RANK-1 SUPPORT FUNCTIONALS AND THE RANK-1 GENERALIZED JACOBIAN, PIECEWISE LINEAR HOMEOMORPHISMS

by

Daniel Ralph

Computer Sciences Technical Report #938

May 1990

RANK-1 SUPPORT FUNCTIONALS AND THE RANK-1 GENERALIZED JACOBIAN, PIECEWISE LINEAR HOMEOMORPHISMS

By

Daniel Ralph

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

(Computer Sciences)

at the

UNIVERSITY OF WISCONSIN – MADISON

1990

Abstract

RANK-1 SUPPORT FUNCTIONALS AND THE RANK-1 GENERALIZED JACOBIAN,

PIECEWISE LINEAR HOMEOMORPHISMS

Daniel Ralph

Under the supervision of Professor Stephen M. Robinson

This research consists of two main topics broadly related as studies of nondifferentiable systems. The first, involving convex and nonsmooth analysis, is aimed at extending the classical Jacobian to nondifferentiable vector functions. The second is on a characterization of the homeomorphisms in a certain class of piecewise linear mappings. This is useful eg. in approximating certain piecewise smooth systems.

For the first, let X, Y be separated, locally convex topological vector spaces over $\mathbb{R}$ with Y semi-reflexive; X^*, Y^* be the respective topological dual spaces of X, Y; and $CL(X, Y)$ be the space of continuous linear mappings from X to Y. We introduce the *rank-1 support functionals* of sets $\Gamma \subset CL(X, Y)$

$$\sigma_\Gamma^1 : X \times Y^* \mapsto \mathbb{R} \cup \{\infty\} : (x, \lambda) \mapsto \sup_{A \in \Gamma} \lambda A x$$

and characterize the extended real functions on $X \times Y^*$ which are rank-1 support functions. The proof uses Hörmander's [Hör] characterization of classical support

functions. As an immediate application we characterize the fans [Iof81, Iof82] which are, up to closed values, spanned by their handles.

Now assume X,Y are normed spaces and let $g : X \to Y$ be Lipschitz near $x_* \in X$. The ***rank-1 generalized Jacobian*** of g, a set valued derivative for g at points where the classical Jacobian may not exist, is studied. We show existence (nonemptiness) of the rank-1 generalized Jacobian

$$\partial^1 g(x_*) \stackrel{\text{def}}{=} \{A \in CL(X,Y) \mid \forall (u,\lambda) \in X \times Y^*,\ \lambda A u \leq (\lambda g)^\circ(x_*; u)\}$$

where $(\lambda g)^\circ(x_*; \cdot)$ is the Clarke generalized directional derivative [Cla] of the real function λg. We actually show that the mapping $(u, \lambda) \mapsto (\lambda g)^\circ(x_*; u)$ is a rank-1 support function of a nonempty set. Existence extends readily to more general spaces. This is a considerable advance on the most general previous existence result in this area, due to Thibault [Thi82]. Some basic properties of the rank-1 generalized Jacobian are explored.

In our second topic we study a class of piecewise linear maps from $\mathbb{R}^n$ to $\mathbb{R}^n$ called *pl-normal* maps. These are the normal maps [Rob90] induced by linear mappings and polyhedral convex sets. Solving such systems is important in many optimization and equilibrium problems. Robinson's [Rob90] homeomorphism theorem characterizes the pl-normal maps that are homeomorphisms, i.e. gives conditions for unique continuous solvability of such systems. We provide a new and shorter proof of this result.

Acknowledgements

I would like to thank my final oral exam committee: professors Stephen M. Robinson, Olvi L. Mangasarian, Michael C. Ferris, Robert R. Meyer and James G. Morris. It has been a pleasure to be associated with them in their rôles as both teachers and colleagues. Special thanks are due to my supervisor Professor Stephen M. Robinson who gave me complete freedom in my research yet provided invaluable guidance and advice. This dissertation has also benefited considerably from his meticulous reading of it.

My debt to my wife Amanda J. Wauchope for her support throughout my doctoral studies is enormous; I can only hope to match it with my gratitude. I must also thank my mother Elvera D. E. Pascoe who, even from my early school days, has always encouraged me in my studies.

The research reported here was sponsored by National Science Foundation under Grant CCR-8801489, and by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant AFOSR-89-0058 and Grant AFOSR-88-0090. The US Government is authorized to reprint and distribute reprints for Government purposes notwithstanding any copyright notation thereon.

Contents

Chapter 1

Introduction

This research consists of two main topics broadly related as studies of nondifferentiable systems. The first, involving convex analysis and nonsmooth analysis of vector valued functions, is aimed at extending the classical Jacobian to nondifferentiable vector functions on general spaces. The second is on a characterization of the homeomorphisms in a certain class of piecewise linear mappings. This is of interest eg. in approximating certain piecewise smooth systems arising from nonlinear programming and economic equilibria problems. Most of the material presented will relate to the former topic (Chapters 1-3); the final chapter is devoted to the latter.

In our first topic we are interested in vector valued functions which are not smooth, i.e. may not have derivatives (Jacobians) at particular points. If we can find a 'reasonable' substitute for the Jacobian when the Jacobian does not exist, we will be in a position to try to emulate some of the successes of calculus and smooth optimization. The primary concern for us is existence: does the substitute Jacobian we define (a set of continuous linear maps called the *rank-1 generalized Jacobian*) actually provide any maps for us to use in place of the (nonexistent) Jacobian? If not, it will be of no use to us.

In finite dimensions, (existence of) the Clarke generalized Jacobian [Cla] has been useful in nonsmooth calculus (eg. the implicit function theorem [Cla, §7.1]) and in nonsmooth optimization, notably in regularity (stability) of constraint systems [Roc85] and sensitivity in parametric programming [CL, Roc85]. Existence alone of the (rank-1) generalized Jacobian in infinite dimensions, which we give in Chapter 3, is the foundation needed before we can even begin to consistently extend these applications of nonsmooth calculus to general spaces. To get existence we need some theory on the properties of sets of continuous linear maps. This theory is set out in Chapter 2.

In Chapter 2, X, Y denote separated, locally convex topological vector spaces over $\mathbb{R}$ with Y semi-reflexive; X^*, Y^* are the respective topological dual spaces of X, Y; and $CL(X, Y)$ is the space of continuous linear mappings from X to Y. Support functionals of nonempty sets $C \subset X^*$

$$\sigma_C : X \to \mathbb{R} \cup \{\infty\} : x \mapsto \sup_{\xi \in C} \xi x$$

are of fundamental importance in convex and nonsmooth analysis [Roc70, Roc74, Cla]. The support functions on X have been characterized in [Hör] as the functions on X taking values in $\mathbb{R} \cup \{\infty\}$ which are lower semicontinuous and sublinear.

We introduce the *rank-1 support functionals* of nonempty sets $\Gamma \subset CL(X, Y)$:

$$\sigma_\Gamma^1 : X \times Y^* \mapsto \mathbb{R} \cup \{\infty\} : (x, \lambda) \mapsto \sup_{A \in \Gamma} \lambda A x$$

The main result of Chapter 2 characterizes the extended real valued functions on $X \times Y^*$ which are rank-1 support functions: these too have properties relating to lower semicontinuity and sublinearity. An immediate application is the characterization of the set valued fans [Iof81, Iof82] which are, up to closed values, 'spanned by their handles'. Of special importance is the fan associated with the *rank-1 generalized Jacobian* described in Chapter 3.

In Chapter 3 we specialize to normed spaces: let X, Y be real normed spaces with Y be reflexive; and $g : X \to Y$ be Lipschitz near $x_* \in X$. The space of bounded linear mappings from X to Y, $BL(X,Y)$, is then identical to $CL(X,Y)$. The main result of Chapter 3 is the existence (nonemptiness) of the so-called rank-1 generalized Jacobian

$$\partial^1 g(x_*) \stackrel{\text{def}}{=} \{A \in BL(X,Y) \mid \forall(u,\lambda) \in (X \times Y^*),\ \lambda A u \leq (\lambda g)^{\circ}(x_*;u)\}$$

where, for $f : X \to \mathbb{R}$ which is Lipschitz near x_*, $f^{\circ}(x_*;\,\cdot)$ is the Clarke generalized directional derivative [Cla]. The rank-1 generalized Jacobian is kind of set valued derivative for g at points where the classical Jacobian may not exist.

The existence proof relies on the characterization of rank-1 support functions shown in Chapter 2. Existence extends readily to metric spaces and, in greatest generality, to separated locally convex topological vector spaces. $\partial^1 g(x_*)$ is a rank-1 (plenary) approximation to the Clarke generalized Jacobian when the latter exists. This is a considerable advance on the most general previous existence result, due to Thibault [Thi82], which required X and Y to be separable Banach spaces with Y reflexive in order to use Haar null sets and Rademacher's theorem [Chr].

Many basic properties of the rank-1 generalized Jacobian are also covered, including its relationship with the Clarke generalized Jacobian, classical derivatives, Sweetser's shields and Ioffe's fans; and some calculus.

In our 4th and final chapter, we study piecewise linear functions from $\mathbb{R}^n$ to $\mathbb{R}^n$ of a certain kind, namely the *pl-normal* maps. These are just the normal maps [Rob90] induced by linear mappings and polyhedral convex sets. Such systems are important in many optimization and equilibrium problems. They arise directly from variational inequalities, or equivalently generalized equations, specified by linear maps and polyhedral sets; and indirectly as approximations to such systems specified by smooth nonlinear functions over polyhedral sets. Robinson's [Rob90] homeomorphism theorem characterizes the pl-normal maps that are homeomorphisms, i.e. gives conditions for unique solvability of such systems. Here we provide a new, shorter proof of this result.

The remainder of the introduction will consist of an outline of the background and results of the related Chapters 2 and 3. We will present the main ideas in familiar settings, i.e. in finite and infinite dimensional normed spaces, without proofs. This is intended to highlight the main ideas of these chapters without the notational and theoretical demands — necessary for a complete investigation — that tend to obscure the development of the material. Section 1.1 will cover material found in Chapter 2 and Section 1.2 will cover material found in Chapter 3.

The following notation will be employed throughout this chapter.

- Let X, Y be normed spaces over $\mathbb{R}$. Often we will use Euclidean spaces $X = \mathbb{R}^n$ and $Y = \mathbb{R}^m$.

- A *form* on X is a linear mapping from X to $\mathbb{R}$; an *operator* from X to Y is a linear mapping from X to Y. The space of all bounded (continuous) operators from X to Y is written $BL(X, Y)$. The *dual space* of X is the space of all bounded forms on X, $BL(X, \mathbb{R})$, and will be denoted by X^*.

  In finite dimensions, we will use the Euclidean inner product $\langle \cdot, \cdot \rangle$ or the transpose operation u^T to consider vectors $u \in \mathbb{R}^n$ as forms on $\mathbb{R}^n$: $\langle u, \cdot \rangle = u^T$. Also, the space of $m \times n$ matrices $\mathbb{R}^{m \times n}$ represents the space of operators from $\mathbb{R}^n$ to $\mathbb{R}^m$.

- Let $f : X \to \mathbb{R}$ and $g : X \to Y$. Recall f is (Gâteaux) differentiable at $x_* \in X$ if there is a bounded form $\nabla f(x_*)$ such that for each $u \in X$,

  $$\lim_{t \to 0} \frac{f(x_* + tu) - f(x_*)}{t} = \nabla f(x_*) u$$

  The derivative $\nabla f(x_*)$ is also called the *gradient* of f at x_*.

Similarly if g is (Gâteaux) differentiable at x_*, its derivative $\nabla g(x_*)$ is a bounded operator from X to Y called the *Jacobian* of g at x_*.

- The set operations of closure and convex hull will be denoted by cl and co respectively.
- The class of all compact, convex, nonempty subsets of X will be denoted by $\mathcal{C}(X)$.
- 'a.e.' will mean 'almost everywhere with respect to Lebesgue measure' (unless another measure is specified).

1.1 Support Functionals for Convex Sets

1.1.1 Convex Sets of Vectors

Convex sets of vectors are of interest in topology [Sch] and mathematical programming [Roc70, Cla], just to mention two areas. The support functions of such sets will provide an important tool for dealing with them.

Definition 1.1 *Let $\emptyset \neq C \subset \mathbb{R}^n$. The support functional of C, denoted σ_C, is given by*

$$\sigma_C(u) \stackrel{\text{def}}{=} \sup_{x \in C} \langle x, u \rangle, \quad \forall u \in \mathbb{R}^n$$

Example 1.2 *Let $n = 1$ and $C = [0, 1]$. Then for $u \in \mathbb{R}$*

$$\sigma_C(u) = u_+ \stackrel{\text{def}}{=} \begin{cases} u & \text{if } u \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Example 1.3 *Let C be the closed unit ball in $\mathbb{R}^n$: $C = \{x \in \mathbb{R}^n | \|x\| \leq 1\}$. Since we are using the 2-norm, $\sigma_C(u) = \|u\|$ for each $u \in \mathbb{R}^n$. In general, σ_C is the dual norm.*

Example 1.4 *[Roc70, Thm. 23.2] If $f : \mathbb{R}^n \to \mathbb{R}$ is convex then it has a directional derivative at any x_* in every direction u, with value*

$$f'(x_*; u) = \sigma_{\partial f(x_*)}(u)$$

Recall that $C \in \mathcal{C}(\mathbb{R}^n)$ iff C is a compact, convex and nonempty subset of $\mathbb{R}^n$. In the 3-dimensional case, the idea of the following result can be traced back to the 1911 paper [Min]. The infinite dimensional result will be quoted later.

Theorem 1.5

a. $C \in \mathcal{C}(\mathbb{R}^n) \iff C = \{x \in \mathbb{R}^n \mid \forall u \in \mathbb{R}^n, \langle x, u\rangle \leq \sigma_C(u)\}$ *and C is nonempty and bounded.*

b. Let $C, D \in \mathcal{C}(\mathbb{R}^n)$. Then

$$\begin{aligned} C \subset D &\iff \sigma_C \leq \sigma_D \\ &\iff \forall u \in \mathbb{R}^n, \ \ \langle C, u\rangle \subset \langle D, u\rangle \end{aligned}$$

where $\langle C, u\rangle \stackrel{\text{def}}{=} \{\langle c, u\rangle \mid c \in C\}$ and similarly for $\langle D, u\rangle$. By symmetry, the statement is true if equality holds throughout in place of the subsets and inequality.

Remark. We know from above that sets in $\mathcal{C}(\mathbb{R}^n)$ can be distinguished from each other by examining their respective actions on n-vectors. This effectively reduces comparisons of sets in $\mathcal{C}(\mathbb{R}^n)$ to comparisons of real intervals.

We have defined the support functions of nonempty sets in $\mathbb{R}^n$, and seen that sets in $\mathcal{C}(\mathbb{R}^n)$ are characterized by their support functions. A related question is: which functions are support functions of sets in $\mathcal{C}(\mathbb{R}^n)$? This was answered in the generality of locally convex topological vector spaces, of which normed spaces are an example, by [Hör]. We quote the finite dimensional version:

Theorem 1.6 *Let* $p : \mathbb{R}^n \to \mathbb{R}$. *Then*

$$\begin{array}{ll} & 1. \quad p \text{ is the support function of a set } C \in \mathcal{C}(\mathbb{R}^n) \\ \iff & \\ & 2a. \quad p \text{ is convex} \\ & 2b. \quad p \text{ is positive homogeneous:} \\ & \qquad \text{for each } u \in \mathbb{R}^n \text{ and } \alpha > 0, \;\; p(\alpha u) = \alpha p(u) \\ & 2c. \quad p \text{ is bounded above: } \sup\{p(u) \mid \|u\| \le 1\} < \infty \end{array}$$

1.1.2 Convex Sets of Matrices

The inner product on $\mathbb{R}^n$ allowed us to treat vectors as forms, and thereby to define support functions of sets in $\mathbb{R}^n$. The matrix analogue of this follows.

Definition 1.7 *Let* $A, B \in \mathbb{R}^{m \times n}$. *Consider* A *as the column vector of length* $m \times n$ *by adjoining the columns of* A *respectively into one long vector:*

$$(A_{11}, A_{21}, \ldots, A_{m1}, A_{12}, \ldots, A_{mn})^T$$

and similarly for B. *The inner product of* A *and* B, *denoted* $\langle A, B \rangle$, *is the standard inner product of the vectors defined from* A, B *respectively.*

Example 1.8 *Let* $m = n = 2$ *and*

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad B = \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix}$$

Then $\langle A, B \rangle = aa' + bb' + cc' + dd'$.

Lemma 1.9 *The inner product on* $\mathbb{R}^{m \times n}$ *satisfies*

$$\langle A, B \rangle = \mathrm{trace}(A^T B)$$

where the trace *of a (square) matrix is the sum of its diagonal entries.*

So applying Theorem 1.5 in the matrix context gives

Corollary 1.10

a. $\Gamma \in \mathcal{C}(\mathrm{IR}^{m\times n}) \iff \Gamma = \{A \in \mathrm{IR}^{m\times n} \mid \forall B \in \mathrm{IR}^{m\times n},\ \langle A, B\rangle \leq \sigma_\Gamma(B)\}$
and Γ *is nonempty and bounded.*

b. Let $\Gamma, \Delta \in \mathcal{C}(\mathrm{IR}^n)$. *Then*

$$\begin{aligned} \Gamma \subset \Delta &\iff \sigma_\Gamma \leq \sigma_\Delta \\ &\iff \forall B \in \mathrm{IR}^{m\times n},\ \langle \Gamma, B\rangle \subset \langle \Delta, B\rangle \end{aligned}$$

where $\langle \Gamma, B\rangle \overset{\text{def}}{=} \{\langle A, B\rangle \mid A \in \Gamma\}$ *and similarly for* $\langle \Delta, B\rangle$. *By symmetry, the statement is true if equality holds throughout in place of the subsets and inequality.*

This kind of result was pointed out in [H-U82, §2].

We will find it convenient to work with an outer approximation to a set $\Gamma \in \mathcal{C}(\mathrm{IR}^{m\times n})$ which is given by considering the action of its support function only on the rank-1 matrices. First we recall the matrices of rank 1.

Lemma 1.11 *For* $B \in \mathrm{IR}^{m\times n}$,

$$B \text{ has rank } 1 \iff B = vu^T \text{ for some } u \in \mathrm{IR}^n, v \in \mathrm{IR}^m$$

In this case, $\langle A, B\rangle = v^T A u$ *for each* $A \in \mathrm{IR}^{m\times n}$.

The general version of our next definition is given as Definition 1.24 (cf. Chapter 2 Definition 2.15).

Definition 1.12 *Let* $\emptyset \neq \Gamma \subset \mathbb{R}^{m\times n}$.

1. *The* rank-1 support functional *of* Γ, σ^1_Γ, *is given by*

$$\sigma^1_\Gamma(u,v) = \sup_{A\in\Gamma} v^T A u, \quad \forall (u,v) \in \mathbb{R}^n \times \mathbb{R}^m$$

2. *The* rank-1 representer *(or closed rank-1 hull) of* Γ *is*

$$\Gamma^1 \overset{\text{def}}{=} \{A \in \mathbb{R}^{m\times n} \mid \forall (u,v) \in \mathbb{R}^n \times \mathbb{R}^m, \;\; v^T A u \leq \sigma^1_\Gamma(u,v)\}$$

If $\Gamma^1 = \Gamma$ *we say* Γ *is a rank-1 representer.*

Example 1.13 *Let* $\Gamma = \{\bar{A}\}$ *for some* $\bar{A} \in \mathbb{R}^{m\times n}$. *Clearly* $\bar{A} \in \Gamma^1$. *Suppose* $A \in \Gamma^1$ *too; then for the unit vectors* $(e_i)_{i=1}^n \subset \mathbb{R}^n, (f_j)_{j=1}^m \subset \mathbb{R}^m$ *we must have*

$$A_{ij} = f_j^T A e_i \leq \sigma^1_\Gamma(e_i, f_j) = f_j^T \bar{A} e_j = \bar{A}_{ij}$$

By taking $-f_j$*'s instead of* f_j*'s we get that* $-A_{ij} \leq -\bar{A}ij$. *Hence* $A = \bar{A}$*; and* $\Gamma^1 = \Gamma$.

Example 1.14 *The interval Jacobian of a differentiable function* g *at* x_* *is an* $m \times n$ *matrix of real intervals* I_{ij},

$$\bar{\nabla} g(x_*) = [I_{ij}]_{m\times n}$$

Each I_{ij} *contains the i-jth partial derivative of* g, $\nabla g(x_*)_{ij} = dg_i(x_*)/dx_j$. *Therefore,* $\nabla g(x_*) \in \bar{\nabla} g(x_*)$. *The interval Jacobian arises naturally when trying to achieve correct error bounds of numerical processes computationally. See [Neu] for example.*

Also

$$\left[\bar{\nabla} g(x_*)\right]^1 = \bar{\nabla} g(x_*)$$

by reasoning similar to that in the previous example.

From the definition, it is clear that the rank-1 representer of a set of matrices contains the original set. Equality does not hold in general, however, as the following counterexample from [Swe77] demonstrates.

Example 1.15

$$\Gamma = \mathrm{cl\,co}\left\{\begin{bmatrix} 1 & 0 \\ -2 & -1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}, \begin{bmatrix} -1 & -2 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 2 \\ 0 & 1 \end{bmatrix}\right\}$$

The set Γ is in fact compact and convex. Note that the identity matrix I is not in Γ because $\mathrm{trace}(A) = 0$ *for each $A \in \Gamma$, whereas* $\mathrm{trace}(I) = 2$.

Nevertheless, we have $u \in \Gamma u$ for each $u \in \mathbb{R}^2$. The case $u = 0$ is trivial, so assume that $u = (u_1, u_2)^T \neq 0$. If $|u_1| \geq |u_2|$ then

$$A = \begin{bmatrix} 1 & 0 \\ 2u_2/u_1 & -1 \end{bmatrix} \in \Gamma$$

with $Au = u$. If $|u_2| \geq |u_1|$ then

$$A = \begin{bmatrix} -1 & 2u_1/u_2 \\ 0 & 1 \end{bmatrix} \in \Gamma$$

with $Au = u$.
We conclude that $I \in \Gamma^1 \backslash \Gamma$.

A summary of the basic properties of the rank-1 representers in $\mathcal{C}(\mathbb{R}^{m\times n})$ follows in Theorem 1.16. Most of these can shown directly using Theorem 1.5, without reference to the inner product on $\mathbb{R}^{m\times n}$, as in [Swe79, Ch IV]. Without boundedness of closed rank-1 representers, however, the situation is more difficult. The general version is quoted later as Theorem 1.25.

Theorem 1.16

a. If $\Gamma \in \mathcal{C}(\mathbb{R}^{m\times n})$ *then* $\Gamma \subset \Gamma^1 \in \mathcal{C}(\mathbb{R}^{m\times n})$ *and* $(\Gamma^1)^1 = \Gamma^1$.

b. $\Delta = \Gamma^1$ *for some* $\Gamma \in \mathcal{C}(\mathbb{R}^{m\times n})$ *iff* $\Delta = \Delta^1$ *and* Δ *is nonempty and bounded.*

c. Let $\Gamma, \Delta \in \mathcal{C}(\mathbb{R}^{m\times n})$.

$$\begin{aligned}
\Gamma^1 \subset \Delta^1 \quad &\iff \quad \forall (u,v) \in \mathbb{R}^n \times \mathbb{R}^m, \;\; \sigma^1_\Gamma(u,v) \le \sigma^1_\Delta(u,v) \\
&\iff \quad \forall (u,v) \in \mathbb{R}^n \times \mathbb{R}^m, \;\; v^T\Gamma u \subset v^T\Delta u \\
&\iff \quad \forall u \in \mathbb{R}^n, \;\; \Gamma u \subset \Delta u \\
&\iff \quad \forall v \in \mathbb{R}^m, \;\; v^T\Gamma \subset v^T\Delta
\end{aligned}$$

The statement remains valid if equality holds throughout in place of the subsets and the inequality.

Aside: [Swe79] has defined the *plenary* sets in $\mathbb{R}^{m\times n}$ as those sets Δ for which $A \in \Delta \iff \forall u,\ Au \in \Delta u$. Similarly we can define the *rank-1* sets as those sets Δ for which $A \in \Delta \iff \forall u, v,\ v^TAu \in v^T\Delta u$. The theorem above says that the class of rank-1 sets in $\mathcal{C}(\mathbb{R}^{m\times n})$ coincides with the class of plenary sets in $\mathcal{C}(\mathbb{R}^{m\times n})$. We note in passing that, generally, a plenary set is rank-1 but not vice versa. See Chapter 2 §2.5 for further details.

The virtue of (closed) rank-1 representers is that their dual description involves only rank-1 matrices, i.e. n- and m-vectors u and v respectively, rather than all matrices. The main theorem on rank-1 representers of $\mathcal{C}(\mathbb{R}^{m\times n})$ characterizes the functionals which are the support functions of rank-1 representers. We will not use this when dealing with the generalized Jacobian in $\mathbb{R}^{m\times n}$, but will need its infinite dimensional analogue for Banach spaces. The result is a corollary of Chapter 2 Theorem 2.21.2.

Theorem 1.17 *Let $P : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, and (u,v) denote a general point of $\mathbb{R}^n \times \mathbb{R}^m$. Then*

1. *P is the rank-1 support function of a set in $\mathcal{C}(\mathbb{R}^{m\times n})$*

$\iff$

2a. *For each (u,v), $P(u,v) = \min\{\sum_i P(u_i,v_i) \mid$
$\forall$ finite sums $\sum_i v_i u_i^T = vu^T, (u_i,v_i)_i \subset \mathbb{R}^n \times \mathbb{R}^m\}$*

2b. *For each (u,v) and $\alpha > 0$, $P(\alpha u, v) = P(u, \alpha v) = \alpha P(u,v)$*

2c. *P is bounded above: $\sup\{P(u,v) \mid \|u\|, \|v\| \leq 1\} < \infty$.*

Compare the hypothesis of this result and Theorem 1.6: the condition 2a here is a convexity requirement, condition 2b relates to positive homogeneity, while 2c is just boundedness again.

1.1.3 Convex Sets of Forms

Here X and Y are considered general normed spaces over $\mathbb{R}$ not necessarily having inner products. Gradients of smooth real functions on X are forms on X so, in anticipation of the generalized gradient, we will be working with support functions of convex sets of forms. The results given here will exactly parallel the results given in the Subsection 1.1.1.

More notation is needed.

- The *dual norm* on the dual space X^* of X is

$$\|\xi\| \stackrel{\text{def}}{=} \sup\{|\xi u| \mid u \in X, \|u\| \leq 1\}$$

  Under this topology the dual space of X^* is denoted X^{**}.

- Y is *reflexive* if it is (isomorphically) equal to its second dual Y^{**}.

- When dealing with convex sets in the dual space X^*, we endow X^* with the weak* topology, also called the $\sigma(X^*, X)$ topology ([Sch]) i.e. the weak topology on X^* determined by X. (See Chapter 2 §2.2 for more on the weak* topology.)

Then $\mathcal{C}(X^*)$ is the class of all convex, weak* compact, nonempty sets in X^*.

Definition 1.18 *Let $\emptyset \neq C \subset X^*$. The support functional of C, denoted σ_C, is given by*

$$\sigma_C(u) \stackrel{\text{def}}{=} \sup_{\xi \in C} \xi u, \quad \forall u \in X$$

Example 1.19 *Let C be the closed unit ball in X^* : $C = \{\xi \in X^* | \|\xi\| \leq 1\}$. Then $\sigma_C(u) = \|u\|$, for each $u \in X$.*

The next two results, due in this generality to [Hör], are the infinite dimensional versions of Theorems 1.5 and 1.6.

Theorem 1.20

a. $C \in \mathcal{C}(X^*) \iff C = \{\xi \in X^* \mid \forall u \in X, \xi u \leq \sigma_C(u)\}$ *is nonempty and bounded.*

b. Let $C, D \in \mathcal{C}(X^*)$. Then

$$\begin{aligned} C \subset D &\iff \sigma_C \leq \sigma_D \\ &\iff \forall u \in X, \ \langle C, u \rangle \subset \langle D, u \rangle \end{aligned}$$

and, by symmetry, the statement is true if equality holds throughout in place of the subsets and inequality.

Theorem 1.21 *Let $p : X \to \mathbb{R}$. Then*

1. *p is the support function of a set $C \in \mathcal{C}(X^*)$*

$\iff$

2a. *p is convex*

2b. *p is positive homogeneous:*
for each $u \in X$ and $\alpha > 0$, $p(\alpha u) = \alpha p(u)$

2c. *p is bounded above: $\sup\{p(u) \mid \|u\| \leq 1\} < \infty$*

1.1.4 Convex Sets of Operators

If g is smooth its Jacobian is a bounded operator from X to Y. To use the results of the previous section on convex sets of operators, and eventually on the generalized Jacobian, we will need a framework in which operators can be used as bounded forms on some other space.

Consider the space of finite rank operators from Y to X.

- $FL(Y,X) \stackrel{\text{def}}{=} \{T \in BL(Y,X) \mid T \text{ has finite rank}\}$. Of special interest are the mappings of rank 1 or 0 in $BL(Y,X)$ given, for $(u,\lambda) \in X \times Y^*$, by

$$u\lambda : Y \to X : y \mapsto u(\lambda y)$$

- The weak* topology on $BL(X,Y)$ is defined as the $\sigma(BL(X,Y), FL(Y,X))$ topology, which is also called the the weak-operator topology. (See Chapter 2 §2.2 for more on the weak* topology, and also Lemma 1.23 below for the duality between $FL(Y,X)$ and $BL(X,Y)$.)

- $\mathcal{C}(BL(X,Y))$ is the family of convex, weak* compact, nonempty sets in $BL(X,Y)$.

The next two lemmas are found in Chapter 2 §2.2.

Lemma 1.22

a. $T \in BL(Y,X)$ *has finite rank* $\iff$ $T = \sum_i u_i\lambda_i$, *a finite sum where* $(u_i,\lambda_i)_i \subset X \times Y^*$.

b. A norm on $FL(Y,X)$ *is given by*

$$\|T\| = \inf\{\sum_i \|u_i\|\|y_i\| \mid \forall \text{ finite sums } \sum_i u_i\lambda_i = T,\ (u_i,\lambda_i)_i \subset X \times Y^*\}$$

for each $T \in FL(Y,X)$.

This space is usually considered (isomorphically) as the tensor product $X \otimes Y^*$ under the *tensor product* or *projective* norm — see [Sch, Ch III §6]. To avoid extra notation we shall not use the tensor product.

Lemma 1.23 *Each operator $A \in BL(X,Y)$ may be treated as a bounded form on $FL(Y,X)$ which maps each $\sum_i u_i\lambda_i \in FL(Y,X)$ to $\sum_i \lambda_i A u_i$. When Y is reflexive, $BL(X,Y)$ is (isomorphic to) the dual space of $FL(Y,X)$.*

The support function of a set in $BL(X,Y)$ is then defined as in Definition 1.18. Again we are especially interested in rank-1 representers of convex sets of operators (cf. Chapter 2 Definition 2.15).

Definition 1.24 *Let $\emptyset \neq \Gamma \subset BL(X,Y)$.*

1. *The* rank-1 support functional *of Γ, σ^1_Γ, is given by*

$$\sigma^1_\Gamma(u,\lambda) = \sup_{A\in\Gamma} \lambda A u, \quad \forall (u,\lambda) \in X \times Y^*$$

2. *The* rank-1 representer *(or w^*-closed rank-1 hull) of Γ is*

$$\Gamma^1 \stackrel{\text{def}}{=} \{A \in BL(X,Y) \mid \forall (u,\lambda) \in X \times Y^*, \quad \lambda A u \leq \sigma^1_\Gamma(u,\lambda)\}$$

   If $\Gamma^1 = \Gamma$ we say Γ is a rank-1 representer.

Just as we needed the weak* topology $\sigma(X^*,X)$ on X^* when dealing with compact sets of bounded forms, here we need to endow $BL(X,Y)$ with the weak* (or weak-operator) topology, $\sigma(BL(X,Y), FL(Y,X))$. Sets in $\mathcal{C}(BL(X,Y))$ will be weak* compact.

The next two results are infinite dimensional versions of Theorems 1.16 and 1.17 respectively. Theorem 1.17 is a corollary of Chapter 2 Theorem 2.21.2.

Theorem 1.25

a. If $\Gamma \in \mathcal{C}(BL(X,Y))$ *then* $\Gamma \subset \Gamma^1 \in \mathcal{C}(BL(X,Y))$ *and* $(\Gamma^1)^1 = \Gamma^1$.

b. $\Delta = \Gamma^1$ *for some* $\Gamma \in \mathcal{C}(BL(X,Y))$ *iff* $\Delta = \Delta^1$ *and* Δ *is nonempty and bounded.*

c. Let $\Gamma, \Delta \in \mathcal{C}(X \times Y^*)$.

$$\begin{aligned}
\Gamma^1 \subset \Delta^1 \quad &\Longleftrightarrow \quad \forall (u,\lambda) \in X \times Y^*, \;\; \sigma^1_\Gamma(u,\lambda) \leq \sigma^1_\Delta(u,\lambda) \\
&\Longleftrightarrow \quad \forall (u,\lambda) \in X \times Y^*, \;\; \lambda\Gamma u \subset \lambda\Delta u \\
&\Longleftrightarrow \quad \forall u \in X, \;\; \Gamma u \subset \Delta u \\
&\Longleftrightarrow \quad \forall \lambda \in Y^*, \;\; \lambda\Gamma \subset \lambda\Delta
\end{aligned}$$

The statement remains valid if equality holds throughout in place of the subsets and the inequality.

Theorem 1.26 *Let* $P : X \times Y^* \to \mathbb{R}$, *and* (u,λ) *denote a general point of* $X \times Y^*$. *Then*

1. *P is the rank-1 support function of a set in* $\mathcal{C}(BL(X,Y))$

$\Longleftrightarrow$

2a. *For each* (u,λ), $P(u,\lambda) = \min\{\sum_i P(u_i,\lambda_i) \mid$
$\forall$ *finite sums* $\sum_i u_i\lambda_i = u\lambda, (u_i,\lambda_i)_i \subset X \times Y^*\}$

2b. *For each* (u,λ) *and* $\alpha > 0$, $P(\alpha u,\lambda) = P(u,\alpha\lambda) = \alpha P(u,\lambda)$

2c. *P is bounded above:* $\sup\{P(u,\lambda) \mid \|u\|, \|\lambda\| \leq 1\} < \infty$

1.2 Nonsmooth Analysis

1.2.1 The Clarke Generalized Gradient

We begin with extensions due to Clarke [Cla] of the classical gradient and Jacobian of smooth functions, in the finite dimensional case: $X = \mathbb{R}^n, Y = \mathbb{R}^m$. We assume that the functions f and g are Lipschitz in a neighborhood of some $x_* \in X$.

The reader may refer to [Cla, Ch 2] for full details. We will, however, work in reverse order to the presentation there (and the review in Chapter 3 §3.3) in which the viewpoint of convex analysis comes first, and the measure theory formulation second.

Definition 1.27 *The* generalized gradient *of* $f : \mathbb{R}^n \to \mathbb{R}$ *at* x_* *is*

$$\partial f(x_*) \stackrel{\text{def}}{=} \operatorname{cl}\operatorname{co}\{\lim \nabla f(x_i) \mid x_i \to x_*, \nabla f(x_i) \textit{ exists}\}$$

Similarly, the generalized Jacobian *of* $g : \mathbb{R}^n \to \mathbb{R}^m$ *at* x_* *is*

$$\partial g(x_*) \stackrel{\text{def}}{=} \operatorname{cl}\operatorname{co}\{\lim \nabla g(x_i) \mid x_i \to x_*, \nabla g(x_i) \textit{ exists}\}$$

The significance of the Lipschitz condition on f and g becomes clear when we have Rademacher's theorem (see [Cla, §2.5]):

Theorem 1.28 *If f (respectively g) is Lipschitz near x_* then it is differentiable a.e. near x_*.*

Since we assume f (g) is Lipschitz near x_*, the generalized gradient (Jacobian) is *nonempty.* It is also closed and convex by definition. Moreover it is bounded because the function gradients (Jacobians) — when these exist — have norms bounded above by the Lipschitz constant of f (g) near x_*. In summary, the generalized gradient (Jacobian) belongs to $\mathcal{C}(\mathbb{R}^n)$ $(\mathcal{C}(\mathbb{R}^{m\times n}))$.

This definition relies on the theory of Lebesgue measure rather than only on the topology of Euclidean space. Therefore it is not an analytic definition.

The main classes of functions motivating the above definition are the smooth functions, and the convex functions.

Example 1.29 *If f (g) is differentiable near x_* and continuously differentiable at x_* then*

$$\partial f(x_*) = \{\nabla f(x_*)\}$$

$$(\partial g(x_*) = \{\nabla g(x_*)\})$$

Example 1.30 *[Cla, Thm. 2.5.1, Prop 2.3.6a] If f is convex then it is Lipschitz near any $x_* \in \mathbb{R}^n$ and the its convex subdifferential at x_* equals the generalized gradient there:*

$$\{x \in \mathbb{R}^n \mid \forall z \in \mathbb{R}^n,\ \langle x, z - x_*\rangle \le f(z) - f(x_*)\} = \partial f(x_*)$$

In optimization we can use the generalized gradient and generalized Jacobian to extend, for example, optimality conditions from the class of smooth problems to the class of problems with nonsmooth data. Nonsmoothness is sometimes inherent in a problem, but often arises naturally from originally smooth problems.

Consider the constrained minimization problem:

$$\text{(P)}\ \min_{x \in X} f(x) \text{ subject to } g(x) \le 0$$

where the vector inequality $g(x) \le 0$ means each component value $g_j(x)$ is nonpositive, $j = 1, \ldots, m$. Other vector inequalities are similarly defined.

Our next two examples were pointed out in [H-U84].

Example 1.31 *Suppose that f and g are C^2 functions.*

Convert (P) to an unconstrained penalty problem: choose $\alpha > 0$ and set $F_\alpha \stackrel{\text{def}}{=} f + (\alpha/2)\|g(\cdot)_+\|^2$, where $[g(x)_+]_j = [g_j(x)]_+$ for each j. It is well known (eg. [Mang, Thm. 2.8]) that if $\alpha_n \uparrow \infty$, x_n minimizes $F_{\alpha n}$ for each n, and x_* is a limit point of the sequence (x_n), then x_* solves (P).

The first order condition for x to minimize F_α is

$$0 = \nabla F_\alpha(x) = \nabla f(x) + \alpha g(x)_+^T \nabla g(x)$$

Although f, g are C^2, F_α is only C^1: ∇F_α is not differentiable at any point where any $(g_j)_+$ is not differentiable. Nevertheless ∇F_α is Lipschitz in a neighborhood of any point.

Example 1.32 *[Mor] Let f be convex. Recall from Example 1.30 that $\partial f(x)$ coincides with the convex subdifferential of f at any $x \in X$.*

Now consider the generalized gradient ∂f as a set valued mapping from $\mathbb{R}^n$ to $\mathbb{R}^n$. The resolvent $(I + \partial f)^{-1}$ is also a set mapping on $\mathbb{R}^n$ given by $x \in (I + \partial f)^{-1}(z) \iff z \in x + \partial f(x)$. Using the fact that (in the convex case) x_* minimizes f iff $0 \in \partial f(x_*)$, it is easy to check that x_* minimizes f iff x_* is a fixed point of the resolvent. Fixed point iteration based on the resolvent mapping is called the proximal point *method.*

Moreau [Mor] actually shows that $(I + \partial f)^{-1}$ is single valued and (globally) Lipschitz, whether or not f is even differentiable. Therefore the resolvent is a candidate for application of the generalized Jacobian.

Example 1.33 *[Roc85, Thm. 8.5] Suppose x_* solves (P), and $J \stackrel{\text{def}}{=} \{j \mid g_j(x_*) = 0\}$. If the constraint qualification*

$$v \in \mathbb{R}^m, v \geq 0 \textit{ and } v_J \neq 0 \implies 0 \notin v^T \partial g(x_*)$$

holds, then there exists a generalized Karush-Kuhn-Tucker point:

$$\exists v \geq 0 \textit{ such that } 0 \in \partial f(x_*) + v^T \partial g(x_*)$$

This is an immediate extension of the existence of KKT points in continuously differentiable nonlinear programming.

It is of interest that the generalized Jacobian $\partial g(x_*)$ never appears independently, but is always premultiplied by some v^T.

Example 1.34 *Consider the nonlinear complementarity problem for a locally Lipschitz (and often smooth) function $F : \mathbb{R}^n \to \mathbb{R}^n$: find $z \in \mathbb{R}^n$ such that*

$$z \geq 0,\ F(z) \geq 0,\ \textit{and } \langle z, F(z) \rangle = 0.$$

An equivalent reformulation of the problem is to find a zero of the locally Lipschitz function $\Phi : \mathbb{R}^n \to \mathbb{R}^n : z \mapsto \min\{z, F(z)\}$, *where the meaning of the minimum is that* $\Phi_j(z) \stackrel{\text{def}}{=} \min\{z_j, F_j(z)\}$. *Another equivalent problem is to find a zero of* $\Psi : \mathbb{R}^n \to \mathbb{R}^n : z \mapsto F(z_+) + z - z_+$, *which is also locally Lipschitz.*

The functions Φ, Ψ *are both candidates for application of the generalized Jacobian.*

To move toward an analytic definition of $\partial f(x_*)$, we need the Clarke generalized directional derivative.

Definition 1.35 *The* generalized directional derivative *of* f *at* x_* *in the direction* $u \in X$ *is*

$$f^\circ(x_*; u) \stackrel{\text{def}}{=} \limsup_{\substack{x \to x_* \\ t \downarrow 0}} \frac{f(x + tu) - f(x)}{t}$$

Note that the function $f^\circ(x_*; \cdot)$ is analytically defined, and resembles the usual definition of a directional derivative. It is important because it turns out to be the support function of the generalized gradient and hence, by Theorem 1.5a, describes the generalized gradient analytically.

Theorem 1.36 *[Cla, Thm 2.5.1] The generalized directional derivative of* f *at* x_* *is the support function of each of the following sets:*

$$G_S \stackrel{\text{def}}{=} \operatorname{cl}\operatorname{co}\{\lim \nabla f(x_i) \,|\, x_i \to x_*,\ x_i \notin \Omega_f \cup S\}$$

where S *is any set of zero Lebesgue measure, and* Ω_f *is the set of points at which* f *is not differentiable. Therefore*

$$\partial f(x_*) = G_S = \{x \in \mathbb{R}^n \mid \forall u \in \mathbb{R}^n,\ \langle x, u\rangle \le f^\circ(x_*; u)\}$$

for each measure zero set S.

The theorem says that the generalized gradient is 'blind' to sets of measure zero.

For more general spaces than $\mathbb{R}^n$, where measure theory may not be available, we are tempted to specify the generalized gradient in terms of the generalized directional derivative as in the above theorem. First we need to demonstrate that $f^\circ(x_*;\cdot)$ is still a support function in the infinite dimensional case. Recall $\mathcal{C}(X^*)$ is the family of convex, weak* compact, nonempty sets in the dual space X^*.

Proposition 1.37 *[Cla, Prop. 2.1.1, Prop 2.1.2] Let X be a Banach space. Then the generalized directional derivative of f at x_* is convex, positive homogeneous and bounded. Hence, by Theorem 1.21, it is the support function of a set in $\mathcal{C}(X^*)$.*

We now proceed with certainty to define the generalized gradient as a non-empty convex and w*-compact set in X^*:

Definition 1.38 *When X is a Banach space, the generalized gradient of f at x_* is given by*

$$\partial f(x_*) \stackrel{\text{def}}{=} \{\xi \in X^* \mid \forall u \in X,\ \xi u \leq f^\circ(x_*; u)\}$$

The generalized gradient under this definition is relatively easy to deal with, often regardless of the dimension of the space X. [Cla, Ch 2] demonstrates a substantial calculus for the generalized gradient including sum and product rules, the mean value theorem etc.

1.2.2 The Rank-1 Generalized Jacobian

We have already seen how to work with convex sets of matrices using support functions. In this section we try to emulate the steps taken by Clarke for real functions to analytically specify the generalized Jacobian of g at x_*. Initially, we again take $X = \mathbb{R}^n, Y = \mathbb{R}^m$.

We note before proceeding that the calculus of the generalized Jacobian is well developed; for example it includes inverse and implicit function theorems — see [Cla, Ch 2, Ch 7].

First, as pointed out in [H-U82], we would like to find a vector analogue to the generalized directional derivative of f. Unfortunately this difficult problem is still unsolved. We have an easier problem if we settle for a 'rank-1 generalized derivative', which maps $(u,v) \in \mathbb{R}^n \times \mathbb{R}^m$ to $(v^T g)^\circ(x_*; u)$ (in infinite dimensions, $(u, \lambda) \in X \times Y^*$ maps to $(\lambda g)^\circ(x_*; u)$ — see Chapter 3 Definition 3.14).

Definition 1.39 *In finite dimensions, the* rank-1 generalized Jacobian *of g at x_* is*

$$\partial^1 g(x_*) \stackrel{\text{def}}{=} \{A \in \mathbb{R}^{m \times n} \mid \forall (u,v) \in \mathbb{R}^n \times \mathbb{R}^m, \;\; v^T A u \leq (v^T g)^\circ(x_*; u)\}$$

In [Swe77] this is called the *plenary* generalized Jacobian.

The following result is well known ([Cla, H-U82, Swe77]) and relies on the existence of the generalized Jacobian, which, in turn, requires Rademacher's theorem.

Theorem 1.40 *In finite dimensions:*

a. $\partial^1 g(x_*) = [\partial g(x_*)]^1$.

b. For each $v \in \mathbb{R}^m$,
$$\begin{aligned} \partial(v^T g)(x_*) &= v^T \partial g(x_*) \\ &= v^T \partial^1 g(x_*). \end{aligned}$$

Example 1.15 showed that the rank-1 representer of a set of operators may strictly contain the original set. Likewise we present a function whose rank-1 generalized Jacobian strictly contains its generalized Jacobian, at a given point, showing that the rank-1 version of the generalized Jacobian is generally not equal to the generalized Jacobian. This example is also due to [Swe77].

Example 1.41 *Let $g : \mathbb{R}^2 \to \mathbb{R}^2$ be the map on points $(x,y)^T$ given by*

$$\begin{array}{ll}
\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} & \text{if } x \geq 0 \\
\begin{bmatrix} 1 & 0 \\ -2 & -1 \end{bmatrix} & \text{if } x \leq 0, y \geq |x| \\
\begin{bmatrix} -1 & -2 \\ 0 & 1 \end{bmatrix} & \text{if } x \leq 0, 0 \leq y \leq |x| \\
\begin{bmatrix} -1 & 2 \\ 0 & 1 \end{bmatrix} & \text{if } x \leq 0, 0 \geq y \geq x \\
\begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix} & \text{if } x \leq 0, y \leq x
\end{array}$$

It is easy to check that g is piecewise linear and continuous, hence Lipschitz; and that $\partial g(0,0)$ is precisely the set Γ given in Example 1.15. Since Γ^1 strictly contains Γ we find that $\partial^1 g(0,0)$ strictly contains $\partial g(0,0)$.

In many instances this deficiency of the rank-1 generalized Jacobian goes unnoticed. Recall the constrained optimization problem (P) as in Example 1.33. From Theorem 1.40b above we see that the constraint qualification and the resulting optimality condition respectively remain unchanged when we replace $\partial g(x_*)$ by $\partial^1 g(x_*)$. In this case the rank-1 generalized Jacobian serves the same purpose as the actual generalized Jacobian.

Following the mold set for the generalized gradient, we would like to show that the rank-1 generalized Jacobian is still a nonempty convex set even in infinite dimensions. Now we consider X and Y as general normed spaces; recall $\mathcal{C}(BL(X,Y))$ is the family of convex, weak* compact, nonempty sets in $BL(X,Y)$. We quote from Chapter 3 Theorem 3.15.1:

Theorem 1.42 *Let X, Y be normed spaces with Y reflexive, and (u,λ) be a general point in $X \times Y^*$. Then the mapping of (u,λ) to $(\lambda g)^\circ(x_*;u)$ satisfies conditions*

2a, b, and c of Theorem 1.21. Hence it is the rank-1 support function of a set in $\mathcal{C}(BL(X,Y))$.

This result may seem surprising given the remark by Hiriart-Urruty [H-U82, §3] that 'the mere question of existence [of the rank-1 generalized Jacobian] is hopeless for X and Y general Banach spaces', even with the requirement of reflexivity on Y.

The theorem also justifies our next, general, definition (see Chapter 3 Definition 3.14).

Definition 1.43 *When X, Y are normed spaces with Y reflexive, the* rank-1 generalized Jacobian *of g at x_* is given by*

$$\partial^1 g(x_*) \stackrel{\text{def}}{=} \{A \in BL(X,Y) \mid \forall (u,\lambda) \in X \times Y^*, \quad \lambda A u \leq (\lambda g)^\circ(x_*;u)\}$$

Thibault [Thi82] has achieved considerable generality in extending the measure theoretic generalized Jacobian to infinite dimensions. The main tool used is a version of Rademacher's theorem for separable Banach spaces, [Chr, Thm. 7.5], where Haar null sets are substituted for sets of zero Lebesgue measure.

Theorem 1.44 *Let X, Y be separable Banach spaces with Y reflexive. If g is Lipschitz near x_* then it is differentiable everywhere near x_* except possibly on a Haar null set.*

Haar null sets retain an important property of sets in $\mathbb{R}^n$ of Lebesgue measure zero: if a Haar null set is deleted from an open set U, the remaining set is dense in U. With this and the previous theorem, we have:

Definition 1.45 *If X, Y are separable Banach spaces with Y reflexive, the generalized Jacobian of g at x_* is*

$$\partial g(x_*) \stackrel{\text{def}}{=} \operatorname{cl}\operatorname{co}\{\lim \nabla g(x_i) : x_i \to x, \nabla g(x_i) \text{ exists}\}$$

This definition has the advantage over $\partial^1 g(x_*)$ of representing g more closely in certain cases, but it is limited to less general spaces than the rank-1 version.

We conclude with [Thi82, Prop 2.3] which shows that when the generalized Jacobian exists in infinite dimensions, its rank-1 representer is still equal to the rank-1 generalized gradient.

Theorem 1.46 *Let X, Y be separable Banach spaces with Y reflexive. Then*

$$\partial^1 g(x_*) = [\partial g(x_*)]^1$$

Chapter 2

Rank-1 Support Functionals of Sets of Operators

2.1 Introduction

Support functions play an important rôle in convex and nonsmooth analysis [Roc70, Roc74, Cla]. If X is, say, a real normed space with continuous dual X^*, the support function σ_C of a nonempty set C in X^* is the possibly infinite valued function given by

$$\sigma_C(x) \stackrel{\text{def}}{=} \sup_{\xi \in C} \xi x, \quad \forall x \in X.$$

So support functions on X are generated by sets of linear forms in X^*. We quote Hörmander's well known characterization [Hör] of the functions on X, with values in $\mathbb{R} \cup \{\infty\}$, which are support functions: they are the functions which are sublinear and lower semicontinuous (see Theorem 2.13.2).

Now suppose Y is another real normed space and $CL(X,Y)$ is the space of continuous linear mappings from X to Y. We consider functions on $X \times Y^*$ which are generated by nonempty sets Γ of $CL(X,Y)$:

$$\sigma^1_\Gamma(x,\lambda) \stackrel{\text{def}}{=} \sup_{A \in \Gamma} \lambda A x.$$

We call this the *rank-1* support function of Γ, as denoted by the superscript 1, because the function domain is essentially restricted to continuous linear mappings from Y to X of rank one or less, that is to mappings of the type $x\lambda : Y \to X : y \mapsto x(\lambda y)$. For reflexive Y the main result of this chapter, Theorem 2.21, characterizes the extended real valued functions on $X \times Y^*$ which are the rank-1 support functions of sets of continuous linear operators from X to Y.

We are motivated by the use of support functions including Hörmander's above result in nonsmooth analysis, particularly in the Clarke generalized gradient [Cla, Ch. 2] of a locally Lipschitz real function on X. With the theory of rank-1 support functions, namely Theorem 2.21, we can extend Clarke's framework to find so-called rank-1 generalized Jacobians of locally Lipschitz functions mapping X to reflexive Y. Exactly this has been carried out in Chapter 3. More generally we consider the *fans* of Ioffe [Iof81, Iof82] which are set valued analogs of sublinear mappings. Again Theorem 2.21 is our main tool, providing a characterization of the closed valued fans from X to reflexive Y which are generated by sets of operators in $CL(X,Y)$. See §2.4 for details.

We will work with separated locally convex topological vector spaces over $\mathbb{R}$, of which real normed spaces are an example. Since technical considerations tend to overwhelm the presentation, we will explain the ideas behind Theorem 2.21 here. To begin, consider a function $p : K \to \mathbb{R} \cup \{\infty\}$, where K is subset of X. Without difficulty we see that p is the restriction to K of a lower semicontinuous, sublinear function iff $p = \bar{p}|_K$, where $\bar{p}$ is the greatest lower semicontinuous, sublinear function on X which lies underneath p. We construct $\bar{p}$ by examining the closed convex hull of the epigraph of p. Together with Hörmander's result, this allows us to exactly specify the extended real functions on K which are restrictions to K of support functions on X. A variant of this result is given later as Theorem 2.14.

To apply these ideas to rank-1 support functionals we will treat $CL(X,Y)$ as the dual of $FL(Y,X)$, where $FL(Y,X)$ consists of finite sums of operators in

$CL(Y,X)$ of rank one or less,

$$\sum x_i\lambda_i, \quad ((x_i,\lambda_i)) \subset X \times Y^*,$$

under an appropriate topology (Definition 2.6, Proposition 2.7). In this scheme, an operator $A \in CL(X,Y)$ acts as a form on $FL(Y,X)$ by mapping each $\sum x_i\lambda_i$ to the scalar $\sum \lambda_i A x_i$. Hence the support function σ_Γ on $FL(Y,X)$ of a nonempty set $\Gamma \subset CL(X,Y)$ is given by

$$\sigma_\Gamma(\sum x_i\lambda_i) \stackrel{\text{def}}{=} \sup_{A\in\Gamma} \sum \lambda_i A x_i.$$

Observe that $\sigma_\Gamma(x\lambda) = \sigma^1_\Gamma(x,\lambda)$ for each (x,λ). So we have, for $p : X \times Y^* \to \mathbb{R} \cup \{\infty\}$,

$$\begin{aligned} & p(x,\lambda) = \sigma^1_\Gamma(x,\lambda), \text{ for some } \Gamma \text{ and each } (x,\lambda) \\ \Longleftrightarrow \quad & p(x,\lambda) = \sigma_\Gamma(x\lambda), \text{ for some } \Gamma \text{ and each } (x,\lambda) \\ \Longleftrightarrow \quad & \text{the mapping } x\lambda \mapsto p(x,\lambda) \text{ on } K \stackrel{\text{def}}{=} \{x\lambda \in FL(Y,X)\} \text{ is} \\ & \text{the restriction to } K \text{ of a support function on } FL(Y,X). \end{aligned}$$

In the previous discussion we therefore only need substitute $FL(Y,X)$ for X and $CL(X,Y)$ for X^*, in order to obtain a characterization of the functions p which are rank-1 support functions on $X \times Y^*$.

The remainder of the chapter is organized as follows.

§2.2 Notation, and topologies on $FL(Y,X)$ such that the dual of $FL(Y,X)$ is $BL(X,Y)$ or $CL(X,Y)$.

§2.3 Rank-1 support functions, including their characterization in Theorem 2.21, and rank-1 representers of sets of operators in $CL(X,Y)$.

§2.4 Application of Theorem 2.21 to characterize the fans that are, up to closed values, spanned by their handles.

§2.5 Discussion of rank-1 and plenary sets.

2.2 Notation and Topology

We formalize our basic notation for use throughout this chapter.

- By a *convex space* we mean a locally convex topological vector space. Let E be a convex space. A *neighborhood* of a point e in E is set whose interior contains e; by a neighborhood (in E) we mean a neighborhood of $0 \in E$. A convex space E has a base of convex neighborhoods (of zero) denoted by $\mathcal{N}(E)$. The family of all bounded subsets of E is written $\mathcal{B}(E)$.

- Assume throughout that E, X, Y, Z denote separated, or Hausdorff, convex spaces over $\mathbb{R}$; and Y is semi-reflexive (see discussion of the bidual, below).

- The linear spaces $\mathcal{F}_i$, $\mathcal{L}_i$ $(i = 1, 2, 3)$ and $\mathcal{F}$, $\mathcal{L}$ to be introduced later (Definition 2.8) will be used without reference in later sections.

- A *form* on E is a linear mapping from E to $\mathbb{R}$; an *operator* from X to Y is a linear mapping from X to Y. The space of all operators A from X to Y is written $L(X, Y)$. The respective spaces of continuous and bounded operators from X to Y are denoted $CL(X, Y)$, and $BL(X, Y)$. The subspace of finite rank operators in $CL(X, Y)$ is $FL(X, Y)$.

- The *algebraic dual space* of E is $L(E, \mathbb{R})$, denoted E'. The *(topological) dual space* of E is $CL(E, \mathbb{R})$, denoted E^*. Unless otherwise specified, E^* is endowed with the strong topology determined by E, or $\beta(E^*, E)$ topology; and is, in this case, called the strong dual of E. $\mathcal{B}(E^*)$ and $\mathcal{E}(E^*)$ are the respective families of strongly bounded and equicontinuous sets in E^*.

  The weak topology on E determined by E^*, or $\sigma(E, E^*)$ topology, is denoted by w-. The weak topology on E^* determined by E, also called the weak* topology, is denoted by w*-.

See the discussion of polar sets below for the standard bases of neighborhoods of the strong and weak* topologies on E^* and the weak topology on E.

- The *bidual* of E is the dual E^{**} of E^*, which we endow with the strong ($\beta(E^{**}, E^*)$) topology. Since E is separated, it has a natural embedding into E^{**} given by $e : E^* \to \mathbb{R} : \phi \mapsto \phi e$ for $e \in E$. E is said to be semi-reflexive if the natural embedding is surjective (maps onto the bidual), and reflexive if the natural embedding is surjective, and is continuous given the strong bidual topology.

- The *polar* of a nonempty set Q in E is given by

$$Q^\circ \stackrel{\text{def}}{=} \{\phi \in E^* \mid \phi e \leq 1, \ \forall e \in Q\}.$$

For example, the family $\{Q^\circ \mid \text{nonempty } Q \in \mathcal{B}(E)\}$ is a strong base of neighborhoods of $0 \in E^*$, and $\{Q^\circ \mid \text{finite, nonempty } Q \in E\}$ is a weak* base of neighborhoods in E^*.

The polar of a nonempty set Φ in E^* may be taken in E^{**}:

$$\Phi^\circ = \{\zeta \in E^{**} \mid \zeta\phi \leq 1, \ \forall \phi \in \Phi\}$$

or in E:

$$\Phi^\circ = \{e \in E \mid \phi e \leq 1, \ \forall \phi \in \Phi\}.$$

For example, the family of sets $\{\Phi^\circ \mid \text{finite, nonempty } \Phi \in E^*\}$, where polars are taken in E, is a base of neighborhoods of the weak topology on E.

- Given $(x, \lambda) \in X \times Y^*$, define the operator $x\lambda \in FL(Y, X)$ by

$$x\lambda : Y \to X : y \mapsto x(\lambda y).$$

We may abuse notation by also considering $x\lambda$ to be a member of $FL(Y^{*\prime}, X)$, by $x\lambda : Y^{*\prime} \to X : \zeta \mapsto x(\zeta\lambda)$, where $Y^{*\prime} \stackrel{\text{def}}{=} (Y^*)'$. This will be clear from the context; in any case both definitions coincide on Y since we consider Y to be a subspace of $Y^{**} \subset Y^{*\prime}$.

For sets $U \subset X$ and $C \subset Y^*$, $UC \stackrel{\text{def}}{=} \{x\lambda \mid (x, \lambda) \in U \times C\}$.

- The set operations of closure and convex hull are denoted by cl and co respectively.

Before proceeding to topologies on $FL(Y, X)$ we will quote two fundamental results for reference in later sections.

Theorem 2.1

1. *If $V \subset E$ is a convex set containing* 0 *then, taking second polars in E,*

$$V^{\circ\circ} = \operatorname{cl} V.$$

2. *Suppose V, W are nonempty convex sets in E. The* support function *of V is*

$$\sigma_V : E^* \to \mathbb{R} \cup \{\infty\} : \phi \mapsto \sup_{e \in V} \phi e$$

and the support function σ_W is similarly defined. Then

$$\operatorname{cl} V \subset \operatorname{cl} U \iff \sigma_V \leq \sigma_W$$

where the inequality is taken pointwise.

Proof The results are both corollaries of the separation of points from convex sets via hyperplanes [Sch, Ch. II §9.2]. Part 1 is actually given as [Sch, Ch. IV §1.5]. □

It is clear that operators in $FL(Y,X)$ of the form $x\lambda$ have rank no greater than 1. The span of $\{x\lambda \mid (x,\lambda) \in X \times Y^*\}$ in $L(Y,X)$ consists of finite rank operators, and according to our first result, equals $FL(Y,X)$.

Lemma 2.2

$$T \in FL(Y,X) \iff T = \sum x_i\lambda_i$$

for some finite set $\{(x_i,\lambda_i)\} \subset X \times Y^*$. *In particular,* T *is a rank-1 operator from* X *to* Y *iff* $T = x\lambda$ *for some* $x \neq 0, \lambda \neq 0$.

Proof Clearly $\sum x_i\lambda_i \in FL(Y,X)$ for each finite set $\{(x_i,\lambda_i)\} \subset X \times Y^*$, so we only need show that every member T of $FL(Y,X)$ has the form of a finite sum.

Let $m \in \mathbb{N}$ be the rank of T, and $\{x_i\}_1^m$ be a basis of $\{Ty|y \in Y\}$. According to [Sch, Ch. II §4.2 Cor.], there are $\xi_1,\ldots,\xi_m \in X^*$ such that $\xi_i x_j$ equals 1 if $i = j$ and 0 otherwise. Note each $\lambda_i \stackrel{\text{def}}{=} \xi_i T$ is a member of Y^*. We get $T = \sum x_i\lambda_i$, as promised. $\square$

It will be convenient to treat $L(X,Y^{*\prime})$, hence $L(X,Y)$, as a space of forms on $FL(Y,X)$.

Definition 2.3 *The natural embedding* $\hat{}$ *of* $A \in L(X,Y^{*\prime})$ *into* $FL(Y,X)'$ *is given by*

$$\begin{aligned}\tilde{A}(\sum x_i\lambda_i) &\stackrel{\text{def}}{=} \sum(Ax_i)\lambda_i \\ &= \sum \lambda_i A x_i \quad \textit{if } A \in L(X,Y).\end{aligned}$$

The first part of our next result shows that $\tilde{A}$ is well defined, i.e. that $\tilde{A}$ maps different representations of T to the same real number.

Lemma 2.4

1. *Let* $\sum x_i\lambda_i$, $\sum x'_j\lambda'_j \in FL(Y,X)$. *If* $A \in L(X, Y^{*\prime})$ *and* $\sum x_i\lambda_i$ *agrees with* $\sum x'_j\lambda'_j$ *on the range of* A, *then*

$$\sum (Ax_i)\lambda_i = \sum (Ax'_j)\lambda'_j. \tag{2.1}$$

In particular, if $\sum x_i\lambda_i = \sum x'_j\lambda'_j$ *then (2.1) holds for each* $A \in L(X, Y^{*\prime})$.

2. *The mapping* $\tilde{}$ *is an algebraic isomorphism of* $L(X, Y^{*\prime})$ *onto* $FL(Y,X)'$, *hence an algebraic isomorphism of* $L(X,Y)$ *into* $FL(Y,X)'$.

Proof

1. Suppose $\sum x_i\lambda_i$ and $\sum x'_j\lambda'_j$ agree on the AX, the range of A, where $A \in L(X, Y^{*\prime})$. Let $\{\zeta_\alpha\}_\alpha \subset Y^{*\prime}$ be a Hamel basis for AX and define forms ξ_α on X by

$$Ax = \sum_\alpha \zeta_\alpha(\xi_\alpha x), \quad \forall x \in X$$

(there are finitely many nonzero summands). Then

$$\begin{aligned} \sum_i (Ax_i)\lambda_i &= \sum_i \left[\sum_\alpha \zeta_\alpha(\xi_\alpha x_i)\right]\lambda_i = \sum_\alpha \xi_\alpha \left[\sum_i x_i\lambda_i\right]\zeta_\alpha \\ &= \sum_\alpha \xi_\alpha \left[\sum_j x'_j\lambda'_j\right]\zeta_\alpha = \sum_j \left[\sum_\alpha \zeta_\alpha(\xi_\alpha x'_j)\right]\lambda'_j \\ &= \sum_j (Ax'_j)\lambda'_j. \end{aligned}$$

2. The mapping $\tilde{}$ is linear. It is also 1-1 for if $A \in L(X, Y^{*\prime})$ and $\tilde{A} = 0$ then $(Ax)\lambda = 0$ for each $x \in X$ and $\lambda \in Y^*$, hence $Ax = 0$ for each x, i.e. $A = 0$. To show surjectivity let $\phi \in FL(Y,X)'$ and define $A : X \to Y^{*\prime}$ by $Ax : \lambda \mapsto \phi(x\lambda)$. Then A is an operator with $\tilde{A} = \phi$.

$\square$

Example 2.5 *Let $X \stackrel{\text{def}}{=} \mathbb{R}^n$, $Y \stackrel{\text{def}}{=} \mathbb{R}^m$ be spaces of column vectors under any separated convex topologies. Then their duals (topological and algebraic) are the spaces of row vectors of the same dimensions n, m respectively. Also*

$$\mathbb{R}^{m\times n} = L(X,Y) = FL(X,Y) = BL(X,Y) = CL(X,Y).$$

Rank-1 operators $x\lambda$ are just outer products xv^T for column vectors $x \in \mathbb{R}^n, v \in \mathbb{R}^m$ (here $\lambda = v^T$), and $FL(Y,X) = \mathbb{R}^{n\times m}$.

For $B \in FL(Y,X)$, we can represent B as the $n \times m$ column vector formed by adjoining all its columns:

$$(B_{11}, B_{21}, \ldots, B_{n1}, B_{12}, \ldots, B_{nm})^T,$$

Then the canonical action of a matrix $A \in \mathbb{R}^{m\times n}$ on $FL(X,Y)$ satisfies

$$\tilde{A}B = \left\langle A^T, B \right\rangle = \text{trace}(A^T \circ B)$$

where $\langle \cdot, \cdot\cdot \rangle$ is the usual inner product on $(n \times m)$-vectors, matrix multiplication is denoted by $\circ$, and the trace *of a matrix is the sum of its diagonal elements.*

It is usual to work with the tensor product $X \otimes Y^*$ instead of $FL(Y,X)$, and the space of bilinear forms $Bil(X \times Y^*, \mathbb{R})$ instead of $L(X,Y)$. To explain, we summarize from [Sch, Ch. III §6]. $Bil(X \times Y^*, \mathbb{R})$ consists of mappings $H : X \times Y^* \to \mathbb{R}$ such that for $(x,\lambda) \in X \times Y^*$, $H(x, \cdot)$ and $H(\cdot, \lambda)$ are forms on Y^* and X respectively. Given $(x,\lambda) \in X \times Y^*$, $x \otimes \lambda$ is defined as the form on $Bil(X \times Y^*, \mathbb{R})$ which maps each H to $H(x,\lambda)$. The *tensor product* of X and Y^*, $X \otimes Y^*$, is the span of $\{x \otimes \lambda \mid (x,\lambda) \in X \times Y^*\}$ in $Bil(X \times Y^*, \mathbb{R})'$. Each element of $X \otimes Y^*$ is a finite sum

$$\sum x_i \otimes \lambda_i, \quad \{(x_i, \lambda_i)\} \subset X \times Y^*,$$

though this representation is not unique in general.

The natural connection between $FL(Y, X)$ and $X \otimes Y^*$ is the identification of each $\sum x_i \lambda_i$ with $\sum x_i \otimes \lambda_i$. This correspondence is an isomorphism because the elements of X can be distinguished from one another by the forms in X^*. There is also a natural embedding $\hat{}$ of $L(X, Y)$ in $Bil(X \times Y^*, \mathbb{R})$: to each $A \in L(X, Y)$, the mapping $\hat{A}$ sends $(x, \lambda) \in X \times Y^*$ to $\lambda A x$. It is easily seen that the mapping $A \mapsto \hat{A}$ preserves the duality between spaces:

$$\check{A}(\sum x_i \lambda_i) = (\sum x_i \otimes \lambda_i)\hat{A};$$

and is an isomorphism of $L(X, Y)$ into $Bil(X \times Y^*, \mathbb{R})$.

To avoid further notation we will not use the tensor product or bilinear forms. Of course, in this chapter, $X \otimes Y^*$ may be used along exactly the same lines as $FL(Y, X)$ to the same effect.

We now consider convex topologies on $FL(Y, X)$ such that the corresponding dual space is identified with $CL(X, Y)$. We call the first topology on $FL(Y, X)$ projective because this is the name given to the corresponding topology of $X \otimes Y^*$ — see [Sch, Ch. III §6].

Definition 2.6

1. *The* projective *topology on $FL(Y, X)$ is defined by the base of neighborhoods consisting of the sets* $\mathrm{co}(UC^\circ)$ *where $U \in \mathcal{N}(X)$ and $C \in \mathcal{B}(Y)$.*

   The associated topological space is denoted $FL(Y, X)_1$, or $\mathcal{F}_1$ for short.

2. *Consider the base of neighborhoods in $FL(Y, X)$ consisting of the sets*

$$\mathrm{co}[\bigcup_{x \in \mathcal{H}} x C_x^\circ \ \cup \bigcup_{V \in \mathcal{N}(Y)} U_V V^\circ]$$

   where $\mathcal{H}$ is a Hamel basis for X, $(C_x)_{x \in \mathcal{H}} \subset \mathcal{B}(Y)$ and $(U_V)_{V \in \mathcal{N}(Y)} \subset \mathcal{N}(X)$.

The associated topological space is denoted $FL(Y,X)_2$, or $\mathcal{F}_2$ for short.

3. *Let Y be metrizable with a base of neighborhoods $\{V_n \mid n \in \mathbb{N}\}$.*

   Consider the base of neighborhoods in $FL(Y,X)$ consisting of the sets

$$\mathrm{co}[\bigcup_{n\in\mathbb{N}} U_n V_n^\circ]$$

   where $(U_n)_{n\in\mathbb{N}} \subset \mathcal{N}(X)$.

   The associated topological space is denoted $FL(Y,X)_3$, or $\mathcal{F}_3$ for short.

Note that, in the above bases of neighborhoods, polars of bounded subsets of Y can be replaced by strong neighborhoods in Y^*. For example, the projective neighborhoods could be expressed as $\mathrm{co}(UW)$ where U and W are neighborhoods of X and Y^* respectively.

Proposition 2.7

1. *$\mathcal{F}_1$ is a separated convex space such that, under the natural embedding $\tilde{}$, $\mathcal{F}_1^*$ is algebraically isomorphic to $BL(X,Y)$.*

2. *$\mathcal{F}_2$ is a separated convex space such that, under the natural embedding $\tilde{}$, $\mathcal{F}_2^*$ is algebraically isomorphic to $CL(X,Y)$.*

3. *Let Y be metrizable.*

   $\mathcal{F}_3$ is a separated convex space such that, under the natural embedding $\tilde{}$, $\mathcal{F}_3^*$ is algebraically isomorphic to $CL(X,Y)$.

Proof That each of $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3$ is a convex space results immediately from [Sch, Ch. I §1.2]. Separation follows if the elements of the dual separate distinct elements of $FL(Y,X)$. Suppose S, T are distinct members of $FL(Y,X)$, i.e. $Sy \neq Ty$ for some $y \in Y$. Since X is separated, there is $\xi \in X^*$ such that $\xi Sy \neq \xi Ty$. The mapping $A : X \to Y : x \mapsto y(\xi x)$ is a bounded, hence continuous, operator from X to Y such that $\tilde{A}S = \xi Sy \neq \xi Ty = \tilde{A}T$. So $\tilde{A}$ separates S and T and, from 1 and 2 below, $\tilde{A}$ is a continuous form on each of $\mathcal{F}_1, \mathcal{F}_2$ and $\mathcal{F}_3$.

In view of Lemma 2.4.2, we will freely pass from forms $\phi \in FL(Y, X)'$ to operators $A \in L(X, Y^{*\prime})$ where $\phi = \tilde{A}$. We will also use the fact that a linear mapping on a topological vector space is continuous if (and only if) it is bounded above on some neighborhood W of the space, for then it is bounded on $W \cap -W$ which is also a neighborhood. Second polars of subsets of Y are taken in Y^{**}.

1. We will show that

$$\mathcal{F}_1^* \cong EL(X, Y^{**}) \tag{2.2}$$

where the latter is the space of all Equicontinuous operators from X to Y^{**}:

$$\begin{aligned} &EL(X, Y^{**}) \\ &\quad \stackrel{\text{def}}{=} \{A \in L(X, Y^{**}) \mid AU \subset D \text{ for some } U \in \mathcal{N}(X), D \in \mathcal{E}(Y^{**})\}. \end{aligned}$$

The result follows because, when Y is semi-reflexive, $\mathcal{E}(Y^{**})$ is identified with $\mathcal{B}(Y)$, hence $EL(X, Y^{**})$ is identified with $BL(X, Y)$.

To prove (2.2), let $\phi \in FL(Y, X)'$ and $A \in L(X, Y^{*\prime})$ be such that $\tilde{A} = \phi$. Then

$$\begin{aligned} & \phi \in \mathcal{F}_1^* \\ \iff \quad & \exists U \in \mathcal{N}(X),\ C \in \mathcal{B}(Y), \quad \phi(UC^\circ) = (AU)(C^\circ) \text{ bounded above} \\ \iff \quad & \exists U \in \mathcal{N}(X),\ C \in \mathcal{B}(Y), \quad AU \subset C^{\circ\circ} \\ \iff \quad & A \in EL(X, Y^{**}). \end{aligned}$$

2. Let $\phi \in FL(Y, X)'$ and $A \in L(X, Y^{*\prime})$ satisfy $\phi = \tilde{A}$. Then

$$\phi \in \mathcal{F}_2^*$$

$$\iff \forall(\, x \in \mathcal{H},\ V \in \mathcal{N}(Y)\,)\, \exists(\, C \in \mathcal{B}(Y),\ U \in \mathcal{N}(X)\,)$$

$$\phi(xC^\circ), \phi(UV^\circ) \text{ bounded above}$$

$$\iff \forall(\, x \in X,\ V \in \mathcal{N}(Y)\,)\, \exists(\, C \in \mathcal{B}(Y),\ U \in \mathcal{N}(X)\,)$$

$$(Ax)C^\circ, (AU)V^\circ \text{ bounded above}$$

$$\iff A \in CL(X,Y)$$

where semi-reflexivity of Y is used to show the final equivalence.

3. From part 2, $\mathcal{F}_2^* \cong CL(X,Y)$. To relate $\mathcal{F}_2$ and $\mathcal{F}_3$, consider the following statements:

   (a) $\mathcal{F}_3^* = \mathcal{F}_2^*$

   (b) If $A \in L(X, Y^{*\prime})$ and $\exists (U_V)_{V \in \mathcal{N}(Y)} \subset \mathcal{N}(X)$ such that $\tilde{A}$ is bounded above on $\bigcup_V U_V V^\circ$, then $\exists (C_x)_{x \in \mathcal{H}} \subset \mathcal{B}(Y)$ such that $\tilde{A}$ is bounded above on $\bigcup_x xC_x^\circ$.

   (c) If $A \in L(X, Y^{*\prime})$ and $\forall V \in N(Y)\ \exists U \in N(X)$ with $(AU)V^\circ$ bounded above, then $A \in L(X, Y^{**})$.

   (d) If $\zeta \in Y^{*\prime}$ is bounded above on each $D \in \mathcal{E}(Y^*)$ then $\zeta \in Y^{**}$.

   Clearly 3a $\iff$ 3b $\iff$ 3c $\Longleftarrow$ 3d. (The converse of the last implication is not hard to see, but is not needed for this proof.)

   If Y is a metric space which is semi-reflexive, then it is also reflexive (use [Sch, Ch. II §8.1, and remarks following Ch. IV §5.6 Cor. 2]). Therefore $\mathcal{E}(Y^{**}) = (\mathcal{B}(Y) =)\mathcal{B}(Y^{**})$. In this situation [Sch, Ch. IV §6.6] says that statement 3d holds, hence 3a holds and we are done.

□

Definition 2.8 *Let Y be semi-reflexive. Let $\mathcal{F}, \mathcal{L}$ be the respective spaces $\mathcal{F}_1$, $BL(X,Y)$ or $\mathcal{F}_2, CL(X,Y)$ or, assuming Y is metrizable, $\mathcal{F}_3, CL(X,Y)$.*

The strong (respectively weak) topology of $\mathcal{L}$ determined by $\mathcal{F}$ is called the strong (w-) topology of $\mathcal{L}$.*

The fact that $\mathcal{L}$ is identified with $\mathcal{F}^*$ explains the notation w*- for the weak topology of $\mathcal{L}$ determined by $\mathcal{F}$.

Example 2.9 *Suppose X and Y normed spaces. Then $\mathcal{F}_1$ is also a normed space (see [Sch, Ch. III §6.4]): for T in $FL(Y,X)$, the projective norm is*

$$\|T\| = \inf\{\sum \|x_i\|\|\lambda_i\| \mid T = \sum x_i\lambda_i \in FL(Y,X)\}$$

The usual operator norm $\|T\|_{op} \overset{\text{def}}{=} \sup\{\|Ty\| \mid y \in Y, \|y\| \leq 1\}$ is no larger than the projective norm though both agree on rank-1 operators:

$$\|x\lambda\| = \|x\|\|\lambda\| = \|T\|_{op}.$$

The operator norm topology on $FL(Y,X)$ is, in general, strictly coarser than the projective topology (so the dual of $FL(Y,X)$ with respect to the operator norm is generally a proper subspace of $BL(X,Y)$, the dual taken with respect to the projective norm).

However, as one can easily check, the dual norm of $FL(Y,X)$ is exactly the operator norm on $BL(X,Y)$:

$$\|\tilde{A}\| = \|A\|_{op}, \quad \forall A \in BL(X,Y).$$

Here all of the topologies on $FL(Y,X)$ in Definition 2.6 are equivalent. In fact these topologies are equivalent so long as Y is normable (though X may not be) in which case we also have $BL(X,Y) = CL(X,Y)$.

Remarks on reflexivity of Y. In the proof of Proposition 2.7 we see that, without the semi-reflexivity condition on Y, the duals of $\mathcal{F}_1$ and $\mathcal{F}_2$ consist of operators from X to Y^{**}, the dual of Y^*. To force dual elements to map X to Y there are two immediate possibilities. Suppose $\mathcal{F}$ is either $\mathcal{F}_1$ or $\mathcal{F}_2$.

1. Enlarge the neighborhoods of $\mathcal{F}$ so that if the mapping $A \in L(X, Y^{**})$ belongs to $\mathcal{F}^*$ then $A \in L(X, Y)$. This can be done using the original base of neighborhoods for $\mathcal{F}$ given in Definition 2.6; denote this by $\mathcal{N}$. The following sets then form the required base of neighborhoods for $FL(Y, X)$:

   $$\mathrm{co}[W \cup x\Lambda],$$

   where $W \in \mathcal{N}$, $x \in X$, and Λ is a w*-neighborhood in Y^*.

   Such a topology on $FL(Y, X)$ seems to be artificially restrictive in practice, however. For example it may be difficult to verify that a sublinear function on $FL(Y, X)$ is (lower semi)continuous in the new topology.

2. If Y itself is the dual of another convex space Z, under some polar topology, then reflexivity may be dropped. To make this precise, let Z be a separated convex space and $\mathcal{D}$ be a subfamily of $\mathcal{B}(Z)$ such that $\{D^\circ \mid D \in \mathcal{D}\}$ defines a base of neighborhoods making Z^* a convex space (see [Sch, Ch. III §3]); endow Z^* with this topology. Let XZ be the span of $\{xz | (x, z) \in X \times Z\}$ in $L(Z^*, X)$, where $xz : Z^* \to \mathbb{R} : \zeta \mapsto x(\zeta z)$. For $A \in L(X, Z^*)$, let $\tilde{A} : XZ \to \mathbb{R} : \sum x_i z_i \mapsto \sum (Ax_i) z_i$. Finally, let $\mathcal{H}$ be a Hamel basis of X.

   Without any more difficulty than the previous proof entails, we obtain the following:

   (a) Let XZ have the topology whose base of neighborhoods consists of sets

   $$\mathrm{co}[\bigcup_{x \in \mathcal{H}} xW_x \;\cup\; UB^\circ]$$

where $(W_x)_x \subset \mathcal{N}(Z)$, $U \in \mathcal{N}(X)$, $B \in \mathcal{B}(Z^*)$ and the polar B° is taken in Z. Then the mapping $\tilde{}$ is an algebraic isomorphism of $BL(X, Z^*)$ onto $(XZ)^*$.

(b) Let XZ have the topology whose base of neighborhoods consists of sets

$$\mathrm{co}[\bigcup_{x \in \mathcal{H}} xW_x \quad \cup \quad \bigcup_{D \in \mathcal{D}} U_D D],$$

where $(W_x)_x \subset \mathcal{N}(Z)$, and $(U_D)_D \subset \mathcal{N}(X)$. Then the mapping $\tilde{}$ is an algebraic isomorphism of $CL(X, Z^*)$ onto $(XZ)^*$.

(c) Suppose that the sets $\cup_x xW_x$ are omitted in either of the above bases of neighborhoods of XZ. The corresponding claim of isomorphism still holds iff any convex subset of Z which absorbs each $D \in \mathcal{D}$ is a neighborhood in Z.

Suppose $\mathcal{D} = \mathcal{B}(Z)$, so that Z^* has the strong dual topology. If the convex sets in Z have the property described in 2c above, then Z is said to be *bornological.* It is not difficult to show [Sch, Ch. II §8.1] that all metrizable spaces are bornological — this is trivial for normed spaces — hence 2c applies to metrizable Z when we let Z^* be the strong dual.

Example 2.10 *Suppose X and Y are normed spaces, with Y possibly non-reflexive. We consider the situation of $BL(X, Y^{**})$ instead of $BL(X, Y)$. Here, of course, boundedness and continuity of linear mappings are equivalent.*

Now $FL(Y, X)(= XY^*)$ is normable with the unit ball $\mathrm{co}[B_X B_Y^*]$. Taking $Z \overset{\mathrm{def}}{=} Y^*$, then, according to 2c and the discussion above, the mapping $\tilde{}$ is an algebraic isomorphism of $BL(X, Y^{**})$ onto $FL(Y, X)^*$.

Example 2.11 *Suppose X is metrizable. We consider $CL(X, X^*)$ as one might do when dealing with gradients of real differentiable functions on X. Endow $FL(X^*, X)(= XX)$ with the topology defined by the base of neighborhoods consisting of all sets*

$$\operatorname{co} \bigcup_{D \in \mathcal{B}(X)} U_D D, \quad \textit{where } (U_D)_D \subset \mathcal{N}(X).$$

Apply 2c with $Z \stackrel{\text{def}}{=} X$ and $\mathcal{D} \stackrel{\text{def}}{=} \mathcal{B}(Z)$: the mapping $\tilde{}$ is an algebraic isomorphism of $CL(X, X^*)$ onto $FL(X^*, X)^*$.

We will not be concerned any further with results at this level of abstraction. The results to follow could, however, be presented in this more general framework.

2.3 Rank-1 Support Functions, Rank-1 Representers

Support functions of closed convex sets are of great importance in functional analysis. By convention, the supremum taken over an empty set is $-\infty$.

Definition 2.12 *Let $\Phi \subset E^*$. The* support functional *of Φ, $\sigma_\Phi : E \to \mathbb{R} \cup \{\pm\infty\}$, is given by*

$$\sigma_\Phi(e) \stackrel{\text{def}}{=} \sup_{\phi \in \Phi} \phi e, \quad \forall e \in E.$$

σ_Φ is called a support function on E.

We are not very interested in the value $-\infty$ or the support function of the empty set, because σ_Φ takes the value $-\infty$ at some point iff $\Phi = \emptyset$ and $\sigma_\Phi \equiv -\infty$.

In the 3-dimensional case, the ideas in part 1 of the next theorem can be traced back to Minkowski's 1911 paper [Min]. Part 2 of the theorem is essentially due to

Hörmander [Hör]. We should point out that an extended real valued function p on E is said to be *sublinear* if it is *positive homogeneous* and *subadditive* (properties 2a and 2b in the theorem below). By convention, nonnegative multiples of infinity are defined by

$$\alpha(\pm\infty) \stackrel{\text{def}}{=} \begin{cases} \pm\infty & \text{if } \alpha > 0 \\ 0 & \text{if } \alpha = 0. \end{cases}$$

For clarity we also state the definition of the limit inferior: given $p : E \to \mathbb{R} \cup \{\pm\infty\}$, $e_\circ \in E$,

$$\liminf_{e \to e_\circ} p(e) \stackrel{\text{def}}{=} \sup_{U \in \mathcal{N}(e_\circ)} \inf_{e \in U} p(e).$$

Theorem 2.13 *Let Φ be a set in E^*.*

1. $\text{w}^*\text{-}\operatorname{cl}\operatorname{co}\Phi = \{\phi \in E^* \mid \forall e \in E,\ \phi e \le \sigma_\Phi(e)\}$.

2. *Let $p : E \to \mathbb{R} \cup \{\pm\infty\}$. Then p is the support function of a nonempty set in E^* iff $p(0) \ne -\infty$ and p is*

   (a) *positive homogeneous: $p(\alpha e) = \alpha p(e)$ for each $\alpha > 0$ and $e \in E$;*

   (b) *subadditive: $p(e_1 + e_2) \le p(e_1) + p(e_2)$ for each $e_1, e_2 \in E$;*

   (c) *lower semicontinuous: $\liminf_{e \to e_\circ} p(e) = p(e_\circ)$ for each $e_\circ \in E$.*

   If p is the support function of a nonempty set, there is a unique (nonempty) w-closed convex set $\Phi \subset E^*$ such that $p = \sigma_\Phi$, and p never takes the value $-\infty$.*

Proof

1. This well known result is a corollary of Theorem 2.1.2, by taking E^* under the w*-topology as the convex superspace.

2. We begin by disposing of $-\infty$. If $p = \sigma_\Phi$ for some $\emptyset \neq \Phi \subset E^*$ then p never takes the value $-\infty$. Suppose instead that $p(0) \neq -\infty$ and p satisfies 2a-2c. Then p never takes the value $-\infty$, otherwise positive homogeneity and lower semicontinuity at 0 would lead to the contradiction $-\infty \geq p(0)$. We may therefore assume without loss of generality that $p : E \to \mathbb{R} \cup \{\infty\}$.

   It turns out that p is positive homogeneous and subadditive iff p is positive homogeneous and convex. Moreover a convex function on E is lower semicontinuous iff it is weakly lower semicontinuous — this follows from [Sch, Ch. IV §3.3] by examining the (convex) epigraph of the function. With these facts, the dual statement of [Hör, Thm. 4] completes the result.

□

Theorem 2.14 *Let $0 \in K \subset E$, and K be a* cone*: $\alpha K \subset K$, $\forall \alpha > 0$. Let $p : K \to \mathbb{R} \cup \{\pm\infty\}$.*

1. *p is the restriction to K of a support function of a nonempty set in E^* iff $p(0) \neq -\infty$ and p is*

   (a) positive homogeneous (on K);

   (b) lower semicontinuous and subadditive in k: $\forall k \in K$,

   $$p(k) = \liminf\{\textstyle\sum p(k_i) \,|\, \textit{finite sums } \sum k_i \to k, (k_i) \subset K\}.$$

   If p is the restriction to K of a support function of a nonempty set, there is a largest (nonempty) w^*-closed convex set $\Phi \subset E^*$ such that $p = \sigma_\Phi|_K$, and p never takes the value $-\infty$.

2. *Suppose $\mathcal{W}$ is a family of sets in K such that $\{coW \,|\, W \in \mathcal{W}\}$ is a base of neighborhoods in E.*

Then p is the restriction to K of a support function of a nonempty equicontinuous set in E^* iff $p(0) \neq -\infty$ and p is

(a) positive homogeneous (on K);

(b) subadditive in k: $\forall k \in K$,

$$p(k) \leq \sum p(k_i) \ \forall \textit{finite sums } \sum k_i = k, (k_i) \subset K;$$

(c) bounded above on some $W \in \mathcal{W}$.

If p is the restriction to K of a support function of a nonempty equicontinuous set, there is a largest (nonempty) w^*-compact convex set $\Phi \subset E^*$ such that $p = \sigma_\Phi|_K$, and p never takes the value $-\infty$.

Proof

1. One way is easy: if $p = \sigma_\Phi|_K$, where $\emptyset \neq \Phi \subset E$, clearly $p(0) \neq -\infty$ and p satisfies 1a and 1b.

   Conversely, suppose $p(0) \neq -\infty$ and 1a and 1b hold. Define $\bar{p}$ by

$$\bar{p}(e) = \begin{cases} \liminf\{\sum p(k_i) \mid \text{finite sums } \sum k_i \to e, (k_i) \subset K\}, & \text{if } e \in \operatorname{cl}\operatorname{co} K \\ \infty, & \text{otherwise.} \end{cases}$$

   Since K is a cone, $\operatorname{co} K = \{\sum k_i \mid \text{finite sums } \sum k_i, (k_i) \subset K\}$ and $\bar{p}$ is a well defined mapping from E to $\mathbb{R} \cup \{\pm\infty\}$.

   By construction of $\bar{p}$ and property 1b we have $\bar{p}|_K = p$; in particular $\bar{p}(0) \neq -\infty$. It is not hard to show from the definition that $\bar{p}$ is both positive homogeneous and subadditive. Finally we show that $\bar{p}$ is lower semicontinuous. If $e_\circ \notin \operatorname{cl}\operatorname{co} K$ then $\bar{p}$ takes the value ∞ on a neighborhood of $e_\circ$, hence is lower semicontinuous there. So let $e_\circ \in \operatorname{cl}\operatorname{co} K$ and $\epsilon > 0$.

For $U \in \mathcal{N}(e_\circ)$ and each $e \in U \cap \operatorname{cl}\operatorname{co} K$ there is, by definition of $\bar{p}$, a finite sum $\sum k_i \in U$ such that $(k_i) \subset K$ and $\sum p(k_i) \leq \bar{p}(e) + \epsilon$. Hence

$$\begin{aligned} & \inf\{\sum p(k_i) \mid \text{ finite sums } \sum k_i \in U,\ (k_i) \subset K\} \\ \leq\ & \inf\{\bar{p}(e) \mid e \in U \cap \operatorname{cl}\operatorname{co} K\} + \epsilon \end{aligned}$$

So

$$\inf\{\sum p(k_i) \mid \text{ finite sums } \sum k_i \in U,\ (k_i) \subset K\} \ \leq\ \inf\{\bar{p}(e) \mid e \in U\} + \epsilon.$$

Taking the supremum over all neighborhoods $U \in \mathcal{N}(e_\circ)$ on both sides of the inequality yields $\bar{p}(e_\circ) \leq \liminf_{e \to e_\circ} \bar{p}(e) + \epsilon$. As ϵ is an arbitrary positive scalar, $\bar{p}$ is lower semicontinuous at $e_\circ$. By Theorem 2.13.2, $\bar{p}$ is the support function of a nonempty set in E^* as required.

The last statement is trivial: the largest such set is

$$\Phi \stackrel{\text{def}}{=} \{\phi \in E^* \mid \phi k \leq p(k),\ \forall k \in K\}.$$

2. We proceed in a similar fashion to the proof of part 1.

   If $p = \sigma_\Phi|_K$ where Φ is a nonempty equicontinuous set in E^*, it is easy to see p has the required properties.

   Conversely, suppose $p(0) \neq -\infty$ and p satisfies 2a- 2c. Note that $\operatorname{co} K = E$ since $\operatorname{co} K$ is a cone containing a neighborhood in E. Now define $\bar{p}$ on E by

$$\bar{p}(e) \stackrel{\text{def}}{=} \inf\{\sum p(k_i) \mid \text{finite sums } \sum k_i = e, (k_i) \subset K\};$$

   then $\bar{p}|_K = p$ by 2b; in particular $\bar{p}(0) \neq -\infty$. Also, with properties 2a and 2b of p, it is easy to see that $\bar{p}$ is positive homogeneous and subadditive, hence convex. Using 2c and convexity of $\bar{p}$ we see that $\bar{p}$ is bounded above on the neighborhood $\operatorname{co} W$. This ensures, by convexity again, that $\bar{p}$ is actually continuous (eg. [Roc74, Thm. 8]).

To sum up: $\bar{p}(0) \neq -\infty$ and $\bar{p}$ is positive homogeneous, subadditive and (lower semi)continuous. Hence, Theorem 2.13.2 says $\bar{p}$ is the support function of a nonempty set Φ in E^*. We know further that $\bar{p}$ is bounded above on a neighborhood in E; consequently, Φ must be equicontinuous.

The largest set Φ with $\sigma_\Phi|_K = p$,

$$\Phi \stackrel{\text{def}}{=} \{\phi \in E^* \mid \phi e \leq p(e), \quad \forall e \in E\}$$

is equicontinuous as well as w*-closed. By Alaoglu-Bourbaki [Sch, Ch. III §4.3 Cor.] it is w*-compact.

□

In view of Theorem 2.7, we can apply Theorem 2.13 taking E as $\mathcal{F}$ and its dual as $\mathcal{L}$. This allows us to characterize the support functions of w*-closed, convex sets of operators. In finite dimensions, where all separated convex topologies are equivalent, this kind of result was noted in [H-U82, §2].

If we are prepared to forego *exact* representation of w*-closed convex sets of operators by support functions, we may use simpler support functionals by restricting function domains to the rank-1 members of $FL(Y,X)$. This involves only pairs (x, λ) in $X \times Y^*$ rather than operators $\sum x_i \lambda_i$ of any finite rank. To this end we have developed Theorem 2.14, which we will apply to $E \stackrel{\text{def}}{=} \mathcal{F}$, $E^* \stackrel{\text{def}}{=} \mathcal{L}$ and $K \stackrel{\text{def}}{=} \{x\lambda \mid (x, \lambda) \in X \times Y^*\}$.

We could equally restrict our attention to any other cone in of $FL(Y,X)$, such as the set of operators of rank less than 3. In fact Theorem 2.14 is easily generalized to the case of a non-cone K, so that restricting mappings to any given subset of $FL(Y, X)$ is possible. The 'natural' applications, however, seem to fall in the rank-1 category as we shall see in §2.4.

The next definition is motivated by the idea of support functions (Definition 2.12) and Theorem 2.13.1.

Definition 2.15 *Let Γ be a set in $\mathcal{L}$.*

1. *The* rank-1 *support functional of Γ, $\sigma^1_\Gamma : X \times Y^* \to \mathbb{R} \cup \{\infty\}$, is given by*

$$\sigma^1_\Gamma(x, \lambda) \stackrel{\text{def}}{=} \sup_{A \in \Gamma} \lambda A x, \quad \forall (x, \lambda) \in X \times Y^*.$$

   σ^1_Γ is called a rank-1 support function on $X \times Y^*$.

2. *The* (maximum) rank-1 representer *of Γ is*

$$\Gamma^1 \stackrel{\text{def}}{=} \{A \in \mathcal{L} \mid \forall (x, \lambda) \in X \times Y^*, \ \lambda A x \leq \sigma^1_\Gamma(x, \lambda)\}.$$

   Γ is said to be a (maximum) rank-1 representer *if $\Gamma = \Gamma^1$.*

Note: Some immediate properties of rank-1 representers are

1. Γ^1 is the largest set whose rank-1 support function coincides with that of Γ. So Γ^1 contains Γ.

2. Γ^1 is the intersection w*-closed convex sets — those of the form $\{A \in \mathcal{L} | \lambda A x \leq \alpha\}$, where $(x, \lambda) \in X \times Y^*$ and $\alpha \in \mathbb{R}$ — hence is w*-closed and convex. Therefore, since $\Gamma^1 \supset \Gamma$, Γ^1 contains the w*-closed convex hull of Γ.

3. For $\bar{A} \in \mathcal{L}$: $\lambda \bar{A} x \in \lambda \Gamma^1 x, \ \forall (x, \lambda) \in X \times Y^* \ \Rightarrow \ \bar{A} \in \Gamma^1$.

4. $(\Gamma^1)^1 = \Gamma^1$.

5. Separation: $A \notin \Gamma^1$ iff there is (x, λ) with $\lambda A x > \sigma^1_\Gamma(x, \lambda)$.

Rank-1 representers have been discussed before as the *plenary subsections* of [Swe79]. An alternative definition is given in terms of *H-convexity* by [KR]. Our terminology emphasizes the rank-1 aspect of such sets because rank-1 representers are the sets in $\mathcal{L}$ which are dually described by the rank-1 operators of $\mathcal{F}$. For equivalent definitions, including one using properties 2 and 3 above, and also details of *rank-1* and *plenary* sets of operators, see the next section. For the moment we are content with some examples:

Example 2.16 *Let $\Gamma = \{A_0\}$, for some $A_0 \in \mathcal{L}$. Clearly $A_0 \in \Gamma^1$. Now suppose $A \in \Gamma^1$, then for each pair $(x, \lambda) \in X \times Y^*$ we have*

$$\tilde{A}(x\lambda) = \lambda A x \leq \lambda A_0 x = \tilde{A}_0(x\lambda),$$

hence, by considering $(-x, \lambda)$ too,

$$\tilde{A}(x\lambda) = \tilde{A}_0(x\lambda).$$

Since the members of $\{\tilde{A} \mid A \in \mathcal{L}\}$, hence the members of $\mathcal{L}$, can be distinguished by the rank-1 mappings $\{x\lambda \mid (x, \lambda) \in X \times Y^*\}$, we must have $A = A_0$, i.e.

$$\Gamma^1 = \Gamma.$$

Example 2.17 *Suppose $g : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at x_*. The* interval Jacobian *of g at x_* is an $m \times n$ matrix of real intervals, J_{ij},*

$$J = [J_{ij}]_{m \times n}$$

where each J_{ij} contains the ij-th partial derivative of g, $\nabla g(x_*)_{ij} = dg_i(x_*)/dx_j$. Therefore, $\nabla g(x_*) \in J$. The interval Jacobian can be used to compute guaranteed error bounds for numerical processes. See [Neu] for example.

Also $J^1 = J$ by reasoning similar to that in the previous example.

The next example shows that convexity and (w*-)compactness of a set are not sufficient to ensure it is a rank-1 representer.

Example 2.18 *[Swe77, Example 6.2]*

$$\Gamma \stackrel{\text{def}}{=} \operatorname{cl}\operatorname{co}\left\{ \begin{bmatrix} 1 & 0 \\ -2 & -1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}, \begin{bmatrix} -1 & -2 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 2 \\ 0 & 1 \end{bmatrix} \right\}.$$

The set Γ is compact and convex. Note that the identity matrix I is not in Γ because trace$(A) = 0$ for each $A \in \Gamma$, whereas trace$(I) = 2$.

Nevertheless, we have $u \in \Gamma u$ for each $u \in \mathbb{R}^2$. The case $u = 0$ is trivial, so assume that $u = (u_1, u_2)^T \neq 0$. If $0 \neq |u_1| \geq |u_2|$ then

$$A = \begin{bmatrix} 1 & 0 \\ 2u_2/u_1 & -1 \end{bmatrix} \in \Gamma$$

with $Au = u$. Otherwise $0 \neq |u_2| \geq |u_1|$ and

$$A = \begin{bmatrix} -1 & 2u_1/u_2 \\ 0 & 1 \end{bmatrix} \in \Gamma$$

with $Au = u$.
We conclude that $I \in \Gamma^1 \backslash \Gamma$.

We now examine sets in $\mathcal{L}$ having one of the rather strong but very useful properties of uniform boundedness, equicontinuity, uniform openness or uniform invertibility. We will show that $\Gamma \subset \mathcal{L}$ satisfies such a property iff its rank-1 representer also satisfies this property in which case, roughly speaking, the 'constants' involved are the same. Thus in some important cases, dealing with the rank-1 representer of Γ is equivalent to dealing with Γ itself.

We review the relevant properties of Γ. For $U \subset X$, let ΓU denote $\cup_{A \in \Gamma} AU$. Recall Γ is uniformly bounded if for some neighborhood U in X, ΓU is bounded in

Y; and Γ is equicontinuous, or uniformly continuous, if for each neighborhood V of Y there is a neighborhood U of X such that $\Gamma U \subset V$. Furthermore, an operator $A \in CL(X,Y)$ is open if it maps open sets to open sets or, equivalently, if for each neighborhood U in X there is a neighborhood V in Y such that $AU \supset V$. (Such a map may also be called a topological homomorphism — see [Sch, Ch. III §1].) A set of operators Γ is uniformly open if for each neighborhood U there is a neighborhood V such that $AU \supset V$ for each A in Γ. Likewise we say Γ is uniformly (continuously) invertible if it is uniformly open and consists of injective, or 1-1, mappings.

Proposition 2.19 *Let Γ be a set in $CL(X,Y)$. Let $\tilde{\Gamma} \stackrel{\text{def}}{=} \{\tilde{A} \mid A \in \Gamma\}$, the canonical embedding of Γ in ${\mathcal{F}_3}^*$. Likewise $\tilde{\Gamma}^1$ is the embedding of Γ^1 in ${\mathcal{F}_2}^*$.*

1. *For each nonempty $B \subset X$, and each closed convex $C \subset Y$ containing 0,*

$$\tilde{\Gamma} \subset (BC^\circ)^\circ \iff \Gamma B \subset C \iff \Gamma^1 B \subset C \iff \tilde{\Gamma}^1 \subset (BC^\circ)^\circ$$

where the polar of C is taken in Y^* and the polar of BC° in ${\mathcal{F}_2}^*$.

If Γ consists of bounded mappings we may substitute $\mathcal{F}_1$ for $\mathcal{F}_2$. If Y is metrizable we may substitute $\mathcal{F}_3$ for $\mathcal{F}_2$.

2. *We have*

$$\begin{aligned} &\Gamma \text{ is uniformly bounded} &&\iff \tilde{\Gamma} \text{ is equicontinuous in } \mathcal{F}_1^* \\ \iff\ &\Gamma^1 \text{ is uniformly bounded} &&\iff \tilde{\Gamma}^1 \text{ is equicontinuous in } \mathcal{F}_1^*. \end{aligned}$$

In fact for each $U \in \mathcal{N}(X)$, and each closed convex $C \in \mathcal{B}(Y)$ containing 0,

$$\tilde{\Gamma} \subset (UC^\circ)^\circ \iff \Gamma U \subset C \iff \Gamma^1 U \subset C \iff \tilde{\Gamma}^1 \subset (UC^\circ)^\circ.$$

3. We have

$$\begin{aligned} \Gamma \textit{ is equicontinuous} &\iff \tilde{\Gamma} \textit{ is equicontinuous in } \mathcal{F}_2^* \\ \iff \Gamma^1 \textit{ is equicontinuous} &\iff \tilde{\Gamma}^1 \textit{ is equicontinuous in } \mathcal{F}_2^*. \end{aligned}$$

In fact for any neighborhood U in X and closed convex neighborhood V in Y

$$\Gamma U \subset V \iff \Gamma^1 U \subset V;$$

and for any N belonging to the neighborhood base of $\mathcal{F}_2$ in Definition 2.6,

$$\tilde{\Gamma} \subset N^\circ \iff \tilde{\Gamma}^1 \subset N^\circ.$$

If Y is metrizable, $\mathcal{F}_3$ may be substituted for $\mathcal{F}_2$.

4. Let X be complete and metrizable and Γ be convex.

Γ is uniformly open iff Γ^1 is uniformly open; in this case, if U and V are closed convex neighborhoods in X and Y then

$$\forall A \in \Gamma,\ AU \supset V \quad \iff \quad \forall A \in \Gamma^1,\ AU \supset V.$$

Moreover, Γ is uniformly invertible iff Γ^1 is.

Proof We note for use below firstly that we can take $\mathcal{N}(Y)$ consisting of closed convex sets ([Sch, Ch. II §4]) hence, secondly, that a set is is bounded in Y iff (it is contained in a closed, convex, bounded set in Y iff) it is contained in a closed, convex, bounded set in Y containing 0.

1. Let B, C be as in the statement of the theorem. Firstly we deduce from Theorem 2.1.1 that $\Gamma B \subset C \iff C^\circ \Gamma B \subset (-\infty, 1]$. To finish, use this and the following:

$$\begin{aligned} \tilde{\Gamma} \subset (BC^\circ)^\circ &\iff C^\circ \Gamma B \subset (-\infty, 1] \\ &\iff C^\circ \Gamma^1 B \subset (-\infty, 1] \iff \tilde{\Gamma}^1 \subset (BC^\circ)^\circ \end{aligned}$$

where the second equivalence holds because $\sigma^1_\Gamma = \sigma^1_{\Gamma^1}$.

2. Given the opening remarks of the proof, this is an immediate corollary of part 1.

3. Part 1 yields, for $U \in \mathcal{N}(X)$ and closed convex $V \in \mathcal{N}(Y)$, that $\Gamma U \subset V \iff \Gamma^1 U \subset V$. This shows that Γ is equicontinuous iff Γ^1 is also equicontinuous.

   Let N be a member of the base of neighborhoods of $\mathcal{F}_2$ in Definition 2.6. Then $N = \operatorname{co} W$ for some nonempty $W \subset \{x\lambda \mid (x,\lambda) \in X \times Y^*\}$. We always have $(\operatorname{co} W)^\circ = W^\circ$, so

$$\begin{aligned} \tilde{\Gamma} \subset N^\circ = W^\circ \quad &\iff \quad \sigma^1_\Gamma(x,\lambda) \le 1, \ \ \forall x\lambda \in W \\ &\iff \quad \sigma^1_{\Gamma^1}(x,\lambda) \le 1, \ \ \forall x\lambda \in W \\ &\iff \quad \tilde{\Gamma}^1 \subset W^\circ = N^\circ. \end{aligned}$$

   Finally,

$$\begin{aligned} &\Gamma \text{ equicontinuous} \\ \iff \ &\Gamma \text{ equicontinuous, } \Gamma x \text{ bounded } \forall x \in \mathcal{H} \\ \iff \ &\forall(\, x \in \mathcal{H},\ V \in \mathcal{N}(Y)\,)\ \exists(\, C \in \mathcal{B}(Y),\ U \in \mathcal{N}(X)\,) \\ &\quad \tilde{\Gamma} \ \subset \ (xC^\circ)^\circ \cup (UV^\circ)^\circ \\ \iff \ &\tilde{\Gamma} \text{ equicontinuous in } \mathcal{F}_2^*, \end{aligned}$$

   where the second equivalence holds by part 1 after again using the opening remarks of the proof.

4. Let U,V be closed convex neighborhoods in X,Y respectively. For A in $\mathcal{L}$ we will use the fact that

$$\sigma_{AU} \ge \sigma_V \iff \operatorname{cl}(AU) \supset V \iff AU \supset V.$$

The first equivalence is given by Theorem 2.1.2. The second, which is related to the Open Mapping Theorem and for which a sufficient condition is that X be complete and metrizable, is given by [Sch, Ch IV §8 Exa. 1, and §8.6]. Then

$$\begin{aligned}
AU \supset V,\ \forall A \in \Gamma &\implies \inf_{A\in\Gamma} \sigma_{AU} \geq \sigma_V \\
&\implies \inf_{\xi\in\lambda\Gamma} \sigma_U(\xi) \geq \sigma_V(\lambda),\ \forall \lambda \in Y^* \\
&\implies \inf_{\xi\in\lambda\Gamma^1} \sigma_U(\xi) \geq \sigma_V(\lambda),\ \forall \lambda \in Y^* \\
&\implies \inf_{A\in\Gamma^1} \sigma_{AU} \geq \sigma_V \\
&\implies AU \supset V,\ \forall A \in \Gamma^1.
\end{aligned}$$

The third implication follows from w*-lower semicontinuity of σ_U on X^* ([Hör]) and Lemma 2.20, below, which says w*- cl$(\lambda\Gamma)$ = w*- cl$(\lambda\Gamma^1)$. So Γ^1 is uniformly open in the manner required.

To show the the final claim, suppose Γ consists of injective operators and let $0 \neq x \in X$. It only remains to be seen that $0 \notin \Gamma^1 x$.

Since X is separated there exists a convex neighborhood U in X containing x such that $\mu_U(x) > 0$, where μ_U is the Minkowski or gauge functional of U

$$\mu_U : X \to [0, \infty) : u \mapsto \inf\{\alpha \geq 0 \,|\, x \in \alpha U\}.$$

By uniform openness of Γ there exists a convex neighborhood V in Y such that $AU \supset V$ for each $A \in \Gamma$. Fix $A \in \Gamma$ and recall A is invertible.

Now consider the gauge functional of V at Ax:

$$\begin{aligned}
\mu_V(Ax) &\stackrel{\text{def}}{=} \inf\{\alpha > 0 \,|\, Ax \in \alpha V\} \\
&= \inf\{\alpha > 0 \,|\, x \in \alpha A^{-1} V\} \\
&\geq \inf\{\alpha > 0 \,|\, x \in \alpha U\} \\
&= \mu_U(x) > 0.
\end{aligned}$$

But $\mu_V = \sigma_{V^\circ}$ — this is well known and can be shown without difficulty using Theorem 2.1.2 — so we have $0 < \mu_U(x) \leq \sigma_{V^\circ}(Ax)$. This is valid for arbitrary elements of Γ, hence

$$\begin{aligned} 0 &< \inf_{A\in\Gamma} \sigma_{V^\circ}(Ax) \\ &= \inf_{y\in\Gamma x} \sigma_{V^\circ}(y) \\ &= \inf_{y\in\Gamma^1 x} \sigma_{V^\circ}(y), \end{aligned}$$

where the final equality follows from lower semicontinuity of σ_{V° ([Hör]) and Lemma 2.20. Clearly then $0 \notin \Gamma^1 x$.

$\square$

Lemma 2.20 *Let $\Gamma \subset \mathcal{L}$ be convex. For each $x \in X$ (respectively $\lambda \in Y^*$),*

$$\begin{gathered} \mathrm{cl}(\Gamma x) = \mathrm{cl}(\Gamma^1 x) \\ (\textit{ resp. } \quad \mathrm{w}^*\text{-}\,\mathrm{cl}(\lambda\Gamma) = \mathrm{w}^*\text{-}\,\mathrm{cl}(\lambda\Gamma^1) \quad). \end{gathered}$$

In particular if Γx (resp. $\lambda\Gamma$) is closed (resp. w^*-closed) then $\Gamma x = \Gamma^1 x$ (resp. $\lambda\Gamma = \lambda\Gamma^1$).

Proof We assume without loss of generality that Γ and its rank-1 representer are both nonempty sets. We have $\sigma_{\Gamma x}(\lambda) = \sigma^1_\Gamma(x,\lambda) = \sigma^1_{\Gamma^1}(x,\lambda) = \sigma_{\Gamma^1 x}(\lambda)$ for each $(x,\lambda) \in X \times Y^*$. Since Γ is assumed convex and Γ^1 is always convex, Theorem 2.1.2 says that the closures of Γx and $\Gamma^1 x$ are the same.

A similar argument, using Theorem 2.13.1 instead of Theorem 2.1.2, shows that the w*-closures of $\lambda\Gamma$ and $\lambda\Gamma^1$ are also identical.

$\square$

Our main result characterizes the rank-1 support functionals on $\mathcal{L}$. It would be reasonable to say that the properties characterizing rank-1 support functions of nonempty sets are "sublinearity and lower semicontinuity in $x\lambda$". Likewise the properties characterizing rank-1 support functions of nonempty sets that are equicontinuous in $\mathcal{F}^*$ are "sublinearity, subadditivity and boundedness in $x\lambda$". The boundedness condition needed is

Condition (A)

When $\mathcal{F} = \mathcal{F}_1$ and $\mathcal{L} = BL(X,Y)$: P is bounded above on $U \times C^\circ$ for some neighborhood U in X and nonempty bounded set C in Y.

or When $\mathcal{F} = \mathcal{F}_2$ and $\mathcal{L} = CL(X,Y)$: For each x in X and neighborhood V in Y, there are a nonempty bounded set C in Y and a neighborhood U in X such that P is bounded above on both $x \times C^\circ$ and $U \times V^\circ$.

or When Y is metrizable, $\mathcal{F} = \mathcal{F}_3$ and $\mathcal{L} = CL(X,Y)$: For each neighborhood V in Y, there is a neighborhood U in X such that P is bounded above on $U \times V^\circ$.

Theorem 2.21 *Let $P : X \times Y^* \to \mathbb{R} \cup \{\pm\infty\}$.*

1. *P is the rank-1 support function of a nonempty set in $\mathcal{L}$ iff $P(0,0) \neq -\infty$ and P is*

   (a) *positive bihomogeneous:*

   $$P(\alpha x, \lambda) = P(x, \alpha\lambda) = \alpha P(x,\lambda), \quad \forall \alpha \geq 0, (x,\lambda) \in X \times Y^*;$$

   (b) *subadditive and lower semicontinuous in $x\lambda$:*

   $$P(x,\lambda) \leq \liminf\{\sum P(x_i,\lambda_i) \mid \mathcal{F} \ni \sum x_i\lambda_i \to x\lambda\}.$$

   If P is the rank-1 support function of a nonempty set, there is a unique (nonempty) rank-1 representer $\Gamma \subset \mathcal{L}$ such that $P = \sigma^1_\Gamma$, and P never takes the value $-\infty$.

2. P is the rank-1 support function of a nonempty set in $\mathcal{L}$ that is equicontinuous in $\mathcal{F}^*$ iff $P(0,0) \neq -\infty$ and P is

(a) positive bihomogeneous;

(b) subadditive in $x\lambda$:

$$P(x,\lambda) \le \sum P(x_i,\lambda_i) \text{ when}$$

(c) satisfies condition (A).

If P is the rank-1 support function of a nonempty set which is equicontinuous in $\mathcal{F}^*$, then there is a unique (nonempty) w^*-compact rank-1 representer $\Gamma \subset \mathcal{L}$ such that $P = \sigma^1_\Gamma$, and P is finite valued.

Proof

1. Observe that

$$\alpha(x\lambda) = (\alpha x)\lambda = x(\alpha\lambda), \quad \forall (x,\lambda) \in X \times Y^*, \alpha \in \mathbb{R}; \tag{2.3}$$

whence $K \stackrel{\text{def}}{=} \{x\lambda \,|\, (x,\lambda) \in X \times Y^*\}$ is a cone in $FL(Y,X)$ containing the zero mapping. Moreover, for $\Gamma \subset \mathcal{L}$, we have

$$\sigma^1_\Gamma(x,\lambda) = \sigma_\Gamma(x\lambda), \quad \forall (x,\lambda) \in X \times Y^*,$$

where in the second support function, σ_Γ, Γ is identified with a subset of $\mathcal{F}^*$. So we have

$$\begin{aligned}
& \exists \text{ nonempty } \Gamma \subset \mathcal{L},\ \forall (x,\lambda) \in X \times Y^*, \quad P(x,\lambda) = \sigma^1_\Gamma(x,\lambda) \\
\iff & \exists \text{ nonempty } \Gamma \subset \mathcal{L},\ \forall (x,\lambda) \in X \times Y^*, \quad P(x,\lambda) = \sigma_\Gamma(x\lambda) \\
\iff & \text{the mapping } x\lambda \mapsto P(x,\lambda) \text{ on } K \text{ is the restriction to } K \text{ of a support} \\
& \text{function of a nonempty set in } \mathcal{L}
\end{aligned}$$

$\iff$ the mapping $p : K \to \mathbb{R} \cup \{\pm\infty\} : x\lambda \mapsto P(x,\lambda)$ is well defined, has $p(0) \neq -\infty$, and satisfies Theorem 2.14 1a and 1b

$\iff$ $P(0,0) \neq -\infty$ and P satisfies the properties 1a and 1b of the theorem.

The third equivalence is given by Theorem 2.14.1.

Uniqueness of Γ in the final statement also follows from Theorem 2.14.1 since, as noted previously, a rank-1 representer of a set Δ is the *largest* set having the same rank-1 support function as Δ.

2. As in part 1, but using Theorem 2.14.2 in place of Theorem 2.14.1.

□

Remark. When X and Y are both normed spaces, $BL(X,Y)$ is a normed space in the dual topology of $\mathcal{F}_1$ hence the classes of equicontinuous (in $\mathcal{F}_l^*$) and w*-compact sets in $BL(X,Y)$ are identical. In this case, part 2 of the theorem characterizes rank-1 support functionals of w*-compact sets in $BL(X,Y)$.

2.4 Application to Fans

We apply rank-1 support functions to Ioffe's *fans* ([Iof81, Iof82]). A fan is a set valued mapping which, in some important ways, generalizes a (sub)linear operator.

Definition 2.22 *Let F be a set valued mapping from X to Y.*

1. *F is called a* fan *if it*

   (a) *takes nonempty, convex values: $F(x)$ is nonempty and convex, for $x \in X$;*

   (b) *is positive homogeneous: $F(\alpha x) = \alpha F(x)$ for $\alpha > 0, x \in X$; and*

   (c) *is subadditive: $F(x_1 + x_2) \subset \mathrm{cl}(F(x_1) + F(x_2))$ for $x_1, x_2 \in X$.*

2. *The* value closure *of F is the set mapping $\bar{F}$ given by $\bar{F}(x) \stackrel{\text{def}}{=} \text{cl}[F(x)]$ for each $x \in X$.*

3. *The* handle *of F is the set*

$$h(F) \stackrel{\text{def}}{=} \{A \in L(X,Y) \mid \lambda Ax \leq \sup_{y \in F(x)} \lambda y, \quad \forall (x,\lambda) \in X \times Y^*\}.$$

F is said to be spanned by its handle *if $F(x) = h(F)x$, for $x \in X$.*

F is spanned by its handle up to closed values *if $\bar{F}(x) = \text{cl}[h(F)x]$, for $x \in X$.*

The *graph* of a set valued mapping (or multifunction) F from X to Y is $\{(x,y) \mid y \in F(x)\}$. We note in passing that the value closure of F is not necessarily the same as the multifunction whose graph is defined as the closure of the graph of F.

The property of a fan's being spanned by its handle (up to closed values) is of special interest because it means the fan is generated by a rank-1 representer, namely its handle. The question of characterization of fans spanned by their handles has been open since at least [AV]. Here we give a characterization of the fans which are spanned by their handles up to closed values. In particular, of the fans with w*-compact handles, we can characterize those which are spanned by their handles.

We recall the terminology of Theorem 2.1: let the support function of a nonempty set V in Y be given by

$$\sigma_V : Y^* \to \mathbb{R} \cup \{\infty\} : \lambda \to \sup_{y \in V} \lambda y.$$

Theorem 2.23 *Let F be a set mapping from X to Y with nonempty values. Define $P : X \times Y^* \to \mathbb{R} \cup \{\infty\}$ for each (x,λ) in $X \times Y^*$ by*

$$P(x,\lambda) \stackrel{\text{def}}{=} \sigma_{F(x)}(\lambda).$$

1. *F is spanned by its handle up to closed values iff $\bar{F}$ is a fan and P satisfies the hypotheses 1a, 1b of Theorem 2.21.*

2. *The handle of F is equicontinuous in $\mathcal{F}^*$ and spans F iff $F = \bar{F}$ is a fan and P satisfies the hypotheses 2a-2c of Theorem 2.21.*

Proof

1. Recall Theorem 2.1.2 which states that nonempty convex sets V, W in Y have the same closure iff their support functions are identical:

$$\operatorname{cl} V = \operatorname{cl} W \iff \sigma_V = \sigma_W.$$

   Also as $h(F)$ is a rank-1 representer, (it and) its images $h(F)x$ are convex. So

$$\begin{aligned} & F \text{ is spanned by its handle up to closed values} \\ \iff \quad & \sigma_{F(x)} = \sigma_{h(F)x} \text{ for each } x \in X \\ \iff \quad & P \text{ is the rank-1 support function of } h(F) \\ \iff \quad & P \text{ satisfies Theorem 2.21 1a, 1b.} \end{aligned}$$

2. The proof follows as above using Theorem 2.21.2 instead of Theorem 2.21.1.

□

Corollary 2.24 *Let F be a fan from X to Y and P be defined as in the theorem.*

1. *Suppose the handle of F is w^*-compact. Then F is spanned by its handle iff $F = \bar{F}$ and P satisfies the hypotheses 1a, 1b of Theorem 2.21.*

2. *Suppose the handle of F is equicontinuous in $\mathcal{F}^*$. Then F is spanned by its handle iff $F = \bar{F}$ and P satisfies the hypotheses 2a-2c of Theorem 2.21.*

A prototype fan comes from the analysis of nondifferentiable vector functions. We will sketch it here in the context of normed spaces X, Y with Y reflexive.

Let $g : X \to Y$ be Lipschitz in a neighborhood U of $x_* \in X$. Given any λ in Y^* notice that the composition λg is a real function on X, Lipschitz in U. The [Cla] *generalized directional derivative* of λg at x_* in the direction $u \in X$ is

$$(\lambda g)^\circ(x_*; u) \stackrel{\text{def}}{=} \limsup_{\substack{x \to x_* \\ t \downarrow 0}} \frac{(\lambda g)(x + tu) - (\lambda g)(x)}{t}.$$

It can be shown easily that for fixed $u \in X$, the function on Y^*

$$\lambda \mapsto (\lambda g)^\circ(x_*; u)$$

is sublinear; hence, by the result dual to Theorem 2.13.2, it is the support function of a nonempty set in Y. This leads to the definition of the set mapping F from X to Y as

$$F(u) \stackrel{\text{def}}{=} \{y \in Y \mid \lambda y \leq (\lambda g)^\circ(x_*; u),\ \forall \lambda \in Y^*\}, \quad \forall u \in X. \tag{2.4}$$

We know already that F has nonempty convex values in Y. With more work in a similar vein, we can show that F is also subadditive and positive homogeneous, hence that F is a fan.

The handle of F in this special case is called the rank-1 generalized Jacobian of g at x_*, denoted $\partial^1 g(x_*)$. This set of continuous linear mappings from X to Y is used as a kind of Jacobian for g at x_* even when the classical Jacobian for g does not exist at x_* (see Chapter 3).

Although F is a fan, it is not obvious that F is actually spanned by its handle $\partial^1 g(x_*)$. From the Lipschitz property of g though we deduce that $\partial^1 g(x_*)$ is at least equicontinuous, so that Theorem 2.23.2 can be brought to bear on this question. Exactly this reasoning is used in Chapter 3 to obtain the following positive result.

Theorem 2.25 *Let X, Y, g be as above and F be the set mapping defined by (2.4). Then F is a fan spanned by its handle:*

$$\{y \in Y \mid \lambda y \leq (\lambda g)^{\circ}(x_*; u), \ \forall \lambda \in Y^*\} = \partial^1 g(x_*)u, \quad \forall u \in X.$$

Extensions of this result to separated locally convex spaces are given in Chapter 3.

2.5 Rank-1 and Plenary sets

We begin by defining rank-1 and plenary sets. The former property is derived from rank-1 representers, specifically the property 3 as listed after Definition 2.15. The latter very similar idea is first found in [Swe77] and attributed to Hubert Halkin. We will make frequent reference to [Swe79] in which a number of properties of plenary sets are displayed. Plenary sets are the same as the *solid* sets of [Rub].

Definition 2.26 *Let $\Gamma \subset \mathcal{L}$.*

1. Γ *is* rank-1 *if for each $A \in \mathcal{L}$*

$$\forall (x, \lambda) \in X \times Y^*, \ \lambda A x \in \lambda \Gamma x \implies A \in \Gamma.$$

The rank-1 hull *of Γ is*

$$\text{rank-1}\Gamma \stackrel{\text{def}}{=} \{A \in \mathcal{L} \mid \forall (x, \lambda) \in X \times Y^*, \ \lambda A x \in \lambda \Gamma x\}.$$

2. Γ *is* plenary *if for each $A \in \mathcal{L}$*

$$\forall x \in X, \ Ax \in \Gamma x \implies A \in \Gamma.$$

The plenary hull *of Γ is*

$$\text{plen}\Gamma \stackrel{\text{def}}{=} \{A \in \mathcal{L} \mid \forall x \in X, \ Ax \in \Gamma x\}.$$

It may be more appropriate to call a rank-1 sets Γ *dually rank-1*, because such sets are not usually composed of rank-1 operators in $L(X, Y)$, rather are dually described by rank-1 operators in $L(Y, X)$.

The rank-1 property depends not only on the set Γ, but also on the superspace $\mathcal{L}$ of Γ and on the dual space $FL(Y, X)$ of $\mathcal{L}$. The definition could equally be given with respect to a superspace of Γ other than $\mathcal{L}$ or in fact to any subset of such a space, and, likewise, using any subset of the space dual to $\mathcal{L}$. A similar statement, and a more general definition on these lines, can be made for the plenary property — see [Swe79, Ch. 4].

The rank-1 and plenary properties coincide when applied to sets in the dual of an 'ordinary' convex space: if Φ is a subset of E^*, it would be rank-1 if

$$(\phi e)\alpha \in (\Phi e)\alpha,\ \forall (e, \alpha) \in E \times \mathbb{R} \implies \phi \in \Phi$$

or, equivalently,

$$(\phi e) \in \Phi e,\ \forall e \in E \implies \phi \in \Phi$$

which is the definition of plenary.

The following relationship between convex and rank-1 or plenary sets is given by [Swe79, Prop. 4.03]:

Proposition 2.27 *Let E^* have its weak* topology. If Φ is a closed set in E^*, then it is convex iff it is pathwise connected and either rank-1 or plenary.*

It is easy to check that the rank-1 (respectively plenary) hull of a set is a rank-1 (resp. plenary) set. Also

Proposition 2.28 *Let Γ be a set in $\mathcal{L}$.*

1. *If Γ is rank-1 then it is also plenary. The converse holds if Γ is w^*-compact, but not in general.*

2. plen$\Gamma \subset$ rank-1Γ. *Equality holds if* Γ *is w*-compact, but not in general.*

Proof

1. Suppose Γ is rank-1. If $Ax \in \Gamma x$ for each x, then $\lambda A x \in \lambda \Gamma x$ for each (x, λ) hence $A \in \Gamma$, i.e. Γ is plenary. Following the proof is a counterexample to the statement of the converse.

   Now suppose Γ is w*-compact, convex and plenary. Then, for each x, Γx is w-closed and convex. If $\lambda A x \in \lambda \Gamma x$ for each (x, λ), then, by Theorem 2.13.1, $Ax \in \Gamma x$ for each x. Therefore, plenarity of Γ implies A belongs to Γ, i.e. Γ is rank-1.

2. Similar to part 1.

$\square$

Example 2.29 *Let*

$$G \stackrel{\text{def}}{=} \{(a,b)^T \in \mathbb{R}^2 \mid (a_+)^2 \leq b \text{ and } (a,b) \neq (0,0)\},$$

where $a_+ \stackrel{\text{def}}{=} \max\{a, 0\}$. *We have*

$$0 \in v^T G, \quad \forall v = (v_1, v_2)^T \in \mathbb{R}^2. \tag{2.5}$$

If either v_1 *or* v_2 *is zero this is clear; otherwise take* $x = (-v_1/v_2, (v_1/v_2)^2) \in G$ *to obtain* $v^T x = 0$.

Define Γ *as the set of matrices in* $\mathbb{R}^{2\times 2}$ *whose columns lie in* G. *Clearly*

$$(0,0)^T \notin G = \Gamma(1,0)^T$$

so plenΓ *does not contain the zero matrix. However using (2.5) we get*

$$0 \in v^T \Gamma u, \quad \forall u, v \in \mathbb{R}^2$$

so the zero matrix lies in rank-1Γ.

Thus plen$\Gamma \neq$ rank-1Γ.

We characterize rank-1 representers in terms of rank-1 and plenary sets in $\mathcal{L}$, and also using sets of operators in $\mathcal{F}$ of rank no greater than 1. We will take the polars of nonempty sets $Q \in \mathcal{F}$ in $\mathcal{L}$:

$$Q^\circ \stackrel{\text{def}}{=} \{A \in \mathcal{L} \mid \tilde{A}T \leq 1, \forall T \in Q\}.$$

Proposition 2.30 *Let Γ be a set in $\mathcal{L}$.*

1. $\Gamma^1 = \tau\text{-cl rank-1 co}\,\Gamma$, *where τ is any topology which is finer than the w^*-topology and makes $\mathcal{L}$ a topological vector space.*

   If Γ has nonempty strong interior or is w^*-compact, then the rank-1 hull operation may be replaced by the plenary hull operation.

2. *If* $0 \in \Gamma^1$, $\Gamma^1 = \{x\lambda \mid \sigma_\Gamma^1(x,\lambda) \leq 1\}^\circ$.

Proof

1. Let Δ be the convex hull of Γ. Since Γ^1 contains Γ and is w*-closed, rank-1 and convex, then Γ^1 contains w*-cl rank-1Δ. Moreover w*-cl rank-1$\Delta \supset$ τ-cl rank-1Δ because the τ-topology is finer than the w*-topology. It is only left to show that $\Gamma^1 \subset \tau$-cl rank-1Δ.

   So choose A in Γ^1 and B in Δ. Then for any (x, λ),

$$\begin{aligned} &\lambda A x \in \text{cl}(\lambda \Gamma x) \subset \text{cl}(\lambda \Delta x), \\ &\lambda B x \in \lambda \Delta x. \end{aligned}$$

   Since $\lambda\Delta x$ is a real interval, it is easy to show that for $0 \leq \alpha < 1$

$$\alpha\lambda A x + (1-\alpha)\lambda B x \in \lambda \Delta x.$$

   This holds for each (x, λ) so $\alpha A + (1-\alpha)B \in$ rank-1Δ. Letting $\alpha \uparrow 1$ yields $A \in \tau$-cl Δ. Hence $\Gamma^1 \subset \tau$-cl rank-1Δ and we are done.

The claim about taking plenary instead of rank-1 hulls is shown by [Swe79, Prop. 4.24, 4.30]. The proof — in the setting of normed spaces — are still valid here.

2. Suppose $0 \in \Gamma^1$. Let Q be the set of $x\lambda$ such that $\sigma_\Gamma^1(x,\lambda) \le 1$. Clearly $\Gamma^1 \subset Q^\circ$.

   Now take $A \notin \Gamma^1$, so $\lambda A x > \sigma_\Gamma^1(x,\lambda)$ for some (x,λ). Since $0 \in \Gamma^1$ (and the rank-1 support functions of Γ and Γ^1 are equal) we may scale by a positive number α such that

   $$\alpha\lambda A x > 1 \ge \sigma_\Gamma^1(x,\alpha\lambda) \ (\ge 0 \ ),$$

   i.e. $A \notin Q^\circ$. Thus $\Gamma^1 \supset Q^\circ$ as needed.

□

Related to part 1 of the proposition is the open question of whether or not the w*-closure of a plenary convex set is a rank-1 representer ([Swe79, Ch. IV, p. 53]).

The above proposition and many of its corollaries to follow may seem, at first, to be operator analogs of important and difficult results in convex analysis. For example, compare Proposition 2.30.1 with Theorem 2.13.1. However, let us rewrite Proposition 2.30.1 in the special case that $\mathcal{L}$ is the dual space E^*, τ is the w*-topology, and $\Gamma = \Phi \subset E$:

$$\text{w*-}\,\mathrm{cl}\{\phi \mid \phi e \in (\mathrm{co}\,\Phi)e,\ \forall e\} \ = \ \{\phi \mid \phi e \le \sigma_\Phi(e),\ \forall e\}. \tag{2.6}$$

This result and the proposition are not nearly so important as Theorem 2.13.1. Likewise, the proofs are trivial by comparison.

Nevertheless, the above proposition and its corollaries are of some interest. For example, (2.6) still holds with any topology τ making E^* a topological vector space such that τ is finer than the w*-topology. This observation and Theorem 2.13.1 gives

$$\tau\text{-cl}\ \{\phi \mid \phi e \in (\mathrm{co}\,\Phi)e,\ \forall e\} = \text{w*-cl}\ \ \mathrm{co}\,\Phi$$

in which our freedom in choosing τ is unusual. Also, as already noted, for this choice of $\mathcal{L}$ the rank-1 hull operation can be replaced with the plenary hull operation.

Corollary 2.31 *The strong or w*- closure of a rank-1 convex set in $\mathcal{L}$ is still a rank-1 (and convex) set.*

Proof According to part 1 of the proposition, if Γ is rank-1 and convex then its strong or w*- closure is Γ^1. Γ^1 is, of course, a rank-1 convex set. □

Another open question is whether or not the strong or w*- closure of a convex plenary set is also plenary ([Swe79, Ch. IV, p. 50]).

Corollary 2.32 *Let Γ be a set in $\mathcal{L}$.*

1. *Γ is a rank-1 representer iff it is strongly or w*- closed, rank-1 and convex. If Γ has nonempty strong interior, or is w*-compact, then the rank-1 property may be replaced by plenarity.*

2. *Γ is a rank-1 representer containing zero iff it is the polar of $\{x\lambda \mid \sigma^1_\Gamma(x,\lambda) \leq 1\} \neq \emptyset$ iff it is the polar of a nonempty set in $\mathcal{F}$ which contains operators of rank not greater than 1.*

Proof The only possible question is in part 2, specifically the claim that $\Gamma = Q^\circ$ implies $\Gamma = \Gamma^1$, where Q is a nonempty set in $\mathcal{F}$ containing operators of rank not greater than 1. This follows from part 1 since Q° is w*-closed, rank-1 and convex. □

From Corollary 2.32.2 we deduce that, in Definition 2.6, polars of sets in the bases of neighborhoods in $FL(Y, X)$ are rank-1 representers.

We come to separation of points and sets in $\mathcal{L}$.

Corollary 2.33 *Let Γ be a set in $\mathcal{L}$. If Γ is a strongly or w^*- closed, rank-1, convex set in $\mathcal{L}$ then*

$$A \notin \Gamma \iff \exists(x, \lambda),\ \lambda A x > \sigma^1_\Gamma(x, \lambda).$$

The rank-1 property may be replaced by plenarity if Γ has nonempty strong interior or is w^*-compact.

To obtain separation of rank-1 sets Γ and Δ via rank-1 operators $x\lambda$, the difference $\Gamma - \Delta \stackrel{\text{def}}{=} \{A - B | A \in \Gamma, B \in \Delta\}$ must also be rank-1. Unfortunately this need not follow even if Γ and Δ are also w*-compact and convex. [Swe79, App. A.05] provides an example of (w*-)compact, convex, plenary sets $\Gamma, \Delta \subset \mathbb{R}^2$ whose difference is not plenary. Applying Lemma 2.28.1 we find that Γ and Δ are also rank-1 sets whose difference is not rank-1.

Here are some elementary properties of rank-1 sets.

Proposition 2.34

1. *Intersections of rank-1 (respectively plenary) sets in $\mathcal{L}$ are rank-1 (plenary).*

2. *The rank-1 (respectively plenary) hull of a set in $\mathcal{L}$ is equal to the intersection of all rank-1 (plenary) sets which contain it.*

3. *The rank-1 and plenary hulls of a convex (respectively w^*-compact) set in $\mathcal{L}$ are convex (w^*-closed).*

4. *Compositions of invertible operators with w^*-compact rank-1 sets yield rank-1 sets: Let Γ be a w^*-compact rank-1 set in $CL(X, Y)$, and Z be a separated*

convex space. If A is a continuously invertible operator in $CL(Y,Z)$ then $A \circ \Gamma \stackrel{\text{def}}{=} \{A \circ G | G \in \Gamma\}$ is rank-1 in $CL(X,Z)$. Likewise, $\Gamma \circ B$ is rank-1 in $CL(Z,Y)$ if B is an invertible mapping in $CL(Z,X)$.

Similarly, compositions of invertible operators with (not necessarily w^*-compact) plenary sets yield plenary sets.

5. *Scaling and translation of rank-1 (plenary) sets yields rank-1 (plenary) sets.*

Proof In the plenary case, most of these results are shown in [Swe79, Ch. VI 4.10, 4.13, 4.16, 4.20]. In the rank-1 case, proofs are nearly identical (and are not, in any case, difficult). $\square$

Chapter 3

The Rank-1 Generalized Jacobian

3.1 Introduction

Let X, Y be real Banach spaces with respective continuous duals X^*, Y^*; and $BL(X, Y)$ be the space of bounded (continuous) linear mappings from X to Y. Let $f : X \to \mathbb{R}$ and $g : X \to Y$ be locally Lipschitz mappings.

Even when f is not differentiable, the Clarke generalized gradient [Cla] is useful in specifying first order information for f. The generalized gradient of f at $x_* \in X$ is a convex set defined using the Clarke generalized directional derivative $f^\circ(x_*; \cdot)$:

$$\partial f(x_*) \stackrel{\text{def}}{=} \{\xi \in X^* \mid \forall u \in X,\ \xi u \leq f^\circ(x_*; u)\}$$

This definition and the proof of existence (nonemptiness) of $\partial f(x_*)$ are given analytically in [Cla], i.e. only with reference to topology, using the theory of support functionals of convex sets ([Hör]). Rockafellar has investigated Clarke calculus for non-Lipschitzian functions (eg. [Roc79, Roc80]); such generality in terms of vector functions is, however, beyond the scope of this thesis.

Analogously we consider the set

$$\partial^1 g(x_*) \stackrel{\text{def}}{=} \{A \in BL(X,Y) \mid \forall (u,\lambda) \in X \times Y^*,\ \lambda A u \le (\lambda g)^\circ(x_*;u)\},$$

as in [Swe77] in finite dimensions, to try to extend the classical notion of a Jacobian to the nonsmooth function g at x_*. Another candidate (see Proposition 3.8.2) is the Clarke generalized Jacobian,

$$\partial g(x_*) \stackrel{\text{def}}{=} \operatorname{cl} \operatorname{co} \{\lim \nabla g(x_i) \mid x_i \to x_*, \nabla g(x_i) \text{ exists}\},$$

where the weak-operator, or weak*, topology of $BL(X,Y)$ is used for infinite dimensions. Existence for the latter requires Rademacher's theorem [Chr], a measure theoretic result which says that g is Gâteaux differentiable almost everywhere with respect to Haar measure if X and Y are separable and Y is reflexive.

Until now, existence of $\partial^1 g(x_*)$ has relied entirely on existence of $\partial g(x_*)$, hence has been limited to separable Banach spaces with Y reflexive ([Swe77, Thi82]). The principal result of this chapter, Theorem 3.15.1, shows existence of $\partial^1 g(x_*)$ for arbitrary normed spaces so long as Y remains reflexive. The existence proof is also novel in the context of the generalized Jacobian: it uses a characterization from Chapter 2 of support functions of *rank-1 representers* in $BL(X,Y)$ similar to the characterization of support functions used for existence of the generalized gradient. We do not require a partial ordering on Y to obtain vector support functionals, as used in [Thi80]. Unlike the existence proof of the generalized gradient, however, Theorem 3.15.1 still needs the finite dimensional (classical) version of Rademacher's theorem. This raises an important question: can Theorem 3.15.1 can be proven in an entirely analytic manner?

Sweetser [Swe77] shows in finite dimensions that $\partial^1 g(x_*)$ is the *plenary* hull of $\partial g(x_*)$ (see also [H-U82]); in infinite dimensions this is given by Thibault [Thi82]. It happens that the plenary and *rank-1* hulls of $\partial g(x_*)$ coincide, so $\partial^1 g(x_*)$ is also the rank-1 hull of $\partial g(x_*)$, when the latter exists. We call $\partial^1 g(x_*)$ the rank-1 generalized Jacobian rather than the plenary generalized Jacobian because its existence, independent of $\partial g(x_*)$, depends on rank-1 support functions.

Nonemptiness of the rank-1 generalized Jacobian is of some interest, even with the proviso that Y be reflexive, given the statement of [H-U82, §3] (also [Thi82, §4]) that 'the mere question of existence [of the rank-1 generalized Jacobian] is hopeless for X and Y general Banach spaces'.

When Y is not reflexive, if we are prepared to consider g as a mapping from X to Y^{**} by embedding Y in its second dual, existence of a rank-1 generalized Jacobian is still assured. Moreover, our existence result still goes through when we use metric spaces instead of normed spaces. In fact if we posit an extra Lipschitz-like condition on g, which is superfluous when Y is metrizable, we can even extend existence to the setting of separable, locally convex topological vector spaces with semi-reflexive Y. (See §3.7.)

Other notable approaches to getting first order information on nonsmooth vector functions, mainly in finite dimensions, include Halkin's *screens* [Hal], Warga's *derivate containers* [War], Sweetser's *shields* [Swe77, Swe79], the generalized Jacobian of Pourciau [Por], and Ioffe's *fans* [Iof81, Iof82]. The first three of these deal with specifications of families of linear mappings related to a general (nonsmooth) function, and properties of these families which guarantee extensions of well known results in smooth functional analysis such as the inverse function theorem. For more on Pourciau's contribution see the discussion after Proposition 3.11. Regarding fans, the prototype seems to be the set mapping from X to Y, $F : u \mapsto \{y \mid \lambda y \leq (\lambda g)^\circ(x_*; u),\ \forall \lambda \in Y^*\}$ which we discuss in §3.5. It turns out that this fan is always generated by the rank-1 generalized Jacobian: $F(u) = \partial^1 g(x_*)u$ (Proposition 3.25).

The remainder of the chapter is organized as follows.

§3.2 Notation and preliminary results, mostly dealing with topology and (rank-1) support functionals of convex sets.

§3.3 Review of the Clarke generalized gradient and generalized Jacobian. The only result crucial to later material will be (the finite dimensional version of) Proposition 3.11, which justifies the generalized Jacobian. Other results are presented primarily for comparison with new results, to follow.

§3.4 The rank-1 generalized Jacobian: existence and basic properties.

§3.5 Relationships between the rank-1 generalized Jacobian and the generalized gradient (when $Y = \mathbb{R}$), the generalized Jacobian (when it exists), Sweetser's shields ([Swe77], [Swe79]), classical derivatives, and Ioffe's fans ([Iof81], [Iof82]).

§3.6 Some calculus for the rank-1 generalized Jacobian, including a chain rule and a mean value theorem.

§3.7 Discussion of extensions to include the cases when Y is not reflexive, and when X, Y are spaces more general than normed spaces.

We note that, apart from the existence proof Theorem 3.15.1 and related material, the ideas behind many of the results given here are well known in less generality. In particular, we recommend Hiriart-Urruty's account [H-U82] in finite dimensions of the (rank-1) generalized Jacobian. Also, to a certain extent, the results and some proofs presented in sections 3.4-3.6 are modeled after the early part of [Cla, Ch. 2].

3.2 Notation and Preliminary Results

We present the notation we will assume throughout this chapter.

- Let X, Y be normed spaces over $\mathbb{R}$ with open unit balls B_X, B_Y respectively. Y is assumed to be reflexive.

- A *form* on X is a linear mapping from X to $\mathbb{R}$; an *operator* from X to Y is a linear mapping from X to Y. The space of all bounded operators A from X to Y is written $BL(X, Y)$, and is endowed with the operator norm $\|A\| \stackrel{\text{def}}{=} \sup\{\|Ax\| \; x \in B_X\}$. The *dual space* of X is $BL(X, \mathbb{R})$ — where $\mathbb{R}$ has the usual Euclidean topology — the normed space of all continuous forms on X. This is denoted by X^*, and its open unit ball by B_X^*.

- The space of bounded finite rank operators from Y to X is denoted $FL(Y, X)$; of special interest will be operators of rank one or less, given by

$$u\lambda : Y \to X : y \mapsto u(\lambda y)$$

  for $(u, \lambda) \in X \times Y^*$. $FL(Y, X)$ is endowed with a norm (different from the operator norm) under which its dual will be $BL(X, Y)$ — see Proposition 3.1, parts 3 and 4, below.

- The weak topology of Y determined by Y^*, or $\sigma(Y, Y^*)$ topology, will be denoted by w-. The weak*, or $\sigma(X^*, X)$, topology of X^* will be denoted by w*-. (See [Sch].)

  Likewise the weak operator, or $\sigma(BL(X, Y), FL(Y, X))$, topology on $BL(X, Y)$ will be denoted by w*- (see Proposition 3.1.4 below).

- The set operations of closure and convex hull will be denoted by cl and co respectively.

- Let $f : X \to \mathbb{R}$ and $g : X \to Y$.

We quote from Chapter 2 §2.2.

Proposition 3.1

1. *Given* $(u, \lambda) \in X \times Y^*$, *define* $u\lambda : Y \to X : y \mapsto u(\lambda y)$. *Then*

$$F \in FL(Y, X) \iff F = \sum u_i \lambda_i$$

for some finite sequence $((u_i, \lambda_i)) \subset X \times Y^*$. *In particular,* F *is a bounded rank-1 operator from* Y *to* X *iff* $F = u\lambda$ *for some* $u \neq 0, \lambda \neq 0$.

2. *Let* $(u, \lambda), (u_1, \lambda_1), \ldots, (u_m, \lambda_m) \in X \times Y^*$, *where* $m \in \mathbb{N}$, *and* $A \in BL(X, Y)$. *If* $u\lambda$ *agrees with* $\sum u_i \lambda_i$ *on the range of* A, *i.e.*

$$\forall x \in X, \quad u(\lambda A x) = \sum u_i(\lambda_i A x),$$

then

$$\lambda A u = \sum \lambda_i A u_i.$$

3. *A norm on* $FL(Y, X)$ *is given by*

$$\|F\| = \inf \left\{ \sum \|u_i\| \|y_i\| \mid \forall \textit{ finite sums } \sum u_i \lambda_i = F, \ ((u_i, \lambda_i))_i \subset X \times Y^* \right\}$$

for each $F \in FL(Y, X)$. *In particular,* $\|u\lambda\| = \|u\| \|\lambda\|$.

4. *Consider each operator* $A \in BL(X, Y)$ *as a form on* $FL(Y, X)$ *by*

$$A : \sum u_i \lambda_i \mapsto \sum \lambda_i A u_i.$$

Then $BL(X, Y)$ *is the dual of* $FL(Y, X)$ *such that the dual norm and the operator norm coincide.*

The typical setting for these results is in terms of the tensor product $X \otimes Y^*$ rather than (the isomorphic space) $FL(Y, X)$; in fact the norm on $X \otimes Y^*$ corresponding to the above norm on $FL(Y, X)$ arises from the well known *projective* topology on the tensor product ([Sch, Ch. 3 §6]). We will avoid the tensor product and the extra notation this requires.

Support functions corresponding to w*-closed, convex sets in X^* are of great importance in functional analysis. When using a support function of a w*-closed convex set in $BL(X,Y)$, we will find it convenient to restrict the function domain to the rank-1 members of $FL(Y,X)$. By convention the supremum over an empty set will be taken as $-\infty$.

Definition 3.2

1. *Let $C \subset X^*$. The* support functional *of C, $\sigma_C : X \to \mathbb{R} \cup \{\infty, -\infty\}$, is given by*

$$\sigma_C(u) \stackrel{\text{def}}{=} \sup_{\xi \in C} \xi u, \quad \forall u \in X.$$

2. *Let $\Gamma \subset BL(X,Y)$. The* rank-1 *support functional of Γ, $\sigma^1_\Gamma : X \times Y^* \to \mathbb{R} \cup \{\infty, -\infty\}$, is given by*

$$\sigma^1_\Gamma(u,\lambda) \stackrel{\text{def}}{=} \sup_{A \in \Gamma} \lambda A u, \quad \forall (u,\lambda) \in X \times Y^*.$$

In the 3-dimensional case, the ideas in part 1 of the next result can be traced back to Minkowski's 1911 paper [Min]. Part 2 is due to Hörmander [Hör] who considers support functions in the general setting of locally convex vector spaces.

Function inequalities such as $\sigma_C \leq \sigma_D$, for subsets C, D of X^*, are taken pointwise. By Cu we mean the set $\{\xi u | \xi \in C\}$, for $C \subset X^*$ and $u \in X$.

Theorem 3.3 *Let C, D be w*-closed, convex sets in X^*.*

1. $C \subset D \iff \sigma_C \leq \sigma_D \iff \forall u \in X,\ Cu \subset Du.$
   By symmetry, the statement is true if equality holds in place of the subsets and inequality.

2. *Let $p : X \to \mathbb{R}$. Then p is the support function of a nonempty, w^*-compact, convex set C in X^* iff p is*

   (a) *positive homogeneous: $p(\alpha u) = \alpha p(u)$ for each $\alpha > 0$ and $u \in X$;*

   (b) *subadditive: $p(u_1 + u_2) \leq p(u_1) + p(u_2)$ for each $u_1, u_2 \in X$;*

   (c) *bounded above on B_X.*

   Such a set C is unique.

We deduce from Theorem 3.3.1 that the w*-closed convex sets C in X^* are characterized by their support functions:

$$C = \{\xi \in X^* \,|\, \xi u \leq \sigma_C(u), \ \forall u \in X\}.$$

This useful property motivates us to define the rank-1 representers in $BL(X,Y)$ as those sets of bounded operators which are characterized by their rank-1 support functions. Also relevant are the rank-1 sets (from Chapter 2 §2.5) and Sweetser's plenary sets [Swe77, Swe79].

Definition 3.4 *Let $\Gamma \subset BL(X,Y)$.*

1. *The* rank-1 representer *of Γ is*

$$\Gamma^1 \stackrel{\text{def}}{=} \{A \in BL(X,Y) \mid \lambda A u \leq \sigma^1_\Gamma(u,\lambda), \ \forall (u,\lambda) \in X \times Y^*\}.$$

   Γ is said to be a rank-1 representer *if $\Gamma = \Gamma^1$.*

2. *Γ is* rank-1 *if for each $A \in BL(X,Y)$,*

$$\forall (u,\lambda) \in X \times Y^*, \ \lambda A u \in \lambda \Gamma u \implies A \in \Gamma.$$

   The rank-1 hull *of Γ is*

$$\text{rank-1}\Gamma \stackrel{\text{def}}{=} \{A \in BL(X,Y) \mid \lambda A u \in \lambda \Gamma u, \forall (u,\lambda) \in X \times Y^*\}.$$

3. Γ *is* plenary *if for each* $A \in BL(X,Y)$,

$$\forall u \in X,\ Au \in \Gamma u \implies A \in \Gamma.$$

The plenary hull *of* Γ *is*

$$\mathrm{plen}\Gamma \stackrel{\mathrm{def}}{=} \{A \in BL(X,Y) \mid Au \in \Gamma u, \forall u \in X\}.$$

Note Γ^1 is a rank-1 set. We quote several other properties of rank-1 sets from Chapter 2, of which the most important is the characterization of rank-1 representers in terms of rank-1 hulls.

Proposition 3.5 *Let* Γ *be a set in* $BL(X,Y)$.

1. $\Gamma^1 = \mathrm{w}^*\text{- cl rank-1 co}\,\Gamma$.

2. *If* Γ *is* w^**-compact and convex, then* $\Gamma u = \Gamma^1 u$ *for* $u \in X$ *and* $\lambda\Gamma = \lambda\Gamma^1$ *for* $\lambda \in Y^*$.

3. *If* Γ *is rank-1 then it is also plenary. The converse holds if* Γ *is* w^**-compact, but not in general.*

Proof

1. Chapter 2 Proposition 2.30.1.

2. Corollary of Chapter 2 Lemma 2.20.

3. Chapter 2 Proposition 2.28.1.

$\square$

Recall from Chapter 2 Theorem 2.21.2 the following rank-1 version of Theorem 3.3.

Theorem 3.6 *Let $P : X \times Y^* \to \mathbb{R}$. Then P is the rank-1 support function of a nonempty, w^*-compact, rank-1 representer Γ in $BL(X, Y)$ iff P is*

1. *positive bihomogeneous: $P(\alpha u, \lambda) = P(u, \alpha\lambda) = \alpha P(u, \lambda)$ for each $\alpha > 0$ and $(u, \lambda) \in X \times Y^*$;*

2. *subadditive in $u\lambda$: $P(u, \lambda) \leq \sum P(u_i, \lambda_i)$ for all finite sums $\sum u_i\lambda_i = u\lambda$;*

3. *bounded above on $B_X \times B_Y^*$.*

Such a set Γ is unique.

3.3 Review: Basics of the Clarke Calculus

We will recall, in infinite dimensions, Clarke's extension of the classical gradient of real functions, and the extension of the classical Jacobian of vector functions due (in greatest generality) to Thibault. Most of the following results involving measure were first shown, in finite dimensions using Lebesgue measure instead of Haar measure, by Clarke. When derivatives are used here, they will be Gâteaux derivatives.

Definition 3.7 *(Clarke)*
The generalized directional derivative *of f at x_* in the direction $u \in X$ is*

$$f^\circ(x_*; u) \stackrel{\text{def}}{=} \limsup_{\substack{x \to x_* \\ t \downarrow 0}} \frac{f(x + tu) - f(x)}{t}.$$

The generalized gradient *of f at x_* is the set*

$$\partial f(x_*) \stackrel{\text{def}}{=} \{\xi \in X^* \mid \forall u \in X,\ \xi u \leq f^\circ(x_*; u)\}.$$

The following theorem justifies the generalized gradient and shows it relation to gradients of f, when these exist. Part 1 relies on the characterization of support functions given by Theorem 3.3.2, whereas part 2 uses integration with respect to Haar measure. A fact of interest noted in [Thi82] is that a finite dimensional set is Haar-null iff it has Lebesgue measure zero.

Proposition 3.8 *Let X be a Banach space, $x_* \in X$, and f be Lipschitz near x_*.*

1. *(Clarke) The generalized directional derivative $f^\circ(x_*;\,\cdot)$ is finite valued, positive homogeneous, subadditive and bounded; hence is the support function of the nonempty, w*-compact, convex set $\partial f(x_*)$.*

2. *(Thibault) If X is also separable, then $f^\circ(x_*;\,\cdot)$ is the support function of each of the following sets:*

$$G_S \stackrel{\text{def}}{=} \text{w}^*\text{-}\operatorname{cl}\operatorname{co}\{\text{w}^*\text{-}\lim \nabla f(x_i) | x_i \to x_*, x_i \notin \Omega_f \cup S\}$$

   where S is any Haar-null subset of X, and Ω_f is the set of points at which f is not differentiable. Therefore

$$\partial f(x_*) = G_S$$

   for each Haar-null set S.

Proof

1. [Cla, Prop. 2.1.1] and either Theorem 3.3.1 or [Cla, Prop. 2.1.2].

2. [Thi82, Prop. 2.2].

□

Part 2 of the theorem says that the generalized gradient is 'blind' to Haar-null sets. We point out the obvious: this characterization of the generalized gradient is measure theoretic whereas the original definition is analytic.

A useful result is the mean value theorem for generalized gradients, [Cla, Thm. 2.3.7].

Proposition 3.9 *(mean value)*

Let f be Lipschitz in a neighborhood U in X. If x_1, x_2 are points of U then there exists a scalar s in $(0,1)$ such that

$$f(x_1) - f(x_2) \in \partial f(sx_1 + (1-s)x_2)(x_1 - x_2).$$

Proposition 3.8.2 motivates a (measure theoretic) definition of the generalized Jacobian.

Definition 3.10 *(Clarke) The* generalized Jacobian *of g at $x_* \in X$ is*

$$\partial g(x_*) \stackrel{\text{def}}{=} \text{w*- cl co}\,\{\text{w*- lim}\,\nabla g(x_i) \mid x_i \to x_*, \nabla g(x_i) \text{ exists}\}$$

The significance of the Lipschitz condition on g becomes clear when we are aware of two facts. Firstly, the extension by Christensen ([Chr, Thm. 7.5]) of Rademacher's theorem — to the case of separable Banach spaces X, Y with Y reflexive — says

$$\begin{array}{ll} & g \text{ is Lipschitz in a neighborhood } U \text{ of } x_* \\ \Longrightarrow & g \text{ is differentiable everywhere in } U \text{ except possibly in a Haar-null subset.} \end{array}$$

Secondly, the Haar measure inherits a property of Lebesgue measure: if a Haar-null set is deleted from the neighborhood U, the remaining set is dense in U ([Thi75]).

Therefore there is a sequence (x_i) converging to x_* such that the Jacobians $\nabla g(x_i)$ exist and are bounded (by the Lipschitz constant of g). Sequential w*-compactness of bounded sequences ensures existence of a w*-limit point of $(\nabla g(x_i))$, whence $\partial g(x_*) \neq \emptyset$. Clearly $\partial g(x_*)$ is bounded by the Lipschitz constant of g and is, by definition, w*-closed and convex. This constitutes a proof of the first part of our final result.

Proposition 3.11 *Let X, Y be separable Banach spaces, with Y reflexive. Suppose g is Lipschitz in a neighborhood U of $x_* \in X$.*

1. *(Thibault) The generalized Jacobian of g at x_* is a nonempty, w^*-compact, convex set in $BL(X, Y)$.*

2. *(Thibault) For each set S in X, let*

$$J_S \stackrel{\text{def}}{=} \text{w}^*\text{-}\operatorname{cl}\operatorname{co}\{\text{w}^*\text{-}\lim \nabla g(x_i) | x_i \to x_*, x_i \notin \Omega_g \cup S\}$$

   where Ω_g is the set of points at which g is not differentiable. Then for each Haar-null set S and $\lambda \in Y^*$,

$$\partial(\lambda g)(x_*) = \lambda \partial g(x_*) = \lambda J_S.$$

3. *(mean value)*

   If x_1, x_2 are points of U, and $\lambda \in Y^*$, then there is a scalar s in $(0,1)$ such that

$$\begin{aligned} \lambda g(x_1) - \lambda g(x_2) &\in \partial(\lambda g)(sx_1 + (1-s)x_2)(x_1 - x_2) \\ &= \lambda \partial g(sx_1 + (1-s)x_2)(x_1 - x_2). \end{aligned}$$

Proof

1. [Thi82].

2. [Thi82, Prop. 2.4].

3. By the mean value result Proposition 3.9 there is $s \in (0,1)$ such that

$$\lambda g(x_1) - \lambda g(x_2) \in \partial(\lambda g)(sx_1 + (1-s)x_2)(x_1 - x_2).$$

   Now use part 2.

□

In finite dimensions, Pourciau [Por] slightly modifies Clarke's definition of the generalized Jacobian to obtain a definition which is blind to sets of zero Lebesgue measure.

The main classes of functions motivating the above definition are the smooth functions, and the convex functions.

Example 3.12 *If f (g) is continuously differentiable near $x_* \in X$ then*
$\partial f(x_*) = \{\nabla f(x_*)\}$
($\partial g(x_*) = \{\nabla g(x_*)\}$*).*

Example 3.13 *[Cla, Thm 2.5.1, Prop 2.3.6a] If f is convex then it is Lipschitz near any $x_* \in X$ and the its convex subdifferential at x_* equals the generalized gradient there:*

$$\{\xi \in X^* \,|\, \forall x \in X,\ \xi(x - x_*) \leq f(x) - f(x_*)\} \;=\; \partial f(x_*).$$

There is a substantial calculus for the generalized gradient and, in finite dimensions, the generalized Jacobian, including chain rules, mean value theorems, and the implicit function theorem — see [Cla, Ch. 2 and 7].

3.4 Rank-1 Generalized Jacobians

Definition 3.14 *The* rank-1 generalized Jacobian *of g at $x_* \in U$ is given by*

$$\partial^1 g(x_*) \stackrel{\text{def}}{=} \{A \in BL(X,Y) \mid \forall (u,\lambda) \in X \times Y^*,\ \ \lambda A u \leq (\lambda g)^\circ(x_*; u)\}$$

In finite dimensions [Swe77] showed that $\partial^1 g(x_*)$ is the plenary hull of the Clarke generalized Jacobian $\partial g(x_*)$. This was extended [Thi82] to separable Banach spaces. We prefer to emphasize the terminology *rank-1* simply because — as we will see — the existence of this object in infinite dimensions relies on the theory of rank-1 support functionals and is independent of the existence of the generalized Jacobian.

Theorem 3.15 *Let g be Lipschitz of modulus K in a neighborhood $U \subset X$. Then for each x_* in U:*

1. *The function $(u, \lambda) \mapsto (\lambda g)^\circ(x_*; u)$ is finite valued, positive bihomogeneous, subadditive in $u\lambda$, and bounded on $B_X \times B_Y^*$; hence is the rank-1 support function of the nonempty, w^*-compact, rank-1 representer $\partial^1 g(x_*)$.*

2. *$(\lambda g)^\circ(x_*; u)$ is finite valued, positive homogeneous, subadditive, and bounded as a function of u (λ respectively) alone.*

3. *$(\lambda g)^\circ(x_*; u)$ is upper semicontinuous in (x_*, u, λ); Lipschitz of modulus $K\|u\|$ as a function of λ alone; and Lipschitz of modulus $K\|\lambda\|$ as a function of u alone.*

   Also, $(\lambda g)^\circ(x_*; u) \leq K\|u\|\|\lambda\|$ and $(-\lambda g)^\circ(x_*; u) = (\lambda g)^\circ(x_*; -u)$.

Proof

1. It is easy to show finiteness, positive bihomogeneity, and boundedness by considering the real function λg — most of this is given by [Cla, Prop. 2.1.1]. If subadditivity in $u\lambda$ also holds, Theorem 3.6 gives the result.

   The subadditivity property is difficult to prove and relies on the finite dimensional version of Proposition 3.11.3.

   Let $u_0 \stackrel{\text{def}}{=} u$, $\lambda_0 \stackrel{\text{def}}{=} \lambda$ and, for some $m \in \mathbb{N}$, $((u_i, \lambda_i))_{i=1}^m$ be a sequence in $X \times Y^*$ such that

$$u_0\lambda_0 = \sum_{i=1}^{m} u_i\lambda_i.$$

   Let $\epsilon > 0$. Then there is $\delta > 0$ such that

(a) for some $\hat{x} \in x_* + \delta B_X$ and $0 < t < \delta$,

$$(\lambda_0 g)^\circ(x_*; u_0) \leq \frac{\lambda_0 g(\hat{x} + tu_0) - \lambda_0 g(\hat{x})}{t} + \frac{\epsilon}{2}$$

(b) for each $x \in \delta(1 + \|u_0\|)B_X$ and $i = 1, \ldots, m$,

$$(\lambda_i g)^\circ(x; u_i) \leq (\lambda_i g)^\circ(x_*; u_i) + \frac{\epsilon}{2m}$$

Only the final statement is possibly in question. This follows from from the upper semicontinuity of $(\lambda_i g)^\circ(\cdot\,; u_i)$, for each $i = 1, \ldots, m$, as given in [Cla, Prop. 2.1.1] or in part 3 to be proven.

Choose $\hat{x}$, t as in (a). We proceed, noting that Lemma 3.16 will be used in advance of its appearance: let $G : \hat{X} \to \mathbb{R}^M$ be the function produced by the lemma, where

$$\begin{aligned} \hat{X} &\stackrel{\text{def}}{=} \text{span}\{u_0, \ldots, u_m, \hat{x}\} \\ \Lambda &\stackrel{\text{def}}{=} \text{span}\{\lambda_0, \ldots, \lambda_m\} \end{aligned}$$

and M is the dimension of Λ. As in the lemma, Ψ is an isometry from Λ to $\mathbb{R}^M$. Let $l_i = \Psi(\lambda_i)^T$ for $i = 0, \ldots, m$. By Lemma 3.16, for each i, $\lambda_i g$ coincides with $l_i G$ on $\hat{X}$. Consequently $u_0 l_0$ coincides with $\sum_{i=1}^m u_i l_i$ on the span of $G(\hat{X})$, because $u_0 \lambda_0 = \sum_{i=1}^m u_i \lambda_i$.Therefore for each x in $\hat{X}$ and matrix A in the generalized Jacobian $\partial G(x)$, $u_0 l_0$ coincides with $\sum_{i=1}^m u_i l_i$ on the range of A. So

$$\begin{aligned} (\lambda_0 g)^\circ(x_*; u_0) &\leq \frac{\lambda_0 g(\hat{x} + tu_0) - \lambda_0 g(\hat{x})}{t} + \frac{\epsilon}{2} \quad \text{by (a) above} \\ &= \frac{l_0 G(\hat{x} + tu_0) - l_0 G(\hat{x})}{t} + \frac{\epsilon}{2} \end{aligned}$$

$$
\begin{aligned}
&= l_0 A u_0 + \frac{\epsilon}{2} \quad \text{for some } A \in \partial G(x) \text{ and some } x \text{ in the} \\
&\qquad \text{interval from } \hat{x} \text{ to } \hat{x} + t u_0, \text{ by Proposition 3.11.3} \\
&= \sum_{i=1}^{m} l_i A u_i + \frac{\epsilon}{2} \quad \text{by Proposition 3.1.2} \\
&\leq \sum_{i=1}^{m} (l_i G)^\circ(x; u_i) + \frac{\epsilon}{2} \quad \text{since } l_i A \in l_i \partial G(x) = \partial(l_i G)(x) \\
&\qquad \text{by Proposition 3.11.2} \\
&\leq \sum_{i=1}^{m} (\lambda_i g)^\circ(x; u_i) + \frac{\epsilon}{2} \quad \text{by Lemma 3.16} \\
&\leq \sum_{i=1}^{m} (\lambda_i g)^\circ(x_*; u_i) + \epsilon \quad \text{given (b) above.}
\end{aligned}
$$

Since ϵ can be chosen arbitrarily close to zero, the required subadditivity must follow.

2. Subadditivity of $(\lambda g)^\circ(x_*; u)$ in $u\lambda$ implies subadditivity in u for fixed λ and subadditivity in λ for fixed u (though not conversely). Similarly, positive bihomogeneity implies positive homogeneity in each variable u and λ separately. These properties can also be shown directly.

3. The required properties can be shown without difficulty along the lines of the proof of [Cla, Prop. 2.1.1]. We will only show upper semicontinuity of $(\lambda g)^\circ(x_*; u)$ in (x_*, u, λ).

   Let $(x_i, u_i, \lambda_i) \to (x_*, u, \lambda)$. Then, for each i, there exist $x_i' \in x_i + (1/i)B_X$ and $t_i' \in (0, 1/i)$ such that

$$
(\lambda_i g)^\circ(x_i; u_i) - \frac{1}{i} \leq \frac{\lambda_i g(x_i' + t_i u_i) - \lambda_i g(x_i')}{t_i}
$$

$$
\begin{aligned}
&= (\lambda_i - \lambda)\frac{g(x_i' + t_i u_i) - g(x_i')}{t_i} \\
&+ \lambda\frac{g(x_i' + t_i u_i) - g(x_i' + t_i u)}{t_i} \\
&+ \frac{\lambda g(x_i' + t_i u) - \lambda g(x_i')}{t_i}
\end{aligned}
$$

Note that for sufficiently large i, the Lipschitz condition on g ensures that the first term is bounded by $\|\lambda_i - \lambda\|K\|u_i\|$ and the second term by $\|\lambda\|K\|u_i - u\|$. Taking upper limits as $i \to \infty$ we obtain

$$\limsup_{i\to\infty}(\lambda_i g)^\circ(x_i; u_i) \leq (\lambda g)^\circ(x_*; u).$$

$\square$

Lemma 3.16 *Assume the hypothesis of Theorem 3.15.*

Suppose $(u_0, \lambda_0), \ldots, (u_m, \lambda_m) \in X \times Y^*$ *and* $\hat{x} \in U$. *Let*

$$
\begin{aligned}
\hat{X} &\stackrel{\text{def}}{=} span\{u_0, \ldots, u_m, \hat{x}\}, \\
\Lambda &\stackrel{\text{def}}{=} span\{\lambda_0, \ldots, \lambda_m\},
\end{aligned}
$$

and M *be the dimension of* Λ.

Let Ψ *be an algebraic isomorphism from* Λ *to* $\mathbb{R}^M$; *and* $\mathbb{R}^M$ *have the norm of the dual of* Λ, *that is*

$$\|e\| \stackrel{\text{def}}{=} \max\{\Psi(\lambda)^T e \mid \lambda \in \Lambda, \|\lambda\| \leq 1\}, \ \ \forall e \in \mathbb{R}^M.$$

Denote the i*th unit vector of* $\mathbb{R}^M$ *— consisting of 0's except for the entry 1 at the* i*th position — by* e_i. *Define* $G : \hat{X} \to \mathbb{R}^M$ *at* $x \in \hat{X}$ *by*

$$G(x) \stackrel{\text{def}}{=} \sum_{i=1}^{M}(\Psi^{-1}(e_i)g(x))e_i.$$

The mapping G has the following properties:

1. *For each (x, λ) in $\hat{X} \times \Lambda$, $\lambda g(x) = \Psi(\lambda)^T G(x)$.*

2. *G is Lipschitz of modulus K in the neighborhood $\hat{U} \stackrel{\text{def}}{=} U \cap \hat{X}$ of $\hat{x}$.*

3. *For each x in $\hat{U}$, (u, λ) in $\hat{X} \times \Lambda$,*
$$(\Psi(\lambda)^T G)^\circ(x; u) \leq (\lambda g)^\circ(x; u).$$

Proof

1. Let $x \in \hat{X}$ and $\lambda \in \Lambda$. Then $\lambda = \sum_{i=1}^{M} \alpha_i \Psi^{-1}(e_i)$ for some scalars $\alpha_1, \ldots, \alpha_M$. So $\Psi(\lambda) = \sum_i \alpha_i e_i$ and
$$\begin{aligned} \lambda g(x) = \sum_{i=1}^{M} \alpha_i \Psi^{-1}(e_i) g(x) &= \left\langle \sum_{j=1}^{M} \alpha_j e_j, \sum_{i=1}^{M} (\Psi^{-1}(e_i) g(x)) e_i \right\rangle \\ &= \Psi(\lambda)^T G(x) \end{aligned}$$

2. Let $x_1, x_2 \in \hat{U}$.
$$\begin{aligned} \|G(x_1) - G(x_2)\| &= \max\{\Psi(\lambda)^T (G(x_1) - G(x_2)) \mid \lambda \in \Lambda, \|\lambda\| \leq 1\} \\ &= \max\{\lambda(g(x_1) - g(x_2)) \mid \lambda \in \Lambda, \|\lambda\| \leq 1\} \quad \text{from 1} \\ &\leq \|g(x_1) - g(x_2)\| \\ &\leq K\|x_1 - x_2\|. \end{aligned}$$

3. Let $x \in \hat{U}$, $u \in \hat{X}$, and $\lambda \in \Lambda$. Then
$$\begin{aligned} (\Psi(\lambda)^T G)^\circ(x; u) &= \limsup_{\substack{z \in \widehat{X}, z \to x \\ t \downarrow 0}} \frac{\lambda g(z + tu) - \lambda g(z)}{t} \\ &\leq \limsup_{\substack{z \to x \\ t \downarrow 0}} \frac{\lambda g(z + tu) - \lambda g(z)}{t} = (\lambda g)^\circ(x; u). \end{aligned}$$

□

We note that Theorem 3.15.1 requires set theoretic tools in finite dimensions, namely Rademacher's theorem. The existence of the generalized gradient, however, can be shown in an entirely analytic fashion ([Cla, Ch. 2]). An open question, important for the completeness of the theory, is whether or not Theorem 3.15.1 also has an analytic proof.

Proposition 3.17 *Let g be Lipschitz of modulus K in a neighborhood $U \subset X$. For $x_* \in U$:*

1. *$\partial^1 g(x_*)$ is a nonempty, w^*-compact, rank-1 (plenary), convex subset of $BL(X,Y)$ and $\|A\| \leq K$ for each $A \in \partial^1 g(x_*)$.*

2. *The set function $\partial^1 g$ is w^*-closed and is w^*-upper semicontinuous in U; hence closed and upper semicontinuous in finite dimensions.*

3. *For each u in X, λ in Y^*,*

$$
\begin{aligned}
(\lambda g)^\circ(x_*; u) &= \max\{\lambda A u \mid A \in \partial^1 g(x_*)\}, \\
\partial(\lambda g)(x_*) &= \lambda \partial^1 g(x_*), \\
F(u) &= \partial^1 g(x_*) u,
\end{aligned}
$$

where $F(u) \stackrel{\text{def}}{=} \{y \in Y \mid (\forall \lambda' \in Y^*)\ \lambda' y \leq (\lambda' g)^\circ(x_*; u)\}$.

4.

$$
\begin{aligned}
\partial^1 g(x_*) &= \bigcap_{\delta > 0} \bigcup_{x \in x_* + \delta B_X} \partial^1 g(x) \\
&= \bigcap_{\lambda} \{A \mid \lambda A \in \partial(\lambda g)(x_*)\} \\
&= \bigcap_{u} \{A \mid Au \in F(u)\}
\end{aligned}
$$

where $F(u)$ is defined as above.

Proof

1. Excepting the norm bound of K, the stated properties of $\partial^1 g(x_*)$ are given by Proposition 3.5.1 after noting Theorem 3.15.1. Now suppose $A \in \partial^1 g(x_*)$. For each (u, λ) in $X \times Y^*$, Theorem 3.15.3 shows

$$\lambda A u \leq (\lambda g)^\circ(x_*; u) \leq K\|u\|\|\lambda\|.$$

By the Hahn-Banach theorem, we may choose λ of unit norm such that $\lambda A u = \|Au\|$. Hence A is bounded above in norm by K.

2. We must first show that the graph of $\partial^1 g|_U$, i.e. the set

$$\Gamma_U \stackrel{\text{def}}{=} \{(x, A) \in U \times BL(X, Y) \mid A \in \partial^1 g(x)\},$$

is closed relative to $U \times BL(X, Y)$, where $BL(X, Y)$ is endowed with the w*-topology. To this end, let $((x_k, A_k))$ be a sequence in Γ_U which converges to a point (x, A) in $U \times BL(X, Y)$. Since $A_k \in \partial^1 g(x_k)$ for each k, we have for $(u, \lambda) \in X \times Y^*$ that

$$\begin{aligned} \lambda A u = \lim_k \lambda A_k u &\leq \limsup_k (\lambda g)^\circ(x_k; u) \\ &\leq (\lambda g)^\circ(x; u) \end{aligned}$$

where the last inequality follows from the upper semicontinuity of $(\lambda g)^\circ(\cdot\,; u)$, as given by Theorem 3.15.3. Therefore $A \in \partial^1 g(x)$, and $(x, A) \in \Gamma_U$ as required.

The operator $\partial^1 g$ cannot be w*-closed without being w*-upper semicontinuous. For suppose W is a w*-open set containing $\partial^1 g(x)$, and (x_k) converges to x in U such that for each k there exists A_k in $\partial^1 g(x_k)\backslash W$. The sequence (A_k) is norm bounded, from the first part of the theorem, hence has a w*-limit point A in $BL(X,Y)\backslash W$. So $((x_k, A_k))$ has a subsequence in Γ_U converging to $(x, A) \in [U \times BL(X,Y)]\backslash\Gamma_U$ — this contradicts the fact that Γ_U is closed relative to $U \times BL(X,Y)$.

3. The first equality is given by Theorem 3.15. We will use this to show the second equality; the third follows similar lines.

   As a function of u only, $(\lambda g)^\circ(x_*; u)$ is the support function of both $\partial(\lambda g)(x_*)$ and $\lambda\partial^1 g(x_*)$. The former set is, by Proposition 3.8.1, nonempty, w*-closed and convex. The latter set also has these properties since $\partial^1 g(x_*)$ is a nonempty, w*-closed, convex set in $BL(X,Y)$, by part 1 above. Appealing to Theorem 3.3.1, the sets $\partial(\lambda g)(x_*)$ and $\lambda\partial^1 g(x_*)$ must be the same.

4. For the first equality observe that

$$\partial^1 g(x_*) \subset \bigcap_{\delta>0} \bigcup_{x\in x_*+\delta B_X} \partial^1 g(x).$$

   Conversely, if A is a member of the set on the right then there is a sequence (x_k) converging to x_* such that each $\partial^1 g(x_k)$ contains A. Weak-operator-closedness of ∂g, above, yields $A \in \partial^1 g(x_*)$. Thus equality holds.

   The remaining equalities are made clear by the following equivalences for any A in $BL(X,Y)$:

$$\begin{array}{lll} & \forall\lambda & \lambda A \in \partial(\lambda g)(x_*) \\ \iff & \forall u,\lambda & \lambda A u \le (\lambda g)^\circ(x_*; u) \ \ (\iff A \in \partial^1 g(x_*)) \\ \iff & \forall u & Au \in F(u) \end{array}$$

□

3.5 Other Derivatives, Shields and Fans

We first point out that the 'rank-1 generalized gradient' coincides with the generalized gradient.

Lemma 3.18 *Let g be Lipschitz near $x_* \in X$. If $Y = \mathbb{R}$,*

$$\partial^1 g(x_*) = \partial g(x_*).$$

Hence $\partial g(x_*)$ has all the properties of $\partial^1 g(x_*)$ listed in Proposition 3.17.

Proof From Theorem 3.15, for each u in X and scalar α

$$(\alpha g)^\circ(x_*; u) = |\alpha| g^\circ(x_*; \operatorname{sgn}(\alpha)u) = g^\circ(x_*; \alpha u).$$

Hence

$$\forall (u, \alpha),\ \alpha\xi u \leq (\alpha g)^\circ(x_*; u) \iff \forall u,\ \xi u \leq g^\circ(x_*; u)$$

□

To compare the rank-1 generalized Jacobian with the generalized Jacobian, when the latter exists, we quote from [Thi82, Prop. 2.3]. This says, in part, that the rank-1 generalized Jacobian is equal to the rank-1 representer of the generalized Jacobian; and, in fact, the rank-1 version is 'blind' to Haar-null sets, as is the generalized gradient.

Proposition 3.19 *Let X and Y be separable Banach spaces, with Y reflexive. Let g be Lipschitz near $x_* \in X$.*

1. *The mapping $(u, \lambda) \mapsto (\lambda g)^\circ(x_*; u)$ is the rank-1 support function of $\partial^1 g(x_*)$ and of (the rank-1 representers of) each of the sets:*

$$J_S \stackrel{\text{def}}{=} \text{w*-}\operatorname{cl}\operatorname{co}\{\text{w*-}\lim \nabla g(x_i) \mid x_i \to x_*, x_i \notin \Omega_g \cup S\}$$

where S is a Haar-null set in X and Ω_g is the set of points of X at which g is not differentiable. Then, for each Haar-null set S in X,

$$\partial^1 g(x_*) = (J_S)^1.$$

For $S = \emptyset$, this gives

$$\partial^1 g(x_*) = [\partial g(x_*)]^1.$$

Hence $\partial^1 g(x_*) \supset \partial g(x_*)$; this inclusion is, in general, strict.

2. *For each $\lambda \in Y^*$,* $\begin{aligned} \partial(\lambda g)(x_*) &= \lambda \partial g(x_*) \\ &= \lambda \partial^1 g(x_*) \end{aligned}$

   Similarly, for each u in X, $\partial g(x_*)u = \partial^1 g(x_*)u$.

Proof

1. Excepting the last statement, this result is given by [Thi82, Prop. 2.3]. For the remainder, first recall that a rank-1 representer always contains the original set, hence $\partial^1 g(x_*) \supset \partial g(x_*)$. The piecewise linear mapping in [Swe77, Example 6.4], quoted in Chapter 1 as Example 1.41, is a counterexample to equality between the rank-1 and the original generalized Jacobian.

2. This is given by [Thi82, Prop. 2.3], except the final statement which can be shown by reasoning similar to that used by Thibault in the proof of his proposition.

□

Now we introduce Sweetser's *shields* [Swe77, Swe79]. A (w*-)shield for g at x_* is a set Γ in $BL(X, Y)$ such that

$\forall$ (w*-)neighborhoods W of Γ

$\exists$ a positive radius δ (of a ball about x_*)

$\forall\ x_1, x_2 \in x_* + \delta B_X$

we have

$g(x_1) - g(x_2) \in W(x_1 - x_2)$.

In finite dimensions, [Swe77] shows that the 'plenary hull' of the generalized Jacobian (= the rank-1 generalized Jacobian) is the minimum closed, plenary, convex shield for g at x_*. We extend this to infinite dimensions.

Proposition 3.20 *Let g be Lipschitz near $x_* \in X$.*

The rank-1 generalized Jacobian $\partial^1 g(x_*)$ is the minimum rank-1 representer which is a w-shield for g at x_*. Equivalently, $\partial^1 g(x_*)$ is the minimum w*-closed, rank-1 (plenary), convex w*-shield for g at x_*.*

Proof First recall, from Theorem 3.15.1, that $\partial^1 g(x_*)$ is a w*-compact, rank-1 representer. Now we show that $\partial^1 g(x_*)$ is a w*-shield for g at x_*. Let W be a w*-neighborhood of $\partial^1 g(x_*)$ which we assume, without loss of generality, is convex and w*-closed. By w*-upper semicontinuity of $\partial^1 g$ (Proposition 3.17.2) there exists $\delta > 0$ such that $\partial^1 g(x_* + \delta B_X) \subset W$. Therefore

$$\Gamma \stackrel{\text{def}}{=} \text{w*-}\operatorname{cl}\operatorname{co}[\partial^1 g(x_* + \delta B_X)] \subset \text{w*-}\operatorname{cl}\operatorname{co} W = W.$$

By the mean value theorem for rank-1 generalized Jacobians, given later as Proposition 3.27, we have for $x_1, x_2 \in x_* + \delta B_X$ that

$$g(x_1) - g(x_2) \in \Gamma(x_1 - x_2) \subset W(x_1 - x_2).$$

So $\partial^1 g(x_*)$ is a w*-shield as claimed.

We have seen that $\partial^1 g(x_*)$ is a rank-1 representer which is a w*-shield for g at x_*. To show that $\partial^1 g(x_*)$ is the minimum such set, let Γ be a rank-1 representer which is a strict subset of $\partial^1 g(x_*)$. Hence there exists $(u, \lambda) \in X \times Y^*$ and $\epsilon > 0$ such that

$$\sigma^1_\Gamma(\lambda, u) + \epsilon < \max_{A \in \partial^1 g(x_*)} \lambda A u - \epsilon = (\lambda g)^\circ(x_*; u) - \epsilon, \tag{3.7}$$

where the equality is given by Proposition 3.17.3.

Now consider the w*-neighborhood

$$W \stackrel{\text{def}}{=} \left\{ A \in BL(X,Y) \mid \lambda A u < \sigma^1_\Gamma(u,\lambda) + \epsilon \right\}$$

of Γ. We choose $x_i \to x_*$, $t_i \downarrow 0$ such that

$$(\lambda g)^\circ(x_*; u) - \epsilon < \frac{\lambda g(x_i + t_i u) - \lambda g(x_i)}{t_i}$$

for each i. We deduce from equation (3.7) that for each i

$$\sigma_W(\lambda, u) \leq \sigma_\Gamma(\lambda, u) + \epsilon < \frac{\lambda g(x_i + t_i u) - \lambda g(x_i)}{t_i}.$$

Therefore

$$\begin{array}{rrcl} & \lambda g(x_i + t_i u) - \lambda g(x_i) & \notin & t_i \lambda W u = \lambda W(x_i + t_i u - x_i) \\ \Rightarrow & g(x_i + t_i u) - g(x_i) & \notin & W(x_i + t_i u - x_i). \end{array}$$

It is clear that, as the sequences $(x_i + t_i u), (x_i)$ both converge to x_*, every neighborhood U of x_* contains points x, x' for which

$$g(x) - g(x') \notin W(x - x'),$$

i.e. Γ is not a w*-shield for g at x_*.

We have shown there exists no proper subset of $\partial^1 g(x_*)$ which is both a rank-1 representer and a shield for g. Proposition 3.5.1 provides the alternative characterization of $\partial^1 g(x_*)$ as the minimum w*-closed, rank-1, convex shield for g at x_*. Proposition 3.5.3 says the rank-1 condition can be replaced by plenarity, because we know $\partial^1 g(x_*)$ is w*-compact.

$\square$

More notation is required in order to discuss relationships between the rank-1 generalized Jacobian, derivatives of a more classical nature, and shields.

We consider X and the dual space Y^* only under their norm topologies. We will consider Y, however, under its two standard topologies, namely the norm topology and the weak (w-) topology. Unless the weak topology is specified, Y is assumed to be a normed space.

The derivatives of interest will be the Gâteaux- (denoted G-), the Hadamard- (H-), and the Fréchet- (F-) derivatives respectively. The following discussion on differentiability should be compared to [Cla, Ch. 2.2] in which *strict differentiability* corresponds to notion, below, of *strict H-differentiability.*

The function g is said to be (w-)G- or (w-)H- or (w-)F- differentiable at x_*, respectively, if there is a map $\nabla g(x_*) \in BL(X, Y)$ such that

$$\text{(w-)} \lim_{t \downarrow 0} \frac{g(x_* + tu) - g(x_*)}{t} = \nabla g(x_*)u,$$

and convergence is uniform for u in finite or compact or bounded sets, respectively. $\nabla g(x_*)$ is called the derivative of g at x_*.

Strict differentiability is stronger than ordinary differentiability. The function g is said to be strictly (w-)G- or strictly (w-)H- or strictly (w-)F-differentiable at x_*, respectively, if there is a map $\nabla g(x_*) \in BL(X, Y)$ such that

$$\text{(w-)} \lim_{\substack{x \to x_* \\ t \downarrow 0}} \frac{g(x + tu) - g(x)}{t} = \nabla g(x_*)u,$$

where convergence is uniform with respect to u in finite or compact or bounded sets, respectively. Here $\nabla g(x_*)$ is called the strict derivative of g at x_*.

It is clear that the weakest of these smoothness properties is w-G-differentiability, while the strongest is strict F-differentiability.

Proposition 3.21 *Suppose g is Lipschitz near $x_* \in X$. If g is differentiable at x_* in any of the ways defined above, then*

$$\nabla g(x_*) \in \partial^1 g(x_*).$$

Proof Clear from definitions. □

Proposition 3.22 *Let $x_* \in X$ and $\bar{A}$ be an operator in $BL(X,Y)$. The following statements are equivalent:*

1. *g is strictly w-H-differentiable at x_* with $\nabla g(x_*) = \bar{A}$.*

2. $$\lim_{\substack{x \to x_* \\ t \downarrow 0}} \frac{\lambda g(x+tu) - \lambda g(x)}{t} = \lambda \bar{A} u, \tag{3.8}$$

   where convergence is uniform for u, λ in compact sets of X, Y^* respectively.

3. *g is Lipschitz near x_*, and for each (u,λ) in $X \times Y^*$, the limit (3.8) holds.*

4. *g is Lipschitz near x_*, and $\partial^1 g(x_*) = \{\bar{A}\}$.*

5. *$\{\bar{A}\}$ is a w^*-shield for g at x_*.*

Proof $(1 \Rightarrow 2)$ Suppose $\bar{A}$ is the strict derivative of g at x_*. Let C, Λ be compact sets in X, Y^* respectively, and ϵ be a positive constant. We will find $\bar{\delta} > 0$ such that

$$\forall (x \in x_* + \bar{\delta} B_X, 0 < t < \bar{\delta}, (u,\lambda) \in C \times \Lambda) \quad |\lambda d(x,t,u)| < \epsilon. \tag{3.9}$$

where $d(x,t,u) \stackrel{\text{def}}{=} (1/t)(g(x+tu) - g(x)) - \bar{A}u$

Now the differences $d(x,t,u)$ weakly converge to $0 \in Y$ as $x \to x_*$ and $t \downarrow 0$, uniformly for u in C; thus for some $\delta_0 > 0$ the set

$$S \stackrel{\text{def}}{=} \{d(x,t,u) \mid x \in x_* + \delta_0 B_X, 0 < t < \delta_0, u \in C\}$$

is w-bounded. By [Sch, Ch. IV §3.2 Cor. 2], S is norm bounded; let this bound be $D > 0$.

A different view of statement 1 is that the limit (3.8) holds, where convergence is uniform for u in compact sets of X and each λ in Y^*. So for each λ there exists $\delta \in (0, \delta_0)$ such that for $x \in x_* + \delta B_X$, $0 < t < \delta$ and $u \in C$, we have $|\lambda d(x,t,u)| \leq \epsilon/2$. Hence, for $\delta' \stackrel{\text{def}}{=} \min\{\delta, \epsilon/(2D)\}$,

$$\forall (x \in x_* + \delta' B_X, 0 < t < \delta', u \in C, \lambda' \in \lambda + \delta' B_Y^*) \quad |\lambda' d(x,t,u)| < \epsilon. \tag{3.10}$$

We can cover Λ with such open sets $\lambda + \delta' B_Y^*$. By compactness of Λ, finitely many of these sets cover it, i.e. finitely many $\lambda_i \in Y^*$ with corresponding radii δ_i' such that (3.10) holds for $\lambda = \lambda_i$ and $\delta' = \delta_i'$ and

$$\Lambda \subset \bigcup_i (\lambda_i + \delta_i' B_Y^*).$$

Let $\bar{\delta} \stackrel{\text{def}}{=} \min_i \delta_i > 0$. Then (3.9) holds and we are done.

($2 \Rightarrow 3$) We only need show that g is Lipschitz near x_*. Assume not; then for each $i \in \mathbb{N}$ there are $x_i, x^i \in x_* + (1/i)B_X$ such that

$$\|g(x_i) - g(x^i)\| \geq i\|x_i - x^i\|.$$

Hence, by the Hahn-Banach theorem, there is a sequence of norm-1 forms (λ_i) in Y^* such that

$$\lambda_i g(x_i) - \lambda_i g(x^i) \geq i\|x_i - x^i\|.$$

For each i, let $u_i \in X$ and $t_i > 0$ satisfy $x^i = x_i + t_i u_i$ where $\|u_i\| = i^{-1/3}$; and $\mu_i = \lambda_i i^{-1/3}$. Then $t_i \downarrow 0$, $\{u_i\} \cup \{0\}$ is compact, and $\{\mu_i\} \cup \{0\}$ is compact. Also

$$\begin{aligned} \frac{\mu_i g(x_i + t_i u_i) - \mu_i g(x_i)}{t_i} &= i^{-1/3} \frac{\lambda_i g(x^i) - \lambda_i g(x_i)}{t_i} \\ &\geq i^{2/3} \frac{\|x^i - x_i\|}{t_i} = i^{2/3}\|u_i\| = i^{1/3} \end{aligned}$$

This contradicts statement 2 because $i^{1/3} \uparrow \infty$ as $i \uparrow \infty$.

(3 $\Rightarrow$ 4) Clearly, for each u and λ, $\lambda \bar{A} u = (\lambda g)^\circ(x_*; u)$. So $\bar{A}$ is a member of $\partial^1 g(x_*)$.

Let $A \in BL(X,Y)\backslash\{\bar{A}\}$. Then for some u in X, $Au \neq \bar{A}u$. By the Hahn-Banach theorem there exists λ in Y^* such that $\lambda Au > \lambda \bar{A} u = (\lambda g)^\circ(x_*; u)$, i.e. $A \notin \partial^1 g(x_*)$.

(4 $\Rightarrow$ 5) This is a corollary of Proposition 3.20.

(5 $\Rightarrow$ 1) By contraposition. Assume statement 1 is false: $\bar{A}$ is not the w-H-derivative of g at x_*. Hence there are a set Λ in Y^* of finite cardinality (which defines a w-neighborhood of $0 \in Y$) and a compact set C in X, for which we have the following: there exist sequences $(x_i) \to x_*$ in X, $(t_i) \downarrow 0$, (u_i) in C, and (λ_i) in Λ such that such that for each i

$$\lambda_i \frac{g(x_i + tu_i) - g(x_i)}{t_i} \geq \lambda_i \bar{A} u_i + 1. \tag{3.11}$$

By compactness of C we may assume without loss of generality that (u_i) converges to some u in C. Similarly, by the finite cardinality of Λ, we assume without loss that (λ_i) is a constant sequence with value always equal to λ.

Using the Lipschitz modulus K of g near x_*, we find for all sufficiently large i that

$$\begin{aligned} \lambda \frac{g(x_i + t_i u) - g(x_i)}{t_i} &\geq \lambda \frac{g(x_i + t_i u_i) - g(x_i)}{t_i} - \|\lambda\| K \|u - u_i\| \\ &\geq \lambda \frac{g(x_i + t_i u_i) - g(x_i)}{t_i} - \frac{1}{2}. \end{aligned}$$

Hence, for sufficiently large i, we use this and equation (3.11) to get

$$\lambda \frac{g(x_i + t_i u) - g(x_i)}{t_i} \geq \lambda \bar{A} u + \frac{1}{2}. \tag{3.12}$$

Now consider the w*-neighborhood

$$W \stackrel{\text{def}}{=} \left\{ A \mid \forall \lambda \in \Lambda,\ \lambda A u < \lambda \bar{A} u + \frac{1}{2} \right\}$$

of $\{\bar{A}\}$. For any positive radius δ we may choose i large enough so that both x_i and $x_i + t_i u$ are within distance δ of x_*, and equation (3.12) holds. From this equation, however, we get

$$g(x_i + t_i u) - g(x_i) \notin t_i W u = W(x_i + t_i u - x_i).$$

Therefore $\{\bar{A}\}$ is not a w*-shield for g at x_*.

$\square$

Corollary 3.23 *Suppose g is differentiable near $x_* \in X$ in any of the ways defined above, and its derivative ∇g is continuous at x_*. Then all the statements 1-5 of Proposition 3.22 are valid.*

Proof We will prove statement 3 of Proposition 3.22. Note that for any λ in Y^* and x_1, x_2 near x_* the classical mean value theorem for λg yields

$$\lambda g(x_1) - \lambda g(x_2) = \lambda \nabla g(s x_1 + (1-s) x_2)(x_1 - x_2)$$

for some $s \in (0, t)$.

The limit (3.8) holds:

$$\begin{aligned} \frac{\lambda g(x + tu) - \lambda g(x)}{t} &= \lambda \nabla g(x + su) u \text{ for some } s \in (0,1) \\ &\to \lambda \nabla g(x_*) u \ \text{ as } x \to 0, t \downarrow 0, \end{aligned}$$

where convergence follows from continuity of ∇g at x_*.

Now use continuity of ∇g to find positive constants δ, D such that for $x \in x_* + \delta B_X$

$$\|\nabla g(x) - \nabla g(x_*)\| < \delta.$$

Then for $x_1, x_2 \in x_* + B_X$, there exists a form $\lambda \in Y^*$ of norm 1 (by Hahn-Banach), and $s \in (0,1)$ such that

$$\begin{aligned} \|g(x_1) - g(x_2)\| &= \lambda g(x_1) - \lambda g(x_2) \\ &= \lambda \nabla g(sx_1 + (1-s)x_2)(x_1 - x_2) \\ &\leq \|\nabla g(sx_1 + (1-s)x_2)\| \|x_1 - x_2\| \\ &\leq (\|\nabla g(x_*)\| + D)\|x_1 - x_2\|. \end{aligned}$$

Hence g is Lipschitz near x_* of modulus $\|\nabla g(x_*)\| + D$.

□

Finally we move to Ioffe's *fans* ([Iof81, Iof82]). A fan is a set valued mapping which, in some important ways, generalizes a (sub)linear operator.

Definition 3.24 *Let F be a set valued mapping from X to Y.*

1. *F is called a* fan *if it*
   (a) *takes nonempty, convex values: $F(u)$ is nonempty and convex, for $u \in X$;*
   (b) *is positive homogeneous: $F(\alpha u) = \alpha F(u)$ for $\alpha \geq 0, u \in X$; and*
   (c) *is subadditive: $F(u_1 + u_2) \subset \mathrm{cl}(F(u_1) + F(u_2))$ for $u_1, u_2 \in X$.*
2. *A fan F is called* odd *if $F(-u) = -F(u)$, for $u \in X$.*

3. Let F be a fan. The handle *of F is the set*

$$h(F) \stackrel{\text{def}}{=} \{A \in BL(X,Y) \mid \lambda A u \leq \sup_{y \in F(u)} \lambda y, \ \ \forall (u,\lambda) \in X \times Y^*\}.$$

F is said to be spanned by its handle *if $F(u) = h(F)u$, for $u \in X$.*

4. The adjoint *of F, when it exists (i.e. has nonempty values), is the fan F^* from Y^* to X^* defined by*

$$F^*(\lambda) \stackrel{\text{def}}{=} \{\xi \in X^* \mid (\forall u \in X)\ \xi u \leq \sup_{y \in F(u)} \lambda y\}.$$

The property of a fan's being spanned by its handle is of special interest, because it means the fan is generated by a rank-1 representer (its handle). The open question of characterizing fans which are spanned by their handles is partially answered in Chapter 2, where a characterization of closed valued fans which are generated by their handles as

$$F(u) = \mathrm{cl}[h(F)u], \tag{3.13}$$

is given. In particular when $h(F)$ is w*-compact, F is generated as in (3.13) iff it is spanned by its handle.

We recall the adjoint A^* of an operator $A \in BL(X,Y)$ which is defined on Y^* by $A^*(\lambda) \stackrel{\text{def}}{=} \lambda A \in X^*$. In fact A^* is a member of $BL(Y^*, X^*)$ where both X^* and Y^* are endowed with their dual norm topology.

Proposition 3.25 *Let g be Lipschitz near $x_* \in X$. Define a set valued mapping F from X to Y by*

$$F(u) \stackrel{\text{def}}{=} \{y \in Y \mid \lambda y \leq (\lambda g)^\circ(x_*; u), \ \ \forall \lambda \in Y^*\}, \quad \forall u \in X.$$

Then F is a fan which is spanned by its handle, hence is odd; and the handle of F is $\partial^1 g(x_*)$. An alternative description of F is

$$F(u) = \partial^1 g(x_*)u, \quad \forall u \in X.$$

Moreover, the adjoint F^* of F exists and has the following properties:

$$F^*(\lambda) = \lambda \partial^1 g(x_*), \quad \forall \lambda \in Y^*;$$

F^* is spanned by its handle, hence is odd; and the handle of F^* is $\{A^* \mid A \in \partial^1 g(x_*)\}$.

Proof $F(u)$ is convex and nonempty by Proposition 3.17.3. Moreover, by Proposition 3.15.2, the mapping $u \mapsto (\lambda g)^\circ(x_*; u)$ is positive homogeneous and sublinear; hence F is positive homogeneous and sublinear. That $h(F) = \partial^1 g(x_*)$ is clear from definitions. That F is spanned by its handle is given by Proposition 3.17.3.

The properties of F^* are easily shown along similar lines. $\square$

The result that the fan F in the theorem is spanned by its handle is quite unexpected given the negative comment of Hiriart-Urruty [H-U82, §3] (see also [Thi82, §4]) that 'the mere question of existence [of the rank-1 generalized Jacobian] is hopeless for X and Y general Banach spaces', even with the requirement of reflexivity on Y.

3.6 Basic Calculus

We begin with the scalar multiplication and sum rules for rank-1 generalized Jacobians.

Lemma 3.26 *Let $x_* \in X$.*

1. *If g is Lipschitz near x_* and s is any scalar, then $\partial^1(sg)(x_*) = s\partial^1 g(x_*)$.*

2. If each function $g_j : X \to Y$ $(j = 1, \ldots, m)$ is Lipschitz near x_* then

$$\partial^1 \left(\sum_{j=1}^{m} g_j \right) (x_*) \subset \left[\sum_{j=1}^{m} \partial^1 g_j(x_*) \right]^1 .$$

Proof

1. The result is not hard to prove given Theorem 3.15: use positive homogeneity of $\lambda \mapsto (\lambda g)^\circ(x_*; u)$ for $s > 0$, and also the fact that $(-\lambda g)^\circ(x_*; u) = (\lambda g)^\circ(x_*; -u)$ when $s < 0$.

2. It is easy to see that

$$(\lambda \sum_j g_j)^\circ(x_*; u) \le \sum_j (\lambda g_j)^\circ(x_*; u),$$

from which Theorem 3.3 gives

$$(\lambda \sum_j g_j)^\circ(x_*; u) \le \max\{\lambda B u \mid B \in \sum_j \partial^1 g_j(x_*)\}.$$

The inclusion follows.

□

Remark. Part 2 of the above lemma was given [Cla, Prop. 2.3.3] for sums of real locally Lipschitz functions on X, $\sum_{j=1}^{m} f_j$, using the generalized gradient in place of the rank-1 generalized Jacobian. Equality is shown to hold [Cla, Prop 2.3.3 Cor. 3] when each of the functions f_j is *regular* at x_*, i.e. the directional derivative $f_j'(x_*; \cdot)$ exists and coincides with the generalized directional derivative $f_j^\circ(x_*; \cdot)$. The class of regular locally Lipschitz functions is 'large', at least from the point of view of optimization, since it includes all (finite valued) convex functions [Cla, Prop. 2.3.6] and the functions defined as the pointwise maxima of suitable (eg. finite [Cla, Prop. 2.3.12]) families of real smooth functions. For the locally Lipschitz function $g : X \to Y$, we too can define regularity: g is *rank-1 regular* at x_* if, for each $\lambda \in Y^*$, the directional derivative $(\lambda g)'(x_*;\cdot)$ exists and coincides with the generalized directional derivative $(\lambda g)^\circ(x_*;\cdot)$. It is not hard to see that equality holds in part 2 of the lemma above if each of the functions g_j is rank-1 regular at x_*. However the class of rank-1 regular functions does not seem to be very rich, a deficiency which we try to explain as follows. If g is rank-1 regular and directionally differentiable at x_*, then it satisfies

$$-g'(x_*;u) = g'(x_*;-u), \quad \forall u \in X,$$

i.e. for each $u \in X$, the mapping of scalars: $t \mapsto g(x_*+tu)$ is differentiable at $t = 0$. This seems to exclude many 'reasonable' nonsmooth functions from consideration (and also exposes a major difference between the rank-1 regular real functions and the regular (real) functions).

Next we have a mean value theorem for the rank-1 generalized Jacobian. This should be compared to [Cla, Prop. 2.6.5] and [Thi82, Prop. 4.3].

Proposition 3.27 *(mean value)*

Let g be Lipschitz in a neighborhood U of $x_* \in X$. For any points x_1, x_2 in U we have

$$\begin{aligned} g(x_1) - g(x_2) &\in \left[\text{w}^*\text{-}\operatorname{cl}\operatorname{co}\partial^1 g(]x_1,x_2[)\right](x_1 - x_2) \\ &= \text{w-}\operatorname{cl}\operatorname{co}\left[\partial^1 g(]x_1,x_2[)(x_1 - x_2)\right] \\ &= \left[\partial^1 g(]x_1,x_2[)\right]^1 (x_1 - x_2), \end{aligned}$$

where $\partial^1 g(]x_1,x_2[) \stackrel{\text{def}}{=} \bigcup_{0<s<1} \partial^1 g(sx_1 + (1-s)x_2)$.

Proof Let $\Gamma \stackrel{\text{def}}{=} \partial^1 g(]x_1, x_2[)$, and $C \stackrel{\text{def}}{=} [\text{w*-}\operatorname{cl}\operatorname{co}\Gamma](x_1 - x_2)$. Note from Proposition 3.17.1 that Γ is bounded, hence $\text{w*-}\operatorname{cl}\operatorname{co}\Gamma$ is w*-compact. Consequently C is w-closed, and we have

$$C \subset \text{w-}\operatorname{cl}\operatorname{co}[\Gamma(x_1 - x_2)] \subset \text{w-}\operatorname{cl} C = C,$$

i.e. $C = \text{w-}\operatorname{cl}\operatorname{co}[\Gamma(x_1 - x_2)]$. Proposition 3.5.2 shows $C = \Gamma^1(x_1 - x_2)$, so it only remains to be seen that $g(x_1) - g(x_2) \in C(x_1 - x_2)$.

For each $\lambda \in Y^*$, the mean value theorem Proposition 3.9 provides $s \in (0,1)$ such that

$$\begin{aligned} \lambda[g(x_1) - g(x_2)] &\in \partial(\lambda g)(sx_1 + (1-s)x_2)(x_1 - x_2) \\ &= \lambda\partial^1 g(sx_1 + (1-s)x_2)(x_1 - x_2), \end{aligned}$$

where equality is given by Proposition 3.17.3. Therefore, for each λ

$$\lambda[g(x_1) - g(x_2)] \in \lambda C.$$

Since C is w-closed and convex, the primal version [Sch, Ch. II §9.1 Cor. 1] of Theorem 3.3.1 assures us that $g(x_1) - g(x_2)$ is a member of C.

□

Now we come to a chain rule for composite vector functions.

Proposition 3.28 *Consider functions $F : X \to \mathbb{R}^n$ and $G : \mathbb{R}^n \to Y$, where F is Lipschitz near $x_* \in X$ and G is Lipschitz near $z_* \stackrel{\text{def}}{=} F(x_*)$.*

The composition $G \circ F : X \to Y$ is Lipschitz near x_*. Also

1. *If $Y = \mathbb{R}$ then both G and $G \circ F$ are real functions with generalized gradients satisfying*

$$\partial(G \circ F)(x_*) \subset \text{w*-}\operatorname{cl}\operatorname{co}[\partial G(z_*) \circ \partial^1 F(x_*)]. \tag{3.14}$$

If $\partial G(z_*)$ is a singleton, equality holds (and the operation w-cl co is superfluous).*

2. $\partial^1(G \circ F)(x_*) \subset [\partial^1 G(z_*) \circ \partial^1 F(x_*)]^1$. *If* $\partial^1 G(z_*)$ *is a singleton, equality holds.*

Proof $F \circ G$ is clearly Lipschitz near x_*.

1. Let u be an arbitrary point of X and

$$\begin{aligned}\alpha_u &\stackrel{\text{def}}{=} \max\{\zeta A u \mid \zeta \in \partial G(z_*),\ A \in \partial^1 F(x_*)\} \\ &= \max\{\xi u \mid \zeta \in \partial G(z_*),\ \xi \in \partial(\zeta F)(x_*)\} \quad \text{by Proposition 3.17.3} \\ &= \max\{(\zeta F)^\circ(x_*; u) \mid \zeta \in \partial G(z_*)\}.\end{aligned}$$

By Theorem 3.3.1, the inclusion (3.14) will hold if

$$(G \circ F)^\circ(x_*; u) \leq \alpha_u.$$

So choose $x_i \to x_*$ and $t_i \downarrow 0$ be such that

$$\lim_i \frac{G \circ F(x_i + t_i u) - G \circ F(x_i)}{t_i} = (G \circ F)^\circ(x_*; u),$$

and consider the quotients $(G \circ F(x_i + t_i u) - G \circ F(x_i))/t_i$. By the mean value theorem Proposition 3.9 there is a point z_i on the line segment from $F(x_i)$ to $F(x_i + t u_i)$ and a generalized subgradient $\zeta_i \in \partial G(z_i)$ such that

$$\frac{G \circ F(x_i + t_i u) - G \circ F(x_i)}{t_i} = \zeta_i \frac{F(x_i + t_i u) - F(x_i)}{t_i}. \tag{3.15}$$

Since the sequence (ζ_i) is bounded (apply Proposition 3.18) and lies in an n-dimensional space, we may assume without loss of generality that it converges in norm to some ζ. In fact $\zeta \in \partial G(z_*)$, because ∂G is a closed set valued mapping (Proposition 3.18) and $z_i \to z_*$.

Similarly, π_z is the projection of $BL(X \times Z, Y)$ onto $BL(Z, Y)$.

Proposition 3.29 *Let G and Z be as above.*

1. *Let $G_x \stackrel{\text{def}}{=} G(\,\cdot, z_*)$. Then $\partial^1 G_x(x_*) \subset \partial^1_x G(x_*, z_*)$.*

2. *$\partial^1_x G(x_*, z_*) = [\pi_x \partial^1 G(x_*, z_*)]^1$.*

Proof

1. This follows from the fact that

$$(\lambda G_x)^\circ(x_*, z_*; u) \leq (\lambda G)^\circ(x_*, z_*; u, 0).$$

2. From Theorem 3.17.3

$$\begin{aligned}(\lambda G)^\circ(x_*, z_*; u, 0) &= \max\{\lambda(Au + B0) \mid (A, B) \in \partial^1 G(x_*, z_*)\} \\ &= \max\{\lambda Au \mid A \in \pi_x \partial^1 G(x_*, z_*)\},\end{aligned}$$

i.e. $(\lambda G)^\circ(x_*, z_*; u, 0)$, as a function of of (u, λ), is the rank-1 support function of $\pi_x \partial^1 G(x_*, z_*)$. By definition of $\partial^1_x G(x_*, z_*)$ we must have

$$\partial^1_x G(x_*, z_*) = [\pi_x \partial^1 G(x_*, z_*)]^1.$$

□

3.7 Extensions

The essential requirement on Y is not really reflexivity, but rather that Y be the dual, under an appropriate polar topology such as the strong topology, of another convex space. Below we discuss the special case of replacing Y by its strong bidual Y^{**}, i.e. the strong dual of Y^*. Using a strong dual also applies in the context of the generalized Jacobian, as seen by the concluding remarks of [Thi82, §3].

We outline a possibility when Y, though normed, is not such a dual space (in particular, Y is not reflexive). Embed each point y of Y in its strong bidual Y^{**} as $\tilde{y} : \lambda \mapsto \lambda y$. Define $\tilde{g} : X \to Y^{**} : x \mapsto \widetilde{g(x)}$; then $\tilde{g}$ is locally Lipschitz iff g is locally Lipschitz. If we omit the reflexivity condition on Y and replace Y by Y^{**}, the latter being a strong dual space, many of our previous results go through. For example, the rank-1 generalized Jacobian would become

$$\partial^1 \tilde{g}(x_*) \stackrel{\text{def}}{=} \{A \in L(X, Y^{**}) \mid \forall (u, \lambda) \in X \times Y^*, \;\; (Au)\lambda \leq (\lambda g)^\circ(x_*; u)\},$$

and Theorem 3.15.1 would be stated:

> If g is Lipschitz near $x_* \in X$, then the mapping $(u, \lambda) \mapsto (\lambda g)^\circ(x_*; u)$ is ... the rank-1 support function of the nonempty, w*-compact, rank-1 representer $\partial^1 \tilde{g}(x_*)$.

We now consider the possibility of X and Y being separated (Hausdorff) locally convex topological vector spaces, not necessarily normable, with Y semi-reflexive. Semi-reflexivity can be dropped as above if we are prepared to consider g as a mapping from X to the bidual Y^{**} under the natural topology such that Y is isomorphically embedded into Y^{**}. $CL(X, Y)$, the space of continuous linear mappings from X to Y, may be strictly larger than the corresponding space of bounded linear mappings. For further details, especially on topology and rank-1 support functionals, see Chapter 2 §3.2.

Let U be a neighborhood in X. The local Lipschitz condition on g generalizes as follows:

g is Lipschitz in U if for each continuous seminorm q on Y there exists a continuous seminorm p on X such that

$$q[g(x_1) - g(x_2)] \leq p[x_1 - x_2], \;\; \forall x_1, x_2 \in U. \tag{3.16}$$

Equivalently we could replace q (p) by the gauge or Minkowski functional μ_V (μ_U) of a convex neighborhood V of $0 \in Y$ (U of $0 \in X$) with $V = -V$ ($U = -U$).

We also need a strong Lipschitz property of g in each direction. In terms of gauge functions:

g is strongly Lipschitz with respect to U in every direction if for each direction $u \in X$ there exists a nonempty bounded set C in Y

$$\mu_C[g(x+tu) - g(x)] \leq t, \quad \forall\, x, x+tu \in U,\ 0 < t \text{ sufficiently small.} \tag{3.17}$$

These conditions are sufficient for obtaining the rank-1 generalized Jacobian, but not always necessary. More subtle conditions specify weak Lipschitz properties, i.e. Lipschitz properties of λg depending on $\lambda \in Y^*$.

Our characterization of rank-1 support functions likewise becomes more complex. Theorem 3.6 is modified as follows.

Let X and Y be separated, locally convex topological vector spaces with Y semi-reflexive, and $P : X \times Y^* \to \mathbb{R}$.

P is the rank-1 support function of a nonempty, equicontinuous, rank-1 representer Γ in $L(X, Y)$ iff

(a) P is positive bihomogeneous;

(b) P is subadditive in $u\lambda$;

(c) for each equicontinuous $\Lambda \subset Y^*$, there is a neighborhood U of X such that P is bounded above on $U \times \Lambda$; and

(d) for each $x \in X$, there is a strong neighborhood W in Y^* such that P is bounded above on $x \times W$.

Such a set Γ is unique and w*-compact. Furthermore, if Y is a metric space then the condition (d) is superfluous.

Theorem 3.15.1 is restated as follows.

> Let X and Y be a separated, locally convex topological vector spaces, with Y semi-reflexive. If g satisfies the Lipschitz conditions (3.16) and (3.17) in a neighborhood U of $x_* \in X$ then the function $(u, \lambda) \mapsto (\lambda g)^\circ(x_*; u)$ is finite valued and satisfies (a)–(d) above; hence is the rank-1 support function of the nonempty, w*-closed, equicontinuous, rank-1 representer $\partial^1 g(x_*)$.
>
> The rank-1 generalized Jacobian is a w*-compact set. If Y is also a metric space, the second Lipschitz condition on g may be omitted.

We sketch the proof, using the original proof of Theorem 3.15.1 as a guide. The boundedness conditions (c) and (d) of the mapping $(u, \lambda) \mapsto (\lambda g)^\circ(x_*; u)$ follow from the first and second Lipschitz conditions on g, respectively. Positive bihomogeneity is clear and, with either (c) or (d), yields finiteness everywhere. We need subadditivity in $u\lambda$ which is shown by the the original proof because

- $(\lambda_i g)^\circ(\cdot\,; u_i)$ is still upper semicontinuous ([Leb]);
- Lemma 3.16 and Proposition 3.11.3 still apply (since, in finite dimensions, all separated, locally convex topologies are equivalent, including norm topologies).

The proof is concluded by applying the modified Theorem 3.6, above.

To finish we give an example combining both of the main ideas above, namely omitting (semi-)reflexivity of Y and going beyond normed spaces.

Example 3.30 *Let $f : X \to \mathbb{R}$ be a differentiable function on a metric space X, and let X^* have the strong dual topology. Suppose that the gradient $\nabla f : X \to X^*$ of f is Lipschitz, as in (3.16), on a neighborhood U of $x_* \in X$.*

Then the rank-1 generalized Jacobian of ∇f *exists at each point of* U *as a nonempty, convex, equicontinuous hence* w^**-compact subset of* $CL(X, X^*)$.

This follows from the above extension of Theorem 3.15.1 for $Y = X^*$, in which the semi-reflexivity requirement on Y can be dropped because Y is a strong dual space.

Chapter 4

A Proof of Robinson's Homeomorphism Theorem for Piecewise Linear Maps

4.1 Introduction

We study a certain class of piecewise linear functions from $\mathbb{R}^n$ to $\mathbb{R}^n$, called, in the terminology of [Rob90], *normal* maps induced by linear mappings and polyhedral convex sets. We will call these *pl-normal* maps for brevity. Such systems are important in many optimization and equilibrium problems. They arise directly from variational inequalities, or equivalently generalized equations, specified by linear maps and polyhedral convex sets; and indirectly as approximations to such systems specified by smooth nonlinear functions over polyhedral convex sets. Perhaps the best known of these is the *linear complementarity problem.* See [Rob90] for further details.

Robinson's homeomorphism theorem [Rob90, Thm. 4.3] characterizes the pl-normal maps which are homeomorphisms, i.e. gives conditions for unique solvability of such systems. Here we provide a new, shorter proof of this result. Some other work on necessary and sufficient conditions for piecewise linear maps to be homeomorphic is found in [FK], [RV], [KS], [Schr] and [KL], of which we will find

[KL] particularly useful.

In this section we will specify our basic notation and quote the characterization theorem for pl-normal maps. Also some minor results will be listed for later application. Later sections are

§4.2 Piecewise Linear Homeomorphism Results

§4.3 When C is a Convex Cone, dealing with the special case of pl-normal maps when the underlying polyhedral convex set C is a cone.

§4.4 The Proof of the Theorem.

Throughout this chapter we will assume that A is a given real $n \times n$ matrix, and C is a nonempty polyhedral convex set in $\mathbb{R}^n$. The projection mapping which sends a point of $\mathbb{R}^n$ to its nearest point in C is denoted π_C.

Definition 4.1 *The* normal *map induced by (A,C) is the mapping from* $\mathbb{R}^n$ *to* $\mathbb{R}^n$ *given by*

$$A_C \stackrel{\text{def}}{=} A\pi_C + I - \pi_C$$

where I *is the* $n \times n$ *identity matrix. Such normal maps are called* pl-normal *maps.*

Proposition 4.5.1 shows that pl-normal maps are indeed piecewise linear.

As mentioned above, questions about solvability of $A_C(x) = y$ can be cast in terms of variational inequalities or generalized equations: let $y \in \mathbb{R}^n$, then

$$\begin{aligned} & \exists x \in \mathbb{R}^n, \quad y = A_C(x) \\ \iff \; & \exists c_0 \in C, \quad \forall c \in C, \quad 0 \leq \langle c - c_0, Ac_0 - y \rangle \\ \iff \; & \exists c_0 \in C, \quad y \in Ac_0 + N_C(c_0) \end{aligned}$$

where the normal cone $N_C(c_0)$ is defined below.

For clarity we specify the following terminology.

Definition 4.2 *Let $F : \mathbb{R}^n \to \mathbb{R}^n$.*

1. *A convex set in $\mathbb{R}^n$ is called* polyhedral *if it is the convex hull of finitely many points. An n-*cell *is a polyhedral convex set with nonempty interior.*

2. *F is* piecewise linear *if there are finitely many n-cells $(C_i)_i$ covering $\mathbb{R}^n$ such that for each i there exist $A_i \in \mathbb{R}^{n\times n}$, $b_i \in \mathbb{R}^n$ satisfying*

$$\forall y \in C_i, \quad F(y) = A_i y + b_i$$

3. *Suppose F is piecewise linear, and the matrices A_i are as above. The* determinants *of F are the determinants of the matrices A_i.*

   F is coherently oriented *if its determinants all have the same (nonzero) sign.*

4. *F is a* Lipschitzian homeomorphism *if it is bijective, and both it and its inverse F^{-1} are Lipschitz mappings. F is a Lipschitzian homeomorphism from $X \subset \mathbb{R}^n$ to $Y \subset \mathbb{R}^n$ if $F(X) \subset Y$ and $F|_X : X \to Y$ is a Lipschitzian homeomorphism.*

Since a piecewise linear map F is continuous on each of finitely many sets covering its domain, each of which is closed, it is easy to see that F must be continuous.

We can now state Robinson's characterization theorem.

Theorem 4.3 *The piecewise linear mapping A_C is a Lipschitzian homeomorphism iff A_C is coherently oriented.*

The result will be demonstrated in §4.4. [KL, Thm. 5.3] gives a similar result for another class of piecewise linear mappings, namely those with *branching number* less than or equal to 4. The branching number of a piecewise linear map is the least natural number b such that every face of the n-cells C_i (as in the above definition) which has codimension two is contained in at most b n-cells. The relationship between this class of mappings and the class of pl-normal maps is not known.

In fact for bijective piecewise linear mappings, (Lipschitz) continuity of the mapping and its inverse is automatic:

Lemma 4.4 *[FK, §2 and Lemma 2] A bijective piecewise linear mapping from* $\mathbb{R}^n$ *to* $\mathbb{R}^n$ *is a Lipschitzian homeomorphism.*

Further notation:

- Let $\emptyset \neq K \subset \mathbb{R}^n$. K is a *cone* if $\alpha K \subset K$ for each $\alpha > 0$. The *polar* cone of K is

$$K^\circ \stackrel{\text{def}}{=} \{x \in \mathbb{R}^n \mid \langle x, k\rangle \leq 0,\ \forall k \in K\}$$

- The *normal* cone to C at $c \in C$ is

$$N_C(x) \stackrel{\text{def}}{=} \{x \in \mathbb{R}^n \mid \langle x, c - \bar{c}\rangle \leq 0,\ \forall c \in C\}$$

  and the *tangent* cone to C at c is $T_C(x) \stackrel{\text{def}}{=} N_C(x)^\circ$.

- A *face* of C is a set $\emptyset \neq F \subset C$ such that for $x_1, x_2 \in C$ and $0 < t < 1$, if $tx_1 + (1-t)x_2$ belongs to F then x_1, x_2 belong to F.

- $\mathcal{N}_C$ is the family of sets $F + N_F$, where F is a face of C and N_F is the set $N_C(f)$ for any relative interior point f of F (see Proposition 4.5.1 below).

- The *critical cone* to C at $x \in \mathbb{R}^n$ is

$$C_x \stackrel{\text{def}}{=} T_C(\pi_C(x)) \cap [x - \pi_C(x)]^\perp.$$

- The *lineality* space of C, $\operatorname{lin} C$, is the largest subspace L in $\mathbb{R}^n$ such that $L + C \subset C$.

The results below, though not too hard to prove, give a lot of insight into the structure of pl-normal maps.

Proposition 4.5 *[Rob90, Prop. 2.1, 2.2, 2.3]*

1. *For each face F of C, and each point f in the relative interior of F, the normal cone to C at f is the same (nonempty) set. Denote this by N_F.*

2. *$\mathcal{N}_C$ consists of finitely many n-cells whose union covers $\mathbb{R}^n$. On each such n-cell $F + N_F$ where F is face of C, A_C acts as an affine mapping*

$$A_C(f + f^*) = Af + f^*, \quad \forall f \in F, f^* \in N_F$$

Part 2 of the proposition is also given as [BM, Thm. 2.3].

Lemma 4.6 *For each $x \in \mathbb{R}^n$ there exists a neighborhood U of $0 \in \mathbb{R}^n$ such that*

$$A_C(x + d) = A_C(x) + A_{C_x}(d), \quad \forall d \in U$$

Proof By [Rob89, Cor. 4.5], there is a neighborhood U of x such that

$$\pi_C(x + d) = \pi(x) + \pi_{C_x}(d),$$

for each d in U. The result is obtained by substituting the expression on the right into the definition of A_C. $\square$

From Lemma 4.6 we see that A_C is one-to-one near x iff A_{C_x} is one-to-one (near 0); and further that the determinants of A_{C_x} are given by the determinants of A_C near x.

4.2 Piecewise Linear Homeomorphism Results

We review some standard definitions.

Definition 4.7

1. *A topological space Y is* simply connected *if it is path-connected and any continuous paths $p : [0,1] \to Y$, $q : [0,1] \to Y$ with the same endpoints are homotopic, i.e. there exists a continuous mapping $H : [0,1] \times [0,1] \to Y$ such that $H(0, \cdot) = p$ and $H(1, \cdot) = q$.*

2. *Let $F : X \to Y$ where X, Y are topological spaces. Let $\hat{Y} \subset Y$. F is* proper *with respect to $\hat{Y}$ if for any set $C \subset \hat{Y}$ which is compact, $F^{-1}(C) \stackrel{\text{def}}{=} \{x \in X \mid F(x) \in C\}$ is compact in X.*

It is well known [Arm] that $\mathbb{R}^n \backslash \{0\}$ is simply connected for $n \geq 3$. We also note that, in the second definition, a set in $\hat{Y}$ is compact in the subspace topology iff it is compact in Y; so we will use whichever compactness condition is convenient.

Our first result lays the foundation for this chapter. Its proof is built almost entirely from [KL], which gives a nice introduction to covering maps [KL, §4]. Recall that a map $F : U \to Y$, where U and Y are sets in $\mathbb{R}^n$, is *positive homogeneous* if for $\alpha > 0$ and $u \in U$, $\alpha u \in U$ implies $F(\alpha u) = \alpha F(u)$.

Theorem 4.8 *Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a piecewise linear mapping and $Y \subset \mathbb{R}^n$ be an open, simply connected set such that $X \stackrel{\text{def}}{=} F^{-1}(Y)$ is connected. Then F is a Lipschitzian homeomorphism from X onto Y iff F is locally one-to-one on X.*

Proof Clearly if F maps X homeomorphically to Y then it is locally one-to-one on X. We set the converse out in stages.

1. Invariance of Domain: Let $x_0 \in X$. X is open because Y is open and F is continuous. By piecewise linearity of F, in a small enough neighborhood U of 0, $x_0 + U \subset X$ and the mapping $F(x_0 + u) - F(x_0)$ for $u \in U$ is positive homogeneous and one-to-one. This mapping on U has a (unique) piecewise linear extension $F_0 : \mathbb{R}^n \to Y$ which is positive homogeneous and one-to-one. Now apply [KL, Lem. 2.2] and [KL, Thm. 5.1] in succession to see that F_0 is actually an open mapping, hence that F is a homeomorphism from U to a neighborhood of $F(x_0)$.

2. F is proper with respect to Y: Let C be compact set in Y, then C is closed and bounded. So $F^{-1}(C)$ is closed by continuity of F. Also F is piecewise linear and locally one-to-one, therefore it can be decomposed into finitely many one-to-one affine maps. It is easy to see that for such a map $F^{-1}(C)$ is bounded so $F^{-1}(C)$ is compact as needed.

3. F is a *covering*: By [KL, §4.2], originally [Bro], F maps X onto Y and X is the disjoint union of countably many open sets, each of which is mapped homeomorphically by F onto Y.

4. F is a Lipschitzian homeomorphism from X onto Y: We use the well known result [Arm] that every covering of a locally and simply connected space is a homeomorphism. Now Lemma 4.4 says that F is actually a Lipschitzian homeomorphism from X onto Y.

$\square$

This kind of result can also be derived from more general results. For example it is a corollary of the next theorem.

Theorem 4.9

1. *Let $F : X \to Y$ where X, Y are topological spaces. Let $\hat{Y}$ be an open set in Y that is locally and simply path-connected, and define $\hat{X} \stackrel{\text{def}}{=} F^{-1}(\hat{Y})$. Then F is a homeomorphism of each path-connected component of $\hat{X}$ onto $\hat{Y}$ if F is a local homeomorphism on $\hat{X}$ and is proper with respect to $\hat{Y}$.*

2. *(Invariance of Domain) Let $F : X \to Y$ be a continuous mapping, where X and Y are open sets in $\mathbb{R}^n$. Then F is a local homeomorphism iff F is locally one-to-one.*

Proof

1. We only need show that F restricted to $\hat{X}$ has the *continuation* property for continuous paths in $\hat{Y}$. For suppose this holds. We observe that $F(\hat{X})$ is open in Y, hence in $\hat{Y}$, because F is a local homeomorphism on $\hat{X}$ where the set $\hat{X} = F^{-1}(\hat{Y})$ is open by continuity of F and openness of $\hat{Y}$. As $\hat{Y}$ is also path connected we deduce [Rhei, Thm 2.11] that $F(\hat{X}) = \hat{Y}$. The result is then given by [Rhei, Thm 3.7] applied to the mapping of F from $\hat{X}$ to $\hat{Y}$.

   For the continuation property, let $p : [0,1] \to \hat{Y}$ be continuous, $0 < t_* \leq 1$, and $q : [0, t_*) \to \hat{X}$ be continuous path such that

   $$F(q(t)) = p(t), \quad \forall t \in [0, t_*).$$

   We must show there is a sequence $(t_n) \uparrow t_*$ such that $q(t_n)$ converges to a point $q_* \in \hat{X}$ with $F(q_*) = p(t_*)$.

   Since $p[0, t_*]$ is a compact set in Y, then $q[0, t_*) \subset F^{-1}(p[0, t_*])$ has compact closure in $\hat{X}$, thus for some sequence $(t_n) \subset [0, t_*), t_n \uparrow t_*$, we know there exists $\lim q(t_n) = q_* \in \hat{X}$. By continuity of F and p we must have $F(q_*) = p(t_*)$, as required.

2. [Dold, Ch. IV.8]

□

[Rhei] provides self contained and even elementary arguments relating to part 1 of Theorem 4.9. Part 2, a variant of the classical Invariance of Domain theorem, is however much deeper and more difficult.

Corollary 4.10 *Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a piecewise linear mapping. Then F is a Lipschitzian homeomorphism iff F is locally one-to-one.*

Proof Take $Y = \mathbb{R}^n$ in Theorem 4.8. □

Corollary 4.11 *Let $n \geq 3$ and $F : \mathbb{R}^n \to \mathbb{R}^n$ be a positive homogeneous piecewise linear mapping. Then F is a Lipschitzian homeomorphism iff it is locally one-to-one except perhaps at 0.*

Proof Necessity of local injectivity is clear. For sufficiency suppose F is locally one-to-one except perhaps at 0. Note that continuity and positive homogeneity of F yield that $F(0) = 0$.

Since $n \geq 3$, $Y \stackrel{\text{def}}{=} \mathbb{R}^n \backslash \{0\}$ is simply connected (and open). Let $X \stackrel{\text{def}}{=} F^{-1}(Y)$ and note that $X = Y$. Otherwise some nonzero point x of $\mathbb{R}^n$ is mapped to zero by F hence, by positive homogeneity of F, every neighborhood of x contains distinct points (eg. on the line segment $[0, x]$) with the same (zero) function value — this contradicts the fact that F is locally one-to-one on $\mathbb{R}^n \backslash \{0\}$. Hence F is locally one-to-one on the connected set X. We apply Theorem 4.8 to show that F is a homeomorphism from $\mathbb{R}^n \backslash \{0\}$ onto $\mathbb{R}^n \backslash \{0\}$.

In fact F is one-to-one since no distinct points of $\mathbb{R}^n \backslash \{0\}$ have the same image under F and the zero vector is the only point whose image is zero. Corollary 4.10 completes the proof.

□

4.3 When C is a Convex Cone

Throughout this section assume K is a nonempty polyhedral convex cone in $\mathbb{R}^n$. We say that K is *pointed* if $\operatorname{lin} K = \{0\}$. We will use Proposition 4.5 without reference, especially the facts that $\mathcal{N}_K$ consists of n-cells covering $\mathbb{R}^n$, and A_K is a piecewise linear mapping which is affine on each of the n-cells of $\mathcal{N}_K$.

The eventual proof that A_K is homeomorphic if it is coherently oriented will rely on an induction argument for $n \geq 3$, the hypothesis for which is

> Let $1 \leq m \leq n-1$, $B \in \mathbb{R}^{m \times m}$, and D be a nonempty polyhedral convex cone in $\mathbb{R}^m$. If B_D is coherently oriented then B_D is a Lipschitzian homeomorphism. (4.18)

We need to start the induction process.

Lemma 4.12 *If n is 1 or 2 and A_K is coherently oriented then A_K is a Lipschitzian homeomorphism.*

Proof The case n=1 is trivial. For $n = 2$ we appeal to the argument at the beginning of [KL, proof of Thm. 5.3] which shows that for a positive homogeneous piecewise linear map $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ which has no more than 4 affine pieces, F is a homeomorphism iff it is proper and coherently oriented. Of course if F is coherently oriented it is proper too (eg. see proof of Theorem 4.8), so the requirement that F be proper is superfluous. Since for $n = 2$ the positive homogeneous map A_K has no more than 4 affine pieces — $\mathcal{N}_K$ contains 1 n-cell if K is 0 or $\mathbb{R}^n$, 2 if K is a ray or a halfspace, and 4 otherwise — we are done.

□

Proposition 4.13 *Suppose (4.18) holds. If A_K is coherently oriented and either K or K° is not pointed then A_K is one-to-one.*

Proof Suppose that K is not pointed, so the dimension of $L \stackrel{\text{def}}{=} \operatorname{lin} K$ is at least 1. The result is given by a straightforward argument in the proof of [Rob90, Thm. 4.3] (with K replacing C). [Rob90, Prop. 4.1] is used to factor out the lineality space of K, then the induction hypothesis (4.18) to find that the reduced mapping (on a space of lower dimension) is one-to-one. From this it follows that A_K is one-to-one as needed.

Now suppose K° is not pointed and recall that $K + N_K$ is an n-cell of $\mathcal{N}_K$. Let the restriction of A_K to $K + N_K$ be represented by $\bar{A} \in \mathbb{R}^{n\times n}$. As $\det\bar{A} \neq 0$ we can define

$$P \stackrel{\text{def}}{=} (\bar{A}^{-1})_{K^\circ}$$

Since $\pi_K = I - \pi_{K^\circ}$ we have

$$\begin{aligned} P &= \bar{A}^{-1}\pi_{K^\circ} + I - \pi_{K^\circ} = \bar{A}^{-1}(I - \pi_K) + \pi_K \\ &= \bar{A}^{-1}(\bar{A}\pi_K + I - \pi_K) = \bar{A}^{-1}A_K \end{aligned}$$

where the last equality follows because $\bar{A}\pi_K = A\pi_K$. As $\det\bar{A}^{-1} \neq 0$ and the determinants of A_K all have the same sign, we see that the determinants of P all have the same sign. By the above P must be one-to-one on $\mathbb{R}^n$, hence $A_K = \bar{A}P$ is one-to-one also.

□

Proposition 4.14 *If A_K is coherently oriented then it is a Lipschitzian homeomorphism.*

Proof The result is valid for dimension $n = 1, 2$ by Lemma 4.12. Now suppose it is valid for dimension $n - 1$, i.e. (4.18) holds, where $n \geq 3$. We will show it holds for dimension n, whence by induction it holds for all natural numbers n.

Suppose $0 \neq x \in \mathbb{R}^n$ and recall the critical cone to K at x,

$$K_x \stackrel{\text{def}}{=} T_K(\pi_K(x)) \cap [x - \pi_K(x)]^{\perp}.$$

We will show that A_{K_x} is one-to-one hence, by Lemma 4.6, A_K is locally one-to-one at x. If $x \notin K^\circ$ then $\pi_K(x) \neq 0$ and the (nontrivial) subspace spanned by $\pi_K(x)$ is contained in K_x. So K_x is not pointed. By Proposition 4.13 A_{K_x} is one-to-one, hence A_K is one-to-one near x. On the other hand if $x \in K^\circ$ then $[x - \pi_K(x)] = x \neq 0$, so $(K_x)^\circ$ contains the line spanned by x. Applying Proposition 4.13 again, we can see that A_K is one-to-one near x.

Therefore the only point at which A_K is perhaps not locally one-to-one is 0. Corollary 4.11 completes the proof.

□

4.4 The Proof of the Theorem

We restate Theorem 4.3:

Theorem *The piecewise linear map A_C is a homeomorphism iff it is coherently oriented.*

Proof To start with, recall that A_C is a piecewise linear mapping (Proposition 4.5.2).

(if) This is not hard to show and in fact holds for general piecewise linear maps [KL, Lemma 2.2].

(only if) Given Corollary 4.10 it is sufficient to show that A_C is locally one-to-one. So let $x \in \mathbb{R}^n$ and consider the critical cone to C at x:

$$K \stackrel{\text{def}}{=} T_C(\pi_C(x)) \cap [x - \pi_C(x)]^{\perp}.$$

First, from Lemma 4.6, A_K inherits coherent orientation from A_C. Then, by Proposition 4.14, A_K is one-to-one. Using Lemma 4.6 again says A_C is one-to-one near x and we are done.

$\square$

Acknowledgement

I would like to thank Stephen M. Robinson for drawing my attention to recent work in this area, especially to the subject of this chapter [Rob90, Thm. 4.3], and for the suggestions he has made to clarify the new proof presented here.

Bibliography

[Arm] Armstrong M.A., *Basic Topology*, Springer, New York, (1983).

[AV] Aubin J.-P., Ioffe's fans and generalized derivatives of vector-valued maps, in J.-P. Aubin and R.B. Vinter, eds., *Convex Analysis and Optimization*, Pitman, Boston, pp. 1-17 (1982).

[BM] Burke J.V. and Moré J.J., On the identification of active constraints, *SIAM J. Numer. Anal.* **25**, pp. 1197-1211 (1988).

[Bré] Brézis H., *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, North-Holland Mathematics Studies No. 5, North-Holland, Amsterdam (1973).

[Bro] Browder F.E., Covering spaces, fibre spaces, and local homeomorphisms, *Duke Math. J.* **21**, pp. 329-336 (1954).

[Chr] Christensen J.P.R., *Topology and Borel Structure.* Notas de Matemática (51), L. Nachbin, ed., American Elsevier, New York (1974).

[Cla] Clarke F.H., *Optimization and Nonsmooth Analysis.* John Wiley and Sons, New York (1983).

[CL] Cornet B. and Laroque G., Lipschitz properties of solutions in mathematical programming, *J. Opt. Th. Appl.* **53**, pp. 407-427 (1987).

[Dold] Dold A., *Lectures in Algebraic Topology.* Springer Verlag, New York (1972).

[FK] Fujisawa T. and Kuh E.S., Piecewise-linear theory of resistive networks, *SIAM J. Applied Math.* **22**, pp. 307-328 (1972).

[H-U82] Hiriart-Urruty J.-B., Characterizations of the plenary hull of the generalized Jacobian matrix, *Math. Prog. Study* **17**, pp. 1-12 (1982).

[H-U84] Hiriart-Urruty J.-B., Strodiot J.-J. and Hien Nguyen V., Generalized Hessian matrix and second-order optimality conditions for problems with $C^{1,1}$ data, *Appl. Math. Optim.* **11**, pp. 43-56 (1984).

[Hal] Halkin H., Interior mapping theorem with set-valued derivative, *J. Analyse Math.* **30**, pp. 200-207 (1976).

[Hör] Hörmander L., Sur la fonction d'appui des ensembles convexes dans un espace localement convexe, *Arkiv för Matematik* **3**, pp. 181-186 (1954).

[Iof81] Ioffe A.D., Nonsmooth analysis: Differential calculus and nondifferentiable mappings, *Trans. American Math. Soc.* **266**, pp. 1-58 (1981).

[Iof82] Ioffe A.D., Nonsmooth analysis and the theory of fans, in J.-P. Aubin and R.B. Vinter, eds., *Convex Analysis and Optimization*, Pitman, Boston, pp. 93-117 (1982).

[KL] Kuhn D. and Löwen R., Piecewise affine bijections of $\mathbb{R}^n$, and the equation $Sx^+ - Tx^- = y$, *Linear Algebra Appl.* **96**, pp. 109-129 (1987).

[KR] Kutateladze S.S. and Rubinov A.M., Minkowski duality and its applications, *Usp. Matem. Nauk.* **27**, pp. 127-176 (1972).

[KS] Kojima M. and Saigal R., On the relationship between conditions that insure a PL mapping is a homeomorphism, *Math. Operations Res.* **5**, pp. 101-109 (1980).

[Leb] Lebourg G., Valeur moyenne pour gradient généralisé, *C. R. Acad. Sc. Paris* **281**, pp. 795-797 (1975).

[Mang] Mangasarian O.L., Some applications of penalty functions in mathematical programming, Lecture Notes in Mathematics No. 1190, *Optimization and Related Fields*, R. Conti et al, eds., Springer-Verlag, Berlin, pp. 307-329 (1986).

[Min] Minkowski H., Theorie der konvexen Körper, insbesondere Begründung ihres Oberflächenbegriffes, *Ges. Abh.* **II**, Leipzig (1911).

[Mor] Moreau J.J., Proximité et dualité dans un espace hilbertien, *Bull. Soc. Math. France* **93**, pp. 273-299 (1965).

[Neu] Neumaier A., *Interval Methods for Systems of Equations.* To be published by Cambridge University Press.

[Por] Pourciau B.H., Analysis and optimization of Lipschitz continuous mappings, *JOTA* **22**, pp. 311-351 (1977).

[Rhei] Rheinboldt W.C., Local mapping relations and global implicit function theorems, *Trans. Amer. Math. Soc.* **138**, pp. 183-198 (1969).

[Rob89] Robinson S.M., An implicit function theorem for a class of nonsmooth functions, to appear in *Mathematics of Operations Research.*

[Rob90] Robinson S.M., Normal maps induced by linear transformations, manuscript, Department of Industrial Engineering, UW-Madison, Madison, WI

[Roc74] Rockafellar R.T., *Conjugate Duality and Optimization.* Regional Conference Series in Applied Mathematics **16**, SIAM, Philadelphia PA (1974).

[Roc79] Rockafellar R.T., Directionally Lipschitzian functions and subdifferential calculus, *Proc. London Math. Soc.* **39**, pp. 331-355 (1979).

[Roc80] Rockafellar R.T., Generalized directional derivatives and subgradients of nonconvex functions, *Canadian J. Math.* **32**, pp. 257-280 (1980).

[Roc85] Rockafellar R.T., Extensions of subgradient calculus with applications to optimization, *Nonlin. Anal. Theory, Methods and Appl.* **9**, pp. 665-698 (1985).

[RV] Rheinboldt W.C. and Vandergraft J.S., On piecewise affine mappings in $\mathbb{R}^n$, *SIAM J. Applied Math.* **29**, pp. 680-689 (1975).

[Rub] Rubinov A.M., Sublinear operators and operator-convex sets, *Siberian Math. J.* **17**, pp. 289-296 (1976).

[Sch] Schaefer H.H., *Topological Vector Spaces.* Macmillan, NY (1966).

[Schr] Schramm R., On piecewise linear functions and piecewise linear equations, *Math. Operations Res.* **5**, pp. 510-522 (1980).

[Swe77] Sweetser T.H. III, A minimal set-valued strong derivative for vector valued Lipschitz functions, *JOTA* **23**, pp. 549-567 (1977).

[Swe79] Sweetser T.H. III, *A set-valued strong derivative in infinite dimensional spaces, with applications in Hilbert spaces*, Ph.D. thesis, University of California, San Diego, CA (1979).

[Thi75] Thibault L., Quelques propriétés des sous-differentiels de fonctions localement Lipschitziennes, *Seminaire d'Analyse Convexe Montpellier 1975*, Exposé no. 16 (1975).

[Thi80] Thibault L., Subdifferentials of compactly lipschitzian vector-valued functions, *Ann. Math., Pura Appl.* **125**, pp. 157-192 (1980).

[Thi82] Thibault L., On generalized differentials and subdifferentials of Lipschitz vector-valued functions, *Nonlin. Anal., Theory, Methods and Appl.* **6**, pp. 1037-1053 (1982).

[War] Warga J., Derivate containers, inverse functions and controllability, in D.L. Russell, ed., *Calculus of Variations and Control Theory*, Academic Press, New York, pp. 13-46 (1976).