ be a sequence in C generated by $x_1 = x \in C$ and*

$$x_{n+1} = R_C(\alpha_n x_n + (1 - \alpha_n)Tx_n), \quad \forall n \in \mathbb{N},$$

where R_C is a sunny generalized nonexpansive retraction of E onto C. Then $\{R_{F(T)}x_n\}$ converges strongly to an element z of $F(T)$, where $R_{F(T)}$ is a sunny generalized nonexpansive retraction of C onto $F(T)$.

Using Lemma 6.1 and the technique developed by [14, 27], we prove the following theorem.

Theorem 6.2. *Let E be a uniformly convex and uniformly smooth Banach space and let C be a nonempty closed convex sunny generalized nonexpansive retract of E. Let $T : C \to C$ be a generalized nonspreading mapping with $F(T) \neq \emptyset$ such that $\phi(Tx, u) \leq \phi(x, u)$ for all $x \in C$ and $u \in F(T)$. Let R be the sunny generalized nonexpansive retraction of E onto $F(T)$. Let $\{\alpha_n\}$ be a sequence of real numbers such that* $0 \leq \alpha_n < 1$ *and* $\liminf_{n\to\infty} \alpha_n(1-\alpha_n) > 0$. *Then, a sequence $\{x_n\}$ generated by $x_1 = x \in C$ and*

$$x_{n+1} = \alpha_n x_n + (1-\alpha_n)Tx_n, \quad \forall n \in \mathbb{N}$$

converges weakly to $z \in F(T)$, where $z = \lim_{n\to\infty} Rx_n$.

Proof. Let $m \in F(T)$. By the assumption, we know that T is a generalized nonexpansive mapping of C into itself. So, we have

$$\begin{aligned}
\phi(x_{n+1}, m) &= \phi(\alpha_n x_n + (1-\alpha_n)Tx_n, m) \\
&\leq \alpha_n \phi(x_n, m) + (1-\alpha_n)\phi(Tx_n, m) \\
&\leq \alpha_n \phi(x_n, m) + (1-\alpha_n)\phi(x_n, m) \\
&= \phi(x_n, m).
\end{aligned}$$

So, $\lim_{n\to\infty} \phi(x_n, m)$ exists. So, we have that the sequence $\{x_n\}$ is bounded. This implies that $\{Tx_n\}$ is bounded. Put $r = \sup_{n\in\mathbb{N}}\{\|x_n\|, \|Tx_n\|\}$. Using Lemma 2.2, we have that

$$\begin{aligned}
\phi(x_{n+1}, m) &= \phi(\alpha_n x_n + (1-\alpha_n)Tx_n, m) \\
&\leq \|\alpha_n x_n + (1-\alpha_n)Tx_n\|^2 - 2\langle \alpha_n x_n + (1-\alpha_n)Tx_n, Jm\rangle + \|m\|^2 \\
&\leq \alpha_n \|x_n\|^2 + (1-\alpha_n)\|Tx_n\|^2 - \alpha_n(1-\alpha_n)g(\|Tx_n - x_n\|) \\
&\quad - 2\alpha_n\langle x_n, Jm\rangle - 2(1-\alpha_n)\langle Tx_n, Jm\rangle + \|m\|^2 \\
&= \alpha_n(\|x_n\|^2 - 2\langle x_n, Jm\rangle + \|m\|^2) \\
&\quad + (1-\alpha_n)(\|Tx_n\|^2 - 2\langle Tx_n, Jm\rangle + \|m\|^2) \\
&\quad - \alpha_n(1-\alpha_n)g(\|Tx_n - x_n\|) \\
&= \alpha_n\phi(x_n, m) + (1-\alpha_n)\phi(Tx_n, m) - \alpha_n(1-\alpha_n)g(\|Tx_n - x_n\|) \\
&\leq \alpha_n\phi(x_n, m) + (1-\alpha_n)\phi(x_n, m) - \alpha_n(1-\alpha_n)g(\|Tx_n - x_n\|) \\
&= \phi(x_n, m) - \alpha_n(1-\alpha_n)g(\|Tx_n - x_n\|).
\end{aligned}$$

Then, we obtain

$$\alpha_n(1-\alpha_n)g(\|Tx_n - x_n\|) \le \phi(x_n, m) - \phi(x_{n+1}, m).$$

From the assumption of $\{\alpha_n\}$, we have

$$\lim_{n\to\infty} g(\|Tx_n - x_n\|) = 0.$$

Since E is reflexive and $\{x_n\}$ is bounded, there exists a subsequence $\{x_{n_i}\}$ of $\{x_n\}$ such that $x_{n_i} \rightharpoonup v$ for some $v \in C$. Since E is uniformly convex and $\lim_{n\to\infty} \|Tx_n - x_n\| = 0$, we have from Proposition 4.1 that v is a fixed point of T. Let $\{x_{n_i}\}$ and $\{x_{n_j}\}$ be two subsequences of $\{x_n\}$ such that $x_{n_i} \rightharpoonup u$ and $x_{n_j} \rightharpoonup v$. We know that $u, v \in F(T)$. Put $a = \lim_{n\to\infty}(\phi(x_n, u) - \phi(x_n, v))$. Since

$$\phi(x_n, u) - \phi(x_n, v) = 2\langle x_n, Jv - Ju\rangle + \|u\|^2 - \|v\|^2,$$

we have $a = 2\langle u, Jv - Ju\rangle + \|u\|^2 - \|v\|^2$ and $a = 2\langle v, Jv - Ju\rangle + \|u\|^2 - \|v\|^2$. From these equalities, we obtain

$$\langle u - v, Ju - Jv\rangle = 0.$$

Since E is strictly convex, it follows that $u = v$; see [32]. Therefore, $\{x_n\}$ converges weakly to an element u of $F(T)$. On the other hand, we know from Lemma 6.1 that $\{R_{F(T)}x_n\}$ converges strongly to an element z of $F(T)$. From Lemma 2.5, we also have

$$\langle x_n - R_{F(T)}x_n, JR_{F(T)}x_n - Ju\rangle \ge 0.$$

So, we have $\langle u - z, Jz - Ju\rangle \ge 0$. Since J is monotone, we also have $\langle u - z, Jz - Ju\rangle \le 0$. So, we have $\langle u - z, Jz - Ju\rangle = 0$. Since E is strictly convex, we have $z = u$. This completes the proof. □

Using Theorem 6.2, we can prove the following weak convergence theorems.

Theorem 6.3. *Let E be a uniformly convex and uniformly smooth Banach space. Let $T : E \to E$ be an $(\alpha, \beta, \gamma, \delta)$-generalized nonspreading mapping such that $\alpha > \beta$ and $\gamma \le \delta$. Assume that $F(T) \neq \emptyset$ and let R be the sunny generalized nonexpansive retraction of E onto $F(T)$. Let $\{\alpha_n\}$ be a sequence of real numbers such that $0 \le \alpha_n < 1$ and $\liminf_{n\to\infty} \alpha_n(1-\alpha_n) > 0$. Then, a sequence $\{x_n\}$ generated by $x_1 = x \in C$ and*

$$x_{n+1} = \alpha_n x_n + (1-\alpha_n)Tx_n, \quad \forall n \in \mathbb{N}$$

converges weakly to $z \in F(T)$, where $z = \lim_{n\to\infty} Rx_n$.

Proof. Since the identity mapping I is a sunny generalized nonexpansive retraction of E onto E, E is a nonempty closed, convex sunny generalized nonexpansive retract of E. We also know that $\alpha > \beta$ together with $\gamma \le \delta$ implies that

$$\phi(Tx, u) \le \phi(x, u)$$

for all $x \in E$ and $u \in F(T)$. So, we have the desired result from Theorem 6.2. □

Theorem 6.4 (Kocourek, Takahashi and Yao [22]). *Let H be a Hilbert space and let C be a nonempty closed convex subset of H. Let $T : C \to C$ be a generalized hybrid mapping with $F(T) \neq \emptyset$ and let P be the metric projection of H onto $F(T)$. Let $\{\alpha_n\}$ be a sequence of real numbers such that $0 \le \alpha_n < 1$ and $\liminf_{n\to\infty} \alpha_n(1-\alpha_n) > 0$. Then, a sequence $\{x_n\}$ generated by $x_1 = x \in C$ and*

$$x_{n+1} = \alpha_n x_n + (1-\alpha_n)Tx_n, \quad \forall n \in \mathbb{N}$$

converges weakly to $z \in F(T)$, where $z = \lim_{n\to\infty} Px_n$.

Proof. Since C is a nonempty closed convex subset of H, there exists the metric projection of H onto C. In a Hilbert space, the metric projection of H onto C is equivalent to the sunny generalized nonexpansive retraction of E onto C. On the other hand, a generalized hybrid mapping $T : C \to C$ with $F(T) \neq \emptyset$ is quasi-nonexpansive, i.e.,

$$\phi(Tx, u) = \|Tx - u\|^2 \le \|x - u\|^2 = \phi(x, u)$$

for all $x \in C$ and $u \in F(T)$. So, we have the desired result from Theorem 6.2. □

Remark. We do not know whether a weak convergence theorem of Mann's type for nonspreading mappings holds or not.

References

1. Alber, Y.I.: Metric and generalized projections in Banach spaces: properties and applications. In: Kartsatos, A.G. (ed.) Theory and Applications of Nonlinear Operators of Accretive and Monotone Type, pp. 15–50. Marcel Dekker, New York (1996)
2. Aoyama, K., Kohsaka, F., Takahashi, W.: Three generalizations of firmly nonexpansive mappings: their relations and continuity properties. J. Nonlinear Convex Anal. **10**, 131–147 (2009)

3. Baillon, J.-B.: Un theoreme de type ergodique pour les contractions non lineaires dans un espace de Hilbert. C. R. Acad. Sci. Paris Ser. A-B **280**, 1511–1514 (1975)
4. Blum, E., Oettli, W.: From optimization and variational inequalities to equilibrium problems. Math. Stud. **63**, 123–145 (1994)
5. Browder, F.E.: Nonexpansive nonlinear operators in a Banach space. Proc. Natl. Acad. Sci. USA **54**, 1041–1044 (1965)
6. Browder, F.E.: Convergence theorems for sequences of nonlinear operators in Banach spaces. Math. Z. **100**, 201–225 (1967)
7. Combettes, P.L., Hirstoaga, A.: Equilibrium problems in Hilbert spaces. J. Nonlinear Convex Anal. **6**, 117–136 (2005)
8. Goebel, K., Kirk, W.A.: Topics in Metric Fixed Point Theory. Cambridge University Press, Cambridge (1990)
9. Goebel, K., Reich, S.: Uniform Convexity, Hyperbolic Geometry, and Nonexpansive Mappings. Marcel Dekker, New York (1984)
10. Ibaraki, T., Takahashi, W.: Convergence theorems for new projections in Banace spaces. RIM Kokyuroku **1484**, 150–160 (2006)
11. Ibaraki, T., Takahashi, W.: A new projection and convergence theorems for the projections in Banach spaces. J. Approx. Theory **149**, 1–14 (2007)
12. Ibaraki, T., Takahashi, W.: Mosco convergence of sequences of retracts of four nonlinear projections in Banach spaces. In: Takahashi, W., Tanaka, T. (eds.) Nonlinear Analysis and Convex Analtsis, pp. 139–147. Yokohama Publishers, Yokohama (2007)
13. Ibaraki, T., Takahashi, W.: Weak and strong convergence theorems for new resolvents of maximal monotone operators in Banach spaces. Adv. Math. Econ. **10**, 51–64 (2007)
14. Ibaraki, T., Takahashi, W.: Weak convergence theorem for new nonexpansive mappings in Banach spaces and its applications. Taiwanese J. Math. **11**, 929–944 (2007)
15. Ibaraki, T., Takahashi, W.: Fixed point theorems for new nonlinear mappings of nonexpansive type in Banach spaces. J. Nonlinear Convex Anal. **10**, 21–32 (2009)
16. Ibaraki, T., Takahashi, W.: Generalized nonexpansive mappings and a proximal-type algorithm in Banach spaces. Nonlinear Analysis and Optimization I: Nonlinear Analysis, Contemporary Mathematics. Amer. Math. Soc. Providence, RI, **513**, 169–180 (2010)
17. Iemoto, S., Takahashi, W.: Approximating fixed points of nonexpansive mappings and nonspreading mappings in a Hilbert space. Nonlinear Anal. **71**, 2082–2089 (2009)
18. Itoh, S., Takahashi, W.: The common fixed point theory of singlevalued mappings and multivalued mappings. Pac. J. Math. **79**, 493–508 (1978)

19. Kamimura, S., Takahashi, W.: Weak and strong convergence of solutions to accretive operator inclusions and applications. Set-Valued Anal. **8**, 361–374 (2000)
20. Kamimura, S., Takahashi, W.: Strong convergence of a proximal-type algorithm in a Banach apace. SIAM J. Optim. **13**, 938–945 (2002)
21. Kamimura, S., Kohsaka, F., Takahashi, W.: Weak and strong convergence theorems for maximal monotone operators in a Banach space. Set-Valued Anal. **12**, 417–429 (2004)
22. Kocourek, P., Takahashi, W., Yao, J.-C.: Fixed point theorems and weak convergence theorems for generalized hybrid mappings in Hilbert spaces. Taiwanese J. Math. **14**, 2497–2511 (2010)
23. Kohsaka, F., Takahashi, W.: Strong convergence of an iterative sequence for maximal monotone operators in a Banach space. Abstr. Appl. Anal. **2004**, 239–249 (2004)
24. Kohsaka, F., Takahashi, W.: Generalized nonexpansive retractions and a proximal-type algorithm in Banach spaces. J. Nonlinear Convex Anal. **8**, 197–209 (2007)
25. Kohsaka, F., Takahashi, W.: Existence and approximation of fixed points of firmly nonexpansive-type mappings in Banach spaces. SIAM J. Optim. **19**, 824–835 (2008)
26. Kohsaka, F., Takahashi, W.: Fixed point theorems for a class of nonlinear mappings related to maximal monotone operators in Banach spaces. Arch. Math. **91**, 166–177 (2008)
27. Matsushita, S., Takahashi, W.: Weak and strong convergence theorems for relatively nonexpansive mappings in Banach spaces. Fixed Point Theory Appl. **2004**, 37–47 (2004)
28. Matsushita, S., Takahashi, W.: A strong convergence theorem for relatively nonexpansive mappings in a Banach space. J. Approx. Theory **134**, 257–266 (2005)
29. Reich, S.: A weak convergence theorem for the alternating method with Bregman distances. In: Kartsatos, A.G. (ed.) Theory and Applications of Nonlinear Operators of Accretive and Monotone Type, pp. 313–318. Marcel Dekker, New York (1996)
30. Takahashi, W.: A nonlinear ergodic theorem for an amenable semigroup of nonexpansive mappings in a Hilbert space. Proc. Am. Math. Soc. **81**, 253–256 (1981)
31. Takahashi, W.: Convex Analysis and Approximation of Fixed Points (Japanese). Yokohama Publishers, Yokohama (2000)
32. Takahashi, W.: Nonlinear Functional Analysis. Yokohama Publishers, Yokohama (2000)
33. Takahashi, W.: Proximal point algorithms and four resolvents of nonlinear operators of monotone type in Banach spaces. Taiwanese J. Math. **12**, 1883–1910 (2008)

34. Takahashi, W.: Introduction to Nonlinear and Convex Analysis. Yokohama Publishers, Yokohama (2009)
35. Takahashi, W.: Fixed point theorems for new nonexpansive mappings in a Hilbert space. J. Nonlinear Convex Anal. **11**, 79–88 (2010)
36. Takahashi, W., Yao, J.C.: Fixed point theorems and ergodic theorems for nonlinear mappings in a Hilbert space. Taiwanese J. Math. (to appear)
37. Takahashi, W., Yao, J.C.: Weak and strong convergence theorems for positively homogeneous nonexpansive mappings in Banach spaces. Taiwanese J. Math. (to appear)
38. Takahashi, W., Yao, J.C.: Nonlinear operators of monotone type and convergence theorems with equilibrium problems in Banach spaces. Taiwanese J. Math. (to appear)
39. Xu, H.K.: Inequalities in Banach spaces with applications. Nonlinear Anal. **16**, 1127–1138 (1981)

Adv. Math. Econ. 15, 89–122 (2011)

Advances in
MATHEMATICAL
ECONOMICS

On the perception and representation of economic quantity in the history of economic analysis in view of the Debreu conjecture*

Akira Yamazaki†

Professor, School and Graduate School of Economics, Meisei University,
2-1-1 Hodokubo, Hino, Tokyo 191-8506, Japan
Emeritus Professor, Graduate School of Economics, Hitotsubashi University,
2-1, Naka, Kunitachi, Tokyo 186-8601, Japan
(e-mail: yamazaki@econ.hit-u.ac.jp)

Received: October 19, 2010
Revised: November 3, 2010

JEL classification: B13, B23, B3

Mathematics Subject Classification (2010): 62P20, 91B02, 91B42, 91B50

Abstract. In this paper, the perception and representation of economic quantity found in the works of Augustin Cournot, Leon Walras, Vilfreto Pareto, and Alfred Marshall will be discussed. An interpretation of the perception and representation of economic quantities and economic variables, specifically relating to the concept of demand, in the fundamental theoretical framework of general equilibrium theory will be provided, particularly from the vantage point of the Debreu conjecture.

Key words: Debreu conjecture, economic quantity, demand, general equilibrium theory, A. Cournot, L. Walras, V. Pareto, A. Marshall

* This paper is the English version of an article written originally in Japanese. The Japanese version was published in both the Mita Journal of Economics 103(1), 2010, and in *Epimetheus in Economics*, edited by T. Maruyama (Chisen Shokan: Tokyo, 2010). The author is heavily indebted to Professor Philip Culbertson for suggesting many elegant improvements to the manuscript in English.

† I acknowledge the support for this research through a JSPS Grant in Aid for Scientific Research, no. 22530188.

S. Kusuoka, T. Maruyama (eds.), *Advances in Mathematical Economics Volume 15*, 89
DOI: 10.1007/978-4-431-53930-8_5,

1. Introduction

General equilibrium theory, as initiated by Leon Walras [18], provides a solid theoretical framework for modern economic analysis. From a purely theoretical point of view, its fundamental mathematical structure was established through a series of papers in the 1950s and 1960s, in which the existence of an equilibrium was proven.

The purpose of this paper is to provide an interpretation of the perception and representation of economic quantities and economic variables, specifically relating to the concept of demand, in the fundamental theoretical framework of general equilibrium theory especially from the vantage point of the Debreu conjecture [8, p. 614].

1.1. The Debreu conjecture

In his presidential address given at the September 1971 meeting of the Econometric Society in Barcelona, Gerard Debreu proffered the following conjecture to the members of the society:

> One expects that if the measure ν is suitably diffused over the space A (of economic agents' characteristics), integration over A of the demand *correspondences* of the agents will yield a total *demand function*, possibly even a total demand function of class C^1.

This address was later published as a paper entitled "Smooth Preferences" (Debreu [8]). In his paper, Debreu introduced a differentiable structure in a space of preference relations, and clarified conditions under which demand functions are differentiable. His above conjecture posed a question to economic theorists: *If* a distribution over preference relations and initial endowments of economic agents that constitute an economy is "diffused" in some appropriate sense so that there is a sufficient variety of characteristics of economic agents or consumers, *then* total demand might become a function, or even a continuously differentiable function, notwithstanding individual demands being correspondences.

The Debreu conjecture had a substantial impact on research projects emanating from the University of California, Berkeley, in the 1970s and 1980s. Since the significance of this conjecture is related essentially to the perception and representation of economic quantity, the purpose of this paper is to review the history of economic analysis from the vantage point of Debreu's conjecture.

But first, the theoretical framework within which this conjecture was posed must be confirmed.

1.2. The related theoretical framework

The discussion in this paper presupposes the theoretical framework of general equilibrium theory as established through a series of work in the 1950s and 1960s. The basic components of the fundamental model representing this framework are as follows:

1. A population of economic agents.
2. A set of commodities.
3. Prices of commodities.
4. A market equilibrium.

Each of these factors assumes the following mathematical representation: A population of economic agents is given by either a finite set or an atomless measure space, i.e., a continuum of economic agents. A set of commodities or the commodity space is given by an ℓ-dimensional Euclidean space $\mathbb{R}^\ell$ if the number of commodities is given by a positive natural number ℓ, and by a linear topological space L if the number of commodities is not finite. Prices of commodities are represented by the dual space of the basic commodity space. In addition, the economic variables that form a market equilibrium are the total demand correspondence (or function) and the total supply correspondence (or function).

The basis of the model is the commodities that form the commodity space and their quantitative representation. One must be clear about how different kinds of commodities are distinguished and how their quantities are expressed. While a number of economists had discussed the typology of commodities, Debreu [7] was most explicit about the way in which commodities must be specified. He distinguished commodities according to (1) their physical characteristics, (2) dates at which they are available, (3) locations where they are available, and (4) events at which they are available ([7, Chaps. 2 and 7]). Debreu's way of specification has subsequently become the standard manner of distinguishing commodities in the literature of economic theory.

Quantities of particular commodities thus distinguished are represented by "any real numbers" without further mathematical restrictions. Hence, as explained earlier, the commodity space is given by an ℓ-dimensional Euclidean space $\mathbb{R}^\ell$ when the number of commodities is limited to a finite natural number ℓ.

1.3. The perception of economic quantity

Notwithstanding the fact that economic quantity has become represented by any real number in the standard theoretical literature since the 1950s, there has been a different perception of economic quantity itself. For instance, even in Debreu's book [7, p. 30], one can find a statement expressing such a view.

In his case, he does not object to using any real number to represent the quantity of commodities such as wheat or liquids. He admits, however, that in the case of commodities such as trucks, their quantity must be an integer. Here is a quote from Debreu [7, p. 30]:

> The quantity of certain kind of wheat is expressed by a number of bushels which can satisfactorily be assumed to be any (non-negative) real number. ...
> ... A quantity of well-defined trucks is an integer; but it will be assumed instead that this quantity can be any real number.

Despite his perception of such an economic quantity, in his theoretical analysis he goes ahead to take the commodity space as $\mathbb{R}^\ell$, and allows economic transactions for real number units of various commodities.

As one sees in his "excuse" quoted below, it is merely for the purpose of seeking analytical simplification.[1]

> This assumption of *perfect divisibility is imposed by the present stage of development of economics*; it is quite acceptable for an economic agent producing or consuming a large number of trucks. Similar goods are machine tools, linotypes, cranes, Bessemer converters, houses,...[2]

The "perfect divisibility" in the quotation references the sense of economics that is different from "divisibility" in the sense of mathematics.[3]

The representation of prices that reflect the valuation of commodities is not discussed in this paper. In the next section, a path in the history of economic analysis concerning the perception and the representation of economic quantity will be traced.

---

[1] In the latter part of the quotation, Debreu does not give an explicit explanation for why the assumption of the perfect divisibility of commodities is acceptable for an economic agent who is transacting a large number of commodities. I would presume that Debreu's perception in this regard is similar to that of Pareto or Walras, which I will take up in the later part of this paper.

[2] Italics are mine.

[3] In general, "divisibility" in arithmetic means that a rational number can be expressed by finite digits. Thus, as an analytical concept, divisibility in the sense of economics stands in polar relationship to its sense in mathematics.

2. The perception and representation of economic quantity in selected literature of economic analysis

In this section, the perception and representation of economic quantity found in the works of Augustin Cournot, Leon Walras, Vilfreto Pareto, and Alfred Marshall will be discussed. Joseph A. Schumpeter admits in his book [17] that these theorists have contributed significantly to the field of economic analysis.

2.1. Perception and representation of economic quantity in Cournot's theory

In establishing a theoretical foundation for explaining the way in which exchange values of commodities are determined in markets, Cournot [4] set out to clarify the concept of demand in markets by introducing an abstract mathematical concept of a function to represent market demand.

> 21. Admettons donc que le débit ou la demande annuelle D est, pour chaque denreé, une fonction particulière $F(p)$ du prix p de cette denreé. ... [4]

It appears that there were no clear-cut distinctions, at the time when he wrote his book, among concepts such as demand, quantity demanded, supply, quantity supplied, equilibrium quantity, etc., which contributed to some confusion in the literature.[5]

Demand function and continuity

Through the use of the mathematical concept of a function, Cournot perceived market demand as a demand function, probably for the first time in the literature of economic theory, and postulated the "continuity" property of the function at an abstract level:

> 22. Nous admettrons que la fonction $F(p)$ qui exprime la loi de la demande ou du débit est une *fonction continue*, c'est-à-dire une fonction qui ne passe pas soudainement d'une valeur à une autre, mais qui prend dans l'intervalle toutes les valeurs intermédiaires. *Il en pourrait être autrement si le nombre des consommateurs était très-limité:*

[4] Cournot [4, p. 37]. The English translation is as follows (Cournot [5, p. 47]): Let us admit therefore that the annual sales or demand D is, for each article, a particular function $F(p)$ of the price p of such article. ...

[5] See Cournot [4, p. 36]: En outre, qu'entend-on par la quantité demandée? Ce n'est sans doute pas celle qui se débite effectivement sur la demande des acheteurs; ...

> ainsi, dans tel ménage, on pourra consommer précisément la même quantité de bois de chauffage, que le bois soit à 10 francs ou à 15 francs le stère; et l'on pourra réduire brusquement la consommation d'une quantité notable, si le prix du stère vient à dépasser cette dernière somme.[6]

The way that Cournot explained the behavior of market demand in the above quotation is that changes in quantity demanded of a market demand function that represents "the law of demand or sale"[7] occur in such a way that one value does not pass to another suddenly, but by taking all intermediate values. In other words, he postulated the continuity property of a market demand function by requiring, so to speak, "the theorem of intermediate values" to hold.

This does not mean, however, that Cournot as a mathematician understood continuous functions in this way. My presumption is that he probably thought his contemporary economists would understand more easily if he explained continuity property by using one of the properties of a continuous function in modern mathematics. In fact, we could quote the following from one of his mathematics books:

> Le caractère propre d'une fonction continue consiste en ce que l'on peu toujours assigner à l'une des variables des valeurs assez voisines pour que la différence entre les valeurs correspondantes de la fonction qui en dépend, tombe au-dessous de toute grandeurs donnée.[8]

[6] Cournot [4, pp. 38–39]; italics added. The English translation is as follows (Cournot [5, pp. 49–50]): 22. We will assume that the function $F(p)$, which expresses the law of demand or of the market, is a *continuous function*, *i.e.* a function which does not pass suddenly from one value to another, but which takes in passing all intermediate values. *It might be otherwise if the number of consumers were very limited*: thus in a certain household the same quantity of firewood will possibly be used whether wood costs 10 francs or 15 francs the stere, and the consumption may suddenly be diminished if the price of the stere rises above the latter figure.

[7] Cournot clearly states that he uses the word "the demand" (la demande) and "the sale"(le débit) synonymously. Cournot [4, pp. 38–39]: Le débit ou la demande (car pour nous ces deux mots sont synonymes, et nous ne voyons pas sous quel rappport la théorie aurait à tenir compte d'une demande qui n'est pas suivie de débit), le débit ou la demande, disons-nous, croît en général quand le prix décroît. The English translation is as follows (Cournot [5, p. 46]): The sales or the demand (for to us these two words are synonymous, and we do not see for what reason theory need take account of any demand which does not result in a sale) – the sales or the demand generally, we say, increases when the price decreases.

[8] Cournot [6, p. 3]. My English translation is as follows: The proper characteristic of a continuous function is that one can always assign, to each of the variables, values sufficiently close to each other such that the difference of the values taken by the function, on which they depend, falls within any given magnitude.

In this quotation, the way he characterized a continuous function can be understood to represent exactly the one using ε-δ in modern analysis or the one using neighborhoods in modern topological analysis.

Although Cournot pursued his analysis based on what he called "the continuity" of a market demand function, he clearly admitted, by way of an example, that the demand function might not well be continuous if the number of consumers is limited. The ground on which Cournot reasoned that individual demand functions are discontinuous was not the indivisibility of commodities such that their quantitative representation should be limited to integers such as $1, 2, \ldots$. The above example from Cournot does not exhibit a behavior of quantity demanded that declines slightly as its price moves up slightly, but that quantity demanded declines abruptly as its price increases to a certain level. If we were to express Cournot's perception in present-day mathematical terms, he would say that it is at most semi-continuous and might not deny its semi-continuity.[9]

It is possible, however, to interpret the meaning of "continuity" in Cournot's book in a different sense. In fact, he continued further discussion of a continuous function on the same page quoted above:

> Si la fonction $F(p)$ est continue, elle jouira de la propriété commune à toutes les fonctions de cette nature, et sur laquelle reposent tant d'applications importantes de l'analyse mathématique: *les variations de la demande seront sensiblement proportionnelles aux variations du prix, tant que celles-ci seront de petites fractions du prix originaire*. D'ailleurs, ces variations seront de signes contraires, c'est-à-dire qu'à une augmentation de prix correspondra une diminution de la demande.[10]

Here, he explained the meaning of the continuity of the demand function $F(p)$. In particular, the quotation in italics shows explicitly that the function $F(p)$ allows "locally linear approximation." In other words, one can deduce that he in fact postulated the *differentiability* of the function in the name of the continuity of the function. This interpretation will be taken up again later with regard to the Debreu conjecture.

---

[9] In the quotation from Cournot, he may be interpreted as denying the lower-semicontinuity of individual demand functions. This will be discussed further in Sect. 4, in addition to explaining the concept of semi-continuity.

[10] Cournot [4, p. 39]. The English translation is as follows (Cournot [5, p. 50]): If the function $F(p)$ is continuous, it will have the property common to all functions of this nature, and on which so many important applications of mathematical analysis are based: *the variations of the demand will be sensibly proportional to the variations in price so long as these last are small fractions of the original price.* Moreover, these variations will be of opposite signs, i.e., an increase in price will correspond with a diminution of the demand.

One of the reasons that he had an interest in the continuity of the demand function is that if the function as the object of main analysis is "continuous," it is very convenient for analytical purposes, as analytical methods in mathematics can be applied.

Continuity by aggregation effects

Thus admitting in general the discontinuity of individual demand functions on the one hand, Cournot claims, on the other hand, that one should regard the market demand function as being continuous. The most interesting part of his views with respect to his perception of economic variables is the following quotation, which could be read as his recognition of the continuity of the market demand function that is induced by aggregation effect. This quotation follows the one cited earlier above.

> Mais plus le marché s'étendra, plus les combinaisons des besoins, des fortunes ou même des caprices, seront variées parmi les consommateurs, plus la fonction $F(p)$ approchera de varier avec p d'une manière continue. Si petite que soit la variation de p, il se trouvera des consommateurs placés dans une position telle que le léger mouvement de hausse ou de baisse imprimé á la denrée influera sur leur consommation, les engagera à s'imposer quelques privations, ou à rèduire leurs exploitations industrielles, ou à substituer une autre denrée à la denrée renchérie, par exemple, la houille au bois, ou l'anthracite à la houille.[11]

Cournot's perceptive recognition in this quotation is that as the market extends wider and the combination of needs, wealth, and preferences are varied among consumers, the closer the market demand function $F(p)$ comes to varying continuously with respect to the market price p. One may note, however, that for Cournot, the market demand function is somewhat statistically conceived and is thought to be derived from available market statistical data. Since, in particular, he did not try to derive the market demand function on a purely theoretical ground, he did not go into an explicit discussion about factors having a bearing on the market demand function, such as the preferences, wealth, etc. of consumers.

[11] Cournot [4, p. 39]. The English translation is as follows (Cournot [5, p. 50]): But the wider the market extends, and the more the combinations of needs, of fortunes, or even of caprices, are varied among consumers, the closer the function $F(p)$ will come to varying with p in a continuous manner. However little may be the variation of p, there will be some consumers so placed that the slight rise or fall of the article will affect their consumptions, and will lead them to deprive themselves in some way or to reduce their manufacturing output, or to substitute something else for the article that has grown dearer, as, for instance, coal for wood or anthracite for soft coal.

2.2. Perception and representation of economic quantity in Walras

Discontinuity of demand curve

Like Cournot, so Walras perceived and very clearly stated that nothing indicates that individual demand curves are continuous. On the contrary, they are discontinuous in general, and in reality their graphs take the form of step curves ("la forme de la courbe en escalier").

> Rien n'indique que les courbes ou les equations partielles $a_{d,1}a_{p,1}$, $d_a = f_{a,1}(p_a)$ et autres soient *continues*, c'est-à-dire qu'une augmentation infiniment petite de p_a y produise une diminution infiniment petite de d_a. Au contraire, ces fonctions seront souvent discontinues. Pour ce qui concerne l'avoine, par exemple, il est certain que notre premier porteur de blé réduira sa demande non pas au fur et à mesure de l'elévation du prix, mais d'une faccon en quelque sorte intermittente chaque fois qu'il se décidera à avoir un cheval de moins dans son écurie. Sa courbe de demande partielle aura donc en réalité la forme de la courbe en escalier passant au point a .... Il en sera de même de tous les autres.[12]

His explanation of a function $d_a = f_{a,1}(p_a)$ as being continuous ("continue") is that an infinitesimally small ("infiniment petite") increase in the price p_a induces an infinitesimally small decrease in d_a. It should be noted that Cournot explained the continuity of the demand function without any appeal to the expression of "infinitesimal smallness."[13]

Immediately following the above quotation, as shown below, Walras claimed that the aggregate or the total demand curve that sums up individual demand curves could possibly be deemed to be continuous by virtue of what he called *the law of large numbers*.

---

[12] Walras [18, pp. 57–58]. The English translation is as follows (Walras [19, p. 169]): There is nothing to indicate that the individual demand curves $a_{d,1}a_{p,1}$ and so on, or the individual demand equations $d_a = f_{a,1}(p_a)$ and so on, are *continuous*, in other words that an infinitesimally small increase in p_a produces an infinitesimally small decrease in d_a. On the contrary, these functions are often discontinuous. In the case of oats, for example, surely our first holder of wheat will not reduce his demand gradually as the price rises, but he will do it in some intermittent way every time he decides to keep one horse less in his stable. His individual demand curve will, in reality, take the *form of a step curve* passing through the point a .... All the other individual demand curves will take the same general form.

[13] In his book on calculus [6], Cournot explained a differentiable function by using the word "infinitesimal smallness." I tend to presume that Cournot [4] avoided the use of these words on purpose. However, in case of Walras, I believe he did not intend to claim the differentiability of the demand function in this expression.

> Et cependant, la courbe totale $A_d A_p \ldots$ peut, en vertu de la *loi* dite *des grands nombres*, étre considérée comme sensiblement continue. En effect, lorsqu'il se produira une augmentation trés petite du prix, l'un au moins des porteurs de (B), *sur le grand nombre*, arrivant à la limite qui l'oblige à se priver d'un cheval, il se produira aussi une diminution trés petite de la demande totale.[14]

Since there is no further explanation pertaining to the law of large numbers to which Walras alluded, it is not clear at all what he meant by his assertion. The law of large numbers, in the sense of statistics or probability theory, asserts that a sample mean converges under a set of specific conditions to the mean of the population as the size of a sample increases. Thus, it is not proper to interpret what Walras called the law of large numbers in the sense of statistics or probability theory. Accordingly, care should be taken in identifying what Walras meant by the law of large numbers.

2.3. Perception and representation of economic quantity in Pareto

Indivisibility of commodities

With regard to economic quantities, Pareto basically perceived units of commodities to be essentially indivisible, and thus he considered units of commodities to be measured by integers.

> 65. Variazioni continue e variazioni discontinue. – Le curve di indifferenza ed i sentieri potrebbero essere discontinui; anzi nel concreto sono realmente tali, cioè le variazioni delle quantità avvengono in modo discontinuo. Un individuo, dallo stato n cui ha 10 fazzoletti passa ad uno stato in cui ne ha 11, e non già agli stati intermedii, in cui avrebbe per esempio 10 fazzoletti e un centesimo di fazzoletto; 10 fazzoletti e due centesimi, ecc.[15]

---

[14] Walras [18, p. 58]. The English translation is as follows (Walras [19, p. 169]): And yet the aggregate demand curve $A_d A_p \ldots$ can, for all practical purposes, be considered as continuous by virtue of the so-called *law of large numbers*. In fact, whenever a very small increase in price takes place, at least one of the holders of (B), out of a large number of them, will then reach the point of being compelled to keep one horse less, and thus a very small diminution in the total demand for (A) will result.

[15] Pareto [15, p. 169]. The English translation is as follows (Pareto [16, p. 122]): 65. Continuous variations and discontinuous variations. The indifference curves and the paths could be discontinuous, and they are in reality. That is, the variations in the quantities occur in a discontinuous fashion. An individual passes from a state in which he has 10 handkerchiefs to a state in which he has 11, and not through intermediate states in which he would have, for example, 10 and 1/100 handkerchiefs, 10 and 2/100 handkerchiefs, etc.

Here, the expression of the “continuous variation” (variazioni continue), or rather, “discontinuous variation” (variazioni discontinue), refers to indifference curves that represent a preference relation in the commodity space. This is in contrast to the previous discussion of Cournot, in which case it was an expression about demand functions. Therefore, Pareto’s argument in the above quotation is not directed to the question of the continuity of demand functions, but is concerned with a question of whether indifference curves can be drawn as continuous or not. In other words, we need to interpret his arguments as being based on the explicit perception of the indivisibility of commodities.

Notwithstanding his fundamental perception of the indivisibility of economic quantities, his arguments following the above quotation proceeded, as below:

> Per avvicinarsi al concreto, occorrerebbe dunque considerare variazioni finite, ma c’è una difficoltà tecnica.
> I problemi aventi per oggetto quantità che variano per gradi infinitesimi sono molto più facili a trattarsi che i problemi in cui le quantità hanno variazioni finite. Giova dunque, ogni qualvolta ciò si possa fare, sostituire quelli a questi; e così effettivamente si opera in tutte le scienze fisico naturali. Si sa che per tal modo si fa un errore; ma si può trascurare, sia quando è piccolo in modo assoluto, sia quando è mimore di altri inevitabili, il che rende inutile di ricercàre da una parte una precisione che sfugge dall’altra. *Tale è appunto il caso per l’economia politica, che considera solo fenomeni medii e che si riferiscono a grandi numeri. Discorriamo dell’individuo, non già per ricercare effettivamente cosa un individuo consuma o produce, ma solo per considerare un elemento di una collettività, e per sommare poi consumo e produzione per molti e molti individui.*[16]

[16] Pareto [15, p. 169]; italics are mine. The English translation is as follows (Pareto [16, p. 123]): In order to come closer to reality, we would have to consider finite variations, but there is a technical difficulty in doing so.

Problems concerning quantities which vary by infinitely small degrees are much easier to solve than problem in which the quantities undergo finite variations. Hence, every time it is possible, we must replace the latter by the former; this is done in all the physiconatural sciences. We know that an error is thereby committed; but it can be neglected either when it is small absolutely, or when it is smaller than other inevitable errors which make it useless to seek a precision which eludes us in other ways. *This is precisely so a in political economy, for there we consider only average phenomena and those involving large numbers. We speak of the individual, not in order actually to investigate what one individual consumes or produces, but only to consider one of the elements of a collectivity and then add up the consumption and the production of a large number of individuals.*

Thus, as in the case of Cournot, Pareto acknowledged that in order to come closer to reality in analysis, one must consider finite variations. Nonetheless, since one faces analytical difficulties in doing so, he proposed replacing quantities going through finite variations ("variazioni finite") with those that vary by infinitesimally small amounts ("gradi infinitesimi"), as is done in the natural sciences. In the first paragraph above, Pareto underlined a technical reason to treat economic quantities as going through continuous variations.

Economic quantity as an "average phenomenon"

In addition to the pretext of its convenience for analysis, and of its conformity with analyses in the physical sciences, in the second paragraph of the above quotation Paerto pointed to a more fundamental ground on which such an analysis could be justified – to wit, when the behavior of individual economic agents such as consumers or producers is analyzed, it is the "average phenomena" ("fenomeni medii") that are examined. Furthermore, the number of economic agents is very large ("grandi numeri"). In such a case, even if the economic quantities that individual consumers or producers face in reality are finite discrete quantities, a possible error resulting from treating them as quantities varying by "infinitesimally small" amounts ("quantità che variano ser gradi infinitesimi"), and hence continuously, can be neglected, as the error is small absolutely, or is smaller than other inevitable errors.

Pareto explained what he called an "average phenomenon" by taking up a concrete example. For instance, in the quotation below, he contended that it would be frivolous to take literally such words as "an individual consumes one and one-tenth watches." Rather, the phrase is to be interpreted as signifying, for example, that "one hundred individuals consume one hundred and ten watches."

> 66. Quando diciamo che un individuo consuma un orologio e un decimo, sarebbe ridicolo il prendere quei termini alla lettera. II decimo dell'orologio è un oggetto sconosciuto e che non ha uso. Ma quei termini esprimono semplicemente che, per esempio, cento individui consumano 110 orologi.
> Quando diciamo che l'equilibrio ha luogo quando un individuo consuma un orologio e un decimo, ciò vuol semplicemente esprimere che l'equilibrio ha luogo quando 100 individui consumano chi uno, chi due o più orologi, e anche punti, in modo che tutti insieme ne consumano 110 circa, e che la media per ciaseuno è 1,1.[17]

[17] Pareto [15, p. 169]. The English translation is as follows (Pareto [16, p. 123]): 66. When we say that an individual consumes one and one-tenth watches, it would be

Pareto went on to remind us that this interpretation of individual behavior as an average phenomenon is not limited to economics, but also prevails in other sciences, such as actuarial science.

> Questo modo non è proprio dell'economia politica, ma appartiene a moltissime scienze. Nelle assicurazioni si discorre di frazioni di viventi; per esempio 27 viventi e 37 centesimi. È pure chiaro che non possono esistere 37 centesimi di un vivente!
> Se non si concede di sosotituire le variazioni continue alle discontinue, conviene rinunciare a dare la teoria della leva. Voi mi dite che una leva a braccia eguali, per esempio una bilancia, è in equilibrio quando porta pesi uguali; io prendo una bilancia che è sensibile solo al centigramma, metto in uno dei piattini un milligramma di più che nell'altro, e vi faccio vedere che, contraddicendo la teoria, sta in equilibrio.
> La bilancia nella quale si pesano i gusti dell'uomo è tale che per alcune merci è sensibile al gramma; per altre solo all'ettogramma; per altre solo al chilogramma, ecc.
> L'unica conclusione da trarne è che da tali bilancie non bisogna richiedere maggiore precisione di quella che possono dare.[18]

---

ridiculous to take those words literally. A tenth of a watch is an unknown object for which we have no use. Rather these words simply signify that, for example, one hundred individuals consume 110 watches.

When we say that equilibrium takes place when an individual consumes one and one-tenth watches, we simply mean that equilibrium takes place when 100 individuals consume – some one, others two or more watches and some even none at all – in such a way that all together they consume about 110, and the average is 1.1 for each.

[18] Pareto [15, p. 169]. The English translation is as follows (Pareto [16, p. 123]): This manner of expression is not peculiar to political economy; it is found in a great number of sciences.

In insurance one speaks of fractions of living persons, for example, 27 and 37 hundredths of living persons. It is quite obvious there is no such thing as thirty-seven hundredths of a living person!

If we did not agree to replace discontinuous variations by continuous variations, the theory of the lever could not be derived. We say that a lever having equal arms, a balance, for example, is in equilibrium when it is supporting equal weights. But I might take a balance which is sensitive to a centigram, put in one of the trays a milligram more than in the other, and state that, contrary to the theory, it remains in equilibrium.

The balance in which we weigh men's tastes is such that, for certain goods it is sensitive to the gram, for others only to the hectogram, for others to the kilogram, etc.

The only conclusion that can be drawn is that we must not demand from these balances more precision than they can give.

2.4. Perception and representation of economic quantity in Marshall

Discontinuity of individual demands

Marshall [12] acknowledged that there are instances of demands at an individual level exhibiting a behavior that is representative of the general demand of an entire market, such as the demand for tea, and for such individual demands, small changes in price induce corresponding small changes in quantities, thus resulting in a continuous variation. In those instances, he observed that their demands are constant ones and the commodities can be purchased in small quantities.

However, as the following quotation shows, Marshall also maintained that individual demands in general are discontinuous, as did Cournot, Walras, and Pareto. He used the example of watches and hats.

> (Marshall [12, pp. 82–83]): Sect. 5. So far we have looked at the demand of a single individual. And in the particular case of such a thing as tea, the demand of a single person is fairly representative of the general demand of a whole market: for the demand for tea is a constant one; and, since it can be purchased in small quantities, every variation in its price is likely to affect the amount which he will buy. But even among those things which are in constant use, there are many for which the demand on the part of any single individual cannot vary continuously with every small change in price, but can move only by great leaps. For instance, a small fall in the price of hats or watches will not affect the action of every one; but it will induce a few persons, who were in doubt whether or not to get a new hat or a new watch, to decide in favour of doing so.

Continuity of aggregate demand in a large market

Marshall observed that there are many commodities for which individuals have "inconstant, fitful, and irregular" needs. He argued that for this genre of commodities, individual demands are irregular and discontinuous. In spite of the discontinuity of individual demands, he clarified his perception of the continuity of market demand by stating clearly that "in large markets, then – where rich and poor, old and young, men and women, persons of all varieties of tastes, temperaments and occupations are mingled together – the peculiarities in the wants of individuals will compensate one another in a comparatively regular gradation of total demand. Marshall's perception matches perfectly that of Cournot, except that, without doubt, Marshall contemplated a variety of possible factors among individuals that induce a regular and continuous variation in the market demand. The following quotation pertains to this. As his arguments are not limited to the one on the continuity of the market demand, but also refer to the law of demand, the strong influence of Cournot on Marshall is probably indicated.

> (Marshall [12, p. 83]): There are many classes of things the need for which on the part of any individual is inconstant, fitful, and irregular. There can be no list of individual demand prices for wedding-cakes, or the services of an expert surgeon. But the economist has little concern with particular incidents in the lives of individuals. He studies rather "the course of action that may be expected under certain conditions from the members of an industrial group," in so far as the motives of that action are measurable by a money price; and in these broad results the variety and the fickleness of individual action are merged in the comparatively regular aggregate of the action of many. In large markets, then – where rich and poor, old and young, men and women, persons of all varieties of tastes, temperaments and occupations are mingled together, – the peculiarities in the wants of individuals will compensate one another in a comparatively regular gradation of total demand. Every fall, however slight in the price of a commodity in general use, will, other things being equal, increase the total sales of it; just as an unhealthy autumn increases the mortality of a large town, though many persons are uninjured by it. And therefore if we had the requisite knowledge, we could make a list of prices at which each amount of it could find purchasers in a given place during, say, a year.

3. The perceptions of Cournot, Walras, Pareto, and Marshall in relation to the literature of the 1950s and 1960s

3.1. Property of economic variables: perception as quantities or functions

We have reviewed how representative theorists such as Cournot, Walras, Pareto, and Marshall, each of whom made significant contributions in the history of economic analysis, perceived and expressed economic quantities through their works in the literature. I now wish to summarize their common perceptions as well as the dissimilarities that exist among them.

Prior to the analytical framework of the 1950s and 1960s on general equilibrium analysis as exemplified in Debreu's prototypical work [7], we seem not to encounter explicit discussions concerning the way in which various commodities that are objects of transactions in markets should be quantitatively or mathematically represented. One of the consequences of the theoretical efforts toward solving the problem of the existence of an equilibrium is that such a basic issue comes to light. Although, in the works in economic analysis prior to the 1950s, issues concerning the perception of functional forms of economic quantities as economic variables and issues concerning the perception of economic quantities themselves are different, they were oftentimes discussed on the same level in a somewhat confused manner.

Common perceptions of the discontinuity of individual demand functions

In order to analyze mathematically the way in which values of commodities are determined in market transactions, Cournot [4] considered quantities demanded, or synonymously in his language, "débit" (sales), as fundamental economic variables in his analysis. He began his analysis by clarifying the demand concept that had been the object of confusion in the literature in his time. He expressed mathematically the market demand of a commodity as a function of its market price. Since market demand was the basis of his analysis, he first discussed continuity as one of its basic properties.[19] But, as we saw in Sect. 2.1, he maintained that demand functions are discontinuous at individual levels. The way he delineated the discontinuity of individual demands is as follows: a slight increase in price does not decrease the quantity demanded in general, but once the price increases up to a certain level, only then does the quantity demanded decline abruptly.

The perception that individual demands in general are discontinuous with respect to price variations was commonly held among the theorists Cournot, Walras, Pareto, and Marshall, with whose works the present paper is concerned. Despite their common perception, we see some essential differences in the grounds for their assertion of individual demands being discontinuous. Their different views reflect their varying perceptions of economic quantities.

Dissimilarities in perception of economic quantities

One cannot find an overt trace of Cournot's perception on the indivisibility of economic quantities. He simply observed that individuals in markets do not behave in such a way as to adjust their purchases a little to meet slight price variations.

[19] As I cautioned earlier, one must be careful to discern whether he actually meant the "continuity" of the market demand, or its "differentiability."

Walras is in the same boat as Cournot insofar as there were no explicit arguments on the perfect divisibility or indivisibility of commodities. Certainly, Walras knew the content of Cournot [4] when he wrote his book [18]. The point of difference between them is as follows: despite the fact that Walras refrained from referring explicitly to the indivisibility of commodities, in my view we should interpret the substance of the quotation from Walras [18, pp. 57–58] as a discussion concerning discontinuity based on the admission of an indivisible commodity (a horse ["cheval"] in case of the quotation). As a matter of fact, Cournot did not enter into a discussion of why individual demands are discontinuous.

Walras's argument, however, seems to stand out as the initiator of general equilibrium analysis, since he essentially argued that even the demand for a *perfectly divisible* commodity (in the quotation's example, "wheat" [blé]) takes the form of a step curve when there is an indivisible commodity. His reasoning is this: as the price of a perfectly divisible commodity gradually declines, unless it goes down to a certain level, one might not reduce the consumption of an indivisible commodity by one unit in order to substitute it for an increase in the consumption of a divisible one. In other words, his insight concerning the discontinuous variation of demands is that it accrues to demands from a substitution between a perfectly divisible commodity and an indivisible one as a result of the utility-maximizing behavior of an individual that has been induced by a change in the prices of the commodities.

Walras explained that an individual demand function (or rather, a demand curve) takes the form of a step curve. We might be tempted to interpret his explanation literally. However, it is highly doubtful whether Walras's true intention was to maintain that individual demand functions are given by step functions in the present-day mathematical sense.

If we were to accept such an interpretation, then individual demand functions would become semi-continuous functions. The diagram given in Walras [18, p. 58], however, shows a graph of a literal step curve. Thus, it would be closer to Walras's true intentions to interpret his argument as meaning that individual demands are "correspondences" (or "multi-valued functions"), the graphs for which are in the form of a step curve rather than semi-continuous functions whose graphs are in the form of a step curve.

We do not find an explicit discussion by Marshall on the perfect divisibility or indivisibility of commodities, as in the case of his predecessors Cournot and Walras. But, since Cournot and Walras differed in their perception of the nature of the discontinuity of individual demands, the way that Marshall explained it is also at variance in details with them. Cournot avoided entering into a discussion about why individual demands should be regarded as discontinuous. He simply accepted their discontinuity based on his observation of individual behaviors. Both Marshall and Walras based their arguments concerning the discontinuity of individual demands on the preferences of individuals. Marshall, however, explained it from what we now call the partial equilibrium analysis point of view, whereas Walras explained it on the basis of a consumption substitution between a perfectly divisible commodity and an indivisible commodity.

By taking the consumption of tea as an example in the previous quotation [12, p. 82], Marshall conceded that individual demand functions are fairly continuous for those commodities that are consumed constantly and can be purchased in small amounts. Nevertheless, by taking hats and watches as an example, he contended that even among those commodities that are in constant use, there are many for which the demand on the part of any single individual may exhibit a great leap at a certain price level. He does not seem to have based his theoretical contention on the observed behavior of individuals, but rather on individual decision-making. Moreover, in his earlier quotation [12, pp. 82–83], he grounded the discontinuity of individual demands on the irregularity of individual needs that would be observed in cases such as wedding cakes or the services of an expert surgeon.

In the case of Pareto, we can see a perspicuous and fundamental difference between his work and that of Cournot, Walras, and Marshall concerning the divisibility of commodities. As was pointed out in Sect. 2, Pareto developed his arguments with a clear understanding of the indivisible nature of commodities. When Pareto took up the issue of continuous or discontinuous variations of economic quantities, it was at the level of indifference curves in the commodity space, and not at the level of demand functions that are derived from an analysis involving indifference curves in the commodity space. However, Pareto followed exactly Cournot's steps in emphasizing *sine qua non* of allowing for "continuous variations" of economic quantities even at individual levels for technical reasons. Pareto's arguments might be construed as his way of understanding or extending the theories of Cournot and Walras, since his work was done under the strong intellectual influences of his predecessors. Yet, as far as the issue of the continuity of demands is concerned, the fact that Pareto began his analysis by discussing the continuity or discontinuity of indifference curves, which represent individual consumers' tastes and preferences in the commodity space, is regarded as a step toward deepening the understanding of the issue.

3.2. Common perception on the continuity of market demand

The discussions here, of Cournot, Walras, Pareto, and Marshall, show that they all maintained that individual demands are "discontinuous" with respect to changes in prices, despite their common understanding that we should nevertheless proceed with our analysis by taking the total market demand to be continuous. The rationale for their common understanding can be classified into two categories.

The first is what Cournot [4, p. 39] emphasized initially, and Marshall [12, pp. 82–83] subsequently elucidated, that as wealth, preferences, and needs among individuals constituting the markets extend so as to encompass their wide varieties, total market demand tends to show continuous variations with respect to price changes.

The second is what Walras initially referred to as "the law of large numbers," and then Pareto subsequently described as an "average phenomenon." A distinctive feature of Pareto's is that he supported his arguments with the understanding of individual economic phenomena as an average phenomenon under large numbers, which gave him a rationale for using continuous real numbers, even at individual levels. It might be taken as a way to understand what Walras called, in an ambiguous manner, the law of large numbers.

It is arguably possible to illuminate Pareto's insight of interpreting economic quantities as an average phenomenon in a model of a "continuum economy" as introduced by Aumann [1], or rather in a model of a "large economy" which is an extensively formalized version of a continuum economy by Hildenbrand [9]. We shall return to this point in the next section.

4. The perception of economic quantity in the history of economic analysis and the Debreu conjecture

4.1. The interpretation of perceptions of "discontinuity" in modern economic analysis

Structural framework of the commodity space and varied perceptions of the discontinuity

In the preceding two sections, the perceptions of the four representative theorists in the history of economic analysis concerning the discontinuity of individual demands were reviewed and discussed. This section will present an interpretation of their perceptions by incorporating a present-day mathematical point of view.

As a theoretical structure of the basic commodity space, let us consider two types. One is typically found in Debreu [7] and is the standard theoretical model structure of general equilibrium analysis from the 1950s and 1960s. With the natural number ℓ representing the finite number of distinguishable commodities, the commodity space is given by the ℓ-dimensional Euclidean space $\mathbb{R}^\ell$, and all the commodities are regarded as perfectly divisible. Hence, any real number may represent a physically possible quantity. The discussion will continue by taking as an example the simplest type of commodity space, that is, the two-dimensional Euclidian space $\mathbb{R}^2$.

The other type of structure of the commodity space is the one in which some among the ℓ commodities are purely indivisible so that only the integer amounts of consumptions are physically viable. In examples with two commodities, the space

$$\{\cdots\cdots, -3, -2, -1, 0, 1, 2, 3, \cdots\cdots\} \times \mathbb{R}$$

will be considered as the commodity space instead of $\mathbb{R}^2$. The first commodity is purely indivisible, and the second perfectly divisible, allowing for any real number amounts as physically possible consumptions.

In the following, various types of perception concerning the "discontinuity" of demand functions will be distinguished. The first type is the "discontinuity" as understood in present-day mathematics. The second type is the one interpreted as a correspondence or set-valued (or a multi-valued) function. The third type is the one understood not only as a correspondence but also by its failure to be upper hemi-continuous.[20] We might add a fourth type: the nonconvex-valuedness of a correspondence.

Interpretation of the "discontinuity" in the commodity space without any explicit indivisible commodities

In his discussion (quoted earlier in Sect. 2.2), Walras argued that individual demand functions are "discontinuous," taking the form of a step curve. If his words are taken literally, he recognized the discontinuity by noticing that they are correspondences, with their graph having the form of a step curve. Following the Walrasian way of expression, the demand is continuous if an infinitesimally small change in prices induces an infinitesimally small change in the quantity demanded; otherwise, it is discontinuous. For an individual demand that is discontinuous having the form of a step curve, Walras displayed a diagram, as in Fig. 1. Of course, it is a literal step curve.

In the commodity space without an explicitly indivisible commodity, let us consider whether one can deduce demand curves having a form that is much the same as the one to which Walras or Cournot referred as discontinuous. In Fig. 2, we tried to exhibit a convex preference relation that induces a step in the graph of its derived demand curve. Taking $\mathbb{R}^2$ as the commodity space, we let the consumption set that represents individual needs

[20] The definition or meaning of the *upper hemi-continuity* is reviewed in footnote 24.

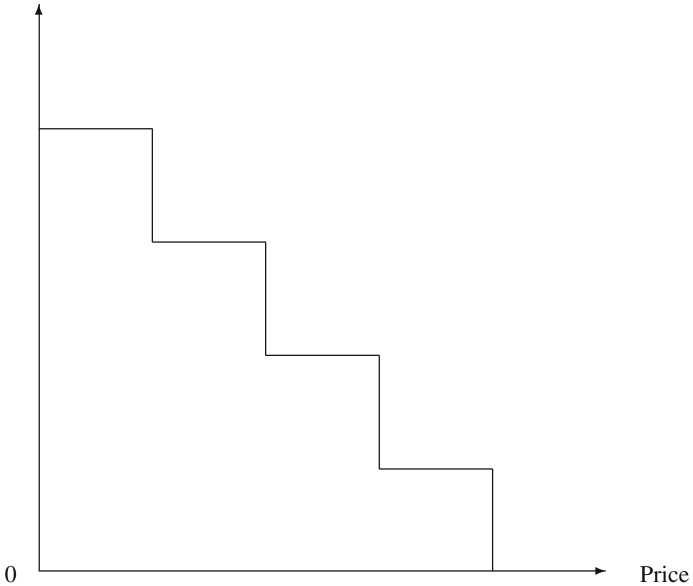

Fig. 1 An individual demand curve

and possible individual consumptions be given by the set $\mathbb{R}^2_+ = \{x = (x_1, x_2) | \, x_1 \geqq 0, x_2 \geqq 0\}$. Piecewise linear lines in the figure show representative indifference curves associated with the individual preference relation.

Price vectors $p^A = (p_1^A, p_2^A)$, $p^B = (p_1^B, p_2^B)$, and $p^C = (p_1^C, p_2^C)$ correspond to the budget lines BL^A, BL^B, and BL^C, respectively. The curve in Fig. 3 is the graph of the consumer's demand for commodity 1, which is deduced from the consumption decisions shown in Fig. 2, with only the price of commodity 1 changing. The consumer's choice under the budget line BL^A is given by point x'. In case the budget line is BL^B, the consumptions that the consumer chooses become any one of the points lying along the line segment between (and including) points x and y. When the budget line moves further to BL^C, the chosen consumptions are given by point y'.

In this case, the induced demand is not a demand function but is a demand correspondence, since all the consumption bundles on the line segment between x and y could be chosen when the price vector is given by p^B. Thus, the demand curve has the form of a step, but not with multiple steps, as in Fig. 1. Moreover, one might question whether Walras (or for that matter, Cournot) had this type of step or jump in mind when either of them talked about the discontinuity of a demand curve.

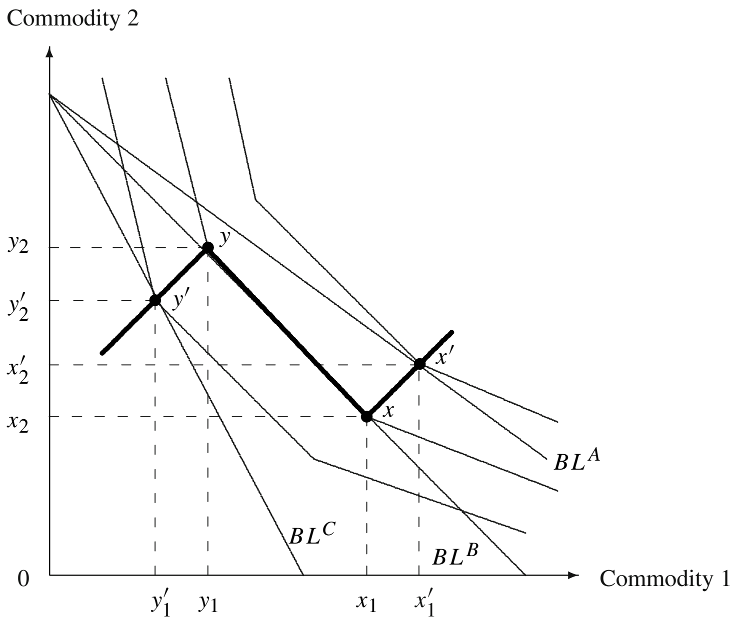

Fig. 2 Consumer's choice

If one really wants to describe a jump in quantity demanded in some sense within the commodity space without an explicitly indivisible commodity, one might need to appeal to a preference relation not satisfying the convexity. Figure 4 exhibits such a preference relation. Piecewise linear lines in the figure show representative indifference curves associated with the individual preference relation. Price vectors $p^A = (p_1^A, p_2^A)$ and $p^B = (p_1^B, p_2^B)$ correspond to the budget lines BL^A and BL^B. The curve in Fig. 5 is the graph of the consumer's demand for commodity 1 induced by the consumption decisions shown in Fig. 4. Commodity bundles x and y in Fig. 4 represent ones that are demanded when market prices are given by the price vector p^A. Hence, in this case also the demand is not a demand function but is given as a demand correspondence, and the quantity of commodity 1 demanded suddenly drops below y_1 if its price rises above the threshold price level p_1^A. Strictly speaking, this example cannot be said to reproduce situations that would fit the descriptions of Walras, Cournot, or Marshall of the discontinuity of individual demands. Nevertheless, it might be said to describe circumstances where the quantity demanded of a commodity changes abruptly at a

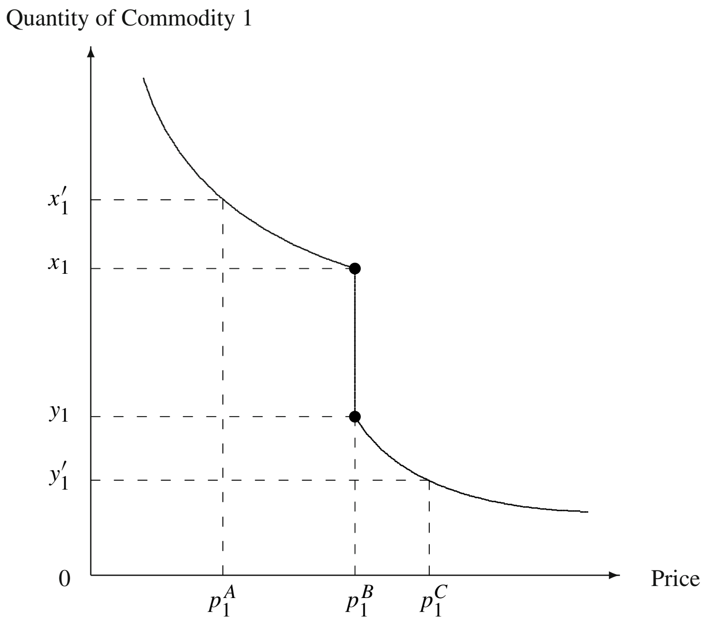

Fig. 3 Consumer's demand curve for commodity 1

certain level of its price within a framework of the commodity space with only perfectly divisible commodities.[21]

Interpretation of the "discontinuity" in the commodity space with explicitly indivisible commodities

The next question to ask, then, is whether one could give an accurate interpretation of the discontinuity that Cournot, Walras, and Marshall perceived by explicitly taking account of indivisible commodities. Let us take as an example the commodity space $\{\cdots\cdots, -3, -2, -1, 0, 1, 2, 3, \cdots\cdots\} \times \mathbb{R}$, where commodity 1 is purely indivisible and commodity 2 is perfectly divisible. Then, we take as the consumption set

$$X = \{0, 1, 2, 3, \cdots\cdots\} \times \mathbb{R}_+$$

[21] Note, however, that here the demands as a correspondence satisfy the property of upper hemi-continuity. It simply represents the lack of convex-valuedness of the correspondence at the price vector p^A.

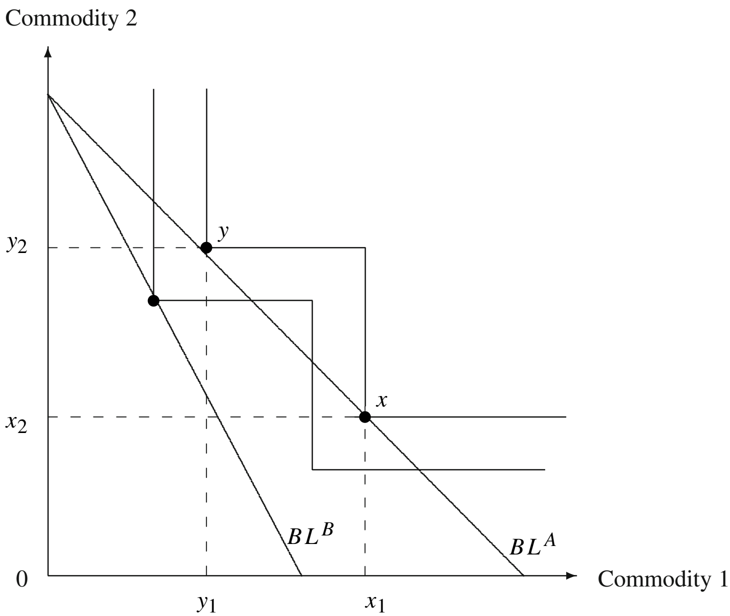

Fig. 4 Consumer's choice

where possible consumptions consist of commodity bundles with nonnegative consumption amounts. $\mathbb{R}_+$ is the set of all nonnegative real numbers. Figure 6 exhibits an example of a consumer's behavior with X as its consumption set. In the figure the consumption set consists of perpendicular half-lines. The preference relation of the consumer is representatively shown by several indifference curves that are given by the consumption vectors lying on both the dotted curves and the perpendicular half-lines.

Let the price vectors $p^A = (p_1^A, p_2^A)$, $p^B = (p_1^B, p_2^B)$, and $p^C = (p_1^C, p_2^C)$ correspond to the budget lines BL^A, BL^B, and BL^C, respectively. The consumer's choice under the budget line BL^A is shown by point x. Under the budget line BL^B, the consumer chooses points x' and y. When the budget line is BL^C, the chosen consumptions are indicated by points y' and z. This consumer's choice behavior in Fig. 6 with respect to the quantity of commodity 1 demanded is shown in Fig. 7 as the curve of its individual demand function. The figure does not exactly matches a piecewise linear step curve as drawn by Walras, reproduced in Fig. 1, but it is possible to regard it as representing what Walras [18, p. 57] explained with respect to the variations

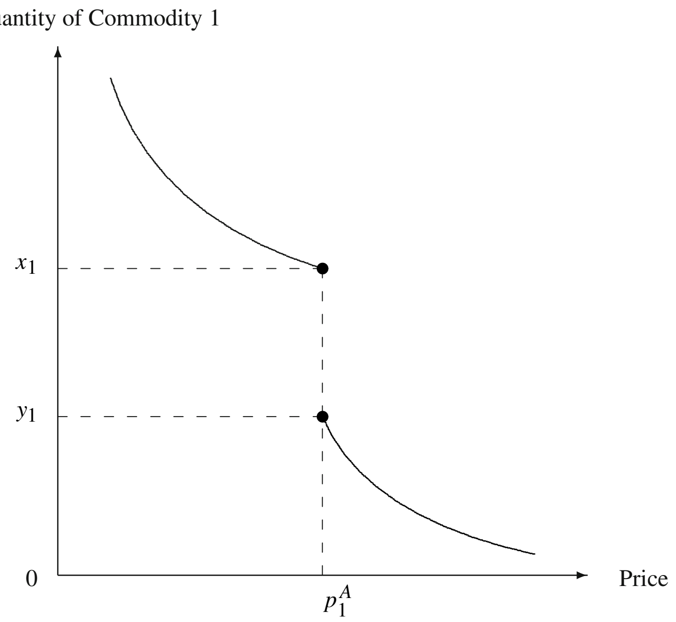

Fig. 5 Consumer's demand curve for commodity 1

of individual demands.[22] Moreover, we might say that it also represents the discontinuity of individual demands discussed by Cournot [4, pp. 38–39] and Marshall [12, p. 82].

Nevertheless, one may wonder whether Cournot's perception of the discontinuity of individual demand functions, as compared with that of Walras, Pareto, or Marshall, might not be rooted in a more profound insight. If this were the case, how could we understand his perception of the discontinuity of the demands?

Since Cournot was a mathematician before he was an economic theorist, and had written, among other things, textbooks on calculus (Cournot [6]) and on probability theory, it would not be appropriate to think that he regarded individual demands to be functions taking the form of a step curve, as in the case of Walras. I believe we could interpret the earlier quotation of Cournot

[22] However, if we take Walras's explanation [18, p. 57] to indicate changes in quantity demanded of a perfectly divisible commodity resulting from its substitution for an indivisible one, then these figures fail to represent his explanation. The Walras' explication might seem persuasive at first sight, but if we depict the consumer's choice in Fig. 6 by the demand function for commodity 2, then, strictly speaking, it seems that his insight may have been misled.

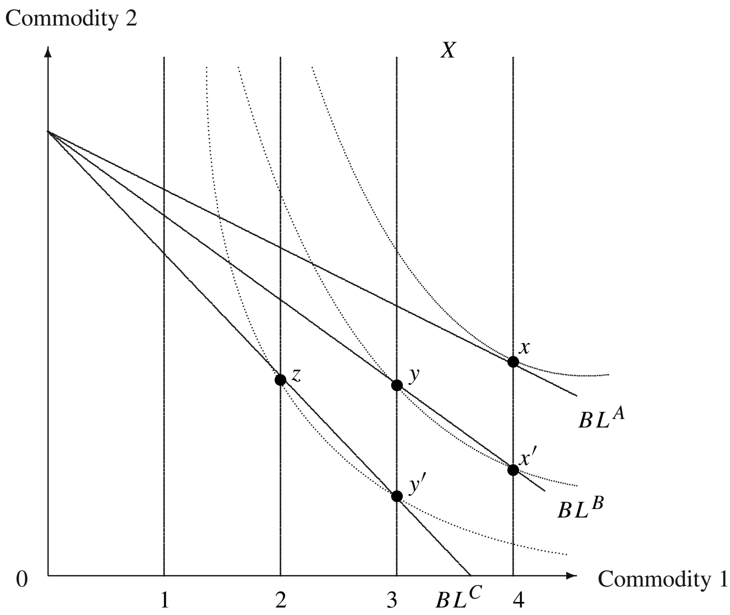

Fig. 6 Consumer's choice

[4, pp. 38–39] to mean that an individual demand for a particular commodity as a real-valued function is at most a *semi-continuous* function.[23]

It is difficult to infer whether Cournot's perception of the discontinuity of individual demand functions was based on a more profound insight or not. Let us be more specific about this point, using Figs. 8 and 9. In Fig. 8, each of

[23] In particular, I would like to call attention to the fact that his explanation in the quotation could be understood to mean that an individual demand function as a real-valued function cannot be (lower semi-) continuous although it might be upper semi-continuous.

A real-valued function $f : X \to \mathbb{R}$ is *upper semi-continuous* at $x \in X$ if the set $\{z | f(z) < f(x)\}$ is open, and f is *upper semi-continuous* if it is upper semi-continuous at all $x \in X$.

$f : X \to \mathbb{R}$ is *lower semi-continuous* at $x \in X$ if the set $\{z | f(z) > f(x)\}$ is open, and f is *lower semi-continuous* if it is lower semi-continuous at all $x \in X$.

f is said to be *semi-continuous* if it is either upper semi-continuous or lower semi-continuous at all $x \in X$.

Note that even if a real-valued function is upper semi-continuous, it need not be upper hemi-continuous when regarded as a correspondence. See a footnote 24 for the concepts of the upper hemi-continuity and the lower hemi-continuity of a correspondence.

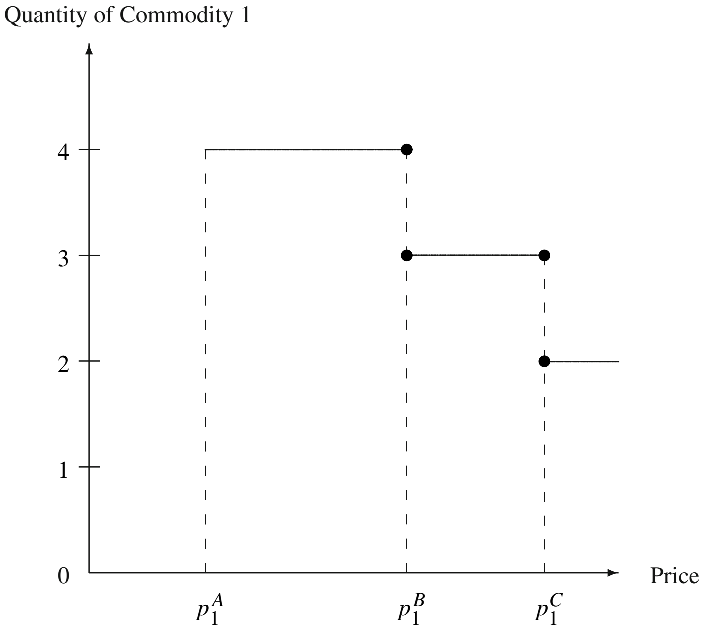

Fig. 7 Consumer's demand curve for commodity 1

the budget lines BL^A, BL^B, BL^C, and BL^D corresponds to the price vector $p^A = (p_1^A, p_2^A)$, $p^B = (p_1^B, p_2^B)$, $p^C = (p_1^C, p_2^C)$, or $p^D = (p_1^D, p_D)$, respectively. Using the price-consumption curve in Fig. 8, these figures show how an individual demand changes in response to the changes in the price of commodity 1 from p_1^A to p_1^D. The demand curve of commodity 1 derived from this price-consumption curve is the graph of the correspondence having the "form of a step curve" in Fig. 9. Blackened points indicate that they are a part of the curve, whereas the points simply circled are not a part of the curve. The correspondence in Fig. 9 is a function except at the price vector p_1^D. Within the region where it becomes a function, it is semi-continuous as a real-valued function. It is not lower semi-continuous but is upper semi-continuous. Regarded as a correspondence in the whole region containing the price vector p_1^D, the demand is not convex-valued at the price p_1^D, but is continuous there, i.e., it is upper hemi-continuous as well as lower hemi-continuous at p_1^D. However, it is not upper hemi-continuous at p_1^B and p_1^C.[24]

[24] A correspondence $F : X \to Y$ is *upper hemi-continuous* at $x \in X$ if for any open set $G \supset F(x)$ in Y, there exists an open set V with $x \in V$ such that for every $z \in V$ one has $F(z) \subset G$. F is *upper hemi-continuous* if it is upper hemi-continuous at every $x \in X$.

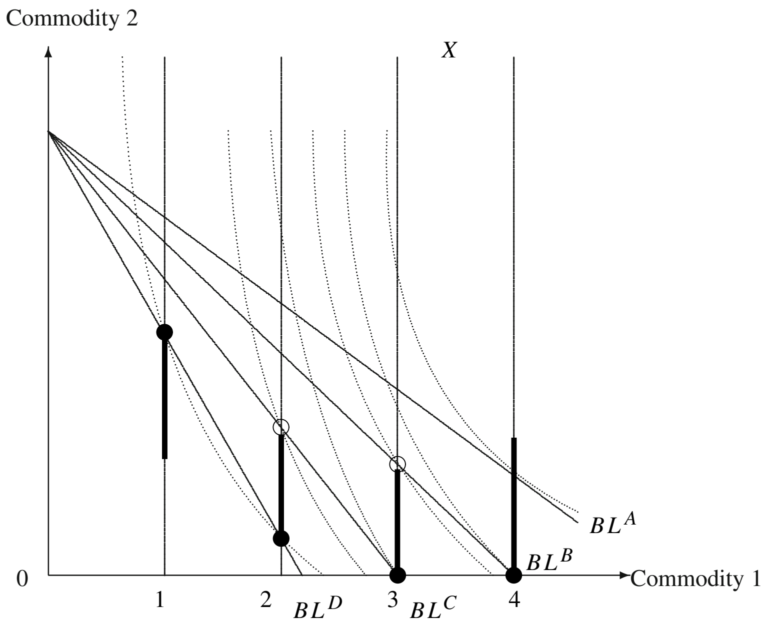

Fig. 8 Consumer's choice

It is not clear whether Cournot's awareness of the discontinuity of individual demands was about their nonconvex-valuedness as exemplified in Figs. 6 and 7 or more insightfully about their lack of upper hemi-continuity. The demand correspondence in Fig. 9 is not convex-valued at p_1^D. If Cournot implicated such a situation in [4, pp. 38–39] as a lower semi-continuous step function as in our earlier footnote, we should admit that he did not apperceive the lack of upper hemi-continuity at p_1^B, p_1^C as a demand correspondence. Since the correspondence in Fig. 9 is single-valued at $p_1^B \mathrm{C} p_1^C$, it is upper semi-continuous at these prices when viewed as a function. One could say that Cournot [4, pp. 38–39] simply meant that demand functions are semi-continuous without further perceptive distinction of upper or lower semi-continuity. In that case it would be possible to say that he had an apprehension of the fact that the demands as a correspondence might fail to be upper hemi-continuous.[25]

F is *lower hemi-continuous* at $x \in X$ if for any open ste G in Y with $F(x) \cap G \neq \emptyset$, there exists an open set V with $x \in V$ such that for any $z \in V$ one has $F(z) \cap G \neq \emptyset$. If F is lower hemi-continuous at every $x \in X$, then F is *lower hemi-continuous.*

[25] Since the proofs of the existence of an equilibrium in the general equilibrium model of 1950s and 1960s were carried out in a framework where individual demand

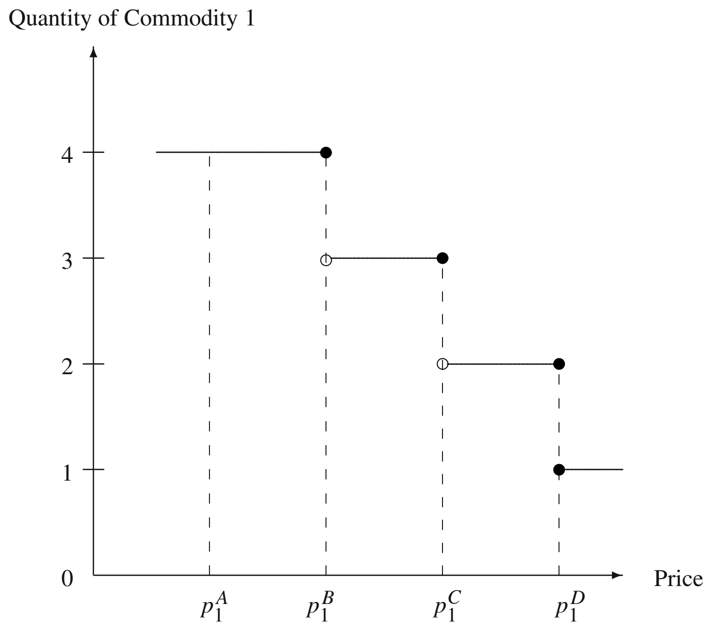

Fig. 9 Consumer's demand curve for commodity 1

To sum up our views on the way in which Cournot, Walras, and Marshall perceived the discontinuity of individual demands and how Pareto perceived the discontinuity:

(1) Except for Pareto, there were no explicit arguments involving indivisible commodities. Further, we did not obtain a demand function with its graph

---

correspondences essentially become upper hemi-continuous, the awareness of circumstances under which individual demand correspondences fail to be upper hemi-continuous seems to be fairly limited.

As we typically see in Debreu [7, 4.8, p. 63], an individual demand correspondence may fail to be upper hemi-continuous if the level of wealth of that individual drops to the minimal level so as to sustain the purchase of the least expensive combinations of commodities among all the possible consumptions. Thus, in most cases, existence proofs have been carried out under conditions in which all the economic agents circumvent the situation of the minimum level of their wealth for possible consumptions.

Now, in Fig. 9, the circumstances under p_1^B or p_1^C do not correspond to the "minimum wealth level" among all the possible consumptions. However, the lack of upper hemi-continuity arose from the existence of an indivisible commodity. In the literature these circumstances were clarified by Broome [3], Mas-Colell [14], and Yamazaki [20, 22].

having the “form of a step curve” that they referred to in their words or in their diagrams within a framework of the commodity space with perfectly divisible commodities only.

(2) By explicitly incorporating an indivisible commodity into the commodity space, one could obtain a demand curve in the “form of a step curve” by interpreting the fact that the demand correspondence is not convex-valued to implicate a situation representing a “step” in the weak sense.

(3) Notwithstanding the interpretation by Cournot, Walras, and Marshall, to the effect that they took account of indivisible commodities, their arguments do not lead to the recognition of the lack of upper hemi-continuity of demand correspondence.

(4) In Cournot’s case, there are some grounds left to believe that his arguments could be taken to imply his perception of the lack of the upper hemi-continuity of individual demand correspondences when some of the commodities are indivisible.

(5) In Pareto’s case, his arguments recognized indivisible commodities in the commodity space in a straightforward way, and he perceived the discontinuity of indifference curves upon which the derivation of a demand curve depends.

4.2. The Debreu conjecture and the interpretation of the common perception of the continuity of market demands

This final section of the paper will now return to the Debreu conjecture and will propose a present-day interpretation of the common perception on the continuity of market demands.

As reviewed in the previous section, Cournot, Walras, Pareto, and Marshall each understood individual demands to vary discontinuously with respect to the change in prices, but they perceived in common that the total market demand varies continuously with respect to price changes.

If we review the content of the Debreu conjecture, we read: “One expects that if the measure ν is suitably diffused over the space A (of economic agents’ characteristics), integration over A of the demand *correspondences* of the agents will yield a total *demand function*, possibly even a total demand function of class C^1.” The first half of this conjecture reflects, in modern mathematical terms, what Cournot [4, p. 38] stated as “the wider the market extends, and the more the combination of needs, of fortunes, or even caprices, are varied among consumers,” and what Marshall [12, pp. 82–83] stated as “large markets ... where rich and poor, old and young, men and women, persons of all varieties of tastes, temperaments and occupations are mingled together ... the peculiarities in the wants of individuals will compensate one another in a comparatively regular gradation of total demand." Debreu did not mention Cournot or Marshall at all in his paper [12] when he promulgated his conjecture. I personally find it very hard to believe that Debreu did not have any knowledge of Cournot's idea about this issue.

The last half of the conjecture consists of two parts. One part is to assert that the total market demand obtained from aggregation over individual demands will possibly become a function, i.e., a single-valued correspondence, regardless of whether individual demands are correspondences or not, provided there is enough "diffusion" or variation of wealth levels, preference relations, needs, etc., among individual agents.[26] The other part is to claim the possibility of having a continuously differentiable function as a total market demand once it becomes a function.

In my view, the first part is definitely a literal translation of Cournot's [4, pp. 38–39] assertion, and the subsequent one is a literal translation of the assertion by Marshall [12, pp. 82–83], both in a large economy representation of general equilibrium analysis, in as much as the fact that the total demand is a function within the framework of the 1950s and 1960s implies that demands change continuously with respect to price changes.[27]

As to the second part of the last half of the conjecture, it might even seem on the surface that no one – not Cournot, Walras, Pareto, or Marshall – referred to it. However, as was pointed out in Sect. 2, we suspect that, from Cournot's description of the property of continuous function, when discussing the *continuity* of demand functions, he might have instead had in mind the *differentiability* of demand functions in the guise of continuity.[28] Indulging in such an interpretation, Cournot's argument [4, pp. 38–39] becomes essentially the Debreu conjecture itself. In other words, by reversing the statement, we could say that the Debreu conjecture succeeds in making a formal statement out of Cournot's idea in terms of the present-day economic theory.

Next, aside from our discussion of the direct significance of the conjecture, we believe it appropriate to discuss a possible common lineage between

---

[26] See Hildenbrand [10] and Yamazaki [21].

[27] It may be better to offer a remark on the question whether a demand correspondence, in being a function, implies de facto its being continuous. Within the framework of the Debreu conjecture, all the commodities are perfectly divisible, and in his general equilibrium with a differentiable structure of preference relations, as noted in an earlier footnote, the circumstances under which individual demand correspondence may fail to be upper hemi-continuous are excluded so that the mere fact of being demand *functions* guarantees the continuity of demand functions.

[28] For, as we remarked in an earlier footnote, he used the property of local linearity as a characteristic of a continuous function. What is more, he described this property as a property of differentiable function in his textbook on calculus [6, pp. 9–10].

the conjecture and the perception of "the law of large numbers" that Walras alluded to, or "the average phenomenon" that Pareto pointed out.

Take a continuum economy as introduced by Aumann [1], and let $I = [0, 1]$ be an index set of the population of individuals composing an economy, and λ be the Lebesgue measure on I with $\lambda(S)$ representing the proportion of individuals belonging to a group S of the people in I. For each $t \in I$, $F(t, p)$ is the value of individual demand correspondences, and it shows the set of demand vectors of individual agent t under the price vector p. In a continuum economy, the integral $\int_I F(t, p)d\lambda$ represents the set, showing the value of the aggregate total demand correspondence at p.[29]

Hildenbrand [9], who systematically expanded the framework of a continuum economy to a model of large economies, referred to an aggregate total demand as a *"mean demand."* The name indicates that units of the aggregate total demand $\int_I F(t, p)d\lambda$ are interpreted to be amounts expressed in terms of units per capita among the population of economic agents in an economy. Now, the value of mean demand $\int_I F(t, p)d\lambda$ as the aggregate total demand is known to be convex-valued in a continuum economy.[30]

In discussing the market total demand by aggregating over individual demands, Walras made reference to "the law of large numbers" without giving any further detailed comments or explanation. Thus, we believe it reasonable to surmise what he meant to address is the fact that mean demand $\int_I F(t, p)d\lambda$ becomes convex-valued. In other words, the law of large numbers in the sense of Walras, is understood to be a phenomenon of straightforward convexifying effects inherent to a process of aggregation itself over a large number of individual demands.

How, then, can what Pareto called "the average phenomenon" of economic quantities be understood? Clearly, one might wish to regard the mean demand representing an aggregate total demand as "the average phenomenon" of economic quantities. However, this forthright interpretation does not seem to be a precise representation of Pareto's average phenomenon in view of the fact that he was concerned about the treatment of quantities in the commodity space. He argued that even though in reality each individual faces choices among discrete amounts, theoretically we may regard each individual as making decisions among continuous economic quantities, since we are only interested in an average phenomenon of those individuals.

---

[29] For the concept of integrals in such a mathematical model, please refer to the books by Hildenbrand [9], Jacobs [11], Maruyama [13], or Yamazaki [22].

[30] This is a consequence of a well-known mathematical theorem attributed to Lyapunov. See, for example, Hildenbrand [9, Theorem 3, p. 62] or Yamazaki [22, Theorem 14.2, p. 186].

For the purpose of clarifying Pareto's arguments, let us consider the commodity space in which some of the commodities are explicitly indivisible. Since quantities are confined to discrete amounts, the commodity space would become nonconvex, as do individual preference relations. Thus, as an analytic convenience, consider the convexification of consumption sets and preference relations. In this procedure of convexification, we can allow individuals to make choices among consumption bundles composed of continuous real numbers not confined to discrete amounts. Then, if the values of the actual mean demand arising from consumptions sets and preference relations without their convexification can be shown to coincide with those of the mean demand resulting from the convexified consumption sets and preference relations, this procedure of convexification can be understood as representing Pareto's ideas.[31]

We are more than willing to concede that the representation as well as the perception of quantities and variables in the economic analysis clearly has a correlation with the development of the mathematical analysis.[32] Hence, as our future research, we would like to deepen our understanding of the way in which the representation of quantities and their variables found in the history of the mathematical analysis might be understood to be related to the perception of economic quantities.

References

1. Aumann, R.J.: Markets with a continuum traders. Econometrica **32**, 39–50 (1964)
2. Bourbaki, N.: Elements of the History of Mathematics, translated by J. Meldrum. Springer, Berlin (1994)
3. Broome, J.: Existence of equilibrium in economies with indivisible commodities. J. Econ. Theory **5**, 224–250 (1972)
4. Cournot, A.A.: Recherches sur les Principe Mathématiques de la Théorie des Richesses. L. Hachette, Paris (1838)
5. Cournot, A.A.: Researches into the Mathematical Principles of the Theory of Wealth, translated by N. Bacon. Macmillan, New York (1927)
6. Cournot, A.A.: Traité Elémentaire de la Théorie des Fonctions et du Calcul Infinitésimal . L. Hachette, Paris (1841)
7. Debreu, G.: Theory of Value. Wiley, New York (1959)

[31] If one is interested in seeing to what extent such a proposition might be shown to be true, see, e.g., Yamazaki [22, Theorem 14.1 or 14.2, pp. 184–187].

[32] See, e.g., Bourbaki [2].

8. Debreu, G.: Smooth preferences. Econometrica **40**, 603–615 (1972)
9. Hildenbrand, W.: Core and Equilibria of a Large Economy. Princeton University Press, Princeton (1974)
10. Hildenbrand, W.: On the uniqueness of mean demand for dispersed families of preferences. Econometrica **48**, 1703–1710 (1980)
11. Jacobs, K.: Measure and Integral. Academic, New York (1978)
12. Marshall, A.: Principles of Economics, 8th edn. Macmillan, London (1920)
13. Maruyama, T.: Integral and Functional Analysis (in Japanese). Springer, Tokyo (2006)
14. Mas-Colell, A.: Indivisible commodities and general equilibrium theory. J. Econ. Theory **16**, 443–456 (1977)
15. Pareto, V.: Manuale di Economia Politica. Società Editrice Libraria, Milano (1906)
16. Pareto, V.: Manual of Political Economy, translated by Ann S. Schweir. MacMillan, New York (1971)
17. Schumpeter, J.A.: History of Economic Analysis, edited from manuscript by Elizabeth B. Schumpeter. Oxford University Press, New York (1954)
18. Walras, L.: Éléments d'Economie Politique Pure (Édition définitive, 1926). Corbaz, Lausanne (1874–1877)
19. Walras, L.: Elements of Pure Economics, translated by William Jaffé. George Allen and Unwin, London (1954)
20. Yamazaki, A.: An equilibrium existence theorem without convexity assumptions. Econometrica **46**, 541–555 (1978)
21. Yamazaki, A.: Continuously dispersed preferences, regular preference-endowment distribution, and mean demand fuction. In: Green, J., Scheinkman, J. (eds.) General Equilibrium, Growth, and Trade, pp. 13–24. Academic, New York
22. Yamazaki, A.: Foundations of Mathematical Economics (in Japanese). Soubunsha, Tokyo (1986)

Adv. Math. Econ. 15, 123–127 (2011)

Advances in
MATHEMATICAL
ECONOMICS

An existence result and a characterization of the least concave utility of homothetic preferences

Yuhki Hosoya*

Graduate School of Economics, Keio University, 2-15-45 Mita, Minato-ku,
Tokyo 108-8345, Japan
(e-mail: stairlimit@moon.cims.jp)

Received: July 24, 2010
Revised: October 22, 2010

JEL classification: D11

Mathematics Subject Classification (2010): 91B16, 91B08

Abstract. This note shows that a utility function of a homothetic preference relation satisfying $u(0) = 0$ is a least concave utility function if and only if it is homogeneous of degree one.

Key words: least concave utility, homothetic preference, homogeneity of degree one

1. Introduction

The aim of this note is to show some characterization of the least concave utility function of homothetic[1] preference relations. We show that if a preference relation is homothetic, then a utility function satisfying $u(0) = 0$ is a least concave utility function if and only if it is homogeneous of degree one.

Debreu [1] defines the least concave utility function of concavifiable preference relations and shows its existence and uniqueness up to a positive

* We are grateful to Toru Maruyama and an anonymous referee for helpful comments and suggestions.

[1] The definition of homothetic preference is in the next section.

DOI: 10.1007/978-4-431-53930-8_6,

affine transformation.[2] The property of "the uniqueness up to a positive affine transformation" just mentioned implies the cardinality of this utility function. Ample possibility of applications is also noteworthy. For example, Kannai [2] proposes a new definition of the substitution and complementarity between two commodities by using this utility function.

However, there seems to be a difficulty in applying the concept of least concave utility, that is, the concrete form of it should be computed and identified. Unfortunately, it is, in general, a hard task.

The aim of this note is to solve this problem partially. We characterize the least concave utility function of a homothetic preference relation as the homogeneous utility function of degree one. Thus, many usual utility functions are, in fact, the least concave utility functions. For example, $\sqrt{xy}$, $\sum_{t=0}^{\infty} \beta^t (\sqrt{x_t} + \sqrt{y_t})^2$, and $\int_0^{\infty} e^{-\rho t} (\sqrt{x(t)} + 2\sqrt{y(t)} + 3\sqrt{z(t)})^2 dt$ are all least concave utility functions.

In next section, we present the formal statement of our theorem. Proofs of the results are provided in Sect. 3.

2. The main result

Let the consumption set Ω be a closed convex cone of some real topological vector space. The binary relation $\succsim$ on Ω is said to be

- Complete if $x \succsim y$ or $y \succsim x$ for any $x, y \in \Omega$.
- Transitive if $x \succsim z$ for any $x, z \in \Omega$ such that there exists $y \in \Omega$ such that $x \succsim y$ and $y \succsim z$.
- Continuous if the graph of $\succsim$ is closed in Ω^2.
- Convex if the set $\{y \in \Omega | y \succsim x\}$ is convex for any $x \in \Omega$.
- Homothetic if $ax \succsim ay$ iff $x \succsim y$ for any $x, y \in \Omega$ and any $a > 0$.

Let $P^{\succsim}$ be the set of all $v \in \Omega$ which satisfy the following two conditions:

(i) $av \succsim bv$ iff $a \geq b$ for any $a, b \geq 0$.
(ii) for any $x \in \Omega$, there exists $a, b \geq 0$ such that $av \succ x$ and $x \succsim bv$.

A function $u : \Omega \to \mathbb{R}$ is called a *utility function* of $\succsim$ if $x \succsim y$ iff $u(x) \geq u(y)$ for any $x, y \in \Omega$. We say that a utility function u of $\succsim$ is *least concave* if (i) it is continuous and concave, and (ii) for any continuous

[2] The uniqueness up to a positive affine transformation means the following fact: if both u_1 and u_2 are the least concave utility function, then there exist some $a > 0$ and $b \in \mathbb{R}$ such that $u_1 = au_2 + b$.

and concave utility function w of $\succsim$, there exists a concave transformation $\phi : w(\Omega) \to \mathbb{R}$ such that $w = \phi \circ u$.[3] Debreu [1] shows that if at least one continuous and concave utility function of $\succsim$ exists, then there also exists a least concave utility function of $\succsim$.

We have finished the preparation for our main result.

Theorem. Let $\succsim$ be a complete, transitive, continuous, convex, and homothetic binary relation on Ω and suppose that $P^{\succsim} \neq \emptyset$. Then, for any $v \in P^{\succsim}$, the function $u_v : x \mapsto \inf\{c > 0 | cv \succsim x\}$ is a least concave utility function of $\succsim$. Moreover, for any utility function u of $\succsim$, the following three statements are equivalent:

(1) $u = u_v$ for some $v \in P^{\succsim}$.
(2) u is a least concave utility function and $u(0) = 0$.
(3) u is homogeneous of degree one.

In the next proposition, we will present certain mild conditions which guarantee $P^{\succsim} \neq \emptyset$.[4]

Proposition. Let $\succsim$ be a complete, transitive, continuous, and homothetic binary relation on Ω. If 0 is not the greatest element on Ω with respect to $\succsim$ and there exists some neighborhood N of 0 such that there exists the greatest element $v \in N$ on N with respect to $\succsim$. Then $v \in P^{\succsim}$ and thus $P^{\succsim} \neq \emptyset$.

Remark. We should display several examples wherein our theorem can be applied.

Example 1. If $\Omega = \mathbb{R}^n_+$, every utility function that satisfies monotonicity, quasi-concavity, and homogeneity of degree one is a least concave utility function.[5] Hence, the followings are all least concave utility functions:

- The Cobb = Douglas utility $\prod_{i=1}^n x_i^{\alpha_i}$. ($\alpha_i > 0$ for all i and $\sum_{i=1}^n \alpha_i = 1$)
- The CES utility $\left(\sum_{i=1}^n \alpha_i x_i^{\rho}\right)^{\frac{1}{\rho}}$. ($\alpha_i > 0$ for all i and $\rho \in]-\infty, 1[\setminus\{0\}$)
- The linear utility $\sum_{i=1}^n \alpha_i x_i$. ($\alpha_i > 0$ for all i)
- The Leontief utility $\min_i\{\alpha_i x_i\}$. ($\alpha_i > 0$ for all i)

[3] Since Ω is convex and w is continuous, $w(\Omega)$ must be connected. In general, any connected subset of $\mathbb{R}$ is convex. Therefore, we have $w(\Omega)$ is convex.

[4] The condition $P^{\succsim} \neq \emptyset$ is so mild that we could not find any example in which $P^{\succsim} = \emptyset$, except a trivial example: $P^{\succsim} = \emptyset$ if $x \sim y$ for any $x, y \in \Omega$.

[5] This result is partially shown in Kihlstrom and Mirman [3].

Example 2. Let $u : \mathbb{R}^n_+ \to \mathbb{R}$ be monotone, quasi-concave, and homogeneous of degree one. Fix any $p \in [1, \infty]$ and $\beta \in]0, 1[$, and let $\Omega = \{(x_t) \in \ell^p(\mathbb{R}^n) | \sum_{t=0}^{\infty} \beta^t u(x_t) < \infty\}$. Then Ω is a closed convex cone of ℓ^p, and

$$U((x_t)) = \sum_{t=0}^{\infty} \beta^t u(x_t)$$

is a least concave utility function.

Example 3. Let $u : \mathbb{R}^n_+ \to \mathbb{R}$ be monotone, quasi-concave, and homogeneous of degree one. Fix any $\rho > 0$, and let $\Omega = \{x \in L^\infty([0, +\infty[, \mathbb{R}^n) | \int_0^\infty e^{-\rho t} u(x(t))dt < \infty\}$. Then Ω is a closed convex cone of L^∞,

$$U(x) = \int_0^\infty e^{-\rho t} u(x(t))dt$$

is a least concave utility function.

3. Proof

Proof of theorem.

It is easy to show that u_v is a utility function for any $v \in P^{\succsim}$.[6] Since $u_v(\Omega) = [0, \infty[$, u_v is continuous. The homogeneity of degree one of u_v is trivial. It is well-known that if a function $u : \Omega \to \mathbb{R}$ is continuous, homogeneous of degree one, and quasi-concave, then u is concave.[7] Hence, u_v is concave and thus $\succsim$ has at least one continuous and concave utility function.

Choose some least concave utility function w. Since u_v is continuous and concave, there exists a monotone concave function $\phi : w(\Omega) \to \mathbb{R}$ such that $u_v = \phi \circ w$. If ϕ is convex, then it is affine. For any concave utility function θ, there exists a monotone concave function $\psi : w(\Omega) \to \mathbb{R}$ such that $\theta = \psi \circ w$. Then $\psi \circ \phi^{-1} : u_v(\Omega) \to \mathbb{R}$ is concave and $\theta = (\psi \circ \phi^{-1}) \circ u_v$. Hence, u_v is least concave.

Therefore, to prove u_v is a least concave utility function, it suffices to show that ϕ is convex. Suppose, on the contrary, that $\phi((1-\alpha)a+\alpha b) > (1-\alpha)\phi(a) + \alpha\phi(b)$ for some $a, b \in w(\Omega)$ and $\alpha \in]0, 1[$. Let $x \in w^{-1}(a)$, $y \in w^{-1}(b)$ and $z \in w^{-1}((1-\alpha)a + \alpha b)$. Then, $x \sim u_v(x)v$, $y \sim u_v(y)v$ and $z \sim u_v(z)v$. By supposition, we have

$$u_v(z) = \phi((1-\alpha)a+\alpha b) > (1-\alpha)\phi(a)+\alpha\phi(b) = (1-\alpha)u_v(x)+\alpha u_v(y).$$

[6] It can be shown by the same argument as the proof of Proposition 3.C.1 of Mas-Colell, Whinston and Green [4].

[7] It can be verify by the same argument as Exercise 2-1 of Stokey and Lucas [5].

Hence, $u_v(z)v \succ [(1-\alpha)u_v(x) + \alpha u_v(y)]v$. Meanwhile, w is concave, and thus,

$$\begin{aligned} w(u_v(z)v) &= w(z) = (1-\alpha)w(x) + \alpha w(y) \\ &= (1-\alpha)w(u_v(x)v) + \alpha w(u_v(y)v) \\ &\leq w(((1-\alpha)u_v(x) + \alpha u_v(y))v). \end{aligned}$$

Thus, we have $u_v(z)v \precsim [(1-\alpha)u_v(x) + \alpha u_v(y)]v$, a contradiction. Hence, ϕ must be a positive affine transformation, and thus, u_v is a least concave utility. Since $u_v(0) = 0$ is trivial, we proved that (1) implies (2).

It is clear that (1) implies (3). Conversely, let u be homogeneous of degree one. Then, $u(0) = 0$ and $u(v) > 0$ for any $v \in P^{\succsim}$. Hence, there exists some $v \in P^{\succsim}$ such that $u(v) = 1$. Then we can easily show that $u = u_v$. Thus, (3) implies (1).

To verify that (2) implies (1), fix any u_v for some $v \in P^{\succsim}$. Then, u_v is a least concave utility function. Therefore, there exists some affine transformation $\phi : c \to ac + b$ such that $u = \phi \circ u_v$. Since $u(0) = u_v(0)$, we have $b = 0$ and thus $u = au_v = u_{\frac{1}{a}v}$, which completes the proof. ■

Proof of proposition. By assumption, there exists $w \in \Omega$ such that $w \succ 0$. Since $\succsim$ is homothetic, we have $tw \succ 0$ for any $t > 0$. Since $tw \in N$ for sufficiently small $t > 0$, we have $v \succ 0$.

Suppose $0 < a < b$ and $av \succsim bv$. Let $a_n = \frac{a^n}{b^{n-1}}$. Then $a = a_1$ and $b = a_0$. Since $\succsim$ is homothetic, $a_n v \succsim a_{n-1}v$. By transitivity, $a_n v \succsim bv$. Since $a_n v \to 0$, we have $0 \succsim bv$, a contradiction. This implies $bv \succsim av$ if and only if $b \geq a$ for any $a, b \geq 0$.

Finally, fix any $x \in \Omega$. If $x = 0$, then $1v \succ x \succsim 0v$. Otherwise, take any $t > 0$ such that $tx \in N$. Then $v \succ \frac{1}{2}tx$ and thus $2t^{-1}v \succ x \succsim 0v$. Therefore, $v \in P^{\succsim}$. ■

References

1. Debreu, G.: Least concave utility functions. J. Math. Econ. **3**, 121–129 (1976)
2. Kannai, Y.: The ALEP definition of complementarity and least concave utility functions. J. Econ. Theory **22**, 115–117 (1980)
3. Kihlstrom, R.E., Mirman, L.J.: Constant, increasing, and decreasing risk aversion with many commodities. Rev. Econ. Stud. **48**, 271–280 (1981)
4. Mas-Colell, A., Whinston, M.D., Green, J.R.: Microeconomic Theory. Oxford University Press, Oxford (1995)
5. Stokey, N., Lucas, R.: Recursive Methods in Economic Dynamics. Harvard University Press, Cambridge, MA (1989)

Subject Index