Tame variational analysis

Dmitriy Drusvyatskiy
Mathematics, University of Washington

Joint work with
Daniilidis (Chile), Ioffe (Technion), and Lewis (Cornell)

May 19, 2015
Theme:

Semi-algebraic geometry is a powerful addition to the Variational Analysis toolkit.
For closed, convex $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, the following are equivalent:

**Quadratic growth:**

$$f(x) \geq f(\bar{x}) + \frac{\alpha}{2}|x - \bar{x}|^2 \quad \text{for } x \text{ near } \bar{x}.$$  

**Error bound:**

$$|x - \bar{x}| \leq \kappa \cdot \text{dist}(0, \partial f(x)) \quad \text{for } x \text{ near } \bar{x}.$$
For closed, convex $f: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, the following are equivalent

**Quadratic growth:**

$$f(x) \geq f(\bar{x}) + \frac{\alpha}{2} |x - \bar{x}|^2$$

for $x$ near $\bar{x}$.

**Error bound:**

$$|x - \bar{x}| \leq \kappa \cdot \text{dist}(0, \partial f(x))$$

for $x$ near $\bar{x}$.

(Proved by Kummer ’91, Aragon-Geoffroy ’08)
For closed, convex $f: \mathbb{R}^n \to \overline{\mathbb{R}}$, the following are equivalent:

**Quadratic growth:**

$$f(x) \geq f(\bar{x}) + \frac{\alpha}{2} |x - \bar{x}|^2 \quad \text{for } x \text{ near } \bar{x}.$$ 

**Error bound:**

$$|x - \bar{x}| \leq \kappa \cdot \text{dist}(0, \partial f(x)) \quad \text{for } x \text{ near } \bar{x}.$$ 

(Proved by Kummer ’91, Aragon-Geoffroy ’08)

Is this true generally?
For closed, convex $f: \mathbb{R}^n \to \overline{\mathbb{R}}$, the following are equivalent:

**Quadratic growth:**

$$f(x) \geq f(\bar{x}) + \frac{\alpha}{2} |x - \bar{x}|^2$$

for $x$ near $\bar{x}$.

**Error bound:**

$$|x - \bar{x}| \leq \kappa \cdot \text{dist}(0, \partial f(x))$$

for $x$ near $\bar{x}$.

(Proved by Kummer ’91, Aragon-Geoffroy ’08)

Is this true generally?

The implication $\uparrow$ always holds (D-Mordukhovich-Nghia ’14).
For closed, convex $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, the following are equivalent:

**Quadratic growth:**
\[
f(x) \geq f(\bar{x}) + \frac{\alpha}{2} |x - \bar{x}|^2 \quad \text{for } x \text{ near } \bar{x}.
\]

**Error bound:**
\[
|x - \bar{x}| \leq \kappa \cdot \text{dist}(0, \partial f(x)) \quad \text{for } x \text{ near } \bar{x}.
\]

(Proved by Kummer ’91, Aragon-Geoffroy ’08)

Is this true generally?

The implication $\uparrow$ always holds (D-Mordukhovich-Nghia ’14). But the converse $\downarrow$ can easily fail.
For closed, convex $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, the following are equivalent

**Quadratic growth:**

$$f(x) \geq f(\bar{x}) + \frac{\alpha}{2} |x - \bar{x}|^2 \quad \text{for } x \text{ near } \bar{x}.$$  

**Error bound:**

$$|x - \bar{x}| \leq \kappa \cdot \text{dist}(0, \partial f(x)) \quad \text{for } x \text{ near } \bar{x}.$$  

(Proved by Kummer '91, Aragon-Geoffroy '08)

Is this true generally?

The implication $\uparrow$ always holds (D-Mordukhovich-Nghia '14). But the converse $\downarrow$ can easily fail.

**Theorem (D-Ioffe)**

*The equivalence holds at local minimizers of semi-algebraic $f$.***
For closed, convex $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, the following are equivalent Quadratic growth:

$$f(x) \geq f(\bar{x}) + \frac{\alpha}{2}|x - \bar{x}|^2$$

for $x$ near $\bar{x}$.

Error bound:

$$|x - \bar{x}| \leq \kappa \cdot \text{dist}(0, \partial f(x))$$

for $x$ near $\bar{x}$.

(Proved by Kummer '91, Aragon-Geoffroy '08)

Is this true generally?

The implication $\uparrow$ always holds (D-Mordukhovich-Nghia ’14). But the converse $\downarrow$ can easily fail.

**Theorem (D-Ioffe)**

The equivalence holds at local minimizers of *semi-algebraic* $f$.

(More in Ioffe’s talk tomorrow.)
Outline

- Basics of semi-algebraic geometry

- Consequences for Variational Analysis:
  - Subgradient descent
  - Sweeping process
  - Sard theorem
  - Size of subdifferential graphs
  - Approximation on singular domains
Semi-algebraic geometry

Semi-algebraic set: finite union of sets

\[
\left\{ x : \begin{array}{l}
p_i(x) < 0 \text{ for } i \in I \\
p_j(x) = 0 \text{ for } j \in J
\end{array} \right\}
\]

where \( p_i, p_j \) are polynomials.
Semi-algebraic geometry

**Semi-algebraic set:** finite union of sets

\[
\left\{ x : p_i(x) < 0 \text{ for } i \in I \right. \\
\left. p_j(x) = 0 \text{ for } j \in J \right\}
\]

where \( p_i, p_j \) are polynomials.

A mapping \( F : \mathbb{R}^n \Rightarrow \mathbb{R}^m \) is **semi-algebraic** if

\[
gph F = \{(x, y) : y \in F(x)\}
\]

is semi-algebraic.
Semi-algebraic geometry

Semi-algebraic set: finite union of sets

\[ \left\{ x : \begin{array}{l} p_i(x) < 0 \text{ for } i \in I \\ p_j(x) = 0 \text{ for } j \in J \end{array} \right\} \]

where \( p_i, p_j \) are polynomials.

A mapping \( F : \mathbb{R}^n \rightarrow \mathbb{R}^m \) is semi-algebraic if

\[ \text{gph } F = \{(x, y) : y \in F(x)\} \] is semi-algebraic.

Robustness:

- **Boolean operations** preserve semi-algebraic sets.
- **Linear mappings** preserve semi-algebraic sets (Tarski-Seidenberg).
Semi-algebraic geometry

Semi-algebraic set: finite union of sets

\[
\left\{ x : \begin{array}{l} p_i(x) < 0 \text{ for } i \in I \\ p_j(x) = 0 \text{ for } j \in J \end{array} \right\}
\]

where \( p_i, p_j \) are polynomials.

A mapping \( F : \mathbb{R}^n \Rightarrow \mathbb{R}^m \) is semi-algebraic if

\[
\text{gph } F = \{(x, y) : y \in F(x)\}
\]
is semi-algebraic.

Robustness:

- Boolean operations preserve semi-algebraic sets.
- Linear mappings preserve semi-algebraic sets (Tarski-Seidenberg).

Eg: \( Q \) semi-algebraic \( \Rightarrow \{x : \exists y \ (x, y) \in Q\} \) semi-algebraic.
Semi-algebraic geometry

Semi-algebraic set: finite union of sets

\[
\left\{ x : \begin{array}{l} p_i(x) < 0 \text{ for } i \in I \\ p_j(x) = 0 \text{ for } j \in J \end{array} \right\}
\]

where \( p_i, p_j \) are polynomials.

A mapping \( F: \mathbb{R}^n \Rightarrow \mathbb{R}^m \) is semi-algebraic if

\[
gph F = \{(x, y) : y \in F(x)\} \text{ is semi-algebraic.}
\]

Robustness:

- Boolean operations preserve semi-algebraic sets.
- Linear mappings preserve semi-algebraic sets (Tarski-Seidenberg).

Eg: \( Q \) semi-algebraic \( \implies \) \( \{x : \exists y \ (x, y) \in Q\} \) semi-algebraic.

Conclusion: \( \partial f, |\nabla f|, \text{sur } F, \text{Lip } F, \ldots \) remain semi-algebraic.
Basic properties

**Curve selection:** Given $x \in \text{cl } Q$, there is an analytic curve $\gamma$ with $\gamma(0) = x$ and $\gamma(0, \eta) \subset Q$. (Bruhat-Cartan ’50, Milnor ’68)
Basic properties

Curve selection: Given $x \in \text{cl } Q$, there is an analytic curve $\gamma$ with $\gamma(0) = x$ and $\gamma(0, \eta) \subset Q$. (Bruhat-Cartan ’50, Milnor ’68)

Semi-algebraic selection: Semialgebraic $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ admit semi-algebraic selections $f \subset F$. 
Basic properties

**Curve selection:** Given $x \in \text{cl } Q$, there is an analytic curve $\gamma$ with $\gamma(0) = x$ and $\gamma(0, \eta) \subset Q$. (Bruhat-Cartan ’50, Milnor ’68)

**Semi-algebraic selection:** Semialgebraic $F: \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ admit semi-algebraic selections $f \subset F$.

**Stratification:** Semialgebraic sets “stratify” into finitely many manifolds $\{M_i\}$, and so have dimension. (Whitney ’65, Łojasiewicz ’71)
Basic properties

**Curve selection:** Given \( x \in \text{cl } Q \), there is an analytic curve \( \gamma \) with \( \gamma(0) = x \) and \( \gamma(0, \eta) \subset Q \). (Bruhat-Cartan ’50, Milnor ’68)

![Diagram of curve selection](image)

**Semi-algebraic selection:** Semialgebraic \( F : \mathbb{R}^n \Rightarrow \mathbb{R}^m \) admit semialgebraic selections \( f \subset F \).

**Stratification:** Semialgebraic sets “stratify” into finitely many manifolds \( \{M_i\} \), and so have dimension. (Whitney ’65, Łojasiewicz ’71)

![Diagram of stratification](image)

**Łojasiewicz inequality:** If \( f \) is semi-algebraic, then on compacta

\[
\text{dist}(x; f^{-1}(0)) \leq C|f(x)|^\alpha.
\]

(Łojasiewicz ’91, Kurdyka ’98, Bolte-Daniilidis-Lewis ’06)
What are consequences for Variational Analysis?
Theorem (D-Ioffe-Lewis)

\( f \) semi-algebraic, \( \bar{x} \) not a local minimizer \( \implies \) there exists a nontrivial solution to

\[
\dot{x} \in -\partial f(x) \quad \text{and} \quad x(0) = \bar{x}.
\]
Theorem (D-Ioffe-Lewis)

\( f \) semi-algebraic, \( \bar{x} \) not a local minimizer \( \implies \) there exists a nontrivial solution to

\[
\dot{x} \in -\partial f(x) \quad \text{and} \quad x(0) = \bar{x}.
\]

Moreover, then

\( \dot{x} \) is the shortest element of \( -\partial f(x) \) a.e.
Theorem (D-Ioffe-Lewis)

If $f$ semi-algebraic, $\bar{x}$ not a local minimizer $\implies$ there exists a nontrivial solution to

\[ \dot{x} \in -\partial f(x) \quad \text{and} \quad x(0) = \bar{x}. \]

Moreover, then

\[ \dot{x} \text{ is the shortest element of } -\partial f(x) \quad \text{a.e.} \]

Theorem (Daniilidis-Bolte-Lewis)

If $f$ is semi-algebraic, then any bounded subgradient curve has finite length.
Subgradient systems

Theorem (D-Ioffe-Lewis)

\( f \) semi-algebraic, \( \bar{x} \) not a local minimizer \( \implies \) there exists a nontrivial solution to

\[
\dot{x} \in -\partial f(x) \quad \text{and} \quad x(0) = \bar{x}.
\]

Moreover, then

\( \dot{x} \) is the shortest element of \( -\partial f(x) \) a.e.

Theorem (Daniilidis-Bolte-Lewis)

If \( f \) is semi-algebraic, then any bounded subgradient curve has finite length.

Many analogues for descent methods; e.g. proximal point, splitting, Gauss-Seidel, etc (Attouch, Bolte, Bot, Noll, Peypouquet, Soubeyran, Svaiter, ...).
Sweeping process

Sweeping process (Moreau ’77):

\[ \dot{x}(t) \in -NS(t)(x(t)) \]

with \( S(t) \) a moving set.
Sweeping process (Moreau ’77):

\[ \dot{x}(t) \in -N_{S(t)}(x(t)) \]

with \( S(t) \) a moving set.

- The monotone case \( S(t) = [f \leq t] \) is subgradient descent.

\[ S(t) \]

\[ -N_{S(t)}(x(t)) \]
Sweeping process

Sweeping process (Moreau ’77):

\[ \dot{x}(t) \in -N_{S(t)}(x(t)) \]

with \( S(t) \) a moving set.

- The monotone case \( S(t) = [f \leq t] \) is
  subgradient descent.

Theorem (D-Daniilidis)

If \( S: \mathbb{R} \rightarrow \mathbb{R}^n \) is semi-algebraic, then every bounded solution of the sweeping process has finite length.
Sweeping process

Sweeping process (Moreau ’77):

\[ \dot{x}(t) \in -N_{S(t)}(x(t)) \]

with \( S(t) \) a moving set.

- The monotone case \( S(t) = [f \leq t] \) is subgradient descent.

**Theorem (D-Daniilidis)**

*If \( S: \mathbb{R} \rightarrow \mathbb{R}^n \) is semi-algebraic, then every bounded solution of the sweeping process has finite length.*

**Key estimate:**

\[ |\dot{x}(t)| \leq \text{Lip } S(t|x(t)) \leq \sup_{x \in S(t) \cap X} \text{Lip } S(t|x) \]

and the upper-bound is integrable by the Łojasiewicz inequality.
Sard Theorem

Mapping $F: \mathbb{R}^n \Rightarrow \mathbb{R}^m$ is **metrically regular** at $(\bar{x}, \bar{y}) \in \text{gph} F$ if

$$\frac{\text{dist}(x, F^{-1}(y))}{\text{dist}(y, F(x))}$$

is bounded near $(\bar{x}, \bar{y})$.

If $F$ is **not** metrically regular at $(\bar{x}, \bar{y})$, then $\bar{y}$ is a **critical value**.
Sard Theorem

Mapping $F: \mathbb{R}^n \Rightarrow \mathbb{R}^m$ is **metrically regular** at $(\bar{x}, \bar{y}) \in \text{gph } F$ if

$$\frac{\text{dist}(x, F^{-1}(y))}{\text{dist}(y, F(x))} \text{ is bounded near } (\bar{x}, \bar{y}).$$

If $F$ is **not** metrically regular at $(\bar{x}, \bar{y})$, then $\bar{y}$ is a **critical value**.

Theorem (Ioffe ’08)

*Semi-algebraic* $F: \mathbb{R}^n \Rightarrow \mathbb{R}^m$ have **almost no** critical values.

- Justifies typical linear convergence of basic schemes, e.g. alternating projections (D-Ioffe-Lewis ’13).
Sard Theorem

Mapping $F: \mathbb{R}^n \Rightarrow \mathbb{R}^m$ is metrically regular at $(\bar{x}, \bar{y}) \in \text{gph} \ F$ if

$$\frac{\text{dist}(x, F^{-1}(y))}{\text{dist}(y, F(x))}$$

is bounded near $(\bar{x}, \bar{y})$.

If $F$ is not metrically regular at $(\bar{x}, \bar{y})$, then $\bar{y}$ is a critical value.

Theorem (Ioffe ’08)

Semi-algebraic $F: \mathbb{R}^n \Rightarrow \mathbb{R}^m$ have almost no critical values.

• Justifies typical linear convergence of basic schemes, e.g. alternating projections (D-Ioffe-Lewis ’13).

Sard Theorem & “gph $\partial f$ is thin”

$\implies$ generic properties of semi-algebraic functions.

(cf. Lewis’ talk)
Consider

$$\min_x f(x) + h(G(x) + y) - \langle v, x \rangle$$

where $f$, $h$, $G$ are semi-algebraic and $G$ is $C^2$-smooth.
Consider

$$\min_x f(x) + h(G(x) + y) - \langle v, x \rangle$$

where $f$, $h$, $G$ are semi-algebraic and $G$ is $C^2$-smooth.

**Optimality conditions:**

$$\begin{bmatrix} v \\ y \end{bmatrix} \in \begin{bmatrix} \nabla G(x)^* \lambda \\ -G(x) \end{bmatrix} + \left( \partial f \times (\partial h)^{-1} \right)(x, y).$$
Consider
\[
\min_x f(x) + h(G(x) + y) - \langle v, x \rangle
\]
where \( f, h, G \) are semi-algebraic and \( G \) is \( C^2 \)-smooth.

**Optimality conditions:**
\[
\begin{bmatrix} v \\ y \end{bmatrix} \in \begin{bmatrix} \nabla G(x)^* \lambda \\ -G(x) \end{bmatrix} + \left( \partial f \times (\partial h)^{-1} \right) (x, y).
\]

Sard theorem & thinness \( \implies \) generic properties:

- qualification conditions, strict complementarity, smooth dependance of \((x, \lambda)\), existence of identifiable manifolds, second order sufficient conditions are necessary at local minimizers.
Generic properties

Consider

\[
\min_{x} f(x) + h(G(x) + y) - \langle v, x \rangle
\]

where \( f, h, G \) are semi-algebraic and \( G \) is \( C^2 \)-smooth.

Optimality conditions:

\[
\begin{bmatrix} v \\ y \end{bmatrix} \in \begin{bmatrix} \nabla G(x)^* \lambda \\ -G(x) \end{bmatrix} + \left( \partial f \times (\partial h)^{-1} \right)(x, y).
\]

Sard theorem & thinness \( \implies \) generic properties:

- qualification conditions, strict complementarity, smooth dependance of \((x, \lambda)\), existence of identifiable manifolds, second order sufficient conditions are necessary at local minimizers.

Remark: Without semi-algebraicity, one needs geometric measure theory and not all properties above are generic.
Approximation of functions

Set-up: \[ Q \xrightarrow{f} \mathbb{R}^n \xrightarrow{\rightarrow} \mathbb{R}. \]

Assume \( Q \) is a disjoint union of manifolds

\[
Q = \mathcal{M}_1 \cup \mathcal{M}_2 \cup \ldots \cup \mathcal{M}_{k-1} \cup \mathcal{M}_k
\]
Approximation of functions

**Motivation:** Integration by parts

\[ \int_Q g \Delta f \, dx \]
Approximation of functions

**Motivation:** Integration by parts

\[
\int_Q g \Delta f \, dx = \int_{\partial Q} g \langle \nabla f, \hat{n} \rangle \, dS - \int_Q \langle \nabla g, \nabla f \rangle \, dx.
\]
Approximation of functions

Motivation: Integration by parts

$$\int_Q g\Delta f \, dx = \int_{\partial Q} g\langle \nabla f, \hat{n} \rangle \, dS - \int_Q \langle \nabla g, \nabla f \rangle \, dx.$$

Goal: approximate $f$ by a $C^2$–smooth $\tilde{f}$ so that the $\nabla \tilde{f} \perp \hat{n}$. 
Motivation: Integration by parts

\[ \int_Q g \Delta f \, dx = \int_{\partial Q} g \langle \nabla f, \hat{n} \rangle \, dS - \int_Q \langle \nabla g, \nabla f \rangle \, dx. \]

Goal: approximate \( f \) by a \( C^2 \)-smooth \( \tilde{f} \) so that the \( \nabla \tilde{f} \perp \hat{n} \).

Theorem (D-Larsson)

Given a continuous \( f : \mathbb{R}^n \to \mathbb{R} \) and any \( \epsilon > 0 \), there exists a \( C^1 \)-smooth \( \tilde{f} \) satisfying

1. **Closeness:** \( |\tilde{f}(x) - f(x)| < \epsilon \) for all \( x \in \mathbb{R}^n \),
2. **Neumann Boundary condition:**

\[ x \in M_i \implies \nabla \tilde{f}(x) \in T_x M_i. \]
Approximation of functions

Motivation: Integration by parts

\[ \int_Q g \Delta f \, dx = \int_{\partial Q} g \langle \nabla f, \hat{n} \rangle \, dS - \int_Q \langle \nabla g, \nabla f \rangle \, dx. \]

Goal: approximate \( f \) by a \( C^2 \)-smooth \( \tilde{f} \) so that the \( \nabla \tilde{f} \perp \hat{n} \).

Theorem (D-Larsson)

Given a continuous \( f : \mathbb{R}^n \to \mathbb{R} \) and any \( \epsilon > 0 \), there exists a \( C^1 \)-smooth \( \tilde{f} \) satisfying

1. Closeness: \(|\tilde{f}(x) - f(x)| < \epsilon \) for all \( x \in \mathbb{R}^n \),
2. Neumann Boundary condition:

\[ x \in \mathcal{M}_i \implies \nabla \tilde{f}(x) \in T_x \mathcal{M}_i. \]

provided \( \{M_i\} \) is a Whitney stratification of \( Q \).
Approximation of functions

Motivation: Integration by parts

\[ \int_Q g \Delta f \, dx = \int_{\partial Q} g \langle \nabla f, \hat{n} \rangle \, dS - \int_Q \langle \nabla g, \nabla f \rangle \, dx. \]

Goal: approximate \( f \) by a \( C^2 \)-smooth \( \tilde{f} \) so that the \( \nabla \tilde{f} \perp \hat{n} \).

Theorem (D-Larsson)

Given a continuous \( f : \mathbb{R}^n \to \mathbb{R} \) and any \( \epsilon > 0 \), there exists a \( C^1 \)-smooth \( \tilde{f} \) satisfying

1. **Closeness:** \( |\tilde{f}(x) - f(x)| < \epsilon \) for all \( x \in \mathbb{R}^n \),
2. **Neumann Boundary condition:**

\[ x \in M_i \implies \nabla \tilde{f}(x) \in T_x M_i. \]

provided \( \{M_i\} \) is a Whitney stratification of \( Q \).

- For semi-algebraic \( Q \), Whitney stratifications always exist!
Conclusion

- Semi-algebraic geometry is a powerful addition to the Variational Analysis toolkit.

- **Applications:** quadratic growth and error bounds, subgradient descent and the sweeping process, Sard theorem, and approximation on singular domains.
Thank you.
References


Available at www.math.washington.edu/~ddrusv