Abstract. Continuous-time projected dynamical systems are an elementary class of discontinuous dynamical systems with trajectories that remain in a feasible domain by means of projecting outward-pointing vector fields. They are essential when modeling physical saturation in control systems, constraints of motion, as well as studying projection-based numerical optimization algorithms. Motivated by the emerging application of feedback-based continuous-time optimization schemes that rely on the physical system to enforce nonlinear hard constraints, we study the fundamental properties of these dynamics on general locally-Euclidean sets. Among others, we propose the use of Krasovskii solutions, show their existence on nonconvex, irregular subsets of low-regularity Riemannian manifolds, and investigate how they relate to conventional Carathéodory solutions. Furthermore, we establish conditions for uniqueness, thereby introducing a generalized definition of prox-regularity which is suitable for non-flat domains. Finally, we use these results to study the stability and convergence of projected gradient flows as an illustrative application of our framework. We provide simple counter-examples for our main results to illustrate the necessity of our already weak assumptions.

1. Introduction. Projected dynamical systems form an important class of discontinuous dynamical system whose trajectories remain in a domain $X$ by projecting outward portions of a vector field $f$ at the boundary of $X$ to prevent a trajectory from leaving the domain. This qualitative behavior is illustrated in Fig. 1a.

Even though projected dynamical systems have a long history in different contexts such as the study of variational inequalities or differential inclusions, new compelling applications in the context of real-time optimization require a holistic study in a more general setting. Hence, this paper is primarily motivated by the renewed interest in dynamical systems that solve optimization problems. Early works in this spirit such as [9] have designed continuous-time systems to solve computational problems such as diagonalizing matrices or solving linear programs. This has further resulted in the study of optimization algorithms over manifolds [2]. Recently, interest has shifted towards analyzing existing iterative schemes with tools from dynamical systems including Lyapunov theory [52] and integral quadratic constraints [20,35]. Most of these have considered unconstrained optimization problems [49] and algorithms that can be modelled with a standard ODE [33] or with variational tools [51]. With this paper we hope to pave the way for the analysis of algorithms for constrained optimization whose continuous-time limits are discontinuous.

Recently, this idea of studying the dynamical aspects of optimization algorithms has given rise to a new type of feedback control design that aims at steering a physical system in real time to the solution of an optimization problem [16,34,40,43,54] without external inputs. Precursors of this idea have been used in the analysis of congestion control in communication networks [32, 38]. More recently, the concept has been widely applied to power systems [19, 22, 27, 36, 41, 50]. This context is particularly challenging, because the physical laws of power flow, saturating components, and other constraints define a highly non-linear, nonconvex feasible domain over which to optimize.

Projected dynamical systems provide a particularly useful framework to model actuation constraints and physical saturation in this context, but existing results are
(a) projected gradient flow on a convex polyhedron, (b) flow on an irregular set with non-unique trajectory, (c) periodic projected trajectory on a subset of a sphere.

Fig. 1: Qualitative behavior of projected dynamical systems: (a) projected gradient flow on a convex polyhedron, (b) flow on an irregular set with non-unique trajectory, (c) periodic projected trajectory on a subset of a sphere.

of limited applicability for complicated problems. Hence, in this paper, we consider new, generalized features for projected dynamical systems. We consider for example irregular feasible domains (Fig. 1b) for which traditional Carathéodory solutions can fail to exist or may not be unique. Furthermore, non-orthogonal projections occur in non-Euclidean spaces and may alter the dynamics. Finally, coordinate-free definitions are required to study projected dynamical systems on subsets of manifolds (Fig. 1c).

Literature review. Different approaches have been reviewed and explored to establish the results in this paper. One of the earliest formulations of projected dynamical systems goes back to [29] which establishes the existence of Carathéodory solutions on closed convex domains. In [17] this requirement is relaxed to $\mathcal{X}$ being Clarke regular (for existence) and prox-regular (for uniqueness). In the larger context of differential inclusions and viability theory [5, 6], projected dynamical systems are often presented as specific examples of more general differential inclusions, but without substantially generalizing the results of [17, 29]. In the context of variational equalities, [42] provides alternative proofs of existence and uniqueness of Carathéodory solutions when the domain $\mathcal{X}$ is a convex by using techniques from stochastic analysis. In [10] various equivalence results between the different formulations are established for convex $\mathcal{X}$. Finally, projected dynamical systems have been defined and studied in the more general context of Hilbert [14] and Banach spaces [15, 23]. The latter, in particular, is complicated by the lack of an inner product and consequently more involved projection operators [53].

The behavior of projected dynamical systems as illustrated in Fig. 1 suggests the presence of switching mechanics that result in different vector fields being active in different parts of the domain and its boundary in particular. This idea is further supported by the fact that in the study of optimization problems with a feasible domain delimited by explicit constraints, it is often useful to define the (finite) set of active constraints at a given point. This suggests that projected dynamical systems should be modeled as switched [37] or even hybrid systems [24] or hybrid automata [39, 48]. However, projected dynamical systems are much more easily (and generally) modeled as differential inclusions without explicitly considering any type of switching.

A special case of projected dynamical systems are subgradient and saddle-point flows arising in non-smooth and constrained optimization. Whereas projection-based algorithms and subgradients are ubiquitous in the analysis of iterative algorithms, work on their continuous-time counterparts is far less prominent has only been studied
Contributions. In this paper, we study a generalized class of projected dynamical systems in finite dimensions that allows for oblique projection directions. These variable projection directions are described by means of a (possibly non-differentiable) metric $g$ and are essential in providing a coordinate-free definition of projected dynamical systems on low-regularity Riemannian manifolds. Compared to previous work, we do not make a-priori assumptions on the regularity (or convexity) of the feasible domain $X$ or the vector field $f$. Instead, we strive to illustrate the necessity of those assumptions that we require by a series of (non-)examples.

Our main contribution is the development of a self-contained and comprehensive theory for this general setup. Namely, we provide weak requirements on the feasible set $X$, the vector field $f$, the metric $g$ and the differentiable structure of the underlying manifold that guarantee existence and uniqueness of trajectories, as well as other properties. Table 1 at the end of the paper concisely summarizes these results.

To be able work with projected dynamical systems on irregular domains and with discontinuous vector fields, we resort to so-called Krasovskii solutions that are a weaker notion than the classical Carathéodory solutions and are commonly used in the study of differential inclusions, because their existence is guaranteed under minimal requirements. We derive this set of regularity conditions in the specific context of projected dynamical system. Under the slightly stronger assumptions involving continuity and Clarke regularity, we show that Krasovskii solutions coincide with the classical Carathéodory solutions, thus recovering (in case of the Euclidean metric) known requirements for the existence of the latter. Finally, we lay out the requirements for uniqueness of solutions which are based on Lipschitz-continuity and a new, generalized definition of prox-regularity which suitable for low-regularity Riemannian manifolds. Our already weak regularity conditions are sharp in the sense that counter-examples can be constructed to show that requirements cannot be violated individually without the respective result failing to hold.

A major appeal of our analysis framework is its geometric nature: All of our notions are preserved by sufficiently regular coordinate transformations, which allows us to extend all of our results to constrained subsets of differential manifolds. A noteworthy by-product of this analysis is the fact that our generalized definition of prox-regularity is an intrinsic property of subsets of $C^{1,1}$ manifolds, i.e., independent of the metric, even though the traditional definition (on $\mathbb{R}^n$) suggests that prox-regularity depends on the choice of metric.

Through a series of examples, we demonstrate the application of our framework to general (nonlinear and nonconvex) optimization problems and study the stability and convergence of projected gradient dynamics under very weak regularity assumptions.

Thus, we believe that our results are not only of interest within the context of discontinuous dynamical systems, but we also envision their use in the analysis of algorithms for nonlinear, nonconvex optimization problems, possibly on manifolds. The properties developed in the present paper also form a solid foundation for constrained feedback control and online optimization in various contexts. Some preliminary results for online optimization in power systems can be found in [26,27].

Paper organization. After introducing notation and preliminary definitions in Sections 2 and 3, we establish the existence of Krasovskii solutions to projected dynamical systems on $\mathbb{R}^n$ in Section 4. In Section 5 we consider Krasovskii solutions of projected gradient systems on irregular domains and study their convergence and stability. Section 6 establishes equivalence of Krasovskii and Carathéodory solutions.
under Clarke regularity. Furthermore, we point out the connection to related work and to continuous-time subgradient flows. In Section 7, we elaborate on the requirements for uniqueness. Finally, in Section 8 we define projected dynamical systems on low-regularity Riemannian manifolds and establish the requirements on the differentiable structure that guarantee existence and uniqueness. Throughout the paper we illustrate our theoretical developments with insightful examples. Finally, Section 9 concisely summarizes our results in the form of Table 1 and concludes the paper. The appendix includes technical definitions and results that are used in proofs but are not required to understand the main results of the paper.

2. Preliminaries.

2.1. Notation. We only consider finite-dimensional spaces. Unless explicitly noted otherwise, we will work in the usual Euclidean setup for $\mathbb{R}^n$ with inner product $\langle \cdot, \cdot \rangle$ and 2-norm $\| \cdot \|$. Whenever it is informative, we make a formal distinction between $\mathbb{R}^n$ and its tangent space $T_x \mathbb{R}^n$ at $x \in \mathbb{R}$, even though they are isomorphic. For a set $A \subset \mathbb{R}^n$ we use the notation $\|A\| := \sup_{v \in A} \|v\|$. The closure, convex hull and closed convex hull of $A$ are denoted by $\text{cl} A$, $\text{co} A$, and $\overline{\text{cl}} A$, respectively. The set $A$ is locally compact if it is the intersection of a closed and an open set. A neighborhood $U \subset A$ of $x \in A$ is understood to be relative neighborhood, i.e., with respect to the subspace topology on $A$. Given a convergent sequence $\{x_k\}$, the notation $x_k \xrightarrow{A} x$ implies that $x_k \in A$ for all $k$. If $x_k \in \mathbb{R}$, the notation $x \xrightarrow{A} 0^+$ means $x_k > 0$ for all $k$ and $x_k$ converges to 0.

Let $V$ and $W$ be vector spaces endowed with norms $\| \cdot \|_V$ and $\| \cdot \|_W$, respectively, and let $A \subset V$. Continuous maps $\Phi : A \rightarrow W$ are denoted by $C^0$. The map $\Phi$ is (locally) Lipschitz (denoted by $C^{0,1}$) if for every $x \in A$ there exists $L > 0$ such that for all $z, y \in A$ in a neighborhood of $x$ it holds that

$$\|\Phi(z) - \Phi(y)\|_W \leq L \|z - y\|_V. \quad (2.1)$$

The map $\Phi$ is globally Lipschitz if (2.1) holds holds for the same $L$ for all $z, y$.

Differentiability is understood in the sense of Fréchet. Namely, if $A$ is open, then the map $\Phi$ is differentiable at $x$ if there is a linear map $D_x \Phi : V \rightarrow W$ such that

$$\lim_{y \rightarrow x} \frac{\|\Phi(y) - \Phi(x) - D_x \Phi(y - x)\|_W}{\|y - x\|_V} = 0. \quad (2.2)$$

The map $\Phi$ is differentiable ($C^1$) if it is differentiable at every $x \in A$. It is $C^{1,1}$ if it is $C^1$ and $D_x \Phi$ is $C^{0,1}$ (as function of $x$). Finally, given bases for $V$ ($\dim V = m$) and $W$ ($\dim W = n$), the Jacobian of $\Phi$ at $x$ is denoted by the $n \times m$-matrix $\nabla \Phi(x)$.

In our context, a set-valued map $F : A \rightrightarrows \mathbb{R}^n$ where $A \subset \mathbb{R}^n$ is a map that assigns to every point $x \in A$ a set $F(x) \subset T_x \mathbb{R}^n$. The set-valued map $F$ is non-empty, closed, convex, or compact if for every $x \in A$ the set $F(x)$ is non-empty, closed, convex, or compact, respectively. It is locally bounded if for every $x \in A$ there exists $L > 0$ such that $\|F(y)\| \leq L$ for all $y \in A$ in a neighborhood of $x$. The same definition also applies to single-valued functions. The map $F$ is bounded if there exists $L > 0$ such that $\|F(y)\| \leq L$ for all $x \in A$. The inner and outer limits of $F$ at $x$ are denoted by $\liminf_{y \rightarrow x} F(y)$ and $\limsup_{y \rightarrow x} F(y)$ respectively (see appendix for a formal definition and summary of continuity concepts which are required for certain proofs only).

2.2. Tangent and Clarke Cones. The ensuing definitions follow [46, Chap 6].
Definition 2.1. Given a set $\mathcal{X} \subset \mathbb{R}^n$ and $x \in \mathcal{X}$, a vector $v \in T_x \mathbb{R}^n$ is a tangent vector of $\mathcal{X}$ at $x$ if there exist sequences $x_k \to x$ and $\delta_k \to 0^+$ such that $\frac{x_k - x}{\delta_k} \to v$. The set of all tangent vectors is the tangent cone of $\mathcal{X}$ at $x$ and denoted by $T_x \mathcal{X}$.

The tangent cone $T_x \mathcal{X}$ (also known as (Bouligand’s) contingent cone [13]) is closed and non-empty (namely, $0 \in T_x \mathcal{X}$) for any $x \in \mathcal{X}$.

In the following definition of Clarke regularity and in most of the paper we limit ourselves to locally compact subsets of $\mathbb{R}^n$. In our context, a more general definition of Clarke regularity does not improve our results and only adds to the technicalities.

Definition 2.2. For a locally compact set $\mathcal{X} \subset \mathbb{R}^n$ the Clarke tangent cone at $x \in \mathcal{X}$ is defined as the inner limit of the tangent cones, i.e., $T^C_x \mathcal{X} := \liminf_{y \to x} T_y \mathcal{X}$.

By definition of the inner limit, we have $T^C_x \mathcal{X} \subseteq T_x \mathcal{X}$. Furthermore, $T^C_x \mathcal{X}$ is closed, convex and non-empty for all $x \in \mathcal{X}$ [46, Thm 6.26].

Definition 2.3. We call a set $\mathcal{X} \subset \mathbb{R}^n$ Clarke regular at $x$ if it is locally compact and $T^C_x \mathcal{X} = T_x \mathcal{X}$. The set $\mathcal{X}$ is Clarke regular if it is Clarke regular for all $x \in \mathcal{X}$.

Example 2.4 (sets defined by inequality constraints). Let $h : \mathbb{R}^n \to \mathbb{R}^m$ be $C^1$ such that $\nabla h(x)$ has full rank for all $x$.\footnote{This rank condition is a standard constraint qualification in nonlinear programming [8]. In general, instead of $\nabla h(x)$ having full rank for all $x$, it suffices that for a given $x$ only the active constraints (i.e., $\nabla h_{i(x)}(x)$) have full rank. Furthermore, equality constraints can be easily incorporated.} Then, the set $\mathcal{X} := \{x \mid h(x) \leq 0\}$ is Clarke regular [46, Thm 6.31]. In particular, let $h$ be expressed componentwise as $h(x) = [h_1(x), \ldots, h_m(x)]^T$, let $I(x) := \{i \mid h_i(x) = 0\}$ denote the set of active constraints at $x \in \mathcal{X}$ and define $h_{I(x)} := [h_i(x)]_{i \in I(x)}$ as the function obtained from stacking the active constraint functions. Then, the (Clarke) tangent cone at $x$ in the canonical basis is given by $T^C_x \mathcal{X} = T_x \mathcal{X} = \{v \mid \nabla h_{I(x)}(x)v \leq 0\}$.

2.3. Low-regularity Riemannian metrics. A natural extension for projected dynamical systems are oblique projection directions. These are conveniently defined via a (Riemannian) metric which defines a variable inner product on $T_x \mathbb{R}^n$ as function
of $x$. Furthermore, the notion of a Riemannian metric is essential to define projected dynamical systems in a coordinate-free setup on manifolds.

We quickly review the definition of bilinear forms and inner products. Let $L^2_n$ denote the space of bilinear forms on $\mathbb{R}^n$, i.e., every $g \in L^2_n$ is a map $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ such that for every $u, v, w \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$ it holds that $g(u + v, w) = g(u, w) + g(v, w)$ and $g(u, v + w) = g(u, v) + g(u, w)$ as well as $g(\lambda v, w) = \lambda g(v, w) = g(v, \lambda w)$. Given the canonical basis of $\mathbb{R}^n$, $g$ can be written in matrix form as $g(u, v) := u^T G v$ where $G \in \mathbb{R}^{n \times n}$. In particular, $L^2_n$ is itself a $n^2$-dimensional space isomorphic to $\mathbb{R}^{n \times n}$.

An inner product $g \in L^2_n$ is a symmetric, positive-definite bilinear form, that is, for all $u, v \in \mathbb{R}^n$ we have $g(u, v) = g(v, u)$. Further, $g(u, u) \geq 0$, and $g(u, u) = 0$ holds if and only if $u = 0$. If $g$ is an inner product we use the notation $\langle u, v \rangle := g(u, v)$. In matrix form, we can write $\langle u, v \rangle := u^T G v$ where $G$ is symmetric positive definite.

We write $\| \cdot \|_g$ given by $\|v\|_g := \sqrt{\langle v, v \rangle}$ to denote the 2-norm induced by $g$. The maximum and minimum eigenvalues of $g$ are denoted by $\lambda^\text{max}_g := \max\{\|v\|_g \mid \|v\| = 1\}$ and $\lambda^\text{min}_g := \min\{\|v\|_g \mid \|v\| = 1\}$ respectively, and the condition number is defined as $\kappa_g := \lambda^\text{max}_g / \lambda^\text{min}_g$.

In this context, also recall that the 2-norms induced by any two inner products on a finite-dimensional vector space are equivalent, that is, for a vector space $V$ with norms $\| \cdot \|_a$ and $\| \cdot \|_b$ there are constants $\ell > 0$ and $L > 0$ such that for every $v \in V$ it holds that $\ell \|v\|_a \leq \|v\|_b \leq L \|v\|_a$. For instance, $\ell = \lambda^\text{min}_b / \lambda^\text{max}_a$ and $L = \lambda^\text{max}_b / \lambda^\text{min}_a$.

Hence, we can define a metric as a variable inner product over a given set.

**Definition 2.5.** Given a set $\mathcal{X} \subset \mathbb{R}^n$, a (Riemannian) metric is a map $g : \mathcal{X} \to L^2_n$ that assigns to every point $x \in \mathcal{X}$ an inner product $\langle \cdot, \cdot \rangle_{g(x)}$. A metric is (Lipschitz) continuous if is (Lipschitz) continuous as a map from $\mathcal{X}$ to $L^2_n$.

If clear from the context at which point $x$ the metric $g$ is applied, we drop the argument in the subscript and write $\langle \cdot, \cdot \rangle_{g}$ or $\| \cdot \|_g$. We always retain the subscript $g$, in order to draw a distinction between the Euclidean norm $\| \cdot \|$.

Since $g$ is positive definite for all $x$ by definition, it follows that $\lambda^\text{max}_{g(x)}$, $\lambda^\text{min}_{g(x)}$ and $\kappa_{g(x)}$ are well-defined for all $x$. However, $\kappa_{g(x)}$ is not necessarily locally bounded (even if $g$ is bounded as a map). In particular, $\lambda^\text{min}_{g(x)}$ might not be bounded below, away from 0. Hence, for metrics we require the following definition of local boundedness.

**Definition 2.6.** A metric $g$ on $\mathcal{X}$ is locally weakly bounded if for every $x \in \mathcal{X}$ there exist $\ell, L > 0$ such that $\ell \leq \kappa(y) \leq L$ holds for all $y \in \mathcal{X}$ in a neighborhood of $x$. It is weakly bounded if $\ell \leq \kappa(x) \leq L$ holds for all $x \in \mathcal{X}$.

A metric $g$ can be locally weakly bounded even if its not locally bounded as a map $\mathcal{X} \to L^2_n$. Furthermore, since maximum and minimum eigenvalues (and hence the condition number) are continuous functions of a metric (or the representing matrix) it follows that a continuous metric is always locally weakly bounded.

**Remark 2.7.** In the following, we will continue to use the Euclidean norm as a distance function on $\mathbb{R}^n$ and use any Riemannian metric only in the context of projection directions. Thereby, we avoid the notational complexity introduced by Riemannian geometry, and more importantly we do not need to make an a priori assumption on the differentiability on the metric $g$ (which is a prerequisite for many Riemannian constructs to exist), thus preserving a high degree of generality.

**2.4. Normal Cones.** Given a metric $g$, we can define (oblique) normal cones induced by $g$ (see Fig. 2c).
DEFINITION 2.8. Let $\mathcal{X} \subset \mathbb{R}^n$ be Clarke regular and let $g$ be a metric on $\mathcal{X}$, then the normal cone at $x \in \mathcal{X}$ with respect to $g$ is defined as the polar cone of $T_x^C \mathcal{X}$ with respect to the metric $g$, i.e.,

$$N_x^g \mathcal{X} := \left( T_x^C \mathcal{X} \right)^* = \left\{ \eta \mid \forall v \in T_x^C \mathcal{X} : \langle v, \eta \rangle_{g(x)} \leq 0 \right\}.$$ 

The normal cone with respect to the Euclidean metric is simply denoted by $N_x \mathcal{X}$.

Remark 2.9. For simplicity, we will use the notion of normal cone only in the context of Clarke regular sets. If $\mathcal{X}$ is not Clarke regular, one needs to distinguish between the regular, general and Clarke normal cones [46].

Example 2.10 (normal cone to constraint-defined sets). As in Example 2.4 consider $\mathcal{X} := \{ x \mid h(x) \leq 0 \}$ where $h : \mathbb{R}^n \to \mathbb{R}^m$ is $C^1$ and $\nabla h(x)$ has full rank for all $x$. Further, let $g$ denote a metric on $\mathcal{X}$ represented by $G(x) \in \mathbb{R}^{n \times n}$. Then, the normal cone of $\mathcal{X}$ at $x$ is given by

$$N_x^g \mathcal{X} = \left\{ \eta \mid \eta = \sum_{i \in I(x)} \alpha_i G^{-1}(x) \nabla h_i(x)^T, \alpha_i \geq 0 \right\}$$

which can be derived by inserting any $\eta$ into (2.3) and using $T_x \mathcal{X}$ in Example 2.4.

3. Projected Dynamical Systems. With the above notions we can now formally define our main object of study.

DEFINITION 3.1. Given a set $\mathcal{X} \subset \mathbb{R}^n$, a metric $g$ on $\mathcal{X}$, and a vector field $f : \mathcal{X} \to \mathbb{R}^n$, the projected vector field of $f$ is defined as the set-valued map

$$\Pi^g_{\mathcal{X}} : \mathcal{X} \ni x \mapsto \arg \min_{v \in T_x \mathcal{X}} \| v - f(x) \|^2_{g(x)}$$

For simplicity, we call $\Pi^g_{\mathcal{X}} f$ a vector field even though $\Pi^g_{\mathcal{X}} f(x)$ might not be a singleton. We will write $\Pi f$ whenever $\mathcal{X}$ and $g$ are clear from the context.

Example 3.2 (pointwise evaluation of a projected vector field). As in Examples 2.4 and 2.10 let $\mathcal{X} := \{ x \mid h(x) \leq 0 \}$ where $h : \mathbb{R}^n \to \mathbb{R}^m$ is $C^1$ and $\nabla h(x)$ has full rank for all $x$ and let $g$ denote a metric on $\mathcal{X}$ represented by $G(x) \in \mathbb{R}^{n \times n}$. Furthermore, consider a vector field $f : \mathcal{X} \to \mathbb{R}^n$. Then, the projected vector field $\Pi^g_{\mathcal{X}} f(x)$ at $x \in \mathcal{X}$ is given by the solution of the convex quadratic program

$$\text{minimize} \quad \langle f(x) - v \rangle^T G(x) (f(x) - v) \quad \text{subject to} \quad \nabla h_i(x) v \leq 0.$$ 

Note that $x$ is not an optimization variable. Hence, the properties of $f$ and $g$ as function of $x$ are irrelevant when doing a pointwise evaluation of $\Pi^g_{\mathcal{X}} f(x)$.

Since $T_x \mathcal{X}$ is non-empty and closed, a minimum norm projection exists, and therefore $\Pi^g_{\mathcal{X}} f(x)$ is non-empty for all $x \in \mathcal{X}$. Hence, a projected dynamical system is described by the initial value problem

$$(3.2) \quad \dot{x} \in \Pi^g_{\mathcal{X}} f(x), \quad x(0) = x_0,$$

where $x_0 \in \mathcal{X}$. If $T_x \mathcal{X}$ is convex for all $x$ then $\Pi^g_{\mathcal{X}} f(x)$ is a singleton for all $x \in \mathcal{X}$ (note that $\| v - f(x) \|^2_{g(x)}$ is always strictly convex as function of $v$). In this case we

---

See, e.g., the first part of the proof of Hilbert’s projection theorem [44, Prop 1.37].
will slightly abuse notation and not distinguish between the set-valued map and its induced vector field, i.e., instead of (3.2) we simply write \( \dot{x} = \Pi^g_X f(x) \), \( x(0) = x_0 \).

An absolutely continuous function \( x : [0, T) \to X \) with \( T > 0 \) and \( x(0) = x_0 \) that satisfies \( \dot{x} \in \Pi^g_X f(x) \) almost everywhere (i.e., for all \( t \in [0, T) \) except on a subset of Lebesgue measure zero) is called a Carathéodory solution to (3.2).

**Remark 3.3.** The class of systems (3.2) can be generalized to \( f \) being set-valued, i.e., \( f : \mathbb{R}^n \Rightarrow \mathbb{R}^n \). This avenue has been explored in [5, 6, 17, 29], albeit only for \( g \) Euclidean and \( X \) Clarke regular. In order not to overload our contributions with technicalities we assume that \( f \) is single-valued, although an extension is possible. ■

As the following example shows, Carathéodory solutions to (3.2) can fail to exist unless various regularity assumptions \( X \), \( f \) and \( g \) hold. Hence, in the next section we propose the use of Krasovskii solutions which exist in more general settings. Furthermore, we will show that the Krasovskii solutions reduce to Carathéodory solutions under the same assumptions that guarantee the existence of the latter.

**Example 3.4 (non-existence of Carathéodory solution).** Consider \( \mathbb{R}^2 \) with the Euclidean metric, the uniform “vertical” vector field \( f = (0, 1) \), and the self-similar closed set \( X \) illustrated in Figure 3 and defined by

\[
X = \left\{ (x_1, x_2) \mid \forall k \in \mathbb{Z} : x_2 = \pm 2x_1 - \frac{2}{\sqrt{k^2 + 2^3 + 2^9}} |x_2| \leq |x_1| \right\} \cup \{0\}.
\]

![Fig. 3](image_url)

(a) Tangent cone and projected vector field at 0, (b) local equilibria for Example 3.4 and (c) Krasovskii regularization for Example 4.4 at 0.

The tangent cone at 0 is given by \( T_0X = \{(v_1, v_2) \mid |v_2| \leq |v_1|\} \). It is not “derivable”, that is, there are no differentiable curves leaving 0 in a tangent direction and remaining in \( X \). However, by definition there is a sequence of points in \( X \) approaching 0 in the direction of any tangent vector. At 0 the projection of \( f \) on the tangent cone is not unique as seen in Figure 3a, namely \( \Pi f(0) = \{(\frac{1}{2}, \frac{1}{2}), (-\frac{1}{2}, \frac{1}{2})\} \).

Furthermore, there is no Carathéodory solution to \( \dot{x} \in \Pi^g_X f(x) \) for \( x(0) = 0 \). To see this, we can argue that any solution starting at 0 can neither stay at 0 nor leave 0. More precisely, on one hand the constant curve \( x(t) = 0 \) for \( t \in [0, T) \) with \( T > 0 \) cannot be a solution since it does not satisfy \( \dot{x} \in \Pi f(0) \). On the other hand, the points \( p_k = \left( \pm \frac{2}{\sqrt{k^2 + 2^3 + 2^9}}, \frac{2}{\sqrt{k^2 + 2^3 + 2^9}} \right) \) illustrated in Figure 3b are locally asymptotically stable equilibria of the system. Namely there is an equilibrium point arbitrarily close to 0. Thus, loosely speaking, any solution leaving 0 would need to converge to an equilibrium arbitrarily close to 0. ■
4. Existence of Krasovskii solutions. The pathology in Example 3.4 can be resolved either by placing additional assumptions on the feasible set $\mathcal{X}$ or by relaxing the notion of a solution. In this section we focus on the latter.

**Definition 4.1.** Given a set-valued map $F : \mathcal{X} \rightrightarrows \mathbb{R}^n$, its Krasovskii regularization is defined as the set-valued map given by

$$K[F] : \mathcal{X} \rightrightarrows \mathbb{R}^n \quad x \mapsto \varlimsup_{y \to x} F(y).$$

Given a set-valued map $F : \mathcal{X} \rightrightarrows \mathbb{R}^n$, an absolutely continuous function $x : [0, T) \to \mathbb{R}^n$ with $T > 0$ and $x(0) = x_0$ is a Krasovskii solution of the inclusion

$$\dot{x} \in F(x), \quad x(0) = x_0$$

if it satisfies $\dot{x} \in K[F](x)$ almost everywhere. In other words, a Carathéodory solution to the regularized set-valued map $K[F]$ is a Krasovskii solution of the original problem.

Hence we can state the following existence result about Krasovskii solutions.

**Theorem 4.2 (existence of Krasovskii solutions).** Let $\mathcal{X} \subseteq \mathbb{R}^n$ be a locally compact set, $f : \mathcal{X} \to \mathbb{R}^n$ a locally bounded vector field and $g$ a locally weakly bounded metric defined on $\mathcal{X}$. Then, for any $x_0 \in \mathcal{X}$ there exists a Krasovskii solution $x : [0, T) \to \mathcal{X}$ for some $T > 0$ to

$$\dot{x} \in \Pi^0_{\mathcal{X}} f(x), \quad x(0) = x_0.$$  \hspace{1cm} (4.1)

In addition, for $r > 0$ such that $U_r := \{x \in \mathcal{X} \mid \|x - x_0\| \leq r\}$ is closed and $L = \max_{x \in U_r} \|K[\Pi^0_{\mathcal{X}} f](y)\|$ exists, the solution is $C^{0,1}$ and exists for $T > r/L$.

**Proof.** We show that the general existence result [25, Cor 1.1] (Proposition A.7) is applicable to Krasovskii regularized projected vector fields. Namely, we need to verify that $K[\Pi^0_{\mathcal{X}} f]$ is convex, compact, non-empty, upper semicontinuous (usc), and

$$K[\Pi^0_{\mathcal{X}} f](x) \cap T_x \mathcal{X} \neq \emptyset \quad \forall x \in \mathcal{X}. \hspace{1cm} (4.2)$$

The fact that $K[\Pi^0_{\mathcal{X}} f]$ is closed and convex is immediate from its definition. It is non-empty since $\Pi^0_{\mathcal{X}} f(x)$ is non-empty and $\Pi^0_{\mathcal{X}} f(x) \subseteq K[\Pi^0_{\mathcal{X}} f](x)$ for all $x \in \mathcal{X}$. Further, we have $\Pi^0_{\mathcal{X}} f(x) \subseteq T_x \mathcal{X}$ by definition for all $x \in \mathcal{X}$ and therefore (4.2) holds.

For the rest of the proof let $F(x) := \limsup_{y \to x} \Pi^0_{\mathcal{X}} f(y)$ (hence, $K[\Pi^0_{\mathcal{X}} f] = \varlimsup F$).

Next, we show that $K[\Pi^0_{\mathcal{X}} f](x)$ is compact for all $x \in \mathcal{X}$. For this, we first introduce an auxiliary metric $\hat{g}$ defined as $\hat{g}(x) := g(x)/\lambda_{g}^{\max}(x)$, that is, we scale the metric at every $x \in \mathcal{X}$ by dividing it by its maximum eigenvalue at that point. This implies that $\|f(x)\|_{\hat{g}(x)} \leq \|f(x)\|$ for all $x \in \mathcal{X}$. Note that the projected vector field is unchanged, i.e., $\Pi^0_{\mathcal{X}} f = \Pi^0_{\mathcal{X}} \hat{f}$, since in (3.1) only the objective function is scaled. Furthermore, $\kappa_{\hat{g}}(x) = \kappa_{g}(x)$ for all $x \in \mathcal{X}$, and consequently $\hat{g}$ is locally weakly bounded since $g$ is locally weakly bounded.

Given any $x \in \mathcal{X}$, since $0 \in T_x \mathcal{X}$ it follows that $\|v\|_{\hat{g}(x)} \leq \|f(x) - 0\|_{\hat{g}(x)}$ for every $v \in \Pi^0_{\mathcal{X}} f(x)$. Consequently, by local boundedness of $\hat{g}$ there exists $L'' > 0$ such that $\|\Pi^0_{\mathcal{X}} f(y)\|_{\hat{g}(y)} \leq L''$ for every $y \in \mathcal{X}$ in a neighborhood of $x$. Furthermore, by weak local boundedness of $\hat{g}$ there exists $L' > 0$ such that $\kappa_{\hat{g}}(x) \leq L'$ in a neighborhood of $x$. Since $\lambda_{g}^{\max}(x) = 1$, it follows that $\lambda_{g}^{\min}(x) \geq 1/L'$ and therefore $\|v\| \leq L'\|v\|_{\hat{g}(y)}$ for all $v \in T_y \mathbb{R}^n$ and all $y \in \mathcal{X}$ in a neighborhood of $x$. Combining these arguments, there exist $L', L'' > 0$ such that for every $y \in \mathcal{X}$ in a neighborhood of $x$ it holds that

$$\frac{1}{L'}\|\Pi^0_{\mathcal{X}} f(y)\| \leq \|\Pi^0_{\mathcal{X}} f(y)\|_{\hat{g}(y)} \leq \|f(y)\|_{\hat{g}(y)} \leq \|f(y)\| \leq L''.$$  \hspace{1cm} (4.3)
Hence, since $\Pi^g_{\mathcal{X}} f = \Pi^g_{\mathcal{X}} f$, it follows that $\Pi^g_{\mathcal{X}} f$ is locally bounded.

Let $U \subset \mathcal{X}$ be a compact neighborhood of $x$ such that (4.3) holds. Consider the graph of $\Pi^g_{\mathcal{X}} f$ restricted to $U$ given by $g \Pi^g_{\mathcal{X}} f|_U := \{(x, v) \mid x \in U, v \in \Pi^g_{\mathcal{X}} f(x)\}$. By definition of the outer limit we have $\text{cl} \, g \Pi^g_{\mathcal{X}} f|_U = g \Pi^g_{\mathcal{X}} f|_U$, i.e., $F$ is the so-called closure of $\Pi^g_{\mathcal{X}} f|_U$ [46, p. 154]. Thus, since $g \Pi^g_{\mathcal{X}} f|_U$ is bounded, $g \Pi^g_{\mathcal{X}} f|_U$ is compact, and consequently $F(y)$ is locally bounded for every $y \in U$. In particular, since $F(x)$ is compact, and the closed convex hull of a bounded set is compact [30, Thm 1.4.3], it follows that $\text{co} F(x) = K[\Pi^g_{\mathcal{X}} f](x)$ is compact for all $x \in \mathcal{X}$.

Finally, we need to show that $K[\Pi^g_{\mathcal{X}} f]$ is usc. For this, note that the map $F$ is outer semicontinuous (osc) and closed by definition. Furthermore, it is locally bounded (as shown above). Consequently, by Lemma A.4, $F$ is also usc. Hence, Lemma A.5 states that $\text{co} F$ is usc as well. Since $F(x)$ is compact for all $x \in \mathcal{X}$, it follows that $\text{co} F(x) = \text{co} F(x)$ [30, Thm 1.4.3], and therefore $K[\Pi^g_{\mathcal{X}} f] = \text{co} F$ is usc.

Thus, $K[\Pi^g_{\mathcal{X}} f]$ satisfies the conditions for Proposition A.7 to be applicable, and therefore the existence of Krasovskii solution to (4.1) is guaranteed for all $x_0 \in \mathcal{X}$. □

Besides weaker requirements for existence, the choice to consider Krasovskii solutions is also motivated by their inherent “robustness” towards perturbations, i.e., solutions to a perturbed system still approximate the solutions of the nominal systems [24, Chap 4]. In the same spirit, one can also establish results about the continuous dependence of solutions on initial values and problem parameters [21].

The existence of solutions for $t \to \infty$ is guaranteed under the following conditions.

**Corollary 4.3 (existence of complete solutions).** Consider the same setup as in Theorem 4.2. If either

(i) $\mathcal{X}$ is closed, $f$ is bounded, and $g$ is weakly bounded, or

(ii) $\mathcal{X}$ is compact, $f$ and $g$ are continuous, or

(iii) $\mathcal{X}$ is closed, $f$ is globally Lipschitz and $g$ is weakly bounded,

then for every $x_0 \in \mathcal{X}$ every Krasovskii solution to (4.1) can be extended to $T \to \infty$.

**Proof.** (i) If $f$ is bounded and $g$ is weakly bounded, then the local boundedness argument of the proof of Theorem 4.2 can be applied globally, i.e., (4.3) holds for all $y \in \mathcal{X}$ for the same $L'$, $L''$ and hence $K[\Pi^g_{\mathcal{X}} f]$ is bounded. Hence, in Theorem 4.2 the constant $L > 0$ exists for $r \to \infty$ and consequently $T \to \infty$.

(ii) Since $f$ is continuous it only takes bounded values on a compact set. Furthermore, continuity of $g$ implies local weak boundedness, i.e., for every $x \in \mathcal{X}$ there exist $\ell_x, L_x > 0$ such that $\ell_x < \kappa_g(y) < L_x$ for all $y \in \mathcal{X}$ in a neighborhood of $x$. Since $\mathcal{X}$ is compact, there exist $\ell := \min_{x \in \mathcal{X}} \ell_x$ and $L := \max_{x \in \mathcal{X}} L_x$ and (4.3) holds for all $y \in \mathcal{X}$. Hence, $g$ is weakly bounded. Then, the same arguments as for (i) apply.

(iii) Assume without loss of generality that $0 \in \mathcal{X}$ (possibly after a linear translation). Global Lipschitz continuity of $f$ implies the existence of $L'' > 0$ such that $\|f(x)\| \leq L''(\|x\| + 1)$ for all $x \in \mathcal{X}$ (linear growth property [5]). To see this, recall that by the reverse triangle inequality and the definition of Lipschitz continuity there exists $L' > 0$ such that $\|f(x) - f(0)\| \leq L'\|x\|$ for all $x, y \in \mathcal{X}$. It follows that $\|f(x)\| \leq L'\|x\| + \|f(0)\|$ and hence $L''$ can be chosen as the maximum of $L'$ and $\|f(0)\|$ to yield the linear growth property.

Since $g$ is weakly bounded, the same arguments used for (4.3) can be used to establish that there exists $L''' > 0$ such that for all $x \in \mathcal{X}$ it holds that

$$L''' \|\Pi^g_{\mathcal{X}} f(x)\| < \|\Pi^g_{\mathcal{X}} f(x)\|_{g(x)} \leq \|f(x)\|_{g(x)} \leq \|f(x)\| < L''(\|x\| + 1).$$

It follows by the same arguments as in the proof of Theorem 4.2 that $\|K[\Pi^g_{\mathcal{X}} f](x)\| \leq L(\|x\| + 1)$ where $L = L''/L'''$, i.e., the linear growth condition applies to $K[\Pi^g_{\mathcal{X}} f]$. 


Hence using standard bounds [5, p. 100], one can conclude that any Krasovskii solution to (4.1) satisfies \( \|x(t)\| \leq ((\|x_0\| + 1)e^{Lt} \). Namely, define \( u(t) := L(\|x(t)\| + 1) \) and note that \( \dot{u}(t) = L^2(\|x(t)\| + 1) = Lu(t) \) holds for all \( t \) where \( \dot{x}(t) \) exists. Hence, Gronwall’s inequality (for discontinuous ODEs) implies the desired bound. It immediately follows that \( x(t) \) cannot have finite escape time and therefore can be extended to \( t \to \infty \), completing the proof of (iii). \( \square \)

**Example 3.4.** The Krasovskii regularization at \( 0 \) of the projected vector field \( \Pi f \) is continuous vector field on \( \hat{N} \) singleton and there is a cone we have \( \Pi f \) and therefore \( \Pi f \) is a singleton and there is \( \hat{N} \) such that the following equivalent statements hold:

1. \( \Pi f \) is Clarke regular at \( x \), then \( \Pi f \) is a singleton and there is \( \hat{N} \) such that the following equivalent statements hold:
2. \( \Pi f \) is Clarke regular at \( x \), then \( \Pi f \) is a singleton and there is \( \hat{N} \) such that the following equivalent statements hold:
3. \( \Pi f \) is Clarke regular at \( x \), then \( \Pi f \) is a singleton and there is \( \hat{N} \) such that the following equivalent statements hold:

Proof. Let \( v \in \Pi f \). As \( T_x X \) is a cone we have \( \lambda v \in T_x X \) for all \( \lambda \geq 0 \). Since \( v \) (locally) minimizes \( \|v - f(x)\|_g(x) \) over \( T_x X \), it follows that \( \lambda = 1 \) minimizes \( M(\lambda) := \frac{1}{2} \|\lambda v - f(x)\|_g(x) \) for \( v \) fixed. Hence, for \( \lambda = 1 \) the optimality condition \( \frac{2M}{g(x)}(\lambda) = \lambda \langle v - f(x), v \rangle_g(x) = 0 \) holds. This proves the first part. The second part follows from Moreau’s Theorem [30, Thm 3.2.5] since \( T_x X \) is convex by Clarke regularity. \( \square \)

**Lemma 4.4.** Consider \( X \subset \mathbb{R}^n \), let \( g \) be a continuous metric on \( X \) and \( f \) a continuous vector field on \( X \). Then, for every \( x \in K[Pi \lambda f](x) \), one has \( \langle f(x), v \rangle_g(x) \geq \|v\|_g(x) \). If in addition \( X \) is Clarke regular, then for \( \hat{N} := f(x) - v \) we have \( \hat{N} \in N^2 X \).

Proof. Let \( F(x) := \limsup_{y \to x} \Pi f(y) \). By definition of the outer limit, there exist sequences \( x_k \to x \) with \( x_k \in X \) and \( v_k \to v \) with \( v_k \in \Pi f(x_k) \) for every \( v \in F(x) \) and every \( x \in X \). In particular, \( \langle f(x_k), v_k \rangle_{g(x_k)} = \|v_k\|_g(x_k) \) holds for every \( k \) by Lemma 4.5. Since \( f \) and \( g \) are continuous the equality holds in the limit, i.e., \( \langle f(x), v \rangle_g(x) = \|v\|_g(x) \) for every \( v \in F(x) \). Taking any convex combination \( v = \sum_i \alpha_i v_i \) with \( v_i \in F(x) \) and \( \alpha_i \geq 0 \) and \( \sum_i \alpha_i = 1 \), we have

\[
\sum_i \langle f(x), \alpha_i v_i \rangle_g(x) = \sum_i \alpha_i \|v_i\|_g(x) \geq \sum_i \alpha_i \|v_i\|_g(x) = \|v\|_g(x),
\]

and therefore \( \langle f(x), v \rangle_g(x) \geq \|v\|_g(x) \) for every \( v \in \mathbb{R} F(x) = K[Pi \lambda f](x) \).

According to Lemma 4.5, if \( X \) is Clarke regular, given a sequence \( x_k \to x \), the sequences \( v_k = \Pi f(x_k) \) and \( \hat{y}_k := N^2 X \) for which \( \hat{y}_k = f(x_k) - \Pi f(x_k) \) are uniquely defined. Since \( g \) is continuous, the mapping \( x \mapsto N^2 X \) is outer semi-continuous (Lemma A.6) and therefore \( \lim_{k \to \infty} \hat{y}_k \in N^2 X \). In other words, for every \( v \in F(x) \) it holds that \( f(x) - v \in N^2 X \). Since by Clarke regularity \( N^2 X \) is convex, it follows that, for any convex combination \( \eta = \sum_i \alpha_i (f(x) - v_i) \) with \( v_i \in F(x) \) and \( \alpha_i \geq 0 \)
and $\sum \alpha_i = 1$, it must hold that $\eta \in N^2\mathcal{X}$, which completes the proof. \hfill \Box

5. Illustration: Stability & Projected Gradient Descent. To illustrate how established stability concepts seamlessly apply to Krasovskii solutions of projected dynamical systems, we consider projected gradient systems, i.e., projected dynamical systems for which the vector field is the gradient of a function. Naturally, these systems are of prime interest for constrained optimization. The same techniques can also be used to assess the stability of equilibria of other vector fields ranging from saddle-point flows [11] to momentum methods [52]. In what follows, we will establish convergence and stability results that generalize our work in [26].

For simplicity, we consider systems defined on a subset of $\mathbb{R}^n$. Extensions to subsets of manifolds will be made possible by the results of the forthcoming Section 8.

5.1. Preliminaries and LaSalle Invariance. In this section, we only consider projected dynamical systems with complete Krasovskii solutions (see Corollary 4.3).

Assumption 5.1. For a feasible set $\mathcal{X}$, a metric $g$ and a vector field $f$ both defined on $\mathcal{X}$, we assume that for every $x_0 \in \mathcal{X}$ every Krasovskii solution $x : [0, T) \to \mathcal{X}$ of

\begin{equation}
\dot{x} \in \Pi^2 \mathcal{X} f(x), \quad x(0) = x_0,
\end{equation}

can be extended to $T \to \infty$. \hfill \Box

We use the usual notions for the limiting behavior of trajectories of discontinuous dynamical systems [18]. Namely, a set $S \subset \mathcal{X}$ is weakly invariant if for every $x_0 \in S$ there exists a solution starting at $x_0$ and remaining in $S$ for all $t \in [0, \infty)$. A set $S$ is strongly invariant if all solutions starting at $x_0$ in $S$ remain in $S$ for all $t \in [0, \infty)$. The union of weakly (strongly) invariant subsets is again weakly (strongly) invariant, hence the notion of largest weakly (strongly) invariant set is well-defined.

A point $\hat{x} \in \mathcal{X}$ is a limit point for a solution $x$ of (5.1) if there exist a sequence $t_k \to \infty$ such that $x(t_k) \to \hat{x}$. The set of all limit points of $x$ is called the $\omega$-limit set and denoted by $\Omega(x)$. Note that $\Omega(x)$ is always weakly invariant. Furthermore, if $x$ is bounded, then $x(t)$ converges to $\Omega(x)$ for $t \to \infty$ [21, §12.4]. The point $\hat{x}$ is a weak equilibrium if the constant function $x(t) = \hat{x}$ for all $t \geq 0$ is a solution of the dynamical system (but possibly not unique). Similarly, $\hat{x}$ is a strong equilibrium if $x(t) = \hat{x}$ is the only solution starting at $\hat{x}$.

A set $S$ is strongly stable if for every neighborhood $U$ of $S$ there exists another neighborhood $V \subset U$ of $S$ such that every solution starting in $V$ remains in $U$ for all $t \in [0, \infty)$. The set $S$ is strongly asymptotically stable if it is strongly stable and every trajectory starting in $V$ converges to $S$.

Given a $C^1$ scalar-valued function $\Psi$ defined on an open neighborhood of $\mathcal{X}$, the set-valued Lie derivative of $\Psi$ with respect to a map $F : \mathcal{X} \to \mathbb{R}^n$ is defined on $\mathcal{X}$ as

$$
\mathcal{L}_F \Psi : \mathcal{X} \ni x \mapsto \{a \in \mathbb{R} | \exists v \in F(x) : D_x \Psi(v) = a\}.
$$

Hence, the following invariance principle is modified from [7] in so far as it requires the dynamical system to be defined only on a (possibly closed) subset of $\mathbb{R}^n$. We provide a proof for completeness. In similar fashion, stability and invariance results for differential inclusions, as found in [5,24,37] and references therein, can be specialized to the case of projected dynamical systems.

Theorem 5.2. [adapted from [7, Thm 3]] Consider a projected dynamical system (5.1) satisfying Assumption 5.1. Furthermore, let $\Psi : V \to \mathbb{R}$ be a $C^1$ function defined on an open neighborhood $V$ of $\mathcal{X}$ such that for every $\ell \in \mathbb{R}$ the set $S_\ell :=$
\( \{ x \mid \Psi(x) \leq \ell \} \cap \mathcal{X} \) is compact. If \( \max \mathcal{L}_{\Pi_x} \Psi(x) \leq 0 \) for all \( x \in \mathcal{X} \), then every solution to (5.1) starting at \( x_0 \in S_t \) will converge to the largest weakly invariant subset of \( \text{cl}\{x \in V \mid 0 \in \mathcal{L}_{\Pi_x} \Psi(x)\} \cap S_t \).

**Proof.** First, we verify that if \( x_0 \in S_t \), then any solution \( x \) of (5.1) remains in \( S_t \), i.e., \( S_t \) is strongly invariant and \( x \) is bounded (since \( S_t \) is compact). Clearly, by definition \( x(t) \in \mathcal{X} \) for all \( t \). Further, assume that there exists \( \tau \) such that \( x(\tau) \notin \{ x \mid \Psi(x) \leq \ell \} \). This, however, contradicts the fact that \( \Psi \) and \( x \) are continuous and \( \mathcal{L}_{\Pi_x} \Psi(x) \leq 0 \) holds almost everywhere. Namely, we must have

\[
\Psi(x(\tau)) = \Psi(x_0) + \int_0^\tau D_x \Psi(\dot{x}(t)) \, dt \leq \Psi(x_0).
\]

Second, we show that \( \Psi \) is constant on \( \Omega(x) \) where \( x \) is a given trajectory. Namely, \( \Psi \circ x \) is continuous and bounded below since \( \Psi \) is continuous and \( x \) is bounded. Further, \( \Psi \circ x \) is non-increasing, and therefore \( \lim_{t \to \infty} (\Psi \circ x)(t) = c \) exists. Furthermore, for any limit point \( \hat{x} \in \Omega(x) \) for which \( t_k \to \infty \) and \( x(t_k) \to \hat{x} \) it must hold that \( \Psi(\hat{x}) = c \) where \( c \) depends on the trajectory \( x \) in general.

Third, we prove that \( \Omega(x) \subset \text{cl} \mathcal{Z} \) where \( \mathcal{Z} := \{ x \in V \mid 0 \in \mathcal{L}_{\Pi_x} \Psi(x) \} \). Since \( \Omega(x) \) is weakly invariant, for very \( \hat{x} \in \Omega(x) \) there exists as solution \( x' \) to (5.1) with \( x'(0) = \hat{x} \) and \( x' \in \Omega(x) \) for all \( t \in [0, \infty) \). Since \( \Psi \circ x' = c \) it follows that \( \frac{d}{dt} (\Psi \circ x')(t) = 0 \) for all \( t \) and therefore \( D_{x'} \Psi(\dot{x}(t)) = 0 \) for almost all \( t \). This implies that \( 0 \in \mathcal{L}_{\Pi_{x'}} \Psi(x(t)) \) and therefore \( x'(t) \in \mathcal{Z} \) for almost all \( t \). Taking a sequence \( t_k \to 0 \) such that \( x'(t_k) \in \mathcal{Z} \) and hence \( x'(t_k) \to \hat{x} \) shows that \( \hat{x} \in \text{cl} \mathcal{Z} \).

Finally, recall that \( \Omega(x) \) is weakly invariant for every solution \( x \) of (5.1), and \( x \) converges to \( \Omega(x) \) since \( x \) is bounded. Hence, every solution converges to the union of all \( \omega \)-limit sets, and hence to the largest weakly invariant subset of \( \text{cl} \mathcal{Z} \). \( \square \)

**Remark 5.3.** The function \( \Psi \) needs to be defined a neighborhood of \( \mathcal{X} \) solely to guarantee that its derivative is well-defined everywhere on \( \mathcal{X} \). For convenience, we thus depart slightly from our principle that projected dynamical systems need only be defined on the feasible set \( \mathcal{X} \). This minor limitation can be avoided by resorting to more general differentiability concepts, e.g., along the same lines as in [7]. \( \blacksquare \)

### 5.2. Stability of Projected Gradient Descent

We turn to the specific case of projected gradient descent. Given a \( C^1 \) potential function \( \Psi : V \to \mathbb{R} \) defined on an open set \( V \), we define the gradient of \( \Psi \) at \( x \in V \) with respect to a metric \( g \) as the unique element \( \text{grad}_g \Psi(x) \in T_x \mathbb{R}^n \) that satisfies

\[
\langle \text{grad}_g \Psi(x), w \rangle_{g(x)} = D_x \Psi(w) \quad \forall w \in T_x \mathbb{R}^n.
\]

In matrix notation we may equivalently write \( \text{grad}_g \Psi(x) = G^{-1}(x) \nabla \Psi(x)^T \).

Hence, in the following we consider *projected gradient systems* of the form

\[
\dot{x} \in \Pi_x \left( - \text{grad}_g \Psi \right)(x), \quad x(0) = x_0 \in \mathcal{X}.
\]

Such systems serve to find local solutions to the optimization problem

\[ \minimize \Psi(x) \text{ subject to } x \in \mathcal{X}. \]

It is reasonable (but important to note) that in general the metric that defines the gradient has to be the same metric that defines the projection.

We use Theorem 5.2 to derive the following stability result for trajectories of (5.2).
Proposition 5.4. Consider $\mathcal{X} \subset \mathbb{R}^n$, a metric $g$ defined on $\mathcal{X}$, and a $C^1$ function $\Psi : V \to \mathbb{R}$ defined on a neighborhood $V$ of $\mathcal{X}$ such that for every $\ell \in \mathbb{R}$ the set $S_\ell := \{x | \Psi(x) \leq \ell\} \cap \mathcal{X}$ is compact. Let Assumption 5.1 be satisfied for the system (5.2). Then, every complete Krasovskii solution of (5.2) converges to the set of weak equilibrium points.

Proof. Let $F(x) := K\left[\Pi^g_X(-\nabla_g \Psi)\right](x)$. In order to apply Theorem 5.2, we first need to show that $\max L_F \Psi(x) \leq 0$ for all $x \in \mathcal{X}$. For this, we first note that for every $a \in L_F \Psi$, we have by definition of the gradient that

$$a = D_x \Psi(w) = \langle \nabla_x \Psi(x), w \rangle_{g(x)} \quad \text{for some } w \in K\left[\Pi^g_X(-\nabla_g \Psi)\right](x).$$

Using Lemma 4.6, we have for any $w \in K\left[\Pi^g_X(-\nabla_g \Psi)\right](x)$ that

$$D_x \Psi(w) = \langle \nabla_x \Psi(x), w \rangle_{g(x)} = -\langle \nabla_x \Psi(x), w \rangle_{g(x)} \leq -\|w\|^2_{g(x)} \leq 0,$$ (5.3)

and consequently $\max L_F \Psi(x) \leq 0$.

Finally, we need to show that $0 \in L_F \Psi(x)$ implies that $0 \in F(x)$, and therefore $x$ is a weak equilibrium point. For this, note that according to (5.3) $0 \in L_F \Psi(x)$ is equivalent to $\langle \nabla_x \Psi(x), w \rangle_{g(x)} = 0$ for some $w \in F(x)$. Using Lemma 4.6 this implies that either $\nabla_x \Psi(x) = 0$ or $w = 0$. Both imply that $0 \in F(x)$. Finally, from $0 \in F(x)$ it follows that $x$ is a weak equilibrium since the constant trajectory starting at $x$ is a solution to $x \in F(x)$.

It is not a priori clear whether equilibria of (5.2) are minimizers of $\Psi$ in $\mathcal{X}$. Hence, the following result connects the two concepts.

Theorem 5.5 (stability of minimizers for projected gradient flows). Let $\mathcal{X}, g$ and $\Psi$ be defined as in Proposition 5.4 and let Assumption 5.1 be satisfied. In addition, assume that $\mathcal{X}$ has a non-empty interior. Then, the following statements hold:

(i) If $\hat{x} \in \mathcal{X}$ is a strongly asymptotically stable equilibrium of (5.2), then it is a strict local minimum of $\Psi$ on $\mathcal{X}$.

(ii) If $\hat{x} \in \mathcal{X}$ is a strict local minimum of $\Psi$ on $\mathcal{X}$, then it is a strongly stable equilibrium (5.2).

It may seem plausible that strict minimizers are strongly asymptotically stable. This, however, is not true in general (even in the unconstrained case) as the counterexample in [1] shows. Similarly, minimizers are not guaranteed to be stable and stable equilibria are not in general minimizers. This can only be guaranteed under additional assumptions, e.g., minimizers being isolated [26] or $\Psi$ being analytic (in the unconstrained case [1]).

Proof. To show (i), let $V \subset \mathcal{X}$ be a neighborhood of $\hat{x}$ such any solution $x(t)$ of (5.2) with $x_0 \in V$ converges to $\hat{x}$. Since $\Psi$ is $C^1$ and $x$ is absolutely continuous, $\Psi \circ x$ is absolutely continuous, and we may write

$$\lim_{t \to +\infty} (\Psi \circ x)(t) = \Psi(\hat{x}) = \Psi(x_0) + \int_{0}^{+\infty} D_x \Psi(x(t)) \, dt.$$ 

Since $D_x \Psi(x(t)) \leq 0$ almost everywhere, it follows that $\int_{0}^{+\infty} L_F \Psi(x(t)) \leq 0$ and hence $\Psi(\hat{x}) \leq \Psi(x(t)) \leq \Psi(x_0)$ for all $t \geq 0$. Since this reasoning applies to all $x_0$ in the region of attraction of $\hat{x}$, it follows that $\hat{x}$ is a local minimizer of $\Psi$.

To see that $\hat{x}$ is a strict minimizer, assume for the sake of contradiction that for some $\bar{x}$ in the region of attraction $U$ of $\hat{x}$ it holds that $\Psi(\bar{x}) \leq \Psi(\hat{x})$. Every solution
\(y(t)\) to (5.2) with \(y(0) = \bar{x}\) nevertheless converges to \(\hat{x}\) by assumption. Therefore, it must hold that \(\int_0^{+\infty} D_y \Psi(\dot{y}(t)) = 0\) and since \(D_y \Psi(\dot{y}(t)) \leq 0\), it follows that \(D_y \Psi(\dot{y}(t)) = 0\) for almost all \(t \geq 0\). But as a consequence of Proposition 5.4, all points \(x\) with \(0 \in L_F \Psi(x)\) are weak equilibrium points, this holds in particular \(\bar{x}\). Consequently \(\hat{x}\) cannot be strongly asymptotically stable in the neighborhood \(U\).

For (ii) note that since \(\mathcal{X}\) has non-empty interior, every (relative) neighborhood of a point \(x \in \mathcal{X}\) has non-empty interior. Hence, consider a neighborhood \(\bar{U} \subset \mathcal{X}\) of \(\hat{x}\), and let \(U \subseteq \bar{U}\) be a compact neighborhood of \(\hat{x}\) in which \(\hat{x}\) is a strict minimizer. Since \(\mathcal{X}\) has non-empty interior, it follows that \(U\) has non-empty interior. Next, we construct a neighborhood \(V \subset U\) such that all trajectories starting in \(V\) remain in \(U\).

Let \(\alpha\) be such that \(\Psi(\hat{x}) < \alpha < \min_{x \in \partial U} \Psi(x)\) where \(\partial U\) is the boundary of \(U\). Define \(V := \{x \in U \mid \Psi(x) \leq \alpha\} \subset U\) which has a non-empty interior because \(\Psi(\hat{x}) < \alpha\). Since for any trajectory, we have \(D_x \Psi(\dot{x}(\tau)) \leq 0\) we conclude that \(V\) is strongly invariant and consequently remains in \(U\), thus establishing strong stability.\[\]

**Example 5.6 (Constrained Newton Flow).** Let \(\mathcal{X} \subset \mathbb{R}^n\) be closed, and let \(\Psi : \mathbb{R}^n \to \mathbb{R}\) be strongly convex and globally Lipschitz continuous and twice differentiable. In particular, the Hessian of \(\Psi\) (denoted by \(\nabla^2 \Psi\)) is continuous and has lower and upper bounded eigenvalues. Hence, we may use \(\nabla^2 \Psi\) to define the weakly bounded metric \(\langle u, v \rangle_{\psi(x)} := u^T \nabla^2 \Psi(x) v\) for \(u, v \in T_x \mathbb{R}^n\). Hence, the projected gradient flow

\[
\dot{x} \in \Pi^\psi_{\mathcal{X}} (-\nabla \psi(x), \quad x(0) = x_0 \in \mathcal{X}
\]

where \(\nabla \psi(x) = (\nabla^2 \Psi(x))^{-1} \nabla \Psi(x)^T\) is a constrained form of a Newton flow, i.e., the continuous-time limit of the well-known Newton method for optimization.

**6. Equivalence of Krasovskii and Carathéodory Solutions.** In this section we study the relation between Carathéodory and Krasovskii solutions. In particular, we show that the solutions are equivalent if the metric is continuous and the feasible domain is Clarke regular, thus recovering (for the Euclidean metric) known existence conditions for Carathéodory solutions. Further, we establish the connection to related work [5,6,17] and highlight the relation between projected gradient flows, as defined in the previous section, and continuous-time subgradient flows for Clarke regular sets [13,18].

**Definition 6.1.** Consider a set \(\mathcal{X} \subset \mathbb{R}^n\), a metric \(g\) and a vector field \(f\), both defined on \(\mathcal{X}\). The sets of Carathéodory and Krasovskii solutions of (3.2) with initial condition \(x_0 \in \mathcal{X}\) are respectively given by

\[
S_C(x_0) := \{x \mid x : [0, T) \to \mathcal{X}, T > 0, x(0) = x_0, \dot{x}(t) \in \Pi^g_{\mathcal{X}} f(x(t)) \text{ a.e.}\}
\]

\[
S_K(x_0) := \{x \mid x : [0, T) \to \mathcal{X}, T > 0, x(0) = x_0, \dot{x}(t) \in K[\Pi^g_{\mathcal{X}} f](x(t)) \text{ a.e.}\}
\]

where a.e. means almost everywhere and \(C^A\) denotes absolutely continuous functions.

Since \(\Pi^g_{\mathcal{X}} f(x) \subset \Pi^g_{\mathcal{X}} f(x)\), it is clear that every Carathéodory solution of (3.2) is also a Krasovskii solution, i.e., \(S_C(x_0) \subset S_K(x_0)\) for all \(x_0 \in \mathcal{X}\). A pointwise condition for the equivalence of the solution sets is given as follows:

**Lemma 6.2.** Given any set \(\mathcal{X}\), metric \(g\) and vector field \(f\), if \(K[\Pi^g_{\mathcal{X}} f](x) \cap T_x \mathcal{X} = \Pi^g_{\mathcal{X}} f(x)\) holds for all \(x \in \mathcal{X}\), then \(S_C(x_0) = S_K(x_0)\) for all \(x_0 \in \mathcal{X}\).

**Proof.** Since \(S_C(x_0) \subset S_K(x_0)\), we only need to consider \(x \in S_K(x_0)\), we only need to consider \(x \in S_K(x_0)\) and show that \(x \in S_C(x_0)\). By Lemma A.1, \(\dot{x}(t) \in T_{x(t)} \mathcal{X}\) holds for \(x(t)\) almost everywhere.
Consequently, \( \dot{x}(t) \in K[\Pi^g_X f](x(t)) \cap T_x X \) almost everywhere, and therefore, by assumption, \( \dot{x}(t) \in \Pi^g_X f(x(t)) \).

The proof of the next result follows ideas from [17]. The requirement that \( g \) and \( f \) need to be continuous deserves particular attention.

**Theorem 6.3** (equivalence of solution sets). If \( X \) is Clarke regular, \( g \) is a continuous metric on \( X \), and \( f \) is continuous on \( X \), then \( S_C(x_0) = S_K(x_0) \) for all \( x_0 \in X \).

**Proof.** It suffices to show that under the proposed assumptions Lemma 6.2 is applicable. By definition of \( \Pi^g_X f(x) \) we have \( \Pi^g_X f(x) \subset K[\Pi^g_X f](x) \cap T_x X \). For the converse, let \( v \in K[\Pi^g_X f](x) \cap T_x X \). By Lemma 4.6, \( v = f(\dot{x}) - \dot{\eta} \) for some \( \dot{\eta} \in N^g_x X \) and \( \|v\|^2_{g(x)} \leq \langle v, f(\dot{x}) \rangle_{g(x)} \). Since \( \langle v, \dot{\eta} \rangle_{g(x)} \leq 0 \) for all \( \dot{\eta} \in N^g_x X \) we have

\[
\|v\|^2_{g(x)} \leq \langle v, f(\dot{x}) \rangle_{g(x)} - \langle v, \dot{\eta} \rangle_{g(x)} \leq \|v\|_{g(x)} \|f(\dot{x}) - \eta\|_{g(x)} \quad \forall \dot{\eta} \in N^g_x X,
\]

where the second inequality is due to Cauchy-Schwarz, and therefore \( \|v - \dot{\eta}\|_{g(x)} \leq \|f(\dot{x}) - \eta\|_{g(x)} \) holds for all \( \dot{\eta} \in N^g_x X \). However, according to Lemma 4.5 the fact that \( \dot{\eta} = \arg \min_{\dot{\eta} \in N^g_x X} \|f(\dot{x}) - \eta\|_{g(x)} \) is equivalent to \( v \in \Pi^g_X f(x) \).

Note that Examples 3.4 and 4.4 show a case where the conclusion of Theorem 6.3 fails to hold because \( X \) is not Clarke regular at the origin. Hence, our sufficient characterization in terms of Clarke regularity is also a sharp one.

Theorem 6.3 also serves as an existence result of Carathéodory solutions, that recovers the conditions derived in [17], but for a general metric.

**Corollary 6.4** (Existence of Carathéodory solutions). If \( X \) is Clarke regular, and \( g \) and \( f \) are continuous on \( X \), then there exists a Carathéodory solution \( x : [0, T] \to X \) of (3.2) with \( x(0) = x_0 \) for some \( T > 0 \), and every \( x_0 \in X \).

Uniqueness, however, requires additional assumptions as will be shown in Section 7. In particular, uniquess of the projection \( \Pi^g_X f(x) \) does not imply uniquess of the trajectory (see forthcoming Remark 7.9).

## 6.1. Related work, alternative formulations, and subgradient flows.

With the statements of Section 6 at hand, we discuss their connection to related literature, and in particular, we point out an important connection with subgradients [13,46] that is well-known in the context of convex analysis, but also extends to Clarke regular domains.

As discussed in the introduction, projected dynamical system have been studied from different perspectives and with various applications in mind. In particular, a number of alternative, but equivalent formulations do exist [10,28], but none considers the case of a variable metric. In the following, we discuss a well-established formulation [5,6,17] that has a number of insightful properties.

Namely, under Clarke regularity of the feasible set \( X \) we may define an alternative differential inclusion given by the initial value problem

\[
(6.1) \quad \dot{x} \in f(x) - N^g_x X, \quad x(0) = x_0 \in X
\]

and define the solution set as

\[
S_N(x_0) := \{ x \mid x : [0, T] \to X, T > 0, x \in C^A, x(0) = x_0, \dot{x} \in f(x) - N^g_x X \text{ a.e.} \}.
\]

The next result is an adaptation of [17, Thm 2.3] to arbitrary metrics. We provide a self-contained proof for completeness.
COROLLARY 6.5. Consider a Clarke regular set $\mathcal{X} \subset \mathbb{R}^n$, a continuous vector field $f$, and a continuous metric $g$, both defined on $\mathcal{X}$. Then, $S_N(x_0) = S_C(x_0)$ holds for systems of the form (3.2) and (6.1), and for all $x_0 \in \mathcal{X}$.

In short, any solution to (6.1) is a Carathéodory solution of (3.2) and vice versa. However, Corollary 6.5 makes no statement about existence of solutions. In fact, the non-compactness of $N_x^g\mathcal{X}$ prevents us from applying the same viability result as for Theorem 4.2.

Proof. We first note that $S_C(x_0) \subset S_N(x_0)$ since $\Pi_x^g f(x) \subset f(x) - N_x^g\mathcal{X}$ for all $x \in \mathcal{X}$ by virtue of Lemma 4.6 and since $\mathcal{X}$ is Clarke regular. Conversely, let $x \in S_N(x_0)$ be defined for $t \in [0, T)$ for $T > 0$. Then for almost all $t$, we have $\dot{x}(t) \in f(x(t)) - N_{x(t)}^g\mathcal{X}$ and $\dot{x}(t) \in T_{x(t)}\mathcal{X} \cap -T_{x(t)}\mathcal{X}$ by Lemma A.1. Thus, for $\dot{x}(t) = f(x(t)) - \eta(x(t))$ with $\eta(x(t)) \in N_{x(t)}^g\mathcal{X}$ it must hold that

$$\langle f(x(t)) - \eta(x(t)), \eta(x(t)) \rangle_{g(x(t))} \leq 0 \quad \text{for all } t \in [0, T).$$

Consequently, $\langle f(x(t)) - \eta(x(t)), \eta(x(t)) \rangle_{g(x(t))} = 0$, and using Lemma 4.5 it follows that $\dot{x}(t) = \Pi_x^g f(x(t))$.

Remark 6.6. Defining inclusions of the form (6.1) for a set $\mathcal{X}$ that is not Clarke regular is possible but technical since one would need to distinguish between different types of normal cones (Remark 2.9). Furthermore, depending on the choice of normal cone the resulting set of solutions can be overly relaxed or too restrictive.

Remark 6.7. Using (ii) in Lemma 4.5 it follows that whenever $\dot{x}$ exists, we have $\dot{x} = \min_{v \in f(x) - N_x^g\mathcal{X}} \|v\|_{g(x)}$. When $g$ is the Euclidean metric, this minimum norm property gives rise to so-called slow solutions of (6.1) [5, Chap 10.1]. For a general metric, the definition of a slow solution generalizes accordingly. However, the property of being “slow” depends on the metric.

Assuming that $f$ is the gradient field of a potential function and $\mathcal{X}$ is Clarke regular, we can show that the connection between projected gradients and subgradients which is well-known for convex functions (and lesser known for regular functions [13, 18]) generalizes to a variable metric. For this, recall that $\Psi : V \rightarrow \mathbb{R}$ where $V \subset \mathbb{R}^n$ is open and $\mathbb{R}^\infty := \mathbb{R} \cup \{\infty\}$ is (subdifferentially) regular if its epigraph $\mathrm{epi} \Psi := \{(x, y) \mid x \in V, y \geq \Psi(x)\}$ is Clarke regular.

DEFINITION 6.8. Let $g$ be a metric on $V$ where $V \subset \mathbb{R}^n$ is open, and let $\Psi : V \rightarrow \mathbb{R}$ be a regular function. A vector $v$ is a subgradient of $\Psi$ with respect to $g$ at $x$, denoted by $v \in \partial\Psi(x)$, if

$$\liminf_{y \to x} \frac{\Psi(y) - \Psi(x) - \langle v, y-x \rangle_{g(x)}}{\|y-x\|} \geq 0.$$ 

In particular, if $\Psi$ is differentiable at $x$, then $\partial\Psi(x) = \{\mathrm{grad}_g \Psi(x)\}$. Further, if $\mathcal{X} \subset V$ is Clarke regular and $I_\mathcal{X} : V \rightarrow \mathbb{R}$ denotes its indicator function, then $\partial I_\mathcal{X}(x) = N_x^g\mathcal{X}$.

The next result is a direct combination of [46, Ex 8.14] and [46, Cor 10.9].

PROPOSITION 6.9. Let $\hat{\Psi} := \Psi + I_\mathcal{X}$ where $\Psi : V \rightarrow \mathbb{R}$ is a $C^1$ function and $I_\mathcal{X}$ is the indicator function of a Clarke regular set $\mathcal{X} \subset V$ where $V \subset \mathbb{R}^n$ is open. Then, for all $x \in \mathcal{X}$ one has

$$\partial\hat{\Psi}(x) = \mathrm{grad}_g \Psi(x) + N_x^g\mathcal{X}.$$
It follows immediately from Corollary 6.5 that under the appropriate assumptions trajectories of projected gradient flows are also solutions to subgradient flows.

**Corollary 6.10 (equivalence with subgradient flows).** Let \( X \) be Clarke regular, let \( g \) be a continuous metric on \( X \), and let \( \Psi \) be a \( C^1 \) potential function on an open neighborhood of \( X \). Then, for any \( x_0 \in X \) there exists a Carathéodory solution \( x : [0,T) \to X \) to the subgradient flow

\[
\dot{x} \in -\partial(\Psi + I_X)(x), \quad x(0) \in X.
\]

Furthermore, \( x \) is a solution if and only if it is a Carathéodory (and Krasovskii) solution to the projected gradient descent (5.2).

**7. Prox-regularity and Uniqueness of Solutions.** Next, we introduce a generalized definition of prox-regular sets on non-Euclidean spaces with a variable metric and show their significance for the uniqueness of solutions of projected dynamical systems. In the Euclidean setting prox-regularity is well-known to be a sufficient condition on the feasible domain \( X \) for uniqueness [17].

The key issue of this section is thus to generalize the definition of prox-regular sets and identify the requirements that lead to unique solutions. By doing so, we also show that prox-regularity of a set is independent of the choice of metric. In the subsequent section this allows us to state that prox-regularity is preserved under \( C^{1,1} \) coordinate transformations and hence well-defined on \( C^{1,1} \) manifolds.

**7.1. Prox-regularity on non-Euclidean spaces.** For illustration, we first recall and discuss the definition of prox-regularity in Euclidean space. Our treatment of the topic is deliberately kept limited. For a more general overview see [3, 45].

**Definition 7.1.** A Clarke regular set \( X \subset \mathbb{R}^n \) is prox-regular at \( x \in X \) if there is \( L > 0 \) such that for every \( z, y \in X \) in a neighborhood of \( x \) and \( \eta \in N_y X \) we have

\[
\langle \eta, z - y \rangle \leq L\|\eta\|\|z - y\|^2.
\]

The set \( X \) is prox-regular if it is prox-regular at every \( x \in X \).

One of the key features of a prox-regular set \( X \) is that for every point in a neighborhood of \( X \) there exists a unique projection on the set [3, Def 2.1, Thm 2.2].

**Example 7.2 (Prox-regularity in Euclidean spaces).** Consider the parametric set

\[
X_\alpha := \{(x_1, x_2) \mid |x_2| \geq \max\{0, x_1\}^\alpha\}
\]

where \( 0 < \alpha < 1 \) and which is illustrated in Figure 4. For \( \alpha \leq 0.5 \) the set is prox-regular everywhere. In particular for the origin, a ball with non-zero radius can be placed tangentially such that it only intersects the set at 0. For \( \alpha > 0.5 \) on the other hand the set is not prox-regular at the origin. In fact, all points on the positive axis have a non-unique projection on \( X_\alpha \) as illustrated in Figure 4c. ■

Definition 7.1 cannot be directly generalized to non-Euclidean spaces since it requires the distance \( \|y - x\| \) between two points in \( X \). Hence, in [31] prox-regularity is defined on smooth (i.e., \( C^\infty \)) Riemannian manifolds resorting to geodesic distances. For our purposes we can avoid the notational complexity of Riemannian geometry, yet preserve a higher degree of generality. Thus, we introduce the following definitions.

**Definition 7.3.** Given a Clarke regular set \( X \subset \mathbb{R}^n \) and a metric \( g \), a normal vector \( \eta \in N_y^x X \) at \( x \in X \) is \( L \)-proximal with respect to \( g \) for \( L \geq 0 \) if for all \( y \in X \)
in a neighborhood of \( x \) we have

\[
(\eta, y - x)_{g(x)} \leq L\|\eta\|_{g(x)}\|y - x\|^2_{g(x)}.
\]

The cone of all \( L \)-proximal normal vectors at \( x \) with respect to \( g \) is denoted by \( \mathcal{N}_x^g L \mathcal{X} \).

A crucial detail in (7.3) is the fact that \( g \) is evaluated at \( x \) and is used as an inner product on \( \mathbb{R}^n \) (which is a slight abuse of notation). In other words, we exploit the canonical isomorphism between \( \mathbb{R}^n \) and \( T_x\mathbb{R}^n \) to use \( g(x) \) as an inner product on \( \mathbb{R}^n \).

**Definition 7.4.** A Clarke regular set \( \mathcal{X} \subset \mathbb{R}^n \) with a metric \( g \) is \( L \)-prox-regular at \( x \in \mathcal{X} \) with respect to \( g \) if \( \mathcal{N}_x^g L \mathcal{X} = \mathcal{N}_x^g \mathcal{X} \) for all \( y \in \mathcal{X} \) in a neighborhood of \( x \). The set \( \mathcal{X} \) is \( \mathcal{L} \)-prox-regular with respect to \( g \) if for every \( x \in \mathcal{X} \) there exists \( L > 0 \) such that \( \mathcal{X} \) is \( \mathcal{L} \)-prox-regular at \( x \) with respect to \( g \).

Note that if \( g \) is the Euclidean metric, Definition 7.4 reduces to Definition 7.1. The following result shows that prox-regularity is in fact independent of the metric. This is the first step towards a coordinate-free definition of prox-regularity.

**Proposition 7.5.** Let \( \mathcal{X} \subset \mathbb{R}^n \) be Clarke regular. If \( \mathcal{X} \) is \( \mathcal{L} \)-prox-regular with respect to a \( C^0 \) metric \( g \), then it is \( \mathcal{L} \)-prox-regular with respect to any other \( C^0 \) metric.

In particular if \( \mathcal{X} \) is prox-regular with respect to the Euclidean metric, i.e., according to Definition 7.1, then it is prox-regular in any other continuous metric on \( \mathbb{R}^n \). For the proof of Proposition 7.5 we require the following lemma.

**Lemma 7.6.** Let \( \mathcal{X} \subset \mathbb{R}^n \) be Clarke regular and consider to metrics \( g, g' \) defined on \( \mathcal{X} \). If for \( x \in \mathcal{X} \) there is \( L > 0 \) such that \( \mathcal{N}_x^g \mathcal{X} = \mathcal{N}_x^{g'} \mathcal{X} \) then \( \mathcal{N}_x^{g'} \mathcal{X} \) holds for \( L' \geq \kappa_{g(x)} \kappa_{g'(x)} L \).

**Proof.** First note that for every \( x \in \mathcal{X} \) the two metrics \( g \) and \( g' \) induce a bijection between \( \mathcal{N}_x^g \mathcal{X} \) and \( \mathcal{N}_x^{g'} \mathcal{X} \). Namely, we define \( q : T_x\mathbb{R}^n \to T_x\mathbb{R}^n \) as the unique element \( q(v) \) that satisfies by \( \langle v, w \rangle_{g(x)} = \langle q(v), w \rangle_{g'(x)} \) for all \( w \in T_x\mathbb{R}^n \). To clarify, in matrix notation we can write \( v^T G(x) w = q(v)^T G'(x) w \) and since \( G(x) \), \( G'(x) \) are symmetric positive definite we have \( q(v) := G'(x)^{-1} G(x) v \). It follows that if \( \eta \in \mathcal{N}_x^g \mathcal{X} \) (hence, by definition \( \langle \eta, w \rangle_{g(x)} \leq 0 \) for all \( w \in T_x\mathbb{R}^n \)), then \( q(\eta) \in \mathcal{N}_x^{g'} \mathcal{X} \). Furthermore, omitting the argument \( x \), we have \( \|q(\eta)\|_{g'} = \eta^T G G^{-1} G \eta \geq 1/\lambda^{\max}_{g'} \|G\eta\| \) and \( \|\eta\|_g = \eta^T G G^{-1} G \eta \leq 1/\lambda^{\min}_{g} \|G\eta\| \), and therefore \( \|q(\eta)\|_{g'(x)} \geq \lambda^{\min}_{g(x)}/\lambda^{\max}_{g(x)} \|\eta\|_{g(x)} \).

Hence, let \( \eta \in \mathcal{N}_x^g \mathcal{X} \setminus \{0\} \) be a \( L \)-proximal normal vector, then

\[
\left\langle \frac{q(\eta)}{\|q(\eta)\|_{g'(x)}}, y - x \right\rangle_{g'(x)} \leq \frac{\lambda^{\min}_{g(x)}}{\lambda^{\max}_{g(x)}} \left\langle \frac{\eta}{\|\eta\|_{g(x)}}, y - x \right\rangle_{g(x)} \leq \frac{\lambda^{\min}_{g(x)}}{\lambda^{\max}_{g(x)}} L \|y - x\|^2_{g(x)}.
\]
Finally, using the equivalence of norms, we have
\[
\frac{\kappa_{g}(x)}{\lambda_{g}(x)} L \| y - x \|_{g(x)}^{2} \leq \frac{\kappa_{g}(x)}{\lambda_{g^\prime}(x)} L \| y - x \|_{g^\prime(x)}^{2} \leq L' \| y - x \|_{g^\prime(x)}^{2},
\]
where \( L' \geq \kappa_{g}(x) \kappa_{g^\prime}(x) L \). Thus, we have shown that if \( v \in N_{x}^{N_{g} L} \mathcal{X} = N_{x}^{g} \mathcal{X} \) then \( q(v) \in N_{x}^{N_{g} L} \mathcal{X} = N_{x}^{g} \mathcal{X} \) which completes the proof. \( \square \)

**Proof of Proposition 7.5.** Since \( g \) and \( g' \) are continuous it follows that \( \kappa_{g}(x) \) and \( \kappa_{g^\prime}(x) \) are continuous in \( x \) and therefore locally bounded. Given any \( x \in \mathcal{X} \) and using the pointwise result in Lemma 7.6, we can choose \( L' > 0 \) such that (7.4) is satisfied for all \( y \in \mathcal{X} \) in a neighborhood of \( x \). \( \square \)

We conclude this section by showing that feasible domains defined by \( C^{1,1} \) constraint functions are prox-regular under the usual constraint qualifications.

**Example 7.7** (prox-regularity of constraint-defined sets). As in Examples 2.4 and 2.10 let \( h : \mathbb{R}^{n} \to \mathbb{R}^{m} \) be \( C^{1} \) and \( \nabla h(x) \) have full rank for all \( x \) and consider \( \mathcal{X} := \{ x \mid h(x) \leq 0 \} \). If in addition, \( h \) is a \( C^{1,1} \) map, then \( \mathcal{X} := \{ x \mid h(x) \leq 0 \} \) is prox-regular with respect to any \( C^{0} \) metric \( g \) on \( \mathbb{R}^{n} \).

To see this, we consider the Euclidean case without loss of generality as a consequence of Proposition 7.5. We first analyze the sets \( \mathcal{X}_{i} := \{ x \mid h_{i}(x) \leq 0 \} \) and then show prox-regularity of their intersection. For this, we only need to consider points \( x \in \partial \mathcal{X}_{i} \) on the boundary of \( \mathcal{X}_{i} \) since for all \( x \notin \partial \mathcal{X}_{i} \) we have \( N_{x} \mathcal{X}_{i} = \{ 0 \} \) and prox-regularity is trivially satisfied. Hence, using the Descent Lemma A.2, for all \( z, y \in \mathbb{R}^{n} \) in a neighborhood of \( x \) and all \( i = 1, \ldots, m \) there exists \( L_{i} > 0 \) such that
\[
-L_{i} \| z - y \|^{2} \leq h_{i}(z) - h_{i}(y) - \langle \nabla h_{i}^{T}(z), z - y \rangle.
\]
In particular, for \( z \in \mathcal{X}_{i} \) (i.e., \( h_{i}(z) \leq 0 \)) and \( y \in \partial \mathcal{X}_{i} \) (i.e., \( h_{i}(y) = 0 \)) in a neighborhood of \( x \) we have
\[
\langle \nabla h_{i}^{T}(y), z - y \rangle \leq h_{i}(z) + L_{i} \| z - y \|^{2} \leq L_{i} \| z - y \|^{2}.
\]

For the set \( \mathcal{X} = \bigcap_{i=1}^{m} \mathcal{X}_{i} \) recall from Example 2.10 that for \( x \in \mathcal{X} \) we have
\[
N_{x} \mathcal{X} = \left\{ \eta \left| \eta = \sum_{i \in I(x)} \alpha_{i} \nabla h_{i}^{T}(x) \right. \right\} \alpha_{i} \geq 0 \right\}.
\]

Consider \( z \in \mathcal{X} \) and \( y \in \partial \mathcal{X} \) in a small enough neighborhood of \( x \). Note that \( y \in \partial \mathcal{X} \) implies that \( y \in \partial \mathcal{X}_{i} \) for all \( i \in I(y) \). Using (7.5), for all \( \eta \in N_{y} \mathcal{X} \) with \( \eta = \sum_{i \in I(y)} \alpha_{i} \nabla h_{i}^{T}(y)/\| \nabla h_{i}(y) \| \) we have
\[
\langle \eta, z - y \rangle = \left( \sum_{i \in I(y)} \alpha_{i} \nabla h_{i}(y)^{T}, z - y \right) \leq \left( \sum_{i \in I(y)} \alpha_{i} L_{i} \right) \| z - y \|^{2}
\]
and therefore \( \langle \eta, z - y \rangle \leq L(y) \| \eta \| \| z - y \|^{2} \), where
\[
L(y) := \frac{\sum_{i \in I(y)} \alpha_{i} L_{i}}{\| \eta \|} = \frac{\sum_{i \in I(y)} \alpha_{i} L_{i}}{\| \sum_{i \in I(y)} \alpha_{i} \nabla h_{i}(y) \|} \leq \max_{i \in I(y)} \frac{\alpha_{i} L_{i}}{\| \nabla h_{i}(y) \|} = \max_{i=1, \ldots, m} \frac{L_{i}}{\| \nabla h_{i}(y) \|}.
\]
The first inequality can be shown by taking the square and proceeding by induction. Since the final bound is with respect to all \( h_{i} \), it is continuous in \( y \) in a neighborhood of \( x \). Consequently, we can choose \( \bar{L} \) such that \( \bar{L} \geq L(y) \) for all \( y \in \mathcal{X} \) in a neighborhood of \( x \), and therefore \( \langle \eta, z - y \rangle \leq \bar{L} \| \eta \| \| z - y \|^{2} \) for \( z \in \mathcal{X} \) in a neighborhood of \( y \). This proves \( L \)-prox-regularity at \( x \) and prox-regularity follows accordingly. \( \square \)
7.2. Uniqueness of solutions to projected dynamical systems. Before formulating our main uniqueness result, we present an example that illustrates the impact of prox-regularity on the uniqueness of solutions.

Example 7.8 (prox-regularity and uniqueness of solutions). We consider the set $X_\alpha := \{(x_1, x_2) \mid |x_2| \geq \max\{0, x_1\}^{\alpha}\}$ for $0 < \alpha < 1$, as in Example 7.2. We study how the value of $\alpha$ affects the uniqueness of solutions of the projected dynamical system defined by the uniform “horizontal” vector field $f(x) = (1, 0)$ for all $x \in X$ and the initial condition $x(0) = 0$ as illustrated in Figure 5.

Since $X_\alpha$ is Clarke regular and closed, since the vector field is uniform, and since we use the Euclidean metric, the existence of Krasovskii solutions and the equivalence of Carathéodory solutions is guaranteed for $t \to \infty$ by Corollary 4.3 and Theorem 6.3, respectively. The prox-regularity of $X_\alpha$ at the origin is however only guaranteed for $0 < \alpha \leq \frac{1}{2}$ (Example 7.2).

A formal analysis reveals that for $0 < \alpha \leq \frac{1}{2}$ the origin is a strong equilibrium, i.e., the constant solution $x(t) = 0$ is the unique solution to the projected dynamical system. For $\frac{1}{2} < \alpha < 1$, however, the origin is only a weak equilibrium point. Namely, a solution may remain at 0 for an arbitrary amount of time before leaving 0 on either upper or lower halfplane, and thus uniqueness is not guaranteed.

Remark 7.9. In general, whether $\Pi f(x_0)$ is a singleton is unrelated to the uniqueness of solutions starting from $x_0$. For instance, in Example 7.8, if $\alpha > 0$ multiple solutions exists even though $\Pi f(x)$ is a singleton at $x = 0$. Conversely, Example 4.4 shows that even if $\Pi f(x_0)$ is not unique, the (Krasovskii) solution starting from $x_0$ is unique.

For the proof of uniqueness under prox-regularity, we require the following lemma.

Lemma 7.10. Let $X$ be $L$-prox-regular at $x$ with respect to a $C^{0,1}$ metric $g$. Then, there exist $\bar{L} > 0$ such that for all $y \in X$ in a neighborhood of $x$ and all $\eta \in N^L_y$ with $\|\eta\|_{g(y)} = 1$ we have $\langle \eta, x - y \rangle_{g(x)} \leq \bar{L}\|y - x\|^2_{g(x)}$.

Proof. We know that $\langle \eta, y - x \rangle_{g(y)} \leq L\|y - x\|^2_{g(y)}$ for $y$ close enough to $x$ because $\eta$ is a $L$-proximal normal vector at $y$ with respect to $g$. Furthermore, by the equivalence of norms there exists $L' > 0$ such that $\langle \eta, y - x \rangle_{g(y)} \leq L'\|y - x\|^2_{g(x)}$.

Next, we show that $|\langle \eta, x - y \rangle_{g(y)} - \langle \eta, x - y \rangle_{g(x)}| \leq M\|y - x\|^2_{g(x)}$ for some...
\(M > 0\). Since \(L^2_n\) is a vector space, we may write
\[
\langle \eta, x - y \rangle_{g(y)} - \langle \eta, x - y \rangle_{g(x)} = \langle \eta, x - y \rangle_{g(y) - g(x)}
\]
which is a slight abuse of notation since \(\langle \cdot, \cdot \rangle_{g(y) - g(x)}\) is not necessarily positive definite and therefore not a metric. Nevertheless, any map of the form \((u, v, g) \mapsto \langle u, w \rangle_g\) where \(g \in L^2_n\) is linear in \(u, v\) and in \(g\) (e.g., \((u, v, g) \mapsto \langle u, w \rangle_{\lambda g} = \lambda \langle u, w \rangle_g\) for any \(\lambda \in \mathbb{R}\)). Therefore, there exist \(M', M > 0\) such that
\[
\left| \langle \eta, x - y \rangle_{g(y) - g(x)} \right| \leq M'\|g(y) - g(x)\|L^2_n\|x - y\|_{g(x)} \leq M\|x - y\|_{g(x)},
\]
where \(\| \cdot \|_{L^2_n}\) denotes any norm on the vector space \(L^2_n\), and the second inequality follows directly from the Lipschitz continuity of \(g\). Hence, we can conclude that that
\[
\langle \eta, x - y \rangle_{g(y)} \leq \langle \eta, x - y \rangle_{g(x)} + \|y - x\|_{g(x)} \leq (L' + M)\|y - x\|_{g(x)}.
\]

Next, we can show the following Lipschitz-type property of projected vector fields.

**Proposition 7.11.** Let \(f\) be a \(C^{0,1}\) field on \(\mathcal{X}\). If \(g\) is a \(C^{0,1}\) metric and \(\mathcal{X}\) is prox-regular, then for every \(x \in \mathcal{X}\) there exists \(L > 0\) such that for all \(y \in \mathcal{X}\) in a neighborhood of \(x\) we have
\[
\langle \Pi^0_{\mathcal{X}} f(y) - \Pi^0_{\mathcal{X}} f(x), y - x \rangle_{g(x)} \leq L\|y - x\|^2_{g(x)}.
\]

**Proof.** As a consequence of Lemma 4.5, we can write
\[
(7.6) \quad \langle \Pi^0_{\mathcal{X}} f(y) - \Pi^0_{\mathcal{X}} f(x), y - x \rangle_{g(x)} = \langle f(y) - f(x), y - x \rangle_{g(x)} + \langle \eta_y, x - y \rangle_{g(x)} + \langle \eta_x, y - x \rangle_{g(x)}.
\]
where \(\eta_y \in N_y^g \mathcal{X} = \Pi^0_{\mathcal{X}} f\) and \(\eta_x \in N_x^g \mathcal{X} = \Pi^0_{\mathcal{X}} f\) for some \(L > 0\).

For the first term, we get \(\langle f(y) - f(x), y - x \rangle_{g(x)} \leq \|f(y) - f(x)\|_{g(x)}\|y - x\|_{g(x)}\), by applying Cauchy-Schwarz. Since \(f\) is Lipschitz and using the equivalence of norms there exists \(L_a > 0\) such that \(\|f(y) - f(x)\|_{g(x)} \leq L_a\|y - x\|_{g(x)}\) for all \(y \in \mathcal{X}\) in a neighborhood of \(x\). Thus, we have \(\langle f(y) - f(x), y - x \rangle_{g(x)} \leq L_a\|y - x\|^2_{g(x)}\).

For the second and third term in (7.6) we have
\[
\langle \eta_y, x - y \rangle_{g(x)} \leq \langle \eta_y, x - y \rangle_{g(x)}\|\eta_y\|_{g(y)} \leq \langle \eta_y, x - y \rangle_{g(x)}\|\eta_x\|_{g(x)} \leq L\|y - x\|^2_{g(x)}.
\]
by Lemma 7.10 and the definition of a \(L\)-proximal normal vector, respectively.

By Lemma 4.5 we know that \(\|\eta_y\|_{g(y)} \leq \|f(y)\|_{g(y)}\) and \(\|\eta_x\|_{g(x)} \leq \|f(x)\|_{g(x)}\). Since \(g\) and \(f\) are continuous we can choose \(M > 0\) such that \(\|f(z)\|_{g(z)} \leq M\) for all \(z \in \mathcal{X}\) in a neighborhood of \(x\). Therefore, (7.6) can be bounded by
\[
(7.6) \quad \langle \Pi^0_{\mathcal{X}} f(y) - \Pi^0_{\mathcal{X}} f(x), y - x \rangle_{g(x)} \leq (L_a + L'M + LM)\|y - x\|^2_{g(x)}
\]
which completes the proof. \(\square\)

Hence, we can state our main result on the uniqueness of solutions which complements results in [17] by considering a variable (but non-differentiable) metric and using our general definition of prox-regularity. In this context, uniqueness is understood in the sense that any two solutions are equal on the interval on which they are both defined.
Theorem 7.12 (uniqueness of solutions). Let \( f \) be a \( C^{0,1} \) vector field on \( X \). If \( g \) is a \( C^{0,1} \) metric and \( X \) is prox-regular, then for every \( x_0 \in X \) there exists \( T > 0 \) such that the initial value problem \( \dot{x} \in \Pi_{T}^{x} f(x) \) with \( x(0) = x_0 \) has a unique Carathéodory solution \( x : [0, T) \rightarrow X \) (which is also the unique Krasovskii solution).

Proof of Theorem 7.12. The proof follows standard contraction ideas [21]. Let \( x(t) \) and \( y(t) \) be two solutions solving the same initial value problem \( \dot{x} \in \Pi_{T}^{x} f(x) \) with \( x(0) = x_0 \in X \), both defined on a non-empty interval \( [0, T) \).

Using Proposition 7.11, there exists \( M > 0 \) and a neighborhood \( V \) of \( x_0 \) such that

\[
\frac{d}{dt} \left( \frac{1}{2} \| y(t) - x(t) \|^2_{g(x_0)} \right) = \langle \Pi_{X}^{x} f(y(t)) - \Pi_{X}^{x} f(x(t)), y(t) - x(t) \rangle_{g(x_0)} \\
\leq M \| y(t) - x(t) \|^2_{g(x_0)}
\]

for all \( t \) in some non-empty subinterval \( [0, T') \subset [0, T) \) for which \( x(t) \) and \( y(t) \) remain in \( V \). Next, consider the non-negative, absolutely continuous function \( q : [0, T') \rightarrow \mathbb{R} \) defined as \( q(t) := \frac{1}{2} \| y(t) - x(t) \|^2_{g(x_0)} e^{-2Mt} \). Note that \( q(0) = 0 \). Furthermore, using (7.7) and applying the product rule we have

\[
\frac{d}{dt} q(t) = \langle \Pi_{X}^{x} f(y(t)) - \Pi_{X}^{x} f(x(t)), y(t) - x(t) \rangle_{g(x_0)} - M \| y(t) - x(t) \|^2_{g(x_0)} e^{-2Mt}
\]

and since \( y(0) = x(0) \) it follows that \( \frac{d}{dt} q(t) \leq 0 \) for \( t \geq 0 \). However, since \( q \) is non-negative and absolutely continuous, we conclude that \( x(t) = y(t) \) for all \( t \in [0, T') \) thus finishing the proof of uniqueness.

Combining all the insights so far, we arrive at the following ready-to-use result:

Example 7.13 (Existence and uniqueness on constraint-defined sets). As in Example 7.7 consider a set \( X := \{ x \in \mathbb{R}^n \mid h(x) \leq 0 \} \) where \( h : \mathbb{R}^n \rightarrow \mathbb{R}^m \) is of class \( C^{1,1} \) and has full rank for all \( x \in \mathbb{R}^n \). Further, consider a globally Lipschitz continuous vector field \( f : \mathbb{R}^n \rightarrow \mathbb{R} \). Then, for every \( x_0 \in X \) there exists a unique and complete Carathéodory solution \( x : [0, \infty) \rightarrow X \) to the initial value problem \( \dot{x} = \Pi_{X}^{x} f(x) \) with \( x(0) = x_0 \) where \( g \) is any weakly bounded \( C^{0,1} \) metric on \( X \).

8. Existence and Uniqueness on low-regularity Riemannian Manifolds.

The major appeal of Theorems 4.2, 6.3, and 7.12 is their geometric nature. Namely, as we will show next, their assumptions are preserved by sufficiently regular coordinate transformations which allows us to give a coordinate-free definition of projected dynamical system on manifolds with minimal degree of differentiability.

Recall that for open sets \( V, W \subset \mathbb{R}^n \) a map \( \Phi : V \rightarrow W \) is a \( C^k \) diffeomorphism if it is a \( C^k \) bijection with a \( C^k \) inverse where, for our purposes, \( C^k \) stands for either \( C^1 \) or \( C^{1,1} \). We employ the usual definition of a \( C^k \) manifold as locally Euclidean, second countable Hausdorff space endowed with a \( C^k \) differentiable structure. In particular, for a point \( p \) on a \( n \)-dimensional manifold \( M \) there exists a chart \((U, \phi)\) where \( U \subset M \) is open and \( \phi : U \rightarrow \mathbb{R}^n \) is a homeomorphism onto its image. For any two charts \((U, \phi), (V, \psi)\) for which \( U \cap V \neq \emptyset \), the map \( \phi \circ \psi^{-1} : \psi(U \cap V) \rightarrow \phi(U \cap V) \) is a \( C^k \) diffeomorphism. A \( C^k \) (Riemannian) metric \( g \) is a map that assigns to every point \( p \in M \) an inner product on the tangent space\(^3\) \( T_p M \) such that in local coordinates \((U, \phi)\) the metric \( g(\phi^{-1}(x)) \) is a \( C^k \) metric for \( x \in \phi(U) \) according to Definition 2.5. A vector field defined on \( M \) is locally bounded at \( x \) if it is locally bounded in any

\(^3\)Note that the definition (and hence the notation) of the tangent space \( T_p M \) of a manifold \( M \) is consistent with the definition of the tangent cone \( T_p X \) of an arbitrary set \( X \) [46, Ex 6.8].
local coordinate domain for \( x \). Similarly, a metric is \textit{locally weakly bounded at} \( x \) if its locally weakly bounded in local coordinates. Given a \( C^k \) manifold \( \mathcal{M} \) with \( k \geq 1 \), a curve \( \gamma : [0, T) \to \mathcal{M} \) is \textit{absolutely continuous} if it is absolutely continuous in any chart domain where it is defined.\(^4\)

The next lemma shows that a \( C^1 \) diffeomorphism maps (Clarke) tangent cones to (Clarke) tangent cones. Hence, Clarke regularity is preserved by \( C^1 \) diffeomorphisms.

**Lemma 8.1.** Let \( V, W \subset \mathbb{R}^n \) be open and consider a \( C^1 \) diffeomorphism \( \Phi : V \to W \). Given \( \mathcal{X} \subset \mathbb{R}^n \) and \( \tilde{\mathcal{X}} := \mathcal{X} \cap V \), for every \( x \in \tilde{\mathcal{X}} \) it holds that

\[
\begin{align*}
T_{\Phi(x)} \Phi(\tilde{x}) &= D_x \Phi (T_x \tilde{x}) \\
T_{\Phi(x)}^C \Phi(\tilde{x}) &= D_x \Phi (T_x^C \tilde{x}).
\end{align*}
\]

Hence, \( \Phi(\tilde{x}) \) is Clarke regular at \( \Phi(x) \) if and only if \( \tilde{x} \) is Clarke regular at \( x \in \tilde{\mathcal{X}} \).

**Proof.** We only need to show that \( T_{\Phi(x)} \Phi(\tilde{x}) \subset D_x \Phi (T_x \tilde{x}) \). Since \( \Phi \) is a \( C^1 \) diffeomorphism the other direction follows by applying the same arguments to \( \Phi^{-1} \).

Let \( v \in T_x \tilde{\mathcal{X}} \). Then, by definition there exist \( x_k \to x \) with \( x_k \in \tilde{\mathcal{X}} \) and \( \delta_k \to 0^+ \) such that \( (x_k - x)/\delta_k \to v \). Furthermore, \( \|x_k - x\|/\delta_k \) converges to \( \|v\| \). According to the definition of the derivative of \( \Phi \), for the same sequence \( \{x_k\} \) we have \( \lim_{k \to \infty} \|\Phi(x_k) - \Phi(x) - D_x \Phi (x_k - x)\|/\|x_k - x\| = 0 \). Since the limit of the element-wise product of convergent sequences equals the product of its limits we can write

\[
\lim_{k \to \infty} \|\Phi(x_k) - \Phi(x) - D_x \Phi (x_k - x)/\delta_k\|/\|x_k - x\| = 0
\]

which, using the fact that \( D_x \Phi \) is linear, simplifies to

\[
\lim_{k \to \infty} \|\Phi(x_k) - \Phi(x) - D_x \Phi (x_k - x)/\delta_k\| = 0.
\]

This implies that \( (\Phi(x_k) - \Phi(x))/\delta_k \to D_x \Phi (v) \), and hence \( D_x \Phi (v) \) is a tangent vector of \( \Phi(\tilde{x}) \) at \( \Phi(x) \). This proves \((8.1)\).

To show \((8.2)\) we use \((8.1)\) together with the definition of the Clarke tangent cone as the inner limit of the surrounding tangent cones (Definition 2.2). We can write

\[
T_{\Phi(x)}^C \Phi(\tilde{x}) = \liminf_{y \to \Phi(x)} T_{y} \Phi(\tilde{x}) = \liminf_{y \to \Phi(x)} D_y \Phi T_y^C \tilde{x}.
\]

Since \( D_x \Phi \) is continuous in \( x \), we have \( \liminf_{y \to x} D_y \Phi (T_y \tilde{x}) = \liminf_{y \to x} D_x \Phi (T_x \tilde{x}) \). Further, Lemma A.3 implies that \( \liminf_{y \to x} D_x \Phi (T_y \tilde{x}) \supset D_x \Phi (\liminf_{y \to x} T_y \tilde{x}) = D_x \Phi (T_x^C \tilde{x}) \) and therefore we have \( T_{\Phi(x)}^C \Phi(\tilde{x}) \supset D_x \Phi (T_x^C \tilde{x}) \). Again, since \( \Phi \) is a diffeomorphism, the opposite inclusion holds by applying the same argument to \( \Phi^{-1} \). This shows \((8.2)\) and completes the proof.

Hence, the notions of (Clarke) tangent cone and Clarke regularity are independent of the coordinate representation on a \( C^1 \) manifold.

**Definition 8.2.** Let \( \mathcal{M} \) be a \( C^1 \) manifold with a metric \( g \) and consider a subset \( \mathcal{X} \subset \mathcal{M} \). The (Clarke) tangent cone \( T_x \mathcal{X} \) \( (T_x^C \mathcal{X}) \) is a subset of \( T_x \mathcal{M} \) such that

\(^4\)Note that local (weak) boundedness of a vector field or metric are properties that are preserved by \( C^1 \) diffeomorphisms. Similarly, absolute continuity is preserved by \( C^1 \) maps [47, Ex 6.44]. Hence, it is sufficient if these properties hold in any local coordinate domain.
\[ D_x φ(T_x Χ) (D_x φ(T_x^C Χ)) \] is the (Clarke) tangent cone of \( φ(Χ ∩ U) \) for any coordinate chart \((U, φ)\) defined at \( x \). The set \( Χ \) is Clarke regular at \( x ∈ Χ \) if it is Clarke regular in any local coordinate domain defined at \( x \).

The next key result establishes that solutions of projected dynamical systems remain solutions of projected dynamical systems under \( C^1 \) coordinate transformations.

**Proposition 8.3.** Let \( V, W ⊂ \mathbb{R}^n \) be open and consider a \( C^1 \) diffeomorphism \( Φ : V \to W \). Let \( Χ ⊂ \mathbb{R}^n \) be locally compact and \( \tilde{Χ} := Χ ∩ V \). Further, let \( g \) be a locally weakly bounded metric on \( W \) and let \( Φ^*g \) denote the pull-back metric along \( Φ \), i.e.,

\[
⟨v, w⟩_{Φ^*g(x)} := (D_x Φ(v), D_x Φ(w))_{g(Φ(x))}
\]

for all \( x ∈ V \) and \( v, w ∈ T_x \mathbb{R}^n \). Further, let \( f : \tilde{Χ} \to \mathbb{R}^n \) be a locally bounded vector field. If \( x : [0, T) → \tilde{Χ} \) for some \( T > 0 \) is a Krasovskii (respectively, Carathéodory) solution to the initial value problem

\[
\dot{x} ∈ Π^Φ^*g f(x), \quad x(0) = x_0,
\]

then \( Φ ∘ x : [0, T) → Φ(\tilde{Χ}) \) is a Krasovskii (respectively, Carathéodory) solution to

\[
\dot{y} ∈ Π^g_{Φ(Φ)} \hat{f}(y), \quad y(0) = y_0,
\]

where \( y_0 := Φ(x_0) \) and \( \hat{f}(y) := D_{Φ^{-1}(y)} Φ(f(Φ^{-1}(y))) \) is the pushforward vector field of \( f \) along \( Φ^{-1} \).

**Proof.** First, note that since \( x \) is absolutely continuous and \( Φ \) is differentiable, \( Φ ∘ x \) is absolutely continuous [47, Ex 6.44]. Second, it holds that \( y(t) ∈ Φ(\tilde{Χ}) \) for all \( t ∈ [0, T) \). Third, using (8.1) we can write for every \( x ∈ Χ \) and \( y := Φ(x) \) that

\[
Π^g_{Φ(Φ)} \hat{f}(y) = \arg\min_{w ∈ T_y Φ(\tilde{Χ})} ∥w − D_x Φ(f(x))∥_g = \arg\min_{w ∈ D_x Φ(T_x Χ)} ∥w − D_x Φ(f(x))∥_g
\]

\[
= D_x Φ \left( \arg\min_{v ∈ T_x Χ} ∥D_x Φ(v) − D_x Φ(f(x))∥_g \right),
\]

where for the last equality we introduce the transformation \( w := D_x Φ(v) \) for \( v ∈ T_x Χ \). Hence, using the definition of the pullback metric (8.3) we continue with

\[
Π^g_{Φ(Φ)} \hat{f}(y) = D_x Φ \left( \arg\min_{v ∈ T_x Χ} ∥v − f(x)∥_{Φ^*g} \right) = D_x Φ \left( Π^Φ^*g f(x) \right).
\]

Consequently, if \( x(\cdot) \) is a Carathéodory solution of (8.4) and hence \( \dot{x}(t) ∈ Π^Φ^*g f(x(t)) \) holds almost everywhere, then \( Φ ∘ x(\cdot) \) satisfies

\[
\frac{d}{dt} (Φ ∘ x) ∈ D_x Φ \left( Π^Φ^*g f(x) \right) = Π^g_{Φ(Φ)} \hat{f}(Φ ∘ x(t))
\]

almost everywhere and hence \( Φ ∘ x(\cdot) \) is a Carathéodory solution to (8.5).

It remains to prove the statement is also true for Krasovskii solutions. For this, we need to show that \( K[Π^g_{Φ(Φ)} \hat{f}](y)] ⊃ D_x Φ(K[Π^Φ^*g f](y)) \). Expanding the definition
of the Krasovskii regularization we get
\[
K \left[ \Pi_{\Phi(X)}^g \hat{f} \right] (y) = \co \limsup_{\hat{y} \to y} \pi_{\Phi(X)}^g \hat{f}(\hat{y}) \\
= \co \limsup_{\hat{x} \to x} D_x \Phi \left( \Pi_X^{p^*} g (\hat{x}) \right) \\
= \co \limsup_{\hat{x} \to x} D_x \Phi \left( \Pi_X^{p^*} \hat{f}(x_k) \right),
\]
where the last equation is due to the fact that \( D_x \Phi \) is continuous in \( x \). Next, with Lemma A.3 we can write
\[
K \left[ \Pi_{\Phi(X)}^g \hat{f} \right] (y) \supset \co \left( \limsup_{\hat{x} \to x} D_x \Phi \left( \Pi_X^{p^*} \hat{f}(x_k) \right) \right) = D_x \Phi \left( K \left[ \Pi_X^{p^*} \hat{f} \right] (x) \right)
\]
where the equation follows from the fact that \( D_x \Phi \) is a linear map and hence commutes with taking the convex closure.

To conclude we can proceed similar to the case of Carathéodory solutions. Let \( x(\cdot) \) be a Krasovskii solution to (8.4) and \( y(\cdot) := \Phi \circ x(\cdot) \). Then, \( \dot{y}(t) = \frac{d}{dt} (\Phi \circ x)(t) = D_xx(t)\Phi(\hat{x}(t)) \) for almost all \( t \in [0, T) \) and we have that
\[
\dot{y}(t) \in D_xx(t) \left( K \left[ \Pi_X^{p^*} \hat{f} \right] (x(t)) \right) \subset K \left[ \Pi_{\Phi(X)}^g \hat{f} \right] (y(t))
\]
for almost all \( t \in [0, T) \), and thus \( y \) is a Krasovskii solution of (8.5).

Hence, Theorems 4.2, 6.3 combined with Proposition 8.3 give rise to our main result on the existence of Krasovskii (Carethodory) solutions to on manifolds.

**Theorem 8.4 (existence on manifolds).** Let \( M \) be \( C^1 \) manifold, \( g \) a locally weakly bounded Riemannian metric, \( \mathcal{X} \subset M \) locally compact, and \( f \) a locally bounded vector field on \( \mathcal{X} \). Then for every \( x_0 \in \mathcal{X} \) there exists a Krasovskii solution \( x : [0, T) \to \mathcal{X} \) for some \( T > 0 \) that solves \( \dot{x}(t) \in \Pi_X^p f(x(t)) \) with \( x(0) = x_0 \). Furthermore, if \( \mathcal{X} \) is Clarke regular, and if \( f \) and \( g \) are continuous, then every Krasovskii solution is a Carathéodory solution and vice versa.

Similarly, Proposition 8.3 directly implies that other results such as Corollary 4.3 extend to \( C^1 \) manifolds. For instance, if \( M \) is compact and \( f \) and \( g \) are continuous, every initial condition admits a complete trajectory. However, to extend our uniqueness results, we require stronger conditions.

**Proposition 8.5.** Let \( V, W \subset \mathbb{R}^n \) be open and \( \Phi : V \to W \) a \( C^{1,1} \) diffeomorphism. Let \( \mathcal{X} \subset \mathbb{R}^n \) be locally compact and consider \( \tilde{\mathcal{X}} := \mathcal{X} \cap V \). If \( \tilde{\mathcal{X}} \) is prox-regular then \( \Phi(\tilde{\mathcal{X}}) \) is prox-regular.

**Proof.** By Proposition 7.5 it suffices to show prox-regularity with respect to a single metric on \( V \) and \( W \) respectively. Hence, let \( W \) be endowed with the Euclidean metric, and let \( e^* \) denote its pullback metric on \( V \) along \( \Phi \), i.e., \( \langle v, w \rangle_{e^*(x)} := \langle D_x\Phi(v), D_x\Phi(w) \rangle \). Similarly to Lemma 8.1, we show that (proximal) normal cones are preserved by \( C^1 \) coordinate transformations, i.e.,
\[
\eta \in N_x^\circ \tilde{\mathcal{X}} \iff D_x \Phi(\eta) \in N_{\Phi(x)}^\circ \Phi(\tilde{\mathcal{X}}) \quad \forall x \in \tilde{\mathcal{X}}
\]
\[
\eta \in \tilde{N}_x^\circ \tilde{\mathcal{X}} \iff D_x \Phi(\eta) \in \tilde{N}_{\Phi(x)}^\circ \Phi(\tilde{\mathcal{X}}) \quad \forall y \in N_x
\]
for some \( L', L > 0 \) where \( N_x \subset \tilde{\mathcal{X}} \) is a neighborhood of \( x \). Since \( \Phi \) is a diffeomorphism it suffices to show one direction only.
prox-regularity is an intrinsic property of subset of C results. In the process, we have established auxiliary findings, such as the fact that uniqueness and other properties of solution trajectories. Table 1 summarizes these vector field, metric and differentiable structure that are required for the existence, directions. We have carved out sharp regularity requirements on the feasible domain, systems on irregular subset on manifolds, including the model of oblique projection of Krasovskii solutions to projected gradient descent—arguably the most prototypical continuous-time constrained optimization algorithm.

This bound can be used to establish

\[ \langle \eta, w \rangle_{e^*(x)} = \langle D_x \Phi(\eta), D_x \Phi(w) \rangle \leq 0 \quad \forall w \in T_x \hat{X} \]

(8.8)

Hence, we define the C^{1,1} function \( \psi(z) := \langle D_y \Phi(\eta), \Phi(z) \rangle \) and note that by linearity we have \( D_z \psi(v) := \langle D_y \Phi(\eta), D_z \Phi(v) \rangle \). This enables us to apply the Desent Lemma A.2 and state that for some \( M > 0 \) it holds that

\[ |\psi(z) - \psi(y) - D_y \psi(z - y)| = \left| \langle D_y \Phi(\eta), \Phi(z) - \Phi(y) - D_y \Phi(z - y) \rangle \right| \leq M \|z - y\|^2. \]

This bound can be used to establish

\[ \langle D_y \Phi(\eta), \Phi(z) - \Phi(y) \rangle \leq \langle D_y \Phi(\eta), D_y \Phi(z - y) \rangle + \gamma(z) \leq (L + M) \|z - y\|^2. \]

Finally note that \( \|z - y\|^2 \leq L' \|\Phi(z) - \Phi(y)\|^2 \) for some \( L' \) since \( \Phi^{-1} \) is Lipschitz continuous. Hence, (8.8) and therefore (8.7) holds for \( L' = L''(L + M) \).

Apart from Proposition 8.5, we note that Lipschitz continuity of a metric and of vector fields is preserved under C^{1,1} coordinate transformations. This allows us to generalize Theorem 7.12 to the following uniqueness result on manifolds.

**Theorem 8.6 (uniqueness on manifolds).** Let \( M \) be C^{1,1} manifold, g a C^{0,1} Riemannian metric, \( X \subset M \) is prox-regular, and \( f \) a C^{0,1} vector field on \( X \). Then, for every \( x_0 \in X \) there exists a unique Carathéodory solution \( x : [0, T) \to X \) for some \( T > 0 \) that solves \( \dot{x}(t) = \Pi^o_X f(x(t)) \) with \( x(0) = x_0 \).

9. Conclusion. We have provided a holistic study of projected dynamical systems on irregular subset on manifolds, including the model of oblique projection directions. We have carved out sharp regularity requirements on the feasible domain, vector field, metric and differentiable structure that are required for the existence, uniqueness and other properties of solution trajectories. Table 1 summarizes these results. In the process, we have established auxiliary findings, such as the fact that prox-regularity is an intrinsic property of subset of C^{1,1} manifolds and independent of the choice of Riemannian metric.

While we believe these results are of general interest in the context of discontinuous dynamical systems, they particularly provide a solid foundation for the study of continuous-time constrained optimization algorithms for nonlinear, nonconvex problems. To illustrate this point, we have included a study the stability and convergence of Krasovskii solutions to projected gradient descent—arguably the most prototypical continuous-time constrained optimization algorithm.
Given a sequence \( x \) that converges to \( x \), it holds that \( x(t) \in T_{x(t)}X \cap T_{x(t)}X \) almost everywhere on \( [0, T) \), where \( T_{x(t)} := \{ v \mid v \in T_{x(t)} \} \).

**Proof.** Let \( t \in [0, T) \) be such that \( x(t) \) exists. This implies that by definition

\[
\dot{x}(t) = \lim_{\tau \to 0^+} \frac{x(t+\tau)-x(t)}{\tau} = \lim_{\tau \to 0^+} \frac{x(t)-x(t-\tau)}{\tau},
\]

Thus, by choosing any sequence \( \tau_k \to 0 \) with \( \tau_k > 0 \), the sequence \( \frac{x(t+\tau_k)-x(t)}{\tau_k} \) converges to a tangent vector and \( \frac{x(t)-x(t-\tau_k)}{\tau_k} \) converges to a vector in \( -T_{x(t)}X \) by definition of \( T_{x(t)}X \) and the fact that \( x(t) \in X \) for all \( t \in [0, T) \).

The following is a local version of [44, Lem 1.30].

**Lemma A.2 (Descent Lemma).** Let \( \Phi : V \to \mathbb{R} \) be a \( C^{1,1} \) map where \( V \subset \mathbb{R}^n \) is open. Given \( x \in V \) there exists \( L > 0 \) such that for all \( z, y \in V \) in a neighborhood of \( x \) it holds that

\[
|\Phi(z) - \Phi(y) - D_y\Phi(z-y)| \leq L\|z-y\|^2
\]

For a comprehensive treatment of the following definitions and results see [6, 30, 44, 46]. Given a sequence \( \{x_k\} \) and a set \( X \), the notation \( x_k \xrightarrow{\text{sub}} X \) denotes the existence of a subsequence \( \{x_{k'}\} \) that converges to \( x \) and \( x_{k'} \in X \) for all \( k' \). Similarly, \( x_k \xrightarrow{\text{ev}} X \) implies that \( x_k \in X \) holds eventually, i.e., for all \( k \) larger than some \( K \), and that \( \{x_k\} \) converges to \( x \). Given a sequence of sets \( \{C_k\} \) in \( \mathbb{R}^n \), its outer limit and inner limit are given as

\[
\limsup_{k \to \infty} C_k := \left\{ x \mid \exists \{i_k\} : x_i \xrightarrow{\text{sub}} C_{i_k} x \right\} \quad \text{and} \quad \liminf_{k \to \infty} C_k := \left\{ x \mid \exists \{i_k\} : x_i \xrightarrow{\text{ev}} C_{i_k} x \right\}
\]
respectively. As a pedagogical example to distinguish between inner and outer limits, consider an alternating sequence of sets given by $C_{2m} := A$ and $C_{2m+1} := B$. Then, we have $\limsup_{k \to \infty} C_k = A \cup B$ and $\liminf_{k \to \infty} C_k = A \cap B$. On the one hand any constant sequence $\{x_k\}$ with $x_k = c \in A \cap B$ for all $k$ satisfies the requirement such that $c \in \liminf_{k \to \infty} C_k$. On the other hand, any sequence $\{x_k\}$ with $x_{2m} = a \in A$ for $m \in \mathbb{N}$ has a trivial (constant) subsequence converging to $a \in A$ and hence $a \in \limsup_{k \to \infty} C_k$. The following result relates the image of an outer (inner) limit to the outer (inner) limit of images of a map $f$.

**Lemma A.3.** [46, Thm 4.26] For a sequence of sets $\{C_k\}$ in $V \subset \mathbb{R}^n$ and a continuous map $f : V \to \mathbb{R}^m$, one has

$$f \left( \liminf_{k \to \infty} C_k \right) \subseteq \liminf_{k \to \infty} f(C_k), \quad f \left( \limsup_{k \to \infty} C_k \right) \subseteq \limsup_{k \to \infty} f(C_k).$$

For a set-valued map $F : V \rightrightarrows \mathbb{R}^m$ with $V \subset \mathbb{R}^n$ and $W \subset \mathbb{R}^m$ its outer limit and inner limit at $x$ are defined respectively as

$$\limsup_{y \to x} F(y) := \bigcup_{x_k \to x} \limsup_{y \to x} F(x_k) \quad \text{and} \quad \liminf_{y \to x} F(y) := \bigcap_{x_k \to x} \liminf_{y \to x} F(x_k).$$

A set-valued map $F : V \rightrightarrows \mathbb{R}^m$ for $V \subset \mathbb{R}^n$ is outer semicontinuous (osc) at $x \in V$ if $\limsup_{y \to x} F(y) \subset F(x)$ [46, Def 5.4]. The map $F$ is upper semicontinuous (usc) at $x$ if for any open neighborhood $A \subset V$ of $F(x)$ there exists a neighborhood $B \subset V$ of $x$ such that for all $y \in B$ one has $F(y) \subset A$ [5, Def 2.1.2]. The map $F$ is outer (upper) semi-continuous if and only if it is osc (usc) at every $x \in V$. For locally bounded, closed set-valued maps outer and upper semicontinuity are equivalent.

**Lemma A.4.** [24, Lem 5.15] Let $F : V \rightrightarrows \mathbb{R}^m$ be closed and locally bounded for $V \subset \mathbb{R}^n$. Then, $F$ is osc at $x \in V$ if and only if it is usc at $x$. Furthermore, $F$ is osc/usc at $x$ if and only if $\text{gph} F := \{(x, v) \mid x \in V, v \in F(x)\}$ locally closed at $x$.

The next result states that upper semicontinuity is preserved by convexification.

**Lemma A.5.** [21, Lem 16, §5] Given a set-valued map $F : V \rightrightarrows \mathbb{R}^m$ with $V \subset \mathbb{R}^n$, if $F$ is usc and $F(x)$ is non-empty and compact for each $x \in V$, then the map $\text{co} F : V \rightrightarrows \mathbb{R}^m$ defined as $x \mapsto \text{co} F(x)$ is usc.

The following result is a generalization of [46, Prop 6.5] to the case of a continuous metric instead of the standard Euclidean metric:

**Lemma A.6.** Let $\mathcal{X}$ be Clarke regular. If the metric $g$ on $\mathcal{X}$ is continuous, then the set-valued map $\mathcal{X} \mapsto N^g_x \mathcal{X}$ is outer semi-continuous.

**Proof.** Consider any two sequences $x_k \to x$ with $x_k \in \mathcal{X}$ and $\eta_k \to \eta$ with $\eta_k \in N^g_x \mathcal{X}$. To complete the proof we need to show that $\eta \in N^g_x \mathcal{X}$. By definition of $N^g_x \mathcal{X}$ we have $\langle v, \eta_k \rangle g(x_k) \leq 0$ for all $v \in T^C_x \mathcal{X}$. Furthermore, by continuity of $g$ we have $\langle v, \eta \rangle g(x) \leq 0$ for all $v \in \limsup_{x_k \to x} T^C_{x_k} \mathcal{X}$. (Namely, we must have $\langle v_k, \eta_k \rangle g(x_k) \leq 0$ for every sequence $v_k \to v$ with $v_k \in T^C_{x_k} \mathcal{X}$, hence the use of lim sup.) By definition of the Clarke tangent cone, we note that $\langle v, \eta \rangle g(x) \leq 0$ holds for all

$$v \in T^C_x \mathcal{X} = \liminf_{x_k \to x} T_{x_k} \mathcal{X} = \liminf_{x_k \to x} T^C_{x_k} \mathcal{X} \subseteq \limsup_{x_k \to x} T^C_{x_k} \mathcal{X},$$

and therefore $\eta \in N^g_x \mathcal{X}$. \(\square\)
The following general existence and viability theorem goes back to [25]. Similar results can also be found in [5, 12, 24].

Proposition A.7 ([25, Cor. 1.1, Rem 3]). Let \( \mathcal{X} \) be a locally compact subset of \( \mathbb{R}^n \) and \( F : \mathcal{X} \to \mathbb{R}^n \) an usc, non-empty, convex and compact set-valued map. Then, for any \( x_0 \in \mathcal{X} \) there exists \( T > 0 \) and a Lipschitz continuous function \( x : [0, T] \to \mathcal{X} \) such that \( x(0) = x_0 \) and \( \dot{x}(t) \in F(x(t)) \) almost everywhere in \( [0, T] \) if and only if the condition \( F(x) \cap T_x \mathcal{X} \neq \emptyset \) holds for all \( x \in \mathcal{X} \). Furthermore, for \( r > 0 \) such that \( U_r := \{ x \in \mathcal{X} \, | \, \| x - x_0 \| \leq r \} \) is closed and \( L = \max_{y \in U_r} \| F(y) \| \) exists, the solution is Lipschitz and exists for \( T > r/L \).

REFERENCES

[1] P.-A. Absil and K. Kurdyka, On the stable equilibrium points of gradient systems, Syst. Control Lett., 55 (2006), pp. 573 – 577.
[2] P.-A. Absil, R. Mahony, and R. Sepulchre, Optimization Algorithms on Matrix Manifolds, Princeton University Press, Princeton, NJ, 2008.
[3] S. Adly, F. Nacry, and L. Thibault, Preservation of prox-regularity of sets with applications to constrained optimization, SIAM Journal on Optimization, 26 (2016), pp. 448–473.
[4] K. J. Arrow, L. Hurwicz, and H. Uzawa, Studies in Linear and Nonlinear Programming, Stanford University Press, Stanford, CA, 1958.
[5] J. P. Aubin, Viability Theory, Systems & Control: Foundations & Applications, Springer, Boston, MA, 1991.
[6] J.-P. Aubin and A. Cellina, Differential Inclusions: Set-Valued Maps and Viability Theory, Grundlehren der mathematischen Wissenschaften, Springer, Berlin, 1984.
[7] A. Bacciotti and F. Ceragioli, Stability and Stabilization of Discontinuous Systems and Nonsmooth Lyapunov Functions, ESAIM: COCV, 4 (1999), pp. 361–376.
[8] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, Nonlinear Programming: Theory and Algorithms, Wiley-Interscience, 3 ed., 2006.
[9] R. W. Brockett, Dynamical systems that sort lists, diagonalize matrices and solve linear programing problems, in Proceedings of the 27th IEEE Conference on Decision and Control, vol. 1, Dec. 1988, pp. 799–803.
[10] B. Brogliato, A. Daniilidis, C. Lemaréchal, and V. Acary, On the equivalence between complementarity systems, projected systems and differential inclusions, Systems & Control Letters, 55 (2006), pp. 45–51.
[11] A. Cherukuri, E. Mallada, and J. Cortés, Asymptotic convergence of constrained primal–dual dynamics, Systems & Control Letters, 87 (2016), pp. 10–15.
[12] F. H. Clarke, Optimization and Nonsmooth Analysis, Classics in Applied Mathematics, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1990.
[13] F. H. Clarke, Y. S. Ledyaev, R. J. Stern, and P. R. Wolenski, Nonsmooth Analysis and Control Theory, Graduate Texts in Mathematics, Springer, New York, NY, 1998.
[14] M.-G. Cojocaru and L. Jonker, Existence of solutions to projected differential equations in Hilbert spaces, Proc. Amer. Math. Soc., 132 (2004), pp. 183–193.
[15] M. G. Cojocaru and S. Pia, Nonpivot and Implicit Projected Dynamical Systems on Hilbert Spaces, Journal of Function Spaces and Applications, 2012 (2012), pp. 1–23.
[16] M. Colombino, E. Dall’Anese, and A. Bernstein, Online Optimization as a Feedback Controller: Stability and Tracking, ArXiv180509877 Math, (2018).
[17] B. Cornet, Existence of slow solutions for a class of differential inclusions, Journal of Mathematical Analysis and Applications, 96 (1983), pp. 130–147.
[18] J. Cortés, Discontinuous dynamical systems, IEEE Control Syst. Mag., 28 (2008), pp. 36–73.
[19] E. Dall’Anese and A. Simonetto, Optimal Power Flow Pursuit, IEEE Trans. Smart Grid, 9 (2018), pp. 942–952.
[20] M. Fazlyab, A. Ribeiro, M. Morari, and V. M. Preciado, Analysis of Optimization Algorithms via Integral Quadratic Constraints: Nonsstrongly Convex Problems, ArXiv17050615 Math, (2017), https://arxiv.org/abs/1705.0615.
[21] A. F. Filippov, Differential Equations with Discontinuous Righthand Sides, Mathematics and its Applications (Soviet Series), Springer, Dordrecht, The Nederlands, 1988.
[22] L. Gan and S. H. Low, An Online Gradient Algorithm for Optimal Power Flow on Radial Networks, IEEE J. Sel. Areas Commun., 34 (2016), pp. 625–638.
[23] S. Giuffrè, G. Idone, and S. Pia, Some classes of projected dynamical systems in Banach
spaces and variational inequalities, J Glob Optim, 40 (2008), pp. 119–128.
[24] R. Goebel, R. G. Sanfelice, and A. R. Teel, Hybrid Dynamical Systems: Modeling, Stability, and Robustness, Princeton University Press, Princeton, NJ, 2012.
[25] G. Haddad, Monotone trajectories of differential inclusions and functional differential inclusions with memory, Israel J. Math., 39 (1981), pp. 83–100.
[26] A. Hauswirth, S. Bolognani, G. Hug, and F. Dörfler, Projected gradient descent on Riemannian manifolds with applications to online power system optimization, in 54th Annual Allerton Conference on Communication, Control, and Computing, Sept. 2016, pp. 225–232.
[27] A. Hauswirth, A. Zanardi, S. Bolognani, F. Dörfler, and G. Hug, Online optimization in closed loop on the power flow manifold, in 2017 IEEE Manchester PowerTech, June 2017, pp. 1–6.
[28] W. P. M. H. Heemels, J. M. Schumacher, and S. Weiland, Projected dynamical systems in a complementarity formalism, Operations Research Letters, 27 (2000), pp. 83–91.
[29] C. Henry, An existence theorem for a class of differential equations with multivalued right-hand side, Journal of Mathematical Analysis and Applications, 41 (1973), pp. 179–186.
[30] J.-B. Hiriart-Urruty and C. Lemaréchal, Fundamentals of Convex Analysis, Grundlehren Text Editions, Springer, Berlin Heidelberg, 2nd printing ed., 2004.
[31] S. Hosseini and M. Pouryayevali, On the metric projection onto prox-regular subsets of Riemannian manifolds, Proc. Amer. Math. Soc., 141 (2013), pp. 233–244.
[32] F. P. Kelly, A. K. Maulloo, and D. K. H. Tan, Rate control for communication networks: Shadow prices, proportional fairness and stability, J Oper Res Soc, 49 (1998), pp. 247–252.
[33] W. Krichene, A. Bayen, and P. L. Bartlett, Accelerated Mirror Descent in Continuous and Discrete Time, in Advances in Neural Information Processing Systems 28, Curran Associates, Inc., 2015, pp. 2845–2853.
[34] L. S. P. Lawrence, J. W. Simpson-Porco, and E. Mallada, The Optimal Steady-State Control Problem, ArXiv181012892 Math, (2018).
[35] L. Lessard, B. Recht, and A. Packard, Analysis and Design of Optimization Algorithms via Integral Quadratic Constraints, SIAM J. Optim., 26 (2016), pp. 57–95.
[36] N. Li, C. Zhao, and L. Chen, Connecting Automatic Generation Control and Economic Dispatch From an Optimization View, IEEE Trans. Control Netw. Syst., 3 (2016), pp. 254–264.
[37] D. Liberzon, Switching in Systems and Control, Systems & Control: Foundations & Applications, Birkhäuser, Basel, Switzerland, 2003.
[38] S. H. Low, F. Paganini, and J. C. Doyle, Internet congestion control, IEEE Control Syst., 22 (2002), pp. 28–43.
[39] J. Lygeros, K. H. Johansson, S. N. Simic, J. Zhang, and S. S. Sastry, Dynamical Properties of Hybrid Automata, IEEE Trans. Autom. Control, 48 (2003), pp. 2–17.
[40] S. Menta, A. Hauswirth, S. Bolognani, G. Hug, and F. Dörfler, Stability of Dynamic Feedback Optimization with Applications to Power Systems, ArXiv181006079 Math, (2018).
[41] D. K. Moltzahn, F. Dörfler, H. Sandberg, S. H. Low, S. Chakrabarti, R. Baldick, and J. Lavaei, A Survey of Distributed Optimization and Control Algorithms for Electric Power Systems, IEEE Trans. Smart Grid, 8 (2017), pp. 2941–2962.
[42] A. Nagurney and D. Zhang, Projected Dynamical Systems and Variational Inequalities with Applications, no. 2 in International Series in Operations Research & Management Science, Springer Science & Business Media, 1 ed., 1996.
[43] Z. E. Nelson and E. Mallada, An integral quadratic constraint framework for real-time steady-state optimization of linear time-invariant systems, ArXiv17100204 Math, (2017).
[44] J. Peypouquet, Convex Optimization in Normed Spaces: Theory, Methods and Examples, Springer, Heidelberg, Germany, 2015.
[45] R. Poliquin and R. Rockafellar, Prox-regular functions in variational analysis, Transactions of the American Mathematical Society, 348 (1996), pp. 1805–1838.
[46] R. T. Rockafellar and R. J.-B. Wets, Variational Analysis, Springer, Berlin Heidelberg, Germany, 3rd printing ed., 1998.
[47] H. Royden and P. Fitzpatrick, Real Analysis, Macmillian, New York, NY, 2 ed., 1969.
[48] S. N. Simić, K. H. Johansson, S. Sastry, and J. Lygeros, Towards a Geometric Theory of Hybrid Systems, in Hybrid Systems: Computation and Control, Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, Mar. 2000, pp. 421–436.
[49] W. Su, S. Boyd, and E. Candes, A differential equation for modeling Nesterov’s accelerated gradient method: Theory and insights, in Advances in Neural Information Processing Systems, 2014, pp. 2510–2518.
[50] Y. Tang, K. Dvijotham, and S. Low, *Real-Time Optimal Power Flow*, IEEE Trans. Smart Grid, 8 (2017), pp. 2963–2973.

[51] A. Wibisono, A. C. Wilson, and M. I. Jordan, A variational perspective on accelerated methods in optimization, PNAS, 113 (2016), pp. E7351–E7358.

[52] A. C. Wilson, B. Recht, and M. I. Jordan, A Lyapunov Analysis of Momentum Methods in Optimization, ArXiv161102635 Cs Math, (2016).

[53] A. Ya.I., Generalized Projection Operators in Banach Spaces: Properties and Applications, in Theory and Applications of Nonlinear Operators of Monotone and Accretive Type, A. Kartsatos, ed., Marcel Dekker, New York, NY, 1996, pp. 15–50.

[54] X. Zhang, A. Papachristodoulou, and N. Li, Distributed Control for Reaching Optimal Steady State in Network Systems: An Optimization Approach, IEEE Trans. Autom. Control, 63 (2018), pp. 864–871.