

Differentiability in Banach Spaces

PhD Thesis

Dean James Ives

Department of Mathematics

University College London

University of London

July 1998

ProQuest Number: 10797889

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10797889

Published by ProQuest LLC (2018). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Abstract

There are three chapters in this work of which the first two contain differentiability results for continuous convex functions on Banach spaces. The final chapter contains differentiability results for Lipschitz isomorphisms of ℓ_2 .

The aim of chapter 1 is to improve on a result of I. Ekeland and G. Lebourg [EL] who show that a Banach space E that admits a Lipschitz Fréchet smooth bump function is an Asplund space. It is shown that if E admits a continuous lower Fréchet smooth bump function then E is an Asplund space.

Chapter 2 contains partial results towards showing that there are Gâteaux differentiability spaces that are not weak Asplund spaces. Suppose that K is a totally ordered first countable Hausdorff compact space. A topology τ_w is defined on $C(K)$ called the wedge topology, and it is shown that if every subdifferential of a continuous convex function f on $C(K)$ contains a measure of finite support then f is Gâteaux differentiable on a τ_w residual set.

Chapter 3 contains three examples of Lipschitz isomorphisms of ℓ_2 to itself for which the derivative fails to be surjective; in the first example the Gâteaux derivative is not surjective at one point, in the second example the weak limit of $\lim_{t \rightarrow 0} (f(th) - f(0))/t$ is zero for all $h \in \ell_2$, and in the third example the Gâteaux derivative is not surjective at all points of the cube $\{x \in \ell_2 : |x_i| < 2^{-i} \text{ for all } i\}$ which is mapped affinely into a hyperplane.

Contents

1	Bump Functions	10
1.1	Introduction	10
1.2	Lemmata	10
1.3	An improvement on Ekeland-Lebourg	16
2	Gâteaux Differentiability on $C(K)$ Spaces	18
2.1	Introduction	18
2.2	Functions on ordered spaces	20
2.3	The oscillation and wedge topology	23
2.4	Directional derivatives	34
2.5	Dirac measures	36
2.6	Measures with finite support	45
2.7	Gâteaux differentiability	62
3	Lipschitz Isomorphisms	67
3.1	Introduction	67
3.2	A preliminary result	68
3.3	The first example	70
3.4	The second example	77
3.5	The third example	80

Acknowledgements

I would like to thank Professor D. Preiss for suggesting that I should study differentiability in Banach spaces. I am especially grateful for his invaluable advice, constant helpfulness and generosity with his time. I would also like to thank Professors D. G. Larman and J. E. Jayne and others at University College London for their help and support.

I acknowledge the generous support of The University of London during my studies.

Basic notation

There follows a list of the notation used in the text, which, although basic, is not completely standard.

A_L	the set of left accumulation points
A_R	the set of right accumulation points
\mathcal{B}_o	the set of basic neighbourhoods of 0 in the oscillation topology
\mathcal{B}_w	the set of basic neighbourhoods of 0 in the wedge topology
$B(x, r)$	the open ball centred at x with radius r
$\text{card } L$	the cardinality of the set L
$C(K)$	the Banach space of continuous functions on the compact space K
$d(g, V)$	the distance of g from the subspace V
$\partial f(\phi)$	the subdifferential of f at ϕ
\overline{G}	the closure of the set G
1_A	the characteristic function of the set A
$\mathcal{M}(K)$	the set of Radon measures on $C(K)$
$\mathcal{M}(F)$	the set of Radon measures on $C(K)$ with support contained in $F \subset K$
$\text{Lip}(f)$	the Lipschitz constant of f
\mathbf{R}	the real numbers
\mathbf{R}^+	the non-negative real numbers
τ_o	the oscillation topology
τ_w	the wedge topology
$U_o(0, \epsilon, \omega)$	a basic neighbourhood of 0 in the oscillation topology (definition 2.3.1)
$U_w(0, \epsilon, \{\phi_l\}_{l \in L}, \{\psi_r\}_{r \in R})$	a basic neighbourhood of 0 in the wedge topology (definition 2.3.2).

Introduction

The two notions of derivative used in this work are the Gâteaux derivative, and the Fréchet derivative. If X and Y are Banach spaces then the Gâteaux derivative of a mapping $\phi : X \mapsto Y$ at $x \in X$, is defined as a continuous linear map $\phi'(x) : X \mapsto Y$ such that

$$\phi'(x)u = \lim_{t \rightarrow 0} \frac{\phi(x + tu) - \phi(x)}{t}$$

for every $u \in X$. For the Fréchet derivative we require in addition that the above limit be uniform for $\|u\| \leq 1$.

We refer to [DGZ3], [Fa], [Gi], and [Ph1] for a comprehensive treatment of Asplund spaces and related concepts. A Banach space E is said to be an Asplund space if every continuous convex function on E is Fréchet differentiable on a residual set. If every continuous convex function on E is Gâteaux differentiable on a residual set then E is called a weak Asplund space, and if every continuous convex function on E is Gâteaux differentiable on a dense set then E is called a Gâteaux differentiability space (GDS).

In chapter 1 we show that if E admits a continuous lower Fréchet smooth bump function then E is an Asplund space. This improves on a result of I. Ekeland and G. Lebourg [EL] that provided E admits a Fréchet smooth bump function then E is an Asplund space. To obtain differentiability points of a continuous convex function on a Banach space one method is to apply a variational principle of which we note Ekeland's variational principle [Ek], the Borwein-Preiss variational principle [BP], and a general variational principle by R. Deville, G. Godefroy, and V. Zizler [DGZ2] from which most results obtained previously by Ekeland's variational principle or the Borwein-Preiss variational principle easily follow. Recently M. Fabian, G. Godefroy, and J. Vanderwerff [FHV] have obtained a smooth variational principle in the case of a Banach space that admits a Fréchet differentiable bump function. All these results imply that a Banach space that admits a Fréchet differentiable bump function is an Asplund space.

For results on Gâteaux differentiability, Deville's version of the Borwein-Preiss variational principle implies that, if E admits a Lipschitz Gâteaux smooth bump function then E is a GDS; it is even a weak Asplund space as shown by a slight adjustment in a result of D. Preiss, R. R. Phelps, and I. Namioka in [PPN] (it is shown that a Banach space with Gâteaux smooth norm is a weak Asplund space) done in M. Fosgerau's thesis. In the Gâteaux case we do not know whether relaxing the condition on the bump function to continuous Gâteaux lower smooth implies that E is a GDS. For the converse it is not known whether an Asplund space admits a lower Fréchet smooth bump function; Richard Haydon [Ha] exhibits an Asplund space that admits no Gâteaux smooth (differentiable at non-zero vectors) equivalent norm, which strongly refutes the converse of the results in [PPN].

We can show, using our method, that if E admits a continuous Gâteaux lower smooth bump function then E is an ϵ -GDS in the sense that for any continuous convex function f and any $y \in E$ the smooth limit

$$\lim_{t \rightarrow 0} \frac{f(x + ty) + f(x - ty) - 2f(x)}{t}$$

is less than ϵ for all points x in a dense subset of E .

In the second chapter we obtain partial results towards showing that there is a GDS that is not weak Asplund. In particular letting K be an ordered first countable Hausdorff space that is compact in the order topology, we define a topology τ_w (the wedge topology) on $C(K)$, for which all continuous convex functions f on $C(K)$ such that each subdifferential contains a Radon measure of finite support, are Gâteaux differentiable on a τ_w -residual subset of $C(K)$. In this connection M. Talagrand [Ta1] has shown that the set of points of Gâteaux differentiability of a convex continuous function f need not be G_δ . M. M. Coban and P. S. Kenderov [CK] have observed that the set of points of Gâteaux differentiability of the sup-norm on the double arrow space D , which may be described as the space of functions on the unit interval with a right limit at 0, a left limit at 1, and left and right limits at every point), is dense but not residual. M. Talagrand [Ta2] gives a proof that there does not exist an equivalent Gâteaux smooth norm on D . More generally it is known from a result of D. Preiss, R. R. Phelps, and I. Namioka in [PPN] that a Banach space that admits a Gâteaux smooth norm is a weak Asplund space; in view of the Coban-Kenderov statement this gives another proof of Talagrand's result.

Chapter three contains three examples of Lipschitz isomorphisms of ℓ_2 for which the derivative fails to be surjective: in the first example the Gâteaux derivative is not surjective at one point, in the second example the weak limit

of $\lim_{t \rightarrow 0} (f(th) - f(0))/t$ is zero for all $h \in \ell_2$, and in the third example the Gâteaux derivative is not surjective at all points of the cube $\{x \in \ell_2 : |x_i| < 2^{-i} \text{ for all } i\}$ which is mapped affinely into a hyperplane.

All these examples have some connection with the linear isomorphism problem for Banach spaces which asks whether Lipschitz isomorphic Banach spaces are necessarily linearly isomorphic. One method of solution is to look for points at which the derivative of a given Lipschitz isomorphism $f : X \rightarrow Y$ exists and is surjective. For Lipschitz maps between finite dimensional Banach spaces Rademacher's theorem states that on \mathbf{R}^n any Lipschitz map into \mathbf{R}^m is differentiable everywhere except on a set of Lebesgue measure zero. Extensions of Rademacher's theorem to infinite dimensions have been found by N. Aronszajn [Ar] using Aronszajn null sets, and by R. R. Phelps [Ph2] using Gaussian null sets; that Gaussian null and Aronszajn null sets are equivalent has been shown by M. Csörnyei [Cs]. With different (weaker) versions of null sets similar results have been proved by P. Mankiewicz ([Mn]) and J. P. R. Christensen ([Cr]). If it were known that Lipschitz homeomorphisms carry null sets to null sets we could obtain a differentiability point x of f for which f^{-1} is differentiable at $f(x)$. It would then follow that $f'(x)$ is an isomorphism. That Aronszajn null sets are not preserved under Lipschitz isomorphisms was shown by V. I. Bogachev [Bo]. There is a recent example by E. Matoušková [Mt], of a Lipschitz isomorphism of a separable Banach space to itself that maps a non Haar null set (see [Cr]) to an Aronszajn null set. Our third example is another example which shows that Aronszajn null sets are not preserved.

N. Aronszajn [Ar] obtains the following extension of Rademacher's theorem.

Let f be a Lipschitz map from a separable Banach space X into a space Y with the Radon-Nykodim property (RNP). Then f is Gâteaux differentiable everywhere except on an Aronszajn null set.

We refer to [DU] for the RNP; a Banach space Y has the RNP if every Lipschitz map $g : \mathbf{R} \rightarrow Y$ is differentiable almost everywhere. Examples of spaces that do not have the RNP include c_0 and $L_1(0, 1)$. Reflexive spaces do have the RNP. The set of *Aronszajn null* sets \mathcal{U} is constructed as follows (see [Ar]). Let E be a Banach space and let $a \in E$ be non-zero then we let

- $\mathcal{U}(a) = \{A \subset E : \text{For all } x \in E \text{ the set } A \cap (x + \mathbf{R}a) \text{ is of Lebesgue measure zero on the line } x + \mathbf{R}a\},$

- for every sequence $(a_n)_{n=1}^\infty \subset E$ with $a_n \neq 0$ we let $\mathcal{U}((a_n)_{n=1}^\infty) = \{ A \subset E : A = \cup A_n, A_n \in \mathcal{U}(a_n) \}$, and
- $\mathcal{U} = \cap \mathcal{U}((a_n)_{n=1}^\infty)$ where the intersection is over all complete sequences in E . (A complete sequence is one whose closed linear span is E .)

Any hyperplane in ℓ_2 is Aronszajn null since in any complete sequence in ℓ_2 there is a line $\mathbf{R}a$ such that any translate of $\mathbf{R}a$ intersects the hyperplane in a one-element set, which is of Lebesgue measure zero.

Chapter 1

Bump Functions

1.1 Introduction

We aim to prove Theorem 1.3.1 which is the following statement.

Let E be a Banach space which admits a continuous lower Fréchet smooth bump function then E is an Asplund space.

We recall some definitions. Let E be a Banach space. A *bump function* on E is a function $b : E \rightarrow \mathbf{R}$ that has bounded non-empty support and attains a positive value. We say that a function $\phi : E \rightarrow \mathbf{R}$ is *Fréchet differentiable* at $x \in E$ if there is a continuous linear functional $\phi'(x)$, called the *Fréchet derivative* of ϕ at x , such that

$$\lim_{\|h\| \rightarrow 0} \frac{\phi(x+h) - \phi(x) - \langle \phi'(x), h \rangle}{\|h\|} = 0.$$

A function $\phi : E \rightarrow \mathbf{R}$ is *lower Fréchet smooth* at x if

$$\liminf_{h \rightarrow 0} \frac{\phi(x+h) + \phi(x-h) - 2\phi(x)}{\|h\|} \geq 0.$$

A set $S \subset E$ is *residual* if it is the complement of a first category set in E . A Banach space E is an *Asplund space* if every continuous convex function on E is Fréchet differentiable on a residual set.

1.2 Lemmata

We aim to establish the four Lemmata needed to prove Theorem 1.3.1. Lemma 1.2.1 is a version of Ekeland's variational principle which has found many applications in non-linear analysis.

Lemma 1.2.1 *Let (T, d) be a complete metric space and let $f : T \rightarrow \mathbf{R}$ be continuous and bounded above, then given $\epsilon > 0$ there is a $z \in T$ such that*

$$f(x) \leq f(z) + \epsilon d(x, z)$$

for all $x \in T$.

PROOF. The point z is the limit of a sequence $(x_n)_{n=1}^\infty$, which we construct inductively as follows. Let x_0 be any point in T . Suppose that for $n \geq 0$ we have constructed the point x_n . We define a subset M_n by

$$M_n = \{ x \in T \mid f(x) - f(x_n) \geq \epsilon d(x, x_n) \}.$$

Since $x_n \in M_n$, we may let $S_n = \sup_{x \in M_n} f(x)$ and choose a point $x_{n+1} \in M_n$ such that $f(x_n) \geq S_n - 2^{-n}$. We claim that

1. the sequence $(f(x_n))_{n=1}^\infty$ is increasing,
2. the sequence $(x_n)_{n=1}^\infty$ is a Cauchy sequence which converges to a point z ,
3. If $m > n$ then $x_m \in M_n$, and
4. $f(x) \leq f(z) + \epsilon d(x, z)$ for all $x \in T$.

For (1), using the definition of M_n and that $x_{n+1} \in M_n$, we have

$$\begin{aligned} f(x_{n+1}) - f(x_n) &\geq \epsilon d(x_{n+1}, x_n) \\ &\geq 0. \end{aligned}$$

For (2), since $(f(x_n))_{n=1}^\infty$ is increasing and bounded above, it is convergent and therefore a Cauchy sequence. Given $\kappa > 0$ we may choose a positive integer N such that for all $m > n \geq N$ we have $f(x_m) - f(x_n) < \kappa$. For all $m > n \geq N$, we have

$$\begin{aligned} \epsilon d(x_m, x_n) &\leq \sum_{i=n}^{m-1} \epsilon d(x_{i+1}, x_i) \\ &\leq \sum_{i=n}^{m-1} [f(x_{i+1}) - f(x_i)] \\ &= f(x_m) - f(x_n) \\ &< \kappa. \end{aligned} \tag{1.1}$$

Therefore $(x_n)_{n=1}^\infty$ is a Cauchy sequence which must converge to some point z .

For (3) we see from (1.1) that $f(x_m) - f(x_n) \geq \epsilon d(x_m, x_n)$. Therefore $x_m \in M_n$.

For (4), supposing that it is not true that $f(x) \leq f(z) + \epsilon d(x, z)$ for all $x \in T$, we may find x such that $f(x) > f(z) + \epsilon d(x, z) + \alpha$ for some $\alpha > 0$. We note that

- f is continuous,
- M_n is closed for all $n \geq 0$,
- $\lim_{m \rightarrow \infty} x_m = z$, and
- $x_m \in M_n$ for all $m > n$.

Therefore $z \in M_n$ for all n . We obtain

$$\begin{aligned} f(x) - f(x_n) &> f(z) - f(x_n) + \epsilon d(x, z) + \alpha \\ &\geq \epsilon(d(z, x_n) + d(x, z)) + \alpha \\ &\geq \epsilon d(x, x_n) + \alpha \end{aligned} \tag{1.2}$$

implying that $x \in M_n$ for all n . Since $\epsilon d(x, x_n) \geq 0$ we have from (1.2) that $f(x) - f(x_n) \geq \alpha$. But if $x \in M_{n-1}$ then $f(x) \leq S_{n-1}$ so that

$$2^{n-1} \geq S_{n-1} - f(x_n) \geq \alpha.$$

□

Lemma 1.2.2 is a version of the Hahn-Banach theorem which we state without proof.

Lemma 1.2.2 *Let E be a Banach space and M a linear subspace of E and suppose that there is a linear functional $x_M^* \in M^*$ such that $\|x_M^*\| = 1$. Then there is a linear functional x^* on E such that $\|x^*\| = 1$ and*

$$\langle x^*, x \rangle = \langle x_M^*, x \rangle$$

for all $x \in M$.

Let $\epsilon > 0$, then we define a norm $\|\cdot\|$ to be ϵ -rough at z if

$$\limsup_{h \rightarrow 0} \frac{\|z+h\| + \|z-h\| - 2\|z\|}{\|h\|} \geq \epsilon. \quad (1.3)$$

We note that, by the triangle inequality, the value of ϵ cannot exceed 2. A norm is said to be ϵ -rough if it is ϵ -rough at all z . Lemma 1.2.3 is a result of E. B. Leach and J. H. M. Whitfield [LW].

Lemma 1.2.3 *A Banach space E is not an Asplund space if and only if for some $\epsilon > 0$ E admits an equivalent ϵ -rough norm.*

In Lemma 1.2.4 we show that if a norm is ϵ -rough at z then for any $\delta > 0$ we can find $\|\hat{h}\| < 2\delta$ for which $\frac{\|z+\hat{h}\| - \|z\|}{\|\hat{h}\|} \geq \epsilon/8$ and $\frac{\|z-\hat{h}\| - \|z\|}{\|\hat{h}\|} \geq \epsilon/8$.

Lemma 1.2.4 *Let $\|\cdot\|$ be a norm on a Banach space E that is ϵ -rough at z . Then for all δ for which $\frac{\epsilon\|z\|}{4} > \delta > 0$ there is a \hat{h} such that*

$$(i) \quad \|z + \hat{h}\| = \|z - \hat{h}\|,$$

$$(ii) \quad \|\hat{h}\| < 2\delta, \text{ and}$$

$$(iii) \quad \frac{\|z+\hat{h}\| - \|z\|}{\|\hat{h}\|} \geq \frac{\epsilon}{8}.$$

PROOF. The point \hat{h} is obtained from h by subtracting a small z component. Applying the Hahn-Banach theorem we find a functional x^* which attains its norm in the z direction and subtracting x^* from the norm we obtain a convex function f with which we may estimate $\frac{\|z+\hat{h}\| - \|z\|}{\|\hat{h}\|}$.

Suppose that h and δ are such that, $\frac{\epsilon\|z\|}{4} > \delta > 0$, $\|h\| < \delta$, $\frac{\|z+h\| + \|z-h\| - 2\|z\|}{\|h\|} \geq \epsilon$ and that $\|z+h\| > \|z-h\|$. To construct \hat{h} we first define $\mathbf{a} : \mathbf{R} \rightarrow E$ by $\mathbf{a}(t) = h - tz$. Defining $a : \mathbf{R} \rightarrow \mathbf{R}$ by $a(t) = \|z + \mathbf{a}(t)\| - \|z - \mathbf{a}(t)\|$, and noting that

- $a(0) > 0$, and
- $a(1) = \|h\| - \|2z - h\| < 0$,

we may use the intermediate value theorem to obtain a real number $t_0 \in (0, 1)$ with $a(t_0) = 0$. We let $\hat{h} = \mathbf{a}(t_0)$ so that

$$\|z - \hat{h}\| = \|z + \hat{h}\| \quad (1.4)$$

which is property (i) of \hat{h} . Substituting $\hat{h} = \mathbf{a}(t_0) = h - t_0 z$ in equation (1.4) gives $\|z - h + t_0 z\| = \|z + h - t_0 z\|$ and applying the triangle inequality we obtain $(1 + t_0)\|z\| - \|h\| \leq (1 - t_0)\|z\| + \|h\|$ which simplifies as

$$t_0\|z\| \leq \|h\|. \quad (1.5)$$

Therefore

$$\begin{aligned} \|\hat{h}\| &= \|h - t_0 z\| \\ &\leq \|h\| + t_0\|z\| \\ &\leq 2\|h\| \\ &< 2\delta. \end{aligned}$$

property (ii) of \hat{h} .

Applying the Hahn-Banach theorem in the form of Lemma 1.2.2 with $M = \text{span } \{z\}$, and x_M^* defined by $\langle x_M^*, tz \rangle = t\|z\|$ for all $tz \in \text{span } \{z\}$, we find a functional x^* with $\|x^*\| = 1$ and $\langle x^*, z \rangle = \|z\|$. Defining $f : E \rightarrow \mathbf{R}$ by $f(x) = \|x\| - \langle x^*, x \rangle$ we claim that f has the following properties:

1. f is convex and non-negative,
2. $\text{Lip } (f) \leq 2$,
3. $f(\lambda z) = 0$ for all $\lambda \geq 0$,
4. $f(\lambda x) = \lambda f(x)$ for all $\lambda \geq 0$,
5. $f(z + h) \leq f(z + \hat{h})$, and
6. $\frac{f(z + \hat{h}) + f(z - \hat{h})}{\|\hat{h}\|} \geq \epsilon/4$.

For (1), f is the sum of $\|\cdot\|$ and $-x^*$ and is therefore convex. Since $\|x^*\| \leq 1$, then f is non-negative.

For (2),

$$\begin{aligned} \text{Lip } (f) &\leq \text{Lip } (\|\cdot\|) + \text{Lip } (x^*) \\ &\leq 2. \end{aligned}$$

For (3) we have for any $\lambda \geq 0$ that

$$\begin{aligned} f(\lambda z) &= \|\lambda z\| - \langle x^*, \lambda z \rangle \\ &= \lambda(\|z\| - \langle x^*, z \rangle) \\ &= 0. \end{aligned}$$

For (4) we have for any $\lambda \geq 0$ that

$$\begin{aligned} f(\lambda x) &= \|\lambda x\| - \langle x^*, \lambda x \rangle \\ &= \lambda(\|x\| - \langle x^*, x \rangle) \\ &= \lambda f(x). \end{aligned}$$

For (5), since $h - \hat{h} = t_0 z$, we have $\langle x^*, h - \hat{h} \rangle = \|h - \hat{h}\|$. Hence

$$\begin{aligned} f(z + \hat{h}) - f(z + h) &= \|z + \hat{h}\| - \|z + h\| + \langle x^*, h - \hat{h} \rangle \\ &= \|z + \hat{h}\| - \|z + h\| + \|h - \hat{h}\| \\ &\geq 0. \end{aligned}$$

For (6) we let $\beta = 1/(1 + t_0)$ so that $0 < \beta \leq 1$ and recalling that $h = \hat{h} + t_0 z$, we obtain

$$\begin{aligned} z - \beta h &= \frac{z(1 + t_0) - h}{1 + t_0} \\ &= \beta(z - \hat{h}). \end{aligned}$$

Applying property (4) of f to $f(z - \beta h)$ with f non-negative gives

$$\begin{aligned} f(z - \beta h) &= \beta f(z - \hat{h}) \\ &\leq f(z - \hat{h}). \end{aligned} \tag{1.6}$$

Using (1.6), and properties (5) and (2) of f , we estimate

$$\begin{aligned} f(z + h) + f(z - h) &\leq f(z + \hat{h}) + f(z - h) \\ &\leq f(z + \hat{h}) + f(z - h) + (f(z - \hat{h}) - f(z - \beta h)) \\ &\leq f(z + \hat{h}) + f(z - \hat{h}) + 2(1 - \beta)\|h\|. \end{aligned} \tag{1.7}$$

From (1.5) we have:

$$\begin{aligned} 1 - \beta &= t_0/(1 + t_0) \\ &\leq t_0 \\ &\leq \|h\|/\|z\|. \end{aligned} \tag{1.8}$$

Therefore substituting (1.8) and $\|\hat{h}\| \leq 2\|h\|$ in (1.7) gives

$$\begin{aligned} \frac{f(z + \hat{h}) + f(z - \hat{h})}{\|\hat{h}\|} &\geq \frac{f(z + \hat{h}) + f(z - \hat{h})}{\|2h\|} \\ &\geq \frac{f(z + h) + f(z - h)}{\|2h\|} - (1 - \beta) \\ &\geq \frac{f(z + h) + f(z - h)}{\|2h\|} - \frac{\|h\|}{\|z\|}. \end{aligned} \tag{1.9}$$

By hypothesis $\|h\| < \delta < \frac{\epsilon\|z\|}{4}$ so that

$$\begin{aligned} \frac{f(z+h) + f(z-h)}{\|2h\|} &= \frac{\|z+h\| + \|z-h\| - 2\|z\|}{\|2h\|} \\ &\geq \frac{\epsilon}{2}. \end{aligned}$$

From (1.9) we have $\frac{f(z+\hat{h})+f(z-\hat{h})}{\|\hat{h}\|} \geq \frac{\epsilon}{4}$ and using $\|z+\hat{h}\| = \|z-\hat{h}\|$ we deduce that

$$\frac{\|z+\hat{h}\| - \|z\|}{\|\hat{h}\|} \geq \frac{\epsilon}{8}.$$

□

1.3 An improvement on Ekeland-Lebourg

The following Theorem improves on the result of Ekeland and Lebourg that a Banach space that admits a Fréchet differentiable bump function is an Asplund space.

Theorem 1.3.1 *Let E be a Banach space which admits an upper semicontinuous lower Fréchet smooth bump function, then E is an Asplund space.*

PROOF. We argue by contradiction. Suppose that E is not an Asplund space so that, by Lemma 1.2.3, E admits an equivalent ϵ -rough norm for some $\epsilon > 0$. Letting $S = \{x : 4b(x) + \|x\| \geq 3\}$ where we may suppose that b is a lower Fréchet smooth bump function such that $b(0) = 1$ and $\text{spt}(b) \subset B(0, 1)$, and applying Ekeland's variational principle (Lemma 1.2.1) to $T = S \cap B(0, 2)$ we obtain an Ekeland maximum point of the norm at which the bump function b is not lower Fréchet smooth.

The set T is non-empty since $0 \in T$. We show that $T \subset B(0, 1)$. Indeed if $x \in T \setminus \text{spt}(b)$ then, since $b(x) = 0$ and $x \in S$, we must have $\|x\| \geq 3$. But $T \subset B(0, 2)$. Therefore $T \setminus \text{spt}(b) = \emptyset$. Hence $T \subset \text{spt}(b) \subset B(0, 1)$.

Define a metric on T by $d(x, y) = \|x - y\|$. Since b is upper semicontinuous S is closed; so T is complete and Ekeland's variational principle (Lemma 1.2.1) is applicable in T . For the continuous function $\|\cdot\|$, which is bounded above, we obtain a point z such that

$$\|x\| \leq \|z\| + \epsilon \frac{\|x - z\|}{16}$$

for all $x \in T$. Since b is lower smooth $T \neq \{0\}$ so that $z \neq 0$.

Let $0 < \delta < \frac{\epsilon\|z\|}{4}$ be such that

$$\frac{b(z+h) + b(z-h) - 2b(z)}{\|h\|} > -\frac{\epsilon}{16}$$

for $0 < \|h\| < 2\delta$. Since the norm is ϵ -rough at z we may apply Lemma 1.2.4 to obtain a point $\hat{h} \in E$ with the following properties:

- $\|\hat{h}\| \leq 2\delta < 1$,
- $\frac{\|z+\hat{h}\| - \|z\|}{\|\hat{h}\|} \geq \epsilon/8$, and
- $\|z + \hat{h}\| = \|z - \hat{h}\|$.

We claim that $z + \hat{h} \notin T$, and $z - \hat{h} \notin T$.

If $z + \hat{h} \in T$ then Ekeland's variational principle implies that $\|z + \hat{h}\| \leq \|z\| + \frac{\epsilon\|\hat{h}\|}{16}$ which contradicts the roughness of the norm, that is, $\|z\| + \frac{\epsilon\|\hat{h}\|}{8} \leq \|z + \hat{h}\|$. Similarly $z - \hat{h} \in T$ would imply that

$$\|z\| + \frac{\epsilon\|\hat{h}\|}{8} \leq \|z - \hat{h}\| \leq \|z\| + \frac{\epsilon\|\hat{h}\|}{16}.$$

Since $\|z\| \leq 1$ and as $\|\hat{h}\| \leq 1$ we have $z + \hat{h} \in B(0, 2)$ and $z - \hat{h} \in B(0, 2)$. From the definition of T we deduce that $z + \hat{h} \notin S$ and $z - \hat{h} \notin S$. From the definition of the set S we obtain $4b(z+\hat{h}) + \|z+\hat{h}\| < 3$ and $4b(z-\hat{h}) + \|z-\hat{h}\| < 3$. It follows that

$$0 < 4 \frac{b(z+\hat{h}) + b(z-\hat{h}) - 2b(z)}{\|\hat{h}\|} + \frac{\|z+\hat{h}\| + \|z-\hat{h}\| - 2\|z\|}{\|\hat{h}\|} < 0$$

which is a contradiction. Hence E is an Asplund space. \square

Chapter 2

Gâteaux Differentiability on $C(K)$ Spaces

2.1 Introduction

We are working on *(totally) ordered spaces* K which we always consider in their order topology. We will assume that K is compact and first countable. We refer to [Na] for a comprehensive treatment of the order topology. For arbitrary a, b we write

$$(a, b) = \{x : \min\{a, b\} < x < \max\{a, b\}\},$$

$$[a, b] = \{x : \min\{a, b\} \leq x \leq \max\{a, b\}\}.$$

The open interval (a, b) is an open set. Our assumption that K is compact implies that K is order complete, that is every monotone sequence is convergent; also every nonempty subset of K has a supremum and infimum. A non-trivial example of an ordered space is given in the text on page 65 where we define the space I_s of signed points $(x, 1), (x, -1)$ on the unit interval $[0, 1]$. $C(K)$ denotes the set of continuous functions on K with the usual norm on $C(K)$ defined by $\|g\| = \sup_{x \in K} |g(x)|$.

In section 2.3 we introduce the *oscillation and wedge topologies*, τ_o and τ_w on $C(K)$, we show that they are equivalent and we deduce the main properties of τ_o . In section 2.4 we obtain a differentiability result for continuous convex functions on a Banach space E for which there is a topology τ finer than the norm topology in which the neighbourhoods of 0 satisfy a geometric condition. This result is used in section 2.7 to show that a continuous convex function f on $C(I_s)$ (which is essentially the space D of functions on $[0, 1]$ with a right limit at 0, a left limit at 1, and left and right limits at all

other points) for which each subdifferential $\partial f(\varphi)$ contains only non-atomic measures is Gâteaux differentiable on a τ_o residual set. Section 2.5 contains results for continuous convex functions f on $C(K)$ for which each subdifferential $\partial f(\varphi)$ contains a Dirac measure. In section 2.6 we extend the results of section 2.5 to the case of subdifferentials that always contain a measure with finite support. In the final section we prove our main result, Theorem 2.6.7, which states that *a continuous convex function f on $C(K)$ for which each subdifferential $\partial f(\varphi)$ contains a measure with finite support is Gâteaux differentiable on a τ_o residual set.* Our results are far short showing that $C(I_s)$ is a Gâteaux differentiability space since we have only considered the two extreme cases.

Recalling the definitions of some topological concepts, we have that a compact Hausdorff space K is *first countable* if for every $x \in K$ there is a sequence $(U_n(x))_{n=1}^\infty$ of neighbourhoods of x such that if U is any neighbourhood of x then there is a neighbourhood U_k in the sequence such that $U_k \subset U$.

A topological space E is called a *Baire space* if for any set $X \subset E$, that is a countable union of nowhere dense sets, the complement $E \setminus X$ is dense in E . The *Banach-Mazur game* [Ox] is a two person game with players A and B as follows. Let S be a subset of a topological space E . A *play* is a decreasing sequence

$$U_1 \supset V_1 \supset U_2 \supset \dots$$

of non-empty open subsets of E which have been chosen alternately by A and B . Player A chooses U_1 , B chooses V_1 , A chooses U_2 , etc. A *strategy for B* is a sequence $f_B = (f_n)_{n=1}^\infty$ of maps f_n where each f_n is defined on the set

$$\{U_1, V_1, U_2, \dots, U_n\}$$

of first $2n - 1$ elements of a play and

$$f_n(U_1, V_1, U_2, \dots, U_n)$$

is a non-empty open subset of U_n . A play is *consistent* with f_B if

$$f_n(U_1, V_1, U_2, \dots, U_n) = V_n$$

for all n . We say that f_B is a *winning strategy* for B if

$$\bigcap_{i=1}^\infty V_i \subset S$$

for every play consistent with f_B . We note the following properties of the Banach-Mazur game.

- A subset S of a topological space E is residual if and only if there is a winning strategy for B [Ox].
- A topological space E is a Baire space if player B has a strategy f_B such that

$$\bigcap_{i=1}^{\infty} V_i \neq \emptyset$$

for all plays consistent with f_B [Ch].

2.2 Functions on ordered spaces

We study the oscillation of a function on K and include two existence lemmata for $C(K)$ functions.

Definition 2.2.1 Let $f : K \rightarrow \mathbf{R}$ be a real valued function defined on K . Let $[a, b] \subset K$ be a closed interval with end points a and b with a not necessarily less than b . We define the oscillation of f on $[a, b]$ as

$$\text{osc}(f, [a, b]) = \sup_{u, v \in [a, b]} |f(u) - f(v)|.$$

Lemma 2.2.2 Let $[a, b] \subset K$ be a closed interval and $g, h \in C(K)$, then

- (i) $\text{osc}(f + g, [a, b]) \leq \text{osc}(f, [a, b]) + \text{osc}(g, [a, b])$ and
- (ii) $\text{osc}(fg, [a, b]) \leq \text{osc}(f, [a, b]) \sup |g| + \text{osc}(g, [a, b]) \sup |f|$
where the supremum is taken over $[a, b]$.

PROOF. Apply the triangle inequality for the modulus of the sum of two functions. This fact coupled with the fact that the supremum of the sum of two functions on a set is less than the sum of their suprema is sufficient to prove both statements. \square

Lemma 2.2.3 Let $(g_n)_{n=1}^{\infty}$ be a sequence of functions in $C(K)$ that converges pointwise to a function $g \in C(K)$. Then for any interval $[a, b] \subset K$

- (i) $\liminf_{n \rightarrow \infty} \text{osc}(g_n, [a, b]) \geq \text{osc}(g, [a, b])$ and
- (ii) if $(g_n)_{n=1}^{\infty}$ converges uniformly to g then

$$\lim_{n \rightarrow \infty} \text{osc}(g_n, [a, b]) = \text{osc}(g, [a, b]).$$

PROOF. For (i) suppose that $(g_n)_{n=1}^\infty$ converges pointwise to g and $[a, b]$ is any interval. Since $[a, b]$ is compact there are points $u, v \in [a, b]$ such that

$$\text{osc}(g, [a, b]) = |g(u) - g(v)|.$$

But for all $n \geq 1$

$$\text{osc}(g_n, [a, b]) \geq |g_n(u) - g_n(v)|.$$

Therefore

$$\liminf_{n \rightarrow \infty} \text{osc}(g_n, [a, b]) \geq \lim_{n \rightarrow \infty} |g_n(u) - g_n(v)| = \text{osc}(g, [a, b]).$$

For (ii) we apply the triangle inequality,

$$|\text{osc}(g_n, [a, b]) - \text{osc}(g, [a, b])| \leq \text{osc}(g_n - g, [a, b]) \leq 2\|g_n - g\|.$$

□

Lemma 2.2.4 is a monotonic version of Urysohn's lemma. This result is well known and proved in [Na] (page 30) under much more general conditions.

Lemma 2.2.4 *Let $a, b \in K$ and $a < b$ then there is a non-decreasing function $g \in C(K)$ such that*

$$g(x) = \begin{cases} 0 & \text{if } x \leq a \\ 1 & \text{if } x \geq b. \end{cases}$$

PROOF. We say that an interval $[a, b]$ has a *gap* if there are points x and y in $[a, b]$ such that $(x, y) = \emptyset$.

Suppose that the interval $[a, b]$ has a gap (x, y) . It suffices to let g be the function defined by

$$g(z) = \begin{cases} 0 & \text{if } z \leq x \\ 1 & \text{if } z \geq y. \end{cases}$$

Suppose that there are no gaps in $[a, b]$. We construct g as follows. Let $\mathbb{Q} \cap [0, 1] = (r_i)_{i=0}^\infty$ be an enumeration of the rationals in the closed unit interval such that $r_0 = 0$ and $r_1 = 1$.

We find a sequence, $(b_i)_{i=0}^\infty$, in K such that

- $b_i < b_j$ whenever $r_i < r_j$,
- $b_0 = a$, and
- $b_1 = b$.

The sequence b_i may be constructed inductively as follows. Suppose that $k \geq 1$ and we have constructed $(b_i)_{i=0}^k$. Let

$$r_l = \max\{r_i : 0 \leq i \leq k \text{ and } r_i < r_{k+1}\}$$

and

$$r_m = \min\{r_i : 0 \leq i \leq k \text{ and } r_i > r_{k+1}\}.$$

Since $r_l < r_m$ and there are no gaps in $[a, b]$ then we may choose a point $x \in (b_l, b_m)$ and let $b_{k+1} = x$.

We claim that the function g defined by

$$g(z) = \begin{cases} \inf_{b_i \geq z} r_i & \text{if } z \leq b \\ 1 & \text{if } z \geq b. \end{cases}$$

is continuous, non-decreasing, $g(z) = 0$ for $z \leq a$, and $g(z) = 1$ for $z \geq b$.

We have $g(a) = \inf_{b_i \geq a} r_i = r_0 = 0$ and $g(b) = \inf_{b_i \geq b} r_i = r_1 = 1$. If $y_1 < y_2$, noting that $\{b_i : y_1 \leq b_i\} \supset \{b_i : y_2 \leq b_i\}$, then $g(y_1) \leq g(y_2)$. It remains to show that g is continuous. Let $L_i = \{x \in K : x < b_i\}$ and $R_i = \{x \in K : x > b_i\}$. Then L_i and R_i are open sets. Given any $z \in [0, 1]$ the function g satisfies the relations

- $g^{-1}([0, z)) = \cup_{r_i < z} L_i$ and
- $g^{-1}((z, 1]) = \cup_{r_i > z} R_i$.

Therefore the inverse image of any open set under g is open and therefore g is continuous. \square

Lemma 2.2.5 asserts that there is a continuous function that dominates a given bounded function.

Lemma 2.2.5 *Let $\hat{\eta} : K \rightarrow \mathbf{R}$ be a bounded function on K such that*

$$\lim_{s \rightarrow t} \hat{\eta}(s) = 0$$

for some $t \in K$. Then there is a function $\eta \in C(K)$ such that

- (i) $\eta(s) > \hat{\eta}(s)$ if $s \neq t$ and
- (ii) $\eta(t) = 0$.

PROOF. We define $\eta(t) = 0$. To define $\eta(s)$ for $s > t$ we distinguish two cases: when t is isolated from the right for which we let

$$\eta(s) = 1 + \sup_{u>t} |\hat{\eta}(u)|$$

for all $s > t$; and when t is not isolated from the right for which we find a sequence $t_1 > t_2 > \dots > t$ such that $\hat{\eta}(u) < 2^{-i-1}$ for each $t < u \leq t_i$ and $i \geq 1$. Letting $\eta(s) = \sup |\hat{\eta}| + 1$ if $s \geq t_1$ and on each interval $[t_{i+1}, t_i]$, $i \geq 2$ using Lemma 2.2.4 we obtain a non-decreasing function g_i such that

$$g_i(z) = \begin{cases} 2^{-i} & \text{if } z = t_i \\ 2^{-i-1} & \text{if } z = t_{i+1}. \end{cases}$$

On $[t_2, t_1]$ similary find g_2 such that

$$g_2(z) = \begin{cases} \sup |\hat{\eta}| + 1 & \text{if } z = t_1 \\ 2^{-2} & \text{if } z = t_2. \end{cases}$$

We let $\eta(s) = g_i(s)$ whenever $s \in [t_{i+1}, t_i]$. The definition of $\eta(s)$ for $s < t$ is similar. \square

2.3 The oscillation and wedge topology

We introduce the *oscillation topology* and the *wedge topology* for $C(K)$. In Lemma 2.3.3 we show that these topologies are equivalent and in Lemma 2.3.5 we list their main properties.

Definition 2.3.1 Suppose that $\epsilon > 0$ and $\omega \in C(K)$ is a non-negative function such that $\text{card} \{ x \in K : \omega(x) = 0 \}$ is finite. Then we denote by B_o the set of subsets $U_o(0, \epsilon, \omega) \subset C(K)$ where $g \in U_o(0, \epsilon, \omega)$ if and only if

- $|g(x)| < \epsilon$ for all $x \in K$ and
- for all $[a, b] \subset K$ such that a is an accumulation point of $[a, b]$, and $\omega(a) = 0$ we have

$$\text{osc}(g, [a, b]) < \sup_{z \in [a, b]} \omega(z).$$

Let \mathbf{A} denote the set of accumulation points of K . We let \mathbf{A}_L denote the set of left accumulation points of K defined as

$$\mathbf{A}_L = \{ l \in K : (l, a) \neq \emptyset \text{ for all } a > l \},$$

and \mathbf{A}_R denote the set of right accumulation points of K defined as

$$\mathbf{A}_R = \{ r \in K : (b, r) \neq \emptyset \text{ for all } b < r \}.$$

Definition 2.3.2 *Suppose that*

- $\epsilon > 0$,
- $L \subset \mathbf{A}_L$ and $\text{card } L$ is finite,
- $R \subset \mathbf{A}_R$ and $\text{card } R$ is finite,
- $\{\phi_l\}_{l \in L} \subset C(K)$, is a family of monotonic non-decreasing non-negative functions such that $\phi_l(y) = 0$ if $y \leq l$ and $\phi_l(y) > 0$ if $y > l$, and
- $\{\psi_r\}_{r \in R} \subset C(K)$ is a family of monotonic non-increasing non-negative functions such that $\psi_r(y) = 0$ if $y \geq r$ and $\psi_r(y) > 0$ if $y < r$.

Then we denote by \mathcal{B}_w the set of subsets $U_w(0, \epsilon, \{\phi_l\}_{l \in L}, \{\psi_r\}_{r \in R}) \subset C(K)$ where $g \in U_w(0, \epsilon, \{\phi_l\}_{l \in L}, \{\psi_r\}_{r \in R})$ if and only if

- $|g(x)| < \epsilon$, for all $x \in K$,
- for all $l \in L$ and all $a > l$ we have $|g(a) - g(l)| < \phi_l(a)$, and
- for all $r \in R$ and all $b < r$ we have $|g(b) - g(r)| < \psi_r(b)$.

For convenience we will refer to $\{\phi_l\}_{l \in L}$ as the left wedge functions of U_w and $\{\psi_r\}_{r \in R}$ as the right wedge functions of U_w .

We may use the sets \mathcal{B}_o and \mathcal{B}_w to define the oscillation topology and the wedge topology respectively. A subset $G \subset C(K)$ is defined to be τ_o -open if for all $g \in G$ there is a $U_g \in \mathcal{B}_o$ such that

$$g + U_g \subset G.$$

To simplify this notation, we shall write $U_o(g, \epsilon, \omega)$ for $g + U_o(0, \epsilon, \omega)$. We define τ_w -open sets similarly. In Lemma 2.3.3 we show that τ_o and τ_w are equivalent.

Lemma 2.3.3 *Every $U \in \mathcal{B}_o$ contains a $V \in \mathcal{B}_w$ and vice versa.*

PROOF. We first show that every non-empty $V \in \mathcal{B}_w$ contains a non-empty $U \in \mathcal{B}_o$. Suppose that $V = U_w(0, \epsilon, \{\phi_l\}_{l \in L}, \{\psi_r\}_{r \in R})$. We let

$$\omega_L(s) = \min_{l \leq s} \phi_l(s)$$

and

$$\omega_R(s) = \min_{r \geq s} \psi_r(s).$$

For each $x \in L \cup R$ we use Lemma 2.2.5 with $\hat{\eta} = 0$ to find $\eta_x \in C(K)$ such that $\eta_x(x) = 0$ and $\eta_x(y) > 0$ for $y \neq x$. We let $\eta(s) = \min_{x \in L \cup R} \eta_x(s)$ and $\omega = 1/2 \min\{\omega_L, \omega_R, \eta\}$. Then ω is continuous, since the only discontinuities of $\min\{\omega_L, \omega_R\}$ occur at points in $x \in L \cup R$ at which $\lim_{y \rightarrow x} \eta(y) = 0$; since $0 \leq \omega \leq \eta$ then $\lim_{y \rightarrow x} \omega(y) = \omega(x)$. Moreover, $\{x \in K : \omega(x) = 0\} = L \cup R$ is finite.

With this definition of ω we find that if $g \in U_o(0, \epsilon, \omega)$ then for $l \in L$ and $a > l$ we have

$$\begin{aligned} |g(a) - g(l)| &\leq \text{osc}(g, [l, a]) \\ &< \sup_{z \in [l, a]} \omega(z) \\ &< \phi_l(a). \end{aligned}$$

Similarly for $r \in R$ and $b < r$ we have $|g(b) - g(r)| < \psi_r(b)$ and hence $g \in V$.

Conversely we must show that every non-empty $U \in \mathcal{B}_o$ contains a non-empty $V \in \mathcal{B}_w$. Given $U_o(0, \epsilon, \omega)$ we require finite families of wedge functions $\{\phi_l\}_{l \in L}$, and $\{\psi_r\}_{r \in R}$ such that, for any $g \in C(K)$ satisfying,

- $|g(a) - g(l)| < \phi_l(a)$ for all $l \in L$ and all $a > l$ and
- $|g(b) - g(r)| < \psi_r(b)$ for all $r \in R$ and all $b < r$,

then g must satisfy $\text{osc}(g, [a, b]) < \sup_{z \in [a, b]} \omega(z)$ for all $[a, b] \subset K$ such that a is an accumulation point of $[a, b]$, and $\omega(a) = 0$.

To construct our wedge functions we use monotonic functions of the form $\phi_l(x) = \sup_{l \leq z \leq x} \omega(z)$. Suppose that $F = \{a_1, \dots, a_k\}$ is the set of accumulation points of K at which the function ω is zero, with $L \subset F$ the set of left accumulation points and $R \subset F$ the set of right accumulation points. The construction is as follows.

- (i) In each interval $[a_i, a_{i+1})$ we find the point t_i such that

$$t_i = \max\{t : \omega(t) = \max_{z \in [a_i, a_{i+1}]} \omega(z)\}.$$

Then either $t_i \in (a_i, a_{i+1})$ with $\omega(t_i) > 0$ or $t_i = a_i$ and $\omega(t_i) = 0$.

- (ii) Find $t_0 \leq a_1$ such that $\omega(t_0) = \max_{z \leq a_1} \omega(z)$ and we find $t_k \geq a_k$ such that $\omega(t_k) = \max_{z \geq a_k} \omega(z)$.

(iii) For $a_i \in L$ we define

$$\phi_{a_i}(x) = \begin{cases} \sup_{a_i \leq z \leq x} \omega(z)/4 & \text{if } a_i \leq x \leq t_i \\ \omega(t_i)/4 & \text{if } x \geq t_i \\ 0 & \text{if } x \leq a_i. \end{cases}$$

(iv) For $a_i \in R$ we define

$$\psi_{a_i}(x) = \begin{cases} \sup_{a_i \geq z \geq x} \omega(z)/4 & \text{if } a_i \geq x \geq t_{i-1} \\ \omega(t_{i-1})/4 & \text{if } x \leq t_{i-1} \\ 0 & \text{if } x \geq a_i. \end{cases}$$

We must verify that each ϕ_l and ψ_r are indeed wedge functions. Since ω is continuous, they are continuous. Suppose that a_i is a right accumulation point of K and that $a_i \geq x_1 \geq x_2 \geq t_{i-1}$. Then $[a_i, x_1] \subset [a_i, x_2]$ so that $\sup\{\omega(z) \mid a_i \geq z \geq x_1\} \leq \sup\{\omega(z) \mid a_i \geq z \geq x_2\}$. Therefore ψ_{a_i} is monotonic non-increasing on $[t_{i-1}, a_i]$. Also $\psi_{a_i}(x)$ is constant on the remaining two intervals, with maximum value $\omega(t_{i-1})/4$ when $x \leq t_{i-1}$ and zero when $x \geq a_i$, so that ψ_{a_i} is a right wedge function. Applying a similar argument we may show that if a_i is a left accumulation point then $\phi_{a_i}(x)$ is a left wedge function. Therefore $U_w(0, \epsilon_1, \{\phi_l\}_{l \in L}, \{\psi_r\}_{r \in R})$ is in \mathcal{B}_w .

We claim that $V = U_w(0, \epsilon, \{\phi_l\}_{l \in L}, \{\psi_r\}_{r \in R}) \subset U_o(0, \epsilon, \omega)$. Suppose that $g \in V$, then clearly $|g(x)| < \epsilon$ for all $x \in K$. Let $[a, b] \subset K$, $\omega(a) = 0$, and let a be an accumulation point of $[a, b]$. If $a < b$, then $a = a_i \in L$. We choose $x, y \in [a, b]$ realising the oscillation of g on $[a, b]$ and estimate that

$$\begin{aligned} \text{osc}(g, [a, b]) &= |g(x) - g(y)| \\ &\leq |g(a) - g(x)| + |g(a) - g(y)| \\ &\leq \phi_a(x) + \phi_a(y) \\ &\leq 2\phi_a(b) \\ &\leq \sup_{a \leq z \leq b} \omega(z)/2 \\ &< \sup_{a \leq z \leq b} \omega(z). \end{aligned}$$

A similar estimate applies when $a > b$. This concludes the proof. \square

1.2.3.4

Let E be an abelian group and \mathcal{B} a collection subsets of E , each element of \mathcal{B} containing 0. Suppose further that

- (i) the intersection of any two elements of \mathcal{B} contains an element of \mathcal{B} ,
- (ii) $U = -U$ for $U \in \mathcal{B}$,
- (iii) given $U \in \mathcal{B}$, there is a $V \in \mathcal{B}$ such that $V + V \subset U$, and that
- (iv) the intersection of all elements of \mathcal{B} is $\{0\}$.

We say that τ is the topology defined by \mathcal{B} if G is in τ if and only if for all $g \in G$ that there is a $U_g \in \mathcal{B}$ such that

$$g + U_g \subset G.$$

This definition of τ together with (i)- (iv) makes (E, τ) a Hausdorff topological space and that addition and inverse are continuous (see[Di] page 35); hence (E, τ) is an abelian topological group.

We make use of the following standard result for abelian topological groups.

Lemma 2.3.4 *Let τ be the topology defined by a collection of sets \mathcal{B} with properties (i)–(iv) above, then*

- 1. *every element of \mathcal{B} is a neighbourhood of 0,*
- 2. *\mathcal{B} is a basis of neighbourhoods of 0,*
- 3. *(E, τ) is a Hausdorff topological space, and*
- 4. *(E, τ) is a topological group.*

Lemma 2.3.5 *The topology τ_o has the following properties:*

- (i) *the intersection of two elements of \mathcal{B}_o contains an element of \mathcal{B}_o ,*
- (ii) *every element of \mathcal{B}_o is convex and symmetric,*
- (iii) *given $U \in \mathcal{B}$, there is a $V \in \mathcal{B}$ such that $V + V \subset U$,*
- (iv) *every element of \mathcal{B}_o is a τ_o neighbourhood of 0,*
- (v) *\mathcal{B}_o is a basis of neighbourhoods of 0 in τ_o ,*
- (vi) *$(C(K), \tau_o)$ is a Hausdorff topological space,*

- (vii) $(C(K), \tau_o)$ is a topological group,
- (viii) if $f \in C(K)$ then the mapping defined by $\lambda \mapsto \lambda f$ from \mathbf{R} to $C(K)$ is continuous if and only if f attains only finitely many values, in particular $(C(K), \tau_o)$ is not a topological vector space unless K is finite,
- (ix) the topology τ_o is finer than the norm topology,
- (x) if $U \in \mathcal{B}_o$, $u \in C(K)$, and $\epsilon > 0$, then $U \cap \bigcup_{0 < t < \epsilon} B(tu, t\epsilon) \neq \emptyset$,
and
- (xi) $(C(K), \tau_o)$ is a Baire space.

PROOF. To show (i), that the intersection of any two elements of \mathcal{B}_o contains an element of \mathcal{B}_o , suppose that $U_o(0, \epsilon_1, \omega_1)$ and $U_o(0, \epsilon_2, \omega_2)$ are given. It suffices to let $\epsilon_3 = \min\{\epsilon_1, \epsilon_2\}$, and $\omega_3 = \min\{\omega_1, \omega_2\}$, then ω_3 is non-negative with finitely many zeros and we have

$$U_o(0, \epsilon_3, \omega_3) \subset U_o(0, \epsilon_1, \omega_1) \cap U_o(0, \epsilon_2, \omega_2).$$

For (ii), that every element of \mathcal{B}_o is convex and symmetric, let $g_1, g_2 \in U_o(0, \epsilon, \omega) \in \mathcal{B}_o$. For each α such that $0 \leq \alpha \leq 1$ we must show that $\alpha g_1 + (1 - \alpha)g_2 \in U_o(0, \epsilon, \omega)$. We have that

$$\|\alpha g_1 + (1 - \alpha)g_2\| \leq \alpha \|g_1\| + (1 - \alpha)\|g_2\| < \epsilon$$

and for any closed interval $[a, b] \subset K$ such that a is an accumulation point of $[a, b]$, and $\omega(a) = 0$, we have (applying Lemma 2.2.2) that

$$\begin{aligned} \text{osc}(\alpha g_1 + (1 - \alpha)g_2, [a, b]) &\leq \alpha \text{osc}(g_1, [a, b]) + (1 - \alpha) \text{osc}(g_2, [a, b]) \\ &< \sup_{x \in [a, b]} \omega(x). \end{aligned}$$

Therefore $U_o(0, \epsilon, \omega)$ is convex. It is symmetric since if $h \in U_o(0, \epsilon, \omega)$ then $-h \in U_o(0, \epsilon, \omega)$.

For (iii), given $U = U_o(0, \epsilon, \omega)$ we let $V = U_o(0, \epsilon/2, \omega/2)$. Then $V + V \subset U$.

For (iv)–(vii) we apply Lemma 2.3.4 using properties (i)–(iii) of Lemma 2.3.5 and noting that the intersection of all $U_o(0, \epsilon, \omega)$ is clearly just $\{0\}$.

For (viii), suppose first that the function f attains only finitely many values $(c_i)_{i=1}^n$. It is sufficient to show continuity of $\lambda \mapsto \lambda f$ at $\lambda = 0$. Given any non-empty neighbourhood of the origin $U_o(0, \epsilon, \omega)$, we must find $\delta > 0$ such that if $|\mu| < \delta$ then we have

- $|\mu| \|f\| < \epsilon$ and
- for any interval $[a, b]$ such that a is an accumulation point of $[a, b]$, and $\omega(a) = 0$ we have $|\mu| \text{osc}(f, [a, b]) < \sup_{z \in [a, b]} \omega(z)$.

Let F be the set of zeros of ω and let $L = F \cap \mathbf{A}_{\mathbf{L}}$ and $R = F \cap \mathbf{A}_{\mathbf{R}}$. Since F is finite and f is continuous, for each $a \in L$ there is $a^+ > a$ such that f is constant on $[a, a^+]$ and $\omega(a^+) > 0$. Similarly for each $a \in R$ there is a^- such that f is constant on $[a^-, a]$ and $\omega(a^-) > 0$. Let $d = \min\{\omega(a^+), \omega(a^-)\}$, $e = \max_{1 \leq i \neq j \leq n} |c_i - c_j|$. Let $\delta = \min\{\frac{\epsilon}{1+\|f\|}, \frac{d}{1+e}\}$ and $|\mu| < \delta$ so that $|\mu| \|f\| < \epsilon$. We must show that for any interval $[a, b]$ such that $a \in F$, and a is an accumulation point of $[a, b]$ we have $|\mu| \text{osc}(f, [a, b]) < \sup_{z \in [a, b]} \omega(z)$. We assume that $b > a$; the remaining case $b < a$ is similar. If $b \leq a^+$, then $|\mu| \text{osc}(f, [a, b]) = 0 < \sup_{z \in [a, b]} \omega(z)$. If $b > a^+$, then

$$\begin{aligned} |\mu| \text{osc}(f, [a, b]) &< \delta e \\ &< d \\ &\leq \omega(a^+) \\ &\leq \sup_{z \in [a, b]} \omega(z). \end{aligned}$$

This proves the sufficiency of f having finitely many values for continuity of scalar multiplication of f .

We must show the necessity of the condition, that if $\lambda \mapsto \lambda f$ is continuous then the function f has only finitely many values. Suppose to the contrary that f does not have finitely many values. Then there are $(f(t_i))_{i=1}^{\infty}$ such that $f(t_i) \neq f(t_j)$ if $i \neq j$. Passing to a subsequence we may suppose that the sequence t_i is strictly monotonic and $\lim_{i \rightarrow \infty} t_i = t$. Assume that t_i is strictly increasing. Because of Lemma 2.3.3 it suffices to find $U_w(0, \epsilon, \{\phi_l\}_{l \in L}, \{\psi_r\}_{r \in R})$ such that for any $\delta > 0$ we have $\delta f \notin U_w(0, \epsilon, \{\phi_l\}_{l \in L}, \{\psi_r\}_{r \in R})$. With this in view we let $L = \emptyset$, $R = \{t\}$ and we construct a function $\psi_t : K \rightarrow \mathbf{R}$ with a unique zero at t as follows. Let $\psi_t(t_i) = 2^{-i} \text{osc}(f, [t, t_i])$ for all $i \geq 1$. For $z \leq t_1$ let $\psi_t(z) = \omega(t_1)$. We may choose, by Lemma 2.2.4, continuous functions $g_i : [t_i, t_{i+1}] \rightarrow \mathbf{R}$ such that

$$g_i(z) = \begin{cases} \omega(t_{i+1}) & \text{if } z = t_{i+1} \\ \omega(t_i) & \text{if } z = t_i. \end{cases}$$

Then we define $\psi_t(z) = g_i(z)$ if $z \in [t_i, t_{i+1}]$. For $z \geq t$ we let $\psi_t(z) = 0$. For any $\delta > 0$ there is an interval $[t, d]$ such that $\delta \text{osc}(f, [t, d]) \geq \sup_{z \in [t, d]} \psi_t(z)$. We need only choose $d = t_i$ for some i such that $2^{-i} < \delta$. So δf does not

belong to our neighbourhood. Therefore $\lambda \mapsto \lambda f$ is not continuous at $\lambda = 0$. The case when t_i is decreasing is similar.

For (ix) , that the topology τ_o is finer than the norm topology, we note that for any ω , non-negative with finitely many zeros, we have $U_o(0, \delta, \omega) \subset B(0, \delta)$ for any open ball $B(0, \delta)$, centre 0 and radius $\delta > 0$.

For (x) , that if $U \in \mathcal{B}_o$, $u \in C(K)$, and $\epsilon > 0$, then $U \cap \bigcup_{0 < t < \epsilon} B(tu, t\epsilon) \neq \emptyset$. We suppose that $U = U_o(0, \kappa, \omega)$ and F_0 is the set of zeros of ω . We must find $t \in \mathbf{R}$, $u' \in C(K)$ such that $0 < t < \epsilon$ and

1. $\|u' - tu\| < t\epsilon$,
2. $\|u'\| < \kappa$, and
3. if $a \in F_0$, and a is an accumulation point of $[a, b]$, then we have

$$\text{osc}(u', [a, b]) < \sup_{z \in [a, b]} \omega(z).$$

Let $F_0 = \{x \in K : \omega(x) = 0\} = \{x_1, \dots, x_k\}$ where $x_1 < x_2 < \dots < x_k$. We construct u' as follows. Choose disjoint intervals $([c_i, d_i])_{i=1}^k$ and a real number $c > 0$ as follows. Let $x_0 = \min K$ and $x_{k+1} = \max K$. For $i = 1, 2, \dots, k$ if $x_i \in \mathbf{A}_{\mathbf{R}}$ we define c_i, d_i as follows. Choose $c_i \in (x_{i-1}, x_i)$ and put $A_i = [x_{i-1}, c_i]$. If $x_i \notin \mathbf{A}_{\mathbf{R}}$ let $c_i = x_i$, and $A_i = [x_{i-1}, c_i]$. Also let $c_{k+1} = x_{k+1}$ so that if $x_i \in \mathbf{A}_{\mathbf{L}}$ choose $d_i \in (x_i, c_{i+1})$ and let $B_i = [d_i, c_{i+1}]$. If $x_i \notin \mathbf{A}_{\mathbf{L}}$ let $d_i = x_i$, and $B_i = (d_i, c_{i+1}]$.

Denote $A = A_1 \cup (B_1 \cap A_2) \cup ((B_2 \cap A_3) \cup \dots \cup (B_{k-1} \cap A_k) \cup B_k$. Then A is a compact set and $A \cap F_0 = \emptyset$. Hence $c = \inf_{x \in A} \omega(x)/2 > 0$.

Choose t such that $0 < t < \min\{\epsilon, \frac{\kappa}{2\epsilon}, \frac{\kappa}{2\|u\|+2}, \frac{c}{2\|u\|+2}, \frac{c}{2\epsilon}, 1\}$. Choosing a function $b \in C(K)$ by Tietze's theorem such that

- $\|b\| < \epsilon$ and
- $b(x) = -u(x) + u(x_i)$ for $x \in [c_i, d_i]$ for each i such that $1 \leq i \leq k$,

we define $u' \in C(K)$ as $u' = tu + tb$. Having defined u' and t we verify relations (1), (2) and (3). For relation (1) we note that $\|u' - tu\| = t\|b\| < t\epsilon$.

For relation (2) using the estimates $\|u'\| \leq \|tu\| + \|tb\|$, $\|b\| < \epsilon$, and $0 < t < \min\{\frac{\kappa}{2\|u\|+2}, \frac{\kappa}{2\epsilon}\}$, we obtain $\|u'\| < \kappa$.

To verify relation (3) it suffices to show that for any x_i , if x_i is an accumulation point of $[x_i, b]$, then $\text{osc}(u', [x_i, b]) < \sup_{z \in [x_i, b]} \omega(z)$. There are two cases for b .

Case (i) $[x_i, b] \subset [c_i, d_i]$. Noting that u' takes the constant value $tu(x_i)$ on $[c_i, d_i]$, we obtain $\text{osc}(u', [x_i, b]) = 0$. Since $\sup_{z \in [x_i, b]} \omega(z) > 0$, then

$$\text{osc}(u', [x_i, b]) < \sup_{z \in [x_i, b]} \omega(z).$$

Case (ii) $[x_i, b] \setminus [c_i, d_i] \neq \emptyset$. Using that $\text{osc}(u', K) \leq 2t\|u\| + 2t\|b\| < 2c$, that $d_i \in [x_i, b]$, and $\omega(d_i) \geq 2c$, we obtain

$$\begin{aligned} \text{osc}(u', [x_j, b] \setminus \bigcup_{i=1}^k [c_i, d_i]) &\leq \text{osc}(u', K) \\ &< 2c \\ &\leq \omega(d_i) \\ &\leq \sup_{x \in [x_i, b]} \omega(x). \end{aligned}$$

This ends the proof of property (x).

For (xi) , that $(C(K), \tau_o)$ is a Baire space, we play the Banach-Mazur game. We recall that in any play, player A chooses a sequence of τ_o open sets $(U_n)_{n=1}^\infty$ and player B chooses a sequence of τ_o open sets $(V_n)_{n=1}^\infty$ so that $U_1 \supset V_1 \supset U_2 \supset V_2 \supset \dots$. We intend to show that there is a strategy for player B such that $\bigcap_{i=1}^\infty V_i$ is non-empty. Suppose that player A chooses a non-empty open subset $U_1 \subset C(K)$. Then player B may choose any basic neighbourhood $U(g_1, \kappa_1, \lambda_1) \subset U_1$ and further chooses $V_1 = U(g_1, \epsilon_1, \omega_1) \subset U(g_1, \kappa_1, \lambda_1)$ where

1. $\epsilon_1 = \frac{\min\{\kappa_1, 1\}}{2}$ and
2. $\omega_1 = \lambda_1/2$.

Suppose that after n turns player A has chosen subsets $(U_i)_{i=1}^n$ and player B has chosen $(V_i)_{i=1}^{n-1}$ so that $U_1 \supset V_1 \supset U_2 \supset V_2 \supset \dots \supset U_{n-1} \supset V_{n-1} \supset U_n$. Player B , choosing any basic neighbourhood $U(g_{n+1}, \kappa_{n+1}, \lambda_{n+1}) \subset U_n$, chooses further $V_{n+1} = U(g_{n+1}, \epsilon_{n+1}, \omega_{n+1}) \subset U(g_{n+1}, \kappa_{n+1}, \lambda_{n+1})$ where

- $\epsilon_{n+1} = \frac{\kappa_{n+1}}{2}$ and
- $\omega_{n+1} = \lambda_{n+1}/2$.

We claim that

1. $(g_i)_{i=1}^\infty$ is a norm Cauchy sequence with limit $h \in C(K)$ and
2. $h \in \bigcap_{i=1}^\infty V_i$.

Verifying claim (1), we have the estimate $\|g_{i+1} - g_i\| < \frac{\epsilon_i}{2} < \frac{\kappa_{i-1}}{2} < \frac{1}{2^i}$. Thus $(g_i)_{i=1}^\infty$ is a Cauchy sequence with limit $h \in C(K)$.

For claim (2) let F_i be the set of zeros of ω_i . We show that for each i we have $\|h - g_{i+1}\| \leq \epsilon_{i+1} < \kappa_{i+1}$, and for any $a \in F_{i+1}$ such that a is an accumulation point of $[a, b]$ that

$$\text{osc}(h - g_{i+1}, [a, b]) \leq \sup_{z \in [a, b]} \omega_{i+1}(z) < \sup_{z \in [a, b]} \lambda_{i+1}(z).$$

It follows that $h \in U_{i+1} \subset V_i$ for each i and hence that $h \in \bigcap_{i=1}^\infty V_i$. For the former inequality, using the triangle inequality, we obtain for all positive integers j that $\|h - g_{i+1}\| \leq \|h - g_j\| + \|g_j - g_{i+1}\|$. Choosing M sufficiently large so that for all $j \geq M$ we have $\|h - g_j\| < \frac{\kappa_{i+1}}{2}$ and $g_j \in V_{i+1}$, we obtain

$$\|h - g_{i+1}\| < \frac{\kappa_{i+1}}{2} + \epsilon_{i+1} = \kappa_{i+1}.$$

Similarly for the latter inequality, using relation (i) of Lemma 2.2.2 we obtain for each positive integer j that $\text{osc}(h - g_{i+1}, [a, b]) \leq \text{osc}(h - g_j, [a, b]) + \text{osc}(g_j - g_{i+1}, [a, b])$. By relation (ii) of Lemma 2.2.3, the uniform convergence of g_j to h , and that $h \in V_{i+1}$, we may choose M sufficiently large that for all $j \geq M$ we have $\text{osc}(h - g_j, [a, b]) < \sup_{z \in [a, b]} \omega_{i+1}(z)$. Then since $g_j \in V_{i+1}$,

$$\begin{aligned} \text{osc}(h - g_{i+1}, [a, b]) &< \sup_{z \in [a, b]} \omega_{i+1}(z) + \sup_{z \in [x, y]} \omega_{i+1}(z) \\ &= \sup_{z \in [a, b]} \lambda_{i+1}(z). \end{aligned}$$

□

In section 2.5 we use the following Lemma to construct τ_0 -open sets.

Lemma 2.3.6 *Let G be an open subset of K , let $t \in G$, let $a \in \mathbf{R}$, and let $\lambda \in C(K)$ be such that $\lambda(t) = 0$. Then the sets*

$$V_1 = \{ \phi \in C(K) : \phi(s) < \phi(t) + \lambda(s) \text{ if } s \in \overline{G} \setminus \{t\} \}$$

and

$$V_2 = \{ \phi \in C(K) : |\phi(s)| < a\phi(t) \text{ if } s \notin G \}$$

of $C(K)$ are τ_0 -open.

PROOF. We prove the simplest case first, V_2 , which is in fact open in the norm topology. If $G = K$ then $V_2 = C(K)$, and so, supposing that $K \setminus G \neq \emptyset$, we let $\phi_1 \in V_2$ and, by compactness of $K \setminus G$, we find a point $s_1 \notin G$ such that $\sup_{s \notin G} |\phi_1(s)| = |\phi_1(s_1)|$. If we choose $0 < \epsilon < \frac{a\phi_1(t) - |\phi_1(s_1)|}{|a|+1}$ then the open ball $B(\phi_1, \epsilon)$ is a subset of V_2 since, if $s \notin G$ and $\varphi \in B(\phi_1, \epsilon)$ then we estimate that

$$\begin{aligned} |\varphi(s)| &= |\phi_1(s)| + \|\varphi - \phi_1\| \\ &< |\phi_1(s_1)| + \epsilon \\ &< a\phi_1(t) - \epsilon|a| \\ &< a\phi(t). \end{aligned}$$

Hence $\varphi \in V_2$.

For V_1 , supposing that $\phi_0 \in V_1$, we require $U(0, \epsilon_0, \omega_0)$ such that $\phi_0 + U(0, \epsilon_0, \omega_0) \subset V_1$. If t is a left and right accumulation point of K then we choose an interval $(a, b) \ni t$ such that $(a, b) \subset \overline{G}$. To find ω_0 we first define $\theta : (a, b) \rightarrow \mathbf{R}$ by

$$\theta(s) = \begin{cases} \inf_{z \in [s, b]} \lambda(z) + \phi_0(t) - \phi_0(z) & \text{if } s \geq t \\ \inf_{z \in (a, s]} \lambda(z) + \phi_0(t) - \phi_0(z) & \text{if } s \leq t. \end{cases}$$

We now define $\omega_0 : K \rightarrow \mathbf{R}$ by

$$\omega_0(s) = \begin{cases} \theta(s) & \text{if } s \in (a, b) \\ \theta(b) & \text{if } s \geq b \\ \theta(a) & \text{if } s \leq a \end{cases}$$

and let

$$\epsilon_0 = 1/2 \inf_{s \in \overline{G} \setminus (a, b)} (\lambda(s) + \phi_0(t) - \phi_0(s)) > 0.$$

Then ω_0 is non-negative with exactly one zero at t . Supposing that $\phi \in \phi_0 + U(0, \epsilon_0, \omega_0)$ we have

- $\|\phi - \phi_0\| < \epsilon_0$ and
- for all $s \neq t$ that $\text{osc}(\phi - \phi_0, [t, s]) < \sup_{z \in [t, s]} \omega_0(z)$.

We show that $\phi \in V_1$. There are two cases for s , either $s \in (a, b)$ or $s \in \overline{G} \setminus (a, b)$. If $s \in \overline{G} \setminus (a, b)$ we have

$$\begin{aligned} \phi(s) - \phi(t) - (\phi_0(s) - \phi_0(t)) &\leq \text{osc}(\phi - \phi_0, [t, s]) \\ &\leq 2\|\phi - \phi_0\| \\ &< 2\epsilon_0 \\ &\leq \lambda(s) + \phi_0(t) - \phi_0(s). \end{aligned}$$

Hence $\phi(s) - \phi(t) < \lambda(s)$.

For the remaining case, that is when $s \in (a, b)$, using the definition of $\theta(s)$ we estimate that

$$\begin{aligned} \phi(s) - \phi(t) - (\phi_0(s) - \phi_0(t)) &\leq \text{osc}(\phi - \phi_0, [t, s]) \\ &< \sup_{z \in [t, s]} \omega_0(z) \\ &\leq \lambda(s) + \phi_0(t) - \phi_0(s). \end{aligned}$$

Hence $\phi(s) - \phi(t) < \lambda(s)$.

In both cases we obtain $\phi(s) < \phi(t) + \lambda(s)$ implying that $\phi \in V_1$ when t is both a right and a left accumulation point.

The remaining cases are:

- t is a right accumulation point but not a left accumulation point;
- t is a left accumulation point but not a right accumulation point;
- t is an isolated point.

If t is a right accumulation point but not a left accumulation point then we can use the preceding argument, replacing the interval (a, b) with an interval of the form $(a, t]$. The case when t is a left accumulation point but not a right accumulation point is dealt with similarly. If t is an isolated point then we have $\inf_{s \in \bar{G}, s \neq t} (\lambda(s) + \phi_0(t) - \phi_0(s)) = \kappa > 0$. Defining $\omega_0 : K \rightarrow \mathbf{R}$ by

$$\omega_0(s) = \begin{cases} 0 & \text{if } s = t \\ \kappa & \text{if } s \neq t \end{cases}$$

and $\epsilon_0 = \kappa/2$ gives $\phi_0 + U(0, \epsilon_0, \omega_0) \subset V_1$. □

2.4 Directional derivatives

Recalling statement (x) of Lemma 2.3.3 we have that any non-empty τ_o open set $U \in \mathcal{B}_o$ intersects any cone of arbitrarily small diameter in at least one point. In Lemma 2.4.1 we deduce from this property that a continuous convex function f on $C(K)$ has directional derivatives in any fixed direction on a τ_o -residual set.

Lemma 2.4.1 *Suppose that a translational invariant topology τ of a Banach space E is finer than the norm topology and is such that*

$$U \cap \bigcup_{0 < t < \epsilon} B(tu, t\epsilon) \neq \emptyset$$

whenever $u \in E$, $\epsilon > 0$ and U is τ open and contains 0. Then for every continuous convex function f on E and every $e \in E$ the set of points $x \in E$ at which the directional derivative $f'(x, e)$ exists is τ -residual.

PROOF. We first define the notion of Δ -differentiability of a continuous convex function f on E . For all $\Delta > 0$ and a fixed element $e \in E$ we say that a point $x \in E$ is a Δ -differentiability point of a continuous convex function $f : E \rightarrow \mathbf{R}$ if

$$\lim_{t \rightarrow 0} \frac{f(x + te) + f(x - te) - 2f(x)}{t} < \Delta.$$

We observe that, since the function f is convex, the function

$$t \mapsto \frac{f(x + te) + f(x - te) - 2f(x)}{t}$$

for $t \in \mathbf{R}^+$ is non-negative and non-decreasing. Therefore a point $x \in E$ is a Δ -differentiability point of f if and only if there is a $t > 0$ such that

$$\frac{f(x + te) + f(x - te) - 2f(x)}{t} < \Delta. \quad (2.1)$$

We show that, given any $\Delta > 0$ and any non-empty τ -open set U there is a non-empty τ -open $V \subset U$ such that every point of V is a Δ -differentiability point of f . Then, by letting $\Delta = 1/n$ for any positive integer n , the set of $1/n$ -differentiability points of f is τ -open and τ -dense. Hence, the set of points $x \in E$ at which the directional derivative $f'(x, e)$ exists is τ -residual in E .

In what follows we assume that $\|e\| = 1$ and that Δ is any positive constant. Supposing that $x \in U$ and using the convexity of f , we have the existence of the limit

$$L = \lim_{t \rightarrow 0^+} \frac{f(x + te) - f(x)}{t}.$$

Then choosing $\delta_1 > 0$ so that for all $t \in (0, 3\delta_1)$, we have

$$0 \leq \frac{f(x + te) - f(x)}{t} - L \leq \frac{\Delta}{8}$$

and for all $t \in (0, \delta_1)$ we estimate that

$$\begin{aligned}
 & \frac{f(x+3te) + f(x+te) - 2f(x+2te)}{t} = \\
 & \frac{f(x+3te) - f(x) - 3tL}{t} + \frac{f(x+te) - f(x) - tL}{t} + \\
 & \frac{4tL + 2f(x) - 2f(x+2te)}{t} \\
 & \leq \frac{3\Delta}{8} + \frac{\Delta}{8} \\
 & = \frac{\Delta}{2}
 \end{aligned}$$

where we use $\frac{(4tL+2f(x)-2f(x+2te))}{t} \leq 0$. Since f is locally Lipschitz there is a constant $K > 0$ and a $\delta_2 > 0$ such that if $z, y \in B(x, \delta_2)$ then $|f(z) - f(y)| \leq K\|z - y\|$. We let $\epsilon = \min\{2\delta_1, \frac{2\delta_2}{3}, \frac{1}{4}, \frac{\Delta}{16K}\}$ and show that the set

$$W = \bigcup_{0 < s < \epsilon} B(x + se, s\epsilon)$$

contains only Δ -differentiability points of f . Let $y \in W$, then $y \in B(x + se, s\epsilon)$ for some $0 < s < \epsilon$. We denote by $s = 2t$ and estimate that

$$\begin{aligned}
 & \frac{f(y+te) + f(y-te) - 2f(y)}{t} \\
 & \leq \frac{f(x+3te) + f(x+te) - 2f(x+2te)}{t} \\
 & \quad + \frac{4K\|x+2te - y\|}{t} \\
 & < \Delta/2 + 8K\epsilon \leq \Delta
 \end{aligned}$$

which, by (2.1), implies that y is a Δ -differentiability point of f . So W contains only Δ -differentiability points. Letting $V = W \cap U$ and using the hypothesis of the Lemma we conclude that V is non-empty. Since W is norm open and τ is finer than the norm topology then the set V is τ open. The set V is a non-empty τ open subset of U that contains only Δ -differentiability points. \square

2.5 Dirac measures

Let \mathcal{F}_D , denote the set of continuous convex functions on $C(K)$ that contain a Dirac measure in each subdifferential. We obtain in Proposition 2.5.3 the following statement.

Suppose that $f \in \mathcal{F}_D$, then given any non-empty τ_o -open set U there is a non-empty τ_o -open set $V \subset U$ and a finite set $F \subset K$ such that for every $\phi \in V$ the subdifferential $\partial f(\phi)$ contains a measure with at most one element support in F .

In section 2.6 we extend this result to the set of continuous convex functions on $C(K)$, \mathcal{F}_F , that contain a measure with finite support in each subdifferential. This result implies that for $\varphi \in V$ the values of $f(\varphi)$ depend only on the values of φ on F so that in section 2.7, using the Fréchet differentiability of convex functions on \mathbf{R}^n , we derive generic Gâteaux differentiability of continuous convex functions $C(K)$ in \mathcal{F}_F .

We recall some definitions. The order relation on $C(K)$ is defined by $f \geq 0$ if and only if for all $x \in K$ we have $f(x) \geq 0$. A Radon measure on K is an element of the dual $C(K)^*$. A Radon measure $\mu \in C(K)^*$ is positive, $\mu \geq 0$, if and only if whenever $f \geq 0$ then $\mu(f) \geq 0$. Let μ and ν be Radon measures, then we define the relation $\nu \leq \mu$ if $\mu - \nu$ is positive. The relation $\nu \leq \mu$ defines an order relation on $C(K)^*$. For any Radon measure $\mu \in C(K)^*$ there is a least positive Radon measure $|\mu| \in C(K)^*$ such that

$$|\mu(f)| \leq |\mu|(|f|)$$

for all $f \in C(K)$. The Radon measure $|\mu|$ is called the *variation* of μ . We will denote the set of Radon measures $C(K)^*$ by $\mathcal{M}(K)$ and the set of positive Radon measures by $\mathcal{M}^+(K)$. For each Radon measure μ there is a Radon measure μ^+ such that μ^+ is the least Radon measure ρ such that $\rho \geq \mu$ and $\rho \geq 0$. Similarly μ^- is the least Radon measure ρ such that $\rho \geq -\mu$ and $\rho \geq 0$. Then we have

$$\mu = \mu^+ - \mu^-,$$

and

$$|\mu| = \mu^+ + \mu^-.$$

Let 1_G denote the characteristic function of G . A set $E \subset K$ is called μ -negligible if for all $\epsilon > 0$ there is an open set $G \supset E$ such that $|\mu|(f) < \epsilon$ for all $f \in C(K)$ such that

$$0 \leq f \leq 1_G.$$

A Radon measure $\mu \in \mathcal{M}(K)$ is *concentrated* on a set A if $K \setminus A$ is μ -negligible.

If $x \in K$ the measure δ_x defined by

$$\delta_x(f) = f(x)$$

for all $f \in C(K)$ is called the *Dirac measure concentrated at $x \in K$* ; it is, of course, concentrated at $\{x\}$.

We denote by $\mathcal{M}_n(K)$, for $n \geq 0$, the subset of $\mathcal{M}(K)$ whose elements are concentrated on an at most n element subset of K . Let F be a subset of K , then we denote by $\mathcal{M}_n(F)$ the subspace of $\mathcal{M}(K)$ whose elements are concentrated on an at most n element subset of F .

Let f be a continuous convex real valued function on $C(K)$ and $\phi_0 \in C(K)$, then the *subdifferential* $\partial f(\phi_0)$ of f at ϕ_0 is defined to be the set

$$\partial f(\phi_0) = \{ \mu \in C(K)^* : f(\xi) \geq f(\phi_0) + \mu(\xi) - \mu(\phi_0) \text{ for all } \xi \in C(K) \}.$$

The subdifferential of f at ϕ_0 is non-empty, convex, and weak* compact for all $\phi_0 \in C(K)$. (See [Ph1].)

The proof of Proposition 2.5.3 relies on two Lemmata for which some notation is required. We first define subsets M_s of the square $[-1, 1]^2 \subset \mathbf{R}^2$, that are related to the subdifferentials of f . Let f be a continuous convex function on $C(K)$, then for each $s \in K$, define M_s to be the subset of the square $[-1, 1]^2 \subset \mathbf{R}^2$ given by

$$M_s = \{ (a, b) \in [-1, 1]^2 : f(\xi) \geq a + b\xi(s) \text{ for all } \xi \in C(K) \}.$$

The sets M_s have the property that if $b\delta_s \in \partial f(\phi)$ then $(f(\phi) - b\delta_s(\phi), b) \in M_s$. We also require a function $\hat{\eta}$ that behaves as a distance between M_s and M_t for some given point t . Suppose that M_t is non-empty. If M_s is non-empty let

$$\hat{\theta}(s) = \sup_{(a,b) \in M_s} \inf_{(a',b') \in M_t} |a - a'| + |b - b'|.$$

Define $\hat{\eta}(s) : K \rightarrow \mathbf{R}$ by

$$\hat{\eta}(s) = \begin{cases} \hat{\theta}(s) & \text{if } M_s \neq \emptyset \\ 0 & \text{if } M_s = \emptyset. \end{cases}$$

Lemma 2.5.1 *Let f be a continuous convex function on $C(K)$. Then*

$$\lim_{s \rightarrow t} \hat{\eta}(s) = 0.$$

PROOF. The proof is by contradiction.

Suppose it is not true that the limit $\lim_{s \rightarrow t} \hat{\eta}(s)$ is zero. Then there is an $\epsilon > 0$, points s_n of K such that $\lim_{n \rightarrow \infty} s_n = t$, and points $(a_n, b_n) \in M_{s_n}$, such that

$$\inf_{(a',b') \in M_t} |a_n - a'| + |b_n - b'| > \epsilon. \quad (2.2)$$

The square $[-1, 1]^2$ is compact, so every sequence in $[-1, 1]^2$ has a convergent subsequence and we may suppose that

$$\lim_{n \rightarrow \infty} (a_n, b_n) = (a, b).$$

We claim that $(a, b) \in M_t$. Choose any $\xi \in C(K)$ then since $(a_n, b_n) \in M_{s_n}$, we have

$$f(\xi) \geq a_n + b_n \xi(s_n)$$

and in the limit as $n \rightarrow \infty$

$$f(\xi) \geq a + b \xi(t).$$

Therefore $(a, b) \in M_t$. Clearly if n is chosen so that $|a_n - a| < \frac{\epsilon}{4}$ and $|b_n - b| < \frac{\epsilon}{4}$ then

$$\inf_{(a', b') \in M_t} |a_n - a'| + |b_n - b'| \leq |a_n - a| + |b_n - b| < \epsilon$$

which contradicts equation (2.2) and completes the proof. \square

Lemma 2.5.2 *Suppose that f is a continuous convex function on $C(K)$, that U is a τ_o -open set containing $U_o(\psi, \epsilon, \omega)$ for which F_0 is the finite set of zeros of ω , that $\phi_0 \in U_o(\psi, \frac{\epsilon}{4}, \frac{\omega}{4})$, and that there is a non-zero multiple of a Dirac measure in $\partial f(\phi_0)$ with support $t \notin F_0$. Then there is a non-empty τ_o -open subset $V \subset U$ such that, for every $\phi \in V$, every multiple of a Dirac measure in $\partial f(\phi)$ is non-zero and has support at t .*

PROOF. The first step in our proof is to define the set V . In order to define it we require some estimates on the range of values of f on U . We may assume that

- $f(U) \subset [-\frac{1}{2}, \frac{1}{2}]$ and that
- $\|\mu\| \leq \frac{1}{2}$ for all $\mu \in \partial f(\xi)$ and for all $\xi \in U$.

The first assumption follows from f being locally Lipschitz, implying that for a τ open set of sufficiently small norm diameter we can assume that the values of f on U lie within an interval of length 1. We need only translate f by a suitable constant to attain the range $[-\frac{1}{2}, \frac{1}{2}]$. For the second assumption, since $\|\mu\| \leq C$ where C is the Lipschitz constant of f restricted to U , then multiplying f by a sufficiently small constant ensures that $\|\mu\| \leq \frac{1}{2}$.

Supposing that $b_0 \delta_t \in \partial f(\phi_0)$ where $b_0 \neq 0$, we have from our assumptions on f that $|b_0| \leq 1$ and $(f(\phi_0) - \phi_0(t), b_0) \in M_t$, so that $M_t \neq \emptyset$. Using Lemma 2.5.1 we obtain that $\lim_{s \rightarrow t} \hat{\eta}(s) = 0$. Letting $\alpha = \frac{10(1 + \|\phi_0\| + \epsilon)}{|b_0|}$, we choose by Lemma 2.2.5 a function $\eta \in C(K)$ with the following properties.

- $\eta(s) > \hat{\eta}(s)$ for $s \neq t$,
- $\eta(t) = 0$, and
- $\eta(s) > \frac{2|\phi_0(s) - \phi_0(t)|}{\alpha}$ for all $s \neq t \in K$.

Since K is a Hausdorff space and F_0 is finite we may choose an open set $H \ni t$ and a positive number c such that

- $\overline{H} \cap F_0 = \emptyset$,
- $\inf_{z \in \overline{H}} \omega(z) > c$, and
- $\eta(s) < \min\{\frac{|b_0|}{2}, \frac{\min\{\frac{\epsilon}{4}, \frac{c}{4}\}}{\alpha}\}$ for all $s \in \overline{H}$.

Let $\kappa = \min\{\frac{\epsilon}{4}, \frac{c}{4}\}$.

We find by Lemma 2.2.4 a function $g \in C(K)$ with the properties that $g(s) = 0$ if $s \notin H$, that $g(s) = 1$ if and only if $s = t$, and that $0 \leq g \leq 1$. Let $G = \{s \in K : g(s)(\kappa - \alpha\eta(s)) > |b_0|\kappa/2\}$. Note that $t \in G$, G is open, and $\overline{G} \subset H$. Finally we define the set V as

$$V = \{\phi \in U : 0 < \text{sign}(b_0)(\phi(s) - \phi_0(s)) \text{ if } s \in \overline{G}, \\ \text{sign}(b_0)(\phi(s) - \phi_0(s)) < \text{sign}(b_0)(\phi(t) - \phi_0(t)) - \alpha\eta(s) \\ \text{if } s \in \overline{G} \text{ and } s \neq t, \\ \text{and } |\phi(s) - \phi_0(s)| < b_0(\phi(t) - \phi_0(t)) \text{ if } s \notin G\}.$$

Noting that $\{\phi \in C(K) : \text{sign}(b_0)(\phi(s) - \phi_0(s)) > 0 \text{ if } s \in \overline{G}\}$ is norm open and applying Lemma 2.3.6 we conclude that the set V is τ_o -open. We show that $\phi \in C(K)$ defined by $\phi(s) = \phi_0(s) + \text{sign}(b_0)g(s)(\kappa - \alpha\eta(s))$ is an element of V . To verify this claim we must show that

- (i) $\|\phi - \psi\| < \epsilon$,
- (ii) if $a \in F_0$, and a is an accumulation point of $[a, b]$, then $\text{osc}(\phi - \psi, [a, b]) < \sup_{z \in [a, b]} \omega(z)$,
- (iii) $0 < \text{sign}(b_0)(\phi(s) - \phi_0(s))$ if $s \in \overline{G}$,
- (iv) $\text{sign}(b_0)(\phi(s) - \phi_0(s)) < \text{sign}(b_0)(\phi(t) - \phi_0(t)) - \alpha\eta(s)$ if $s \in \overline{G}$ and $s \neq t$, and
- (v) $|\phi(s) - \phi_0(s)| < b_0(\phi(t) - \phi_0(t))$ if $s \notin G$.

For (i) we have

$$\begin{aligned} \|\phi - \psi\| &\leq \|\phi_0 - \psi\| + \sup_{s \in K} |g(s)(\kappa - \alpha\eta(s))| \\ &\leq \frac{\epsilon}{4} + \kappa < \epsilon. \end{aligned}$$

For (ii), in the case $[a, b] \cap G = \emptyset$, and noting that $\phi_0 \in U(\psi, \frac{\epsilon}{4}, \frac{\omega}{4})$, we obtain the estimate

$$\begin{aligned} \text{osc}(\phi - \psi, [a, b]) &= \text{osc}(\phi_0 - \psi, [a, b]) \\ &\leq \sup_{z \in [a, b]} \frac{\omega}{4} < \sup_{z \in [a, b]} \omega(z). \end{aligned}$$

For the case $[a, b] \cap G \neq \emptyset$ we obtain

$$\begin{aligned} \text{osc}(\phi - \psi, [a, b]) &\leq \text{osc}(\phi_0 - \psi, [a, b]) + \text{osc}(\text{sign}(b_0)g(\kappa - \alpha\eta), [a, b]) \\ &\leq \frac{\sup_{z \in [a, b]} \omega(z)}{4} + 2\kappa \\ &\leq \frac{\sup_{z \in [a, b]} \omega(z)}{4} + \frac{c}{2} \\ &< \frac{\sup_{z \in [a, b]} \omega(z)}{4} + \inf_{z \in \overline{H}} \frac{\omega(z)}{2} \\ &< \sup_{z \in [a, b]} \omega(z). \end{aligned}$$

For (iii), we have from the definition of G and ϕ that

$$\text{sign}(b_0)(\phi(s) - \phi_0(s)) \geq |b_0|\kappa/2 > 0$$

for all $s \in \overline{G}$.

For (iv), when $s \in \overline{G}$ and $s \neq t$ noting that $\phi(t) = \phi_0(t) + \text{sign}(b_0)\kappa$, we obtain

$$\begin{aligned} \text{sign}(b_0)(\phi(s) - \phi_0(s)) &= \text{sign}(b_0)g(s)(\kappa - \alpha\eta(s)) \\ &< \kappa - \alpha\eta(s) \\ &= \text{sign}(b_0)(\phi(t) - \phi_0(t)) - \alpha\eta(s). \end{aligned}$$

For (v), when $s \notin G$ we have from the definition of G that

$$|\phi(s) - \phi_0(s)| \leq |b_0|\kappa/2 < |b_0|\kappa = b_0(\phi(t) - \phi_0(t)).$$

To complete the proof we show that for all $s \neq t$, for all $(a, b) \in M_s$, and for all $\phi \in V$ we have $a + b\phi(s) < f(\phi)$. We deduce that for all $s \neq t$, for all b such that $|b| \leq 1$, and for all $\phi \in V$, that

$$b\delta_s \notin \partial f(\phi). \quad (2.3)$$

Otherwise $b\delta_s \in \partial f(\phi)$ for some b with $|b| \leq 1$ and some $\phi \in V$ and therefore $(f(\phi) - b\delta_s(\phi), b) \in M_s$, so that with $a = f(\phi) - b\delta_s(\phi)$ we obtain a pair $(a, b) \in M_s$ such that $a + b\phi(s) \leq f(\phi)$ which is a contradiction. Using the hypothesis that each subdifferential of f contains a multiple of a Dirac measure we conclude that for all $\phi \in V$ there is a b such that $b\delta_t \in \partial f(\phi)$. Further we must have $b \neq 0$ since if $0\delta_t \in \partial f(\phi)$ for some $\phi \in V$ then $0\delta_s \in \partial f(\phi)$ for all $s \neq t$ and this contradicts (2.3). For the proof of our statement there are six cases.

Case (I). $s \notin G$.

We estimate $a + b\phi(s)$ with $|b| \leq 1$, $|(\phi(s) - \phi_0(s))| < b_0(\phi(t) - \phi_0(t))$, and $b_0\delta_t \in \partial f(\phi_0)$, to obtain

$$\begin{aligned} a + b\phi(s) &= a + b\phi_0(s) + b(\phi(s) - \phi_0(s)) \\ &\leq f(\phi_0) + b(\phi(s) - \phi_0(s)) \\ &< f(\phi_0) + b_0(\phi(t) - \phi_0(t)) \\ &\leq f(\phi). \end{aligned}$$

Case (II). $s \in \overline{G}$, $s \neq t$, and $b \geq b_0 > 0$.

We may find $(a', b') \in M_t$ such that $|a - a'| + |b - b'| < \eta(s)$. Then we make the following list of estimates to substitute in

$$a + b\phi(s) = (a - a') + (b - b')\phi(s) + b'(\phi(s) - \phi(t)) + a' + b'\phi(t).$$

We have $a - a' \leq \eta(s)$, $\|\phi\| \leq \|\phi_0\| + \epsilon$, (since $\phi \in U$) and $(b - b')\phi(s) \leq \eta(s)\|\phi\|$. Using $b' \geq \frac{b_0}{2}$, (since $|b - b'| < \eta(s) < \frac{b_0}{2}$ and $b \geq b_0 > 0$) we have $b'(\phi(s) - \phi(t)) < b'(\phi_0(s) - \phi_0(t) - \alpha\eta(s)) < -b'\frac{\alpha\eta(s)}{2} < -b_0\frac{\alpha\eta(s)}{4}$, and $a' + b'\phi(t) \leq f(\phi)$ (since $(a', b') \in M_t$).

The result is

$$\begin{aligned} a + b\phi(s) &< (1 + \|\phi_0\| + \epsilon)\eta(s) - \frac{|b_0|\alpha\eta(s)}{4} + f(\phi) \\ &= -\frac{3}{2}(1 + \|\phi_0\| + \epsilon)\eta(s) + f(\phi) \\ &< f(\phi) \end{aligned}$$

where we use that $\alpha = \frac{10(1 + \|\phi_0\| + \epsilon)}{|b_0|}$.

Case (III). $s \in \overline{G}$, $s \neq t$, and $b \leq -b_0 < 0$.

Using $b\phi(s) < b\phi_0(s)$ (since $b_0 > 0$ gives $0 < \phi(s) - \phi_0(s)$) and $b_0(\phi(t) - \phi_0(t)) > 0$ we estimate that,

$$\begin{aligned} a + b\phi(s) &< a + b\phi_0(s) \\ &\leq f(\phi_0) \\ &< f(\phi_0) + b_0(\phi(t) - \phi_0(t)) \\ &\leq f(\phi). \end{aligned}$$

Case (IV). $s \in \overline{G}$, $s \neq t$, and $b \geq -b_0 > 0$.

Using $b\phi(s) < b\phi_0(s)$ (since $b_0 < 0$ gives $\phi(s) - \phi_0(s) < 0$) and $b_0(\phi(t) - \phi_0(t)) > 0$ we estimate that

$$\begin{aligned} a + b\phi(s) &< a + b\phi_0(s) \\ &\leq f(\phi_0) \\ &< f(\phi_0) + b_0(\phi(t) - \phi_0(t)) \\ &\leq f(\phi). \end{aligned}$$

Case (V). $s \in \overline{G}$, $s \neq t$, and $b \leq b_0 < 0$.

This case is similar to case (III). There are, however, some differences due to changes of sign. We find $(a', b') \in M_t$ such that $|a - a'| + |b - b'| < \eta(s)$ and we substitute the following list of estimates in

$$a + b\phi(s) = (a - a') + (b - b')\phi(s) + b'(\phi(s) - \phi(t)) + a' + b'\phi(t).$$

We have $a - a' \leq \eta(s)$, $\|\phi\| \leq \|\phi_0\| + \epsilon$ (since $\phi \in U$), $(b - b')\phi(s) \leq \eta(s)\|\phi\|$, $b'(\phi(s) - \phi(t)) < b'\frac{\alpha\eta(s)}{2} \leq b_0\frac{\alpha\eta(s)}{4} < 0$ (noting that for all $\phi \in V$ we have, $(\phi(s) - \phi_0(s)) > (\phi(t) - \phi_0(t)) + \alpha\eta(s)$ giving $\phi(s) - \phi(t) - \frac{\alpha\eta(s)}{2} > \phi_0(s) - \phi_0(t) + \frac{\alpha\eta(s)}{2} > 0$, that $b' \leq \frac{b_0}{2} < 0$ since $|b - b'| < \eta(s) < -\frac{b_0}{2}$ and $b \leq b_0 < 0$, and that η is chosen so that $\eta(s) > \frac{2\|\phi_0(s) - \phi_0(t)\|}{\alpha}$), and finally $a' + b'\phi(t) \leq f(\phi)$ (since $(a', b') \in M_t$). We estimate that

$$\begin{aligned} a + b\phi(s) &< (1 + \|\phi_0\| + \epsilon)\eta(s) - \frac{|b_0|\alpha\eta(s)}{4} + f(\phi) \\ &= -\frac{3}{2}(1 + \|\phi_0\| + \epsilon)\eta(s) + f(\phi) \\ &< f(\phi) \end{aligned}$$

where $\alpha = \frac{10(1 + \|\phi_0\| + \epsilon)}{|b_0|}$.

Case (VI). $s \in \overline{G}$, and $|b| < |b_0|$.

For all $\phi \in V$ we have

$$0 < \text{sign}(b_0)(\phi(s) - \phi_0(s)) < \text{sign}(b_0)(\phi(t) - \phi_0(t)) - \alpha\eta(s)$$

so that $b(\phi(s) - \phi_0(s)) < b_0(\phi(t) - \phi_0(t))$. Hence

$$\begin{aligned} a + b\phi(s) &= a + b\phi_0(s) + b(\phi(s) - \phi_0(s)) \\ &< f(\phi_0) + b_0(\phi(t) - \phi_0(t)) \\ &\leq f(\phi). \end{aligned}$$

□

Proposition 2.5.3 *Suppose that $f \in \mathcal{F}_D$, then given any non-empty τ_o -open set U there is a non-empty τ_o -open set $V \subset U$ and a finite set $F \subset K$ such that*

$$\partial f(\phi) \cap \mathcal{M}_1(F) \neq \emptyset$$

for every $\phi \in V$.

PROOF. Suppose that $U = U(\psi, \epsilon, \omega)$ and that $F_0 = \{z \in K : \omega(z) = 0\}$. In the case that $\partial f(\phi) \cap \mathcal{M}_1(F_0) \neq \emptyset$ for every $\phi \in U(\psi, \frac{\epsilon}{4}, \frac{\omega}{4})$ we let $V = U(\psi, \frac{\epsilon}{4}, \frac{\omega}{4})$ and $F = F_0$. Otherwise we find $\phi_0 \in U(\psi, \frac{\epsilon}{4}, \frac{\omega}{4})$ such that $\partial f(\phi_0) \cap \mathcal{M}_1(F_0) = \emptyset$. Then since $f \in \mathcal{F}_D$ we must have $b_0\delta_t \in \partial f(\phi_0)$ for some $b_0 \in \mathbb{R}$ and some $t \notin F_0$. Further we can assume that $b_0 \neq 0$ since this is handled by the first case with $0 \in \mathcal{M}_1(F_0)$. Applying Lemma 2.5.2 we obtain a τ_o -open subset $V \subset U_0$ such that for all $\phi \in V$ there is a $b \neq 0$ such that $b\delta_t \in \partial f(\phi)$. Letting $F = \{t\}$ gives us $\partial f(\phi) \cap \mathcal{M}_1(F) \neq \emptyset$ for all $\phi \in V$. □

In Lemma 2.5.4 we give a well known example of a function in the set \mathcal{F}_D (see [DGZ3]).

Lemma 2.5.4 *For all $\phi \in C(K)$ the subdifferential at ϕ of the supremum norm on $C(K)$, $\partial\|\phi\|$, contains a measure with at most one element support.*

PROOF. We have that $\delta_{x_0} \in \partial\|\varphi\|$ if $\varphi(x_0) = \|\varphi\|$ and that $-\delta_{x_0} \in \partial\|\varphi\|$ if $\varphi(x_0) = -\|\varphi\|$. □

2.6 Measures with finite support

We first reformulate the definition of the wedge topology in somewhat simpler notation. Whenever $F \subset K$ is finite, $\{\eta_x : x \in F\}$ are non-negative functions such that $\eta_x(z) = 0$ if and only if $z = x$, and $\epsilon > 0$. We denote by $U_w(0, \epsilon, \{\eta_x : x \in F\})$ the set of those $\phi \in C(K)$ such that $|\phi(y)| < \epsilon$ for all $y \in K$ and $|\phi(y) - \phi(x)| < \eta_x(y)$ for all $y \neq x$. It is easy to see that these neighbourhoods of the origin define our original wedge topology. Indeed, if $\epsilon, L, R, \{\phi_l\}, \{\psi_r\}$ are as in Definition 2.3.2, we let $F = L \cup R$, use for each $x \in F$, Lemma 2.2.5 with $\hat{\eta} = 0$ to find $\hat{\eta}_x$ such that $\hat{\eta}_x(x) = 0$ and $\hat{\eta}_x(y) > 0$ for $y \neq x$, and define

- $\eta_x(z) = \max\{\phi_x(z), \psi_x(z)\}$ if $x \in L \cap R$,
- $\eta_x(z) = \phi_x(z)$ for $z \geq x$, $\eta_x(z) = \hat{\eta}_x(z)$ for $z < x$ if $x \in L \setminus R$, and
- $\eta_x(z) = \psi_x(z)$ for $z \leq x$, $\eta_x(z) = \hat{\eta}_x(z)$ for $z > x$ if $x \in R \setminus L$.

Then $U_w(0, \epsilon, \{\eta_x : x \in F\}) \subset U_w(0, \epsilon, \{\phi_l\}_{l \in L}, \{\psi_r\}_{r \in R})$. Conversely, if $\epsilon > 0$ and $\{\eta_x : x \in F\}$ are given, we find $0 < \epsilon_1 < \epsilon$ such that

$$\eta_x(z) \geq \epsilon_1 \text{ if } z < x \in F \text{ and } x \notin \mathbf{A}_R, \text{ and}$$

$$\eta_x(z) \geq \epsilon_1 \text{ if } z > x \in F \text{ and } x \notin \mathbf{A}_L.$$

Letting $L = F \cap \mathbf{A}_L$ and $R = F \cap \mathbf{A}_R$, and defining

- $\phi_x(s) = \inf_{z \leq s} \eta_x(z)$ and
- $\psi_x(s) = \inf_{z \geq s} \eta_x(z)$,

we get that

$$U_w(0, \epsilon_1, \{\phi_l\}_{l \in L}, \{\psi_r\}_{r \in R}) \subset U_w(0, \epsilon, \{\eta_x : x \in F\}).$$

We also note that the sets $U_w(0, \epsilon, \{\eta_x : x \in F\})$ are open in the wedge topology: if $\phi \in U_w(0, \epsilon, \{\eta_x : x \in F\})$, let $\hat{\epsilon} = \epsilon - \|\phi\|$, $\hat{\eta}_x(y) = \eta_x(y) - |\phi(y) - \phi(x)|$, and observe that $\phi \in U_w(0, \hat{\epsilon}, \{\hat{\eta}_x : x \in F\}) \subset U_w(0, \epsilon, \{\eta_x : x \in F\})$. Again, we will use the notation $U_w(\phi, \epsilon, \{\eta_x : x \in F\})$ for $\phi + U_w(0, \epsilon, \{\eta_x : x \in F\})$. We show that norm relatively open subsets of $U_w(\phi, \epsilon, \{\eta_x : x \in F\})$ are second category in themselves; this statement is needed in the proof of Lemma 2.6.3.

Lemma 2.6.1 *The sets $U_w(\phi, \epsilon, \{\eta_x : x \in F\})$ have the following properties.*

- (i) $U_w(\varphi, \epsilon, \{\eta_x : x \in F\})$ is of type G_δ in the norm topology.
- (ii) Norm relatively open subsets of $U_w(\varphi, \epsilon, \{\eta_x : x \in F\})$ are second category in themselves.

PROOF. For (i) we note that a wedge neighbourhood is an intersection of a norm ball (which is G_δ) and of a finite number of sets of the form

$$U = \{g : |g(y) - g(x)| < \eta(y) \text{ for all } y \neq x\}$$

where $x \in K$, $\eta : K \rightarrow [0, \infty)$ is continuous and such that $\eta(z) = 0$ iff $z = x$. So it is enough to show that U is G_δ . For this observe that

$$U_n = \{g : |g(y) - g(x)| < \eta(y) \text{ for all } y \text{ such that } \eta(y) \geq 1/n\}$$

is norm open (if $g \in U_n$, the function $y \rightarrow \eta(y) - |g(y) - g(x)|$ attains its minimum, say m , on $\{y : \eta(y) \geq 1/n\}$, so $m > 0$ and the norm ball around g with radius $m/2$ lies in U_n). Since

$$U = \bigcap_{n=1}^{\infty} U_n,$$

U is G_δ .

Assertion (ii) is a corollary of the first statement. Each norm relatively open subset V of a basic wedge open set $U_w(0, \epsilon, \{\eta_x : x \in F\})$ is a G_δ subset of $C(K)$. So V is completely metrisable (see [Ch]) and, as such, is second category in itself. \square

The remainder of this section is devoted to generalising the results of the previous section to the case of continuous convex functions for which there is a measure with finite support in each subdifferential. We state Proposition 2.6.6 which is the main result of this section. Let $\mathcal{M}(F)$ be the set of measures that have support contained in a given set F .

Suppose that U_0 is non-empty and τ open, that f is convex and norm Lipschitz on U_0 . Then there is a non-empty τ open subset $U \subset U_0$, and a finite set $F \subset K$, such that for every $\varphi \in U$,

$$\partial f(\varphi) \cap \mathcal{M}(F) \neq \emptyset.$$

The proof of Proposition 2.6.6 relies on Lemmata 2.6.2 to 2.6.5. First we fix a positive integer n and a norm relatively open subset W of U_0 for which each subdifferential contains a measure with at most n element support. Finally in Lemma 2.6.5, we find a wedge open subset U of W for which all the measures of n element support have support contained in F .

The following notation is used throughout this section.

- U_0 is the wedge open set $U_0 = U_w(\hat{\varphi}_0, \epsilon_0, \{\hat{\eta}_x : x \in F_0\})$.
- The function f is convex, norm Lipschitz on U_0 , and such that for every $\varphi \in U_0$ the set $\partial f(\varphi)$ contains a measure with finite support.
- For any non-negative integer k , \mathcal{M}_k is the set of measures of at most k element support.
- For any integer $k \geq 0$ and any finite set $F \subset K$, $\mathcal{N}_{k,F}$ is the set of measures ν whose support contains at most k points outside F .
- For each integer $0 \leq j \leq k$, $\mathcal{M}_{k,j}$ is the set of measures ν whose support contains exactly k points out of which exactly j points lie outside F_0 , $\mathcal{M}_{k,j,F}$ is the set of measures ν whose support is contained in F and has exactly k points out of which exactly j points lie outside F_0 , and $\mathcal{M}_{k,F}$ is the set of measures ν whose support is contained in F and has exactly k points.
- If $W \subset U_0$ then $\mathcal{M}_{k,j}(W)$ is the set of measures that belong to $\mathcal{M}_{k,j} \cap \partial f(\varphi)$ for some $\varphi \in W$.
- $\overline{\mathcal{M}_{k,j}(W)}$ is the closure of $\mathcal{M}_{k,j}(W)$ in the weak* topology.

The purpose of Lemma 2.6.2 is to fix an integer n and a subset $W \subset C(K)$ such that measures of finite support in the subdifferentials of f on W have n elements in their support outside of a given finite set F .

Lemma 2.6.2 *Suppose that*

- $F \subset K$ is a finite set,
- $\mathcal{M} \subset \mathcal{M}(K)$ is weak* closed,
- $V \subset U_0$ is non-empty and second category in itself in the norm topology, and
- for each $\varphi \in V$, $\mathcal{M} \cap \bigcup_{m=0}^{\infty} \mathcal{N}_{m,F} \cap \partial f(\varphi) \neq \emptyset$.

Then there is a non-negative integer n and a non-empty norm relatively open subset W of V such that

$$\mathcal{M} \cap \mathcal{N}_{n,F} \cap \partial f(\varphi) \neq \emptyset$$

for every $\varphi \in W$ and

$$\overline{\bigcup_{\varphi \in W} \mathcal{M} \cap \mathcal{N}_{n,F} \cap \partial f(\varphi)} \subset \mathcal{N}_{n,F} \setminus \mathcal{N}_{n-1,F}.$$

PROOF. We claim that for each non-negative integer m the set

$$A_m = \{\varphi \in V : \mathcal{M} \cap \partial f(\varphi) \cap \mathcal{N}_{m,F} \neq \emptyset\}$$

is norm relatively closed in V . We first show that $\mathcal{M} \cap \partial f(\varphi) \cap \mathcal{N}_{m,F}$ is weak* closed. To show that each $\mathcal{N}_{m,F}$ is weak* closed let μ be a weak* limit point of a sequence μ_j in $\mathcal{N}_{k,F}$. If $\mu \notin \mathcal{N}_{k,F}$ then there are $m+1$ points in the support of μ outside F , so that for all sufficiently large j , $\text{spt}(\mu_j)$ has $m+1$ points outside F which contradicts μ_j in $\mathcal{N}_{k,F}$. Therefore $\mu \in \mathcal{N}_{k,F}$. From \mathcal{M} and $\partial f(\varphi)$ being weak* closed we get that $\mathcal{M} \cap \partial f(\varphi) \cap \mathcal{N}_{m,F}$ is weak* closed. Suppose that $\varphi_n \in A_m$ converges in norm to $\varphi \in C(K)$. The subdifferential map $\varphi \mapsto \partial f(\varphi)$ is norm to weak* upper semicontinuous on $C(K)$ and therefore each weak* open set $V^* \supset \partial f(\varphi)$ also contains $\partial f(\varphi_n)$ for all sufficiently large n . Suppose that $\mathcal{M} \cap \partial f(\varphi) \cap \mathcal{N}_{m,F} = \emptyset$. Then the weak* open set $C(K)^* \setminus (\mathcal{M} \cap \mathcal{N}_{m,F})$ contains $\partial f(\varphi)$ and therefore contains $\partial f(\varphi_n)$ for all sufficiently large n which contradicts that $\varphi_n \in A_m$. Therefore $\varphi \in A_m$ and hence A_m is norm relatively closed in V . This proves the claim.

We deduce that there is a non-negative integer k and a non-empty norm relatively open subset $\hat{V} \subset V$ such that $\hat{V} \subset A_k$. If not then, since $V = \bigcup_{i=1}^{\infty} A_i$, we deduce that V is first category in itself; but this contradicts the hypothesis that V is not first category in itself.

Let k be the least integer such that there is a non-empty norm relatively open subset $\hat{V} \subset V$ with $\hat{V} \subset A_k$. We claim that there is a non-empty norm relatively open subset W of \hat{V} such that

$$\overline{\bigcup_{\varphi \in W} \mathcal{M} \cap \partial f(\varphi) \cap \mathcal{N}_{k,F}} \subset \mathcal{N}_{k,F} \setminus \mathcal{N}_{k-1,F}.$$

Since $\mathcal{N}_{k,F}$ is weak* closed, for all norm relatively open subsets W of \hat{V} we have $\overline{\bigcup_{\varphi \in W} \mathcal{M} \cap \partial f(\varphi) \cap \mathcal{N}_{k,F}} \subset \mathcal{N}_{k,F}$. We must find a norm relatively open subset W of \hat{V} such that

$$\overline{\bigcup_{\varphi \in W} \mathcal{M} \cap \partial f(\varphi) \cap \mathcal{N}_{k,F} \cap \mathcal{N}_{k-1,F}} = \emptyset.$$

Assuming that no such set W exists, we choose for each norm relatively open subset H of \hat{V} , a measure $\mu_H \in \overline{\bigcup_{\varphi \in H} \mathcal{M} \cap \partial f(\varphi) \cap \mathcal{N}_{k,F} \cap \mathcal{N}_{k-1,F}}$, and we let Δ denote the weak* closure of the set of such measures.

Since $\mathcal{N}_{k-1,F}$ is weak* closed we deduce that $\Delta \subset \mathcal{N}_{k-1,F}$. We show that $\partial f(\varphi) \cap \Delta \neq \emptyset$ for all $\varphi \in \hat{V}$ concluding that $M \cap \partial f(\varphi) \cap \mathcal{N}_{k-1,F} \neq \emptyset$ for all $\varphi \in \hat{V}$, which contradicts k being minimal with this property. We seek a contradiction by assuming that $\partial f(\varphi) \cap \Delta = \emptyset$ for some $\varphi \in \hat{V}$. Choosing a weak* open neighbourhood O of $\partial f(\varphi)$ whose weak* closure does not meet Δ , and noting that the subdifferential map $\varphi \mapsto \partial f(\varphi)$ is norm to weak* upper semicontinuous on $C(K)$, we may find a norm open neighbourhood H_0 of φ such that $\partial f(H_0) \subset O$. Therefore the weak* closure of $\bigcup_{\psi \in H_0} \partial f(\psi)$ does not meet Δ . By assumption there is a measure $\mu_{H_0} \in (\bigcup_{\psi \in H_0} \partial f(\psi)) \cap \Delta$ giving a contradiction and proving the claim. \square

In Lemma 2.6.3 we find integers k and j for which the subdifferentials of f on a non-empty norm relatively open subset W of U_0 contain measures of k element support out of which j are not in F_0 .

Lemma 2.6.3 *There are integers $0 \leq j \leq k$ and a non-empty norm relatively open subset W of U_0 such that*

$$\mathcal{M}_{k,j} \cap \partial f(\varphi) \neq \emptyset$$

for all $\varphi \in W$, and furthermore

$$\overline{\mathcal{M}_{k,j}(W)} \subset \mathcal{M}_{k,j}.$$

In particular there are numbers $0 < c < C < \infty$ such that each measure $\nu \in \overline{\mathcal{M}_{k,j}(W)}$ has support consisting of exactly k elements out of which exactly j are outside F_0 , and $c \leq |\nu(x)| \leq C$ for each element x in $\text{spt}(\nu)$.

PROOF. Applying Lemma 2.6.2 with $F = \emptyset$ so that $\mathcal{N}_{k,F} = \mathcal{M}_k$, with $\mathcal{M} = \mathcal{M}(K)$, and with $V = U_0$, we find a non-empty norm relatively open subset W_0 of U_0 and a non-negative integer k such that

$$\partial f(\varphi) \cap \mathcal{M}_k \neq \emptyset$$

for all $\varphi \in W_0$ and

$$\bigcup_{\varphi \in W_0} \overline{\mathcal{M}_k \cap \partial f(\varphi)} \subset \mathcal{M}_k \setminus \mathcal{M}_{k-1}.$$

Applying Lemma 2.6.2 a second time with $F = F_0$, with $\mathcal{M} = \mathcal{M}_k$, and with $V = W_0$ we find a non-empty norm relatively open subset W of W_0 and a non-negative integer j such that

$$\mathcal{M}_k \cap \partial f(\varphi) \cap \mathcal{N}_{j, F_0} \neq \emptyset$$

for all $\varphi \in W$, and

$$\overline{\bigcup_{\varphi \in W} \mathcal{M}_k \cap \partial f(\varphi) \cap \mathcal{N}_{j, F_0}} \subset \mathcal{N}_{j, F_0} \setminus \mathcal{N}_{j-1, F_0}.$$

This proves the first part of the statement since $(\mathcal{M}_k \setminus \mathcal{M}_{k-1}) \cap (\mathcal{N}_{j, F_0} \setminus \mathcal{N}_{j-1, F_0}) = \mathcal{M}_{k, j}$.

That each measure $\nu \in \overline{\mathcal{M}_{k, j}(W)}$ has support consisting of exactly k elements out of which exactly j are outside F_0 , is a reformulation of the inclusion $\overline{\mathcal{M}_{k, j}(W)} \subset \mathcal{M}_{k, j}$.

The existence of C such that $|\nu(x)| \leq C$ for all $\nu \in \partial f(\varphi)$ for all $\varphi \in U_0$ and all $x \in \text{spt}(\nu)$ follows from the assumption that f is Lipschitz on U_0 . Suppose that there is no positive c such that for all $\nu \in \overline{\mathcal{M}_{k, j}(W)}$ we have $c \leq |\nu(x)|$ for each element $x \in \text{spt}(\nu)$. Then for each positive integer p we may find a measure $\nu_p \in \overline{\mathcal{M}_{k, j}(W)}$ such that $|\nu_p(x_p)| < 1/p$ for some $x_p \in \text{spt}(\nu_p)$. Since the ν_p have exactly k element support and K is compact we obtain a contradiction with a weak* accumulation point of the sequence ν_p that belongs to $\overline{\mathcal{M}_{k, j}(W)}$ and has less than k element support. \square

For Proposition 2.6.6 we need to find a wedge neighbourhood $U = U_w(\varphi_0, \epsilon, \{\eta_x : x \in F\})$ such that, if $\varphi \in U$ then every measure in $\partial f(\varphi)$ of finite support necessarily has support in F . We choose some $\varphi_0 \in W$ and some measure $\mu_0 \in \mathcal{M}_{k, j} \cap \partial f(\varphi_0)$, and let $F = \text{spt}(\mu_0)$. The functions η_x are defined on decreasing sequence of open sets G_p whose intersection is F . The sequence G_p is given in the following Lemma.

Lemma 2.6.4 *Suppose that W is the wedge open set of Lemma 2.6.3, that $\varphi_0 \in W$, and that*

- x_1, \dots, x_k are distinct points of K , such that for some integer $j \geq 1$ we have $x_i \notin F_0$ for $1 \leq i \leq j$, and $x_i \in F_0$ for $i > j$,
- $G \ni x_1, \dots, x_j$ is an open set such that $G \cap F_0 = \emptyset$, and
- d_p is an arbitrary fixed sequence of positive real numbers,

Then there are open sets $G_{i,p} \subset K$ ($i = 1, \dots, k$, $p = 1, \dots$) with the following properties:

- (i) $x_i \in G_{i,p+1} \subset \overline{G_{i,p+1}} \subset G_{i,p}$,
- (ii) $G_{i,p} \subset G$ if $i \leq j$,
- (iii) $G_{i,p} \cap F_0 = \{x_i\}$ if $i > j$,
- (iv) $G_{i_1,p} \cap G_{i_2,p} = \emptyset$ for $i_1 \neq i_2$,
- (v) $|\varphi_0(x) - \varphi_0(x_i)| < d_p$ for every $x \in G_{i,p}$,
- (vi) the intersection of the sets $G_p = \bigcup_{i=1}^k G_{i,p}$ is $\{x_1, \dots, x_k\}$,
- (vii) if $\nu \in \overline{\mathcal{M}_{k,j}(W)}$ is concentrated in G_p , then $\text{spt}(\nu) = \{y_1, \dots, y_k\}$, where $y_i \in G_{i,p}$ for all i and $y_i = x_i$ if $i > j$, and
- (viii) if $\varphi \in W$ and $\nu \in \mathcal{M}_{k,j} \cap \partial f(\varphi)$ is concentrated in G_p , then there is $\mu \in \overline{\mathcal{M}_{k,j}(W)} \cap \mathcal{M}(\{x_1, \dots, x_k\})$ such that $\sum_{i=1}^k |\mu(G_{i,p}) - \nu(G_{i,p})| < d_p$ and $\nu(\varphi - \varphi_0) > \mu(\varphi - \varphi_0) - d_p$.

PROOF. Since K is first countable for each $i = 1, \dots, k$ we may choose a basis $G_{i,p}$ of open neighbourhoods of x_i such that properties (i) – (vi) hold.

For (vii) suppose that $\nu \in \mathcal{M}_{k,j}$. We claim that there is an integer p_0 such that for $p \geq p_0$ the support of every measure concentrated on G_p has non-empty intersection with every $G_{i,p}$. Indeed, if this were not the case there would be arbitrarily large p for which we could find $\nu_p \in \overline{\mathcal{M}_{k,j}(W)}$ concentrated on G_p such that $|\nu_p|(G_{i_p,p}) = 0$ for suitable i_p . We can pass to a subsequence on which i_p is constant, say $i_{p_j} = i$. Using (i), (vi), and that $\text{spt}(\nu_p) \subset G_p$ we infer that there is a weak* accumulation point ν of ν_p with $\text{spt}(\nu) \subset \{x_1, \dots, x_k\} \setminus \{x_i\}$. We have a contradiction since $\nu \in \mathcal{M}_{k,j}(W)$ and it has exactly k element support.

Replacing $G_{i,p}$ by $G_{i,p+p_0}$ if necessary, we may assume that $p_0 = 1$. By the previous claim we have that every measure $\nu \in \overline{\mathcal{M}_{k,j}(W)}$ concentrated on G_p has $\text{spt}(\nu) = \{y_1, \dots, y_k\}$ where $y_i \in G_{i,p}$ for $i = 1, \dots, k$. By Lemma 2.6.3 exactly j of $\{y_1, \dots, y_k\}$ are outside F_0 , but $y_1, \dots, y_j \notin F_0$ because of (ii). We infer that $y_i \in F_0$ for $i > j$. By (iii) we must have that $y_i = x_i$ for $i > j$, which concludes the proof of (vii).

For (viii) let $A_p \subset \mathbf{R}^{k+2}$ be defined by

$$A_p = \{(\nu(G_{1,p}), \dots, \nu(G_{k,p}), f(\varphi) - \nu(\varphi), \nu(\varphi_0)) : \varphi \in W, \nu \in \partial f(\varphi) \cap \mathcal{M}_{k,j}, \text{spt}(\nu) \subset G_p\}.$$

Then A_p is a decreasing sequence of bounded non-empty subsets of \mathbf{R}^{k+2} and therefore the set

$$A = \bigcap_{p=1}^{\infty} \overline{A_p}$$

is non-empty and compact. Moreover, the sequence

$$\gamma_p = \sup_{a \in A_p} \inf \left\{ \sum_{j=1}^{k+2} |a_j - b_j| : b \in A \right\}$$

tends to zero. We may assume that $\gamma_p < d_p/3$, since, if necessary, we may replace each $G_{i,p}$ with G_{i,q_p} , for some q_p so large that $\gamma_{p_q} < d_p/3$. The validity of (i) – (vii) would be unaffected after such a replacement.

Suppose that $\varphi \in W$ and that $\nu \in M_{k,j} \cap \partial f(\varphi)$ is concentrated in G_p , then since $\gamma_p < d_p/3$, there is $b = (b_1, \dots, b_{k+2}) \in A$ such that

$$|f(\varphi) - \nu(\varphi) - b_{k+1}| + |\nu(\varphi_0) - b_{k+2}| + \sum_{j=1}^k |\nu(G_{j,p}) - b_j| < d_p/3.$$

Since $A = \bigcap_{p=1}^{\infty} \overline{A_p}$, for each $q = 1, 2, \dots$ we may find $\varphi_q \in W$ and $\mu_q \in M_{k,j} \cap \partial f(\varphi_q)$ such that $\text{spt}(\mu_q) \subset G_q$ and $|f(\varphi_q) - \nu(\varphi_q) - b_{k+1}| + |\mu_q(\varphi_0) - b_{k+2}| + \sum_{j=1}^k |\mu_q(G_{j,q}) - b_j|$ is so small that

$$\begin{aligned} & |(f(\varphi) - \nu(\varphi)) - (f(\varphi_q) - \mu_q(\varphi_q))| + |\nu(\varphi_0) - \mu_q(\varphi_0)| \\ & + \sum_{j=1}^k |\nu(G_{j,p}) - \mu_q(G_{j,q})| < \frac{d_p}{3}. \end{aligned}$$

We infer from (i), (vi), and (vii) that all weak* accumulation points μ of the sequence μ_q , (which exist by compactness of K) belong to $\overline{\mathcal{M}_{k,j}(W)} \cap \mathcal{M}(\{x_1, \dots, x_k\})$, and satisfy

$$|\nu(\varphi_0) - \mu(\varphi_0)| \leq d_p/3$$

and

$$\sum_{j=1}^k |\nu(G_{j,p}) - \mu(G_{j,q})| \leq \frac{d_p}{3} < d_p.$$

Using also that $f(\varphi_q) - \mu_q(\varphi_q) > f(\varphi) - \nu(\varphi) - d_p/3$ and that $\mu_q \in \partial f(\varphi_q)$, we infer that

$$f(\varphi) \geq f(\varphi_q) - \mu_q(\varphi_q) + \mu_q(\varphi) \geq f(\varphi) - \nu(\varphi) - d_p/3 + \mu_q(\varphi).$$

So letting $q \rightarrow \infty$ we obtain $\nu(\varphi) \geq \mu(\varphi) - d_p/3$. Recalling that $|\nu(\varphi_0) - \mu(\varphi_0)| \leq d_p/3$ we conclude that

$$\nu(\varphi - \varphi_0) \geq \mu(\varphi - \varphi_0) - 2d_p/3 > \mu(\varphi - \varphi_0) - d_p.$$

□

In the following Lemma, we obtain the wedge open set U on which each subdifferentials of f contains a measure with fixed finite element support.

Lemma 2.6.5 *There are integers $0 \leq j \leq k$, a non-empty τ open subset $U \subset U_0$, and a finite set $F \subset K$ such that*

$$\partial f(\varphi) \cap \mathcal{M}_{k,j,F} \neq \emptyset$$

for every $\varphi \in U$.

PROOF. The proof has two parts. In the first part we construct the functions $\{\eta_x\}_{x \in F}$ that we require to define U . In the last part we verify that the measures of finite support in the subdifferentials of f on U do indeed have support contained in the finite set F .

We may apply Lemma 2.6.3 to obtain a non-empty norm relatively open subset W of U_0 and integers $0 \leq j \leq k$ such that

$$\partial f(\varphi) \cap \mathcal{M}_{k,j} \neq \emptyset$$

for all $\varphi \in W$. If $j = 0$ then all of the k element support of $\nu \in \partial f(\varphi) \cap \mathcal{M}_{k,j}$ are in F_0 and therefore we can set $U = W$ and $F = F_0$.

For $j \geq 1$ we use the second part of Lemma 2.6.3 to obtain $\overline{\mathcal{M}_{k,j}(W)} \subset \mathcal{M}_{k,j}$ and constants $0 < c < C < \infty$ such that each measure $\nu \in \overline{\mathcal{M}_{k,j}(W)}$ has support consisting of exactly k elements out of which exactly j are outside F_0 , and $c \leq |\nu(x)| \leq C$ for each element x in $\text{spt}(\nu)$.

Since for all measures $\mu \in \overline{\mathcal{M}_{k,j}(W)}$ we have that $|\mu(x)| \leq C$ for each element x in $\text{spt}(\mu)$ then letting $s = \sup_{\mu \in \overline{\mathcal{M}_{k,j}(W)}} |\mu|(K \setminus F_0)$, we have $s < \infty$ and, since $K \setminus F_0$ is open, $s = \sup_{\mu \in \mathcal{M}_{k,j}(W)} |\mu|(K \setminus F_0)$. So there are $\varphi_0 \in W$ and $\mu_0 \in \mathcal{M}_{k,j} \cap \partial f(\varphi_0)$ such that $|\mu_0|(K \setminus F_0) > s - c/4$.

We may suppose that

- $\text{spt}(\mu_0) = \{x_1, \dots, x_k\}$ where x_1, \dots, x_k are distinct points such that x_i is not in F_0 for $1 \leq i \leq j$ and x_i is in F_0 for $i > j$,
- $G \ni x_1, \dots, x_j$ is an open set such that $\overline{G} \cap F_0 = \emptyset$,
- functions η_x are chosen so that $\eta_x(y) = \frac{1}{2}(\hat{\eta}_x(y) - |(\varphi_0(y) - \varphi_0(x)) - (\hat{\varphi}_0(y) - \hat{\varphi}_0(x))|)$, for $x \in F_0$,
- $\kappa = \frac{1}{2} \min\{\delta, \epsilon_0 - \|\varphi_0 - \hat{\varphi}_0\|, \inf\{\eta_x(y) : x \in F_0, y \in G\}\}$ is positive where $\delta > 0$ is chosen so that $\{\psi \in U_0 : \|\psi - \varphi_0\| < \delta\} \subset W$,
- $d_p = 2^{-p-1} \frac{\kappa c}{1+2k\kappa}$, and

- $G_{i,p} \subset K$ ($i = 1, \dots, k$, $p = 1, \dots$) and G_p are the open sets of Lemma 2.6.4.

Then applying Lemma 2.2.4 we may find for each integer $1 \leq i \leq k$ and each integer $p \geq 1$ a function $g_{i,p} \in C(K)$ such that,

- $g_{i,p}(x) = 0$ for $x \in K \setminus G_{i,p}$,
- $g_{i,p}(x) = \text{sign}(\mu_0(x_i))$ for $x \in G_{i,p+1}$, and
- $0 \leq |g_{i,p}| \leq 1$.

We make the following claims.

Claim (i) Let $g = \kappa \sum_{p=1}^{\infty} 2^{-p} \sum_{i=1}^j g_{i,p}$ then

- $g(x_i) = \kappa \text{sign}(\mu_0(x_i))$ for $1 \leq i \leq j$,
- $\kappa(1 - 2^{-p+1}) \leq |g(x)| \leq \kappa(1 - 2^{-p})$ for $x \in \cup_{i=1}^j G_{i,p} \setminus \cup_{i=1}^j G_{i,p+1}$, and
- $g = 0$ outside $\cup_{i=1}^j G_{i,1}$, in particular $g(x_i) = 0$ for $i > j$.

Claim (ii) For $1 \leq i \leq j$ let $\eta_{x_i}(y) = \sup_{p \geq 1} 2^{-p-1} \kappa(1 - |g_{i,p}|(y))$, then

- $\eta_{x_i} \in C(K)$,
- $\eta_{x_i}(x_i) = 0$ and $\eta_{x_i}(y) > 0$ if $y \neq x_i$, and
- $\eta_{x_i}(y) \leq 2^{-p-1} \kappa$ if $y \in G_p$.

Claim (iii) Choosing $0 < \epsilon < 1/4$ such that

$$(1 - \epsilon)(s - c/4) > (1 + \epsilon)s - c/2 + 2k\epsilon C, \quad (2.4)$$

and $F = F_0 \cup \{x_1, \dots, x_j\}$, we let $U = U(\varphi_0 + g, \epsilon\kappa, \{\eta_x \mid x \in F\})$. Then $U \subset W$ and if $\varphi_0 + h \in U$ then,

- $|h(y)| < \kappa(1 + \epsilon)$ for all y ,
- $|h(y)| < \kappa(1/2 + \epsilon)$ for $y \notin G_2$,
- $|h(x)| < \kappa\epsilon$ for $x \in F_0$, and
- if $1 \leq i \leq j$ then $\text{sign } h(y) = \text{sign}(\mu_0(x_i))$ for all $y \in G_{i,2}$.

For claim (i), noting that $g_{i,p}(x_i) = \text{sign}(\mu_0(x_i))$ for all i and p , and that $g_{m,p}(x_i) = 0$ if $m \neq i$, we obtain

$$\begin{aligned} g(x_i) &= \kappa \sum_{p=1}^{\infty} 2^{-p} \text{sign}(\mu_0(x_i)) \\ &= \kappa \text{sign}(\mu_0(x_i)). \end{aligned}$$

Let $x \in \cup_{i=1}^j G_{i,p} \setminus \cup_{i=1}^j G_{i,p+1}$, then $g_{i,q}(x) = 0$ for $q \geq p+1$ and since there is exactly one i_x such that $1 \leq i_x \leq j$ and that $x \in G_{i_x,p} \setminus G_{i_x,p+1}$ we have that $g_{i_x,q}(x) = \text{sign}(\mu_0(x_{i_x}))$ for $q \leq p-1$ and $g_{m,q}(x) = 0$ if $m \neq i_x$. Therefore

$$\begin{aligned} g(x) &= \kappa \sum_{q=1}^{p-1} 2^{-q} \text{sign}(\mu_0(x_{i_x})) + \kappa 2^{-p} g_{i_x,p}(x) \\ &= \kappa(1 - 2^{-p+1}) \text{sign}(\mu_0(x_{i_x})) + \kappa 2^{-p} g_{i_x,p}(x). \end{aligned}$$

Since $g_{i_x,p}(x)$ takes values between 0 and $\text{sign}(\mu_0(x_{i_x}))$ we have that

$$\kappa(1 - 2^{-p+1}) \leq |g(x)| \leq \kappa(1 - 2^{-p}).$$

If $x \notin \cup_{i=1}^j G_{i,1}$ then $g_{i,p}(x) = 0$ for all i and p . Therefore $g = 0$ outside $\cup_{i=1}^j G_{i,1}$, in particular $g(x_i) = 0$ for $i > j$.

For claim (ii), if $y \in G_{i,p}$, then $|g_{i,q}(y)| = 1$ for $q < p$. So $\eta_{x_i}(y) = \sup_{q \geq p} 2^{-p-1} \kappa(1 - |g_{i,q}(y)|) \leq 2^{-p-1} \kappa$, which is the last statement of (ii). It also shows that $\eta_{x_i}(x_i) = 0$ and that η_{x_i} is continuous at x_i . If $y \notin \overline{G_{i,p}}$, then $|g_{i,q}(y)| = 0$, so $\eta_{x_i}(y) \geq 2^{-p-1} \kappa$. Hence

$$\eta_{x_i}(y) = \sup_{1 \leq q \leq p} 2^{-p-1} \kappa(1 - |g_{i,q}(y)|)$$

and so η_{x_i} is positive and continuous at y . Finally, we recall that $\cap_{p=1}^{\infty} \overline{G_{i,p}} = \{x_i\}$ to infer that η_{x_i} is positive and continuous at every point of $K \setminus \{x_i\}$.

For claim (iii), we first show that $U \subset W$. If $\varphi \in U$ then using

- $\|\varphi - (\varphi_0 + g)\| < \epsilon \kappa$,
- $\|g\| \leq \kappa$, and
- $\kappa \leq \frac{1}{2} \min\{\delta, \epsilon_0 - \|\varphi_0 - \hat{\varphi}_0\|\}$,

we estimate that

$$\|\varphi - \varphi_0\| \leq \epsilon \kappa + \kappa < 2\kappa \leq \delta.$$

It remains to show that $\varphi \in U_0 = U(\hat{\varphi}_0, \epsilon_0, \{\hat{\eta}_x : x \in F_0\})$. For the norm of $\varphi - \hat{\varphi}_0$ we have

$$\|\varphi - \hat{\varphi}_0\| \leq \|\varphi - \varphi_0\| + \|\varphi_0 - \hat{\varphi}_0\| < 2\kappa + \|\varphi_0 - \hat{\varphi}_0\| < \epsilon_0.$$

For the oscillation of $\varphi - \hat{\varphi}_0$ we need only consider $x \in F_0$. Recall that $\eta_x(y) = \frac{1}{2}(\hat{\eta}_x(y) - |(\varphi_0(y) - \varphi_0(x)) - (\hat{\varphi}_0(y) - \hat{\varphi}_0(x))|)$. In the case when $y \notin G_1$ then $g(y) = g(x) = 0$, and we estimate that,

$$\begin{aligned} |(\varphi(y) - \varphi(x)) - (\hat{\varphi}_0(y) - \hat{\varphi}_0(x))| &< \eta_x(y) + |g(y) - g(x)| + \\ &|(\varphi_0(y) - \varphi_0(x)) - (\hat{\varphi}_0(y) - \hat{\varphi}_0(x))| \\ &< \hat{\eta}_x(y). \end{aligned}$$

In the case that $y \in G_1$, using $g(x) = 0$ so that $|g(y) - g(x)| \leq \kappa$, and using $\kappa \leq \frac{1}{2} \inf\{\eta_x(y) : x \in F_0, y \in G\}$, we estimate that

$$\begin{aligned} |(\varphi(y) - \varphi(x)) - (\hat{\varphi}_0(y) - \hat{\varphi}_0(x))| &< \eta_x(y) + \\ &|(\varphi_0(y) - \varphi_0(x)) - (\hat{\varphi}_0(y) - \hat{\varphi}_0(x))| + \kappa \\ &\leq 2\eta_x(y) + |(\varphi_0(y) - \varphi_0(x)) - (\hat{\varphi}_0(y) - \hat{\varphi}_0(x))| \\ &= \hat{\eta}_x(y). \end{aligned}$$

Therefore $\varphi \in W$. Hence $U \subset B(\varphi_0, \delta) \cap U_0 \subset W$.

Finally, suppose that $\varphi_0 + h \in U$. Since $U = U(\varphi_0 + g, \epsilon\kappa, \{\eta_x : x \in F\})$, we have that $|h(y)| < |g(y)| + \epsilon\kappa \leq \kappa(1 + \epsilon)$ for all y . If $y \notin G_2$, then $|g(y)| \leq \kappa/2$ so that $|h(y)| < \kappa(1/2 + \epsilon)$. If $x \in F_0$ then $g(x) = 0$ so that $|h(x)| < \kappa\epsilon$. Suppose that $y \in G_{i,2}$ and $1 \leq i \leq j$. Noting that $\|h - g\| \leq \kappa\epsilon$, that $\epsilon < 1/4$, and that $|g(y)| \geq \kappa/2$, we deduce that $|h(y) - g(y)| < \kappa/4$ and hence that $\text{sign}(h(y)) = \text{sign}(g(y)) = \text{sign}(\mu_0(x_i))$.

We finish the proof of the Lemma by showing that, *whenever $\varphi \in U$ and $\nu \in \mathcal{M}_{k,j} \cap \partial f(\varphi)$ then ν is concentrated on the set $F_0 \cup \{x_1, \dots, x_j\}$* . Assuming, in order to find a contradiction, that this is not the case, we use that $K \setminus (F_0 \cup \{x_1, \dots, x_j\}) = \cup_{q=1}^{\infty} (K \setminus G_{q+1})$ to find the least q such that $|\nu|(K \setminus G_{q+1}) > 0$.

Let $\varphi = \varphi_0 + h$. Since $\nu \in \partial f(\varphi)$ and $\mu_0 \in \partial f(\varphi_0)$, we have $f(\varphi) \geq f(\varphi_0) + \mu_0(h) \geq f(\varphi) - \nu(h) + \mu_0(h)$, so that $\nu(h) \geq \mu_0(h)$ (this property is known as monotonicity of the subdifferential). Since $|h(x_i) - g(x_i)| < \epsilon\kappa$,

$$\mu_0(x_i)h(x_i) \geq \mu_0(x_i)g(x_i) - \epsilon\kappa|\mu_0(x_i)| = \kappa|\mu_0(x_i)|(1 - \epsilon)$$

for $1 \leq i \leq j$ and $|h(x_i)| \leq \kappa\epsilon$ for $i > j$, we estimate that

$$\begin{aligned}
 \nu(h) &\geq \mu_0(h) \\
 &= \sum_{i=1}^k \mu_0(x_i) h(x_i) \\
 &\geq \sum_{i=1}^j \kappa(1-\epsilon) |\mu_0(x_i)| + \sum_{i=j+1}^k -\kappa\epsilon |\mu_0(x_i)| \\
 &= \kappa(1-\epsilon) |\mu_0|(K \setminus F_0) - \kappa\epsilon |\mu_0|(F_0) \\
 &\geq \kappa(1-\epsilon)(s - c/4) - k\kappa\epsilon C.
 \end{aligned} \tag{2.5}$$

We distinguish two cases; the case when ν is non zero on the set where $|h|$ has small values, that is when $q = 1$, and the remaining case when $q \geq 2$.

Case (1) If $q = 1$, then we use Lemma 2.6.3 and that $\text{spt}(\nu) \cap (K \setminus G_2)$ has at least one element, to infer that $|\nu|(K \setminus G_2) \geq c$. Using claim (iii), that $|h(x)| < \kappa(1+\epsilon)$ for all x , that $|h(x)| < \kappa(1/2+\epsilon)$ for $x \notin G_2$, and that $|h(x)| < \kappa\epsilon$ for $x \in F_0$, we estimate that,

$$\begin{aligned}
 \nu(h) &\leq \kappa(1+\epsilon) |\nu|(G_2) + \kappa(1/2+\epsilon) |\nu|(K \setminus (G_2 \cup F_0)) \\
 &\quad + \kappa\epsilon |\nu|(F_0) \\
 &= \kappa(1+\epsilon) |\nu|(G_2) + \kappa(1+\epsilon) |\nu|(K \setminus (G_2 \cup F_0)) \\
 &\quad - \frac{\kappa |\nu|(K \setminus (G_2 \cup F_0))}{2} + \kappa\epsilon |\nu|(F_0) \\
 &= \kappa(1+\epsilon) |\nu|(K \setminus F_0) - \\
 &\quad \frac{\kappa |\nu|(K \setminus (G_2 \cup F_0))}{2} + \kappa\epsilon |\nu|(F_0) \\
 &\leq \kappa(1+\epsilon)s - \frac{\kappa c}{2} + k\kappa\epsilon C.
 \end{aligned}$$

From (2.5) we deduce that

$$(1-\epsilon)(s - c/4) \leq (1+\epsilon)s - c/2 + 2k\epsilon C,$$

which contradicts (2.4).

Case (2) If $q \geq 2$, noting that $|\nu|(K \setminus G_q) = 0$ and since $|\nu|(K \setminus G_{q+1}) > 0$ we infer that $|\nu|(G_q \setminus G_{q+1}) \geq c$. In particular ν is concentrated in G_q and, by (viii) of Lemma 2.6.4, we have that $\text{spt}(\nu) = \{y_1, \dots, y_k\}$, where $y_i \in G_{i,q}$ for all i and $y_i = x_i$ if $i > j$, and there is $\mu \in \overline{\mathcal{M}_{k,j}(W)} \cap \mathcal{M}(\{x_1, \dots, x_k\})$ such that $\sum_{i=1}^k |\mu(G_{i,q}) - \nu(G_{i,q})| < d_q$ and $\nu(h) > \mu(h) - d_q$.

The measure μ may be close to μ_0 or it may be far from it. Since the treatment of these two situations is different, we distinguish two subcases and consider the far away case first.

Sub-case (a) Suppose that there is an integer m , $1 \leq m \leq j$, such that $\mu(x_m)$ and $\mu_0(x_m)$ have opposite signs. We have $|\nu(y_m) - \mu(x_m)| < d_q \leq c \leq |\nu(y_m)|$. Therefore $\text{sign}(\nu(y_m)) = \text{sign}(\mu(x_m))$. By claim (iii) $h(y_m)$ has the same sign as $\mu_0(x_m)$ and we conclude that $\nu(y_m)$ and $h(y_m)$ have opposite signs. Hence $\int_{G_{m,q}} h d\nu = \nu(y_m)h(y_m) \leq 0$, and we obtain the following estimate for $\nu(h)$:

$$\begin{aligned}
 \nu(h) &= \int_{G_q \setminus G_{m,q}} h d\nu + \int_{G_{m,q}} h d\nu \\
 &\leq \int_{G_q \setminus G_{m,q}} h d\nu \\
 &= \int_{G_q \setminus (G_{m,q} \cup F_0)} h d\nu + \int_{(G_q \setminus G_{m,q}) \cap F_0} h d\nu \\
 &\leq \kappa(1 + \epsilon)|\nu|(G_q \setminus (G_{m,q} \cup F_0)) + \kappa\epsilon|\nu|((G_q \setminus G_{m,q}) \cap F_0) \\
 &\leq \kappa(1 + \epsilon)[|\nu|(K \setminus F_0) - |\nu|(G_{m,q})] + \kappa\epsilon|\nu|(F_0) \\
 &\leq \kappa(1 + \epsilon)(s - c) + k\kappa\epsilon C.
 \end{aligned}$$

From (2.5) we deduce that

$$\begin{aligned}
 (1 - \epsilon)(s - c/4) &\leq (1 + \epsilon)(s - c) + 2k\epsilon C \\
 &\leq (1 + \epsilon)s - \frac{c}{2} + 2k\epsilon C
 \end{aligned}$$

which contradicts (2.4).

The following final sub-case, although the proof is as short as the previous ones, is the one for which most of the above work has been done. It is only here that the use of the wedge topology is essential, and also where we use more than just the monotonicity of the subdifferential.

Sub-case (b) Suppose that for all $1 \leq i \leq j$ that $\mu(x_i)$ and $\mu_0(x_i)$ have the same sign. If $\text{spt}(\nu) = \{y_1, \dots, y_k\}$, where $y_i \in G_{i,q}$ for all i and $y_i = x_i$ if $i > j$, then using the minimality of q we have $y_m \in G_{m,q} \setminus G_{m,q+1}$ for some $1 \leq m \leq j$. We note that $h + \varphi_0 \in U = U(\phi_0 + g, \epsilon\kappa, \{\eta_x : x \in F\})$ implies that

$$|(h(y_i) - h(x_i)) - (g(y_i) - g(x_i))| < \eta_{x_i}(y_i).$$

If $1 \leq i \leq j$ and $y_i \in G_{i,l} \setminus G_{i,l+1}$ then $l \geq q$ and by claim (i) and claim (ii) we have $|g(x_i) - g(y_i)| \geq \kappa 2^{-l-1}$ and $\eta_{x_i}(y_i) \leq \kappa 2^{-l-1}$. Therefore $h(x_i) - h(y_i)$ has the same sign as $g(x_i) - g(y_i)$, namely, sign $(\mu_0(x_i))$. In the only other possible case we have $y_i = x_i$, so that $h(y_i) = h(x_i)$.

Furthermore for $y_m \in G_{m,q} \setminus G_{m,q+1}$ we have

$$\begin{aligned} |h(x_m) - h(y_m)| &> |g(x_m) - g(y_m)| - \eta_{x_m}(y_m) \\ &\geq (\kappa - \kappa(1 - 2^{-q})) - \eta_{x_m}(y_m) \\ &\geq \kappa 2^{-q} - \eta_{x_m}(y_m) \\ &\geq \kappa 2^{-q} - \kappa 2^{-q-1} = \kappa 2^{-q-1}. \end{aligned}$$

We recapitulate that for each $i = 1, \dots, j$ we have

- $\mu(x_i), \mu_0(x_i)$, and $h(x_i) - h(y_i)$ all have the same sign, or $h(y_i) = h(x_i)$,
- $|\nu(y_i) - \mu(x_i)| < d_q$,
- $y_i = x_i$ if $i > j$,
- $|h(x_i)| \leq (1 + \epsilon)\kappa < 2\kappa$, and
- for $y_m \in G_{m,q} \setminus G_{m,q+1}$ we have $|h(x_m) - h(y_m)| \geq \kappa 2^{-q-1}$.

We estimate that

$$\begin{aligned} \nu(h) &= \sum_{i=1}^k \nu(y_i) h(y_i) \\ &\leq \sum_{i=1}^k \mu(x_i) h(y_i) + 2k\kappa d_q \\ &= \sum_{i=1}^j \mu(x_i) h(y_i) + \sum_{i=j+1}^k \mu(x_i) h(y_i) + 2k\kappa d_q \\ &= \sum_{i=1}^j \mu(x_i) h(x_i) + \sum_{i=1}^j \mu(x_i) (h(y_i) - h(x_i)) + \\ &\quad \sum_{i=j+1}^k \mu(x_i) h(x_i) + 2k\kappa d_q \\ &= \sum_{i=1}^k \mu(x_i) h(x_i) - \sum_{i=1}^j |\mu(x_i)| |h(y_i) - h(x_i)| + 2k\kappa d_q \\ &\leq \mu(h) - \kappa 2^{-q-1} c + 2k\kappa d_q. \end{aligned}$$

From (viii) of Lemma 2.6.4, we have $\nu(h) > \mu(h) - d_q$, and therefore $-d_q < -\kappa 2^{-q-1}c + 2k\kappa d_q$ so that $d_q > 2^{-q-1} \frac{\kappa c}{1+2k\kappa}$ which contradicts our choice of $d_q = 2^{-q-1} \frac{\kappa c}{1+2k\kappa}$.

□

Proposition 2.6.6 *Suppose that $U_0 \subset C(K)$ is non-empty and τ open, that f is convex and norm Lipschitz on U_0 and that for each $\varphi \in C(K)$ there is a $\mu \in \partial f(\varphi)$ that has finite support. Then there is a non-empty τ open subset $U \subset U_0$, and a finite set $F \subset K$, such that for every $\varphi \in U$,*

$$\partial f(\varphi) \cap \mathcal{M}(F) \neq \emptyset.$$

PROOF. We may apply Lemma 2.6.5 to find integers $0 \leq j \leq k$, a non-empty τ open subset $U \subset U_0$, and a finite set $F \subset K$ such that

$$\partial f(\varphi) \cap \mathcal{M}_{k,j,F} \neq \emptyset$$

for every $\varphi \in U$. Since $\mathcal{M}_{k,j}(F) \subset \mathcal{M}_k(F)$ then

$$\partial f(\varphi) \cap \mathcal{M}_k(F) \neq \emptyset$$

for every $\varphi \in U$. This ends the proof. □

We complete this section with an example of a function f such that each subdifferential has a measure of finite support. In particular, if \mathbf{V} is an n dimensional subspace of $C(K)$ then the function $d(\psi, \mathbf{V})$, which is defined as the distance of ψ from the subspace \mathbf{V} , has a measure with at most $n + 1$ element support in each subdifferential.

Suppose that $\mathbf{V} = \text{span} \{\varphi_1, \dots, \varphi_n\}$ is an n dimensional subspace of $C(K)$, then we define

$$d(\psi, \mathbf{V}) = \inf_{\lambda_i \in \mathbb{R}} \|\psi - \sum_{i=1}^n \lambda_i \varphi_i\|$$

for all $\psi \in C(K)$.

Lemma 2.6.7 *The continuous convex function $d(\psi, \mathbf{V})$ has a measure with at most $n + 1$ element support in each subdifferential.*

PROOF. For ease of notation let $f(\varphi) = d(\varphi, \mathbf{V})$ where \mathbf{V} is the span of n linearly independent functions $\varphi_1, \dots, \varphi_n$. If $\varphi \notin \mathbf{V}$, find $\psi \in \mathbf{V}$ such that $f(\varphi) = \|\varphi - \psi\|$, and note that $\partial f(\varphi)$ consists of those Radon measures μ that satisfy

1. $\mu(\varphi_i) = 0$ for $i = 1, \dots, n$,
2. $\|\mu\| \leq 1$,
3. μ^+ is concentrated on the set $A = \{x : \varphi(x) - \psi(x) = f(\varphi)\}$, and
4. μ^- is concentrated on the set $B = \{x : \varphi(x) - \psi(x) = -f(\varphi)\}$.

To see this suppose first that (1)-(4) hold. Whenever $\eta \in \mathbf{V}$ and $h \in C(K)$, then $\|\varphi + h - \eta\| \geq \mu(\varphi + h - \eta) = \mu(\varphi + h) = \mu(\varphi) + \mu(h)$, since $\|\mu\| \leq 1$ and $\mu(\eta) = 0$ by (1). By (3), by (4), and using also that $\mu(\eta) = 0$, we have that $\mu(\varphi) = f(\varphi)$. So $\|\varphi + h - \eta\| \geq f(\varphi) + \mu(\varphi)$, and taking infimum over $\eta \in \mathbf{V}$ gives that $f(\varphi + h) \geq f(\varphi) + \mu(\varphi)$, hence $\mu \in \partial f(\varphi)$. Conversely, assume that $\mu \in \partial f(\varphi)$. For every $t \in \mathbf{R}$, $f(\varphi + t\varphi_i) = f(\varphi)$, so $0 = f(\varphi + t\varphi_i) - f(\varphi) \geq t\mu(\varphi_i)$, which shows that $\mu(\varphi_i) = 0$. Since the Lipschitz constant of f is one, $\|\mu\| \leq 1$. To prove (3) and (4), assume first that $|\mu|(K \setminus A \cup B) > 0$. Since A and B are compact, there is $g \in C(K)$ such that $g = 0$ on some open set containing $A \cup B$ and $\mu(g) \neq 0$. Then for $|t|$ sufficiently small $\|\varphi - \psi\| = \|(\varphi + tg) - \psi\|$, so that $f(\varphi) \geq f(\varphi + tg) \geq f(\varphi) + t\mu(g)$, which is impossible. Assume next that $\mu^-(A) > 0$. Find an open set G such that $\varphi(x) - \psi(x) > 0$ for $x \in G$ and $\mu^-(A \cap G) > 0$, the latter fact may be used to find $g \in C(K)$, $g \geq 0$, $\text{spt}(g) \subset G$, and $\mu(g) < 0$. For sufficiently small $t > 0$ we then set $\|\varphi - tg - \psi\| \leq \|\varphi - \psi\|$, so $f(\varphi) \geq f(\varphi - tg) \geq f(\varphi) - t\mu(g) > f(\varphi)$. A similar contradiction is obtained in the case when $\mu^+(B) > 0$.

Let $\mu \in \partial f(\varphi)$ be extreme. If μ is not a combination of $n + 1$ Dirac measures, then there are disjoint Borel sets $S_1, \dots, S_k \subset A$, $S_{k+1}, \dots, S_{n+2} \subset B$ covering $A \cup B$ such that $\mu(S_i) \neq 0$ for all $i = 1, \dots, n + 2$. Consider the system of n linear equations

$$\sum_{i=1}^{n+2} \alpha_i \int_{S_i} \varphi_j d\mu = 0 \quad (j = 1, \dots, n)$$

together with the equation

$$\sum_{i=1}^{n+2} \alpha_i \mu(S_i) = 0.$$

This is a system of $n + 1$ linear equations for $n + 2$ unknowns, so it has a non-trivial solution $\alpha_1, \dots, \alpha_{n+2}$. If $|t| > 0$ is sufficiently small, we infer that the measures

$$\nu_t(E) = \sum_{i=1}^{n+2} (1 + t\alpha_i) \mu(E \cap S_i)$$

verify (1)–(4), so they belong to $\partial f(\varphi)$. Since they are different from μ (if $\alpha_i \neq 0$, then $\nu_t(S_i) \neq \mu(S_i)$), and $\mu = (\nu_t + \nu_{-t})/2$, we conclude that μ is not an extreme point of $\partial f(\varphi)$, which is a contradiction since $\partial f(\varphi)$ is weak* compact, it has extreme points. Hence $\partial f(\varphi)$ contains a measure with at most $n + 1$ element support. \square

2.7 Gâteaux differentiability

We use the results of section (2.6) to obtain the following differentiability result for continuous convex functions on $C(K)$.

Theorem 2.7.1 *Let f be a continuous convex function on $C(K)$ such that each subdifferential $\partial f(\varphi)$ contains a measure with finite support. Then f is Gâteaux differentiable on a τ_o -residual set.*

PROOF. Our method of proof is to use the notion of ϵ -Gâteaux differentiability (as defined below the directional derivatives form, within ϵ , a linear mapping) and we show that, for all $\epsilon > 0$, the function f is ϵ -Gâteaux differentiable on a τ_o -dense open set V_ϵ . Consequently f is Gâteaux differentiable on the τ_o -residual set $\bigcap_{n=1}^{\infty} V_{1/n}$. We use Lemma 2.6.6 for each non-empty τ_o -open set U to find a non-empty τ_o -open set $V \subset U$ such that for all $\varphi \in V$, f depends only on a finite set of values $\varphi(x_1), \dots, \varphi(x_k)$. From this, and using generic Fréchet differentiability of convex functions on \mathbf{R}^k we deduce ϵ -Gâteaux differentiability of f on a non-empty τ_o -open subset of V . This will finish the proof, since it follows that the τ_o -interior of the set of points of ϵ -Gâteaux differentiability is dense in $C(K)$. (Otherwise the above applied to the complement of its closure would give a contradiction.)

We say that a function g is ϵ -Gâteaux differentiable on \mathcal{U} if there is $g' \in C(K)^*$ such that for all $\phi \in U$ and all ψ we have

$$\left| \lim_{t \rightarrow 0} \frac{g(\phi + t\psi) - g(\phi)}{t} - \langle g', \psi \rangle \right| \leq \epsilon \|\psi\|.$$

Let U be τ_o -open. Applying Lemma 2.6.6, there are a non-empty τ_o -open subset $V \subset U$, a finite set $F \subset K$, and a non-negative integer n such that

$$\partial f(\varphi) \cap \mathcal{M}(F) \neq \emptyset \quad (2.6)$$

for every $\varphi \in V$. We may suppose that $V = U_w(\hat{\varphi}_1, \kappa, \{\eta_x : x \in F\})$ where each η_x is non-negative, with exactly one zero at x , and $F = \{x_1, \dots, x_k\}$. We

claim that for all $\varphi \in V$, $f(\varphi)$ depends only on $\varphi(x_1), \dots, \varphi(x_k)$. Supposing that the claim is not true, there are functions $\varphi_1 \in V$ and $\varphi_2 \in V$ such that $(\varphi_1(x_1), \dots, \varphi_1(x_k)) = (\varphi_2(x_1), \dots, \varphi_2(x_k))$ and

$$f(\varphi_1) \neq f(\varphi_2). \quad (2.7)$$

From (2.6) there are $b_1, \dots, b_k \in \mathbf{R}$ such that

$$f(\xi) \geq f(\varphi_1) + b_1(\xi(x_1) - \varphi_1(x_1)) + \dots + b_k(\xi(x_k) - \varphi_1(x_k)) \quad (2.8)$$

for all ξ . Similarly there are $b'_1, \dots, b'_k \in \mathbf{R}$ such that

$$f(\xi) \geq f(\varphi_2) + b'_1(\xi(x_1) - \varphi_2(x_1)) + \dots + b'_k(\xi(x_k) - \varphi_2(x_k)) \quad (2.9)$$

for all ξ . Substituting $\xi = \varphi_2$ in (2.8) we have $f(\varphi_1) \leq f(\varphi_2)$. Substituting $\xi = \varphi_1$ in (2.9) we have $f(\varphi_1) \geq f(\varphi_2)$ and we deduce that $f(\varphi_1) = f(\varphi_2)$ which contradicts (2.7).

Next we find an ϵ -Gâteaux derivative for f on a nonempty subset of V . We define a subset A of \mathbf{R}^k and a function g on A as follows:

$$A = \{(y_1, \dots, y_k) \in \mathbf{R}^k : \text{there is } \varphi \in V \text{ such that} \\ (y_1, \dots, y_k) = (\varphi(x_1), \dots, \varphi(x_k))\}.$$

Defining $h : V \rightarrow A$ by $h(\varphi) = (\varphi(x_1), \dots, \varphi(x_k))$, and noting that $f(\varphi)$ depends on $\varphi(x_1), \dots, \varphi(x_k)$, we define g on A by $g(y_1, \dots, y_k) = f(\varphi)$ for any $\varphi \in h^{-1}(y_1, \dots, y_k)$. We show that A contains an open ball B_0 . We denote the standard basis in \mathbf{R}^k by $\{e_i\}_{1 \leq i \leq k}$. Let $y \in A$ and $g(y) = f(\hat{\phi})$. We claim that for each i there is a positive constant c_i and $\beta_i \in C(K)$ such that $\hat{\phi} + \beta_i \in V$ and $h(\hat{\phi} + \beta_i) = y + c_i e_i \in A$. Noting that V is convex and h is linear, we deduce from this claim that A contains the convex hull of the $k+1$ affinely independent vectors $y, y + c_i e_i$, hence A contains an open ball.

To prove the claim, using Lemma 2.2.4, let $\gamma_i \in C(K)$ be such that $0 \leq \gamma_i \leq 1$, $\gamma_i(x_j) = 0$ for $j \neq i$, and $\gamma_i(x) = 1$ for all x belonging to an open neighbourhood G_i of x_i . Then $\beta_i = c_i \gamma_i$ has the required property provided that $c_i > 0$, that

- $c_i < \kappa$ since this gives $\|(\hat{\phi} + \beta_i) - \hat{\phi}\| < \kappa$, and
- $c_i < \inf_{s \in K \setminus G_i} \eta_{x_i}(s)$

since this gives $|(\hat{\phi} + \beta_i - \hat{\phi})(s) - (\hat{\phi} + \beta_i - \hat{\phi})(x_i)| = |\beta_i(s) - \beta_i(x_i)| \leq c < \eta_{x_i}(s)$ for $s \notin G_i$ and since this inequality is obvious if $s \in G_i$.

We also note that g is convex on A . Indeed, if $y = h(\varphi)$, and $z = h(\psi)$, and $\varphi, \psi \in V$, then for all $0 \leq t \leq 1$ we have $t\varphi + (1-t)\psi \in V$ and $h(t\varphi + (1-t)\psi) = ty + (1-t)z$. So $g(ty + (1-t)z) = f(t\varphi + (1-t)\psi) \leq tf(\varphi) + (1-t)f(\psi) = tg(y) + (1-t)g(z)$.

Since a convex function is differentiable a.e in \mathbf{R}^k we obtain a Fréchet derivative g'_a of g at some point $\mathbf{a} = (a_1, \dots, a_k) \in B_0$. Given $\epsilon > 0$ there is a $\delta > 0$ such that

$$\left| \frac{g(\mathbf{y}) - g(\mathbf{x}) - g'_a(\mathbf{y} - \mathbf{x})}{\|\mathbf{y} - \mathbf{x}\|} \right| < \epsilon$$

for all $\mathbf{y}, \mathbf{x} \in B = B(\mathbf{a}, \delta) \subset B_0$. Therefore

$$|f(\psi) - f(\varphi) - g'_a(h(\psi) - h(\varphi))| \leq \epsilon \|h(\psi) - h(\varphi)\|$$

for all $\psi, \varphi \in h^{-1}(B)$.

Let $\psi \in C(K)$ be constant on a neighbourhood of each $x \in F$. Since $h^{-1}(B)$ is a norm relatively open subset of V , for any $\varphi \in h^{-1}(B)$ there is a $\delta > 0$ such that $\varphi + t\psi \in h^{-1}(B)$ whenever $|t| < \delta$; for such t we get, using linearity of h , that

$$|f(\varphi + t\psi) - f(\varphi) - tg'_a(h(\psi))| \leq \epsilon |t| \|h(\psi)\| \leq \epsilon |t| \|\psi\|.$$

Hence

$$\lim_{t \rightarrow 0} \left| \frac{f(\varphi + t\psi) - f(\varphi)}{t} - g'_a(h(\psi)) \right| < \epsilon \|\psi\|$$

for all such ψ . Since the set of all such ψ is norm dense in $C(K)$ and since f is locally Lipschitz, this inequality extends to all $\psi \in C(K)$. Thus $g'_a \circ h \in C(K)^*$ is an ϵ -Gâteaux derivative of f at $\varphi \in h^{-1}(B)$. (it does not depend on φ). Therefore f is ϵ -Gâteaux differentiable on $h^{-1}(B)$ which is non-empty and τ_o -open. This ends the proof. \square

Our final result of this section concerns the double arrow space D (see [Ta2] and [Fa]). The double arrow space is defined as follows. We equip the unit interval of signed points,

$$I_s = \{(\sigma, 1) : \begin{array}{l} x = 0 \text{ and } \sigma = 1, \\ \text{or } 0 < x < 1 \text{ and } \sigma \in \{-1, 1\}, \\ \text{or } x = 1 \text{ and } \sigma = -1 \}, \end{array}$$

with an order topology by defining a basis of neighbourhoods $B^\lambda(x, \sigma)$ where

$$B^\lambda(x, 1) = \{(y, \sigma) : x < y < x + \lambda, \sigma \in \{-1, 1\}, \\ \text{or } y = x \text{ and } \sigma = 1\},$$

and

$$B^\lambda(x, -1) = \{(y, \sigma) : x > y > x - \lambda, \sigma \in \{-1, 1\}, \\ \text{or } y = x \text{ and } \sigma = -1\}.$$

We can identify $C(I_s)$ with the Banach space D of functions on $[0, 1]$ that are right continuous at every $0 \leq x < 1$, left continuous at $x = 1$, and have left limits at every $0 < x < 1$, equipped with the supremum norm as follows.

If $f \in C(I_s)$, we let $g(x) = f(x, 1)$ if $0 \leq x < 1$ and $g(1) = f(1, -1)$. It is easy to see that the function g on $[0, 1]$ has the required properties.

Conversely, if $g \in D$, then we let

$$f(x, 1) = g(x^+) \text{ for } 0 \leq x < 1, \text{ and} \\ f(x, -1) = g(x^-) \text{ for } 0 < x \leq 1.$$

Let $\pi : I_s \rightarrow I = [0, 1]$ be defined by $\pi(x, \sigma) = x$. If $f \in C(I_s)$ and g is defined as above, then the set $\{(x, \sigma) \in I_s : f(x, \sigma) \neq g(\pi(x, \sigma))\}$ is countable. Moreover, g is bounded and continuous except at a countable set; so it is of the first class, and so there is a bounded sequence g_n of continuous functions such that $g_n(x) \rightarrow g(x)$ for every $x \in I$ (see [Ku]). Consequently, $g_n \circ \pi(x, \sigma) \rightarrow f(x, \sigma)$ for all $(x, \sigma) \in I_s$ except possibly a countable set.

We conclude that there is a countable set $S \subset C(I_s)$ such that for every $f \in C(I_s)$ there is a bounded sequence $f_n \in S$ such that $f_n \rightarrow f$ except at a countable set; indeed, it sufficed to take $S = \{h \circ \pi\}$ as h runs through a countable norm dense subset of $C(I)$. We observe that non-atomic measures on I_s are determined by their values on S : if μ, ν are non-atomic and $\mu(h) = \nu(h)$ for $h \in S$ then for every $f \in C(I_s)$ the sequence $f_n \in S$ described above converges to f almost everywhere with respect to μ as well as ν , so $\mu(f) = \lim_{n \rightarrow \infty} \mu(f_n) = \lim_{n \rightarrow \infty} \nu(f_n) = \nu(f)$.

Theorem 2.7.2 *Suppose $f : C(I_s) \rightarrow \mathbf{R}$ is a continuous convex function such that the subdifferential of f contains only non-atomic measures at every ϕ . Then f is Gâteaux differentiable on a τ_0 -residual set.*

PROOF. Let $S \subset C(I_s)$ be the countable set defined before the Theorem. By Lemma 2.3.6 the set R of those $\phi \in C(I_s)$ at which $f'(\phi, \psi)$ exists for all $\psi \in S$ is τ_0 -residual. Therefore if $\phi \in R$ and $\mu \in \partial f(\phi)$ then $\langle \mu, \psi \rangle = f'(\phi, \psi)$ for all $\psi \in S$. But the non-atomic measure μ is uniquely determined by its values on S and therefore the subdifferential $\partial f(\phi)$ contains only one measure. Hence $f'(\phi)$ exists. \square

Chapter 3

Lipschitz Isomorphisms

3.1 Introduction

In this chapter we construct three Lipschitz isomorphisms of ℓ_2 to itself for which the derivative is not an isomorphism. We recall that, if X and Y are Banach spaces, then the Lipschitz constant $\text{Lip}(f)$ of a map $f : X \rightarrow Y$ is defined as $\text{Lip}(f) = \sup_{x,y \in X} \frac{\|f(x)-f(y)\|}{\|x-y\|}$; a map $f : X \rightarrow Y$ is a *Lipschitz isomorphism* provided that it is a bijection and that f and f^{-1} have bounded Lipschitz constants.

All of these constructions have some relevance to the Lipschitz classification of Banach spaces (see[BL]), in particular the linear isomorphism problem for Banach spaces which asks the question:

Given Lipschitz isomorphic Banach spaces X and Y , are they linearly isomorphic?

We may seek a solution to this problem by looking for points $x \in X$ at which the derivative $f'(x)$ exists and is an isomorphism. The authors N. Aronszajn [Ar], P. Mankiewicz [Mn], and J. P. R. Christensen [Cr] have each obtained a Radamacher type theorem (using different notions of null sets):

If f is a Lipschitz map from a separable Banach space X into a space Y with RNP then f is Gâteaux differentiable almost everywhere.

The Radon Nikodym Property (RNP) may be defined by saying that a Banach space Y has the RNP if every Lipschitz map $f : \mathbf{R} \rightarrow Y$ is differentiable everywhere except on a set of Lebesgue measure zero. For an extensive treatment of the Radon Nikodym Property (RNP) we refer to [DU], for the

various notions of null sets we refer to the forthcoming book [BL], and for the equivalence of Gaussian null sets and Aronszajn null sets we refer to [Cs].

Unfortunately none of the presently known notions of null sets that satisfy a Radamacher type theorem are preserved by Lipschitz isomorphisms. For an example of a Lipschitz isomorphism that does not preserve Aronszajn null sets we refer to [Bo] and for an example of a Lipschitz isomorphism that maps a non Haar null set to an Aronszajn null set we refer to [Mt]. If it were possible to find null sets that satisfied a Radamacher type theorem and were preserved by Lipschitz isomorphisms then there would exist a Gâteaux differentiable point x such that f^{-1} is Gâteaux differentiable at $f(x)$, and hence that $f'(x)$ would be a linear isomorphism.

Our examples show that the derivative of a Lipschitz isomorphism f on ℓ_2 to itself may not be surjective at all points where it exists. Example 1 is everywhere Gâteaux differentiable and the Gâteaux derivative of f at zero maps ℓ_2 into a hyperplane. Example 2 is such that the weak limit, $\lim_{t \rightarrow 0} \frac{f(tx)}{t}$, is zero for all $x \in \ell_2$. Example 3 maps a non-Aronszajn null set into an Aronszajn null set; in particular f maps a cube into a hyperplane. This cube, which has empty interior, is not Aronszajn null.

Section 2 contains preliminary work in which we show that the Gâteaux derivative of a Lipschitz isomorphism is a linear isomorphism onto a necessarily closed subspace of Y . Sections 3, 4 and 5 contain the examples.

3.2 A preliminary result

In Theorem 3.2.1 we derive some properties of Lipschitz isomorphisms between Banach spaces X and Y , from which we may conclude that the Gâteaux derivative of a Lipschitz isomorphism f is Lipschitz and is a linear isomorphism onto its range $f'(X)$.

Theorem 3.2.1 *Let X and Y be Banach spaces, let $f : X \rightarrow Y$ be a Lipschitz isomorphism, and let $g : Y \rightarrow X$ be the Lipschitz inverse of f . Suppose that $f'(x) : X \rightarrow Y$ exists at a point $x \in X$ then*

- (i) $\|f'(x)\| \leq \text{Lip}(f)$,
- (ii) $\|f'(x)z\| \geq \|z\|/\text{Lip}(g)$, consequently, $f'(x)$ is a linear isomorphism onto its range and, in particular $\text{Ker}(f'(x)) = \{0\}$,
- (iii) $f'(x)$ has an inverse $h : \text{Im}(f'(x)) \rightarrow X$, such that $h = \hat{g}'(f(x))$, where \hat{g} is the restriction of g to $\text{Im} f'(X)$,

(iv) $\|\hat{g}'(f(x))\| \leq \text{Lip}(g)$, and

(v) $\text{Im}(f'(x))$ is a closed subspace of Y .

PROOF. For (i), estimating the dual norm of $f'(x)$, we obtain

$$\begin{aligned} \frac{\|f'(x)z - f'(x)y\|}{\|z - y\|} &\leq \sup_{\|w\|=1} \|f'(x)w\| \\ &= \sup_{\|w\|=1} \left\| \lim_{t \rightarrow 0} \frac{f(x + tw) - f(x)}{t} \right\| \\ &\leq \text{Lip}(f). \end{aligned}$$

Hence $\|f'\| \leq \text{Lip}(f)$.

For (ii), since $f^{-1} = g : Y \rightarrow X$ is Lipschitz, we have

$$\begin{aligned} \|z\| &= \left\| \frac{g(f(x + tz)) - g(f(x))}{t} \right\| \\ &\leq \text{Lip}(g) \left\| \frac{f(x + tz) - f(x)}{t} \right\|. \end{aligned}$$

Taking the limit as $t \rightarrow 0$, we get that $\|z\| \leq \text{Lip}(g)\|f'(x)z\|$.

For (iii), from (ii) we have $f'(x)$ is injective and so $f'(x)$ has an inverse $h : \text{Im}(f'(x)) \rightarrow X$. We must show that $h = \hat{g}'(f(x))$.

For each $y \in \text{Im } f'(x)$, we may choose $u \in X$ such that $y = f'(x)u$. Given any $\epsilon > 0$, since $f'(x)$ exists, there is a $\delta > 0$ such that for all $t \in (-\delta, \delta)$ there is a point $a(t) \in Y$ such that $\|a(t)\| < \epsilon$ and $f(x + tu) = f(x) + tf'(x)u + ta(t)$. Therefore

$$\begin{aligned} &\left\| \frac{g(f(x) + ty) - g(f(x))}{t} - h(y) \right\| \\ &= \left\| \frac{g(f(x) + tf'(x)u) - g(f(x + tu)) + g(f(x + tu)) - g(f(x))}{t} - hf'(x)u \right\| \\ &= \left\| \frac{g(f(x + tu) - ta(t)) - g(f(x + tu)) + (x + tu) - x}{t} - u \right\| \\ &\leq \text{Lip}(g)\|a(t)\| \\ &\leq \text{Lip}(g)\epsilon. \end{aligned}$$

For (iv) we may apply (i) to g .

Property (v) of f is an immediate consequence of the fact that $f'(x)$ is an isomorphism onto its range. \square

3.3 The first example

Example 1 is an everywhere Gâteaux differentiable Lipschitz isomorphism f of ℓ_2 to itself such that $f'(0)$ is not surjective. The isomorphism $f = \lim_{n \rightarrow \infty} T_n \circ \dots \circ T_1$ is obtained by composing a sequence of Lipschitz isomorphisms T_n where each T_n has the property that, if $\|x\| \leq 3^{n+1}$ then T_n has the same action on the coordinates of x as the cycle $p_n = (1, 2n, 2n+1)$. The result of composing N such cycles is the cycle $(1, 2, \dots, 2N+1)$ and we obtain

$$\lim_{t \rightarrow 0} \frac{f(th) - f(0)}{t} = \lim_{N \rightarrow \infty} \sum_{i=1}^{2N} h_i e_{i+1} + h_{2N+1} e_1 + \sum_{i=2N+2}^{\infty} h_i e_i$$

for all h . Hence $f'(0)h = \sum_{i=1}^{\infty} h_i e_{i+1}$ maps ℓ_2 onto the hyperplane $x_1 = 0$.

Throughout this chapter we make use of Lemma 3.3.1 to estimate the Lipschitz constants of our mappings.

Lemma 3.3.1 *If C is a convex set in a normed linear space X , Y is a metric space, $h : C \rightarrow Y$ is continuous and C can be covered by countably many sets on each of which the Lipschitz constant of h does not exceed L , then $\text{Lip}(h) \leq L$.*

PROOF. It suffices to consider the case when $C = [a, b] \subset \mathbf{R}$ (since to estimate $d(h(x), h(y))$ we consider the line $tx + (1-t)y$ for $0 \leq t \leq 1$ contained in C) and to show that $\text{dist}(h(b), h(a)) \leq L(b-a)$. Suppose that $\text{dist}(h(a), h(b)) > L(b-a)$. Let $[a, b] = \bigcup_{i=1}^{\infty} M_i$, where M_i are sets on which the Lipschitz constant of h does not exceed L . Let $S = \{\sup(M_i) : i = 1, 2, \dots\}$. The function $g(t) = \text{dist}(h(a), h(t)) - L(t-a)$ is continuous on $[a, b]$ and $g(a) = 0 < g(b)$. Using that $g(S)$ is countable, we choose $c \in [g(a), g(b)] \setminus g(S)$ and use the intermediate value theorem to find the last $t \in [a, b]$ such that $g(t) = c$. Whenever $t < s \leq b$, then $g(s) > g(t)$, which gives $\text{dist}(h(s), h(t)) \geq \text{dist}(h(a), h(s)) - \text{dist}(h(a), h(t)) \geq g(s) - g(t) + L(s-t) > L(s-t)$. Finding M_i containing t , we infer that t is the maximum of M_i , so $t \in S$, which contradicts $g(t) = c \notin g(S)$. \square

All of our constructions make use of the rotation map $R_\theta : \ell_2^2 \rightarrow \ell_2^2$ which is defined by $R_\theta(x_1, x_2) = (x_1 \cos \theta - x_2 \sin \theta, x_1 \sin \theta + x_2 \cos \theta)$. We note that as an operator on ℓ_2^2 , $\|R_\theta x\| = \|x\|$ for all x , $R_\theta R_\varphi = R_{\theta+\varphi}$, and $R_\theta^{-1} = R_{-\theta}$.

Lemma 3.3.2 *The rotation map R_θ has the properties*

1. $\|R_\theta - I\| \leq |\theta|$ for all θ and

2. $\|R_\theta - R_\varphi\| \leq |\theta - \varphi|$ for all θ and φ .

PROOF. We identify ℓ_2^2 with the complex numbers \mathbb{C} by setting $(x_1, x_2) = x_1 + ix_2$. Let $x = (x_1, x_2) = (\cos \alpha, \sin \alpha) = e^{i\alpha}$. Then $R_\theta x = e^{i(\alpha+\theta)}$. Now for all θ ,

$$\begin{aligned} |e^{i\theta} - 1| &= |e^{i\theta/2}| |2i \sin(\theta/2)| \\ &\leq 2|\theta/2| = |\theta|. \end{aligned}$$

Hence

$$\begin{aligned} |R_\theta x - x| &= |e^{i\alpha}(e^{i\theta} - 1)| \\ &= |e^{i\theta} - 1| \\ &\leq |\theta| \end{aligned}$$

which proves (1), and (2) follows directly from (1) using $\|R_\theta\| = 1$, $R_\theta R_\varphi = R_{\theta+\varphi}$, and $R_\theta^{-1} = R_{-\theta}$. \square

We use R_θ as follows. Let P and Q be the projections of ℓ_2 defined by $Px = x_1 e_1 + x_2 e_2$ and $Qx = \sum_{i=3}^\infty x_i e_i$. Suppose that

- $\theta : \ell_2 \rightarrow \mathbb{R}$ is Lipschitz, $\text{Lip}(\theta) \leq K$ and $\theta(u) = 0$ if $\|u\| \geq R$, and that
- $\phi : \ell_2 \rightarrow \mathbb{R}$ is Lipschitz, $\text{Lip}(\phi) \leq K$ and $\phi(u) = 0$ if $\|u\| \geq R$.

We identify $x_1 e_1 + x_2 e_2$ with (x_1, x_2) and define maps $T : \ell_2 \rightarrow \ell_2$ and $S : \ell_2 \rightarrow \ell_2$ by

$$T(x) = R_{\theta(x)}(Px) + Qx.$$

and

$$S(x) = R_{\phi(x)}(Px) + Qx.$$

Lemma 3.3.3 $\text{Lip}(T), \text{Lip}(S) \leq KR + 2$.

PROOF. To estimate $\text{Lip}(S)$ choose x, y with $\|x\| \leq \|y\|$. Write $\theta(x) = \theta_1$, $\theta(y) = \theta_2$. Then

$$|\theta_1 - \theta_2| \leq K\|y - x\| \quad \text{by assumption}$$

so that

$$\begin{aligned} \|R_{\theta_2}(Py) - R_{\theta_1}(Px)\| &= \|R_{\theta_2}(Py - Px) + (R_{\theta_2} - R_{\theta_1})(Px)\| \\ &\leq \|y - x\| + |\theta_1 - \theta_2| \|Px\| \\ &\leq (K\|Px\| + 1)\|y - x\| \end{aligned} \tag{3.1}$$

If $\|Px\| \leq R$ then $\|R_{\theta_2}(Py) - R_{\theta_1}(Px)\| \leq (KR + 1)\|y - x\|$ so that $\|\mathcal{S}(x) - \mathcal{S}(y)\| \leq \|R_{\theta_2}(Py) - R_{\theta_1}(Px)\| + \|Qx - Qy\| \leq (KR + 2)\|y - x\|$. If $\|Px\| > R$ then $\mathcal{S}(x) = x$ and $\mathcal{S}(y) = y$. We note that $C = \ell_2 = B(0, R) \cup (\ell_2 \setminus B(0, R))$, the restriction of \mathcal{S} to $B(0, R)$ has Lipschitz constant at most $KR + 2$, and the restriction of \mathcal{S} to $\ell_2 \setminus B(0, R)$ has Lipschitz constant $1 \leq KR + 2$ to get that $\text{Lip}(\mathcal{S}) \leq KR + 2$.

To estimate $\text{Lip}(\mathcal{T})$ we use that $\|u\| \geq \|Pu\|$. If $\|x\| \leq R$, we have $\|R_{\phi_2}(Py) - R_{\phi_1}(Px)\| \leq (K\|Px\| + 1)\|y - x\| \leq (KR + 1)\|y - x\|$ so that $\|\mathcal{T}(x) - \mathcal{T}(y)\| \leq \|R_{\phi_2}(Py) - R_{\phi_1}(Px)\| + \|Qx - Qy\| \leq (KR + 2)\|y - x\|$. If $\|Px\| \geq R$ then $\mathcal{T}(x) = x$ and $\mathcal{T}(y) = y$. We note that since $C = \ell_2 = \{x : \|Px\| \leq R\} \cup (\ell_2 \setminus \{x : \|Px\| \leq R\})$, the restriction of \mathcal{T} to $\{x : \|Px\| \leq R\}$ has Lipschitz constant at most $KR + 2$, and the restriction of \mathcal{T} to $\ell_2 \setminus \{x : \|Px\| \leq R\}$ has Lipschitz constant $1 \leq KR + 2$, we get $\text{Lip}(\mathcal{T}) \leq KR + 2$. \square

In this section and in section 3.4 we apply Lemma 3.3.3 with $\theta : \ell_2 \rightarrow \mathbf{R}$ defined as follows. Let $0 < R_1 < R_2$ and $\lambda : \mathbf{R}^+ \rightarrow \mathbf{R}$ be a Lipschitz function such that

$$\lambda(u) = \begin{cases} \pi/2 & \text{if } u \leq R_1 \\ 0 & \text{if } u \geq R_2, \end{cases}$$

with $\text{Lip}(\lambda) \leq K \leq 2/(R_2 - R_1)$. Let $\theta(x) = \lambda(\|x\|)$ so that $\theta(x) = \theta(y)$ whenever $\|x\| = \|y\|$ and $\text{Lip}(\theta) \leq K$. Then it is easy to see that T has the following properties:

- (i) $\|T(x)\| = \|x\|$ for all x ,
- (ii) $T(x) = x$ for all x such that $\|x\| \geq R_2$,
- (iii) $T(x) = -x_2e_1 + x_1e_2 + Qx$ for all x such that $\|x\| \leq R_1$,
- (iv) $T(x)$ is obtained from x by a rotation about $\{y \in \ell_2 : y_1 = y_2 = 0\}$,
- (v) $T^{-1}(y) = R_{-\theta(y)}(Py) + Qy$, and
- (vi) $\text{Lip}(T), \text{Lip}(T^{-1}) \leq KR_2 + 2$.

For (v), using (i) and that θ is norm invariant, we get that $\theta(T(T^{-1}(y))) = \theta(T^{-1}(y))$ and that $\theta(T(x)) = \theta(x)$. We may then easily verify that $T(T^{-1}(y)) = y$ and $T^{-1}(T(x)) = x$. Finally we get (vi) applying Lemma 3.3.3 to T^{-1} .

Lemma 3.3.4 *There is a constant $K_1 \geq 1$ such that for all 3-cycles p on the natural numbers and all $r > 0$ we can find a Lipschitz isomorphism $T_{p,r}$ on ℓ_2 such that*

- (i) $T_{p,r}(x) = x$ for $\|x\| \geq r$,
- (ii) $T_{p,r}(x) = \sum_{i=1}^{\infty} x_i e_{p(i)}$ for all x such that $\|x\| \leq r/2$,
- (iii) $\|T_{p,r}(x)\| = \|x\|$ for all x ,
- (iv) If $x_i = x_j = x_k = 0$ and $p = (i, j, k)$ then $T_{p,r}(x) = x$,
- (v) $\text{Lip}(T_{p,r}), \text{Lip}(T_{p,r}^{-1}) = K_1 \leq 144$, and
- (vi) $T_{p,r}$ is Fréchet differentiable at all non zero points.

PROOF. We first handle the case when $r = 2$. We recall the Lipschitz isomorphism T of Lemma 3.3.3. Let $\lambda : \mathbf{R}^+ \rightarrow \mathbf{R}$ be defined by

$$\lambda(u) = \begin{cases} \pi/2 & \text{if } u \leq 1 \\ \pi/2 \sin^2(\pi u/2) & \text{if } 1 \leq u \leq 2 \\ 0 & \text{if } u \geq 2. \end{cases}$$

Define $\theta : \ell_2 \rightarrow \mathbf{R}$ by $\theta(x) = \lambda(\|x\|)$; then θ is differentiable at $x \neq 0$ and that θ is Lipschitz with $\text{Lip}(\theta) \leq K = \pi^2/2 \leq 5$. Replacing (x_1, x_2) with (x_i, x_j) we obtain for each $i \neq j$, (with $R_2 = 2$ and $R_1 = 1$) $S_{ij} : \ell_2 \rightarrow \ell_2$, given by

$$S_{i,j}(x) = R_{\theta(x)}(P_{i,j}x) + Q_{i,j}x,$$

where $P_{i,j}x = x_i e_i + x_j e_j$, and $Q_{i,j}x = \sum_{k \neq i,j}^{\infty} x_k e_k$. It is clear from the properties of T that

1. $\|S_{ij}(x)\| = \|x\|$ for all x ,
2. $S_{ij}(x) = x$ for all x such that $\|x\| \geq 2$,
3. $S_{ij}(x) = -x_j e_i + x_i e_j + Q_{i,j}x$ for all x such that $\|x\| \leq 1$,
4. $S_{ij}^{-1}(y) = R_{-\theta(y)}(Py) + Qy$, and
5. $\text{Lip}(S_{ij}) \leq 2K + 2 \leq 12$ (and the same for S_{ij}^{-1}).

We note also that $S_{i,j}$ is Fréchet differentiable at $x \neq 0$, in particular

$$S'_{i,j}(x)h = \lambda'(\|x\|) \frac{\langle x, h \rangle}{\|x\|} R_{\theta(x)-\pi/2}(P_{i,j}x) + R_{\theta(x)}P_{i,j}h + Q_{i,j}h.$$

Let $T_{p,2} = S_{ki} \circ S_{i,j}$. For general $r > 0$ it suffices to let $T_{p,r} = \alpha_r^{-1} \circ T_{p,2} \circ \alpha_r$ where $\alpha_r(x) = 2x/r$ and $\alpha_r^{-1} = rx/2$. Then $\text{Lip}(\alpha_r) = 2/r$ and $\text{Lip}(\alpha_r^{-1}) = r/2$ so that $\text{Lip}(T_{p,r}) = (2K+2)^2 = K_1 = 144$. Similarly for $T_{p,r}^{-1}$, which proves (v). Properties (i)-(iv) follow directly from those of T . Property (vi) follows since $S_{i,j}$ and α_r are Fréchet differentiable at non-zero points. \square

Example 1 is a map f defined by composing the sequence $T_i = T_{p_i, r_i}$ of the Lipschitz isomorphisms of Lemma 3.3.4 where $r_i = 3^{-i}$ and $p_i = (1, 2i, 2i+1)$. We observe that for each $N \geq 1$, $T_N \circ \cdots \circ T_1$ is a Lipschitz isomorphism with Lipschitz constant $K_1 \leq 144$ and has the same action on the coordinates of x as the cycle $(1, 2, 3, \dots, 2N+1)$ whenever $\|x\| \leq 3^{-N-1}$. The map f is defined by $f(x) = \lim_{N \rightarrow \infty} T_N \circ \cdots \circ T_1$ with the result that $f'(0)$ is the shift operator.

Example 1 *There is an everywhere Gâteaux differentiable Lipschitz isomorphism f of ℓ_2 such that $f'(0)$ is not surjective.*

PROOF. For each integer $i \geq 1$ let $r_i = 3^{-i}$ so that we have $r_{i+1} = r_i/3$ for all i , and let p_i be the 3-cycle $(1, 2i, 2i+1)$. Using the Lipschitz isomorphisms $T_{p,r}$ of Lemma 3.3.4 with $r = r_i$ and $p = p_i$, we define the map $f : \ell_2 \mapsto \ell_2$ by

$$f(x) = \lim_{N \rightarrow \infty} T_N \circ T_{N-1} \circ \cdots \circ T_1(x)$$

where $T_i = T_{p_i, r_i}$.

We claim that f is well defined and

1. $f(0) = 0$,
2. $f(x) = T_N \circ \cdots \circ T_1(x)$ if $\|x\| \geq 3^{-N}$,
3. $f^{-1}(x) = \lim_{k \rightarrow \infty} T_1^{-1} \circ \cdots \circ T_k^{-1}(x)$,
4. f is a Lipschitz isomorphism such that $\text{Lip}(f) \leq K_1$ and $\text{Lip}(f)^{-1} \leq K_1$ where K_1 is the constant in Lemma 3.3.4,
5. $f(x) = \sum_{i=1}^{2N} x_i e_{i+1}$ for all x such that $\|x\| \leq 2^{-1}3^{-N}$ and $x \in \text{span}\{e_1, \dots, e_{2N}\}$

6. f is everywhere Gâteaux differentiable, and

7. $f'(0)$ is not surjective.

That f is well defined will follow directly from (1) and (2).

For (1) since $T_{p_N, r_N}(0) = 0$ for all $N \geq 1$ then $f(0) = 0$.

For (2) using (iii) of Lemma 3.3.4 we get that $\|T_N \circ \dots \circ T_1(x)\| = \|x\| \geq 3^{-N}$. By (i) of that Lemma we have that if $\|x\| \geq r_n$ then $T_{p_n, r_n}(x) = x$ so that with $\|x\| \geq 3^{-N}$ we get for all $n \geq N+1$ that $T_n \circ \dots \circ T_N \circ \dots \circ T_1(x) = T_N \circ \dots \circ T_1(x)$. Hence $f(x) = T_N \circ \dots \circ T_1(x)$.

For (3) if $x \neq 0$ using (2) we see that if $r_N = 3^{-N} \leq \|x\|$ then $f^{-1}(x) = T_1^{-1} \circ \dots \circ T_N^{-1}(x)$. The same argument used in (2) gives $f^{-1}(x) = T_1^{-1} \circ \dots \circ T_n^{-1}(x)$ for all $n \geq N+1$. We deduce that for all $x \neq 0$ that $f^{-1}(x) = \lim_{k \rightarrow \infty} T_1^{-1} \circ \dots \circ T_k^{-1}(x)$. If $x = 0$ then $f^{-1}(0) = \lim_{k \rightarrow \infty} T_1^{-1} \circ \dots \circ T_k^{-1}(0) = 0$. Hence

$$f^{-1}(x) = \lim_{k \rightarrow \infty} T_1^{-1} \circ \dots \circ T_k^{-1}(x).$$

for all x .

For (4), using (2), we choose an integer N so that, if not both x and y are zero, 3^{-N} is less than the least non zero of $\|x\|, \|y\|$. Otherwise we let $N = 1$, so that there is some integer N such that

$$\|f(x) - f(y)\| \leq \|T_N \circ \dots \circ T_1(x) - T_N \circ \dots \circ T_1(y)\|. \quad (3.2)$$

To estimate the Lipschitz constant of $T_N \circ \dots \circ T_1$ we need only note that by Lemma 3.3.4 $\text{Lip}(T_k) = K_1$ for each k and T_k has its own region $R_k = \{x : 2^{-1} \cdot 3^{-k} \leq \|x\| \leq 3^{-k}\}$ on which it is non-isometric and is isometric on each of $\{x : 3^{-k} \leq \|x\|\}$ and $\{x : \|x\| \leq 2^{-1} \cdot 3^{-k}\}$. We apply Lemma 3.3.1 with $h = T_N \circ \dots \circ T_1$ and $C = \ell_2 = B(0, 3^{-N-1}) \cup \{x : \|x\| \geq 3^{-1}\} \cup \bigcup_{k=1}^N R_k \cup (S_k \setminus R_k)$ where $S_k = \{x : 3^{-k-1} \leq \|x\| \leq 3^{-k}\}$ for $1 \leq k \leq N$. It is clear (from the properties of T_k and $\text{Lip}(h) = \prod_{i=1}^N \text{Lip}(T_i)$) that

- the restriction of h to R_k has Lipschitz constant at most K_1 for each $1 \leq k \leq N$,
- the restriction of h to $S_k \setminus R_k$ has Lipschitz constant at most 1 for each $1 \leq k \leq N$,
- the restriction of h to $B(0, 3^{-N-1})$ has Lipschitz constant at most 1, and
- the restriction of h to $\{x : \|x\| \geq 3^{-1}\}$ has Lipschitz constant at most 1.

We conclude that $\text{Lip}(T_N \circ \cdots \circ T_1) \leq K_1$, so that, by (3.2), $\text{Lip}(f) \leq K_1$. Similarly $\text{Lip}(f^{-1}) \leq K_1$. Hence f is a Lipschitz isomorphism.

For (5) applying (ii) of Lemma 3.3.4 (that if $\|x\| \leq r/2$ then $T_{p,r}(x) = \sum_{i=1}^{\infty} x_i e_{p(i)}$) we obtain for all x such that $\|x\| \leq r_N/2 = 2^{-1}3^{-N}$ that $T_N \circ \cdots \circ T_1(x) = x_{2N+1}e_1 + \sum_{i=1}^{2N} x_i e_{i+1} + \sum_{i=2N+2}^{\infty} x_i e_i$. If also $x \in \text{span}\{e_i\}_{i=1}^{2N}$ then $T_N \circ \cdots \circ T_1(x) = \sum_{i=1}^{2N} x_i e_{i+1}$. Applying (iv) of Lemma 3.3.4, that for all $k \geq 1$, if $x_1 = x_{2k} = x_{2k+1} = 0$ then $T_k(x) = x$ we see that if $k \geq N+1$ and $x \in \text{span}\{e_i\}_{i=2}^{2N+1}$ then $T_k(x) = x$. Therefore

$$T_k \circ \cdots \circ T_{N+1} \circ T_N \circ \cdots \circ T_1(x) = T_N \circ \cdots \circ T_1(x).$$

whenever $x \in \text{span}\{e_i\}_{i=2}^{2N+1}$. Hence

$$f(x) = \sum_{i=1}^{2N} x_i e_{i+1}$$

whenever $x \in \text{span}\{e_i\}_{i=2}^{2N+1}$ and $\|x\| \leq 3^{-N}$.

For (6), when $x \neq 0$, we apply (2) so that if $\|x\| > 3^{-N}$ then $f(x) = T_N \circ \cdots \circ T_1(x)$. Each T_i is Fréchet differentiable at x so that f is Fréchet and hence Gâteaux differentiable at x . We may calculate the Gâteaux derivative of f at zero. If $h \in \text{span}\{e_i\}_{i=1}^{2N}$ then

$$\begin{aligned} f'(0)h &= \lim_{t \rightarrow 0} \frac{f(th) - f(0)}{t} \\ &= \sum_{i=1}^{2N} h_i e_{i+1} \end{aligned} \tag{3.3}$$

Since (3.3) is true for all N , that $\bigcup_{N=1}^{\infty} \text{span}\{e_i\}_{i=1}^{2N}$ is norm dense in ℓ_2 , and f is Lipschitz, given $\epsilon > 0$ and $h \in \ell_2$, we let $h^{(N)} = (h_1, h_2, \dots, h_N, 0, \dots)$ and choose N so that $\|h - h^{(N)}\| \leq \epsilon$ to get that

$$\begin{aligned} &\left\| \lim_{t \rightarrow 0} \frac{f(th) - f(0)}{t} - (0, h_1, h_2, \dots) \right\| \\ &= \left\| \lim_{t \rightarrow 0} \frac{f(th) - f(th^{(N)}) + f(th^{(N)}) - f(0)}{t} - (0, h_1, h_2, \dots) \right\| \\ &\leq (K_1 + 1)\|h - h^{(N)}\| + \left\| \lim_{t \rightarrow 0} \frac{f(th^{(N)}) - f(0)}{t} - (0, h_1, h_2, \dots, h_N, 0, \dots) \right\| \\ &\leq (K_1 + 1)\epsilon. \end{aligned}$$

Therefore

$$f'(0)h = \sum_{i=1}^{\infty} h_i e_{i+1}$$

for all h . Hence $f'(0)$ exists, it is the shift operator which is not surjective. This completes (6) and (7) and ends the proof. \square

3.4 The second example

Our second example is constructed in a similar way to the first example; the major difference is the choice of coordinates used in the rotation map. At the n 'th stage we choose k and m such that $n = 2^m + k, m \geq 0$, and $0 \leq k < 2^m$ and perform the rotation on the coordinates x_{k+1} and x_{2^m+k+1} .

Example 2 *There is a Lipschitz isomorphism f of ℓ_2 such that*

$$w \lim_{t \rightarrow 0} \frac{f(tx)}{t} = 0$$

for all x .

PROOF. For each integer $n \geq 1$ choose integers k and m such that $n = 2^m + k, m \geq 0$, and $0 \leq k < 2^m$. Define θ as follows. Let $\lambda : \mathbf{R}^+ \rightarrow \mathbf{R}$ be a Lipschitz function such that

$$\lambda(x) = \begin{cases} \pi/2 & \text{if } 0 \leq x \leq 2^{-n-1} \\ 0 & \text{if } x \geq 2^{-n} \end{cases}$$

and $\text{Lip}(\lambda) \leq 2/2^{-n-1} = 2^{n+2}$. Let $\theta(x) = \lambda(\|x\|)$ so that $\text{Lip}(\theta) \leq 2^{n+2} = K_2$ and $\theta(x) = 0$ if $\|x\| \geq 2^{-n} = R_2$. We apply Lemma 3.3.3 replacing coordinates x_1 and x_2 with x_{k+1} and x_{2^m+k+1} respectively, to obtain for each n , Lipschitz isomorphisms $f_n(x) = R_{\theta(x)}(Px) + Qx$ of ℓ_2 to itself such that

- (i) $f_n(0) = 0$,
- (ii) $f_n(x) = x$ for all x such that $\|x\| \geq 2^{-n}$,
- (iii) $f_n(x) = -x_{2^m+k+1}e_{k+1} + x_{k+1}e_{2^m+k+1} + \sum_{i \neq k+1, 2^m+k+1} x_i e_i$ for all x such that $\|x\| \leq 2^{-n-1}$,
- (iv) $f_n(x)$ is obtained from x by a rotation about $\{y \in \ell_2 : y_{k+1} = y_{2^m+k+1} = 0\}$,
- (v) $\|f_n(x)\| = \|x\|$ for all x , and
- (vi) $\text{Lip}(f_n) \leq R_2 K_2 + 2 \leq 2^{-n} 2^{n+2} + 2 \leq 6$ (and the same for f_n^{-1}).

Define $f : \ell_2 \rightarrow \ell_2$ by

$$f(x) = \lim_{N \rightarrow \infty} f_N \circ f_{N-1} \circ \cdots \circ f_1(x).$$

We claim that f is well defined and that

1. $f(0) = 0$,
2. $f(x) = f_N \circ \cdots \circ f_1(x)$ if $\|x\| \geq 2^{-N}$,
3. $f^{-1} = \lim_{k \rightarrow \infty} f_1^{-1} \circ \cdots \circ f_k^{-1}(x)$,
4. f is a Lipschitz isomorphism such that $\text{Lip}(f) \leq 6$ and $\text{Lip}(f)^{-1} \leq 6$,
5. if $u \in \text{span} \{e_i\}_{1 \leq i \leq 2^{p-1}}$ then $f(tu) \in \text{span} \{e_i\}_{i \geq 2^p}$ whenever $|t| < 2^{-2^{p+1}-2}/\|u\|$, and
6. $w \lim_{t \rightarrow 0} \frac{f(tx)}{t} = 0$ for all $x \in \ell_2$.

To verify (1)-(6) of f we proceed as in the first example. That $f(x)$ is well defined will follow from (1) and (2).

For (1) since $f_n(0) = 0$ for all $n \geq 1$ then $f(0) = 0$.

For (2) using (v) of f_n and that $\|x\| \geq 3^{-N}$ we get that $\|f_N \circ \cdots \circ f_1(x)\| = \|x\|$. By (ii) for all $n \geq N+1$ we have $f_n \circ \cdots \circ f_N \circ \cdots \circ f_1(x) = f_N \circ \cdots \circ f_1(x)$. Hence $f(x) = f_N \circ \cdots \circ f_1(x)$.

For (3) if $x \neq 0$, using (2), we see that if $2^{-N-1} \leq \|x\|$ then $f^{-1}(x) = f_1^{-1} \circ \cdots \circ f_N^{-1}(x)$. By (ii), $f_n^{-1}(x) = x$ for $n \geq N+1$, so that $f^{-1} = \lim_{k \rightarrow \infty} f_1^{-1} \circ \cdots \circ f_k^{-1}(x)$. If $x = 0$ then as $f_n^{-1}(0) = 0$ for all n we have $f^{-1}(0) = \lim_{k \rightarrow \infty} f_1^{-1} \circ \cdots \circ f_k^{-1}(0) = 0$.

For (4), using (2), we choose an integer N so that, if not both x and y are zero, 3^{-N} is less than the least non-zero of $\|x\|$ or $\|y\|$, otherwise we let $N = 1$, so that

$$\|f(x) - f(y)\| \leq \|f_N \circ \cdots \circ f_1(x) - f_N \circ \cdots \circ f_1(y)\|.$$

For $\text{Lip}(f_N \circ \cdots \circ f_1)$ we note that, for each $1 \leq n \leq N$, f_n is isometric in each of the sets $\{x \in \ell_2 : \|x\| \leq 2^{-n-1}\}$ and $\{x \in \ell_2 : 2^{-n} \leq \|x\|\}$. The Lipschitz constant of f_n restricted to $U_n = \{x \in \ell_2 : 2^{-n-1} \leq \|x\| \leq 2^{-n}\}$ is at most 6. Letting $F_N = f_N \circ \cdots \circ f_1$, we get that,

- the restriction of F_N to U_n has Lipschitz constant at most $\text{Lip}(f_n) \leq 6$ for each $1 \leq n \leq N$,
- the restriction of F_N to $B(0, 2^{-N-1})$ has Lipschitz constant at most 1, and
- the restriction of F_N to $\{x \in \ell_2 : 2^{-1} \leq \|x\|\}$ has Lipschitz constant at most 1.

We apply Lemma 3.3.1 with $C = \ell_2 = B(0, 2^{-N-1}) \cup \{x \in \ell_2 : 2^{-1} \leq \|x\|\} \cup_{n=1}^N U_n$ from which we deduce that $\text{Lip}(f_N \circ \dots \circ f_1) \leq 6$. Hence $\text{Lip}(f) \leq 6$. A similar estimate applies to $\text{Lip}(f^{-1})$ and we conclude that f is a Lipschitz isomorphism.

For (5) suppose that $u \in \text{span}\{e_i\}_{1 \leq i \leq 2^p-1}$. Since the mappings f_n for $n < 2^p$ can change only first $2^p - 1$ coordinates, all coordinates beyond the first $2^p - 1$ of $f_{2^p-1} \circ \dots \circ f_1(tu)$ are still zero. We may suppose that

$$f_{2^p-1} \circ \dots \circ f_1(tu) = \sum_{i=1}^{2^p-1} x_i e_i$$

for some x_i not necessarily zero. We claim that if $|t|\|u\| \leq 2^{-2^{p+1}-2}$ then $f_{2^{p+1}-1} \circ \dots \circ f_1(tu)$ has zero as the first $2^p - 1$ coordinates as well as all coordinates from 2^{p+1} onwards. By (iii) f_{2^p} interchanges coordinates at e_{2^p+1} and e_1 and since the e_{2^p+1} coordinate of $f_{2^p-1} \circ \dots \circ f_1(tu)$ is zero, we get that

$$f_{2^p} \circ \dots \circ f_1(tu) = x_1 e_{2^p+1} + \sum_{i=2}^{2^p-1} x_i e_i.$$

Repeating this argument another $2^p - 2$ times, we see that

$$f_{2^{p+1}-1} \circ \dots \circ f_1(tu) = \sum_{i=1}^{2^p-1} x_i e_{2^p+i}$$

which proves our claim. Using (iv), we then see by induction that if $m \geq p+1$ then $f_{2^m} \circ \dots \circ f_1(tu)$ can have non-zero coordinates only for indices $2^p \leq j < 2^m$, which shows that the first $2^p - 1$ coordinates of $f(tu)$ are zero. The case $m = p+1$ is as claimed above. If $f_{2^M} \circ \dots \circ f_1(tu)$ has non-zero coordinates only for indices $2^p \leq j < 2^M$ then the coordinates at e_2 and e_{2^M+2} are zero and by (iv) $f_{2^{M+1}} \circ \dots \circ f_1(tu)$ is obtained from $f_{2^M} \circ \dots \circ f_1(tu)$ by a rotation about $\{y \in \ell_2 : y_2 = y_{2^M+2} = 0\}$, so that $f_{2^{M+1}} \circ \dots \circ f_1(tu)$ has non-zero coordinates only for indices $2^p \leq j < 2^M + 1$. We may repeat this argument $2^M - 1$ times so that $f_{2^{M+1}} \circ \dots \circ f_1(tu)$ has non-zero coordinates only for indices $2^p \leq j < 2^{M+1}$. This completes the induction. Hence the first $2^p - 1$ coordinates of $f(tu)$ are zero.

For (6) we have from (5), that for each integer $i \geq 1$, that $\langle \frac{f(tx)}{t}, e_i \rangle = 0$ whenever $|t|\|x\| \leq 2^{-2^{p+1}-2}$ for any integer p such that $2^p \geq i$. Hence

$$w \lim_{t \rightarrow 0} \frac{f(tx)}{t} = 0$$

for all x . This ends the proof. \square

3.5 The third example

It is shown that there is a Lipschitz isomorphism f of ℓ_2 onto itself such that $f(x) = (0, x_1, x_2, \dots)$ whenever $x \in \ell_2$ satisfies $|x_j| \leq 2^{-j}$ for each j .

The construction is a further application of the method used in Example 1 of section 3.3. The map f is defined by composing a sequence of Lipschitz isomorphisms $(g_k)_{k=1}^\infty$. Each g_k is again constructed using a rotation in the plane to achieve the interchange of two coordinates. The major difference here is that, instead of spherical annuli, the non-isometric regions of the isomorphisms g_k consist of a region between sets U_k and $\ell_2 \setminus U_{k-1}$ where each U_k is a cylinder of a set that is the product of a disc and a cube.

Lemma 3.5.1 *For each integer $k \geq 2$, let*

$$U_k = \{x \in \ell_2 : x_1^2 + x_{k+1}^2 \leq 2^{-2k+4} \text{ and } |x_j| \leq 2^{-j+1} + 2^{-k} \text{ for } 2 \leq j \leq k\},$$

and $W_k = \text{span } \{e_i\}_{i \neq 1, k}$. Then for each $k \geq 2$ there is a Lipschitz isomorphism $g_k : \ell_2 \rightarrow \ell_2$ such that

1. $g_k(u) = u$ whenever $\text{dist}(u, U_k) \geq 2^{-k}$, in particular whenever $u = xe_1 + ye_{k+1} + w$ with $w \in W_k$ and $x^2 + y^2 \geq 2^{-2k+6}$,
2. $g_k(xe_1 + ye_{k+1} + w) = ye_1 - xe_{k+1} + w$ whenever $x, y \in \mathbf{R}$ and $w \in W_k$ are such that $xe_1 + ye_{k+1} + w \in U_k$,
3. $g_k(u)$ is obtained from u by a rotation about W_k ,
4. the set U_k is g_k invariant, that is $g_k(U_k) = U_k$,
5. g_k has an inverse,
6. the restriction of g_k to U_k is an isometry,
7. $U_k \supset U_{k+1}$,
8. $g_{k+1}(x) = x$ for every $x \in \ell_2 \setminus U_k$,
9. $\lim_{k \rightarrow \infty} \sup_{x \in \ell_2} \text{dist}(x, \ell_2 \setminus U_k) = 0$, and
10. $\text{Lip}(g_k), \text{Lip}(g_k^{-1}) = L \leq 15$.

PROOF. That g_k is a Lipschitz isomorphism will follow from (5) and (10). For $k \geq 2$ we define g_k in the following way. Define $\theta : \ell_2 \rightarrow \mathbf{R}$ by

$$\theta(u) = \max\{0, 1 - \text{dist}(u, U_k)2^k\}.$$

Note that $\text{Lip}(\text{dist}(u, U_k)) \leq 1$ so that θ is Lipschitz with $\text{Lip}(\theta) = \max\{0, 1, 2^k\} = 2^k$, and if $u = xe_1 + ye_{k+1} + w$ with $w \in W_k$ and $(x^2 + y^2)^{1/2} \geq 2^{-k+3} = R_3$, then $\text{dist}(u, U_k) \geq 2^{-k+3} - 2^{-k+2} \geq 2^{-k}$ which implies that $\theta(u) = 0$. Let $\text{span}\{e_1, e_{k+1}\} = V_k = \ell_2^2$ so that $\ell_2 = V_k \oplus W_k$. We recall the rotation map R_θ of Lemma 3.3.2 and define $g_k : \ell_2 \rightarrow \ell_2$ by

$$g_k(u) = R_{-\pi\theta(u)/2}z + w \text{ if } u = z + w, z \in V_k, w \in W_k.$$

We verify (1)-(10) for g_k .

To see (1), it suffices to note that whenever $u \in \ell_2$ satisfies $\text{dist}(u, U_k) \geq 2^{-k}$ then $\theta(u) = 0$ so that $g_k(u) = u$. As noted above, if $u = xe_1 + ye_{k+1} + w$ with $w \in W_k$ and $x^2 + y^2 \geq 2^{-2k+6}$, then $\text{dist}(u, U_k) \geq 2^{-k}$.

For (2), we observe that $\theta(u) = 1$ for $u \in U_k$, so that, if $u = xe_1 + ye_{k+1} + w$ where $xe_1 + ye_{k+1} \in V_k$ and $w \in W_k$, then $g_k(u) = ye_1 - xe_{k+1} + w$.

For (3), in the definition of g_k , R_θ is a rotation in the plane orthogonal to W_k .

To prove (4), it suffices to note that $g_k(u)$ is obtained from u by a rotation about W_k and to use the rotational invariance of U_k .

For (5), note that the function $u \rightarrow \text{dist}(u, U_k)$ is invariant under rotations about W_k . Since $g_k(u)$ is obtained from u by a rotation about W_k , it follows that $\text{dist}(u, U_k) = \text{dist}(g_k(u), U_k)$, which, according to the definition of θ , implies $\theta(u) = \theta(g_k(u))$. Letting $h_k(u) = R_{\pi\theta(u)/2}z + w$ if $u = z + w, z \in V_k, w \in W_k$ we get that

$$h_k \circ g_k(u) = R_{\pi\theta(g_k(u))/2} \circ R_{-\pi\theta(u)/2}z + w = z + w.$$

Similarly for $g_k \circ h_k$; hence $g_k^{-1} = h_k$.

For (6), if $u_1 = z_1 + w_1$, and $u_2 = z_2 + w_2$ are in U_k where $z_1, z_2 \in V_k$ and $w_1, w_2 \in W_k$ then using property (2), the orthogonality of V_k and W_k , and that $R_{-\pi/2}$ is an isometry, we have $\|g_k(u_1) - g_k(u_2)\|^2 = \|R_{-\pi/2}z_1 - R_{-\pi/2}z_2 + w_1 - w_2\|^2 = \|R_{-\pi/2}(z_1 - z_2) + w_1 - w_2\|^2 = \|R_{-\pi/2}(z_1 - z_2)\|^2 + \|w_1 - w_2\|^2 = \|z_1 - z_2\|^2 + \|w_1 - w_2\|^2 = \|u_1 - u_2\|^2$.

For (7) and (8), we show that, if $x \in \ell_2 \setminus U_k$, then $\text{dist}(x, U_{k+1}) \geq 2^{-k-1}$. This inequality shows that $U_k \supset U_{k+1}$ and, by (1), that $g_{k+1}(x) = x$ for all $x \in \ell_2 \setminus U_k$. If $x \in \ell_2 \setminus U_k$, then $x_1^2 + x_{k+1}^2 > 2^{-2k+4}$ or $|x_j| > 2^{-j+1} + 2^{-k}$ for

some $2 \leq j \leq k$. Let $y \in U_{k+1}$ then $y_1^2 + y_{k+2}^2 \leq 2^{-2k+2}$ and $|y_j| \leq 2^{-j+1} + 2^{-k}$ for each $2 \leq j \leq k+1$. We estimate that

$$y_1^2 + y_{k+1}^2 \leq 2^{-2k+2} + 2^{-2k+1} + 2^{-2k-2} = 5^2 \cdot 2^{-2k-2}$$

so that if $x_1^2 + x_{k+1}^2 > 2^{-2k+4}$ then

$$\|x - y\| > 2^{-k+2} - (5^2 \cdot 2^{-2k-2})^{1/2} = 3 \cdot 2^{-k-1} > 2^{-k-1}.$$

If $|x_j| > 2^{-j+1} + 2^{-k}$ then

$$\|x - y\| \geq |x_j| - |y_j| > 2^{-k} - 2^{-k-1} = 2^{-k-1}$$

so that in both cases $\|x - y\| \geq 2^{-k-1}$, and we infer that $\text{dist}(x, U_{k+1}) \geq 2^{-k-1}$.

For (9) if $x \in U_k$ then $x_1^2 + x_{k+1}^2 \leq 2^{-2k+4}$ so that we may choose any $y \in \ell_2 \setminus U_k$ with $y_1^2 + y_{k+1}^2 > 2^{-2k+4}$ to infer that $\inf_{y \in \ell_2 \setminus U_k} \|x - y\| \leq 2^{-k+2}$. Therefore $\sup_{x \in \ell_2} \inf_{y \in \ell_2 \setminus U_k} \|x - y\| \leq 2^{-k+2}$. Hence

$$\lim_{k \rightarrow \infty} \sup_{x \in \ell_2} \text{dist}(x, \ell_2 \setminus U_k) = 0.$$

Finally for (10) we apply Lemma 3.3.3 with $\phi = -\pi\theta/2$ and x_1, x_k replacing x_1, x_2 , so that, as noted above for θ , we get $\phi(x) = 0$ whenever $\|Px\| = (x_1^2 + x_{k+1}^2)^{1/2} \geq 2^{-k+3} = R_3$, and $\text{Lip}(\phi) = (\pi/2)\text{Lip}(\theta) \leq \pi 2^{k-1} = K_3$. We get for $S = g_k$ that

$$\text{Lip}(g_k) \leq K_3 R_3 + 2 \leq \pi 2^{k-1} 2^{-k+3} + 2 \leq 15.$$

We estimate $\text{Lip}(g_k^{-1})$ similarly. □

Lemma 3.5.2 *Suppose that h_1, \dots, h_n are Lipschitz mappings of a Banach space X onto itself and that for each k there is a set $A_k \subset X$ such that*

1. *the restriction of h_k to A_k has Lipschitz constant at most one,*
2. *$h_k(X \setminus A_k) \subset X \setminus A_{k+1}$ whenever $k < n$, and*
3. *the restriction of h_{k+1} to $h_k(X \setminus A_k)$ has Lipschitz constant at most one whenever $k < n$.*

Then

$$\text{Lip}(h_n \circ \dots \circ h_1) \leq \max(\text{Lip}(h_n), \dots, \text{Lip}(h_1)).$$

PROOF. Let g_0 be the identity. For $1 \leq j \leq n+1$ let $g_j = h_j \circ \dots \circ h_1$, $B_j = X \setminus A_j$

$$M_j = \bigcap_{k=1}^{j-1} g_{k-1}^{-1}(A_k) \cap \bigcap_{k=j}^n g_{k-1}^{-1}(X \setminus A_k) = \bigcap_{k=1}^{j-1} g_{k-1}^{-1}(A_k) \cap g_{j-1}^{-1}(B_j)$$

$$(\text{since } g_{j-1}^{-1}(B_j) \subset g_{j-1}^{-1}(B_{j-1}) \subset \dots \subset B_1)$$

with $\bigcap_{k=1}^0 g_{k-1}^{-1}(A_k) = X$ and $\bigcap_{k=n+1}^n g_{k-1}^{-1}(X \setminus A_k) = X$. The sets M_j , $1 \leq j \leq n+1$, cover X . To see this, let $x \in X$. Suppose that there is a $j \leq n+1$ for which $g_{j-1}(x) \in X \setminus U_j$ and let j be least with this property. We have by the second assumption that $h_k(X \setminus A_k) \subset X \setminus A_{k+1}$ for all $k < n$. Therefore for any $k \geq j$ we get that $g_{k-1}(x) = h_{k-1} \circ \dots \circ h_j \circ g_{j-1}(x) \in X \setminus A_k$. As j is least, whenever $k \leq j-1$, we have $g_{k-1}(x) \in A_k$ and therefore $x \in M_j$. If $j > n+1$ then $g_{j-1}(x) \in A_j$ for all $1 \leq j \leq n+1$ and $x \in M_{n+1}$.

The restriction of g_n to each such M_j is a composition of the restriction of h_1 to A_1 , \dots , h_{j-1} to A_{j-1} , which all have Lipschitz constant at most one according to the first assumption, followed by h_j whose Lipschitz constant we estimate by $\text{Lip}(h_j)$, and followed by the restriction of h_{j+1} to $h_j(X \setminus A_j)$, \dots , h_n to $h_{n-1}(X \setminus A_{n-1})$, which all have Lipschitz constant at most one according to the last assumption. Hence the restriction of g_n to each M_j has Lipschitz constant at most $\max(\text{Lip}(h_n), \dots, \text{Lip}(h_1))$. Since g_n is continuous (it is even Lipschitz), by Lemma 3.3.1 it has Lipschitz constant at most $\max(\text{Lip}(h_n), \dots, \text{Lip}(h_1))$. \square

Example 3 *There is a Lipschitz isomorphism f of ℓ_2 onto itself such that $f(x) = (0, x_1, x_2, \dots)$ whenever $x \in \ell_2$ satisfies $|x_j| \leq 2^{-j}$ for each j .*

PROOF. Let g_1 be the identity of ℓ_2 and $U_1 = \{x \in \ell_2 : x_1^2 + x_2^2 \leq 4\} \supset U_2$. Using the Lipschitz isomorphisms $g_k : \ell_2 \rightarrow \ell_2$ of Lemma 3.5.1 we define f by

$$f(x) = - \lim_{k \rightarrow \infty} g_k \circ \dots \circ g_1(x)$$

for all $x \in \ell_2$. We show that the limit exists for each x , that f is a Lipschitz isomorphism of ℓ_2 onto itself, and that $f(x) = (0, x_1, x_2, \dots)$ whenever $|x_j| \leq 2^{-j}$ for each j .

First let $f_n = g_n \circ \dots \circ g_1$. We apply Lemma 3.5.2 with $h_k = g_k$ and $A_k = U_k$ to infer that $\text{Lip}(f_n) = L \leq 15$; the assumptions of the Lemma are satisfied since from Lemma 3.5.1 with g_1 the identity and $U_1 = \{x \in \ell_2 : x_1^2 + x_{k+1}^2 \leq 4\}$, we get that for all $n+1 > k \geq 1$,

- (by (6)) the restriction of g_k to U_k has Lipschitz constant at most 1,
- (by (8)) the restriction of g_{k+1} to $g_k(\ell_2 \setminus U_k)$ has Lipschitz constant at most 1 whenever $k < n$,

- (by (10)) $\text{Lip}(g_k) \leq L$,
- (4) and (7) imply that $g_k(\ell_2 \setminus U_k) = \ell_2 \setminus U_k \subset \ell_2 \setminus U_{k+1}$ whenever $k < n$, and
- g_k is surjective.

We apply Lemma 3.5.2 again with $h_k = g_{n-k+1}^{-1}$, $A_k = \ell_2 \setminus U_{n-k}$ for $k < n$, h_n the identity, and $A_n = \ell_2 \setminus U_1$ to infer that the Lipschitz constant of $f_n^{-1} = g_1^{-1} \circ \dots \circ g_n^{-1}$ also does not exceed $L \leq 15$; to apply that Lemma we recall from Lemma 3.5.1 that for all $n+1 > k \geq 1$,

- (8) implies that the restriction of $h_k = g_{n-k+1}^{-1}$ to $A_k = \ell_2 \setminus U_{n-k}$ is the identity, so it has Lipschitz constant at most 1,
- (8) implies that the restriction of g_{n-k+1}^{-1} to $\ell_2 \setminus U_{n-k}$ is an isometry so that U_{n-k} is g_{n-k+1}^{-1} invariant, and with (7) we get that $h_k(\ell_2 \setminus A_k) = g_{n-k+1}^{-1}(U_{n-k}) = U_{n-k} \subset U_{n-k-1} = \ell_2 \setminus A_{k+1}$ whenever $k < n$ (for $k = n-1$ we get that $h_{n-1}(\ell_2 \setminus A_{n-1}) = g_2^{-1}(U_1) = U_1 = \ell_2 \setminus A_n$),
- (6) implies that the restriction of $h_{k+1} = g_{n-k}^{-1}$ to $h_k(\ell_2 \setminus A_k) = g_{n-k+1}^{-1}(U_{n-k}) = U_{n-k}$ is an isometry, so it has Lipschitz constant at most 1, whenever $k < n$,
- by (10), $\text{Lip}(h_k) = \text{Lip}(g_{n+k-1}^{-1}) \leq L$, and
- $h_k = g_{n-k+1}^{-1}$ is surjective.

From property (9) of Lemma 3.5.1, given $\varepsilon > 0$, we find n such that for any $x \in \ell_2$ there is $z \in \ell_2 \setminus U_n$ such that $\|z - x\| < \varepsilon$. From (7) and (8) of Lemma 3.5.1 we infer that

$$g_m \circ g_{m-1} \circ \dots \circ g_{n+1}(z) = z$$

for $m > n$. Applying f_m^{-1} , we get that

$$f_n^{-1}(z) = f_m^{-1}(z)$$

and so we estimate that

$$\begin{aligned} \|f_m^{-1}(x) - f_n^{-1}(x)\| &\leq \|f_m^{-1}(x) - f_m^{-1}(z)\| + \|f_m^{-1}(z) - f_n^{-1}(x)\| \\ &= \|f_m^{-1}(x) - f_m^{-1}(z)\| + \|f_n^{-1}(z) - f_n^{-1}(x)\| \\ &\leq 2L\|z - x\| \\ &< 2L\varepsilon, \end{aligned}$$

and we see that the sequence f_n^{-1} is uniformly convergent. Similarly, letting $y = f_n^{-1}(z)$, we have that $f_n(y) = f_m(y)$ for $m > n$, and we estimate

$$\begin{aligned} \|f_m(x) - f_n(x)\| &\leq \|f_m(x) - f_m(y)\| + \|f_m(y) - f_n(x)\| \\ &= \|f_m(x) - f_m(y)\| + \|f_n(y) - f_n(x)\| \\ &\leq 2L\|y - x\| \\ &< 2L^2\varepsilon, \end{aligned}$$

which gives that the sequence f_n is uniformly convergent. Let $g = \lim_{n \rightarrow \infty} f_n$ and $h = \lim_{n \rightarrow \infty} f_n^{-1}$.

We claim that $h = g^{-1}$. Since

$$\|f_n(f_n^{-1}(x)) - g(h(x))\| \leq L\|f_n^{-1}(x) - h(x)\| + \|f_n(h(x)) - g(h(x))\| \rightarrow 0,$$

then $g(h(x)) = x$. Similarly $h(g(x)) = x$. Noting that a pointwise limit of a sequence of functions with uniform bound on their Lipschitz constants is Lipschitz we have that g , and g^{-1} are Lipschitz.

Let $C_1 = \{x \in \ell_2 : |x_j| \leq 2^{-j} \text{ for all } j\}$ and, for $k \geq 2$ let

$$\begin{aligned} C_k &= \{x \in \ell_2 : |x_1| \leq 2^{-k}, |x_j| \leq 2^{-j+1} \text{ for } 2 \leq j \leq k, \\ &\quad \text{and } |x_j| \leq 2^{-j} \text{ for } j > k\}. \end{aligned}$$

Then $C_k \subset U_k$, so the expression for g_k on U_k gives that $g_k(C_k) = C_{k+1}$. We infer that for every $x \in C_1$,

$$g_k \circ \dots \circ g_1(x) = (x_{k+1}, -x_1, -x_2, \dots, -x_k, x_{k+2}, x_{k+3}, \dots),$$

which in the limit as $k \rightarrow \infty$ shows that $f(x) = -g(x) = (0, x_1, x_2, \dots)$. \square

We deduce that $f'(x)$ is the shift operator whenever $|x_j| < 2^{-j}$ for all j . For each integer $n \geq 1$ and any $y \in \text{span} \{e_i\}_{i=1}^n$ we have $f(x + ty) = (0, x_1 + ty_1, x_2 + ty_2, \dots, x_n + ty_n, x_{n+1}, \dots)$ for all sufficiently small $|t|$. Therefore

$$\lim_{t \rightarrow 0} \frac{f(x + ty) - f(x)}{t} = (0, y_1, \dots, y_n, 0, \dots).$$

Since f is Lipschitz and $\bigcup_{n=1}^{\infty} \text{span} \{e_i\}_{i=1}^n$ is norm dense in ℓ_2 we deduce that $f'(x)y = (0, y_1, y_2, \dots)$ for all $y \in \ell_2$. To see this, given $y \in \ell_2$ and $\varepsilon > 0$, for

each integer $N \geq 1$ let $y^{(N)} = (y_1, \dots, y_N, 0, \dots)$ and recall that $\text{Lip}(f) \leq L$. We may choose N sufficiently large that $\|y - y^{(N)}\| < \epsilon$ and estimate that

$$\begin{aligned}
 & \left\| \lim_{t \rightarrow 0} \frac{f(x + ty) - f(x)}{t} - (0, y_1, y_2, \dots) \right\| \\
 = & \left\| \lim_{t \rightarrow 0} \frac{f(x + ty) - f(x + ty^{(N)}) + f(x + ty^{(N)}) - f(x)}{t} - (0, y_1, y_2, \dots) \right\| \\
 \leq & (L + 1)\|y - y^{(N)}\| + \\
 & \left\| \lim_{t \rightarrow 0} \frac{f(x + ty^{(N)}) - f(x)}{t} - (0, y_1, y_2, \dots, y_N, 0, \dots) \right\| \\
 \leq & (L + 1)\epsilon.
 \end{aligned}$$

Bibliography

- [As] Asplund, E.: *Fréchet differentiability of convex functions*, Acta. Math. **121** (1968), 31–47.
- [Ar] Aronszajn, N.: *Differentiability of Lipschitzian mappings between Banach spaces*, Studia Math. **57** (1976), 147–160.
- [BL] Benyamini, Y.; Lindenstrauss, J.: *Geometric nonlinear functional analysis*, A book in preparation.
- [Bo] Bogachev, V. I.: *Some results on differentiable measures*, Math. USSR-Sb., 55:335-349, 1986.
- [BP] Borwein, J. M.; Preiss, D.: *Smooth variational principle with applications to subdifferentiability of convex functions*, Trans. Amer. Math. Soc. **303** (1987), 517–527.
- [Ch] Choquet, G.: *Lectures on Analysis. Volume 1. Integration and Topological Vector Spaces*, W.A.Benjamin, Inc. New York, New York 10016. (1969).
- [CK] Coban, M. M.; Kenderov, P. S.: *Dense Gâteaux Differentiability of the sup-norm in $C(T)$* , C. R. Acad. Bulgare. Sci. **38** (1985), 1603–1604.
- [Cr] Christensen, J. P. R.: *Measure theoretic zero sets in infinitely dimensional spaces and application to differentiability of Lipschitz mappings*, Actes du Deuxieme Colloque d'Analyse Fonctionelle de Bordeaux. (1973), Vol 2, 29–39.
- [Cs] Csörnyei, M.: *Equivalence of null sets*, Israel J. Math
- [Di] Dieudonné, J.: *Treatise on Analysis Volume II*, Academic Press. New York and London (1970).

- [DGZ1] Deville, R.; Godefroy, G.; Zizler, V.: *Une principe variationnel utilisant des fonctions bosse*, C. R. Acad. Sci. Paris. **312 Serie 1** (1991), 281–286.
- [DGZ2] Deville, R.; Godefroy, G.; Zizler, V.: *A smooth variational principle with applications to Hamilton-Jacobi equations in infinite dimensions*, J. Functional Anal. **111** (1993), 197–212.
- [DGZ3] Deville, R.; Godefroy, G.; Zizler, V.: *Smoothness and renormings in Banach spaces*, Pitman Monographs and Surveys in Pure and Applied Mathematics. Longman Scientific & Technical, Harlow, copublished in the United States with John Wiley & Sons, Inc., New York. **64** (1993).
- [DU] Diestel, J.; Uhl, J. J., Jr. : *Vector measures*, Math. Surveys 15, Amer. Math. Soc. (1977).
- [Ek] Ekeland, I.: *Nonconvex minimisation problems*, Bull. Amer. Math. Soc. (New Series) **1** (1979), 443–474.
- [EL] Ekeland, I.; Lebourg, G.: *Generic Fréchet differentiability and perturbed optimization problems in Banach spaces*, Trans. Amer. Math. Soc. **224** (1976), 193–216.
- [Fa] Fabian, M. J.: *Gâteaux Differentiability of Convex Functions and Topology (Weak Asplund Spaces)* Canadian Math. Soc. series of monographs and advanced texts. A Wiley-Interscience publication (1997).
- [FHV] Fabian, M.; Hajek, P.; Vanderwerff, J.: *On smooth variational principles in Banach spaces*, J. Math. Anal. Appl. **197** (1996), 153–172.
- [Fo] Fosgerau, M.: *PhD thesis, Three papers in analysis and general topology* University of London.(1992).
- [Gi] Giles, J. R.: *Convex analysis with application in the differentiation of convex functions*, Research Notes in Mathematics 58, Pitman (Advanced Publishing Program), Boston, Mass.-London 1982.
- [Ha] Haydon, R.: *A counterexample to several questions about scattered compact spaces*, Bull. London Math. Soc. **22** (1990), 261–268.
- [Ku] Kuratowski, K.: *Topology Volume1* Academic Press New York and London. PWN–Polish Scientific Publishers Warszawa. (1966).

- [LW] Leach, E. B.; Whitfield, J. H. M.: *Differentiable norms and rough norms on Banach spaces*, Proc. Amer. Math. Soc. **33** (1972), 120–126.
- [Mn] Mankiewicz, P.: *On the differentiability of Lipschitz mappings in Fréchet spaces*, Studia Math. **45** (1973), 15–29.
- [Mt] Matoušková, E.: *Lipschitz images of Haar null sets*, to appear.
- [Mz] Mazur, S.: *Über konvexe Mengen in linearen normierten Räumen*, Studia Math. **4** (1932), 70–84.
- [Na] Nachbin, L.: *Topology and order*, Van Nostrand Mathematical studies, No. 4. D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto, Ont.-London, (1965).
- [Ox] Oxtoby, J. O.: *The Banach-Mazur game and Banach category theorem*, Contributions to the Theory of Games, vol III, Annals of Math. Studies, Princeton, N. J. **39** (1957), 159–163.
- [Ph1] Phelps, R. R.: *Convex functions, monotone operators, and differentiability*, Second edition. Lecture Notes in Mathematics, 1364. Springer-Verlag, Berlin. (1993).
- [Ph2] Phelps, R. R.: *Gaussian null sets and differentiability of Lipschitz maps on Banach spaces*, Pacific. J. Math. **77** (1978), 523–531.
- [PPN] Preiss, D.; Phelps, R. R.; Namioka, I.: *Smooth Banach spaces, weak Asplund spaces and monotone or usco mappings*, Israel J. Math. **72** (1990), 257–279.
- [Ta1] Talagrand, M.: *Deux exemples de fonctions convexes* C. R. Acad. Sci. Paris. **288** (1979), 461–464.
- [Ta2] Talagrand, M.: *Renormages de quelques $C(K)$* , Israel. J. Math. **54** (1986), 257–279.