MULTI-AGENT NAVIGATION FUNCTIONS: HAVE WE MISSED SOMETHING?

Herbert G. Tanner, {btanner@udel.edu}
Adithya Boddu, {adithyab@udel.edu}

UD MEEG Technical Report
Number MEEG TR-2010-0001
August 2010

Department of Mechanical Engineering
University of Delaware
Newark, Delaware
URL: http://www.me.udel.edu

MULTI-AGENT NAVIGATION FUNCTIONS: HAVE WE MISSED SOMETHING?

Herbert G. Tanner, {btanner@udel.edu}
Adithya Boddu, {adithyab@udel.edu}

University of Delaware, Department of Mechanical Engineering
MEEG TR-2010-0001

August 2010

Abstract

This paper presents a methodology for designing (centralized) control laws which can provably steer a group of robotic agents to fall into a formation of arbitrary shape, while following collision free trajectories. The scheme is based on the concept of navigation functions, a special type of artificial potential functions without local minima, and the paper describes how this idea can be generalized from its original formulation for single-robot systems, to multi-robot formations. We indicate why existing solutions that have appeared in literature, although potentially functional, may not have been accompanied with sufficient guarantees against the possibility of the system ever getting stuck in non-optimal configurations. The problem is therefore revisited here under a new set of assumptions, a new construction is proposed, and the properties of this new centralized potential function are analytically demonstrated.

1 Introduction

With the advent of inexpensive and efficient portable computation and communication electronics that enable the effective application of multi-agent robotic systems, interest in developing a comprehensive theory for formation control design has intensified. Among several different solution approaches [1, 2, 3, 4, 5], there is one that generalizes a potential field approach to (single) robot navigation which offers global stability guarantees. This particular approach involves the construction of a specific type of artificial potential function without local minima, known as *navigation function*. The (single-robot) navigation function was first introduced in [6] for environments with sphere-world topology, and was then generalized to star-world environments in [7] through a series of diffeomorphic transformations. It applied to robots with negligible volume (point-mass robots). Generalizations of this approach have appeared in literature, both in the form of centralized and decentralized schemes.

A centralized architecture typically involves a single potential function, and information about its construction is assumed to be shared among all agents in the group. In the centralized approach [8] the gradient of a single navigation function generates trajectories for multiple robots in their composite configuration space.

In the decentralized case [9, 10, 11, 12, 13], individual agents use their own potential function, which they can calculate using information that is locally available (either by their own sensors or by direct communication with their nearest neighbors). Robot trajectories are then calculated independently based on these agent-specific functions. While [9] and [10] are a direct extension of single robot navigation described in [6] to multiple agents, the construction in [9] requires knowledge of the number of agents in the workspace. In both of these papers, there were no stationary obstacles in the environment. In [11], the emphasis is on formation control, and the workspace includes agents as well as obstacles, with each agent having a limited communication region within which the agent can properly communicate with any other agent. In [12] and [13], the problem of formation control is treated in a case where agent relative position specifications are encoded in the form of a formation control graph. In [12], the degree of decentralization is limited because each agent has a copy of the global navigation function where as in [13], the decentralization is complete, in that, the navigation function is composed of local navigation functions built on each agent. In both [12] and [13], the communication capability of each agent is assumed limited.

As it may be expected, all the above generalizations of the single-robot navigation function to the multi-robot case follow the analysis steps of the original construction and make very similar assumptions. It turns out, however, that some of these assumptions unavoidably break down in the multi-agent case, and very special care is needed to work around the technical challenges presented because of this reason. One of these critical assumptions that has been overlooked is the one that requires that "obstacles are isolated." In the single-robot case where this assumption was originally imposed, the statement implies that as the system approaches a collision configuration (with an obstacle), there is a single obstacle the distance to which continuously decreases, whereas distances to other obstacles remain bounded above zero. In the multi-robot case collisions can occur between robots, and since all robots can move, it is not clear why only two robots can collide with each other at any given time; in fact, it is theoretically possible that all robots collapse on each other simultaneously. We show that the breakdown of such an assumption has serious implications on the properties of potential functions constructed as direct generalizations of the single-robot navigation function.

2 Single Agent Navigation Functions

In its original "model" form, the *navigation function* [7] was defined for a single robot agent,

$$\dot{x} = u \ ,$$

moving in a workspace $\mathscr{W} \triangleq \{x \in \mathbb{R}^n : \|x\|^2 \leq \rho_0^2\}$, populated by M spherical obstacles $\mathscr{O}_i \triangleq \{x \in \mathbb{R}^n : \|x - c_i\|^2 < \rho_i^2\}$ (note that these sets are open). Here, c_i and ρ_i are the center and radius of obstacle i, respectively. The boundary of $\mathscr{W}$ is referred to as the zero$^{\text{th}}$ obstacle, centered at the origin. Then the free workspace for the robot is the set

$$\mathscr{F} \triangleq \mathscr{W} \setminus \bigcup_{i=1}^{M} \mathscr{O}_i \ .$$

In this paper we use the notation $\delta\mathscr{F}$ to denote the boundary of a set —in this case, $\mathscr{F}$— reserving the symbol ∂ for (generalized) gradients.

This workspace $\mathscr{F}$ is said to be *valid* if the closures of all $\mathscr{O}_i$ are in the interior of $\mathscr{F}$, and that none of them intersect:

$$\|c_i - c_j\| > \rho_i + \rho_j \tag{1}$$

Definition 1 ([7]). *Let $\mathscr{F} \subset \mathscr{R}^n$ be a compact connected analytic manifold with boundary. A map $\varphi : \mathscr{F} \to [0,1]$ is a navigation function if φ is*

1. *Analytic on $\mathscr{F} - \mathscr{M}$, where $\mathscr{M}$ is a set of measure zero.*
2. *Polar on $\mathscr{F}$, with minimum at x_d.*
3. *Morse on $\mathscr{F}$;*
4. *Admissible on $\mathscr{F}$.*

Rimon and Koditschek [6] have shown that under certain conditions, the following function can become a navigation function

$$\varphi = \frac{\gamma_d}{[\gamma_d{}^k + \beta]^{1/\kappa}} \ . \tag{2}$$

In the above, $\gamma_d(x)$ is a scalar function that serves as a metric of how close the robot is to its desired configuration, $\beta(x)$ is another scalar function that quantifies how close the robot is to obstacles, and k is a tuning parameter, which needs to be sufficiently large in order for the function to have the properties of Definition 1. In addition, γ_d and β are assumed to have a special structure, namely

$$\gamma_d(x) \triangleq \|x - x_d\|^2 \qquad \beta(x) \triangleq \prod_{i=1}^{M} \beta_i(x) \qquad \beta_i(x) \triangleq \|x - c_i\|^2 - \rho_i^2 \ , \tag{3}$$

where x_d is the desired configuration, i ranges from 0 to M (β_0 is the boundary of the whole workspace).

So in the particular case of φ as given by (2) and with the assumption that the workspace is valid, one can analytically show that φ is a navigation function [6]. The workspace $\mathscr{F}$ is first decomposed into regions (ε is thought of as a small parameter):

- $\{x_d\}$: the destination point,
- $\mathscr{B}_i(\varepsilon)$: Region "close" to the boundary of obstacle i, where which $0 < \beta_i < \varepsilon$.
- $\delta\mathscr{F}$: Boundary of free space,
- $\mathscr{F}_0(\varepsilon) \triangleq \bigcup_{i=1}^{M} \mathscr{B}_i \setminus \{x_d\}$: Region "close" to the obstacle boundary,
- $\mathscr{F}_2(\varepsilon) \triangleq \mathscr{F} \setminus (\{x_d\} \cup \delta\mathscr{F} \cup \mathscr{F}_0 \cup \mathscr{F}_1)$: away from obstacles,

and then a sequence of Propositions establish that *if the workspace is valid*:

1. The destination x_d is a non-degenerate local minimum of φ,
2. The critical points of φ are in the interior of the free space,
3. For every ε one can choose a k so that $\frac{\gamma}{\beta}$ has no critical points away from obstacles,
4. There exists a lower bound for ε, below which the critical points of $\frac{\gamma}{\beta}$ close to the obstacle boundary are not local minima,
5. There is another lower bound for ε, below which there are no critical points close to the workspace boundary, and
6. There is a last lower bound for ε, below which whatever critical points besides the destination are inside the free space, are non-degenerate (and therefore, they have to be saddles).

3 Multi-Robot Navigation Functions

3.1 The initial approach

Consider a homogeneous group of N mobile agents, moving in an n-dimensional obstacle-free environment, each with dynamics given by

$$\dot{p}_i = u_i, \quad p_i \in \mathbb{R}^n, \quad i = 1, \dots, N \tag{4}$$

where p_i and u_i are the position and control input of agent i, respectively. Let p denote the stack vector of all p_i, and define $\mathscr{P}$ as the set of p for which $\|p_i - p_j\| \geq \rho$, for all $i, j = 1, \dots, N$.

A straightforward approach followed in recent literature is to start with (2), define a goal function γ_d and the obstacle function β is a similar way as in (3):

$$\varphi = \frac{\gamma}{(\gamma + \beta)^{1/\kappa}} \quad .$$

These choices have advantages and disadvantages: on one hand, this γ_d requires each robot to achieve a pre-specified position in the workspace and thus does not allow the formation to "float" freely in space; on the other hand it ensures that the destination point will eventually turn out to be a non-degenerate critical point. The choice of β, however, as the product of individual obstacle functions that vanish at a collision configuration, has more severe implications: it threatens to destroy the validity of workspace $\mathscr{F}$, because now the obstacles are *not* isolated. In fact, there is always this pesky configuration where every agent touches at least two other agents, and more than one individual obstacle functions vanish simultaneously. In more extreme situations where a sufficient number of individual obstacle functions vanish concurrently, one can have $\nabla^2\beta$ vanishing too, which makes the task of proving claim 4 in the list of the previous section, at least using the known approach, impossible. This is because, at a critical point,

$$\nabla^2\left(\frac{\gamma}{\beta}\right) \propto 2\frac{\|\nabla\beta\|}{\|\nabla\gamma\|} I - \nabla^2\beta \quad , \tag{5}$$

and with $\nabla^2\beta \to 0$ the hope of finding a negative eigenvalue for this matrix evaporates. Having said this, it is not necessarily the case that a potential function built in this way may fail to stabilize a multi-agent formation —in fact, reported numerical results point to the opposite. This could be because the potential problems may arise only in extreme situations that cannot be predicted beforehand, and it could also be the case that a critical point in question is not a local minimum but rather just a degenerate critical point. Nonetheless, without the analytic convergence guarantees of the navigation function approach one can never definitively exclude the possibility of failure.

3.2 The new suggestion

In this paper, the agents are treated as autonomous, identical sphere-shaped robots. There are no static obstacles in their environment which means that and the only collisions that can occur are between them.

At the cost of not being able to generalize the proposed method to star-shaped words, we relax the requirement for having analytic and admissible functions. It appears that there can be benefits in using nonsmooth functions to construct φ, which we intend to exploit.

The goal is to steer the agents from any relative initial configuration so that they fall into a pre-specified formation, without colliding with each other and without fixing the location of that formation in space. The formation itself is described by means of a graph:

Definition 2 (Formation graph [13])**.** *The formation graph $\mathscr{G} = \{\mathscr{V}, \mathscr{E}, \mathscr{C}\}$ is a directed labeled graph consisted of:*

- *a set of vertices $\mathscr{V} = \{v_1, \ldots, v_N\}$, indexed by the mobile agents,*
- *a set of edges $\mathscr{E} = \{(i, j) \in \{1, \ldots, N\} \times \{1, \ldots, N\}\}$, containing ordered pairs of nodes that represent inter-agent position specifications, and*
- *a set of labels (formation specifications) $\mathscr{C} = \{c_{ij} \mid (i, j) \in \mathscr{E}\}$*

Whenever there is a $c_{ij} \in \mathscr{C}$, it implies that the desired relative position between agent i and agent j is c_{ij}, that is, ideally, we should have $\|p_i - p_j - c_{ij}\| = 0$. If graph $\mathscr{G}$ is (weakly) connected, then the formation is uniquely specified.

Since meeting the formation specifications depends on relative positions only, our analysis is performed in the space of relative differences

$$q_{ij} \triangleq p_i - p_j \quad .$$

If we denote p the stack vector of absolute agent positions, and q the stack vector of relative agent positions (differences), then

$$q = Bp,$$

where B is the incidence matrix of graph $\mathscr{G}$. Let us Our decision to shift the analysis on the space of relative positions exclusively is that mixing absolute and relative coordinates (e.g., when considering both formation specifications and static obstacle avoidance) leads to both analytic and conceptual problems when it comes to establishing the properties of a potential function.

We will work with a function φ, which is not a navigation function in the strict sense, but it is just polar, Morse, and it blows up at the boundary of the collision free space:

$$\varphi(q) = \frac{\gamma(q)}{\beta(q)} \quad ,$$

where γ is the "goal" function, a positive semi-definite scalar function attaining zero only when the agents are at their desired relative configurations

$$\gamma(q) \equiv {\gamma_d}^{\kappa}(q) \triangleq \sum_{(i,j)\in\mathscr{E}} \left\|q_{ij} - c_{ij}\right\|^2 \quad , \tag{6}$$

and $\beta(q)$ is another positive semi-definit scalar function, varying in the interval $[0, 1]$ which vanishes when any agents are in contact and is maximal when all agents are *sufficiently* far from each other. We suggest the following choice of the scalar function β

$$\beta = \log(2 - a\, e^{-\left(-r+d+d^2\right)^2}) \quad , \tag{7}$$

with a and r positive scalar parameters that are used to set the location where β vanishes and its derivative there. Specifically, if one needs β to vanish when $d = d_0$ and have a derivative equal to c, then the choice

$$r = \frac{4{d_0}^3 + 6{d_0}^2 + 2d_0 - c}{2(2d_0 + 1)}, \tag{8a}$$

$$a = \exp\left([{d_0}^2 + d_0 - r]^2\right) \quad , \tag{8b}$$

meets the requirement. The argument of β is the minimum out of all pairwise distances between agents

$$d(q) \triangleq \min_{\substack{i\in\{1,\dots,N\}\\ j\neq i}} \left\{ \left\| p_i - p_j \right\| \right\} \quad , \tag{9}$$

measured between the centers of their spherical shapes, and where the minimum is taken over every combination of $i, j \in \{1,\dots,N\}$. If the agents' radii are equal to ρ, then it follows that one needs to choose r and a so that β vanishes when $d = 2\rho$. The properties of this distance function are important in the present analysis and will be discussed in some detail in Section 3.3

The choice of β as in (7) gives the obstacle function the following attributes:

1. it vanishes whenever any two agents collide and remains positive otherwise;
2. it approaches a constant asymptotically as agents grow further apart;
3. it is non-differentiable.

The need to resort to a function of non-smooth nature if one desires to establish convergence properties for φ while maintaining the local nature of interaction between agents, arises once the following list of requirements for β is adopted:

1. $q \to \delta\mathscr{F} \Rightarrow \beta \to 0$: at collision configurations, β needs to vanish.
2. $\|q\| \to \infty \Rightarrow \beta \to c < \infty$: when agents distance themselves, their repulsion effect should weaken.
3. $\beta > 0$ for $q \in \mathscr{F}$.
4. $\frac{\partial \beta}{\partial x} > 0$ for β to serve as a metric of distance to collision configurations.
5. $\nabla\beta \neq 0$ at $\delta\mathscr{F}$ so that we do not end up with critical points on, or have to slide along the boundary of the free space (needed in Proposition 3).
6. $\frac{\partial^2 \beta}{\partial x^2} > 0$ near the critical points, to avoid running into the situation illustrated in (5).

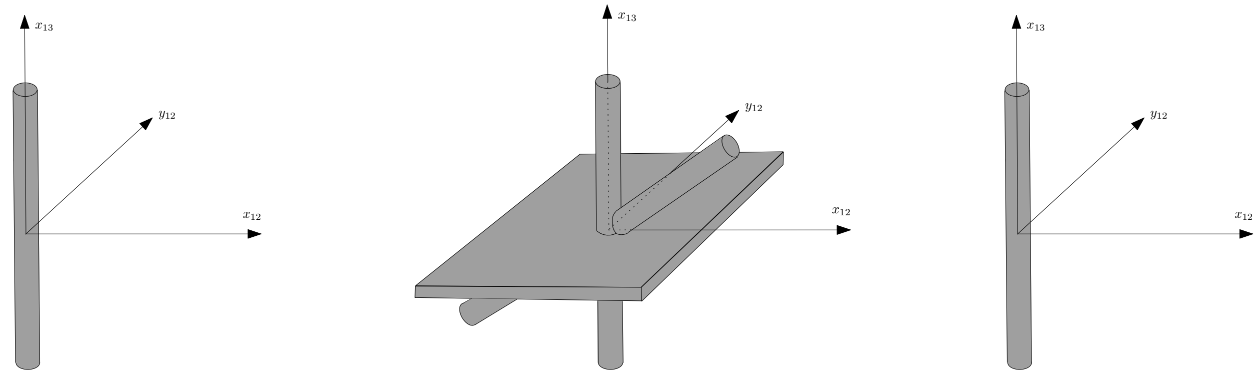
One can observe that the requirement for β being strictly increasing and bounded in $[0,1]$ is inconsistent with the specification that its second derivative is positive, at least for differentiable functions; the curvature has to switch. It is true that these particular conditions are not necessary in the form stated above, but it is not obvious how to relax them either. On the other hand, one might give up the requirement for a bounded β, forcing the agents to "feel" each other's influence anywhere they are in their workspace (no localized collision avoidance interaction). In this work, we attempt to address the requirement for localized interaction directly, by adopting a nonsmooth structure for β that fits within the specifications set above.

3.3 The distance function

The distance function (of a point to a set) is typically defined as the minimum norm of the difference between the point and any other point in the set. In our case, the point in question is q and the set is the manifold where any of the components of q becomes zero. Since this is quite hard to visualize in $n > 3$ dimensions we prefer to define d more intuitively as in (9), but in this section we attempt to provide some insight into the nature and the properties of this function.

First we need to note that in all but trivial cases, vector q cannot be visualized in a three-dimensional graphic. As an illustrative example, we take the case of three agents moving in a two-dimensional workspace. We thus have p_1, p_2, and p_3 denoting the positions of the agents. Their relative position vectors are denoted $q_{12} = p_1 - p_2$, $q_{13} = p_1 - p_3$ and $q_{23} = p_2 - p_3$. Note, however, that if q_{12} and q_{13} are set q_{23} is derived directly, so to represent the formation configuration we only need two of these vectors, say q_{12} and q_{13}. (In graph theoretic terms, we would say that we pick the edges of a minimum spanning tree from the formation graph.) For simplicity, therefore, let us think of q be consisted of only the components of q_{12} and q_{13}:

$$q = \begin{bmatrix} {q_{12}}^T & {q_{13}}^T \end{bmatrix}^T \quad .$$



(a) $y_1 - y_3 < 0$ (b) $y_1 - y_3 = 0$ (c) $y_1 - y_3 > 0$

Figure 1: Three 3-D slices of an obstacle function defined in a four-dimensional space. The collision configurations between agents 1 and 2 are marked by the cylinder that contains the (hyper)line $x_{12} = y_{12} = 0$. In the three-dimensional slices where the dimension y_{13} is not pictured, the collision configurations between agents 1 and 3 are shown as the "thick" hyperplane passing through the origin on the $y_1 - y_3 = 0$ slice. Note the diagonal cylinder with axis on the x_{13}-x_{12} plane: this represents collisions between agents 2 and 3 (although x_{23} and y_{23} are not mapped). This diagonal collision region expresses the fact that when $q_{12} = q_{13}$, agents 2 and 3 overlap; at the slice where $y_{13} = 0$, therefore, and on the plane where $y_{12} = 0 = y_{13}$, the diagonal line $x_{12} = x_{13}$ maps configurations where all three agents have the same y coordinate, and agent 2 is on top of agent 3. These three graphs illustrate that pairwise obstacle functions (i.e., collision between 1 and 2, or collision between 1 and 3) define regions in the relative position space which are not isolated, and irrespectively of the agents' volume the origin of this space will always be a point common to all regions.

When trying to visualize q we run into a dimensionality problem. Fortunately, for this particular case we "only" have four dimensions: x_{12}, y_{12}, x_{13}, and y_{13}. We can therefore attempt to picture what happens using three-dimensional slices of this four-dimensional space (see Fig. 1).

As Fig. 1 indicates, the distance-to-collision function defines complex nonintuitive regions in the configuration space which partially overlap. In addition, just as any typical distance function, it is nonsmooth in general; not only at the origin, as in the case of a classical vector norm, but also throughout the configuration space as the pairs of agents that are closest to each other compared to any other pair within the group, may change. Fortunately, the literature reports several interesting properties of the distance function and its (generalized) derivative. Here, the generalized (directional) derivative is being understood in the Clarke sense:

Definition 3 (Generalized derivative [14])**.** *The (Clarke) generalized derivative of $f(x)$ in the direction v is defined as*

$$f^{\circ}(x;v) = \limsup_{\substack{y \to x \\ h \downarrow 0}} \frac{f(y + hv) - f(x)}{h} \quad .$$

Definition 4 (Generalized gradient [14])**.** *The (Clarke) generalized gradient of $f(x)$ at $x \in X$ is defined as*

$$\partial f(x) = \{\zeta \in X^* \mid f^{\circ}(x;v) \geq \langle \zeta, v \rangle \rangle\} \quad ,$$

where X^* is the dual space of continuous linear functionals on X, and $\langle \zeta, v \rangle$ denotes the value of functional ζ at v.

In the setting of the problem we are addressing in this paper, let us define the discrete sets of points

$$\begin{aligned} \Omega &\triangleq \{p_i \mid 1 \leq i \leq N\}, \\ \Omega_i &\triangleq \{p_j \mid j \neq i\}, \quad \text{for } i \in \{1, \ldots, N\} \quad , \end{aligned}$$

which includes all agent position vectors except for the position of agent i. These two sets allow us to express the distance function we introduced in (9),

$$d(q) \triangleq \min_{\substack{i,j \in \{1,\ldots,N\} \\ i \neq j}} \|q_{ij}\| \quad .$$

in terms of the distances between a point $p_i \in \Omega$ and the set Ω_i

$$d_i(p_i) \triangleq \min_{z\in\Omega_i} \|p_i - z\| = \min_{\substack{j=1,\dots,N\\ i\neq j}} \|q_{ij}\| \tag{10}$$

as follows:

$$d(q) = \min_{1\leq i\leq N} d_i(p_i) = \min_{i=1,\dots,N} \min_{\substack{j=1,\dots,N\\ i\neq j}} \|q_{ij}\| \ . \tag{11}$$

With reference to the point-to-set distance d_i, we define the set

$$Q(p_i) = \{z \in \Omega_i \mid \|p_i - z\| = d_i(p_i)\} \ ,$$

which identifies the nearest neighbor(s) of agent i. The indices of the nearest neighbors (the agents equidistant from i with the smallest distance from i) are contained in the sets

$$I_i = \left\{ j \in \{1,\dots,N\} : p_j \in Q(p_i) \right\} \ .$$

Existing literature [15, Theorem 1] allows us to state a the following fact about the distance d_i, which is defined in (10) in terms of the Euclidean norm:

Corollary 1. *For $p_i \neq p_j$, for all $j \in \{1,\dots,N\}\setminus\{i\}$, (i.e., when $q_{ij} \neq 0$),*

$$\partial d_i(p_i) = \mathrm{co}\left\{ \partial\|q_{ij}\| \mid j \in I_i \right\} \ .$$

Note that away from $q_{ij} = 0$, $\partial\|q_{ij}\| = \nabla\|q_{ij}\| = \frac{p_j - p_i}{\|q_{ij}\|}$: the unit vector that points away from p_i and toward p_j.

Remark 1. *One may be tempted to show the equality of Corollary 1 alternatively, using regularity arguments. In many general cases, the distance function from a point to a set is indeed convex (e.g., this happens when the set is convex). Convexity implies regularity, which in turns sharpens the inclusions to equalities. In the case considered here, however, one cannot use this argument, because the set with respect to which the distance is measured is finite and nonconvex. Without being able to resort to regularity, however, in our next statement we relinquish the equality just achieved, since the distance function used for the potential function is not d_i but rather d.*

The following statement follows from [14, Proposition 2.3.12] (the case of pointwise maxima) —but we were not able to locate it in literature in the general form stated below. Stronger versions of it exist for the case where the functions within the min are convex or regular [16].

Proposition 1 (Pointwise minima)**.**

$$\partial \min_{i=1,\dots,n} f_i(x) \subset \mathrm{co}\{\partial f_i(x) : i = 1,\dots,n\} \ .$$

Proof.

$$\begin{aligned}
\partial \min_{1\leq i\leq n} f_i(x) &= \partial\left(-\max_{1\leq i\leq n}\left\{ -f_i(x)\right\}\right) = -\partial \max_{1\leq i\leq n}\left\{ -f_i(x)\right\} \\
&\overset{[14,\text{ Proposition 2.3.12}]}{\subset} -\mathrm{co}\left\{ \partial\left(-f_i(x)\right) : 1 \leq i \leq n \right\} = -\mathrm{co}\left\{ -\partial f_i(x) : 1 \leq i \leq n \right\} \\
&\overset{[16,\text{ Proposition A.1}]}{=} \mathrm{co}\left\{ \partial f_i(x) : 1 \leq i \leq n \right\}
\end{aligned}$$

□

As a result, we can write

$$\partial d(q) \subset \mathrm{co}\{\partial\|q_{ij}\| : (i,j) \in \mathscr{E}, j \in I_i\} \ ,$$

where the derivatives are taken with respect to q_{ij}. Away from points where there exists a $(i,j) \in \mathscr{E}$ such that $q_{ij} = 0$, $\partial\|q_{ij}\|$ is a singleton (recall that gradients are taken with respect to q_{ij}, so in each case the derivative is that of the distance of a vector from the origin) and therefore

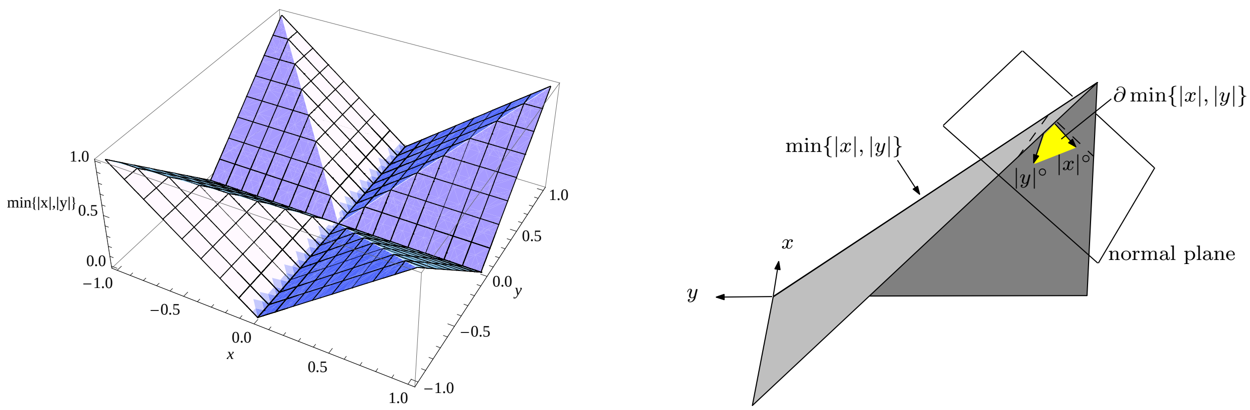
$$q_{ij} \neq 0\,, \forall (i,j) \in \mathscr{E} \Rightarrow \partial d(q) \subset \mathrm{co}\{\nabla\|q_{ij}\| : (i,j) \in \mathscr{E}\} \ . \tag{12}$$

It may not be obvious, but these convex hulls never contain the zero vector away from $q_{ij} = 0$. This may become clear with the following two-dimensional example.

Let

$$f(x,y) = \min\{|x|,|y|\} \ ,$$

in which case we have $f_1(x,y) = |x|$, and $f_2(x,y) = |y|$. The graph of $f(x,y)$ is shown in Fig. 2(a)



(a) Graph of $\min\{|x|,|y|\}$

(b) Generalized gradient along the $x = y$ line

Figure 2: The $\min\{|x|,|y|\}$ function and its generalized derivative. Fig. 2(a) shows that the function has minima along the x and y axes. Along the lines $x = y$ and $x = -y$ the function is not differentiable. Fig. 2(b) illustrates why the generalized gradient along the $x = \pm y$ lines does not contain zero: $(0,0) \notin \mathrm{co}\{(0,1),(1,0)\}$! There is no positive λ for which $(0,0) = \lambda(0,1) + (1-\lambda)(1,0)$.

The generalized gradient of f_1 away from $x = 0$ is given by

$$x \neq 0 \Rightarrow \partial|x| = \nabla|x| = \begin{cases} (1,0), & x > 0 \\ (-1,0), & x < 0 \end{cases} \ ,$$

while for f_2 we have

$$y \neq 0 \Rightarrow \partial|y| = \nabla|y| = \begin{cases} (1,0), & y > 0 \\ (-1,0), & y < 0 \end{cases} \ .$$

Therefore, away from the axes and the lines $x = \pm y$ the generalized gradient of f will be a singleton:

$$\partial f(x,y) = \begin{cases} (0,1), & 0 < x < y \\ (1,0), & 0 < y < x \\ (-1,0), & -y < x < 0 \\ (0,1), & x < -y < 0 \\ (0,-1), & x < y < 0 \\ (-1,0), & y < x < 0 \\ (1,0), & 0 < x < -y \\ (0,-1), & 0 < -y < x. \end{cases} \ , \quad x \neq \pm y, x \neq 0, y \neq 0 \ .$$

As Fig. 2(b) illustrates, along the lines $x = \pm y$ for $x, y \neq 0$, ∂f is a convex set that does not contain the origin. The distance function $d(q)$ is practically a generalization of $f(x,y)$ in multiple dimensions and behaves very similarly in terms of its (generalized) derivatives. In fact, for functions expressed as pointwise minima (the case of maxima can be treated similarly) where Proposition 1 applies, we have the following result.

Lemma 1 ([16]). *The origin is contained in the interior of the convex hull of a set of n arbitrary vectors $\{v_i \in \mathbb{R}^m, i = 1,\ldots,n\}$ iff there exists a v_i such that for all $w \in \mathbb{R}^m$, $\langle w, v_i \rangle > 0$.*

3.4 Properties of the proposed potential function

The analysis of this section focuses specifically on the properties of the gradient of φ, and particularly on the nature and behavior of its critical points as a function of the potential function parameter κ, as well as the obstacle function parameters a and r.

The discussion is organized in two parts. Section 3.4.1 treats the case where the distance function d is differentiable, namely away from configurations where inter-agent distances become equal to each other, and away from configurations where $q_{ij} = 0$ (the latter configurations are infeasible since the agents are assumed to have a non-zero spherical volume, and a configuration where $q_{ij} = 0$ corresponds to a situation where they overlap in their workspace). Section 3.4.2 discusses specifically the case where more than one pair of agents achieve the smallest inter-agent distance among the group. Then, the distance function does not have a derivative in the classical sense, but admits a generalized derivative (and gradient) which has the properties demonstrated in Section 3.3.

3.4.1 When the distance is differentiable

In the following proposition, we show that the destination configuration q_d is a non-degenerate local minimum of φ.

Proposition 2. *(cf. [6, Proposition 3.1]) If the workspace is valid, the destination point, q_d, is a non-degenerate local minimum of φ.*

Proof. We have

$$\varphi = \frac{\gamma}{\beta} \ ,$$

where $\beta = \log(2 - a\, e^{-\left(r+d+d^2\right)^2})$. At configurations where d is differentiable, we can write

$$\nabla\varphi = \frac{1}{\beta^2}\left(\beta\nabla\gamma - \gamma\nabla\beta\right) = \frac{1}{\beta^2}\left(\beta\kappa{\gamma_d}^{\kappa-1}\nabla\gamma_d - {\gamma_d}^{\kappa}\nabla\beta\right) \ . \tag{13}$$

Evaluating $\nabla\varphi$ at the destination q_d gives

$$\nabla\varphi\,(q_d) = \frac{\nabla\gamma(q_d)}{\beta(q_d)} \ .$$

The Hessian of ϕ, on the other hand, is

$$\begin{aligned}
\nabla^2\varphi &= \frac{1}{\beta^4}\left[\beta^2\,\nabla(\beta\nabla\gamma - \gamma\nabla\beta) - (\beta\nabla\gamma - \gamma\nabla\beta)\,\nabla\left(\beta^2\right)\right] \\
&= \frac{1}{\beta^4}\left[\beta^2\left(\beta\nabla^2\gamma + \nabla\beta\nabla\gamma^T - \nabla\gamma\nabla\beta^T - \gamma\nabla^2\beta\right) - (\beta\nabla\gamma - \gamma\nabla\beta)2\beta\nabla\beta^T\right] \\
&= \frac{1}{\beta^3}\left[\beta\left(\beta\nabla^2\gamma - \gamma\nabla^2\beta + \nabla\beta\nabla\gamma^T - \nabla\gamma\nabla\beta^T\right) - (\beta\nabla\gamma - \gamma\nabla\beta)2\beta\nabla\beta^T\right] \ .
\end{aligned} \tag{14}$$

At q_d, we have $\nabla\gamma(q_d) = 0$ and $\gamma(q_d) = 0$. We thus have

$$\nabla\varphi(q_d) = 0 \ ,$$

and the Hessian therefore reduces to

$$\nabla^2\varphi = \frac{1}{\beta^2}\left[\beta\nabla^2\gamma - \gamma\nabla^2\beta\right] = \frac{\nabla^2\gamma}{\beta} \ .$$

Since $\gamma_d = \sum \left\|q_{ij} - c_{ij}\right\|^2$ we have

$$\nabla^2\varphi = \frac{2}{\beta} I \ ,$$

where I here denotes the $\frac{N(N-1)}{2}$-dimensional identity matrix. With $\beta(q_d) > 0$, it follows that q_d is a non-degenerate critical point. □

We can also show that there are no critical points on the obstacle boundary $\delta\mathscr{F}$:

Proposition 3. *(cf. [6, Proposition 3.2]) If workspace is valid, all critical points of φ are in interior of free space $\mathscr{F}_2$.*

Proof. Recalling (13)

$$\nabla\varphi = \frac{\beta\nabla\gamma - \gamma\nabla\beta}{\beta^2} \ ,$$

note that on the boundary of the collision-free configuration space (i.e., on $\delta\mathscr{F}$) we have $\beta = 0$, while both γ and $\nabla\beta$ (the latter by choosing $r > 0$ in (7)) do not vanish. Therefore, as $q \to \delta\mathscr{F}$, the magnitude of $\nabla\varphi$ blows up while it aligns with that of $\nabla\beta \neq 0$. □

In the following proposition, we show that critical points can be made to move arbitrarily close to the collision configurations by varying the tuning parameter κ, i.e we can push the critical points from $\mathscr{F}_2$ into $\mathscr{F}_0$.

Proposition 4. *(cf. [6, Proposition 3.3]) For every $\varepsilon > 0$ there exists a positive integer $N(\varepsilon)$ such that if $\kappa \geq N(\varepsilon)$ then there are no critical points of φ in $\mathscr{F}_2(\varepsilon)$.*

Proof. Since at a critical point $\nabla\varphi = 0$, from (13) it follows that at this configuration

$$\beta\nabla\gamma = \gamma\nabla\beta \ .$$

Substituting for γ from (6) we get

$$\beta\kappa{\gamma_d}^{\kappa-1}\nabla\gamma_d = {\gamma_d}^{\kappa}\nabla\beta \Rightarrow \beta\kappa\nabla\gamma_d = \gamma_d\nabla\beta \ .$$

Taking norms on both sides

$$\kappa\beta\,\|\nabla\gamma_d\| = \gamma_d\|\nabla\beta\| \ .$$

A sufficient condition for the above equality *not* to hold is

$$\kappa > \frac{\gamma_d\|\nabla\beta\|}{\beta\,\|\nabla\gamma_d\|} \ . \tag{15}$$

The gradient of the obstacle function defined in (7) is expressed as

$$\nabla\beta = \frac{2(2d+1)(-r+d+d^2)\,e^{-(-r+d+d^2)^2}}{2-e^{-(-r+d+d^2)^2}}\nabla d \ ,$$

where ∇d is the gradient of the distance function (9). We therefore have

$$\|\nabla\beta\| = \left|\frac{2(2d+1)(-r+d+d^2)\,e^{-(-r+d+d^2)^2}}{2-e^{-(-r+d+d^2)^2}}\right|\|\nabla d\| \ .$$

It is thus seen that both $\|\nabla\gamma_d\|$ (as long as q_d is away from collision configurations) and $\|\nabla\beta\|$ are bounded in $\mathscr{F}_2$. In fact, since $\nabla\beta$ is upper bounded everywhere in $\mathscr{F}_2$ and β attains a minimum of ε at the boundary of $\mathscr{F}_2$, it follows that the ratio $\frac{\|\nabla\beta\|}{\beta}$ is upper bounded in $\mathscr{F}_2$. In addition, since γ_d and $\|\nabla\gamma_d\|$ are continuous functions that must attain their extremum points in $\mathscr{F}_2$, and given that as $q \to q_d$, $\frac{\gamma_d}{\|\nabla\gamma_d\|} \to 0$, the bound for κ on the right hand side of (15) is finite anywhere in $\mathscr{F}_2$. □

Proposition 5. *(cf. [6, Proposition 3.4]) For any valid workspace, there exists an $\varepsilon_0 > 0$ such that φ has no local minimum in $\mathscr{F}_0$ as long as $\varepsilon < \varepsilon_0$.*

Proof. At the critical point where $\beta\nabla\gamma - \gamma\nabla\beta = 0$, (14) after plugging in (6) reduces to

$$\nabla^2\varphi = \frac{1}{\beta^2}\left[\beta\nabla^2{\gamma_d}^{\kappa} - {\gamma_d}^{\kappa}\nabla^2\beta\right] = \frac{{\gamma_d}^{\kappa-2}}{\beta^2}\left[\kappa\beta\left(\gamma_d\nabla^2\gamma_d + (\kappa-1)\nabla\gamma_d\nabla{\gamma_d}^T\right) - {\gamma_d}^2\nabla^2\beta\right] \tag{16}$$

Using the equation for the vanishing gradient at the critical point and (6),

$$\kappa\beta\nabla\gamma_d = \gamma_d\nabla\beta \ , \tag{17}$$

and taking the outer product on both sides,

$$(\kappa\beta)^2\nabla\gamma_d\nabla\gamma_d{}^T = \gamma_d{}^2\nabla\beta\nabla\beta^T \tag{18}$$

substituting for $\nabla\gamma_d\nabla\gamma_d{}^T$ in (16) yields

$$\nabla^2\varphi = \frac{\gamma_d^{\kappa-1}}{\beta^2}\left[\kappa\beta\nabla^2\gamma_d + \left(1-\frac{1}{\kappa}\right)\frac{\gamma_d}{\beta}\nabla\beta\nabla\beta^T - \gamma_d\nabla^2\beta\right] \tag{19}$$

From (17), by taking norms on both sides we have,

$$\kappa\beta = \frac{\gamma_d\|\nabla\beta\|}{\|\nabla\gamma_d\|}$$

substituting for $\kappa\beta$ in (19), we have

$$\nabla^2\varphi = \frac{\gamma_d^{\kappa-1}}{\beta^2}\left[\frac{\gamma_d\|\nabla\beta\|}{\|\nabla\gamma_d\|}\nabla^2\gamma_d + \left(1-\frac{1}{\kappa}\right)\frac{\gamma_d}{\beta}\nabla\beta\nabla\beta^T - \gamma_d\nabla^2\beta\right] = \frac{\gamma_d{}^\kappa}{\beta^2}\left[\frac{\|\nabla\beta\|}{\|\nabla\gamma_d\|}\nabla^2\gamma_d + \left(1-\frac{1}{\kappa}\right)\frac{1}{\beta}\nabla\beta\nabla\beta^T - \nabla^2\beta\right]$$

Let $\tilde{\upsilon}$ be any unit vector orthogonal to $\nabla\beta$, that is $\tilde{\upsilon}^T\cdot\nabla\beta = 0$ with $\|\tilde{\upsilon}\| = 1$. Then the quadratic form $\tilde{\upsilon}^T\left(\nabla^2\varphi\right)\tilde{\upsilon}$ expands to

$$\tilde{\upsilon}^T\left(\nabla^2\varphi\right)\tilde{\upsilon} = \frac{\gamma_d{}^\kappa}{\beta^2}\tilde{\upsilon}^T\left(\frac{\|\nabla\beta\|}{\|\nabla\gamma_d\|}\cdot\nabla^2\gamma_d - \nabla^2\beta\right)\tilde{\upsilon} = \frac{\gamma_d{}^\kappa}{\beta^2}\tilde{\upsilon}^T\left(2I\frac{\|\nabla\beta\|}{\|\nabla\gamma_d\|} - \nabla^2\beta\right)\tilde{\upsilon}\ , \tag{20}$$

where I is the $\frac{N(N-1)}{2}$-dimensional identity matrix. Now, the right hand of (20) to be negative, the following condition suffices

$$\max\left\{2\frac{\|\nabla\beta\|}{\|\nabla\gamma_d\|}\right\} - \min\left\{\sigma\left(\nabla^2\beta\right)\right\} < 0 \tag{21}$$

where $\sigma(\cdot)$ denotes the spectrum of a matrix. Toward this end, we recall (7),

$$\beta = \log(2 - a\,e^{-\left(-r+d+d^2\right)^2})$$

and we treat it as a function of two (not necessarily independent) variables, $x_1 \triangleq d$ and $x_2 \triangleq d^2$: $\beta = \beta(x_1,x_2)$. Then the first partial derivative of β with respect to q can be written

$$\nabla\beta = \frac{\partial\beta}{\partial x_1}\frac{\partial x_1}{\partial q} + \frac{\partial\beta}{\partial x_2}\frac{\partial x_2}{\partial q}\ .$$

while the second is

$$\nabla^2\beta = \frac{\partial^2\beta}{\partial x_1{}^2}\left(\frac{\partial x_1}{\partial q}\left(\frac{\partial x_1}{\partial q}\right)^T\right) + \frac{\partial\beta}{\partial x_1}\left(\frac{\partial^2 x_1}{\partial q^2}\right) + \frac{\partial^2\beta}{\partial x_2{}^2}\left(\frac{\partial x_2}{\partial q}\left(\frac{\partial x_2}{\partial q}\right)^T\right) + \frac{\partial\beta}{\partial x_2}\left(\frac{\partial^2 x_2}{\partial q^2}\right)$$

With reference to (7) we have,

$$\frac{\partial\beta}{\partial x_1} = \frac{\partial\beta}{\partial x_2} = \frac{2a\,e^{-(-r+x_1+x_2)^2}(-r+x_1+x_2)}{2-a\,e^{-(-r+x_1+x_2)}} > 0, \quad \text{for } x_1+x_2-r > 0\ , \tag{22}$$

and

$$\frac{\partial^2\beta}{\partial x_1{}^2} = \frac{\partial^2\beta}{\partial x_2{}^2} = \frac{-2a[a+2e^{(x_1+x_2-r)^2}\left(2r^2-4r(x_1+x_2)+2(x_1+x_2)^2-1\right)]}{\left(a-2e^{(x_1+x_2-r)^2}\right)^2}\ ,$$

which for $(x_1+x_2) \to r^+$, in the region where the critical point is expected, converges to $\frac{2a}{2-a} > 0$.

To determine $\min\left\{\sigma\left(\nabla^2\beta\right)\right\}$, we first note that with the partial derivatives of β being positive, and write

$$\min\left\{\tilde{\upsilon}^T\nabla^2\beta\,\tilde{\upsilon}\right\} = \frac{\partial^2\beta}{\partial {x_1}^2}\min\left\{\tilde{\upsilon}^T\frac{\partial x_1}{\partial q}\left(\frac{\partial x_1}{\partial q}\right)^T\tilde{\upsilon}\right\} + \frac{\partial\beta}{\partial x_1}\min\left\{\tilde{\upsilon}^T\frac{\partial^2 x_1}{\partial q^2}\tilde{\upsilon}\right\}$$
$$+\frac{\partial^2\beta}{\partial {x_2}^2}\min\left\{\tilde{\upsilon}^T\frac{\partial x_2}{\partial q}\left(\frac{\partial x_2}{\partial q}\right)^T\tilde{\upsilon}\right\} + \frac{\partial\beta}{\partial x_2}\min\left\{\tilde{\upsilon}^T\frac{\partial^2 x_2}{\partial q^2}\tilde{\upsilon}\right\} \quad . \tag{23}$$

The first and third term in (23) involve rank-one matrices made of the same vector, and thus their minimum eigenvalue is zero. With this observation, (23) we write

$$\min\left\{\tilde{\upsilon}^T\nabla^2\beta\,\tilde{\upsilon}\right\} \geq \frac{\partial\beta}{\partial x_1}\min\left\{\tilde{\upsilon}^T\frac{\partial^2 x_1}{\partial q^2}\tilde{\upsilon}\right\} + \frac{\partial\beta}{\partial x_2}\min\left\{\tilde{\upsilon}^T\frac{\partial^2 x_2}{\partial q^2}\tilde{\upsilon}\right\} \quad ,$$

and given that $\frac{\partial\beta}{\partial x_1} = \frac{\partial\beta}{\partial x_2}$,

$$\min\left\{\tilde{\upsilon}^T\nabla^2\beta\,\tilde{\upsilon}\right\} \geq \frac{\partial\beta}{\partial x_1}\left[\min\left\{\tilde{\upsilon}^T\left(\frac{\partial^2 x_1}{\partial q^2}+\frac{\partial^2 x_2}{\partial q^2}\right)\tilde{\upsilon}\right\}\right]$$

Rewrite $x_1 = d = \min\|q_{ij}\| = \sqrt{\min q_{ij}^T q_{ij}}$ and name the relative vector q_{ij} with the minimum norm w for convenience. Since d is assumed to be differentiable, around the critical point it will hold:

$$\frac{\partial^2 x_1}{\partial q^2} = \frac{\partial^2\left(\sqrt{\|w\|^2}\right)}{\partial^2 q^2} = \frac{\partial}{\partial q}\left(\frac{\partial}{\partial q}\left(\sqrt{\|w\|^2}\right)\right) = \frac{\partial}{\partial q}\left(\frac{[0\cdots 0\; w\; 0\cdots 0]^T}{\sqrt{\|w\|^2}}\right) = \frac{\partial}{\partial q}\left(\frac{[0\cdots 0\; w\; 0\cdots 0]^T}{\|w\|}\right)$$
$$= \frac{\|w\|\frac{\partial[0\cdots 0\; w\; 0\cdots 0]^T}{\partial q} - [0\cdots 0\; w\; 0\cdots 0]^T\frac{\partial\|w\|}{\partial q}}{\|w\|^2} = \frac{\|w\|\operatorname{diag}\{0,\dots,0,I_n,0,\dots,0\} - [0\cdots 0\; w\; 0\cdots 0]^T\left(\frac{[0\cdots 0\; w\; 0\cdots 0]^T}{\|w\|}\right)^T}{\|w\|^2}$$
$$= \frac{1}{\|w\|}\operatorname{diag}\left\{0,\dots,0,\frac{\|w\|^2 I_n - ww^T}{\|w\|^2},0,\dots,0\right\} \quad .$$

On the other hand,

$$\frac{\partial^2 x_2}{\partial q^2} = \frac{\partial^2\|w\|^2}{\partial q^2} = 2\operatorname{diag}\{0,\dots,0,I_n,0,\dots,0\} \quad .$$

Putting it together,

$$\min\tilde{\upsilon}^T\frac{\partial^2\beta}{\partial q^2}\tilde{\upsilon} \geq \frac{1}{\|w\|}\frac{\partial\beta}{\partial x_1}\min\left\{\tilde{\upsilon}^T\operatorname{diag}\left\{0,\dots,0,2\|w\|I_n + \frac{\|w\|^2 I_n - ww^T}{\|w\|^2},0,\dots,0\right\}\tilde{\upsilon}\right\}$$
$$= \frac{1}{\|w\|}\frac{\partial\beta}{\partial x_1}\min\left\{2\|w\| + 1 - \tilde{\upsilon}^T\frac{ww^T}{\|w\|^2}\tilde{\upsilon}\right\} = 2\frac{\partial\beta}{\partial x_1}$$

With reference to (21), a sufficient condition for the quadratic form $\tilde{\upsilon}^T\nabla^2\varphi\,\tilde{\upsilon}$ to be negative is that in $\mathscr{F}_0$ is that

$$\max_{\mathscr{F}_0}\frac{\|\nabla\beta\|}{\|\nabla\gamma_d\|} < \min_{\mathscr{F}_0}\left\{\frac{\partial\beta}{\partial d}\right\} \quad . \tag{24}$$

With $\frac{\partial d^2}{\partial q}$ growing as fast as $\|\nabla\gamma_d\|$, and with the magnitude of $\frac{\partial\beta}{\partial d}$ being regulated arbitrary through the choice of r and a, there is always an appropriate choice of parameters r and a so that (24) is satisfied. Practically, the further the desired formation encoded in γ_d is from $\mathscr{F}_0$ (the near collision configurations), the easier (24) is to satisfy. □

To show that φ is Morse we use the same lemma as [6], which says that the non-singularity of a linear operator folows from the fact that its associated quadratic form is sign definite on complementary subspaces.

Lemma 2. *([6, cf. Lemma 3.1]) Let $\mathbb{R}^n = \mathscr{P} \oplus \mathscr{N}$ and let the symmetric matrix $Q \in \mathbb{R}^{n\times n}$ define a quadratic form on $\mathbb{R}^n$.*

$$\xi(\upsilon) \stackrel{\triangle}{=} \upsilon^T Q \upsilon$$

If $\xi|_{\mathscr{P}}$ (the restriction of ξ in $\mathscr{P}$) is positive definite and $\xi|_{\mathscr{N}}$ is negative definite, then Q is non-singular and

$$\text{index}(Q) = \dim(\mathscr{P}) \ .$$

Let $\xi_q(\upsilon)$ denote $\upsilon^T \nabla^2 \varphi(q) \upsilon$, where $q \in \mathscr{F}_0(\varepsilon)$, and υ is a vector in the tangent space T_q of $\mathscr{F}_0(\varepsilon)$ at q.

Proposition 6. *(cf. [6, Proposition 3.6]) There exists an $\varepsilon_2 > 0$ such that for every $\varepsilon < \varepsilon_2$ at each critical point of φ in $\mathscr{F}_0(\varepsilon)$, there is a direct sum decomposition $T_q = \mathscr{P}_q \oplus \mathscr{N}_q$, for which $\xi_q|_{\mathscr{P}_q}$ is positive definite, $\xi_q|_{\mathscr{N}_q}$ is negative definite, and* $\dim(\mathscr{P}_q) = 1$.

Proof. Assume $q \in \mathscr{B}(\varepsilon)$ where $\mathscr{B}(\varepsilon) = \{q : 0 < \beta < \varepsilon\}$. Define $\mathscr{P}_q = \text{span}\{\nabla\beta(q)\}$ and let $\mathscr{N}_q$ be the orthogonal component of $\mathscr{P}_q$ in T_q. In Proposition 5 it was shown $\xi_q|_{\mathscr{N}_q}$ is negative definite as long as $\varepsilon < \varepsilon_0$; the goal now is to show that $\xi_q|_{\mathscr{P}_q}$ is positive definite.

Let $\upsilon = \frac{\nabla\beta}{\|\nabla\beta\|}$. From (13) and (6), at a critical point we recall (17):

$$\kappa\beta\nabla\gamma_d = \gamma_d\nabla\beta \ .$$

Just as in the proof of Proposition (5), taking outer products in (17) yields (18), and substituting in (16) we get (19). Here, we instead take squared norms on both sides of (17) to solve for $\kappa\beta$ as follows

$$(\kappa\beta)^2\|\nabla\gamma_d\|^2 = \gamma_d^2\|\nabla\beta\|^2 \Rightarrow \kappa\beta = \frac{\gamma_d}{4\kappa\beta}\|\nabla\beta\|^2 \ , \tag{25}$$

exploiting the fact that $\|\nabla\gamma_d\|^2 = {\gamma_d}^2$. Substituting for $\kappa\beta$ as found in (25) in (16)

$$\begin{aligned}\nabla^2\varphi(q) &= \frac{{\gamma_d}^{\kappa-1}}{\beta^2}\left[\frac{\gamma_d}{4\kappa\beta}\|\nabla\beta\|^2\nabla^2\gamma_d + \left(1-\frac{1}{\kappa}\right)\frac{\gamma_d}{\beta}\nabla\beta\nabla\beta^T - \gamma_d\nabla^2\beta\right]\\ &= \frac{{\gamma_d}^{\kappa}}{\beta^2}\left[\frac{\|\nabla\beta\|^2}{2\kappa\beta}I + \left(1-\frac{1}{\kappa}\right)\frac{1}{\beta}\nabla\beta\nabla\beta^T - \nabla^2\beta\right]\end{aligned}$$

Evaluating $\xi_q|_{\mathscr{P}_q}$ with $\upsilon = \frac{\nabla\beta}{\|\nabla\beta\|}$

$$\begin{aligned}\upsilon^T\nabla^2\varphi\,\upsilon &= \frac{{\gamma_d}^{\kappa}}{\beta^2}\left[\frac{1}{2\kappa\beta}\|\nabla\beta\|^2 + \left(1-\frac{1}{\kappa}\right)\frac{1}{\beta}\upsilon^T\nabla\beta\nabla\beta^T\upsilon - \upsilon^T\nabla^2\beta\,\upsilon\right]\\ &= \frac{{\gamma_d}^{\kappa}}{\beta^2}\left[\frac{1}{2\kappa\beta}\|\nabla\beta\|^2 + \left(1-\frac{1}{\kappa}\right)\frac{1}{\beta}\left(\frac{\nabla\beta^T\nabla\beta}{\|\nabla\beta\|}\right)^2 - \upsilon^T\nabla^2\beta\,\upsilon\right]\\ &= \frac{{\gamma_d}^{\kappa}}{\beta^2}\left[\frac{2\kappa-1}{2\kappa\beta}\|\nabla\beta\|^2 - \upsilon^T\nabla^2\beta\,\upsilon\right]\end{aligned}$$

Thus, $\upsilon^T\nabla^2\varphi\upsilon$ is positive, if

$$\min\left\{\frac{2\kappa-1}{2\kappa\beta}\|\nabla\beta\|^2\right\} - \max\left\{\upsilon^T\nabla^2\beta\,\upsilon\right\} > 0 \tag{26}$$

Expanding $\nabla^2\beta$ as in the proof of Proposition 5,

$$\begin{aligned}\max\left\{\upsilon^T\nabla^2\beta\,\upsilon\right\} \le{}& \frac{\partial^2\beta}{\partial {x_1}^2}\max\left\{\upsilon^T\frac{\partial x_1}{\partial q}\left(\frac{\partial x_1}{\partial q}\right)^T\upsilon\right\} + \frac{\partial\beta}{\partial x_1}\max\left\{\upsilon^T\frac{\partial^2 x_1}{\partial q^2}\upsilon\right\}\\ &+ \frac{\partial^2\beta}{\partial {x_2}^2}\max\left\{\upsilon^T\frac{\partial x_2}{\partial q}\left(\frac{\partial x_2}{\partial q}\right)^T\upsilon\right\} + \frac{\partial\beta}{\partial x_2}\max\left\{\upsilon^T\frac{\partial^2 x_2}{\partial q^2}\upsilon\right\}\\ \le{}& \frac{\partial^2\beta}{\partial {x_1}^2}(1+4\|w\|^2) + \frac{\partial\beta}{\partial x_1}\left(\frac{2\|w\|+1}{\|w\|}\right) \ ,\end{aligned} \tag{27}$$

The terms $\frac{\partial^2\beta}{\partial {x_1}^2}$ and $\frac{\partial\beta}{\partial x_1}$ that appear in the right hand side of (27), are both positive and bounded when $q \in \mathscr{B}(\varepsilon)$, with bounds dependent on parameters c and d_0 as indicated in (8). In view of (26), it therefore suffices to show that an appropriately small choice of ε can make the first term of (26) sufficiently large so as the whole difference is positive. To this end, note first that $\min\left\{\frac{2\kappa-1}{2\kappa}\frac{\|\nabla\beta\|^2}{\beta}\right\} > 0$ when $\kappa > \frac{1}{2}$, and that $\frac{2\kappa-1}{2\kappa}$ is strictly increasing and upper bounded by 1 as $\kappa \to \infty$. The factor involving β and its gradient can be bounded as follows. First recall that

$$\nabla\beta = \frac{\partial\beta}{\partial x_1}\cdot\frac{\partial x_1}{\partial q} + \frac{\partial\beta}{\partial x_2}\cdot\frac{\partial x_2}{\partial q} = \frac{\partial\beta}{\partial x_1}\frac{w}{\|w\|} + 2\frac{\partial\beta}{\partial x_2}w \enspace .$$

Then, given that $\frac{\partial\beta}{\partial x_1} = \frac{\partial\beta}{\partial x_2}$, write

$$\begin{aligned}\min\left\{\frac{\|\nabla\beta\|^2}{\beta}\right\} &= \min_{\mathscr{F}_0}\left\{\frac{1}{\beta}\left(\frac{\partial\beta}{\partial x_1}\right)^2\left(2w+\frac{w}{\|w\|}\right)^2\right\} \geq (4{d_0}^2+4d_0+1)\min\left\{\frac{1}{\beta}\left(\frac{\partial\beta}{\partial x_1}\right)^2\right\} \\ &= \frac{4{d_0}^2+4d_0+1}{\varepsilon}\min_{\mathscr{F}_0}\left\{\left(\frac{\partial\beta}{\partial x_1}\right)^2\right\} \enspace ,\end{aligned}$$

since $\max_{\mathscr{F}_0}\beta = \varepsilon$. On the other hand, $\frac{\partial\beta}{\partial x_1}$ is lower bounded in $\mathscr{F}_0$, and in fact for sufficiently large c, namely

$$c \geq \frac{1-\sqrt{64{d_0}^4+128{d_0}^3+96{d_0}^2+32d_0+5}}{2(2d_0+1)} \enspace ,$$

and sufficiently small ε, this lower bound is exactly c. It thus follows that for $\varepsilon \to 0$,

$$\min_{\mathscr{F}_0}\left\{\frac{\|\nabla\beta\|^2}{\beta}\right\} \geq \frac{c^2(4{d_0}^2+4d_0+1)}{\varepsilon} \to \infty \enspace ,$$

and (26) can always be satisfied. □

3.4.2 When the distance is non-differentiable

To analyze the stability of the formation in the case when the distance function d is not differentiable, it is necessary to introduce some terminology and results from nonsmooth analysis and control theory. The treatment of this section is centered around an invariance principle result for differential inclusions that is due to Ryan [17]. Based on this nonsmooth invariance principle, and the characterization of the generalized gradient of the distance function discussed in Section 3.3, we conclude the (almost[1]) global attractive properties of the desired formation configuration.

When the right-hand side of the differential equation (4) is only piecewise continuous, the solution $p(t)$ (and therefore $q(t)$) for $t \geq 0$, cannot be defined in the classical sense anymore. Instead, it takes the form of a Filippov solution to the differential inclusion

$$\dot{q} = F\big(q(t)\big), \quad q(t) \in \mathscr{F} \enspace , \tag{28}$$

where $F(\cdot)$ is the Filippov set-valued map which in our case will be given by

$$F(q) = \overline{\text{co}}\left\{\lim_{k\to\infty} Bu\big(q[k]\big) : q[k] \to q, q' \neq \mathscr{M}\right\} \enspace ,$$

where $\overline{\text{co}}$ denotes the convex closure of a set, $q[k]$ is any sequence of points at which φ is differentiable, and $\mathscr{M}$ can be any set of measure zero. Such Filippov solutions are absolutely continuous curves $q : [0,\infty) \to \mathscr{F}$, which satisfy (28) for almost all $t \in [0,\infty)$. For the existence and uniqueness properties of Filippov solutions, we refer to [18, 19].

We say that a solution of $q(t)$ of (28) is maximal [17], if it can be extended for all positive time. The solution $q(t)$ is *precompact* if it is maximal, and the closure of the image of any interval $[0,\omega)$ under $q(t)$, written $\text{cl}\left(q\big([0,\omega)\big)\right)$, is a compact set in $\mathscr{F}$. Essentially, a precompact Filippov solution is the nonsmooth equivalent of a differentiable solution that is bounded within a compact set of the state space.

[1]With the exception of a set of configurations of measure zero, that correspond to the unstable critical points of the potential function.

Definition 5 (Limit set [17]). *Let $q(t)$ be a maximal solution of* (28). *A point $z \in \mathscr{F}$ is an ω-limit point of $q(t)$ if there exists an increasing sequence $\{t_n\} \subset [0,\omega)$ such that $t_n \to \omega$ with $n \to \infty$ implies that $q(t_n) \to z$. The set $\Omega(q)$ of all limit points of the solution $q(t)$ is the ω-limit set of $q(t)$.*

Definition 6 (Weak invariance [17]). *With respect to* (28)*, a set $\mathscr{S} \subset \mathscr{F}$ is weakly invariant if for each $q(0) \in \mathscr{S} \cap \mathscr{F}$, there exists at least one maximal solution $q(t)$ of* (28) *with $\omega = \infty$ and with $q\big([0,\omega)\big) \in \mathscr{S}$.*

The following fact is the nonsmooth equivalent of the known result that states that differentiable solutions bounded in a compact set have a non-empty attractive invariant set.

Proposition 7 ([17]). *If $q(t)$ is a precompact solution of* (28) *then $\Omega(q)$ is a non-empty, compact, connected subset of $\mathscr{F}$. Moreover, $\Omega(q)$ is the smallest closed set approached by $q(t)$ and it is weakly invariant.*

Theorem 1 ([17]). *Let $V : \mathscr{F} \to \mathbb{R}$ be locally Lipschitz. Define (the map)*

$$y : \mathscr{F} \to \mathbb{R}; z \mapsto y(z) \triangleq \max\{V^\circ(z;v) \mid v \in F(q)\} \ .$$

Suppose $\mathscr{Y} \subset \mathscr{F}$ is nonempty and that $y(z) \leq 0$ for all $z \in \mathscr{Y}$. If $q(t)$ is a precompact solution of (28) *then, for some constant $c \in V(\mathrm{cl}(\mathscr{Y}) \cap \mathscr{F})$, q approaches the largest weakly-invariant set in $\Sigma \cap V^{-1}(c)$ where*

$$\Sigma = \{z \in \mathrm{cl}(\mathscr{Y}) \cap \mathscr{F} \mid y(z) \geq 0\} \ .$$

Take $\varphi(q)$ as $V(q)$. Function φ is continuous and locally Lipschitz. At every q where $d(q)$ is differentiable, $\dot{q} = -\nabla\varphi$, so $V^\circ(q,\dot{q}) = \dot{V}(q) = -(\nabla\varphi)^2$. Where $d(q)$ is not differentiable, using the chain rule [14, Theorem 2.3.9(ii)], we get

$$\partial\varphi = \frac{1}{\beta^2}\left(\beta\kappa{\gamma_d}^{\kappa-1}\,\nabla\gamma_d - {\gamma_d}^{\kappa}\,\frac{\partial\beta}{\partial d}\,\partial d\right) \ . \tag{29}$$

.

As one approaches collision configurations, $\beta \to 0$, (29) suggests that $\partial\varphi \to -\gamma\frac{\partial\beta}{\partial d}\partial d$, which does not contain the zero vector. Therefore, the arguments of Proposition 3 apply.

If $0 \in \partial\varphi$, (29) yields

$$\frac{2\beta\kappa}{\gamma_d\frac{\partial\beta}{\partial d}}(q - q_d) \in \partial d \ .$$

Away from the destination and collision configurations, β and $q - q_d$ are bounded from below by ε and $\sqrt{\varepsilon}$, respectively. In any domain containing q_d, the terms in the denominator, γ_d and $\frac{\partial\beta}{\partial d}$ are upper bounded too. Thus, repeating the argument for a sufficiently large κ establishes the statement of Proposition 4, namely that the critical points of φ can be pushed in the ε-neighborhood of the collision configurations.

To extend Proposition 5 to the case where d is not differentiable, we do not have employ the Hessian of φ. The next proposition is a dual to [16, Proposition 3], with slightly more relaxed conditions.

Proposition 8. *(cf. [16, Proposition 3]) Let $V : \mathscr{F} \to \mathbb{R}$ be a mapping where $V(q) = \max_i f_i(q)$ where all f_i are smooth functions. If $0 \in \mathrm{int}\,(\partial V(y))$, then y is a local minimum of V.*

Proof. If f_i are smooth, then they are regular and from the pointwise maxima Theorem [14] $\partial V(q) = \mathrm{co}\{\nabla f_i : i \in I(q)\}$, where $I(q)$ is the set of indices for which $f_i(q) = V(q)$. Based on Lemma 1, for the origin to belong in the interior of $\partial V(q)$, it is necessary that there exists an i such that $\langle\nabla f_i(q), w\rangle > 0$, for every $w \in T_q\mathbb{F}$ (the tangent space of $\mathscr{F}$ at q). Following the same reasoning as in the proof of [16, Proposition 3], there must be a function f_i that increases along any direction w from point q. This implies that $V(q)$ is a local minimum. □

The proofs of the following two statements can be constructed in a straightforward manner, similarly to how the proposition above was established from [16, Proposition 3].

Proposition 9. *([16, Proposition 4]) At a saddle point, V is nonsmooth, and the origin is contained in $\delta\partial V$.*

Proposition 10. *([16, cf. Proposition 5]) At a local maximum of V, $0 = \partial\varphi$.*

In view of the above statements, there is a simple argument that shows why nonsmooth critical points of φ can only be saddles: note that the interior of a convex hull of $k \leq n$ vectors in an n-dimensional vector space is empty; all points are on the boundary. In the case of $d(q)$, even when every q_{ij} is one-dimensional, the number of any nontrivial different distances between agents cannot be more than the dimension of $\mathscr{F}$; in other words, and with reference to (11), $|\{(i,j) : i = 1,\ldots,N, j = 1,\ldots,N, i \neq j\}| \leq \dim\mathscr{F}$. With this in mind, it is clear that any point of ∂d (or $\partial\varphi$, for that matter) will be a boundary point, and according to Proposition 9, a nonsmooth critical point of φ will necessarily be a saddle point.

We now design the right-hand side of (28) by setting

$$Bu = \arg\min_{\zeta\in-\partial\varphi} \varphi^{\circ}(q;\zeta) \ .$$

This happens if the agent inputs are chosen as

$$u = B^{\dagger} \arg\min_{\zeta\in-\partial\varphi} \varphi^{\circ}(q;\zeta) \ , \tag{30}$$

where $B^{\dagger} = (B^T B)^{-1} B^T$ is the Moore-Penrose generalized inverse of the incidence matrix of the formation graph $\mathscr{G}$, B. With this choice of control inputs,

$$F(q) = \overline{\mathrm{co}}\{-\lim_{q'\to q} \nabla\varphi(q'),\ \arg\min_{\zeta\in-\partial\varphi} \varphi^{\circ}(q;\zeta) : \varphi \text{ differentiable at } q'\} \ ,$$

implying that $F(q)$ is either a singleton, if φ is differentiable, or the convex hull of the element of the negated generalized gradient that produces the smallest generalized derivative, and the negated gradients of φ around the point of non-differentiability. Note that at the saddle points of φ, the directions along which φ is increasing are orthogonal to the negated gradient around the saddle point. The only way these directions may appear in the convex hull $F(q)$ therefore, is if they coincide with the element of the generalized gradient of φ at q which minimizes φ°. Since, however, the only type of critical point φ can have at $q \neq q_d$ is a saddle, there are always be tangent vector directions in $-\partial\varphi$ that produce negative generalized derivatives of φ, and thus vector $\zeta \in -\nabla\varphi$ is such that $\varphi^{\circ}(q;\zeta) \leq 0$. In either case, the map $y(q)$ of Theorem 1 is such that $y(q) \leq 0$ whenever q is a saddle point. The fact that $q(t)$ is a pre-compact solution of (28) is established by invoking the invariance of the level sets of φ: since the directional derivative of φ along the system's trajectories is non-positive, level sets that contain the initial configuration $q(0)$ are positively invariant (strongly, in the sense that all solutions do not increase φ). With φ being positive definite in $\mathscr{F}$ once the coordinate system is shifted at q_d, this implies that the invariant level sets are also compact. The closedness of $q(t)$ is given by definition [17]; the boundedness follows by the positive invariance of the level sets. According to Theorem 1 therefore, $q(t)$ approaches the largest weakly-invariant set in the level set of φ for which $y(q) = 0$. The points in $\mathscr{F}$ that satisfy $y(q)$ are the saddle points of φ and q_d. Since the saddle points of φ have attraction regions of measure zero, we conclude that $q(t)$ converges to q_d from *almost everywhere* in $\mathscr{F}$.

This discussion is summarized in the following statement:

Theorem 2. *The solutions of the dynamical system*

$$\dot{q} = \arg\min_{\zeta\in-\partial\varphi} \varphi^{\circ}(q;\zeta) \ ,$$

approach q_d which is such that $\varphi(q_d) = 0$, from almost everywhere in $\mathscr{F}$ (with the exception of a set of measure zero).

4 Conclusion

Generalizations the classic navigation function approach by Rimon and Koditschek [7] to a multi-agent setting, typically employ the product of positive semidefinite functions as a metric of the distance of the system from collision configurations. We indicate that within the same context, and following the proof techniques that have appeared in literature, one may fail to establish the Morse character of the function's critical points, on which the convergence properties of the potential field that the function generates, rely upon. In this document, we propose an alternative construction for the potential function for which we can establish the properties sufficient for the (almost global) convergence of the system to the desired configuration. The new construction is a nonsmooth positive definite function, and the proof techniques are based on nonsmooth analysis, and control theory for dynamical systems expressed in the form of (finite-dimensional) differential inclusions.

Acknowledgments

This work is supported by NSF under grant #0822845.

References

[1] R. Olfati-Saber, W. B. Dunbar, and R. M. Murray, "Cooperative control of multi-vehicle systems using cost graphs and optimization," in *In Proceedings of the American Control Conference*, 2003, pp. 1–15.

[2] T. Balch and R. C. Arkin, "Behavior-based formation control for multirobot teams," *IEEE Transactions on Robotics and Automation*, vol. 14, no. 6, pp. 926–939, December 1998.

[3] P. Ogren and N. Leonard, "Obstacle avoidance in formation," in *Robotics and Automation, 2003. Proceedings. ICRA '03. IEEE International Conference on*, vol. 2, 14-19 2003, pp. 2492 – 2497 vol.2.

[4] H. Tanner, G. Pappas, and V. Kumar, "Leader-to-formation stability," *Robotics and Automation, IEEE Transactions on*, vol. 20, no. 3, pp. 443 – 455, june 2004.

[5] M. Egerstedt and X. Hu, "Formation constrained multi-agent control," *Robotics and Automation, IEEE Transactions on*, vol. 17, no. 6, pp. 947 –951, dec 2001.

[6] D. E. Koditschek and E. Rimon, "Robot navigation functions on manifolds with boundary," *Advances in Applied Mathematics*, vol. 11, pp. 412–442, 1990.

[7] E. Rimon and D. Koditschek, "Exact robot navigation using artificial potential functions," *IEEE Transactions on Robotics and Automation*, vol. 8, no. 5, pp. 501–518, October 1992.

[8] S. G. Loizou and K. J. Kyriakopoulos, "Closed loop navigation for multiple holonomic vehicles," in *Proc. of IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2002, pp. 2861–2866.

[9] M. M. Zavlanos and K. J. Kyriakopoulos, "Decentralized motion control of multiple mobile agents," in *Proceedings of the 11th Mediterranean Conference on Control and Automation, Rhodes,Greece*, 2003.

[10] D. Dimarogonas, K. Kyriakopoulos, and D. Theodorakatos, "Totally distributed motion control of sphere world multi-agent systems using decentralized navigation functions," in *Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on*, 15-19 2006, pp. 2430 –2435.

[11] M. De Gennaro and A. Jadbabaie, "Formation control for a cooperative multi-agent system using decentralized navigation functions," in *American Control Conference, 2006*, 14-16 2006, p. 6 pp.

[12] H. G. Tanner and A. Kumar, "Towards decentralization of multi-robot navigation functions," in *Proceedings of the IEEE International Conference on Robotics and Automation*. Barcelona, Spain: IEEE, 2005, pp. 4143–4148.

[13] ——, "Formation stabilization of multiple agents using decentralized navigation functions," in *Robotics: Science and Systems I*, S. Thrun, G. Sukhatme, S. Schall, and O. Brock, Eds. Springer, 2005, pp. 49–56.

[14] F. H. Clarke, *Optimization and nonsmooth analysis*. SIAM, 1990.

[15] S. I. Dubov, "On the generalized gradient of the distance function," *Journal of mathematical sciences*, vol. 100, no. 6, pp. 2593–2600, 2000.

[16] H. Choset, "Nonsmooth analysis, convex analysis, and their applications to motion planning," *International Journal of Computational Geometry & Applications*, vol. 9, no. 4-5, pp. 447–469, 1999.

[17] E. P. Ryan, "An integral invariance principle for differential inclusions with applications in adaptive control," *SIAM Journal on Control and Optimization*, vol. 36, no. 3, pp. 960–980, 1998.

[18] A. F. Filippov, *Differential Equations with Discontinuous Righthand Sides*. Kluwer Academic, 1988.

[19] J. Cortes, "Discontinuous dynamical systems," *Control Systems Magazine*, pp. 36–73, June 2008.