Università degli Studi di Padova

Facoltà di Ingegneria

Corso di Laurea in Ingegneria dell'Informazione

Consensus Problem in Nonlinear Spaces

Laureando

Martina Zambelli

Relatore

Prof. Michele Pavon

Anno Accademico 2010/2011

To my parents

Contents

Chapter 1

Introduction

Over the past decade, particular attention has been devoted to the study of collective problems where interacting agents must reach a common objective under information and communication constraints. These problems arise in a variety of disciplines including physics, biology, computer science and systems and control theory. Analysis and design efforts have been devoted to understand how a group of moving agents (e.g. flocks of birds, schools of fish or autonomous robots) can reach a consensus without an external reference and in a decentralized way.

Recent results have contributed to a good understanding of synchronization of interacting agents in Euclidean space, based on the linear consensus algorithm

$$\frac{d}{dt}x_k(t) = \sum_{j=1}^{N} a_{jk}(t)(x_j(t) - x_k(t)), \quad k = 1, 2, ..., N \tag{1.1}$$

where a_{jk} is the weight of link $j \rightsquigarrow k$ in the graph representing the communication links between the N agents, and $x_k \in \mathbb{R}^n,\ k = 1, 2, ..., N$ the state of agent k. Global exponential synchronization is ensured even with varying a_{jk}, as long as the agents are uniformly connected.
However, many interesting applications involve manifolds that are not homeomorphic to an Euclidean space, like the circle S^1 for (e.g. oscillator) phase variables or the group of rotations $SO(n)$ for rigid body orientations.
The celebrated Kuramoto model, which deals with the study of synchronization phenomena in populations of coupled oscillators, has been adopted to investigate consensus problem on the circle.

An essential difference between linear consensus algorithms and their nonlinear extensions is the non-convex nature of symmetric spaces like the circle. This property is what makes the convergence analysis graph dependent when the state space is nonlinear.

Stability analysis of these algorithms deals with the Lyapunov theory: stability (also called stability in the sense of Lyapunov or Lyapunov stability) means that the system trajectory can be kept arbitrarily close to the origin (assuming the origin be an equilibrium point) by starting sufficiently close to it. It is therefore a form of continuity of solutions with respect to their initial conditions.
In many engineering applications, Lyapunov stability is not enough. Some types of engineering requirement is captured by the concept of asymptotic stability. Frequently, it is still not sufficient to know that a system will converge to the equilibrium point after infinite time. There is a need to estimate how fast the system trajectory approaches the origin. The concept of exponential stability can be used for this purpose.

Practical applications include autonomous swarm/formation operation, distributed decision making, neural and communication networks, clustering and other reduction methods, optimal covering or coding, and other fields where averaging/synchronizing or distributing a set of points appears as a subproblem. For instance, formation control of autonomous vehicles (e.g. flocking or rendezvous problems), or even sensor networks; consider for examples a room where there are several temperature sensors: a possible solution to ensure that the sensors reach a common goal without adopting a central unit, is to use distributed control and consensus algorithms.

The thesis is outlined as follows. *Chapter 2* is devoted to recalling some mathematical preliminaries, such as concepts of stability, Lyapunov theory, graph theory; in this chapter, we also give some (intuitive) notions of critical concepts concerning nonlinear spaces (e.g. the concept of manifold, geodesic and geodesic distance, Lie group). *Chapter 3* deals with the goal of the thesis: the consensus problem; we firstly introduce the problem in linear space and then we present consensus in nonlinear spaces, focusing our attention on the circle. In *Chapter 4* we give some critical example of application of consensus in engineering problems (e.g. autonomous ocean sampling network, phase synchronization of oscillator networks, vehicle formations).

Chapter 2

Mathematical Preliminaries

2.1 Stability and Lyapunov Theory

Given a control system, the first and most important question about its various properties is whether it is stable. Qualitatively, a system is described as stable if starting the system somewhere near its desired operating point implies that it will stay around the point ever after.
Basic concepts and results of this section may be found in references [1, 2, 3, 4, 11].

2.1.1 Stability of Linear Systems

We consider a system described by a linear homogeneous time-invariant differential equations

$$\dot{x}(t) = Fx(t) \tag{2.1}$$

where $F : \mathcal{X} \to \mathcal{X}$ is a linear map and $\mathcal{X}$ an n-dimensional space. The equation is called *linear* because the right-hand side is a linear function of $x(t)$. It is called *homogeneous* because the right-hand side is zero for $x(t) = 0$ (an equation of the form $\dot{x} = Fx + G$ with $G \neq 0$ is called *inhomogeneous*). The equation is called *time invariant* because F is independent of t.

In order to specify a solution of (2.1), one has to provide an initial value. The solution of the *initial value problem*: find a function x satisfying

$$\dot{x} = Fx, \quad x(0) = x_0, \tag{2.2}$$

is denoted by $x(t, x_0)$. In the scalar case ($n = 1$, $F = f$), it is well known that $x(t, x_0) = e^{ft}x_0$. In order to have a similar result for the multivariable case, we introduce the matrix exponential function

$$e^{Ft} := \sum_{k=0}^{\infty} \frac{t^k F^k}{k!}. \tag{2.3}$$

It follows that $x(t, x_0) = e^{Ft}x_0$.

In mathematics, stability theory addresses the stability of solutions of differential equations and of trajectories of dynamical systems under small perturbations of initial conditions. The simplest kind of an orbit is a fixed point, or an equilibrium. The stability of fixed points of a system of constant coefficient linear differential equations of first order can be analyzed using the eigenvalues of the corresponding matrix so that the concept of stability is strictly related to the asymptotic behavior of each elementary mode. The general expression of each mode is

$$t^h e^{\lambda t} = t^h e^{(\sigma+i\omega)t} = t^h e^{\sigma t}[\cos(\omega t) + i \sin(\omega t)]. \tag{2.4}$$

where $\lambda = \sigma + i\omega$ is a complex eigenvalue of multiplicity h.

We can summarize all cases as follows

- convergence (to zero) of a mode if and only if $\sigma < 0$, i.e. the corresponding eigenvalue λ satisfies $Re[\lambda] < 0$;
- boundness if and only if a) we are in the previous case (convergence to zero) or b) $\sigma = h = 0$, i.e. the corresponding eigenvalue λ satisfies $Re[\lambda] = 0$ and moreover we are considering only the first of the possibly multiple modes associated to the same eigenvalue;
- divergence in all the other cases.

Thus the following property holds:

Property 2.1 *e^{Ft} converges to zero if and only if all the eigenvalues λ_i of F have $Re[\lambda_i] < 0$, while it is limited if and only if all the eigenvalues λ_i of F have $Re[\lambda_i] \leq 0$, and eigenvalues with $Re[\lambda_i] = 0$ have equal geometric and algebraic multiplicity.*

What is important in control theory is the concept of *asymptotic stability* from the theory of ordinary differential equations. For linear time-invariant systems, this concept can be defined as follows

Definition 2.1 *The system (2.1) is called* stable *if every solution tends to zero for $t \to \infty$.*

Equilibrium point and Stability

We define *equilibrium point* of a dynamical system a point in which the state remains "anchored", in the sense that the corresponding trajectory is a point. In other words, it must be verified that

$$x(t) = x(0), \quad \forall t \geq 0 \tag{2.5}$$

which is equivalent to the fact that $x(t)$ is constant and therefore its derivative is zero.
The equilibrium points of a system are obtained by solving the equation $Fx = 0$. We note that this homogeneous system always admits the solution $x = 0$ and that this is the only solution to the system only if $\det F \neq 0$, whereas if F is not invertible, infinite solutions appear.

The stability around the equilibrium points is defined by the following

Definition 2.2 *An equilibrium points x_{eq} is called* **simply stable** *if* $\forall \epsilon > 0\ \exists \delta > 0$ *such that if* $||x(0) - x_{eq}|| \leq \delta$ *then we have* $||x(t) - x_{eq}|| \leq \epsilon$, $\forall t \geq 0$*, while it is called* **asymptotically stable** *if:*
1. it is simply stable;
2. $\lim_{t\to+\infty} x(t) = x_{eq}$ *as long as* $x(0)$ *is chosen sufficiently close to* x_{eq}.

Finally we can conclude with this result:

Property 2.2 *:*

- e^{Ft} *is limited* $\Leftrightarrow$ $x_{eq} = 0$ *is simply stable ;*
- $\lim\limits_{t\to+\infty} e^{Ft} = 0$ $\Leftrightarrow$ $x_{eq} = 0$ *is asymptotically stable.*

2.1.2 Stability of Nonlinear Systems

The most useful and general approach for studying the stability of nonlinear control systems is the theory introduced in the late 19$^{\text{th}}$ century by the Russian mathematician Alexandr Mikhailovich Lyapunov. Lyapunov's work includes two methods for stability analysis: the so-called linearization method and the direct method. The linearization method draws conclusions about a nonlinear system's local stability around an equilibrium point from the stability properties of its linear approximation. The direct method is not restricted to local motion, and determines the stability properties of a nonlinear system by constructing a scalar "energy-like" function for the system and examining the function's time variation.

A nonlinear dynamic system can usually be represented by a set of nonlinear differential equations in the form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t) \tag{2.6}$$

where $\mathbf{f}$ is a $n \times 1$ nonlinear vector function, and $\mathbf{x}$ is the $n \times 1$ state vector. A solution $\mathbf{x}(\mathrm{t})$ of the equation (2.6) usually corresponds to a curve in state space as t varies from zero to infinity. This curve is generally referred to as a *state trajectory* or a *system trajectory.* Linear systems are a special class of nonlinear systems. In the more general context of nonlinear systems, the adjectives "time-varying" or "time-invariant" referred to linear systems are traditionally replaced by "non-autonomous" and "autonomous".

Autonomous Systems

Definition 2.3 *The nonlinear system (2.6) is said to be* autonomous *if* $\mathbf{f}$ *does not depend explicitly on time, i.e. if the system's state equation can be written*

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) \tag{2.7}$$

Otherwise, the system is called non-autonomous.

The fundamental difference between autonomous and non-autonomous systems lies in the fact that the state trajectory of an autonomous system is independent of the initial time, while that of a non-autonomous system generally is not.

Equilibrium points

As for linear systems, it is possible for a system trajectory to correspond to only a single point. Such a point is called an equilibrium point.

Definition 2.4 *A state* $\mathbf{x}^*$ *is an* equilibrium state *(or* equilibrium point*) of the system if once* $\mathbf{x}(t)$ *is equal to* $\mathbf{x}^*$*, it remains equal to* $\mathbf{x}^*$ *for all future time. Mathematically, this means that the constant vector* $\mathbf{x}^*$ *satisfies*

$$\mathbf{0} = \mathbf{f}(\mathbf{x}^*) \tag{2.8}$$

A nonlinear system can have several (or infinitely many) isolated equilibrium point, e.g. the well-known Logistic (Riccati) Equation $\dot{x} = x(1-x)$.

Stability

We introduced the intuitive notion of stability as a kind of well-behavedness around a desired operating point. However, since nonlinear systems may have much more complex behavior than linear systems, the mere notion of stability is not enough to describe the essential features of their motion.

We now give some notations.
Let $\mathbf{B}_R$ denote the spherical region (or ball) defined by $||\mathbf{x}|| < R$ in state-space, and $\mathbf{S}_R$ the sphere itself, defined by $||\mathbf{x}|| = R$.

Definition 2.5 *Suppose state* $\mathbf{x} = \mathbf{0}$ *is an equilibrium state. It is said to be* **stable** *if, for any* $R > 0$*, there exists* $r > 0$*, such that if* $||\mathbf{x}(0)|| < r$*, then* $||\mathbf{x}(t)|| < R$ *for all* $t \geq 0$*. Otherwise, the equilibrium point is* unstable.

Essentially, stability (also called **stability in the sense of Lyapunov** or **Lyapunov stability**) means that the system trajectory con be kept arbitrarily close to the origin by starting sufficiently close to it. It is therefore a form of continuity of solutions with respect to their initial conditions.
More formally, the definition states that the origin is stable, if, given that

we do not want the state trajectory $\mathbf{x}(\mathrm{t})$ to get out of a ball of arbitrarily specified radius $\mathbf{B}_R$, a value $r(R)$ can be found such that starting the state from within the ball $\mathbf{B}_r$ at time 0 guarantees that the state will stay within the ball $\mathbf{B}_R$ thereafter.
The definition 2.5 can be written

$$\forall R > 0,\ \exists r > 0,\ ||\mathbf{x}(0)|| < r\ \Rightarrow\ \forall t \geq 0,\ ||\mathbf{x}(t)|| < R$$

or, equivalently

$$\forall R > 0,\ \exists r > 0,\ \mathbf{x}(0) \in \mathbf{B}_r\ \Rightarrow\ \forall t \geq 0,\ \mathbf{x}(t) \in \mathbf{B}_R$$

In many engineering applications, Lyapunov stability is not enough. Some types of engineering requirement is captured by the concept of *asymptotic stability.*

Definition 2.6 *An equilibrium point* **0** *is* **asymptotically stable** *if it is stable, and if in addition there exists some* $r > 0$ *such that* $||\mathbf{x}(0)|| < r$ *implies that* $\boldsymbol{x}(t) \to \mathbf{0}$ *as* $t \to \infty$*.*

Asymptotic stability means that the equilibrium is stable, and that in addition, states started close to **0** actually converge to **0** as time t goes to infinity. Thus, system trajectories starting from within the ball $\mathbf{B}_r$ converge to the origin. The ball $\mathbf{B}_r$ is called a *domain of attraction* of the equilibrium point (while *the* domain of attraction of the equilibrium point refers to the largest such region, i.e. to the set of all points such that trajectories initiated at these points eventually converge to the origin). An equilibrium point which is Lyapunov stable but not asymptotically stable is called *marginally stable.*

In many engineering applications, it is still not sufficient to know that a system will converge to the equilibrium point after infinite time. There is a need to estimate how fast the system trajectory approaches **0**. The concept of *exponential stability* can be used for this purpose.

Definition 2.7 *The equilibrium point* **0** *is* **exponentially stable** *if there exist two strictly positive numbers* α *and* λ *such that*

$$||\mathbf{x}(t)|| \leq \alpha ||\mathbf{x}(0)|| e^{-\lambda t} \quad \forall t > 0 \tag{2.9}$$

in some ball $\mathbf{B}_r$ *around the origin.*

In words, (2.9) means that the state vector of an exponentially stable system converges to the origin like a decreasing exponential function. The positive number λ is often called the *rate* of exponential convergence.

All the results examined above can be generalized for a non-specific equilibrium point x^*, simply translating from the origin to the equilibrium point x^*.

Moreover, the above definitions are formulated to characterize the *local* behavior of systems, i.e. how the state evolves after starting near the equilibrium point. Local properties tell little about how the system will behave when the initial state is at some distance away from the equilibrium. *Global* concepts are required for this purpose.

Definition 2.8 *If asymptotic (or exponential) stability holds for any initial states, the equilibrium point is said to be asymptotically (or exponentially) stable* in the large*. It is also called* globally *asymptotically (or exponentially) stable.*

Non-Autonomous Systems

For non-autonomous systems, of the form $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$, *equilibrium points* $\mathbf{x}^*$ are defined by

$$\mathbf{f}(\mathbf{x}^*, t) \equiv \mathbf{0} \quad \forall t \geq t_0 \tag{2.10}$$

Definition 2.9 *The equilibrium point* $\mathbf{0}$ *is* **stable** *at* t_0 *if for any* $R > 0$*, there exists a positive scalar* $r(R, t_0)$ *such that*

$$||\mathbf{x}(t_0)|| < r \quad \Rightarrow \quad ||\mathbf{x}(t)|| < R \quad \forall t \geq t_0 \tag{2.11}$$

Otherwise, the equilibrium point $\mathbf{0}$ *is* unstable*.*

Definition 2.10 *The equilibrium point* $\mathbf{0}$ *is* **asymptotically stable** *at time* t_0 *if*

- *it is stable*
- $\exists r(t_0) > 0$ *such that* $||\mathbf{x}(t_0)|| < r(t_0) \quad \Rightarrow \quad ||\mathbf{x}(t)|| \to \mathbf{0}$ *as* $t \to \infty$

Here, the asymptotic stability requires that there exists an attractive region for *every* initial time t_0. The size of the attractive region and the speed of trajectory convergence may depend on the initial time t_0.

Definition 2.11 *The equilibrium point* $\mathbf{0}$ *is* **exponentially stable** *at time* t_0 *if there exist two positive numbers* α *and* λ *such that for sufficiently small* $\mathbf{x}(t_0)$*,*

$$\forall t \geq t_0, \quad ||\mathrm{x}(t)|| \leq \alpha ||\mathrm{x}_0|| e^{-\lambda(t-t_0)} \tag{2.12}$$

Definition 2.12 *The equilibrium point* $\mathbf{0}$ *is* **globally asymptotically stable** *if* $\forall \mathbf{x}(t_0)$

$$\mathbf{x}(t) \to \mathbf{0} \quad \text{as } t \to \infty \tag{2.13}$$

As for autonomous systems, the results presented for non-autonomous systems can be generalized to a non-specific equilibrium point x^*.

2.1.3 Lyapunov Analysis

Linearization and Local Stability

Lyapunov's linearization method is concerned with the *local* stability of a nonlinear system. It is a formalization of the intuition that a nonlinear system should behave similarly to its linearized approximation for small range motions.

Consider the autonomous system in (2.7), and assume that $\mathbf{f(x)}$ is continuously differentiable. Then the system dynamics can be written as

$$\dot{\mathbf{x}} = \left(\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right)_{\mathbf{x=0}} \mathbf{x} + \mathbf{f}_{\mathbf{h.o.t.}}(\mathbf{x}) \tag{2.14}$$

where $f_{h.o.t.}(x)$ stands for higher-order terms in $\mathbf{x}$. Note that the above Taylor expansion starts directly with the first-order term, due to the fact that $\mathbf{f(0)=0}$, since $\mathbf{0}$ is an equilibrium point. Let us use the constant matrix $\mathbf{A}$ to denote the Jacobian matrix of $\mathbf{f}$ with respect to $\mathbf{x}$ at $\mathbf{x=0}$ (an $n \times n$ matrix of elements $\frac{\partial f_i}{\partial x_j}$)

$$\mathbf{A} = \left(\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right)_{\mathbf{x=0}}$$

Then, the system

$$\dot{\mathbf{x}} = \mathbf{A}\ \mathbf{x} \tag{2.15}$$

is called the *linearization* (or *linear approximation*) of the original nonlinear system at the equilibrium point $\mathbf{x=0}$.

Theorem 2.1 Lyapunov's linearization method

- *If the linearized system is strictly stable (i.e. if all eigenvalues of* $\boldsymbol{A}$ *are strictly in the left-half complex plane), then the equilibrium point is asymptotically stable (for the actual nonlinear system).*

- *If the linearized system is unstable (i.e. if at least one eigenvalue of* $\boldsymbol{A}$ *is strictly in the right-half complex plane), then the equilibrium point is unstable (for the nonlinear system).*

- *If the linearized system is marginally stable (i.e. all eigenvalues of* $\boldsymbol{A}$ *in the left-half complex plane, but at least one of them is on the* $\boldsymbol{j\omega}$ *axis), then one cannot conclude anything from the linear approximation (the equilibrium point may be stable, asymptotically stable, or unstable for the nonlinear system).*

Lyapunov's Direct Method

The basic philosophy of Lyapunov's direct method is the mathematical extension of a fundamental physical observation: if the total *energy* of a mechanical (or electrical) system is continuously dissipated, then the system, *whether linear or nonlinear*, must eventually settle down to an equilibrium point. Thus, we may conclude the stability of a system by examining the variation of a *scalar* function.

The basic procedure of Lyapunov's direct method is to generate a scalar "energy-like" function for the dynamic system, and examine the time variation of the scalar function. This function has two properties: the first one is that the function is strictly positive unless all the state variables (x, $\dot{x}$, $\ddot{x}$, ...) are zero; the second one is a property associated with the dynamics, i.e. the function is monotonically decreasing when the variables vary according to the dynamic equation.

In Lyapunov's direct method, the first property is formalized by the notion of *positive definite functions*, the second is formalized by the so-called *Lyapunov functions.*

Lyapunov Direct Method for Autonomous Systems

Definition 2.13 *A scalar continuous function $V(\mathbf{x})$ is said to be* locally positive definite *if $V(\mathbf{0}) = 0$ and, in a ball $\boldsymbol{B}_{R_0}$ $\quad \mathbf{x} \neq \mathbf{0} \;\Rightarrow\; V(\mathbf{x}) > 0$.*
If $V(\mathbf{0}) = 0$ and the above property holds over the whole state space, then $V(\mathbf{x})$ is said to be globally positive definite.

This definition implies that the function V has a unique minimum at the origin $\mathbf{0}$.

With $\mathbf{x}$ denoting the state of the system (2.7), a scalar function V($\mathbf{x}$) actually represent an implicit function of time t. Assuming that $V(\mathbf{x})$ is differentiable, its derivative with respect to time is

$$\dot{\boldsymbol{V}} = \frac{\boldsymbol{dV}(\mathbf{x})}{\boldsymbol{dt}} = \frac{\boldsymbol{\partial V}}{\boldsymbol{\partial}\mathbf{x}}\dot{\mathbf{x}} = \frac{\boldsymbol{\partial V}}{\boldsymbol{\partial}\mathbf{x}}\mathbf{f}(\mathbf{x})$$

We see that, since $\mathbf{x}$ is required to satisfy the autonomous state equation (2.7), $\dot{V}$ only depends on $\mathbf{x}$. It is often referred to as "the derivative of V along the system trajectory" - in particular, $\dot{V} = 0$ at an equilibrium point.

Definition 2.14 *If, in a ball $\boldsymbol{B}_{R_0}$, the function $V(\mathbf{x})$ is positive definite and has continuous partial derivatives, and if its time derivative along any state trajectory of system (2.7) is negative semi-definite, i.e. $\dot{V}(\mathbf{x}) \leq 0$ then $V(\mathbf{x})$ is said to be a* **Lyapunov function** *for the system (2.7).*

Theorem 2.2 ***Lyapunov Theorem for Local Stability***
If, in a ball $\boldsymbol{B}_{R_0}$, there exists a scalar function $V(\mathbf{x})$ with continuous first partial derivatives such that

1. *$V(\mathbf{x})$ is positive definite (locally in $\boldsymbol{B}_{R_0}$)*
2. *$\dot{V}(\mathbf{x})$ is negative semi-definite (locally in $\boldsymbol{B}_{R_0}$)*

then the equilibrium point $\mathbf{0}$ is stable. If, actually, the derivative $\dot{V}(\mathbf{x})$ is locally negative definite in $\boldsymbol{B}_{R_0}$, then the stability is asymptotic.

Theorem 2.3 ***Lyapunov Theorem for Global Stability***
Assume that there exists a scalar function V of the state $(\mathbf{x})$, with continuous first order derivatives such that

1. *$V(\mathbf{x})$ is positive definite*
2. *$\dot{V}(\mathbf{x})$ is negative definite*
3. *$V(\mathbf{x}) \to \infty$ as $||\mathbf{x}|| \to \infty$*

then the equilibrium at the origin is globally asymptotically stable.

Lyapunov Direct Method for Non-Autonomous Systems

We give some preliminary definitions.

Definition 2.15 *A scalar time-varying function $V(\mathbf{x}, t)$ is* locally positive definite *if $V(\mathbf{x}, t) = 0$ and there exists a time-invariant positive definite function $V_0(\mathbf{x})$ such that*

$$\forall t \geq t_0, \quad V(\mathbf{x}, t) \geq V_0(\mathbf{x}). \tag{2.16}$$

In Lyapunov analysis of non-autonomous systems, the concept of decreasing functions is also necessary.

Definition 2.16 *A scalar function $V(\mathbf{x}, t)$ is said to be* decreasing *if $V(\mathbf{x}, t) = 0$, and if there exists a time-invariant positive definite function $V_1(\mathbf{x})$ such that*

$$\forall t \geq 0, \quad V(\mathbf{x}, t) \leq V_1(\mathbf{x}). \tag{2.17}$$

The main Lyapunov stability results for non-autonomous systems can be summarized by the following theorem.

Theorem 2.4 ***(Lyapunov theorem for non-autonomous systems)***

Stability**:** *If, in a ball $\boldsymbol{B}_{R_0}$ around the equilibrium point $\mathbf{0}$, there exists a scalar function $V(\mathbf{x}, t)$ with continuous partial derivatives such that*

1. *V is positive definite*
2. *$\dot{V}$ is negative semi-definite*

then the equilibrium point $\mathbf{0}$ is stable in the sense of Lyapunov.

***Uniform stability and uniform asymptotic stability*:** *If, furthermore,*

3. V is decreasing

then the origin is uniformly stable. If condition 2 is strengthened by requiring that $\dot{V}$ be negative definite, then the equilibrium point is uniformly asymptotically stable.

***Global uniform asymptotic stability*:** *If the ball $\boldsymbol{B}_{R_0}$ is replaced by the whole state space, and condition 1, the strengthened condition 2, condition 3, and the condition*

4. $V(\mathbf{x}, t)$ is radially unbounded[1]

are all satisfied, then the equilibrium point at $\mathbf{0}$ is globally uniformly asymptotically stable.

Lyapunov Functions for Linear Time-Invariant (LTI) Systems

Given a linear system of the form $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$, let us consider a quadratic Lyapunov function candidate

$$V = \mathbf{x}^T \mathbf{P} \mathbf{x}$$

where $\mathbf{P}$ is a given symmetric positive definite matrix[2]. Differentiating the positive definite function V along the system trajectory yields another quadratic form

$$\dot{V} = \dot{\mathbf{x}}^T \mathbf{P} \mathbf{x} + \mathbf{x}^T \mathbf{P} \dot{\mathbf{x}} = -\mathbf{x}^T \mathbf{Q} \mathbf{x} \qquad (2.18)$$

where

$$\mathbf{A}^T \mathbf{P} + \mathbf{P} \mathbf{A} = -\mathbf{Q} \qquad (2.19)$$

The question, thus, is to determine whether the symmetric matrix $\mathbf{Q}$ defined by the so-called ***Lyapunov equation*** (2.19) above, is itself positive definite. If this is the case, then V satisfies the conditions of the basic theorem of the previous section, and the origin is globally asymptotically stable.
A more useful way of studying a given linear system using scalar quadratic functions is, instead, to derive a positive definite matrix $\mathbf{P}$ from a given positive definite matrix $\mathbf{Q}$, i.e.

- choose a positive definite matrix $\mathbf{Q}$
- solve for $\mathbf{P}$ from the Lyapunov equation (2.19)
- check whether $\mathbf{P}$ is p.d.

[1] We say that a function $V(\mathbf{x})$ is *radially unbounded* if $V(\mathbf{x}) \to \infty$ as $||\mathbf{x}|| \to \infty$

[2] A square symmetric $n \times n$ matrix $\mathbf{M}$ is *positive definite* (p.d.) if $\mathbf{x} \neq \mathbf{0} \;\Rightarrow\; \mathbf{x}^\mathbf{T} \mathbf{M} \mathbf{x} > 0$

If $\mathbf{P}$ is p.d., then $\mathbf{x}^T\mathbf{P}\mathbf{x}$ is a Lyapunov function for the linear system and global asymptotical stability is guaranteed.
Unlike the previous approach of going from a given $\mathbf{P}$ to a matrix $\mathbf{Q}$, this technique of going from a given $\mathbf{Q}$ to a matrix $\mathbf{P}$ always leads to conclusive results for stable linear systems.

Theorem 2.5 *A necessary and sufficient condition for a LTI system $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$ to be strictly stable is that, for any symmetric p.d. matrix $\mathbf{Q}$, the unique matrix $\mathbf{P}$ solution of the Lyapunov equation (2.19) be symmetric positive definite.*

This theorem shows that *any* positive definite matrix $\mathbf{Q}$ can be used to determine the stability of a linear system. A simple choice of $\mathbf{Q}$ is the identity matrix. Physical concepts like energy may lead us to some uniquely effective choices of Lyapunov functions.

Existence of Lyapunov Function

There exists also converse theorem for Lyapunov stability theorem.

Theorem 2.6 ***(Stability)***
If the origin of (2.6) is stable, there exists a positive definite function $V(\mathbf{x}, t)$ with a non-positive derivative.

This theorem indicates the existence of a Lyapunov function for every stable system.

Theorem 2.7 ***(Uniform Asymptotic Stability)***
If the equilibrium point at the origin is uniformly asymptotically stable, there exists a positive definite and decreasing function $V(\mathbf{x}, t)$ with a negative definite derivative.

This theorem is theoretically important because it is useful in establishing robustness of uniform asymptotic stability to persistent disturbance.

2.1.4 Stability Definitions for Multiagent Dynamics

The dynamical equations used to model a network of n interacting agents that share a common state space $\mathcal{X}$ (which is now assumed to be finite-dimensional Euclidean) is

$$x(t+1) = f(t, x(t)) \tag{2.20}$$

or, expressed in terms of the individual agents' states

$$\begin{aligned} x_1(t+1) &= f_1(t, x_1(t), ..., x_n(t)) \\ &\vdots \\ x_n(t+1) &= f_n(t, x_1(t), ..., x_n(t)). \end{aligned}$$

In order to enable a clear and precise formulation of the stability and convergence properties of the discrete-time system (2.20) we extend the familiar stability concepts of Lyapunov theory to the present framework. Notice that we are interested in the agents' states converging to a common, constant value and that we expect this common value to depend continuously on the initial states. In other words, we are dealing with a continuum of equilibrium points. This means that the classical stability concepts developed for the study of individual, typically isolated, equilibria are not well-adapted to the present situation. Alternatively, one may shift attention away from the individual equilibria and consider the stability properties of the set of equilibria instead.

The stability notions that we introduce below incorporate, by definition, the requirement that all trajectories converge to one of the equilibria. In the following definition, we make a conceptual distinction between equilibrium solutions and equilibrium points: an equilibrium point is an element of the state space which is the constant value of an equilibrium solution. By referring explicitly to equilibrium solutions in the following definition we distinguish the present stability concepts from the more familiar set stability concepts.

Definition 2.17 *(Stability):*
Let $\mathcal{X}$ be a finite-dimensional Euclidean space and consider a continuous map $f : \mathbb{N} \times \mathcal{X} \to \mathcal{X}$ giving rise to the discrete-time system

$$x(t+1) = f(t, x(t)) \tag{2.21}$$

Consider a collection of equilibrium solutions of (2.21) and denote the corresponding set of equilibrium points by Φ. With respect to the considered collection of equilibrium solutions, the dynamical system (2.21) is called

1. stable *if for each $\phi_1 \in \Phi$, for all $c_2 > 0$ and for all $t_0 \in \mathbb{N}$ there is $c_1 > 0$ such that every solution ξ of (2.21) satisfies: if $|\xi(t_0) - \phi_1| < c_1$ then there is $\phi_2 \in \Phi$ such that $|\xi(t) - \phi_2| < c_2$ for all $t \geq t_0$;*

2. bounded *if for each $\phi_1 \in \Phi$, for all $c_1 > 0$ and for all $t_0 \in \mathbb{N}$ there is $c_2 > 0$ such that every solution ξ of (2.21) satisfies: if $|\xi(t_0) - \phi_1| < c_1$ then there is $\phi_2 \in \Phi$ such that $|\xi(t) - \phi_2| < c_2$ for all $t \geq t_0$;*

3. globally attractive *if for each $\phi_1 \in \Phi$, for all $c_1, c_2 > 0$ and for all $t_0 \in \mathbb{N}$ there is $T \geq 0$ such that every solution ξ of (2.21) satisfies: if $|\xi(t_0) - \phi_1| < c_1$ then there is $\phi_2 \in \Phi$ such that $|\xi(t) - \phi_2| < c_2$ for all $t \geq t_0 + T$;*

4. globally asymptotically stable *if it is stable, bounded and globally attractive.*

Definition (2.17) may be interpreted as follows. Stability and boundedness require that any solution of (2.21) which is initially close to Φ remains close to one of the equilibria in Φ, thus excluding, for example, the possibility of drift along the set Φ. Global attractivity implies that every solution of (2.21) converges to one of the equilibria in Φ.
If the collection of equilibrium solutions is a singleton consisting of one equilibrium solution, then the notions of stability, boundedness, global attractivity and global asymptotic stability of Definition (2.17) coincide with the classical notions that have been introduced for the study of individual equilibria. In general, however, the notions introduced here are strictly stronger than their respective counterparts from set stability theory.

The following theorem provides a sufficient condition for uniform stability, uniform boundedness and uniform global asymptotic stability in terms of the existence of a set-valued[3] Lyapunov function.

Theorem 2.8 *(Lyapunov Characterization):*
Let $\mathcal{X}$ be a finite-dimensional Euclidean space and consider a continuous map $f : \mathbb{N} \times \mathcal{X} \to \mathcal{X}$ giving rise to the discrete-time system (2.21). Let Ξ be a collection of equilibrium solutions of (2.21) and denote the corresponding set of equilibrium points by Φ. Consider an upper semicontinuous[4] set-valued function $V : \mathcal{X} \rightrightarrows \mathcal{X}$ satisfying

1. $x \in V(x), \ \forall x \in \mathcal{X}$;
2. $V(f(t,x)) \subseteq V(x), \ \forall t \in \mathbb{N}, \ \forall x \in \mathcal{X}$.

If $V(\phi) = \{\phi\}$ for all $\phi \in \Phi$, then the dynamical system (2.21) is uniformly stable with respect to Ξ. If $V(x)$ is bounded for all $x \in \mathcal{X}$, then the dynamical system (2.21) is uniformly bounded with respect to Ξ. Consider in addition a function $\mu : \mathrm{Image}(V) \to \mathbb{R}_{\geq 0}$ and a lower semicontinuous function $\beta : \mathcal{X} \to \mathbb{R}_{\geq 0}$ satisfying

3. $\mu \circ V : \mathcal{X} \to \mathbb{R}_{\geq 0} : x \mapsto \mu(V(x))$ *is bounded on bounded subsets of* $\mathcal{X}$;
4. β *is positive definite with respect to* Φ; *that is,* $\beta(\phi) = 0$ *for all* $\phi \in \Phi$ *and* $\beta(x) > 0$ *for all* $x \in \mathcal{X} \setminus \Phi$;

---

[3] In mathematics, a multivalued function (shortly multifunction), also named *set-valued function*, is a left-total relation, i.e. every input is associated with one or more outputs. E.g., let $\dot{x} = f(x,u)$, with $u \in \mathcal{U}$ be a function in the two variables x and u; then $\dot{x} \in f(x,\mathcal{U})$ is a multifunction, where the variable u has been replaced by the set $\mathcal{U}$.

[4] In mathematical analysis, a set-valued map $F : X \to Y$ is called upper semicontinuous at $x \in \mathrm{Dom}(F)$ if and only if for any neighborhood $\mathcal{U}$ of $F(x)$, $\exists \eta > 0$ *such that* $\forall x' \in B_x(x,\eta)$, $F(x') \subset \mathcal{U}$. It is said to be upper semicontinuous if and only if it is upper semicontinuous at any point of $\mathrm{Dom}(F)$.
A set-valued map $F : X \to Y$ is called lower semicontinuous at $x \in \mathrm{Dom}(F)$ if and only if for any $y \in F(x)$ and for any sequence of elemnts $x_n \in \mathrm{Dom}(F)$ converging to x, there exists a sequence of elements $y_n \in F(x_n)$ converging to y. It is said to be lower semicontinuous if it is lower semicontinuous at every point $x \in \mathrm{Dom}(F)$.

5. *$\mu(V(f(t,x))) - \mu(V(x)) \leq -\beta(x),\ \forall t \in \mathbb{N},\ \forall x \in \mathcal{X}$.*

If $V(\phi) = \{\phi\}$ for all $\phi \in \Phi$ and $V(x)$ is bounded for all $x \in \mathcal{X}$ then the dynamical system (2.21) is uniformly globally asymptotically stable with respect to Ξ.

The set-valued function V plays the role of a Lyapunov function which is nonincreasing along the solutions of (2.21). The set-valued nature of V is important: unlike a real-valued function, a set-valued function allows for a continuum of minima which are not comparable with each other. For this reason, a set-valued Lyapunov function, unlike a real Lyapunov function, may be used to conclude that each trajectory converges to one equilibrium out of a continuum of equilibria. The function μ serves as a measure for the size of the values of V.

The strict decrease condition (Condition 5) of Theorem 2.8 may be considerably relaxed. Consider, for example, the following condition which requires that $\mu \circ V$ decreases over time-intervals of length τ.

6. *There is a time $\tau \in \mathbb{N}$ such that $\forall t \in \mathbb{N},\ \forall x \in \mathcal{X}$*

$$\mu(V(f(t+\tau-1, ...f(t+1, f(t,x))...))) - \mu(V(x)) \leq -\beta(x).$$

Theorem 2.8 is still true if Condition 5) is replaced by Condition 6).

2.2 Some Notions for Nonlinear Spaces

For a more detailed discussion we refer the reader to [5, 6, 9, 17, 18].

2.2.1 Manifold

In mathematics (specifically in differential geometry and topology), a manifold is a topological space that on a small enough scale resembles the Euclidean space of a specific dimension, called the dimension of the manifold. Thus, a line and a circle are one-dimensional manifolds, a plane and a sphere are two-dimensional manifolds, and so on into high-dimensional space. Informally, a manifold is a space that is "modeled on" Euclidean space. There are many different kinds of manifolds and generalizations. In geometry and topology, all manifolds are topological manifolds, possibly with additional structure, most often a differentiable structure. More formally, every point of an n-dimensional manifold has a neighborhood homeomorphic to an open subset of the n-dimensional space $\mathbb{R}^n$.
Roughly speaking, a homeomorphism is a continuous stretching and bending of an object (a topological space can be regarded as a geometric object) into a new shape. More formally, a function $f : X \to Y$ between two topological spaces is said to be a homeomorphism if it is a bijection, continuous and

has continuous inverse function f^{-1}. If such a function exists between two spaces, we say they are homeomorphic.

Although manifolds resemble Euclidean spaces near each point (*locally*), the global structure of a manifold may be more complicated. For example, any point on the usual two-dimensional surface of a sphere is surrounded by a circular region that can be flattened to a circular region of the plane, as in a geographical map. However, the sphere differs from the plane *in the large*: in the language of topology, they are not homeomorphic. The structure of a manifold is encoded by a collection of charts that form an atlas, in analogy with an atlas consisting of charts of the surface of the Earth.

The broadest common definition of manifold is a topological space locally homeomorphic to a topological vector space over the reals. This omits finite dimension, allowing structures such as Hilbert manifolds to be modeled on Hilbert spaces, Banach manifolds to be modeled on Banach spaces.
The study of manifolds combines many important areas of mathematics: it generalizes concepts such as curves and surfaces as well as ideas from linear algebra and topology.

Homogeneous Manifold

Informally, a homogeneous manifold $\mathcal{M}$ can be seen as a manifold on which "all points are equivalent". The present work considers connected compact homogeneous (CCH) manifolds satisfying the following embedding property.

Assumption 2.1 *$\mathcal{M}$ is a CCH manifold smoothly embedded in $\mathcal{E} \subseteq \mathbb{R}^n$ with the Euclidean norm $||y|| = r_{\mathcal{M}}$ constant over $y \in \mathcal{M}$.*

It is sometimes preferred to represent $y \in \mathcal{M}$ by a matrix instead of a vector. The corresponding norm is the Frobenius norm $||B|| = \sqrt{\text{trace}(B^T B)}$.

Riemannian Manifold

To measure distances and angles on manifolds, the manifold must be Riemannian.

In Riemannian geometry and differential geometry of surfaces, a Riemannian manifold or Riemannian space $(\mathcal{M}, g)$ is a real differentiable manifold $\mathcal{M}$ in which each tangent space is equipped with an inner product g, a Riemannian metric, which varies smoothly from point to point. A Riemannian metric makes it possible to define various geometric notions on a Riemannian manifold, such as angles, lengths of curves, areas (or volumes), curvature, gradients of functions and divergence of vector fields.

2.2.2 Geodesic Distance

In metric geometry, a geodesic is a curve which is everywhere locally a distance minimizer. More precisely, a curve $\gamma : I \to \mathcal{M}$ from an interval I of the reals to the metric space $\mathcal{M}$ is a geodesic if there is a constant $\nu \geq 0$ such that for any $t \in I$ there is a neighborhood J of t in I such that for any $t_1, t_2 \in J$ we have

$$d(\gamma(t_1), \gamma(t_2)) = \nu|t_1 - t_2|.$$

This generalizes the notion of geodesic for Riemannian manifolds. However, in metric geometry the geodesic is often considered with natural parametrization, i.e. in the above identity $\nu = 1$ and

$$d(\gamma(t_1), \gamma(t_2)) = |t_1 - t_2|.$$

If the last equality is satisfied for all $t_1, t_2 \in I$ the geodesic is called a *minimizing geodesic* or *shortest path.*

The distance $d(p, q)$ between two points p and q of a Riemannian manifold $\mathcal{M}$ is defined as the infimum of the length taken over all continuous, piecewise continuously differentiable curves $\gamma : [a, b] \to \mathcal{M}$ such that $\gamma(a) = p$ and $\gamma(b) = q$. With this definition of distance, geodesics in a Riemannian manifold are then the locally distance-minimizing paths.

2.2.3 Lie Groups

Lie groups, named after Sophus Lie, are differentiable manifolds that carry also the structure of a group which is such that the group operations are defined by smooth maps.
A simple example of a compact Lie group is the circle: the group operation is simply rotation. This group, known as $U(1)$, can be also characterized as the group of complex numbers of modulus 1 with multiplication as the group operation.
Another example is the special orthogonal Lie group $SO(n)$. This can be viewed as the set of positively oriented orthonormal bases of $\mathbb{R}^n$, or equivalently, as the set of rotation matrices in $\mathbb{R}^n$; it is the natural state space for the orientation of a rigid body in $\mathbb{R}^n$.

2.3 Notions of Graph Theory

In the framework of coordination with limited interconnections between agents, it is customary to represent communication links by means of a *graph*.

Definition 2.18 *A* direct graph $\mathbb{G}(\mathcal{V},\mathcal{E})$ *(short* digraph $\mathbb{G}$*) is composed of a finite set* $\mathcal{V}$ *of* vertices*, and a set* $\mathcal{E}$ *of* edges *which represent interconnections among the vertices as ordered pairs (j,k) with* $j,k \in \mathcal{V}$*.*

A *weighted digraph* $\mathbb{G}(\mathcal{V},\mathcal{E},\mathcal{A})$ is a digraph associated with a set $\mathcal{A}$ that assigns a positive *weight* $a_{jk} \in \mathbb{R}_{>0}$ to each edge $(j,k) \in \mathcal{E}$.

A digraph is said to be *undirected* if $a_{jk} = a_{kj}\ \forall j,k \in \mathcal{V}$. If $(j,k) \in \mathcal{E}$ whenever $(k,j) \in \mathcal{E}\ \forall j,k \in \mathcal{V}$ but $a_{jk} \neq a_{kj}$ for some $j,k \in \mathcal{V}$, then the graph is called *bidirectional*.

For the consensus problem, each agent is identified with a vertex of a graph; the N agents=vertices are indicated by positive integers $1,2,...,N$, so $\mathcal{V} = \{1,2,...,N\}$. The presence of edge (j,k) has the meaning that agent j sends information to agent k, or equivalently agent k measures quantities concerning agent j. It is assumed that no communication link is needed for an agent to get information about itself, so $\mathbb{G}$ contains no self-loops: $(k,k) \notin \mathcal{E}\ \forall k \in \mathcal{V}$.

A frequent alternative notion for $(j,k) \in \mathcal{E}$ is $j \rightsquigarrow k$. We say that j is an *in-neighbor* of k and k is an *out-neighbor* of j.

For an undirected graph, all arrows are bidirectional; therefore we simply say that j and k are *neighbors* and write $j \sim k$.

The *in-degree* of vertex k is $d_k^{(i)} = \sum_{j=1}^{N} a_{jk}$. The *out-degree* of vertex k is $d_k^{(o)} = \sum_{j=1}^{N} a_{kj}$. A digraph is said to be *balanced* if $d_k^{(i)} = d_k^{(o)}\ \forall k \in \mathcal{V}$; in particular, undirected graphs are balanced.

The *adjacency matrix* $A \in \mathbb{R}^{N\times N}$ of a graph $\mathbb{G}$ contains a_{jk} in row j, column k; it is symmetric if and only if $\mathbb{G}$ is undirected. The in- and out-degrees of vertices $1,2,...N$ can be assembled in diagonal matrices $D^{(i)}$ and $D^{(o)}$.

The Laplacian matrix $\mathcal{L} \in \mathbb{R}^{m\times m}$ is defined as $\mathcal{L}_{ii} = \sum_{j\neq i} a_{ij}$, $\mathcal{L}_{ij} = -a_{ij}$ for $i \neq j$. Matrix $\mathcal{L}$ is symmetric if the graph is undirected, and also satisfies the following property:

- All the eigenvalues of $\mathcal{L}$ have nonnegative real parts. Zero is an eigenvalue of $\mathcal{L}$, with $\mathbf{1}$ as the corresponding right eigenvector;
- Zero is a simple eigenvalue of $\mathcal{L}$ if and only if graph $\mathbb{G}$ has a directed spanning tree[5];

[5] A spanning tree $\mathbb{T}$ of a connected, undirected graph $\mathbb{G}$ is a tree composed of all the vertices and some (or perhaps all) of the edges of $\mathbb{G}$. Informally, a spanning tree of $\mathbb{G}$ is a

- If graph $\mathbb{G}$ contains a directed spanning tree, then, with proper permutation, $\mathcal{L}$ can be reduced to the Frobenius normal form

$$\mathcal{L} = \begin{bmatrix} L_{11} & L_{12} & \dots & L_{1k} \\ 0 & L_{22} & \dots & L_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & L_{kk} \end{bmatrix}$$

where $\mathcal{L}_{ii},\ i = 1, ..., k-1$, are irreducible, each $\mathcal{L}_{ii}$ has at least one row with positive row sum, and $\mathcal{L}_{kk}$ is irreducible or is a zero matrix of dimension one.

The *in-Laplacian* of $\mathbb{G}$ is $\mathcal{L}^{(i)} = D^{(i)} - A$. Similarly, the associated *out-Laplacian* is $\mathcal{L}^{(o)} = D^{(o)} - A$. For a balanced graph $\mathbb{G}$, the *Laplacian* $\mathcal{L} = \mathcal{L}^{(i)} = \mathcal{L}^{(o)}$. The standard definition of Laplacian $\mathcal{L}$ is for undirected graph. For the latter, $\mathcal{L}$ is symmetric. For general digraphs, by construction, $(\mathbf{1}_N)^T \mathcal{L}^{(i)} = 0$ and $\mathcal{L}^{(o)} \mathbf{1}_N = 0$ where $\mathbf{1}_N$ is the column vector of N ones. The spectrum of the Laplacian reflects several interesting properties of the associated graph. In particular, it reflects its *connectivity* properties.

A *directed path* of length l from vertex j to vertex k is a sequence of vertices $v_0, v_1, ..., v_l$ with $v_0 = j$ and $v_l = k$ and such that $(v_m, v_{m+1}) \in \mathcal{E}$ for $m = 0, 1, ..., l-1$. An *undirected path* between vertices j and k is a sequence of vertices $v_0, v_1, ..., v_l$ with $v_0 = j$ and $v_l = k$ and such that $(v_m, v_{m+1}) \in \mathcal{E}$ or $(v_{m+1}, v_m) \in \mathcal{E}$ for $m = 0, 1, ..., l-1$.

A digraph $\mathbb{G}$ is *strongly connected* if it contains a directed path from every vertex to every other vertex (and thus also back to itself). A digraph $\mathbb{G}$ is *root-connected* if it contains a node k, called *root*, from which there is a path to every other vertex (but not necessarily back to itself). A digraph $\mathbb{G}$ is *weakly-connected* if it contains an undirected path between any two of its vertices. For an undirected graph $\mathbb{G}$, all these notions become equivalent and are simply summarized by the term *connected*.

For $\mathbb{G}$ representing interconnections in a network of agents, coordination can only take place if $\mathbb{G}$ is connected. If this is not the case, coordination will only be achievable separately in each connected component of $\mathbb{G}$.

When the graph $\mathbb{G}$ can vary with time, the communication links are represented by a *time-varying graph* $\mathbb{G}(t)$ in which the vertex set $\mathcal{V}$ is fixed (by convention), but edges $\mathcal{E}$ and weights $\mathcal{A}$ can depend on time. To prevent edges from vanishing or growing indefinitely, the present work considers δ - digraphs, for which the element of $A(t)$ are bounded and satisfy the threshold $a_{jk}(t) \geq \delta > 0\ \forall (j,k) \in \mathcal{E}(t)$, for all t.

---

selection of edges of $\mathbb{G}$ that form a tree spanning every vertex. That is, every vertex lies in the tree, but no cycles (or loops) are formed. On the other hand, every bridge of $\mathbb{G}$ must belong to $\mathbb{T}$.

In addition, in continuous-time $\mathbb{G}$ is assumed to be piecewise continuous. For δ-digraphs $\mathbb{G}(t)$, it is intuitively clear that coordination may be achieved if information exchange is "sufficiently frequent", without requiring it to take place all the time.

Definition 2.19 *In discrete-time, for a δ-digraph $\mathbb{G}(\mathcal{V}, \mathcal{E}(t), \mathcal{A}(t))$ and some constant $T \in \mathbb{Z}_{\geq 0}$, define the graph $\overline{\mathbb{G}}(\mathcal{V}, \overline{\mathcal{E}}(t), \overline{\mathcal{A}}(t))$ where $\overline{\mathcal{E}}(t)$ contains all edges that appear in $\mathbb{G}(\tau)$ for $\tau \in [t, t+T]$ and $\overline{a}_{jk}(t) = \sum_{\tau=t}^{t+T} a_{jk}(\tau)$.*
Similarly, in continuous-time, for a δ-digraph $\mathbb{G}(\mathcal{V}, \mathcal{E}(t), \mathcal{A}(t))$ and some constant $T \in \mathbb{R}_{>0}$, define the graph $\overline{\mathbb{G}}(\mathcal{V}, \overline{\mathcal{E}}(t), \overline{\mathcal{A}}(t))$ by

$$\overline{a}_{jk}(t) = \begin{cases} \int_t^{t+T} a_{jk}(\tau)d\tau & \text{if } \int_t^{t+T} a_{jk}(\tau)d\tau \geq \delta \\ 0 & \text{if } \int_t^{t+T} a_{jk}(\tau)d\tau < \delta \end{cases}$$

$(j, k) \in \overline{\mathcal{E}}(t)$ if and only if $\overline{a}_{jk}(t) \neq 0$.
Then $\mathbb{G}(t)$ is said to be uniformly connected over *T if there exists a time horizon T and a vertex $k \in \mathcal{V}$ such that $\overline{\mathbb{G}}(t)$ is root-connected with root k for all t.*

Chapter 3

The Consensus Problem

The distributed computation of means/averages of datasets (in an algorithmic setting) and the synchronization of a set of agents (in a control setting) - i.e., driving all the agents to a common point in state space - are ubiquitous tasks in current engineering problems.

Practical applications include autonomous swarm/formation operation, distributed decision making, neural and communication networks, clustering and other reduction methods, optimal coding, and other fields where averaging/synchronizing or distributing a set of points appears as a subproblem.

Synchronization algorithms are well understood in Euclidean spaces. They are based on the natural definition and distributed computation of the centroid in $\mathbb{R}^m$. However, many of the applications above involve manifolds that are not homeomorphic to an Euclidean space. Even for formations moving in $\mathbb{R}^2$ or $\mathbb{R}^3$, the agents' orientations evolve in a manifold $SO(2) \cong S^1$ or $SO(3)$[1].

Most of the work related to synchronization and balancing on manifolds concerns the circle S^1. The most extensive literature on the subject derives from the Kuramoto model.

In accordance with the consensus approach, the agents are reduced to kinematic models and the focus is on (almost) global convergence for various agent interconnections, without any leader or external reference. Consensus among a group of agents depends on the available communication links. When considering limited agent interconnections, it is customary to represent communication links by means of a graph.

[1] We denote with $SO(n)$ the group of rotations, with $SE(2)$ the group of rigid motions in the plane, and with $SE(3)$ the group of rigid motions in the space

3.1 Consensus in Linear Spaces

Linear consensus algorithm describes the behavior of N agents *locally* exchanging information about their state $x_k \in \mathbb{R}^n$, $k \in \mathcal{V} = \{1, 2, ..., N\}$, in order to asymptotically reach a *global* consensus, i.e. a common value of agreement.

In continuous-time, the update is

$$\frac{d}{dt} x_k(t) = \sum_{j=1}^{N} a_{jk}(t)(x_j(t) - x_k(t)), \quad k = 1, 2, ..., N \tag{3.1}$$

where a_{jk} is the weight of link $j \rightsquigarrow k$: it is the entry, at time t, of adjacency matrix $A \in \mathbb{R}^{n \times n}$ associated with graph $\mathbb{G}$ representing the communication topology. The state of agent k evolves towards to the (positively weighted) arithmetic mean of its neighbors, $\frac{1}{d_k^i} \sum_{j \rightsquigarrow k} a_{jk} x_j$ where $d_k^{(i)}$ is the in-degree $d_k^{(i)} = \sum_{j \rightsquigarrow k} a_{jk}$.

The corresponding update in discrete-time is

$$x_k(t+1) = \frac{1}{\beta_k(t) + d_k^{(i)}(t)} \left(\sum_{j \rightsquigarrow k} a_{jk}(t) x_j(t) + \beta_k(t) x_k(t) \right), \tag{3.2}$$

with non-vanishing weight $\beta_k(t) \geq \beta_0 > 0$, which represent inertia values. In a more compact form, the update in discrete-time can be written as

$$x_k(t+1) = \sum_{j=1}^{N} a_{jk}(t) x_j(t) = x_k(t) + \sum_{j \neq k} a_{jk}(t)(x_j(t) - x_k(t)). \tag{3.3}$$

The weights a_{jk} induce a communication graph between the agents. They can be asymmetric (leading to a *directed* communication graph) and/or depend on time (leading to a *time-varying* communication graph).

In matrix form, the continuous-time algorithm is the linear time-varying system

$$\dot{x}(t) = -\mathcal{L}(t) x(t) \tag{3.4}$$

where $\mathcal{L}(t) = D^{(i)}(t) - A(t)$ is the Laplacian matrix associated with $\mathbb{G}$, while the discrete-time algorithm is the linear time-varying system

$$x(t+1) = \mathcal{M}(t) x(t) \tag{3.5}$$

where $\mathcal{M}(t) = (D^{(i)}(t) + B(t))^{-1}(A(t) + B(t))$ and $B(t)$ is a diagonal matrix with elements $B_{kk} = \beta_k$. The matrices $\mathcal{M}(t)$ are stochastic, i.e. the elements of each row sum to one.

Example 3.1 *Let the set of states of a discrete-time system be represented by the (balanced) digraph in the figure below, with $x_1 = 1,\ x_2 = 2,\ x_3 = 3$ be the corresponding values.*

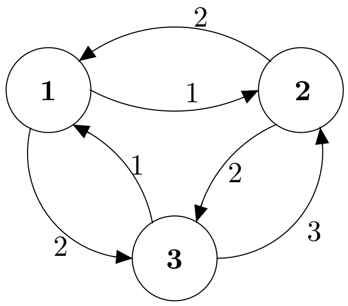

Figure 3.1: Consensus in geodetically and non-geodetically convex set.

The matrix of the inertia values β is

$$B = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

The adjacency matrix A is

$$A = \begin{bmatrix} 0 & 1 & 2 \\ 2 & 0 & 2 \\ 1 & 3 & 0 \end{bmatrix}$$

The Laplacian (as the graph is balanced)

$$\mathcal{L} = \mathcal{L}^{(i)} = \mathcal{L}^{(i)} = \begin{bmatrix} 3 & -1 & -2 \\ -2 & 4 & -2 \\ -1 & -3 & 4 \end{bmatrix}$$

Thus,

$$D = \mathcal{L} + A = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

Finally, we we have that

$$\begin{aligned} M &= (D+B)^{-1}(A+B) = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 7 \end{bmatrix}^{-1} \begin{bmatrix} 2 & 1 & 2 \\ 2 & 4 & 2 \\ 1 & 3 & 3 \end{bmatrix} = \\ &= \begin{bmatrix} 1/5 & 0 & 0 \\ 0 & 1/8 & 0 \\ 0 & 0 & 1/7 \end{bmatrix} \begin{bmatrix} 2 & 1 & 2 \\ 2 & 4 & 2 \\ 1 & 3 & 3 \end{bmatrix} = \begin{bmatrix} 2/5 & 1/5 & 2/5 \\ 2/8 & 4/8 & 2/8 \\ 1/7 & 3/7 & 3/7 \end{bmatrix} \end{aligned}$$

which is a stochastic matrix.

The update guarantees asymptotic consensus under minimal connectivity assumptions of the underlying communication graph.

Property 3.1 *Consider a set of N agents evolving on $\mathbb{R}^n$ according to (continuous-time) (3.1) or according to (discrete-time) (3.2). Then the agents globally and exponentially converge to a consensus value $\overline{x} \in \mathbb{R}^n$ if the communication among agents is characterized by a (piecewise) continuous δ-digraph which is uniformly connected.*

If in addition, $\mathbb{G}$ is balanced for all times, then the consensus value is the arithmetic mean of the initial values: $\overline{x} = \frac{1}{N}\sum_{k=1}^{N} x_k(0)$.

When interconnections are not only balanced, but also undirected and fixed, then the linear consensus algorithm is a gradient descent[2] algorithm for the **disagreement cost function**

$$V_{vect}(x) = \frac{1}{2}\sum_{k=1}^{N}\sum_{j=1}^{N} a_{jk}||x_j - x_k||^2 = x^T(\mathcal{L} \otimes I_n)x \tag{3.6}$$

where $||z||$ denotes the Euclidean norm $\sqrt{z^T z}$ of $z \in \mathbb{R}^m$, and $\mathcal{L}$ is the Laplacian matrix of $\mathbb{G}$, $x \in \mathbb{R}^{Nn}$ denotes the vector whose elements $(k-1)n+1$ to kn contain x_k, and $\otimes I_n$ is the Kronecker product[3] by the $n \times n$ identity matrix.

The disagreement cost function can also be written as

$$V(x) = \frac{1}{2(N-1)}\sum_{k=1}^{N}\sum_{j\neq k} ||x_k - x_j||^2 \tag{3.7}$$

where $x_k \in \mathbb{R}^n,\ k = 1, 2, ..., N$ are the states of the N agents. Then the gradient descent is

$$\dot{x}_k = -x_k + \frac{1}{N-1}\sum_{j\neq k} x_j \tag{3.8}$$

[2] Gradient descent is a first-order optimization algorithm. To find a local minimum of a function using gradient descent, one takes steps proportional to the negative of the gradient (or of the approximate gradient) of the function at the current point. If instead one takes steps proportional to the positive of the gradient, one approaches a local maximum of that function; the procedure is then known as gradient ascent.

[3] If A is an $m \times n$ matrix and B is a $p \times q$ matrix, then the Kronecker product $A \otimes B$ is the $mp \times nq$ block matrix

$$A \otimes B = \begin{bmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \dots & a_{mn}B \end{bmatrix}$$

We see that the term $\frac{1}{N-1}\sum_{k\neq j} x_j$ in the above equation is the arithmetic average of other agents values, that is argmin $\sum_{j\neq k} ||x_k - x_j||^2$.

The generalized update for linear consensus over a communication graph $\mathbb{G}$ is then

$$\dot{x}_k = -x_k + \frac{1}{d_k}\sum_{j\rightsquigarrow k} x_j \tag{3.9}$$

where $\frac{1}{d_k}\sum_{j\rightsquigarrow k} x_j$ is the arithmetic average of neighbors, i.e. $\text{argmin}_x \sum_{j\rightsquigarrow k} ||x - x_j||^2$.

The gradient can no longer be considered, as in the previous equation, if the graph $\mathbb{G}$ is directed and/or time-varying.

Consensus is now **"move towards the average of your neighbors"**:

$$\begin{aligned} \dot{x}_k &= -x_k + mean(x_j : j \rightsquigarrow k) \\ x_k^+ &= \alpha x_k + (1-\alpha) mean(x_j : j \rightsquigarrow k),\ 0 < \alpha < 1. \end{aligned}$$

Consensus algorithms provide a basic model of *distributed* computation: a given agent at a given time only performs *local average* computations but the spread of information over time eventually enables the computation of a centralized quantity (the arithmetic average of initial states under a balancing assumption).

An essential feature of consensus algorithms is their symmetry, i.e. their invariance properties to certain transformations: the dynamics are invariant under reordering of the agents (discrete permutation symmetry) and under uniform translation of the state (continuous symmetry): for any $a \in \mathbb{R}^n$, a shifted initial condition $y_k(0) = x_k(0) + a\ \forall k \in \mathcal{V}$ yields the shifted solution $y_k(t) = x_k(t) + a\ \forall k \in \mathcal{V}$ and $\forall t \geq 0$. This is because the exchange of information only involves *relative* quantities $(x_j - x_k)$. As a consequence, the distributed computation of a centralized quantity does not require that all agents share a (centralized) common reference frame.

Convergence Analysis

The convergence analysis of linear consensus algorithms is the convergence analysis of a time-varying linear system. For consensus algorithms defined in $\mathbb{R}$, the early analysis of Tsitsiklis (1984) [7] rests on the basic but fundamental observation that the (time-invariant) Lyapunov function

$$V(x) = \max_{1\leq k\leq N} x_k - \min_{1\leq k\leq N} x_k \tag{3.10}$$

is non-increasing along the solutions.

In discrete-time, the flow of the algorithm involves products of stochastic matrices and its convergence properties is related to ergodicity theorems in the theory of Markov chains.

In 2004-2005, Moreau emphasizes that the convergence result relies in an essential way on the *convexity* rather than the *linearity* of the update law: the position of each agent k for $t > \tau$ always lies in the convex hull of the $x_j(\tau),\ j = 1, 2, ..., N$. The permanent contraction of this convex hull, at some nonzero minimal rate because weights are non-vanishing, allows to conclude that the agents end up at a consensus value. This approach extends the Lyapunov function (3.10) to vector-valued algorithms and allows for the convergence analysis of nonlinear consensus algorithms provided that the convexity property is retained.

This convergence analysis rests on the fact that the convex hull $[\min x_k,\ \max x_k]$ cannot expand along the solutions. Convergence is ensured if $\mathbb{G}$ is root-connected. If $\mathbb{G}$ is time-varying, root-connectedness over a uniform horizon is still sufficient to ensure exponential convergence.

3.2 Consensus in Nonlinear Spaces

3.2.1 Riemannian Consensus

A geometric interpretation of the linear consensus algorithms (3.1) and (3.2) is to view the state $x_k(t)$ as the estimate at time t by agent k of the consensus value. At each time step, each agent updates its current estimate of the consensus value towards a (weighted) average of its neighbors estimates. Moreover, the weighted arithmetic average can be given the geometric interpretation of the point that minimizes the sum of the (weighted) squared distances:

$$\sum_{k=1}^{N} a_{jk} x_k = \min_{z \in \mathbb{R}^n} \sum_{k=1}^{N} a_{jk} ||z - x_k||^2. \tag{3.11}$$

With this geometric interpretation, consensus algorithms can be defined on arbitrary Riemannian manifolds. The Riemmanian (or Karcher 3.15) mean on a manifold $\mathcal{M}$ is defined by substituting the Riemannian (geodesic) distance for the Euclidean distance in (3.11):

$$\text{mean}(x_1, ..., x_N) = \min_{z \in \mathbb{R}^n} \sum_{k=1}^{N} a_{jk} d^2_{\mathcal{M}}(z, x_k). \tag{3.12}$$

Furthermore, on a Riemannian manifold, “updating a point towards a new point” simply translates as “moving along the geodesic path connecting the two points”. This approach yields an intrinsic definition of consensus algorithms on Riemannian manifolds: agents reach consensus if each point is the mean of its neighbors:

$$x_k \in \underset{z}{\operatorname{argmin}} \sum_{j \rightsquigarrow k} d^2(z, x_j). \tag{3.13}$$

Therefore, consensus minimizes the cost $\frac{1}{N} \sum_{k=1}^{N} \sum_{k \rightsquigarrow j} d^2(x_j, x_k)$ and the consensus algorithm is **"direct each point towards the mean of its neighbors"**.

However, this approach suffers both a fundamental limitation and a practical limitation. The fundamental limitation is that the uniqueness of a geodesic is only ensured locally. If several geodesics connect two points, both the concepts of mean and the concepts of "moving along shortest paths connecting two points" become non-unique. A practical limitation is that the computation of geodesics at each time step in a distributed algorithm might represent a formidable computational task.
Those limitations can be overcome to a large extent if the manifold is embedded in an Euclidean space and if mean and distance calculations are carried out in the Euclidean geometry of the ambient space.

An additional desired property of our generalized consensus algorithms is to retain the symmetry properties. In essence, distributed algorithms should be defined on spaces where all points "look alike". This is the case for Lie groups, and, more generally, homogeneous spaces.

Those considerations led us to consider state spaces that satisfy assumption 2.1.

The embedding space $\mathcal{E}$ denotes the linear vector space $\mathbb{R}^n$ or the linear matrix space $\mathbb{R}^{n \times n}$. The additional condition $||y|| = r_{\mathcal{M}}$ is in agreement with the fact that all points on $\mathcal{M}$ should be equivalent.
On a manifold that satisfies Assumption 2.1, a convenient alternative to the intrinsic generalization of consensus algorithms is to base the calculations on the distance of the ambient Euclidean space. Replacing the distance on $\mathcal{M}$ by the Euclidean distance in $\mathcal{E}$ in 3.12 leads to the **induced arithmetic mean** $IAM \subseteq \mathcal{M}$ of N agents of weights $a_{ij} > 0$ and positions $x_k \in \mathcal{M},\ k = 1, ..., N$:

$$IAM(x_1, ..., x_N) = \underset{z \in \mathcal{M}}{\arg\min} \sum_{k=1}^{N} a_{jk} ||\widehat{x}_k - \widehat{z}||^2. \tag{3.14}$$

The notation $\widehat{z}$ denotes the (vector) embedding of z in the linear space $\mathcal{E}$.

The IAM on $SO(3)$[4] is called the *projected arithmetic mean.* The point in its definition is that distances are measured in the embedding space $\mathbb{R}^m$. It

[4] $SO(3)$ denotes the rotation group of 3×3 orthogonal matrices.

thereby differs from the canonical definition of mean of N agents on $\mathcal{M}$, the *Karcher mean*, which uses the geodesic distance $d_{\mathcal{M}}$ along the Riemannian manifold $\mathcal{M}$ (with, in the present setting, the Riemannian metric induced by the embedding of $\mathcal{M}$ in $\mathbb{R}^m$) as follows:

$$C_{Karcher} = \operatorname*{argmin}_{c \in \mathcal{M}} \sum_{k=1}^{N} a_{jk} d^2_{\mathcal{M}}(x_k, c). \tag{3.15}$$

The IAM has the following properties:

1. The IAM of a single point y_1 is the point itself.
2. The IAM is invariant under permutations of agents of equal weights.
3. The IAM commutes with the symmetry group of the homogeneous manifold.
4. The IAM does not always reduce to a single point.

The main advantage over the Karcher mean is computational. The IAM is closely related to the weighted centroid in the ambient space, defined as

$$C_e(\widehat{x}_1, ..., \widehat{x}_N) = \frac{\sum_{k=1}^{N} a_{jk} \widehat{x}_k}{\sum_{k=1}^{N} a_{jk}}. \tag{3.16}$$

Since $||c|| = r_{\mathcal{M}}$ by Assumption 2.1, an equivalent definition for the IAM is

$$IAM = \operatorname{argmax}_{z \in \mathcal{M}} (\widehat{z}^T C_e). \tag{3.17}$$

Hence, computing the IAM just involves a search for the global maximizers of a linear function on $\mathcal{E}$. Moreover,

Assumption 3.1 *The local maxima of a linear function $f(\widehat{z}) = \widehat{z}^T b$ over $z \in \mathcal{M}$, with b fixed in $\mathcal{E}$, are all global maxima.*

Assumptions 2.1 and 3.1 are satisfied for a number of nonlinear spaces encountered in applications. Meaningful examples include the circle S^1, the unit sphere of $\mathcal{E}$ and the orthogonal group $SO(n)$.

The generalization of linear consensus algorithms to state spaces that satisfy Assumptions 2.1 and 3.1 is straightforward: the update is simply taken as a linear update towards the centroid in the ambient space, and then projected to the closest point of the manifold $\mathcal{M}$.
In continuous-time, this amounts to the differential equation

$$\dot{x}_k(t) = \text{Proj}_{T\mathcal{M}_{x_k}} \Big(\sum_j a_{jk} (\widehat{x}_j - \widehat{x}_k) \Big), \; k = 1...N. \tag{3.18}$$

Similarly, the discrete-time update is

$$x_k(t+1) = \mathrm{Proj}_{\mathcal{M}}\left(\frac{1}{\beta_k + d_k^{(i)}}\left(\sum_{j \rightsquigarrow k} a_{jk}\widehat{x}_j + \beta_k \widehat{x}_k\right)\right). \tag{3.19}$$

which can also be written as

$$x_k(t+1) \in \mathrm{IAM}(\{x_j(t) |\ j \rightsquigarrow k \text{ in } G(t)\} \cup \{x_k(t)\}). \tag{3.20}$$

3.2.2 Consensus on the Circle

The circle is a fundamental and representative example of nonlinear space. If consensus problem is well-defined on a geodetically convex set, it is fundamentally different on a non-geodetically one.

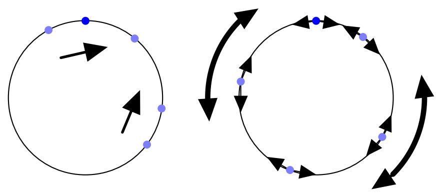

Figure 3.2: Consensus in geodetically and non-geodetically convex set.

Consider a set of N agents evolving on the circle S^1; the position of agent k on the circle is denoted by the angular variable $\theta_k \in S^1$. The interconnection (or communication) among agents is represented with a directed graph $\mathbb{G}$. A nonvanishing weight $a_{jk} \geq a_m$ can be associated to the link from node j to node k, for some fixed $a_m > 0$, and $a_{jk} = 0$ for $j \not\rightsquigarrow k$.

A classical (continuous-time) consensus algorithm in the vector space $\mathbb{R}$ reads

$$\frac{d}{dt}x_k = \sum_{j \rightsquigarrow k} a_{jk}(x_j - x_k), \quad x_k \in \mathbb{R},\ k = 1...N. \tag{3.21}$$

A natural adaptation of (3.21) on the circle is

$$\frac{d}{dt}\theta_k = \sum_{j \rightsquigarrow k} a_{jk}\sin(\theta_j - \theta_k), \quad k = 1...N. \tag{3.22}$$

This is in fact the celebrated Kuramoto model with equal natural frequencies.

Defining $z_k = e^{i\theta_k}$, (3.22) is equivalent to

$$\frac{d}{dt} z_k = \mathrm{Proj}_{z_k}\left(\sum_{j \rightsquigarrow k} a_{jk}(z_j - z_k)\right) \tag{3.23}$$

where $\mathrm{Proj}_{z_k}(r_k)$ denotes the orthogonal projection of $r_k \in \mathbb{C}$ onto the direction tangent to the unit circle at $z_k = e^{i\theta_k}$, that is $\mathrm{Proj}_{z_k}(r_k) = iz_k\langle iz_k, r_k\rangle$.

The geometric interpretation is that (3.23) defines a consensus update similar to (3.21) but constrained to the manifold where $||z_k|| = 1$.
On the circle and other nonlinear manifolds, there are graph-dependent consensus configurations: it has been shown that model (3.22) (almost) globally converges towards synchronization for tree graphs and the equally weighted complete graph. For other graphs, (3.22) may fail to converge to fixed positions (e.g. "cyclic pursuit" problem for a directed cycle graph) or may locally converge to a stable configuration that is different from synchronization (e.g. agents uniformly distributed around the circle for an undirected cycle graph).
However, stable consensus configurations different from synchronization are not exceptional: in fact, any configuration sufficiently "spread" on the circle is a stable consensus for a well-chosen directed graph.

Property 3.2 *Consider N agents distributed on the circle in a configuration $\{\theta_k\}$ such that for every k, there is at least one agent located in $(\theta_k, \theta_k + \pi/2)$ and one agent located in $(\theta_k - \pi/2, \theta_k)$; such a configuration requires $N \geq 5$. Then there exists a positively weighted, directed and root-connected interconnection graph making this configuration locally exponentially stable under (3.22).*

It is thus possible to identify how specific configurations can be made locally exponentially stable by choosing appropriate weights for the directed graph edges. For any of these choices, synchronization is also exponentially stable but thus only locally; in particular, agents initially located within a semicircle always converge towards synchronization.

As the circle embedded in the (complex) plane with its center at the origin satisfies Assumptions 2.1 and 3.1, the IAM is simply the central projection of C_e onto the circle. It corresponds to the entire circle if $C_e = 0$ and reduces to a single point in other situations.
The IAM uses the chordal length between points, while the Karcher mean would use arclength distance.

For the circle, the embedding of a point $\theta_k \in S^1$ in $\mathbb{C}$ is the vector $e^{i\theta_k}$. The continuous-time update is

$$\dot{\theta}_k = \mathrm{Im}\Big(\sum_j a_{jk}(e^{i(\theta_j-\theta_k)}-1)\Big) = \sum_j a_{jk}\sin(\theta_j-\theta_k) \tag{3.24}$$

while the discrete-time update is

$$\theta_k(t+1) = \arg\left(\frac{1}{\beta_k + d_k^{(i)}}\left(\sum_{j\rightsquigarrow k} a_{jk}e^{i\theta_j} + \beta_k e^{i\theta_k}\right)\right) \tag{3.25}$$

which can also be written in a more compact form as

$$\theta_k(t+1) = \arg\left(e^{\imath\theta_k} + \frac{1}{d_k}\sum_{j\rightsquigarrow k} e^{\imath\theta_j}\right) \tag{3.26}$$

Those expressions establish a clear connection between consensus algorithms on the circle and phase synchronization models. For the equally-weighted complete graph, the continuous-time update 3.24 is the model of Kuramoto with identical (zero) natural frequencies.
The discrete-time update 3.25 is Vicsek's phase update law [13] governing the headings of a set of particles in the plane.

Kuramoto Model

The collective behaviors of limit-cycle oscillators appear in many biological phenomena such as flashing of fireflies. A mathematical study on the collective behavior of limit-cycle oscillators was pioneered by Winfree and Kuramoto using simple dynamical systems for phase evolution.

The Kuramoto model consists of a population of N coupled phase oscillators, $\theta_k(t)$, having natural frequencies ω_k distributed with a given probability density $g(\omega)$, and whose dynamics is governed by

$$\dot{\theta}_k = \omega_k + \sum_{j=1}^{N} K_{jk}\sin(\theta_j-\theta_k), \quad k = 1,...,N. \tag{3.27}$$

When the coupling is sufficiently weak, the oscillators run incoherently whereas beyond a certain threshold collective synchronization emerges spontaneously.

Many different models for the coupling matrix K_{jk} have been considered such as nearest-neighbor coupling, hierarchical coupling, random long-range coupling, or even state dependent interactions. Among other phase models for synchronization phenomena, our main interest in this work lies on a simple mean-field[5] model.

[5] With the term "mean-field" we refer to a method to analyze physical systems with

Taking $K_{jk} = K/N > 0$ in Eq. (3.27), the model is then written as

$$\dot{\theta}_k = \omega_k + \frac{K}{N}\sum_{j=1}^{N} \sin(\theta_j - \theta_k), \quad t > 0, \ k = 1, ..., N. \tag{3.28}$$

where K measures the coupling strength (relative to heterogeneity).

The transformation that allows this model to be solved exactly (at least in the $N \to \infty$ limit) is as follows. We define the "order" parameters r and ψ as

$$re^{i\psi} = \frac{1}{N}\sum_{j=1}^{N} e^{i\theta_j},$$

where r represents the phase-coherence of the population of oscillators, and ψ indicates the average phase. Applying this transformation, the governing equation becomes

$$\dot{\theta}_k = \omega_k + Kr\sin(\psi - \theta_k). \tag{3.29}$$

A further transformation is usual for a rotating frame in which the statistical average of phases over all oscillators is zero (i.e. $\psi = 0$): the governing equation becomes

$$\dot{\theta}_k = \omega_k - Kr\sin(\theta_k). \tag{3.30}$$

A formal way to define the complete phase/frequency synchronization is given by the following definition:

Definition 3.1 *Let $\mathcal{P} = \{\theta_k\}_{k=1}^{N}$ be the system of oscillators whose dynamics is governed by the Kuramoto system (3.27).*
1. The system $\mathcal{P}$ has asymptotic complete phase synchronization if and only if the following condition holds.

$$\lim_{t\to\infty} |\theta_k(t) - \theta_j(t)| = 0, \quad \forall k \neq j,$$

2. The system $\mathcal{P}$ has asymptotic complete frequency synchronization if and only if the following condition holds.

$$\lim_{t\to\infty} |\omega_k(t) - \omega_j(t)| = 0, \quad \forall k \neq j,$$

where $\omega_k := \dot{\theta}_k$ is the instantaneous frequency of k-th oscillator.

For further details we refer the reader to [32, 14, 15, 33].

multiple bodies. The main idea is to replace all interactions to any one body with an average or effective interaction. This reduces any multi-body problem into an effective one-body problem.

3.2.3 Coordination in Vector Spaces

Coordination is closely related to consensus and synchronization: while *synchronization* refers to consensus among positions or configuration variables and invariance with respect to the reference frame (i.e. only relative positions matter), *coordination* refers to consensus among velocities (e.g. a consensus problem for the velocity model is $\dot{v}_j = u_j$) and implies N points moving like a single one (i.e. relative positions are constant).
For instance, on the circle, phase synchronization refers to consensus on S^1, while coordination would refer to a situation where all phase variables evolve at the same speed, a situation commonly referred to as phase locking.

Both synchronization and coordination can be considered as consensus problems, but in different spaces: the configuration space for the former, and the tangent bundle[6] for the latter.
When the configuration space is linear, there is a natural identification between the configuration space and the tangent space. Velocity vectors are treated as vectors of the configuration space by translating them to the origin. Coordination and synchronization are then equivalent in the sense that they both reduce to a linear consensus problem.
The situation is different when the configuration space is nonlinear. Velocity vectors of different agents then belong to different tangent spaces that can non-longer be identified to the configuration space. This raises the issue of comparing velocities and to ensure the analog invariance properties of a coordination control law. Lie groups offer a convenient setting for coordination on nonlinear spaces. For more details, readers can see [8, 17].

[6] In mathematics, the tangent bundle of a differentiable manifold $\mathcal{M}$ is the disjoint union of the tangent spaces of $\mathcal{M}$.

3.2.4 Convergence Analysis of Consensus Algorithms

In this section we investigate conditions and methods to analyze the convergence of consensus algorithms. We only give some concepts and invite readers to see [4] for more details.

Convergence Analysis for Time-invariant Communication Topologies

We now investigate conditions under which the information states of consensus algorithm (3.1) converge when the communication topology is time invariant and the gains a_{jk} are constant, that is the nonsymmetric Laplacian matrix $\mathcal{L}$ is constant. Because $\mathcal{L}$ has zero row sums, 0 is an eigenvalue of $\mathcal{L}$ with the associated eigenvector $\mathbf{1}$, the $n \times 1$ column vector of ones. If 0 is a simple eigenvalue of $\mathcal{L}$, then $x(t) \to \overline{x}\mathbf{1}$, where $\overline{x}$ is a scalar constant, which implies that $|x_j(t) - x_k(t)| \to 0$, as $t \to \infty$, for all $j, k = 1, ..., N$. Convergence analysis therefor focuses on conditions that ensure that zero is a simple eigenvalue of $\mathcal{L}$. Otherwise the kernel of $\mathcal{L}$ includes elements that are not in span$\{\mathbf{1}\}$, in which case consensus is not guaranteed. Zero is a simple eigenvalue of $\mathcal{L}$ if and only if the associated directed graph has a directed spanning tree. This result implies that (3.1) achieves consensus if and only if the directed communication topology has a directed spanning tree or the undirected communication topology is connected.

Analogous conditions hold for discrete-time consensus algorithm (3.2). Gershgorin's theorem implies that all eigenvalues of $\mathcal{M}$ are either in the open unit disk or at 1. If 1 is a simple eigenvalue of $\mathcal{M}$, then $|x_j(t) - x_k(t)| \to 0$, as $t \to \infty$. Perron-Frobenius' theorem implies that 1 is a simple eigenvalue of stochastic matrix A if directed graph $\Gamma(A)$ is strongly connected, or equivalently, if A is irreducible. Thus, (3.2) achieves consensus if and only if either the directed communication topology has a directed spanning tree or the undirected communication topology is connected.

Requiring a directed spanning tree is less stringent than requiring a strongly connected and balanced graph. However, the consensus equilibrium is a function only of the initial information states of those agents that have a directed path to all other agents.

Convergence Analysis for Dynamic Communication Topologies

Communication topologies are often dynamic. Therefore, in this section, we investigate conditions under which consensus algorithms converge under random switching of the communication topologies.

One approach to analyzing switching topologies is to use algebraic graph theory, which associates each graph topology with an algebraic structure of corresponding matrices. Because (3.1) is linear, its solution can be written as $x(t) = \Phi(t,0)x(0)$, where $\Phi(t,0)$ is the transition matrix corresponding to $-\mathcal{L}(t)$. $\Phi(t,0)$ is a stochastic matrix with positive diagonal entries for all $t \geq 0$. Consensus is achieved if $\lim_{t\to\infty} \Phi(t,0) \to \mathbf{1}\mu^T$, where μ is a column vector. It is typical to assume that the communication topology is piecewise constant over finite lengths of time, called *dwell times*, and that dwell times are bounded below by a positive constant. Thus, convergence analysis involves the study of infinite products of stochastic matrices and the results for SIA matrices (i.e. indecomposable and aperiodic stochastic matrices).

Nonlinear analysis con also be used to study consensus algorithms.

For continuous-time consensus algorithm (3.1), we consider the Lyapunov function $V(x) = \max\{x_1, ..., x_N\} - \min\{x_1, ..., x_N\}$. It can be shown that the equilibrium set span$\{\mathbf{1}\}$ is uniformly exponentially stable if there is an interval length $T > 0$ such that, for all t, the directed graph of $-\int_t^{t+T} \mathcal{L}(s)ds$ has a directed spanning tree.

For discrete-time consensus algorithm (3.2), a set-valued Lyapunov function V is defined as $V(x_1, ..., x_N) = (\text{conv}\{x_1, ..., x_N\})^N$, where $\text{conv}\{x_1, ..., x_N\}$ denotes the convex hull of $\{x_1, ..., x_N\}$, and $X^N := X \times \cdots \times X$. It can be shown that for all $t_2 \geq t_1$ $V(t_2) \subseteq V(t_1)$, and that $x(t)$ approached an element of the set span$\{\mathbf{1}\}$, which implies that consensus is reached.

Finally, information consensus is also studied from a stochastic point of view, which considers a random network, where the existence of an information channel between a pair of agents at each time is probabilistic and independent of other channels, resulting in a time-varying undirected communication topology. For example, adjacency matrix $A = [a_{jk}] \in \mathbb{R}^{N\times N}$ for an undirected random graph is defined as $a_{jj}(p) = 0,\ a_{jk}(p) = 1$ with probability p, $a_{jk}(p) = 0$ with probability $1-p$ for all $j \neq k$. In this case, consensus over undirected random network is addressed by notions from stochastic stability.

Convergence based on the Contraction of the Convex Hull

Moreau's convergence analysis of consensus algorithms is based on the contraction of the convex hull of agents' states. This analysis is not restricted to linear updates and it applies to the nonlinear consensus algorithms over arbitrary convex sets of the manifold $\mathcal{M}$.

In the case of the circle, for instance, the largest convex subset is a semicircle. As a consequence, uniform convergence to consensus under a uniform connectedness assumption holds provided that the states are initially contained in a semicircle.

Because none of the considered manifolds is globally convex, the linear convergence result never provides a global convergence analysis.

For continuous-time algorithm (3.18), whenever the communication graph

is fixed and undirected, all solutions converge to the set of critical points and all strict minima of the disagreement cost function $V(x)$ are stable equilibria. Synchronization corresponds to the global minimum of the disagreement cost function.
The descent property extends to discrete-time algorithms for sufficiently small step sizes (i.e. for β_k close to one) or for an asynchronous version of (3.19): in the latter case, the value of the different agents is updated one at the time; then the Lyapunov function can only decrease, in contrast to the synchronous situation where the entire vector of agents' states is updated at once.
The global properties of consensus algorithms defined on nonlinear spaces are graph dependent.

Solutions to recover (almost) Global Convergence on the Circle

The negative fact that a global analysis of the proposed consensus algorithms seems elusive on nonlinear spaces, even on the circle, is compensated for by the fact that the algorithms can be modified in such a way that convergence is guaranteed under a mere uniform connectedness assumption, such as in linear spaces. Three solutions have been investigated (on the circle) to recover global convergence.

Gossip Algorithm A possible solution is to introduce randomness in the link selection of the consensus algorithm, following the idea of a Gossip consensus algorithm. At each time instant, a given agent selects randomly one (or none) of its neighbors. The update is then taken as if this neighbor was the only one, disregarding the information from others.
An extreme version of this Gossip algorithm is when the update is chosen with no inertia ($\beta = 0$): each agent selects one neighbor randomly and replaces its current value by the neighbor's value with a certain probability. In this case, the consensus value is the initial condition of one of the agents.
The convergence property of this Gossip algorithm is very general and does not require any geometric structure on the underlying configuration space. The convergence property of the algorithm only relies on the probabilistic time evolution of N states switching among at most N different symbols.
If $\mathbb{G}(t)$ is uniformly connected, gossiping achieves global asymptotic synchronization with probability one.
Favoring the probability of convergence to synchronization through randomness comes with a price: convergence can be arbitrarily slow. Moreover, the probabilistic setting breaks the symmetry.
See [8, 10, 25] for more details.

Dynamic Consensus Another solution is to increase the amount of information exchanged by the agents. The non-convexity of S^1 can be circumvented if the agents are able to communicate auxiliary variables in addition to their positions on the circle. An interpretation is that the limited number of communication links for information flow is compensated by sending larger communication packets along existing links. Such strategies allow to recover the synchronization properties of vector spaces for almost all initial conditions. Their potential interest lies more in engineering applications.
The proposed dynamic consensus algorithm exploits the embedding of the algorithm in a linear space. Consensus among the embedding variables relies on convergence of linear consensus algorithms, leading to asymptotic consensus on the manifold as well. When the manifold is a Lie group such as the circle, the communication of auxiliary variables can be implemented in such a way that it respects the symmetry constraints of the algorithm, i.e. without the need of a common reference frame. [8]

Potential Shaping The last solution exposed in this work only concerns the case of a fixed undirected graph. Convergence is guaranteed to a local minimum of the disagreement cost function. It is possible to shape the potential in such a way that the only stable equilibrium corresponds to synchronization. The only graph information needed to construct this potential is an upper bound on the number of nodes. A descent algorithm for this shaped potential guarantees almost global convergence to synchronization if the (fixed undirected) graph is connected. [8]

3.3 Conclusion: Main Difference

An essential difference between linear consensus algorithms and their nonlinear extensions is the non-convex nature of symmetric spaces like the circle. This property is what makes the convergence analysis graph dependent when the state space is nonlinear.
From a design viewpoint, it is of interest to reformulate consensus algorithms on nonlinear spaces in such a way that they converge (almost) globally under the same assumptions as linear consensus algorithms.

Chapter 4

Applications

In this section, we briefly describe a few applications of consensus algorithm. Many of these applications concern either biological issues, e.g. animal behavior, or engineering problems, e.g. ocean sampling, phase synchronization, distributed control for large telescopes and multivehicle coordination (the reader is referred to [21, 22, 19, **?**, 20, 4, 16, 30]).
Other applications of consensus deal with optimal coding and neural networks. In this thesis, we only refer the reader's attention to [23, 24].

Autonomous Ocean Sampling Network

The ***Autonomous Ocean Sampling Network*** (AOSN) and ***Adaptive Sampling And Prediction*** (ASAP) projects aim to develop a sustainable, portable, adaptive ocean observing and prediction system for use in coastal environments. These projects employ, among other observation platforms, autonomous underwater vehicles (AUVs) that carry sensors to measure physical and biological signals in the ocean. Critical to this effort are reliable, efficient and adaptive control strategies to enable the mobile sensor platforms to collect data autonomously. Other details can be found in [21, 22].

Phase Synchronization of Oscillator Networks

Agreement on the circle appears for phase synchronization of oscillator networks.
For example, certain species of ***fireflies*** show a group behavior of synchronous flashing. Their synchronized and rhythmic flashing has received much attention among many researchers, and there has been a study of biological models for their entrainment of flashing.

Buck (1988):

> "In the region from India east to the Philippines and New Guinea, enormous aggregations of fireflies gather in trees and flash in near-perfect synchrony. While different species have slightly different methods for flashing in rhythm, the behavior that is modeled here is governed by the following rules:
>
> - Each firefly has an intrinsic flashing frequency, and when left alone it will flash at periodic intervals.
> - The flashes are timed by the progressive excitation of a chemical within each firefly; the excitation increases until it reaches a certain threshold, at which point a flash is emitted and the excitation is reset to zero.
> - If a firefly senses a certain amount of luminescence from its neighbors, it will reset its excitation to zero in order to flash simultaneously with those neighbors in the future; however, if the excitation is close enough to the flashing threshold, the flash has already been started and will proceed as planned even though the excitation is reset to zero."

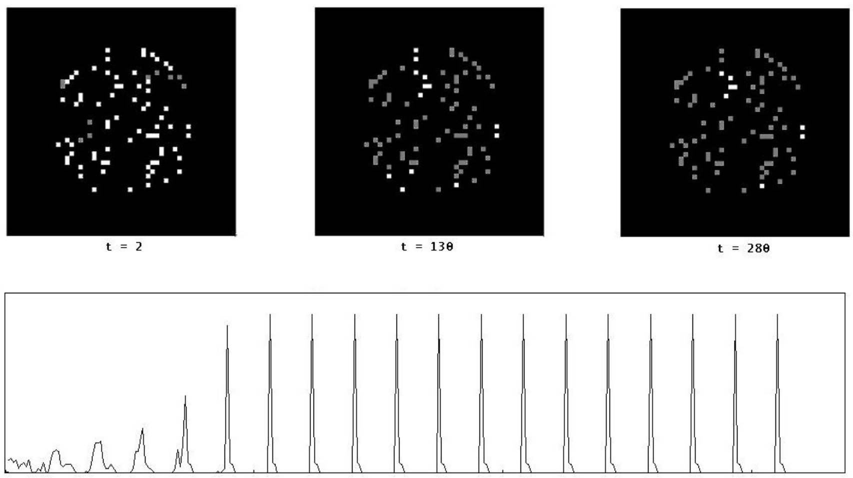

Figure 4.1: Screen shots and time series from simulation of firefly flashing behavior. In the top panels, gray squares indicate flashing fireflies, white squares dormant fireflies. Time series shows the number of fireflies "on" at each time step. (http://skyeome.net/wordpress/?p=56)

The reader can find more information for example in [19].

The European Extremely Large Telescope

The ***E-ELT primary mirror*** is composed of 984 hexagonal segments of 0.7m edge length. Each segment is supported by 3 unidimensional position actuators which move, in first approximation, neglecting small local curvature of the mirror perpendicularly to the mirror surface. This allows to separately control piston, tip and tilt (PTT) of each segment. Discontinuities in the mirrors reflecting surface are accurately monitored by 5604 so-called Edge Sensor pairs, which measure the relative vertical displacement at two positions on adjacent segment edges. The induced coupling of adjacent segments in the measurement values is a fundamental property giving rise to a distributed system. Consensus is regarded as distributed regulation over the relative state of agents (the mirrors).
Relative sensing is a key feature of distributed system theory and it is also the fundamental source of performance limitation. For further details see [20]

Multivehicle Coordination Problems

A first example is the so-called ***Rendezvous Problem***. This problem requires that a group of vehicles in a network meet at a time or a location determined through team negotiation. Consensus algorithms can be used to perform the negotiation in a way that is robust to environment disturbances.
A second example is the ***Formation Stabilization***, that requires that vehicles collectively maintain a prescribed geometric shape. In the decentralized version, each vehicle knows the desired formation shape, but the location of the formation needs to be negotiated among team members. The information state for this problem includes the center of the formation, which is negotiated by the team of vehicles employing consensus algorithms.
Another interesting problem is ***Flocking***. Using biologically observed motions of flocks of birds, it is possible to define three rules of flocking (that are collision avoidance, velocity matching, and flock centering) and apply them for coordination strategies in multivehicle robotic systems. More details can be found in [4, 16, 30].

Chapter 5

Conclusion

This thesis studied the problem of consensus, considering a set of N agents locally exchanging information about their state in order to asymptotically reach a common value of agreement: a global consensus.
This type of collective problem arises in a variety of disciplines, including physics, biology, and system and control theory: consider, for example, flocks of birds, schools of fish, sensor networks, etc.
Both coordination and synchronization have been considered as consensus issues, the former in the configuration space (consensus among positions) and the latter in the tangent bundle (consensus among velocities).

The consensus problem has been analyzed both in linear and in non-linear spaces. Its analysis required some mathematical preliminaries. First of all the stability conditions and the Lyapunov theory for non-linear systems. Then some (intuitive) notions for non-linear spaces, such as concepts of manifold, geodesic distance and Lie group. Finally some notions of graph theory, required in the framework of coordination between agents.

The update algorithms in the linear space (in the continuous-time case) is:

$$\frac{d}{dt}x_k(t) = \sum_{j=1}^{N} a_{jk}(t)(x_j(t) - x_k(t))$$

Under weak assumptions on the agents communication links (represented by mean of a graph), the consensus value is the arithmetic mean of the initial values. Under stronger constraints, the linear consensus algorithm is a gradient descent algorithm for the disagreement cost function.

A geometric interpretation of consensus allows us to consider an extension of the linear algorithm in the non-linear space, where "updating a point towards a new point" simply translates into "moving along the geodesic path connecting the two points". Thus, substituting Euclidean distance with

Riemannian (geodesic) distance, the (weighted) arithmetic average can be rewritten as

$$\text{mean}(x_1, ..., x_N) = \min_{z \in \mathbb{R}^n} \sum_{j=1}^{N} a_{ij} d^2_{\mathcal{M}}(z, x_j).$$

A more useful and convenient approach is on manifolds that are smoothly embedded in an ambient Euclidean space. For these manifolds, Riemannian distances can be substituted by Euclidean distance, leading to the induced arithmetic mean:

$$IAM(x_1, ..., x_N) = \operatorname*{argmin}_{z \in \mathcal{M}} \sum_{j=1}^{N} a_{ij} ||\widehat{x}_j - \widehat{z}||^2.$$

Finally, the attention has been focused on the simplest case of non-linear space: the circle. A natural adaptation of linear consensus on the circle is

$$\frac{d}{dt}\theta_k = \sum_{j \rightsquigarrow k} a_{jk} \sin(\theta_j - \theta_k), \quad k = 1...N.$$

that is, in fact, the celebrated Kuramoto model (briefly recalled in a specific section). Considering the circle embedded in the (complex) plane with its center at the origin, the IAM turns out to be the central projection of the centroid onto the circle. In addition, IAM uses the chordal length between points.

An essential difference between linear consensus algorithms and their nonlinear extensions is the non-convex nature of symmetric spaces like the circle. This property is what makes the convergence analysis graph dependent when the state space is nonlinear.

The last section of the thesis highlighted some examples of application of the consensus problem, both in biological (i.e. flashing fireflies) and engineering (AOSN, E-ELT, etc.) issues.

This thesis is based on the existing scientific literature, in particular Sepulchre's papers. Many other works, (e.g. Moreau, Jadbabaie, Sarlette, Ren, Bonilla, Georgiou's papers or books) has been used to deeply investigate fundamental topics related to consensus and exposed in the thesis, such as stability, Kuramoto model, coordination, synchronization, etc. Others have been cited as references for the readers.

Finally, to reach a more complete knowledge of consensus problem, one should consider also, for example, optimization strategies, more complex algorithms, analysis on Lie groups and on more complex manifolds, different techniques for convergence analysis.

Bibliography

[1] Jean-Jacques E.Slotine, Weiping Li, *Applied Nonlinear Control*, Prentice-Hall, Inc., Upper Saddle River, New Jersey, 1991

[2] Harry L. Trentelman, Anton A. Stoorvogel and Malo Hautus, *Control Theory for Linear Systems*, Springer-Verlag London Limited, Great Britain, 2001

[3] Mauro Bisiacco, Simonetta Braghetto, *Teoria dei sistemi dinamici*, Progetto Leonardo, Bologna, 2010

[4] Wei Ren, Randal W. Beard, *Distributed Consensus in Multi-vehicle Cooperative Control: Theory and Applications*, Springer-Verlag London Limited, Great Britain, 2008

[5] S. Amari, *Differential-Geometrical Methods in Statistics, Lecture notes in Statistics*, Springer-Verlag, Berlin, Germany, 1985

[6] W. Rudin, *Real and Complex Analysis*, 3rd ed. McGraw-Hill, New York, 1987

[7] J. Tsitsiklis *Problems in decentralized decision making and computation* PhD thesis, Department of EECs, MIT, 1984

[8] R. Sepulchre, *Consensus on Nonlinear Spaces*, Annual Reviews in Control vol. **35**, pp. 56-64, 2011

[9] Alain Sarlette, Rodolphe Sepulchre, *Consensus on homogeneous manifolds*, Joint 48th IEEE Conference on Decision and Control vol. [**1-4244-3871-3**], pp. 6438 -6443, 2009

[10] Alain Sarlette, S. Emre Tuna, Vincent D. Blondel, Rodolphe Sepulchre, *Global synchronization on the circle*, IFAC Proceedings Volumes (IFAC-PapersOnline) vol. **17 iss:1**, pp. [1474-6670], 2008

[11] Luc Moreau, *Stability of Multiagent Systems With Time-Dependent Communication Links*, IEEE Transactions on Automatic Control vol. **50, no.2**, pp. 169-182, 2005

[12] Alain Sarlette, Rodolphe Sepulchre and Naomi Ehrich Leonard, *Autonomous rigid body attitude synchronization*, Automatica vol. **45**, pp. 572-577, 2009

[13] Alireza Tahbaz-Salehi and Ali Jadbabaie, *On recurrence of graph connectivity in Vicsek's model of motion coordination for mobile autonomous agents*, Proceedings of the American Control Conference vol. [**1-4244-0988-8**], pp. 699 -704 , 2009

[14] Juan A. Acebrón, L. L. Bonilla, Conrad J. Pérez Vicente and Félix Ritort, Renato Spigler, *The Kuramoto model: A simple paradigm for synchronization phenomena*, Reviews of modern physics vol. [**0034-6861**] **Acebrón vol:77 iss:1**, pp. 137 -185 , 2005

[15] Seung-Yeal Ha, Taeyoung Ha, Jong-Ho Kim, *On the complete synchronization of the Kuramoto phase model*, Physica D vol. **239**, pp. 1692-1700, 2010

[16] Wei Ren, Randal W. Beard, Ella M. Atkins, *A survey of Consensus Problems in Multi-agent Coordination*, American Control Conference vol. [**0-7803-9098-9**], pp. 1859-1864, 2005

[17] Alain Sarlette, Silvère Bonnabel and Rudolphe Sepulchre, *Coordinated Motion Design on Lie Groups*, IEEE Transactions on Automatic Control vol. **55, no.5**, pp. 1047-1058, 2010

[18] Xianhua Jiang, Lipeng Ning, and Tryphon T. Georgiou, *Distances and Riemannian metrics for multivariate spectral densities*, IEEE Transactions on Signal Processing vol. **55, no. 8**, pp. 3995-4003, 2007

[19] DaeEun Kim, *A spiking neuron model for synchronous flashing of fireflies*, BioSystems vol. **76**, pp. 7-20, 2004

[20] Alain Sarlette, Christian Bastin, Martin Dimmler, Babak Sedghi, Toomas Erm, Bertrand Bauvir and Rodolphe Sepulchre *Integral control from distributed sensing: an Extremely Large Telescope case study*

[21] Thomas B. Curtin, James G. Bellingham, Josko Catipovic and Doug Webb, *Autonomous Oceanographic Sampling Networks*, Oceanography vol. **6 No 3**, pp. 86-94, 1993

[22] P.Bhatta, E.Fiorelli, F.Lekien, N.E.Leonard, D.A.Paley, F.Zhang, and R.Bachmayer, R.E.Davis, D.M.Fratantoni, R.Sepulchre *Coordination of an underwater glider fleet for adaptive ocean sampling*

[23] Mehmet Ercan Yildiz and Anna Scaglione, *Coding With Side Information for Rate-Constrained Consensus*, IEEE Transactions on Signal Processing vol. **56, no.8**, pp. 3753-3764, 2008

[24] Lars Kai Hansen and Peter Salamon, *Neural Network Ensembles*, IEEE Transactions on Pattern Analysis and Machine Intelligence vol. **12, no.10**, pp. 993-1001, 1990

[25] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, *Randomized gossip algorithms*, IEEE Trans. Inf. Theory vol. **52**, pp. 2508-2530, 2006

[26] F. Garin, L. Schenato, *A survey on distributed estimation and control applications using linear consensus algorithms*, Networked Control Systems vol. **vol. 406**, pp. 75-107, 2011

[27] Jadbabaie, J. Lin and A. S. Morse, *Coordination of groups of mobile autonomous agents using nearest neighbor rules* IEEE Trans. Aut. Control 48 988-1001 2003

[28] A. Okubo, *Dynamical aspects of animal grouping: swarms, schools, flocks, and herds* Advances in Biophysics 22 1-94 1986

[29] R. Olfati-Saber and R. M. Murray, *Consensus problems in networks of agents with switching topology and time-delays* IEEE Trans. Aut. Control 49 1520-1533 2004

[30] R. Olfati-Saber, *Flocking for Multi-Agent Dynamic Systems: Algorithms and Theory* IEEE Trans. Aut. Control 51 401-420 2006

[31] R. Olfati-Saber, J. A. Fax, and R. M. Murray, *Consensus and Cooperation in Networked Multi-Agent Systems* Proceedings of the IEEE 95 215-233 2007

[32] S. H. Strogatz, *From Kuramoto to Crawford: exploring the onset of synchronization in populations of coupled oscillators* Physics D: Nonlinear Phenomena 143 1-20 2000

[33] M. Marodi, F. d'Ovidio, and T. Vicsek, *Synchronization of oscillators with long range interaction: Phase transition and anomalous finite size effects* Physical Review E 66 011109 2002