**RESEARCH ACCOMPLISHMENTS AND PLAN**


# PANG CHIN HOW JEFFREY


CONTENTS



I work on theoretical and numerical aspects of continuous optimization. My recent work has focused on variational analysis (more specifically, set-valued analysis and semi-algebraic variational analysis) and numerical methods for the mountain pass problem. Applications of my work in set-valued analysis include sensitivity analysis, optimal control and the calculus of variations, which will be explained in Sections 1 to 3. Section 1 focuses on my work on the foundations of set-valued maps. Section 2 focuses on differential inclusions (which in turn forms the basis for modern calculus of variations and optimal control), and Section 3 focuses on semi-algebraic variational analysis. Applications of my work in the numerical mountain pass problem include numerical PDEs and computational chemistry, which will be explained in Section 4. Section 5 describes my first project on the variational analysis of pseudospectra, which led to my current work in the two separate directions. Finally, Section **??** describes some future plans.

## 1. SET-VALUED ANALYSIS AND APPLICATIONS IN OPTIMIZATION

Variational analysis is the "broad spectrum of modern mathematical theory that has grown in connection with the study of problems of optimization, equilibrium, control and stability of linear and nonlinear systems" as presented in the Lanchester prize winning work [RW98]. An important theme in variational analysis is that of handling the nonsmoothness commonly appearing in such problems. Several other works like [BZ05, Cla83, CLSW98, DR09a, Mor06] deal with other important aspects of variational analysis.

Much of my work involves set-valued maps. See also [AF90]. We say that $S$ is a *set-valued map* or a *multifunction*, denoted by $S : X \rightrightarrows Y$, if $S(x) \subset Y$ for all $x \in X$. There are many examples of set-valued maps in optimization and related areas. For example, the generalized derivatives of a nonsmooth function, the feasible set of a parametric optimization problem, and the set of optimizers to a parametric optimization problem may be profitably viewed as set-valued maps. The *graph* of $S$, denoted by $\mathrm{Graph}(S)$, is the set $\{(x,y)\in X\times Y \mid y\in S(x)\}$.

1.1. **Background: Marginal functions in optimization.** For a set-valued map $S:\mathbb{R}^n\rightrightarrows\mathbb{R}^m$ and function $\varphi:\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}$, the *marginal function* $f:\mathbb{R}^n\to\mathbb{R}$ is defined by

$$f(x):=\min_{y\in S(x)}\varphi(x,y). \tag{1}$$

The generalized differentiability properties of $f$ can be deduced from the generalized differentiability properties of both $\varphi$ and $S$. As an example, in the case of a linear $\varphi$ and $S:\mathbb{R}^n\rightrightarrows\mathbb{R}^m$ defined by

$$\begin{aligned} S(x) \quad &:= \quad \{y:F_i(y)\le x_i \text{ for } i=1,\dots,n\},\\ &\qquad\text{where } F:\mathbb{R}^m\to\mathbb{R}^n \text{ is smooth.}\end{aligned}$$

This set-valued map represents the dependence of the feasible set in a linear programming problem on its right hand side. Using the theory in [Pan11a], we can calculate that if the Mangasarian-Fromovitz constraint qualification is satisfied and $H:\mathbb{R}^m\rightrightarrows\mathbb{R}^n$ is a prefan, then $S$ is pseudo strictly $H$-differentiable at $(\bar u,\bar x)$ if $H$ is such that for all $p\in\mathbb{R}^m\backslash\{0\}$, there exists $q\in -H(-p)$ such that $\nabla F(\bar y)q-p\in T_{\mathbb{R}^n_-}(F(\bar y)-\bar x)$.

We move on to the general problem on the sensitivity analysis of $f$ with respect to $x$ in (1). Such sensitivity analysis can be analyzed using the theory of set-valued maps. We denote the composition $S_2\circ S_1:X\rightrightarrows Z$ of set-valued maps $S_1:X\rightrightarrows Y$ and $S_2:Y\rightrightarrows Z$ in the usual way by

$$S_2\circ S_1(x):=\bigcup_{y\in S_1(x)}S_2(y).$$

The epigraphical mapping $E_\varphi:\mathbb{R}^n\times\mathbb{R}^m\rightrightarrows\mathbb{R}$ is defined by $E_\varphi(x):=[\varphi(x,y),\infty)$, and $E_f:\mathbb{R}^n\rightrightarrows\mathbb{R}$ is defined similarly. Define the map $\bar S:\mathbb{R}^n\rightrightarrows\mathbb{R}^n\times\mathbb{R}^m$ by $\bar S(x)=\{x\}\times S(x)$. We then have

$$E_f(x)=E_\varphi\circ\bar S(x). \tag{2}$$

We still need the theory of subdifferentials and coderivatives to study the sensitivity analysis of $f$. The subdifferential mapping $\partial f:\mathbb{R}^n\rightrightarrows\mathbb{R}^n$ in convex and variational analysis is a generalization of the gradient $\nabla f:\mathbb{R}^n\to\mathbb{R}^n$ for smooth functions $f$. (Specifically, $\partial f(x)=\{\nabla f(x)\}$.) The coderivative of $S:\mathbb{R}^n\rightrightarrows\mathbb{R}^m$ at $(\bar x,\bar y)\in\mathrm{Graph}(S)$ is denoted by $D^*S(\bar x\mid\bar y):\mathbb{R}^m\rightrightarrows\mathbb{R}^n$, which is a generalization of the adjoint operator. (Specifically, for $F:\mathbb{R}^n\to\mathbb{R}^m$, $D^*F(\bar x\mid F(\bar x))\equiv\nabla^T F(\bar x)$.) For convex $\varphi$ and graph-convex $S$, the chain rules of set-valued maps applied to (2) gives

$$\begin{aligned}\partial f(\bar x) \quad &= \quad \{x^*+D^*S(\bar x\mid\bar y)(y^*)\mid(x^*,y^*)\in\partial\varphi(\bar x,\bar y)\}\\ &\qquad\text{for any } \bar y\in S(\bar x) \text{ s.t. } \varphi(\bar x,\bar y)=f(\bar x).\end{aligned} \tag{3}$$

1.2. **Generalized differentiation of set-valued maps** [Pan11b]**.** We say that a set-valued map $S$ has the *Aubin property at* $(\bar x,\bar y)\in\mathrm{Graph}(S)$ *with modulus* $\kappa\ge0$ if for every $\delta>0$, there exists neighborhoods $U_\delta$ of $\bar x$ and $V_\delta$ of $\bar y$ such that

$$S(x)\cap V_\delta\subset S(x')+(\kappa+\delta)\|x-x'\|\mathbb{B} \text{ for all } x,x'\in U_\delta,$$

where $\mathbb{B}$ denotes the unit ball. The Aubin property was first defined by Aubin [Aub84] to study the Lipschitz dependence of parameters in optimization problems.

In [Pan11b], I extended the Aubin property: For a positively homogeneous map $H:X\rightrightarrows Y$, we say that $S$ is *pseudo strictly $H$-differentiable at* $(\bar x,\bar y)\in\mathrm{Graph}(S)$ if for every $\delta>0$, there exists neighborhoods $U_\delta$ of $\bar x$ and $V_\delta$ of $\bar y$ such that

$$S(x)\cap V_\delta\subset S(x')+H(x-x')+\delta\|x-x'\|\mathbb{B} \text{ for all } x,x'\in U_\delta. \tag{4}$$

The case where $H(x)=\kappa\|x\|\mathbb{B}$ for some $\kappa\ge0$ reduces to the Aubin property. The rest of this section describes further theory and applications of this generalized derivative.

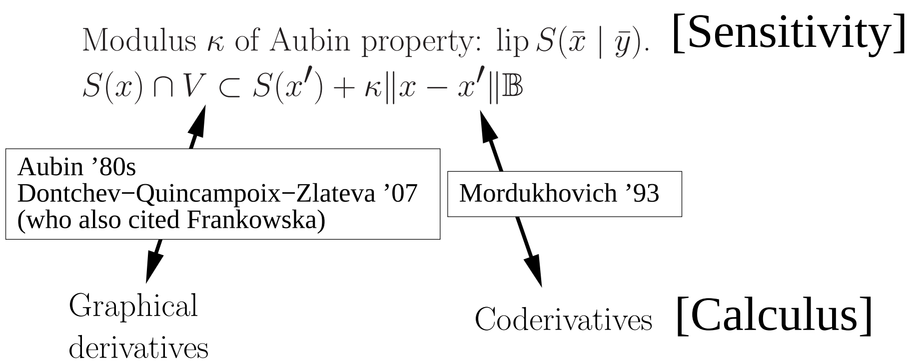


FIGURE 1. The above diagram summarizes the relationship between the Aubin property (defined formally in Subsection 1.2), graphical derivatives (derived from tangent cones of Graph($S$)) and coderivatives (derived from normal cones of Graph($S$)). The relationship between the Aubin property and the graphical derivative is known as the Aubin criterion, while the relationship between the Aubin property and the coderivative is known as the Mordukhovich criterion. The graphical derivatives and coderivatives are easier to calculate and establish calculus rules, while the Aubin property is directly related to sensitivity analysis. Other than those mentioned in the figure, others who have made progress on the relation between the Aubin property and the graphical derivatives and coderivatives include Robinson, Ursescu, Rockafellar, Ioffe, Jourani, Thibault and Shao.

1.3. **Characterizing the generalized derivatives** [Pan11a]**.** Refer to Figure 1 for a background between the relationship between the Aubin property, graphical derivatives and coderivatives. In [Pan11a], I generalized the Aubin and Mordukhovich criterions for characterizing the generalized derivatives in the sense of [Pan11b]. We say that $H : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is a *prefan* if $H$ is positively homogeneous, convex-valued, compact-valued and $H(\mathbb{B}) \subset K\mathbb{B}$ for some $K < \infty$.

Let the collection of possible graphical limits of graphical derivatives $DS(x \mid y) : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ of $S$ as $(x,y) \xrightarrow[\text{Graph}(S)]{} (\bar{x},\bar{y})$ be given by $\underset{(x,y)\xrightarrow[\text{Graph}(S)]{}(\bar{x},\bar{y})}{\text{g-LIMSUP}} DS(x \mid y)$. For example, the graphical limit $\underset{(x,y)\xrightarrow[\text{Graph}(S)]{}(\bar{x},\bar{y})}{\text{g-LIMSUP}} DS(x \mid y)$ in the notation of [RW98, Mor06] can be written as

$$\text{Graph}\Big(\underset{(x,y)\xrightarrow[\text{Graph}(S)]{}(\bar{x},\bar{y})}{\text{g-limsup}} DS(x \mid y)\Big) = \bigcup \Big\{\text{Graph}(G) : G \in \underset{(x,y)\xrightarrow[\text{Graph}(S)]{}(\bar{x},\bar{y})}{\text{g-LIMSUP}} DS(x \mid y)\Big\}.$$

A result compatible to the Aubin criterion can be stated as follows:

**Theorem 1.1.** [Pan11a] *(Extended Aubin criterion) For $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, and a prefan $H : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, $S$ is pseudo strictly $H$-differentiable if and only if*

$$G(p) \cap [-H(-p)] \neq \emptyset \text{ for all } p \neq 0 \text{ and } G \in \underset{(x,y)\xrightarrow[\text{Graph}(S)]{}(\bar{x},\bar{y})}{\text{g-LIMSUP}} DS(x \mid y).$$

*The result still holds if*

$$G \in \underset{(x,y)\xrightarrow[\text{Graph}(S)]{}(\bar{x},\bar{y})}{\text{g-LIMSUP}} D^{**}S(x \mid y)$$

*instead, where $D^{**}S(x \mid y) : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is defined via* $\text{Graph}(D^{**}S(x \mid y)) = \overline{\text{co}}\,\text{Graph}(DS(x \mid y))$.

The result in terms of coderivatives is even more appealing. For a positively homogeneous map $D:\mathbb{R}^m \rightrightarrows \mathbb{R}^n$, define $\mathscr{H}(D)$ by

$$\begin{aligned}\mathscr{H}(D) \quad := \quad & \{H:\mathbb{R}^n \rightrightarrows \mathbb{R}^m : H \text{ a prefan, and for all } p \neq 0 \text{ and } u \in \mathbb{R}^m, \\ & \min_{y\in H(p)} \langle u,y\rangle \leq \min_{v\in \overline{\text{co}}\, D(u)} \langle v,p\rangle\},\end{aligned}$$

**Theorem 1.2.** [Pan11a] *(Extended Mordukhovich Criterion) For $S:\mathbb{R}^n \rightrightarrows \mathbb{R}^m$, and a prefan $H:\mathbb{R}^n \rightrightarrows \mathbb{R}^m$, $S$ is pseudo strictly $H$-differentiable at $(\bar{x},\bar{y})$ if and only if $H \in \mathscr{H}(\overline{\text{co}}D^*S(\bar{x}\mid\bar{y}))$.*

Note that for any positively homogeneous map $D:\mathbb{R}^m \rightrightarrows \mathbb{R}^n$, we have $\mathscr{H}(\overline{\text{co}}D) = \mathscr{H}(D)$. Moreover, $\mathscr{H}$ has a strict reverse inclusion property, i.e.

$$\mathscr{H}(D_1) \begin{array}{c}\subset\\ \subsetneq\\ =\end{array} \mathscr{H}(D_2) \text{ if and only if } \overline{\text{co}}D_1 \begin{array}{c}\supset\\ \supsetneq\\ =\end{array} \overline{\text{co}}D_2.$$

The proof in [Pan11a] also highlights relations between the graphical derivative and the limiting coderivative, and unifies the Aubin and Mordukhovich criterions under a common hood. It also shows the limitations of generalizing the Mordukhovich criterion to the infinite dimensional case. The current method of proof does not seem to be easily generalized to Asplund spaces, where the Mordukhovich criterion also applies.

On the practical side, Theorem 1.2 tells us that the coderivatives are not only useful in the calculus of marginal functions in (3), they are also related to the pseudo strict differentiability of set-valued maps.

1.4. **Implicit multifunction theorems with positively homogeneous maps** [Pan11c]**.** The implicit function theorem for single-valued maps can be generalized for set-valued maps to address applications arising from generalized equilibria. For $G:X\times Y \rightrightarrows Z$, consider the set $I:X\rightrightarrows Y$ defined by

$$I(x) = \{y \mid 0 \in G(x,y)\}.$$

The condition $0\in G(x,y)$ can refer to an optimality condition for a nonlinear programming problem with parameter $x$ involving $f(x,\cdot)$ if the partial subdifferential $\partial_y f(x,\cdot)$ satisfies $\partial_y f(x,y) = G(x,y)$. More generally, for $G(x,y) = \partial_y \varphi(x,y) + N_{S(x)}(y)$, $0\in G(x,y)$ reflects the optimality condition of marginal function:

$$P(x):\ \min_{y\in S(x)} \varphi(x,y),$$

giving us sensitivity analysis of the solution set. The condition $0\in G(x,y)$ can also address some other form of equilibria, for example in variational inequalities and complementarity problems. Robinson [Rob80] was the first to study the Lipschitz continuity of the solution set $S$ if $S$ is a single-valued map, and the set-valued generalizations were pursued in [LZ99, DR09b], and summarized in the recent book [DR09a].

With the theory of generalized derivatives in Section 1.2, the implicit multifunction framework can be substantially extended. Using appropriate definitions of the partial generalized derivatives and partial linear openness (where linear openness is equivalent to generalized differentiation of the inverse set-valued map), we have the following result in [Pan11c]: Consider positively homogeneous maps $H_1:X\rightrightarrows Z$ and $H_2:Z\rightrightarrows Y$. If $0\in G(\bar{x},\bar{y})$, $G$ is partially pseudo strictly $H_1$-differentiable w.r.t. $x$ uniformly in $y$ at $(\bar{x},\bar{y},0)$, and $G$ is partially $H_2$-linearly open w.r.t. $y$ uniformly in $x$ at $(\bar{x},\bar{y},0)$, then under mild conditions, $I$ is partially pseudo strictly $(-H_2\circ H_1)$-differentiable at $(\bar{x},\bar{y})$. The classical implicit function theorem can be seen as a particular case of this implicit multifunction theorem.

## 2. Differential inclusions

For a set-valued map $F:[0,T]\times\mathbb{R}^n \rightrightarrows \mathbb{R}^n$, a function $x:[0,T]\to\mathbb{R}^n$ is said to satisfy a differential inclusion if

$$x'(t) \in F(t,x) \text{ a.e.}$$

In the case where $F$ is single-valued, we have the familiar differential equation. The differential inclusion framework is being increasingly promoted as a unifying framework for modern optimal control and the calculus of variations in the recent texts [Cla83, Mor06, Smi02, Vin00]. Other classical texts on differential inclusions include [AC84, AF90]. For $\varphi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ and $C \subset \mathbb{R}^n \times \mathbb{R}^n$, the optimization problem below is often studied;

$$\begin{array}{rl} \min_{x(\cdot)\in AC([0,1],\mathbb{R}^n)} & \varphi\big(x(0),x(T)\big) \\ \text{s.t.} & x'(t) \in F\big(t,x(t)\big) \text{ a.e.} \\ & \big(x(0),x(T)\big) \in C. \end{array} \tag{5}$$

Here, $AC([0,1],\mathbb{R}^n)$ stands for the set of absolutely continuous functions $x : [0,T] \to \mathbb{R}^n$. A huge part of research in optimal control and differential inclusions is on necessary conditions for optimality, starting from the classical Pontryagin Maximum Principle in the 60s. Recent work on necessary conditions in differential inclusions were obtained through the work of Clarke, Loewen, Rockafellar, Vinter, Mordukhovich, Kaskosz and Lojasiewicz, Milyutin, Smirnov, Zheng, Zhu and others. (See [Cla05] or [Mor06, Chapter 6] for a survey.) One type of optimality conditions are the (nonsmooth) Euler Lagrange condition and the transversality condition.

2.1. **Coderivatives of the reachable map (in progress).** The reachable map $R : X \rightrightarrows X$, which is of independent interest in optimal control, is defined by

$$\begin{array}{rcl} R(x) & = & \big\{y \mid \exists x(\cdot) \in AC([0,T],\mathbb{R}^n) \text{ s.t.} \\ & & x'(t) \in F\big(t,x(t)\big) \text{ a.e., } x(0) = x \text{ and } x(T) = y\big\}. \end{array}$$

We introduce conditions so that the convexified limiting coderivative of the reachable map satisfies

$$\begin{array}{rcl} \overline{\text{co}}D^*R(\bar{x} \mid \bar{y})(v) & \subset & \overline{\text{co}}\{u : \exists x(\cdot), p(\cdot) \in AC([0,T],\mathbb{R}^n) \text{ s.t.} \\ & & x'(t) \in F\big(t,x(t)\big), p'(t) \in -\overline{\text{co}}D^*_xF\big(t,x(t) \mid x'(t)\big)\big(p(t)\big), \\ & & \bar{x} = x(0), \bar{y} = x(T), u = p(0) \text{ and } v = p(T)\} \text{ for all } v \in \mathbb{R}^n. \end{array} \tag{6}$$

The proof first calculates the coderivatives of $R_N : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, the reachable map of the discretized differential inclusion problem, and then uses coderivative calculus and the material in Subsection 1.3 to approximate $\overline{\text{co}}D^*R(x \mid y)$ from $D^*R_N(x \mid y)$.

Using the map $\tilde{\varphi}(x,y) = \varphi(x,y) + \iota_C(x,y)$ in place of $\varphi(x,y)$ if necessary, where $\iota_C(\cdot)$ is the indicator function, we can reduce (5) to minimizing $v : \mathbb{R}^n \to \mathbb{R}$, where

$$\begin{array}{rcl} v(x) = & \min_y & \varphi(x,y) \\ & \text{s.t.} & y \in R(x) \end{array} \tag{7}$$

Note that (7) is in the framework of marginal functions in (1). We can deduce an upper bound of the Clarke subdifferential $\partial_C v(x) := \overline{\text{co}}\partial v(x)$. The optimality condition of $0 \in \partial v(x)$ combined with (3) implies the Euler Lagrange and transversality conditions.

A related question does not seem to have been studied: Given a path $x(\cdot) \in AC([0,T],\mathbb{R}^n)$, how does one perturb $x(\cdot)$ so that there is an improvement in the objective? We can obtain insight by studying the problem

$$\begin{array}{rl} \min\limits_{(x,y)} & \varphi(x,y) \\ \text{s.t.} & (x,y) \in \text{Graph}(R). \end{array} \tag{8}$$

The convexified limiting coderivative $\overline{\text{co}}D^*R(x \mid y)$ in turn gives some insight on the normal cone $N_{\text{Graph}(R)}(x,y)$ (which is related to $\text{Graph}(D^*R(x \mid y))$ by a linear transformation).

## 3. Semi-algebraic variational analysis

Semi-algebraic sets and functions (defined in terms of finitely many polynomials) are studied because they are a large class of objects used in practice, and are absent of much of the pathological behavior which, while of theoretical interest, do not appear often in practice. Examples of how simpler results can be obtained for the semi-algebraic case include the semismoothness of Lipschitz semi-algebraic functions [BDL09] and a Sard's theorem for nonsmooth, set-valued semi-algebraic functions [Iof08].

Examples of semi-algebraic functions are piecewise polynomials, rational functions, and the map from the set of matrices to its eigenvalues. Sharper results that may be useful in practice may be obtainable in the semi-algebraic setting. Recent references on semi-algebraic sets and functions, and the more general theory of tame topology and o-minimal structures, are [BR90, Cos99, Cos02, vdD98]. See for example [Iof09] and the lecture notes [Dan09] for a survey of the literature.

3.1. **Generic set-valued differentiability of semi-algebraic maps** [DP11]**.** In [DP11], I proved that if $X \subset \mathbb{R}^n$ is a semi-algebraic set, a semi-algebraic set-valued map $S : X \rightrightarrows \mathbb{R}^m$ is *set-valued differentiable* everywhere except on a subset of $X$ of smaller dimension. More precisely, there is a set $X' \subset X$ satisfying $\dim(X') < \dim(X)$ such that for any $\bar{x} \in X \backslash X'$ and $\bar{y} \in S(\bar{x})$, we can find a (non necessarily unique) linear map $L : \mathbb{R}^n \to \mathbb{R}^m$ such that for any $\delta > 0$, there are neighborhoods $U_\delta$ of $\bar{x}$ and $V_\delta$ of $\bar{y}$ such that

$$S(x) \cap V_\delta \subset S(x') + L(x - x') + \delta\|x - x'\|\mathbb{B} \text{ for all } x, x' \in U_\delta \cap X.$$

The example of $i_{\mathbb{Q}} : \mathbb{R} \to \mathbb{R}$, where $i_{\mathbb{Q}}(x)$ equals 1 if $x$ is rational and 0 otherwise, shows that this theorem cannot be true for maps in general. Applications of this continuity result include justifying the soundness of algorithms involving set-valued maps, and in the design of algorithms exploiting the presence of generic continuity.

3.2. **Lipschitz properties in semi-algebraic robust regularization** [LP09]**.** In [LP09], I proved that given a semi-algebraic function $f : \mathbb{R}^n \to \mathbb{R}$ and any point $x \in \mathbb{R}^n$, the robust regularization $f_\varepsilon : \mathbb{R}^n \to \mathbb{R}$ defined by $f_\varepsilon(x) := \max_{|x'-x| \leq \varepsilon} f(x')$ is locally Lipschitz at $x$ for all small $\varepsilon > 0$. This result shows that Lipschitz methods in optimization can almost always be implemented in robust optimization. An example of a non semi-algebraic function that does not have this property is the Cantor function.

## 4. Numerical methods for finding saddle points

For $f : X \to \mathbb{R}$, we say that $c \in X$ is a critical point if $\nabla f(c) = 0$. While critical points that are local maxima and minima can be found using optimization, saddle points (i.e., critical points that are not local extrema) are found using algorithms which trace their roots to the mountain pass theorem [AR73] and their variants. The mountain pass theorem states that under suitable conditions, given two local minima $a, b \in X$, there is a critical point $c \in X$ such that $f(c) > \max\{f(a), f(b)\}$, and the critical value $f(c)$ equals

$$\inf_{p \in \Gamma(a,b)} \sup_{t \in [0,1]} f \circ p(t), \tag{9}$$

where $\Gamma(a,b)$ is the set of paths $p : [0,1] \to X$ such that $p(0) = a$ and $p(1) = b$.

The problem of finding saddle points numerically is important in the problem of finding weak solutions to partial differential equations numerically. The original paper of a mountain pass algorithm to solve partial differential equations is [CM93], and it contains several semilinear elliptic problems. Particular applications in numerical partial differential equations include finding periodic solutions of a boundary value problem modeling a suspension bridge [Fen94] (introduced by [LM91]), studying a system of Ginzburg-Landau type equations arising in the thin film model of superconductivity [GM08], the choreographical 3-body problem [ABT06], and cylinder buckling [HLP06].

The problem of finding saddle points numerically is by now well entrenched in the chemistry curriculum. In transition state theory, the problem of finding the least amount of energy to transition between two stable states is equivalent to finding an optimal mountain pass between these two stable states. The highest point on the optimal mountain pass can then be used to determine the reaction kinetics. The foundations of transition state theory was laid by Marcelin, and important work by Eyring and Polanyi in 1931 and by Pelzer and Wigner a year later established the importance of saddle points in transition state theory. We cite the Wikipedia entry on transition state theory for more on its history and further references. Numerous methods for computing saddle points were suggested through the years, and we refer to the surveys [HJJ00, HS05, Sch11, Wal06] as well as the recent text [Wal03]. A software for computing saddle points in chemistry is Gaussian[1]. Tools for computing transition states[2] are also included in VASP[3]. Though the entire optimal mountain pass is needed for such an application, the process of computing saddle points often gives hints on an optimal mountain pass.

The rest of this section describes my work on numerical methods for finding saddle points. My first motivation for working on numerical methods for finding saddle points was to find the Wilkinson distance as explained in Subsection 5.2.

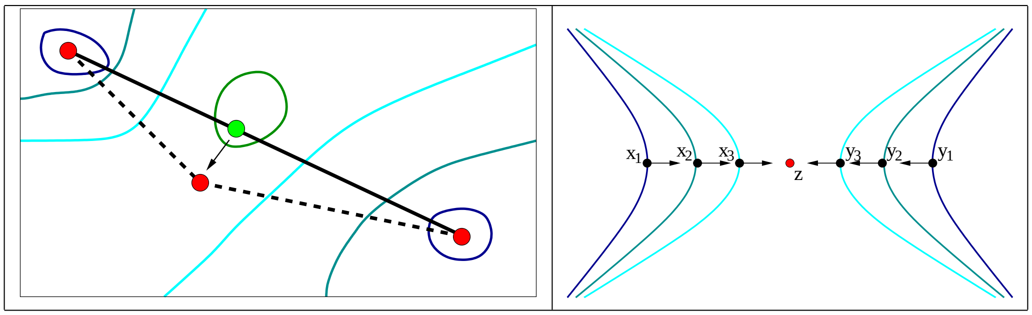


FIGURE 2. The diagram on the left shows the classical method of perturbing paths for the mountain pass problem, while the diagram on the right shows convergence to the critical point by looking at level sets.

4.1. **Level set methods for finding critical points of mountain pass type** [LP11]**.** While most current methods for finding saddle points discretize the paths in the set $\Gamma(a,b)$, I analyzed a different characterization of $f(c)$: The critical level $f(c)$ equals

$$\sup\left\{l \geq \max\left(f(a), f(b)\right) \mid a,b \text{ lie in different path components of } \{x \mid f(x) \leq l\}\right\}. \tag{10}$$

In other words, the critical level is the supremum of all levels $l$ such that there is no path connecting $a$ and $b$ in the set

$$\text{lev}_{\leq l} f := \{x \mid f(x) \leq l\}.$$

The difference between the methods are contrasted in Figure 2. This method is identical to the step and slide method in [MF01].

Among the advantages of this characterization are that we can find the closest points in different path components of $\text{lev}_{\leq l} f$ instead of keeping track of an entire path, as is currently done. The distance between the points in the path components indicate the progress of the algorithm. The key contribution in [LP11] was to prove local superlinear convergence to the critical point $\bar{x}$ when $f : \mathbb{R}^n \to \mathbb{R}$ has the quadratic approximation

$$f(x) = f(\bar{x}) + (x-\bar{x})^T A(x-\bar{x}) + o(|x-\bar{x}|^2). \tag{11}$$

4.2. **Global mountain pass algorithms using the parallel distance (in progress).** Building on earlier work, we identify two criteria for a global mountain pass algorithm to be effective:

(1) The global algorithm should achieve fast convergence once close to the saddle point.
(2) The global algorithm should find one saddle point.

The global mountain pass algorithms that rely on perturbing paths do not achieve fast local convergence, emphasizing property (2) at the expense of property (1). On the other hand, methods that emphasize property (1) and concentrate on building the quadratic model (11) do not satisfy property (2). In order to achieve property (1), one needs to make use of the quadratic approximation (11). If a global mountain pass algorithm aims to decrease the distance between components of $\text{lev}_{\leq l} f$ instead, one can concentrate computational efforts close to the saddle point, and the process of exploiting the quadratic approximation is more natural.

[1]http://www.gaussian.com/
[2]http://theory.cm.utexas.edu/vtsttools/neb/
[3]http://cms.mpi.univie.ac.at/vasp/vasp/vasp.html

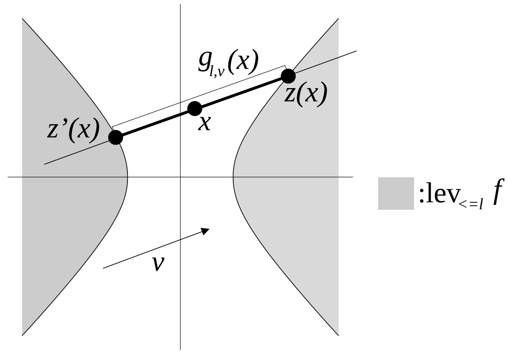


FIGURE 3. Illustration of the parallel distance.

In work currently in progress resulting from a UROP (undergraduate research) with Justin T Brereton in summer 2011, we studied the properties of the parallel distance: For a vector $v \neq 0$ and a level $l$, the parallel distance $g_{l,v}(x)$ is defined by

$$\begin{aligned} g_{l,v}(x) &:= \operatorname{diam}\big(S_{l,v}(x)\big). \\ S_{l,v}(x) &:= [\{x\} + \mathbb{R}\{v\}] \cap \operatorname{lev}_{\geq l} f. \end{aligned}$$

See Figure 3 for an illustration. In the case where $f$ is an exact quadratic, it turns out that $g_{l,v}(\cdot)^2$ is an exact quadratic. Much of the properties carry over to the smooth case once close enough to the saddle point as well. We are currently exploring how the parallel distance can be combined with other steps studied in [LP11] to create a robust global mountain pass algorithm that is also successful locally.

4.3. **Level set methods for finding critical points of general Morse index** [Pan]**.** The method for finding saddle points in the previous section is only sufficient for finding critical points of Morse index 1 (i.e., the maximal dimension of the subspace of $X$ on which the Hessian of $f : X \to \mathbb{R}$ is negative definite). For critical points $c$ of finite Morse index, the multidimensional mountain pass theorem shows that under nice conditions, critical levels $f(c)$ have a min-max characterization similar to that of (9). In [Pan], I showed that level set characterizations similar to that of (10) are possible, preserving some of the favorable properties of the algorithm in [LP11].

## 5. PSEUDOSPECTRA

The *pseudospectrum* $\Lambda_\varepsilon : \mathbb{C}^{n\times n} \rightrightarrows \mathbb{C}$ is defined by

$$\Lambda_\varepsilon(A) := \{z \mid z \text{ is an eigenvalue of } A+E \text{ for some } E \text{ such that } \|E\| \leq \varepsilon\}.$$

Pseudospectra gives better insight than eigenvalues in applications where the matrices and operators are nonnormal. For example, Lipschitz continuity of the pseudospectra (in the set-valued sense) is related to the transient behavior of the eigenvalues through the Kreiss matrix theorem. See [TE05] for more details.

5.1. **Lipschitz behavior of the pseudospectra** [LP08]**.** In [LP08], I proved conditions for the Lipschitz continuity of pseudospectra as a set-valued map. This work has led to my other work on variational analysis. Pseudospectra and eigenvalues provide a rich source of examples for semi-algebraic variational analysis.

5.2. **Numerical methods for the Wilkinson distance.** Another direction that I pursued from my research on pseudospectra was to study numerical methods for finding the *Wilkinson distance*: Given a matrix $A \in \mathbb{C}^{n\times n}$, the Wilkinson distance is the distance of $A$ to the closest matrix with repeated eigenvalues. The Wilkinson distance gives the region for which eigenvalues are differentiable in the entries of the matrix, and is important in stability issues of eigenvalues and eigenvectors. The Wilkinson problem can be traced back to Wilkinson [Wil65].

However, a simple guaranteed way to compute the Wilkinson distance is elusive [Dem90]. Recent work by Alam and Bora [AB05] reduced the Wilkinson distance problem to a global mountain pass problem through pseudospectra, which led to my work on computing saddle points. See also [ABBO11]. In addition to the local mountain pass problem addressed in [LP11], the global mountain pass problem still needs to be addressed to obtain an algorithm for the Wilkinson distance.

## REFERENCES


[AB05] R. Alam and S. Bora. On sensitivity of eigenvalues and eigendecompositions of matrices. *Linear Algebra Appl.*, 396:273–301, 2005.

[ABBO11] R. Alam, S. Bora, R. Byers, and M.L. Overton. Characterization and construction of the nearest defective matrix via coalescence of pseudospectral components. *Linear Algebra Appl.*, 435:3:494–513, 2011.

[ABT06] G. Arioli, V. Barutello, and S. Terracini. A new branch of mountain pass solutions for the choreographical 3-body problem. *Commun. Math. Phys.*, 268:439–463, 2006.

[AC84] J.-P. Aubin and A. Cellina. *Differential Inclusions: Set-Valued Maps and Viability Theory*. Springer, New York, 1984. Grundlehren der mathematischen Wissenschaften, Vol 264.

[AF90] J.-P. Aubin and H. Frankowska. *Set-Valued Analysis*. Birkhäuser, Boston, 1990. Republished as a Modern Birkhäuser Classic, 2009.

[AR73] A. Ambrosetti and P.H. Rabinowitz. Dual variational methods in critical point theory and applications. *J. Funct. Anal.*, 14:349–381, 1973.

[Aub84] J.-P. Aubin. Lipschitz behavior of solutions to convex minimization problems. *Math. Oper. Res.*, 1:225–245, 1984.

[BDL09] J. Bolte, A. Daniilidis, and A.S. Lewis. Tame functions are semismooth. *Math. Program., Ser. B*, 117:5–19, 2009.

[BR90] R. Benedetti and J.-J. Risler. *Real algebraic and semi-algebraic sets*. Hermann, Paris, 1990.

[BZ05] J.M. Borwein and Q.J. Zhu. *Techniques of Variational Analysis*. Springer, NY, 2005. CMS Books in Mathematics.

[Cla83] F.H. Clarke. *Optimization and Nonsmooth Analysis*. Wiley, Philadelphia, 1983. Republished as a SIAM Classic in Applied Mathematics, 1990.

[Cla05] F.H. Clarke. Necessary conditions in dynamic optimization. *Mem. Amer. Math. Soc.*, 173, no. 816, 2005.

[CLSW98] F.H. Clarke, Yu. S. Ledyaev, R.J. Stern, and P.R. Wolenski. *Nonsmooth Analysis and Control Theory*. Springer, NY, 1998.

[CM93] Y.S. Choi and P.J. McKenna. A mountain pass method for the numerical solution of semilinear elliptic equations. *Nonlinear Anal.*, 20:4:417–437, 1993.

[Cos99] M. Coste. *An Introduction to O-minimal Geometry*. Instituti Editoriali e poligrafici internazionali, Universita di Pisa, 1999. Available electronically at http://perso.univ-rennes1.fr/michel.coste/.

[Cos02] M. Coste. *An Introduction to Semialgebraic Geometry*. Instituti Editoriali e poligrafici internazionali, Universita di Pisa, 2002. Available electronically at http://perso.univ-rennes1.fr/michel.coste/.

[Dan09] A. Daniilidis. Gradient dynamical systems, tame optimization and applications. 2009. Lecture Notes, Spring School on Variational Analysis (Paseky nad Jizerou, Czech Republic, April 20-24, 2009) Available at http://pareto.uab.es/ adaniilidis/Articles.htm.

[Dem90] J.W. Demmel. Nearest defective matrices and the geometry of ill-conditioning. In *Reliable Numerical Computation*, pages 35–55. Oxford Univ. Press, UK, 1990.

[DP11] A. Daniilidis and C.H.J. Pang. Continuity and differentiability of set-valued maps revisited in the light of tame geometry. *J. London Math. Soc.*, 83(3):637–658, 2011.

[DR09a] A.L. Dontchev and R.T. Rockafellar. *Implicit Functions and Solution Mappings: A View from Variational Analysis*. Springer, New York, 2009. Springer Monographs in Mathematics.

[DR09b] A.L. Dontchev and R.T. Rockafellar. Robinson's implicit function theorem and its extensions. *Math. Program.*, 117:129–147, 2009.

[Fen94] Y. Feng. The study of nonlinear flexings in a floating beam by variational methods. oscillations in nonlinear systems: Applications and numerical aspects. *J. Comput. Appl. Math.*, 52:1–3:91–112, 1994.

[GM08] D. Glotov and P.J. McKenna. Numerical mountain pass solutions of Ginzburg-Landau type equations. *Comm. Pure Appl. Anal.*, 7:6:1345–1359, 2008.

[HJJ00] G. Henkelman, G. Jóhannesson, and H. Jónsson. Methods for finding saddle points and minimum energy paths. In S.D. Schwartz, editor, *Progress in Theoretical Chemistry and Physics*, pages 269–300. Kluwer, Amsterdam, 2000.

[HLP06] J. Horák, G.J. Lord, and M.A. Peletier. Cylinder buckling: The mountain pass as an organizing center. *SIAM J. Appl. Math.*, 66:5:1793–1824, 2006.

[HS05] H.P. Hratchian and H.B. Schlegel. Finding minima, transition states, and following reaction pathways on ab initio potential energy surfaces. In C.E. Dykstra, K.S. Kim, G. Frenking, and G.E. Scuseria, editors, *Theory and Applications of Computational Chemistry: The First 40 Years*, pages 195–259. Elsevier, The Netherlands, 2005.

[Iof08] A.D. Ioffe. Critical values of set-valued maps with stratifiable graphs: Extensions of Sard and Smale-Sard theorems. *Proc. Amer. Math. Soc.*, 136:3111–3119, 2008.

[Iof09] A.D. Ioffe. An invitation to tame optimization. *SIAM J. Optim.*, 19:1894–1917, 2009.

[LM91] A.C. Lazer and P.J. McKenna. Nonlinear flexings in a periodically forced floating beam. *Math. Method. Appl. Sci.*, 14:1:1–33, 1991.

[LP08] A.S. Lewis and C.H.J. Pang. Variational analysis of pseudospectra. *SIAM J. Optim.*, 19-3:1048–1072, 2008.

[LP09] A.S. Lewis and C.H.J. Pang. Lipschitz behavior of the robust regularization. *SIAM J. Control Optim.*, 8-45:3080–3104, 2009.

[LP11] A.S. Lewis and C.H.J. Pang. Level set methods for finding critical points of mountain pass type. *Nonlinear Anal.*, 74:12:4058–4082, 2011.

[LZ99] Yu. S. Ledyaev and Q.J. Zhu. Implicit multifunction theorems. *Set-Valued Anal.*, 7:209–238, 1999.

[MF01] R.A. Miron and K.A. Fichthorn. The step and slide method for finding saddle points on multidimensional potential surfaces. *J. Chem. Phys.*, 115:9:8742–8747, 2001.

[Mor06] B.S. Mordukhovich. *Variational Analysis and Generalized Differentiation I and II*. Springer, Berlin, 2006. Grundlehren der mathematischen Wissenschaften, Vols 330 and 331.

[Pan] C.H.J. Pang. Level set methods for finding saddle points of general Morse index.

[Pan11a] C.H.J. Pang. Characterizing generalized derivatives of set-valued maps: Extending the Aubin and Mordukhovich criterions. 2011. Available in http://arxiv.org/abs/1106.2338.

[Pan11b] C.H.J. Pang. Generalized differentiation with positively homogeneous maps: Applications in set-valued analysis and metric regularity. *Math. Oper. Res.*, 36:3:377–397, 2011.

[Pan11c] C.H.J. Pang. Implicit multifunction theorems with positively homogeneous maps. *Nonlinear Anal.*, 2011. Available at http://www.sciencedirect.com/science/article/pii/S0362546X1100277X.

[Rob80] S.M. Robinson. Strongly regular generalized equations. *Math. Oper. Res.*, 5:43–62, 1980.

[RW98] R.T. Rockafellar and R.J.-B. Wets. *Variational Analysis*. Springer, Berlin, 1998. Grundlehren der mathematischen Wissenschaften, Vol 317.

[Sch11] H.B. Schlegel. Geometry optimization. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 1:5:790–809, 2011.

[Smi02] G. Smirnov. *Introduction to the theory of differential inclusions*. Amer. Math. Soc., Providence, RI, 2002. Graduate Studies in Mathematics, Volume 41.

[TE05] L.N. Trefethen and M. Embree. *Spectra and Pseudospectra*. Princeton, NJ, 2005.

[vdD98] L. van den Dries. *Tame Topology and o-minimal Structures*. Cambridge, UK, 1998.

[Vin00] R.B. Vinter. *Optimal control*. Birkhäuser, Boston, 2000.

[Wal03] D.J. Wales. *Energy Landscapes: With applications to clusters, biomolecules and glasses*. Cambridge, Cambridge, UK, 2003.

[Wal06] D.J. Wales. Energy landscapes: calculating pathways and rates. *International Reviews in Physical Chemistry*, 25:1-2:237–282, 2006.

[Wil65] J.H. Wilkinson. *The Algebraic Eigenvalue Problem*. Oxford, UK, 1965.