# AN OVERVIEW OF VARIATIONAL ANALYSIS

# 2. VARIATIONAL GEOMETRY

**Terry Rockafellar**
University of Washington, Seattle

Forum on Developments and Origins of Operations Research

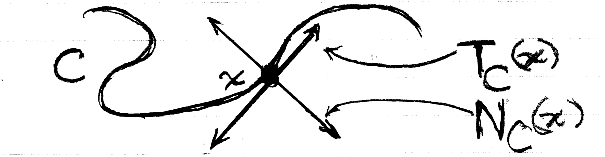
Organized by the *Mathematical Programming Branch of the Operations Research Society of China* and the *Southern University of Science and Technology* (SUSTech)

Shenzhen, China, 6–10 December 2021 (virtual)

# Origins in Classical Differential Geometry

**Motivational framework:** sets specified by nice equation systems

$C$ = some "smooth manifold" $\subset \boldsymbol{R}^n$ curve, hypersurface ...




**Tangent subspace:** $T_C(x)$ furnishes the linearization of $C$ at $x$

$w \in T_C(x) \iff \exists \xi : (-\varepsilon, \varepsilon) \to C$ with $\xi(0) = x$, $\xi'(0) = w$

i.e., $w$ is a tangent vector at $x$ to a curve within $C$

**Normal subspace:** $N_C(x)$ = orthogonal complement of $T_C(x)$

$v \in N_C(x) \iff v \cdot w = 0$ for all $w \in T_C(x)$

i.e., $v$ is $\perp$ (orthogonal) to every tangent vector $w$ at $x$

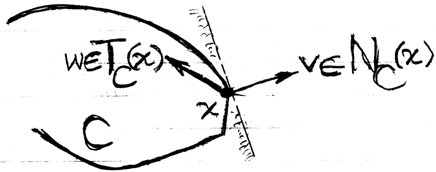
(these subspaces move "smoothly" as $x$ moves in $C$)

# Later Ideas in Convex Geometry

**Motivational framework:** sets given by linear inequality systems

$C$ = some closed convex set $\subset \boldsymbol{R}^n$ maybe a polyhedron ...

a one-sided approach to tangents and normals becomes essential




**Tangent vectors:** the elements of $T_C(x)$ as "feasible variations"

- consider first all $w$ at $x$ such that $x + \tau w \in C$ for $\tau \in [0, \varepsilon)$
- then pass to the vectors $w$ in the <u>closure</u> of that set

**Normal vectors:** the elements of $N_C(x)$

- all vectors $v$ such that $v \cdot x' \leq v \cdot x$ for all $x' \in C$

$$v \cdot (x' - x) \leq 0$$

# Polarity Replaces Orthogonality in Convex Geometry

**Cones:** a set $K \subset R^n$ is a cone $\iff$
$0 \in K$ and $\tau w \in K$ for all $\tau > 0$ when $w \in K$

**Polarity:** the polar of $K$ is $K^* = \{ v \mid v \cdot w \leq 0, \forall w \in K \}$

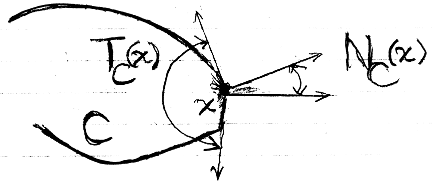
$K^*$ is a closed convex cone, $(K^*)^* =$ closed convex hull of $K$

**Orthogonality as a special case:** $K^* = K^\perp$ if $K$ is a subspace

**Tangent-normal polarity for convex sets $C$**

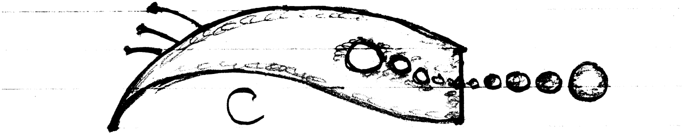
$T_C(x)$ and $N_C(x)$ are closed convex cones polar to each other

$N_C(x) = T_C(x)^*, \qquad T_C(x) = N_C(x)^*$



# Aiming Beyond Classical and Convex Analysis

**Challenges:** can't define tangent vectors using line segments
can't define normal vectors by linear inequalities




**Approach to "variations" in the early days of optimization:**

obtain vectors $w$ from "curves" $\xi$ that enter $C$ at $x$,

$\exists \xi : [0, \varepsilon) \to C$ such that $\xi(0) = x$, $\xi'(0) = w$

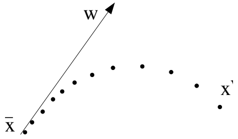
the set of such $w$ = the cone of "**derivable**" tangent vectors

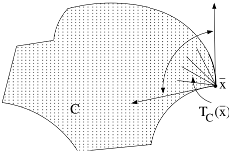
**Status:** this concept produces the right tangent cones for many kinds of $C$ (convex, smooth manifold, NLP feasible set)
but it has serious limitations in a larger framework of theory

# Tangent Vectors: "General" and "Regular"

**General tangent cone:** $T_C(x)$ from "directional" convergence

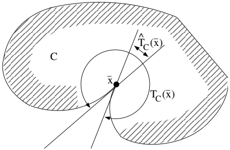
$$w \in T_C(x) \iff \exists\, \tau^\nu \searrow 0,\ x^\nu \underset{C}{\rightarrow} x, \text{ such that } \tfrac{1}{\tau^\nu}[x^\nu - x] \rightarrow w$$




**Regular tangent cone:** $\widehat{T}_C(x)$ demanding local "stability"

$$w \in \widehat{T}_C(x) \iff \forall x^\nu \underset{C}{\rightarrow} x,\ \exists\, w^\nu \in T_C(x^\nu) \text{ with } w^\nu \rightarrow w$$




**Equivalently:** "$\widehat{T}_C(x) = \liminf_{x' \underset{C}{\rightarrow} x} T_C(x')$" "inner" set limit

## Normal Vectors: "General" and "Regular"

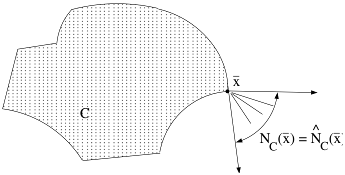
**Regular normal cone:** $\widehat{N}_C(x)$ from local inequalities

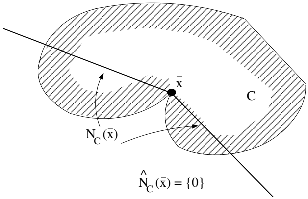
$$v \in \widehat{N}_C(x) \iff v{\cdot}x' \le v{\cdot}x + o(\,|x'-x|\,) \text{ for } x' \in C$$

classical "o" notation: $\frac{o(t)}{t} \to 0$ as $t \searrow 0$, $\quad |\cdot| =$ euclidean norm

**General normal cone:** $N_C(x)$ absorbing nearby information

$$v \in N_C(x) \iff \exists\, x^\nu \underset{C}{\to} x,\ v^\nu \in \widehat{N}_C(x^\nu), \text{ such that } v^\nu \to v$$




**Equivalently:** $$\text{“}N_C(x) = \limsup_{x' \underset{C}{\to} x} \widehat{N}_C(x')\text{”}$$ "outer" set limit

**important history to explain: Clarke, Mordukhovich,. . .**

# Fundamental Tangent-Normal Relations

**Cone notation:** $T_C$ general, $\widehat{T}_C$ regular, $N_C$ general, $\widehat{N}_C$ regular

**Central theorem of first-order variational analysis**

- $\widehat{T}_C(x) \subset T_C(x)$ with $\widehat{T}_C(x)$ convex, both cones closed
- $\widehat{N}_C(x) \subset N_C(x)$ with $\widehat{N}_C(x)$ convex, both cones closed

$$\widehat{N}_C(x) = T_C(x)^* \text{ (polar)}, \quad \widehat{T}_C(x) = N_C(x)^* \text{ (polar)}$$

$$\widehat{T}_C(x) = T_C(x) \iff \widehat{N}_C(x) = N_C(x)$$

**Variational regularity:** the case where **regular = general**

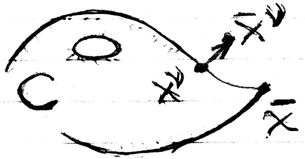
- holds for $C$ convex or smooth manifold or nice NLP feasible set
- a basic property of sets unknown before variational analysis!

**Pioneers:** Clarke on tangent side, Mordukhovich on normal side

"Clarke regularity" $\iff$ "Mordukhovich regularity"

# Implication for Boundaries and Interiors

boundary points $\bar{x}$ of $C$ are signaled by $\exists$ nonzero $v \in N_C(\bar{x})$




$$\bar{x}^\nu \to \bar{x} \in C, \quad \bar{x}^\nu \notin C$$
$$x^\nu = \mathrm{proj}_C\, \bar{x}^\nu, \quad x^\nu \to \bar{x}$$
$$v^\nu = \frac{\bar{x}^\nu - x^\nu}{|\bar{x}^\nu - x^\nu|} \in N_C(x^\nu), \ |v^\nu| = 1$$

**Corresponcing characterization of set interiors**

$\bar{x} \in \mathrm{int}\, C \subset \boldsymbol{R}^n \iff \widehat{T}_C(\bar{x}) =$ all of $\boldsymbol{R}^n$

$\iff$ the only normal $v \in N_C(\bar{x})$ is $v = 0$

# Regular Normals From an Optimization Perspective

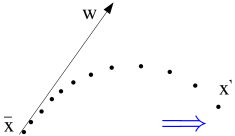
**Consider:** maximizing a $\mathcal{C}^1$ function $h$ over a closed set $C \subset \boldsymbol{R}^n$

local max at $\bar{x}$?

suppose $x^\nu \to \bar{x}$ in $C$ and

$\exists\, \tau^\nu \searrow 0,\ \frac{1}{\tau^\nu}[x^\nu - \bar{x}] \to w$

i.e., $w \in T_C(\bar{x})$

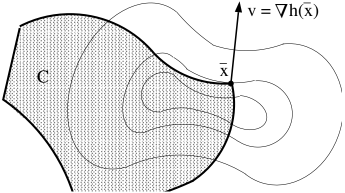



$\Longrightarrow\ \nabla h(\bar{x}) \cdot w \leq 0$

$\Longrightarrow\ \nabla h(\bar{x})$ belongs to the polar cone $T_C(\bar{x})^* = \widehat{N}_C(\bar{x})$

## Local max characterization of regular normals

$v \in \widehat{N}_C(\bar{x}) \iff \exists h$ with local max on $C$ at $\bar{x}$ and $\nabla h(\bar{x}) = v$



# First-Order Optimality and Its Calculus

**Basic necessary condition in minimization**

If a $\mathcal{C}^1$ function $f_0$ has a local minimum relative at $\bar{x}$ relative to a closed set $C$, then $-\nabla f_0(\bar{x}) \in \widehat{N}_C(\bar{x})$, hence $-\nabla f_0(\bar{x}) \in N_C(\bar{x})$

**How useful?** there must be a good "calculus" of normal cones

*general normals work better for this than regular normals*

**Example: normal cones to an intersection of closed sets**

Let $C = C_1 \cap \cdots \cap C_s$ and let $\bar{x} \in C$ satisfy

$$v_i \in N_{C_i}(\bar{x}),\ v_1 + \cdots + v_s = 0 \quad \Longrightarrow \quad v_1 = \cdots = v_s = 0$$

(constraint qualification). Then

$$N_C(\bar{x}) \subset N_{C_1}(\bar{x}) + \cdots + N_{C_s}(\bar{x}) = \left\{ v_1 + \cdots v_s \,\middle|\, v_i \in N_{C_i}(\bar{x}) \right\}$$

Moreover, if every $C_k$ is variationally **regular** at $\bar{x}$, then $C$ is variationally **regular** at $\bar{x}$, and the inclusion holds as an equation

# Connection With Lagrange Multipliers

**Constraints:** $C = \{x \in \mathbb{R}^n \,|\, (f_1(x), \ldots, f_m(x)) \in D\}$, $f_i \in \mathcal{C}^1$

**NLP case:** $D = (-\infty, 0] \times \cdots \times (-\infty, 0] \times \{0\} \times \cdots \times \{0\}$

**Multiplier vectors:** $y = (y_1, \ldots, y_m) \in \mathbb{R}^m$

**Constraint qualification (CQ):** at a point $\bar{x} \in C$

only $y = (0, \ldots, 0)$ satisfies $\begin{cases} y \in N_D(f_1(\bar{x}), \ldots, f_m(\bar{x})) \text{ with} \\ y_1 \nabla f_1(\bar{x}) + \cdots + y_m \nabla f_m(\bar{x}) = 0 \end{cases}$

### Normal cone formula with multipliers, under (CQ)

Let $D$ be variationally **regular** at the point $(f_1(\bar{x}), \ldots, f_m(\bar{x}))$ **(e.g., convex)**. Then $C$ is variationally **regular** at $\bar{x}$ and

$$v \in N_C(\bar{x}) \iff \begin{cases} \exists\, y \in N_D(f_1(\bar{x}), \ldots, f_m(\bar{x})) \text{ with} \\ y_1 \nabla f_1(\bar{x}) + \cdots + y_m \nabla f_m(\bar{x}) = v \end{cases}$$

# Prospects

**Key feature of variational geometry:** **great generality**

Although inspired by the one-sided effects of constraints it covers **all closed sets, regardless of any special structure**

---

**Deep applicability:** **from "feasible sets" to much more**

Tangent cones and normal cones can be employed in the study of closed sets that are the **epigraphs of extended-real-valued functions**, or on the other hand **graphs of set-valued mappings**

$\longrightarrow$ **generalized differentiation of such functions/mappings**

---

**Example:** **second-order variational geometry**

Question: how can concepts of **curvature** of $C$ be articulated very generally? Answer: by means of tangent and normals to the graph of the **normal cone mapping** $N_C : x(\in C) \mapsto N_C(x)$

# Further Study

- Basics about convexity, including topics like polar cones, are available in the book *Convex Analysis* (1970). But although many of the novel ideas of variational geometry came first from the study of convex sets, they are somewhat hidden in that book, because it focuses more on convex functions.

  The polarity between $N_C(x)$ and $T_C(x)$ is embedded in a result about subgradients and directional derivatives. Indeed, the notation $N_C(x)$ and $T_C(x)$ came only later, and the term "tangent cone" isn't even included in the index of *Convex Analysis*.

- The main place to learn more about the variational geometry in this lecture is Chapter 6 of the book *Variational Analysis* (1998). Chapter 2 there offers support about convexity.

  It's not necessary to study Chapters 4 and 5 deeply before getting into Chapter 6. To the extent that ideas in those chapters are needed, they will be introduced in Lecture 4 of this overview.