Formation Control of Autonomous Robots Based on Cooperative Behavior

Florian Dörfler and Bruce Francis

Abstract— **This paper considers the formation control problem for autonomous robots, where the target formation is specified as a minimally rigid formation. A distributed control law based on potential functions is derived from a directed sensor graph and relies on the graph matrices only. By methods of inverse optimality a certain class of sensor graphs is identified that is related to a cooperative behavior among the robots. These graphs are referred to as cooperative graphs, and undirected graphs, directed cycles, and directed open chain graphs can be identified as such graphs. Cooperative graphs admit a local stability result of the target formation together with a guaranteed region of attraction, that depends on the rigidity properties of the formation.**

I. Introduction

The control of a network of autonomous mobile robots is an interesting instance of distributed control. This paper contributes to formation control, that is, getting autonomous mobile robots into a formation when each robot has only locally sensed information about the others.

In this problem graph theory plays a natural role, both to define a formation and to describe the sensor relationships–who can "see" whom. Early work used the graph theoretic concept of rigidity [1], [2]. Rigidity was first introduced as a design tool to construct undirected graphs, but more recent references extend the rigidity concept to directed graphs [3] and employ it as an analysis tool for the stability of an undirected formation [4]. For an overview of the applications of rigidity in formation control, see [5].

Typically a potential function approach is used in the formation control problem [2], [6], [4], [7]. Other approaches first construct distributed control laws and then later relate them to potential functions [8], [9], [10], [11]. Historically the idea of potential functions emerged for undirected graphs, but has recently been extended to directed graphs [4]. A typical stability analysis of the desired formation usually makes use of the potential function as a Lyapunov function in combination with the invariance principle, that usually does not guarantee convergence of the robots to a particular geometric configuration. Reference [4] specifies the desired formation as an infinitesimally rigid framework and proves its stability by methods of linearization and center manifold theory and making use of infinitesimal rigidity.

We follow the approach of [4], specifying a minimally rigid target formation and deriving a potential function based control law for a general directed sensor graph. Although the control laws are designed for each robot independently, for certain graphs the resulting closed-loop dynamics reflect a cooperative behavior among the robots. This becomes obvious for undirected graphs, where the overall dynamics can be related to an optimal control problem based on rigidity and the target formation, but also for certain directed graphs where the multi-robot system shows the same behavior. This cooperative behavior combined with the infinitesimal rigidity of the target formation then directly paves the way for our stability analysis. Compared to [4], which has an infinitesimal result, we pursue a Lyapunov-type analysis in the space of inter-agent distances leading to an exponential stability result with a guaranteed region of attraction.

The paper is organized as follows: In Section II some graph theoretic preliminaries are given and the formation control problem is specified. In Section III a potential function based control law is proposed and in Section IV the cooperative behavior resulting from this control is specified. A stability result is presented in Section V and finally some some conclusions and an outlook are given in Section VI.

II. The Formation Control Problem

A. Basic Notation and Definitions

A *directed graph* $\mathcal{G} = (V, E)$ consists of a finite set of *nodes* $V = \{1, \ldots, n\}$ and *edges* $E \subset V \times V$. We assume the edges are ordered, that is, $E = \{1, \ldots, m\}$, and exclude the possibility of self loops—an edge from a node to itself. The *neighbour set* $\mathcal{N}_i$ of the node i is the set of all nodes j where there is an edge from i to j. In this case i is the *source node* of the edge and j is the *sink node*; the edge is then also called an *outgoing edge* of node i. With a directed graph $\mathcal{G}$ we associate two different matrices. The matrix relating the nodes to the edges is called the *incidence matrix* $H = \{h_{ij}\} \in \mathbb{R}^{m \times n}$ of the graph and is defined component-wise as $h_{ij} = 1$ if node j is the sink node of edge i and as $h_{ij} = -1$ if node j is the source node of edge i; all other elements are zero. Let $\mathbf{1}$ denote the vector of 1's. Then $\mathbf{1}$ lies in the kernel of H. Furthermore, $\operatorname{rank}(H) = n - 1$ iff $\mathcal{G}$ is connected ([12], Proposition 4.3) and thus in this case $\operatorname{Ker}(H) = \operatorname{span}\{\mathbf{1}\}$. For the remainder of this work we will assume that all graphs are connected. In addition to the incidence matrix, of interest for us is also the *outgoing edge matrix* $O = \{o_{ij}\} \in \mathbb{R}^{n \times m}$ with components $o_{ij} = -1$ if node i has outgoing edge j and $o_{ij} = 0$ else.

We take an *undirected graph* to be a special case of a directed graph, where, for every i, j, if there's an edge from i to j, there is also a reverse one from j to i. For such a graph the edges can be ordered such that the edge set is partitioned

This work was supported by the Landesstiftung Baden-Württemberg.

Florian Dörfler is with the Center for Control, Dynamical Systems and Computation, University of California at Santa Barbara, Santa Barbara, CA 93106, USA. Bruce Francis is with the Electrical and Computer Engineering Department, University of Toronto, M5S3G4 Toronto, Canada doerfler@engineering.ucb.edu, bruce.francis@utoronto.ca

978-3-9524173-9-3

as $E = \{E_+, E_-\}$, where elements of E_- have the opposite orientation from the elements of E_+. The incidence matrix is then obtained as $H = \left[H_u^T, -H_u^T\right]^T$ and the outgoing edge matrix as $O = \left[O_u, O_u - H_u^T\right]$, where H_u and O_u are the incidence and outgoing edge matrix corresponding to E_+.

B. Frameworks and Infinitesimal Rigidity

Rigidity is a property that refers to undirected graphs. Let $\mathcal{G} = (V, E)$ be an undirected graph with n nodes and m edges. We embed $\mathcal{G}$ into the plane $\mathbb{R}^2$ by assigning to each node i a point $z_i \in \mathbb{R}^2$. A *framework* is then a pair $(\mathcal{G}, z)$, where[1] $z = (z_1, \ldots, z_n) \in \mathbb{R}^{2n}$.

It is convenient to introduce the notation $\hat{A} := A \otimes I_2$, where A is a matrix, I_2 is the two dimensional identity matrix, and $\otimes$ is the Kronecker product. In rigidity we don't want to count link lengths twice, so we define the vector $e = \hat{H}_u z$ and partition it as $e = (e_1, \ldots, e_{m/2})$, $e_i \in \mathbb{R}^2$. Thus e is the concatenated vector of links with the orientation of the edges in E_+. We define the *rigidity function* $r_\mathcal{G} : \mathbb{R}^{2n} \to \mathbb{R}^m$ as

$$r_\mathcal{G}(z) = \frac{1}{2}(\|e_1\|^2, \ldots, \|e_m\|^2), \tag{1}$$

where the norm is the standard Euclidean norm. The framework $(\mathcal{G}, z)$ is then said to be *rigid* if there is an open neighbourhood $\mathcal{U}$ of z such that, if $q \in \mathcal{U}$ and $r_\mathcal{G}(z) = r_\mathcal{G}(q)$, then $(\mathcal{G}, z)$ is congruent to $(\mathcal{G}, q)$.

Although rigidity is a very intuitive concept, its definition does not provide a condition that is easy to check. Luckily there is a linearized version of the concept that involves a rank condition on the Jacobian $\partial r_\mathcal{G}(z)/\partial z \in \mathbb{R}^{m \times 2n}$, called the *rigidity matrix* of the framework $(\mathcal{G}, z)$. A framework is *infinitesimally rigid* if the rank of the rigidity matrix equals $2n - 3$ (it can't be more). If $(\mathcal{G}, z)$ is infinitesimally rigid, so is $(\mathcal{G}, z')$ for a generic (open and dense) set of z'. If a framework is infinitesimally rigid, then it is also rigid but the converse is not necessarily true. Finally, note that an infinitesimally rigid framework must have at least $2n - 3$ links. If it has exactly $2n-3$ links, then we say it is *minimally rigid*.

The links e obtained by the positions z as $e = \hat{H}_u z$ are not independent, but are located in the space $\mathrm{Im}\hat{H}_u$, the image (range space) of $\hat{H}_u$, considered as a map; this space is called the *link space*. It is a subspace of $\mathbb{R}^{2n}$ whose normal vectors are in $\mathrm{Ker}\left(\hat{H}_u^T\right)$ and thus correspond to cycles in the graph ([12], Theorem 4.5). The rigidity function $r_\mathcal{G}(z)$, that maps the positions to the squared link lengths, then induces the function $v : \mathrm{Im}\hat{H}_u \to \mathbb{R}^m$ via

$$v(e) = \frac{1}{2}(\|e_1\|^2, \ldots, \|e_m\|^2).$$

Thus $r_\mathcal{G}(z) = v(\hat{H}_u z)$. Defining $D(e) = \mathrm{diag}(e_k)$, we arrive at the following simple form for the rigidity matrix:

$$\frac{\partial r_\mathcal{G}(z)}{\partial z} = \frac{\partial v(e)}{\partial e}\frac{\partial e}{\partial z} = D(e)^T \hat{H}_u =: R_\mathcal{G}(e). \tag{2}$$

[1] Vectors are written either as n-tuples or columns vectors.

Infinitesimal rigidity is a property of a framework, but since it obviously depends only on the links, we can equivalently talk about the infinitesimal rigidity of a pair $(\mathcal{G}, e)$, which we refer to as a *formation*, and to $R_\mathcal{G}(e)$ as its rigidity matrix.

In the general case of a *directed* graph $\mathcal{G}$, everything carries over naturally: a framework $(\mathcal{G}, z)$ is defined in the same way, and the equation relating e and z is $e = \hat{H}z$, where H is the incidence matrix of the directed graph $\mathcal{G}$. The link space is thus again $\mathrm{Im}\hat{H}$ and the pair $(\mathcal{G}, e)$ is a formation. We say that the framework, respectively the formation, resulting from a directed graph is infinitesimally, respectively minimally, rigid if the corresponding undirected framework is infinitesimally, respectively minimally, rigid.

C. Distributed Control of Autonomous Robots

For our purposes an autonomous robot is a wheeled, actuated vehicle in the plane that has no communication devices and is equipped with a compass and an onboard camera. The camera is assumed to have no sensing limitations regarding range and direction. By methods of feedback linearization many standard models of wheeled robots can be transformed to a *kinematic point* whose motion is fully actuated, that is, it has dynamics $\dot{z}_i = u_i$, where $z_i \in \mathbb{R}^2$ is the position of robot i in the plane and $u_i \in \mathbb{R}^2$ is a direct velocity command. Altogether we consider n such robots and with the concatenated vectors $z = (z_1, \ldots, z_n)$ and $u = (u_1, \ldots, u_n)$ in $\mathbb{R}^{2n}$ the overall dynamics are simply $\dot{z} = u$.

The topology of information exchange between the robots is expressed in the *sensor graph* $\mathcal{G}$, a directed graph with n nodes and m edges. Node i of the graph $\mathcal{G}$ corresponds to robot i, and a directed edge from robot i to robot $j \in \mathcal{N}_i$ means that robot i can sense robot j via its onboard camera. Thus robot i can sense the relative distance and direction of robot j in a global reference frame (due to the compass). If we consider the concatenated vector $e = (e_1, \ldots, e_m) \in \mathbb{R}^{2m}$ and the incidence matrix H of the graph $\mathcal{G}$, then the links are obtained by $e = \hat{H}z$. Thus the sensor graph $\mathcal{G}$ and the positions z define the framework $(\mathcal{G}, z)$ with the links e in $\mathrm{Im}\hat{H}$. Suppose the robots should now perform certain tasks in a distributed way. In such a distributed control approach the control input u_i of robot i depends exclusively on local sensory information, which are the links $e_k = z_j - z_i$ with $j \in \mathcal{N}_i$. That is to say, $u_i = u_i(e_k)$, where k is an outgoing edge of node i.

D. Problem Statement

Given the sensor graph $\mathcal{G}$ and a set of distance constraints $d_k > 0$, where $k \in \{1, \ldots, m\}$, the desired formation is to have $\|e_k\| = d_k$ for all k. Ideally the robots should converge to this formation from any starting point. If we define $\mathbf{d} \in \mathbb{R}^m$ to be the vector with components $d_k^2/2$, then the goal is $r_\mathcal{G}(z) = \mathbf{d}$ and we refer to the set of frameworks $\left(\mathcal{G}, r_\mathcal{G}^{-1}(\mathbf{d})\right)$ as the *target formation*. References [1], [2] show that, in order to guarantee cohesion of the target formation, it has to be specified as an infinitesimally rigid formation, and the recent reference [4] shows that infinitesimal rigidity of the target formation is not only a necessary but in certain cases also a sufficient condition to stabilize the robots to the target formation. Because of this and in order to minimize sensor costs, we *assume* that the target formation is minimally rigid for every $z \in r_{\mathcal{G}}^{-1}(\mathbf{d})$. The formation control problem is then to find a distributed control law u such that the robots converge to a stationary formation where $r_{\mathcal{G}}(z) = \mathbf{d}$. This can also be formulated as a set stabilization problem:

> Find a control law u such that each $u_i = u_i(z_j - z_i)$, $j \in \mathcal{N}_i$, i.e., each control law can be implemented by onboard sensing, $z(t)$ converges as $t \to \infty$, and $r_{\mathcal{G}}(\lim_t z(t)) = \mathbf{d}$.

It is known that the goal $r_{\mathcal{G}}(\lim_t z(t)) = \mathbf{d}$ cannot be achieved for every initial position $z(0)$, for example, the references [8], [9], [10], [11], [13] show that three robots obeying potential function based control laws cannot form a triangle from an initially collinear position.

Remark 2.1: Reference [3] shows that in the directed graph case a property called constraint consistence is also a necessary graphical condition for the robots to attain a formation. That's why strictly speaking $\left(\mathcal{G}, r_{\mathcal{G}}^{-1}(\mathbf{d})\right)$ should be specified as a minimally persistent (i.e., minimally rigid and constraint consistent) formation. All graphs treated in this paper are constraint consistent and that's why we omit this characterization.

III. A Potential Function Based Control Law

We introduce the simple quadratic function $T(\omega) = \omega^2/8$. So $T(\omega)$ and its derivative are both 0 iff $\omega = 0$. All the following results in this paper can also be proved for more general functions as proposed in [9], but we omit doing this for the sake of simplicity; the reader is referred to [14]. For each robot a potential function is constructed that is zero whenever the robot has the desired distance from its neighbour and is positive otherwise. A potential function can then be interpreted as a cost that each robot has to pay for violating its distance constraints. For robot i we take $W_i : \mathbb{R}^{2n} \to \mathbb{R}$ defined as

$$W_i(z) = \sum_{k, o_{ik} \neq 0} T\left(\|e_k\|^2 - d_k^2\right), \tag{3}$$

where the summation is taken over all outgoing edges k of node i. In order to minimize its cost, robot i should move in the direction of the steepest descent of its potential function:

$$u_i = -\left[\frac{\partial}{\partial z_i} W_i(z)\right]^T = \frac{1}{2} \sum_{k, o_{ik} \neq 0} e_k \left(\|e_k\|^2 - d_k^2\right). \tag{4}$$

In terms of the outgoing edge matrix O, the overall closed-loop *z-dynamics* are then obtained as

$$\dot{z} = u = -\hat{O} D(e)[v(e) - \mathbf{d}]. \tag{5}$$

The initial condition of the z-dynamics is the initial location $z(0) = z_0 \in \mathbb{R}^{2n}$ of the robots. The control law is illustrated in Figure 1, where $c : \operatorname{Im}\hat{H} \to \mathbb{R}^{2m}$ is defined as

$$c(e) = D(e)[v(e) - \mathbf{d}]. \tag{6}$$

Different approaches to the formation control problem analyzing the z-dynamics in the state space $\mathbb{R}^{2n}$ have been proposed [2], [4]. One obstacle is that the target formation parametrized in the state space as $\mathcal{E}_z = \{z : r_{\mathcal{G}}(z) = \mathbf{d}\}$ is invariant under translations and rotations in the plane. Such a set is non-compact, which complicates an analysis based on differential geometry, set stability or invariance concepts. In addition, the formation specification is in the link space. Fortunately, the target formation parametrized in the link space, $\mathcal{E}_e = \{e : v(e) = \mathbf{d}\}$, *is* compact.

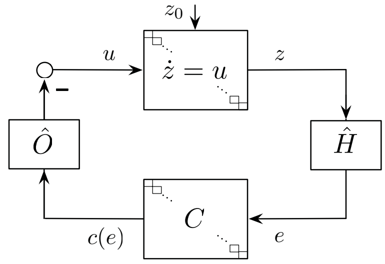

Fig. 1. Overall system's closed-loop dynamics for a directed sensor graph

For these obvious reasons we approach the formation control problem as a set stability problem in the link space. The closed-loop *e-dynamics* resulting from (5) are

$$\dot{e} = \hat{H} u = -\hat{H}\hat{O} D(e)[v(e) - \mathbf{d}]. \tag{7}$$

Equation (7), together with the initial condition $e(0) = e_0 = \hat{H} z_0$, defines a dynamical system evolving on the link space $\operatorname{Im}\hat{H}$ and will be simply called the *link dynamics* with solution $\phi(t, e_0)$. The target formation $\mathcal{E}_e$ is not the only equilibrium set of the link dynamics, since the matrices $\hat{O}$ and $\hat{H}$ have non-trivial kernels and we have the additional term $D(e)$. An intriguing approach to prove stability of $\mathcal{E}_e$, using the somewhat natural set-Lyapunov function candidate

$$V(e) = \sum_{k=1}^{m} T\left(\|e_k\|^2 - d_k^2\right), \tag{8}$$

does not provide the desired result—quite the contrary, $V(e)$ in some cases even increases along trajectories of the link dynamics. Thus stability of the target formation with respect to the link dynamics is far from obvious.

Classically, the function V was used together with the invariance principle in the undirected graph case [2], [6], [7]. For undirected graphs we can express the z-dynamics in terms of the links $e = \hat{H}_u z$, that is, the links corresponding to edges in E_+, and the z-dynamics (5) simplify to

$$\dot{z} = u = -\hat{H}_u^T D(e)\,[v(e) - \mathbf{d}] \tag{9}$$

(see [14]). Equation (9) corresponds in Figure 1 simply to an exchange of $\hat{O}$ with $\hat{H}_u^T$, similar to what is shown in [7]. Furthermore, note that the right-hand side of equation (9) can be derived as the gradient control

$$\dot{z} = u = -\left[\frac{\partial}{\partial z} V\left(\hat{H}_u z\right)\right]^T. \tag{10}$$

From (10) the link dynamics in the undirected graph case can then also be obtained from a gradient control law as

$$\dot{e} = \hat{H}_u\, u = -\hat{H}_u\, \hat{H}_u^T\, D(e)\, [v(e) - \mathbf{d}] \quad (11)$$

$$= -\hat{H}_u R_{\mathcal{G}}(e)^T [v(e) - \mathbf{d}] = -\hat{H}_u \hat{H}_u^T \left[\frac{\partial}{\partial e} V(e)\right]^T . \quad (12)$$

Note that the link dynamics (12) correspond to edges with orientation defined in E_+, and the links with the inverse orientation defined in E_- follow the same dynamics. The gradient structure of the overall control law for undirected sensor graphs strongly resembles the nonlinear control design concept of inverse optimality [15] and can indeed be related to an optimal control problem, as shown in the next section.

IV. Inverse Optimality and Cooperative Graphs

Let us illuminate the closed-loop dynamics from a game theoretic viewpoint. Each robot has its own individual strategy that consists of minimizing its potential function $W_i(z)$, that is, optimizing all its outgoing links to the desired length. On the other hand the robots do not follow a common protocol. This leads to the fact that each robot follows its own strategy regardless of what other non-neighbouring robots are doing. A valid question to ask is whether or not the robots indeed act cooperatively and achieve a common goal. If we look at the example in Figure 2 it does not necessarily seem so. Figure 2 shows two distinct robots that are embedded in a larger network and are interconnected by a directed link e_1 from robot 1 to robot 2. While the strategy of robot 1 is to meet its distance constraint on the link e_1, robot 2 follows its own strategy, which does not involve link e_1. In fact, the overall graph could be set up such that robot 2 is moving away from robot 1, as illustrated by the dashed line. Thus the robots do not act in a way that we would call cooperative.

A. Inverse Optimality of an Undirected Setup

Intuitively such a scenario cannot happen in an undirected graph, where for each link is also a reverse one. Loosely speaking, this implies that the strategy of both robots involves optimizing their distance to a desired length. Therefore, robots interconnected in an undirected graph should altogether pursue a common goal, that is, the overall system should be inverse optimal with respect to a meaningful cost functional. This is established by the following theorem.

Theorem 4.1: For an undirected graph the gradient control law $u = -\hat{H}_u^T [\partial V(e)/\partial e]^T$ is an inverse optimal control law that optimizes the cost functional

$$J(e_0, u) = \frac{1}{2}\int_0^\infty \left\|R_{\mathcal{G}}(e)^T[v(e) - \mathbf{d}]\right\|^2 + \|u\|^2 \, d\tau \quad (13)$$

$$\text{s.t.}\quad \dot{e} = \hat{H}_u\, u \ ,\ e(0) = e_0 \in \operatorname{Im} \hat{H}_u \quad (14)$$

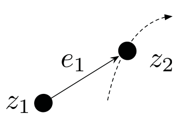

Fig. 2. Robots do not necessarily act cooperatively

with the terminal set $\mathcal{E}_e$. The value function corresponding to the optimal control problem (the minimum cost-to-go starting from $e(t)$) is, with $V(e(t))$ as defined in (8), given by

$$\int_t^\infty \left\|R_{\mathcal{G}}(e)^T [v(e) - \mathbf{d}]\right\|^2 d\tau \equiv V(e(t)) \quad (15)$$

Before we move on to the proof of Theorem 4.1, let us discuss the implications of the optimality of the overall strategy u. These become clear from the value function (15), that describes the optimal cost-to-go from any $e \in \operatorname{Im}\hat{H}_u$ at time t and, interestingly enough, corresponds to $V(e)$. Since $d/dtV(e)$ is given by the negative integrand of the functional on the right-hand side of (15) and is clearly negative, $V(e)$ is bounded along trajectories of the link dynamics. But then also the functional on the right-hand side of (15) is finite and we conclude that the integrand itself converges to zero. This again implies that the solution $\phi(t, e_0)$ either converges to a set where the rigidity matrix $R_{\mathcal{G}}(e)$ has a permanent rank loss or that $\phi(t, e_0) \to \mathcal{E}_e$. In the second case the terminal condition is also satisfied, while it is not in the first one. Hence, the robots either converge to the target formation, which is the global minimum of $V(e)$, or get stuck in a non-rigid formation corresponding to a local minimum of $V(e)$. From the viewpoint of the infinite horizon optimal control problem with terminal set $\mathcal{E}_e$, these two cases correspond to initial conditions from where the optimal control problem is either feasible or not.

Proof of Theorem 4.1: We derive a solution to the optimal control problem with cost functional (13) via dynamic programming. The Hamilton-Jacobi-Bellman equation is

$$0 = \min_{u\in\mathbb{R}^{2m}} \left\{ \frac{1}{2}\, [v(e) - \mathbf{d}]^T R_{\mathcal{G}}(e)\, R_{\mathcal{G}}(e)^T\, [v(e) - \mathbf{d}] + \frac{1}{2}\, u^T u + \left[\frac{\partial}{\partial e}\tilde{V}(e)\right] \hat{H}_u\, u \right\} , \quad (16)$$

where $\tilde{V} : \operatorname{Im}\hat{H}_u \to \mathbb{R}$ is the value function that fulfills the boundary condition ([16], Section 10.23)

$$\tilde{V}(e) = 0 \quad \Leftrightarrow \quad e \in \mathcal{E}_e\, . \quad (17)$$

The control input u^* minimizing the right-hand side of (16) is then given by $u^* = -\hat{H}_u^T [\partial\tilde{V}(e)/\partial e]^T$. Plugging u^* back into (16) gives the PDE

$$0 = [v(e) - \mathbf{d}]^T R_{\mathcal{G}}(e)\, R_{\mathcal{G}}(e)^T [v(e) - \mathbf{d}] - \left[\frac{\partial}{\partial e}\tilde{V}(e)\right] \hat{H}_u\, \hat{H}_u^T \left[\frac{\partial}{\partial e}\tilde{V}(e)\right]^T . \quad (18)$$

An intriguing solution for $\tilde{V}(e)$ that fulfills the PDE (18) and the boundary condition (17) is the sum of the potential functions $V(e)$. Hence, the optimal control law is given by $u = -\hat{H}_u^T [\partial V(e)/\partial e]^T$ and $V(e)$ is the value function. ■

B. Cooperative Graphs

In the general directed graph case each robot has its own potential function that it is trying to optimize and we are not able to derive the overall control law from the function $V(e)$. In this case we don't know if the overall control can be related to a meaningful optimal control problem. However, an idea that might generalize is that the robots act cooperatively in the sense that they behave as if the sensor graph were undirected. We refer to such graphs as cooperative graphs and define them as follows.

Definition 4.1: Consider a directed graph $\mathcal{G}$ with matrices H and O and the matrix H_u of the corresponding undirected graph. The graph $\mathcal{G}$ is said to be *cooperative* if

$$HO + O^T H^T = c\, H_u H_u^T\,, \tag{19}$$

where $c = 2$ if $\mathcal{G}$ is itself undirected, and $c = 1$ else.

Remark 4.1: Note that cooperative graphs strongly resemble the class of *balanced graphs* [17] for which holds (in our notation) $OH + H^T O^T = 2\, H_u^T H_u$ ([17], Theorem 7).

Unlike the inverse optimality result in Theorem 4.1, Definition 4.1 does not, at first glance, seem to relate the sensor graph to a cooperative behavior among the robots. This becomes obvious by an equivalent description of a cooperative graph in terms of a differential dissipation inequality.

Lemma 4.1: Under the assumptions of Definition 4.1 the graph $\mathcal{G}$ is cooperative if and only if for every $e \in \operatorname{Im} \hat{H}$

$$\frac{\partial V(e)}{\partial e}\, \dot{e} \equiv -\frac{c}{2}\, [v(e) - \mathbf{d}]^T\, R_{\mathcal{G}}(e)\, R_{\mathcal{G}}(e)^T\, [v(e) - \mathbf{d}]\,, \tag{20}$$

where $V(e)$ is defined in (8) and c as in Definition 4.1.

Proof: For a directed graph $\mathcal{G}$ the derivative of $V(e)$ along trajectories of the link dynamics (7) is given by

$$\begin{aligned}\frac{\partial V(e)}{\partial e}\, \dot{e} &= -[v(e) - \mathbf{d}]^T D(e)^T \hat{H}\hat{O} D(e)[v(e) - \mathbf{d}] \\ &= -\frac{1}{2}[v(e) - \mathbf{d}]^T D(e)^T \left(\hat{H}\hat{O} + \hat{O}^T\hat{H}^T\right) D(e)[v(e) - \mathbf{d}].\end{aligned}$$

With $R_{\mathcal{G}}(e)$ written out the right-hand side of (20) is

$$(c/2)\, [v(e) - \mathbf{d}]^T D(e)^T \hat{H}_u \hat{H}_u^T D(e)[v(e) - \mathbf{d}] \tag{21}$$

and clearly equals $(\partial V(e)/\partial e)\dot{e}$ iff (19) holds with $c = 1$.

In the case of an undirected graph we look at the dynamics (12), that is, the link dynamics corresponding to edges with orientation given in E_+. By straightforward calculation it can be verified that both (19) and (20) always hold for $c = 2$. ■

Remark 4.2: The apparently artificial distinction of undirected graphs from directed ones via the factor c arises because edges are not counted twice for undirected graphs and it is sufficient to look the dynamics of one set of edges.

Suppose solutions of the link dynamics exist on an unbounded time interval (shown in the next section), then (20) may be integrated over $[t, \infty)$ to obtain exactly the closed-loop functional (15) with a possibly non-vanishing but finite terminal cost $V(e(\infty))$. That is to say, for a cooperative graph the closed-loop link dynamics optimize the same common cost functional as in the undirected graph case. The following theorem identifies besides undirected graphs also directed cycles and open chains as cooperative graphs.

Theorem 4.2: Every undirected graph, every directed cycle, and every directed open chain is a cooperative graph.

Proof: The proof in the case of undirected graphs is trivial. We will prove it in the case of directed cycles. A directed cyclic graph has the same number of nodes as links $(n = m)$ and the graph matrices $H = P_m - I_m \in \mathbb{R}^{m\times m}$ and $O = -I_m \in \mathbb{R}^{m\times m}$, where I_m is the m-dimensional identity matrix, and P_m is the orthogonal and circulant matrix $P_m = \operatorname{circ}\begin{bmatrix}0 & 1 & 0 & \dots & 0\end{bmatrix}$. For the corresponding undirected cyclic graph the edges can be labeled, such that $H_u = H$. We then have for (19) with $c = 1$

$$\begin{aligned}HO + O^T H^T &= -(P_m - I_m) - (P_m - I_m)^T \\ &= (P_m - I_m)(P_m - I_m)^T = HH^T = H_u H_u^T\,.\end{aligned}$$

Thus, by Definition 4.1, a directed cycle is a cooperative graph. The proof for directed open chains is similar and based on the fact that such graphs are almost directed cycles with one link missing (see [14]). ■

In light of Theorem 4.2 both a directed and an undirected cyclic graph are cooperative. Due to inverse optimality this is not surprising for the undirected graph, but in the case of a directed cycle, each robot focuses on its leader robot only and does not care about the interconnecting link to its follower. Nevertheless, the robots behave as if there would be no leader-follower structure but rather an information flow in both directions. So from the viewpoint of optimality, both setups optimize the very same cost functional, although the robots in the directed graph case are not meant to do so. This quite astonishing fact also can be found for undirected and directed open chains.

V. Cooperative Graphs and Stability of the Formation

The notion of a cooperative graph reflects not only the cooperative behavior among the robots but, together with the infinitesimal rigidity of the target formation, it also admits a direct stability result. Let us introduce the notation

$$\Omega(c) := \left\{ e \in \operatorname{Im}\hat{H} : V(e) \le c \right\}, \tag{22}$$

and establish a theorem on the limit sets of the link dynamics.

Theorem 5.1: Consider a cooperative graph $\mathcal{G}$. For every initial condition $e_0 \in \operatorname{Im}\hat{H}$ the link dynamics (7) are forward complete and bounded in the sublevel set $\Omega(V(e_0))$, and their solution $\phi(t, e_0)$ converges to the largest invariant set contained in

$$\mathcal{W}_e = \left\{ e \in \Omega(V(e_0)) : \left\| R_{\mathcal{G}}(e)^T\, [v(e) - \mathbf{d}] \right\| = 0 \right\}. \tag{23}$$

Moreover, consider the sublevel set $\Omega(\rho)$ where ρ is sufficiently small, such that for every $e \in \Omega(\rho)$, $(\mathcal{G}, e)$ is minimally rigid. Then for every $e_0 \in \Omega(\rho)$ the set $\mathcal{E}_e$ is exponentially stable with respect to the link dynamics.

Proof: Note that the right-hand side of the link dynamics (7) is locally Lipschitz. Because the function $V(e)$ is positive definite with respect to the compact set $\mathcal{E}_e$ and its derivative along trajectories $(\partial V(e)/\partial e)\dot{e}$ given by (20) is negative semidefinite, $V(e)$ is a suitable set Lyapunov function candidate for the target formation $\mathcal{E}_e$. For every initial condition $e_0 \in \hat{H}\left(\mathbb{R}^{2n}\right)$ the sublevel set $\Omega(V(e_0))$ is a compact and invariant set and thus the link dynamics are forward complete ([18], Theorem 3.3). Note that the assumptions for the invariance principle ([18], Theorem

4.4) are satisfied and thus $\phi(t, e_0)$ converges to the largest invariant set in $\Omega(V(e_0))$ where $(\partial V(e)/\partial e)\dot{e} = 0$, i.e., $\mathcal{W}_e$.

Due to minimal rigidity of the target formation the matrix $R_{\mathcal{G}}(e)^T \in \mathbb{R}^{2n \times m}$ has full rank $m \quad \forall\, e \in \mathcal{E}_e$, or said differently $R_{\mathcal{G}}(e)R_{\mathcal{G}}(e)^T$ has no zero eigenvalues $\forall\, e \in \mathcal{E}_e$. The eigenvalues of $R_{\mathcal{G}}(e)R_{\mathcal{G}}(e)^T$ are continuous functions of the matrix elements and thus of $e \in \hat{H}\left(\mathbb{R}^{2n}\right)$. The minimal eigenvalue of $R_{\mathcal{G}}(e)R_{\mathcal{G}}(e)^T$ is positive $\forall\, e \in \mathcal{E}_e$ and, due to continuity, also in an open neighbourhood of $\mathcal{E}_e$. Let $\mathcal{Q}$ be the set where the matrix $R_{\mathcal{G}}(e)R_{\mathcal{G}}(e)^T$ has a zero eigenvalue. In order to continue, consider the level set $\Omega(\rho)$, where ρ is small enough that $\Omega(\rho)$ does not intersect the set $\mathcal{Q}$. Since $\Omega(\rho)$ is compact, we define λ as

$$\lambda := \min_{e \in \Omega(\rho)} \text{eig}\left(R_{\mathcal{G}}(e)R_{\mathcal{G}}(e)^T\right) > 0. \tag{24}$$

We then have for the derivative of $V(e)$ along trajectories

$$(\forall\, e \in \Omega(\rho))\, \dot{V}(e) \leq -\frac{\lambda c}{2}\, \|[v(e) - \mathbf{d}]\|^2 = -\lambda c\, V(e), \tag{25}$$

and thus $\forall e \in \Omega(\rho)$, $\dot{V}(e)$ is negative definite with respect to $\mathcal{E}_e$. Therefore, by standard arguments of set stability theory [19], the set $\mathcal{E}_e$ is asymptotically stable with $\Omega(\rho)$ as guaranteed region of attraction. Moreover, by the Bellman-Gronwall Lemma ([18], Lemma A.1), we have that $\forall\, e_0 \in \Omega(\rho)$, $V(e(t)) \leq V(e_0)e^{-\lambda c t}$ and, after applying some inequalities, we can indeed show that the point-to-set distance to the set $\mathcal{E}_e$ is exponentially decreasing. ∎

Although the link dynamics (7) and the z-dynamics (5) both have the target formation as an equilibrium set, the convergence $\phi(t, e_0) \to \mathcal{E}_e$ in the link space does not imply convergence to a finite point in the state space but only a convergence of the point-to-set distance to the set $\mathcal{E}_z$.

The exponential convergence rate of the link dynamics allows us to remove this obstacle. For every initial condition $z_0 \in \hat{H}^{-1}\left(\Omega(\rho)\right)$ the link dynamics with initial condition $e_0 = \hat{H} z_0$ converge exponentially to $\mathcal{E}_e$. Hence, the function

$$f(t) := -\hat{O}D(\phi(t, e_0))[v(\phi(t, e_0)) - \mathbf{d}] \tag{26}$$

is exponentially decreasing in time and thus a $\mathcal{L}_1$ function. The z-dynamics can then be written in integrated form as

$$z(t) = z_0 + \int_0^t f(\tau)\, d\tau. \tag{27}$$

Since $f \in \mathcal{L}_1$, the integral on the right-hand side of (27) exists even in the limit as $t \to \infty$ and thus a solution of the z-dynamics converges to a finite point in $\mathcal{E}_z$.

We are now ready to state our final theorem.

Theorem 5.2: Consider a cooperative sensor graph $\mathcal{G}$. For every initial condition $z_0 \in \hat{H}^{-1}\left(\Omega(\rho)\right)$ the control law (4) solves the formation control problem.

VI. Conclusions and Outlook

This paper related the formation control problem of autonomous robots in the undirected graph case to an optimal control problem of the overall system, which is related to rigidity and to the target formation. The definition of a cooperative graph allowed us to extend the implications of this result to directed graphs; directed cycles and directed open chains are identified as such graphs. The two ingredients, a cooperative graph and minimal rigidity of the target formation, led directly to a stability result with a guaranteed region of attraction.

The definition of a cooperative graph imitates the behavior of the robots as if they were interconnected in an undirected graph. This quite restrictive definition can be modified to capture a wider variety of graphs. Cooperative graphs also do not extend to acyclic graphs whose analysis is mainly based on the cascade structure of the overall system [4], [10]. A different way to capture such graphs or a situation, such as in Figure 2, could be a non-cooperative game approach. Finally, for a global stability analysis of the link dynamics, it has to be shown that $\mathcal{E}_e$ is the only positively invariant set in $\mathcal{W}_e$. This can be done by methods of differential geometry, which will be shown in the journal version of this paper.

References

[1] T. Eren, P. N. Belhumeur, B. D. O. Anderson, and A. S. Morse, "A framework for maintaining formations based on rigidity," *Proceedings of the 15th IFAC World Congress , Barcelona*, 2002.

[2] R. Olfati-Saber and R. M. Murray, "Distributed cooperative control of multiple vehicle formations using structural potential functions," *Proceedings of the 15th IFAC World Congress , Barcelona*, 2002.

[3] J. M. Hendrickx, B. D. O. Anderson, J.-C. Delvenne, and V. D. Blondel, "Directed graphs for the analysis of rigidity and persistence in autonomous agent systems," *International Journal of Robust and Nonlinear Control*, no. 17, pp. 960–981, 2007.

[4] L. Krick and M. Broucke and B. Francis, "Stabilization of infinitesimally rigid formations of multi-robot networks," *International Journal of Control*, Submitted January 2008.

[5] B. D. O. Anderson, C. B. Yu, and F. Baris, "Information architecture and control design for rigid formations," *Proceedings of the 26th Chinese Control Conference*, pp. 2–10, 2007.

[6] H. G. Tanner, A. Jadbabaie, and G. Pappas, "Stable Flocking of Mobile Agents, Part I: Fixed Topology," *42nd IEEE Conference on Decision and Control, Maui Hawaii*, pp. 2010–2015, December 2003.

[7] M. Arcak, "Passivity as a Design Tool for Group Coordination," *IEEE Transactions on Automatic Control*, vol. 52, no. 8, pp. 1380–1390, 2007.

[8] M. Cao, C. B. Yu, A. S. Morse, B. D. O. Anderson, and S. Dasgupta, "Controlling a Triangular Formation of Mobile Autonomous Agents," in *46th IEEE Conference on Decision and Control (CDC), New Orleans, LA, USA*, December 2007, pp. 3603–3608.

[9] ——, "Generalized Controller for Directed Triangle Formations," in *17th International Federation of Automatic Control World Congress (IFAC), Seoul, Korea*, July 2008.

[10] M. Cao, B. D. O. Anderson, A. S. Morse, and C. B. Yu, "Control of acyclic formations of mobile autonomous agents," *preprint*, 2008.

[11] S. Smith, M. Broucke, and B. Francis, "Stabilizing a multi-agent system to an equilateral polygon formation," *MTNS*, 2006.

[12] N. Biggs, *Algebraic Graph Theory*, 2nd ed. Cambridge University Press, 1993.

[13] B. D. O. Anderson, C. B. Yu, S. Dasgupta, and A. S. Morse, "Control of three co-leader formation in the plane," *System and Control Letters*, pp. 573–578, 2007.

[14] F. Dörfler, "Geometric Analysis of the Formation Control Problem for Autonomous Robots," Diploma Thesis, University of Toronto, 2008.

[15] R. Sepulchre, M. Janković, and P. Kokotović, *Constructive Nonlinear Control*. Springer, 1997.

[16] R. Boudarel, J. Delmas, and P. Guichet, *Dynamic Programming and its Application to Optimal Control*. Acadmic Press, 1971.

[17] R. Olfati-Saber and R. M. Murray, "Concensus Problems in Networks of Agents with Switching Topology and Time-Delays," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1520–1533, 2004.

[18] H. Khalil, *Nonlinear Systems*, 3rd int. ed. Prentice Hall, 2000.

[19] N. Bathia and G. Szegö, *Stability Theory of Dynamical Systems*. Springer, 1970.