Emergent behaviors of Cucker-Smale flocks on Riemannian manifolds

Seung-Yeal Ha, Doheon Kim, and Franz Wilhelm Schlöder

***Abstract*—In this paper, we present a new Cucker-Smale (CS) model on smooth Riemannian manifolds using the concepts of covariant derivative and parallel transport, and we also study its emergent dynamics under an a priori assumption on the energy functional. For Euclidean space, our proposed model coincides with the original Cucker-Smale model. As concrete examples, we consider three Riemannian manifolds: the unit 2-sphere, the unit circle and the Poincaré half-plane, and provide explicit reductions from the proposed general model to aforementioned manifolds via explicit formulas for the covariant derivative and parallel transport.**

***Index Terms*—Cucker-Smale model, flocking, geometric flocking, multi-agent systems, Riemannian manifold.**

I. Introduction

Emergent behaviors of many-body(particle) systems are ubiquitous in our natural and man-made complex systems, e.g., synchronous flashing of fireflies [5], [15], flocking of fish [13], neuronal synchronization in a brain [11], arrays of Josephson junctions [26], application to unmanned aerial vehicles (UAVs) such as aircrafts or satellites [16], [30], [33] etc. We also refer to [1] for a brief survey of the emergent dynamics of classical many-body system. Among them, our main interest lies on the collective behavior of CS particles on a complete Riemannian manifold $(M, \langle\cdot,\cdot\rangle_M)$. To motivate our discussion, we first begin with the CS model [10] on Euclidean space. Let $\boldsymbol{x}_i$ and $\boldsymbol{v}_i$ be the position and velocity of the i-th CS particle, respectively. Then, the CS model on $\mathbb{R}^d$ reads as follows.

$$\begin{aligned} &\dot{\boldsymbol{x}}_i = \boldsymbol{v}_i, \quad t>0, \quad i = 1,\cdots,N, \\ &\dot{\boldsymbol{v}}_i = \frac{\kappa}{N}\sum_{j=1}^N \psi(\|\boldsymbol{x}_i-\boldsymbol{x}_j\|)(\boldsymbol{v}_j-\boldsymbol{v}_i), \end{aligned} \tag{I.1}$$

where $\psi = \psi(r)$ is a Lipschitz continuous communication weight function which quantifies the degree of communications between particles, and it is assumed to be a function of relative distances. Since the vector field generated by the R.H.S. of (I.1) is Lipschitz continuous, the global well-posedness of (I.1) directly follows from the standard Cauchy-Lipschitz theory. Hence, an interesting issue is the identification of possible asymptotic patterns arising from initial data along the CS flow and their basin of attraction. From this point of view, there has been extensive research from various angles, e.g., stochastic (CS) model [6], [12], rigorous mean-field limit and its kinetic model [7], [24], generalization and thermo-mechanical extension [23], [29], etc. (see [2], [9] for review articles). As far as the authors know, all aforementioned works on the CS model deal with the Euclidean space setting so that there are no curvature effects arising from constraints on the positions of particles. In this paper, we are mainly interested in the *flocking realization problem* [8] on a Riemannian manifold: for a given Riemannian manifold $(M, \langle\cdot,\cdot\rangle_M)$ and a collection of N particles,

- (Q1): Design a continuous dynamical system on $(M, \langle\cdot,\cdot\rangle_M)$ exhibiting flocking behavior.
- (Q2): Once question (Q1) is realized, establish sufficient frameworks for emergent dynamics in terms of initial data and system parameters.

Once the flocking realization problem is resolved, we will see how the curvature and topology of the ambient Riemannian manifold $(M, \langle\cdot,\cdot\rangle_M)$ play key roles in the flocking dynamics of the CS particles. In this direction, the literature is very sparse. For example, a first-order aggregation model on the unit sphere, namely the swarm sphere model and its emergent dynamics have been proposed and studied in a series of papers [8], [21], [25], [28], [31]. Recently, the first author and his collaborator provided a second-order extension model for the swarm sphere model in [19]. Consensus and anti-consensus algorithms on a compact homogeneous manifold were studied in [35], and models for opinion and aggregation dynamics on Riemannian manifolds were also proposed in [3], [17], [18]. In [36], a discrete-time consensus algorithm on a general Riemannian manifold is proposed. In [27], synchronization on Riemannian manifolds, including extensions of the Kuramoto model and its high-dimensional analogues, is studied. However, to the authors' knowledge, there is no previous literature modeling velocity alignment dynamics of particles on a Riemannian manifold which is a key feature of flocking behavior. Our goal in this paper is to provide answers to the flocking realization problem. More precisely, our main results can be summarized as follows.

Let $(M, \langle\cdot,\cdot\rangle_M)$ be a complete smooth Riemannian manifold. Then, for $\boldsymbol{x}_i, \boldsymbol{x}_j \in M$, let P_{ij} be the parallel transport along the length-minimizing geodesic from the point $\boldsymbol{x}_j$ to $\boldsymbol{x}_i$, and we set $\frac{D}{dt}\boldsymbol{v}_i := \nabla_{\boldsymbol{v}_i}\boldsymbol{v}_i$, where ∇ denotes covariant derivative. Under this setting, our first result is the proposal of the extension of the CS model (I.1) on $(M, \langle\cdot,\cdot\rangle_M)$ (see

S.-Y. Ha is with the Department of Mathematical Sciences and Research Institute of Mathematics, Seoul National University, Seoul 08826, and the School of Mathematics, Korea Institute for Advanced Study, Seoul 02455, Republic of Korea. e-mail: syha@snu.ac.kr

D. Kim is with the School of Mathematics, Korea Institute for Advanced Study, Seoul 02455, Republic of Korea. e-mail: doheonkim@kias.re.kr

F. W. Schlöder is with the Department of Mathematics and Applications, University of Milano-Bicocca, Via Cozzi 55, 20125 Milan, Italy. e-mail: franz.schloeder@unimib.it

Section II-B):

$$\begin{cases} \dfrac{d\boldsymbol{x}_i}{dt} = \boldsymbol{v}_i, \quad t>0, \quad 1 \le i \le N, \\ \dfrac{D\boldsymbol{v}_i}{dt} = \dfrac{\kappa}{N}\displaystyle\sum_{k=1}^{N}\psi_{ik}(t)\big(P_{ik}\boldsymbol{v}_k - \boldsymbol{v}_i\big), \\ \boldsymbol{x}_i(0) = \boldsymbol{x}_i^0, \quad \boldsymbol{v}_i(0) = \boldsymbol{v}_i^0, \end{cases} \tag{I.2}$$

where $\psi_{ik}(t) = \psi(\mathrm{dist}(\boldsymbol{x}_i(t), \boldsymbol{x}_k(t)))$, with $\mathrm{dist}(\cdot,\cdot)$ denoting the geodesic distance. A common feature of (I.2) and one of two models in [3] is the use of geodesic distance. However unlike [3], we use the parallel transport to describe the alignment dynamics of particle's velocities. We elaborate on well-definedness issue for (I.2) in Section II.

Note that the tangent space of $\mathbb{R}^d$ is $\mathbb{R}^d$ itself, and $\frac{D}{dt}$ and parallel transport are the usual time derivative and identity map. Hence system (I.2) is the same as the CS model (I.1). Unlike the CS model (I.1), system (I.2) is not translation invariant in general. Hence the total momentum will not be conserved. This can be understood as a nontrivial curvature effect of the manifold. However, we can still derive a dissipative estimate for (I.2). More precisely, we define the total energy functional $\mathcal{E}$:

$$\mathcal{E} := \sum_{i=1}^{N}\|\boldsymbol{v}_i\|_{\boldsymbol{x}_i}^2, \quad \boldsymbol{v}_i \in T_{\boldsymbol{x}_i}M,$$

where $\|\cdot\|_{\boldsymbol{x}_i}$ is the norm in $T_{\boldsymbol{x}_i}M$ induced by the Riemannian metric $\langle\cdot,\cdot\rangle_{\boldsymbol{x}_i}$. Then, the isometry property of parallel transport enables us to derive a dissipative energy estimate (see Lemma II.2):

$$\frac{d\mathcal{E}}{dt} = -\frac{\kappa}{N}\sum_{i,k}\psi(\mathrm{dist}(\boldsymbol{x}_i, \boldsymbol{x}_k))\|P_{ik}\boldsymbol{v}_k - \boldsymbol{v}_i\|_{\boldsymbol{x}_i}^2 \le 0. \tag{I.3}$$

Finally, thanks to (I.3) and a priori uniform bound for $\mathcal{E}''$, one can derive a velocity alignment estimate (see Proposition II.1):

$$\lim_{t\to\infty}\|P_{ik}\boldsymbol{v}_k(t) - \boldsymbol{v}_i(t)\|_{\boldsymbol{x}_i}^2 = 0, \quad \forall\ i,k = 1,\cdots,N.$$

Second, we consider three explicit Riemannian manifolds: the unit 2-sphere in $\mathbb{R}^3$, the unit circle and the Poincaré half-plane. For these explicit manifolds, we can explicitly compute $\frac{D}{dt}$ and parallel transport in terms of state variables $(\boldsymbol{x}_i, \boldsymbol{v}_i)$ of an ambient space $\mathbb{R}^d$. Thus, the abstract CS model (I.2) can be rewritten as an explicit dynamical system. For example, (I.2) reduces to the following explicit form for the unit sphere in $\mathbb{R}^3$:

$$\begin{cases} \dot{\boldsymbol{x}}_i = \boldsymbol{v}_i, \quad t>0, \quad i=1,\cdots,N, \\ \dot{\boldsymbol{v}}_i = -\|\boldsymbol{v}_i\|^2\boldsymbol{x}_i + \dfrac{\kappa}{N}\displaystyle\sum_{k=1}^{N}\psi_{ik}\Big(\langle\boldsymbol{x}_k, \boldsymbol{x}_i\rangle\boldsymbol{v}_k \\ \quad + \dfrac{\langle\boldsymbol{x}_k\times\boldsymbol{x}_i, \boldsymbol{v}_k\rangle}{1+\langle\boldsymbol{x}_k, \boldsymbol{x}_i\rangle}(\boldsymbol{x}_k\times\boldsymbol{x}_i) - \langle\boldsymbol{v}_k, \boldsymbol{x}_i\rangle\boldsymbol{x}_k - \boldsymbol{v}_i\Big). \end{cases} \tag{I.4}$$

Note that for two particles with geodesic distance π on the unit 2-sphere, there is an infinite number of great circles passing through them, hence there will be some ambiguity in the velocity coupling $P_{ik}\boldsymbol{v}_k - \boldsymbol{v}_i$. However, as long as there is no podal-antipodal relation in the ensemble, the shortest geodesic between two particles' positions is uniquely defined as the short segment of the unique great circle connecting them, and parallel transport is nothing but a rotation along the great circle. These simple geometric observations enable us to provide a quantitative analysis for the reduced model (I.4) (see Theorem III.1). On the other hand, for the two-particle system on $\mathbb{S}^2$, we can even remove a priori assumption on the nonexistence of podal-antipodal relations, and we can provide a sufficient framework only in terms of the initial configuration and the coupling strength (see Theorem III.2). After a further reduction to the unit circle, our reduced CS model becomes a second-order model which can be related to the Kuramoto type model after integration. Finally, we perform a similar analysis for the Poincaré half-plane as in the unit sphere case, and provide emergent dynamics only in terms of initial data and communication weight function.

The rest of this paper is organized as follows. In Section II, we review the CS model on Euclidean space and then present a new CS model on a Riemannian manifold. In Section III, we provide a reduction for the CS model on a 2-sphere $\partial B_r(0)$ in $\mathbb{R}^3$ from the general CS model by explicitly calculating the parallel transport. In Section IV, we provide a reduction for the CS model on the unit circle $\mathbb{S}^1$ and discuss the emergent dynamics. In Section V, we again provide a reduction of the CS model to the Poincaré half plane $\mathbb{H}$ and study the emergent dynamics of the derived model in terms of initial data. Finally, Section VI is devoted to a brief summary of our main results and discussions on some open problems.

II. Preliminaries and the general model

In this section, we briefly review the classical CS model [10] on the Euclidean space, and discuss the generalized CS model (I.2) on a smooth Riemannian manifold $(M, \langle\cdot,\cdot\rangle_M)$.

A. *The CS model on Euclidean space*

Let $\boldsymbol{x}_i$ and $\boldsymbol{v}_i$ be the position and velocity of the i-th particle in $\mathbb{R}^d$, respectively. Then, the CS model reads as follows.

$$\begin{aligned} \dot{\boldsymbol{x}}_i &= \boldsymbol{v}_i, \quad t>0, \quad i=1,2,\cdots,N, \\ \dot{\boldsymbol{v}}_i &= \frac{\kappa}{N}\sum_{j=1}^{N}\psi(\|\boldsymbol{x}_j - \boldsymbol{x}_i\|)(\boldsymbol{v}_j - \boldsymbol{v}_i), \end{aligned} \tag{II.1}$$

where κ and $\|\cdot\|$ are a nonnegative coupling strength and the Euclidean ℓ_2-norm in $\mathbb{R}^d$. The function ψ represents the communication weight measuring the degree of communications between particles and it is assumed to satisfy the following relations:

$$\begin{aligned} &\psi > 0, \quad \|\psi\|_{L^\infty} \le 1, \quad [\psi]_{Lip} < \infty, \\ &(\psi(r_1) - \psi(r_2))(r_1 - r_2) \le 0, \quad r_1, r_2 > 0. \end{aligned} \tag{II.2}$$

Note that the R.H.S. of (II.1) is Lipschitz continuous in $(\boldsymbol{x}, \boldsymbol{v})$ and bounded in any compact set in phase space. Hence, the standard Cauchy-Lipschitz theory guarantees a global well-posedness in any finite-time interval. Moreover, the R.H.S. of (II.1) is invariant under Galilean transformation $(\boldsymbol{x}, \boldsymbol{v}) \to (\boldsymbol{x} + \boldsymbol{c}t, \boldsymbol{v} + \boldsymbol{c})$ so that the total momentum is a constant of motion. Next, we recall a definition of global flocking (mono-cluster flocking) as follows.

Definition II.1. [10], [22], [24] *Let $\mathcal{G} := \{(\boldsymbol{x}_i, \boldsymbol{v}_i)\}_{i=1}^N$ be a global smooth solution to* (II.1) - (II.2). *Then $\mathcal{G}$ exhibits asymptotic global flocking if and only if the following two relations hold.*

1) *(Velocity alignment): The relative velocities approach to zero asymptotically.*
$$\lim_{t\to\infty} \max_{i,j} \|\boldsymbol{v}_i(t) - \boldsymbol{v}_j(t)\| = 0.$$
2) *(Spatial coherence): The relative positions are uniformly bounded:*
$$\sup_{0\le t<\infty} \max_{i,j} \|\boldsymbol{x}_i(t) - \boldsymbol{x}_j(t)\| < \infty.$$

Next, we review the propagation of velocity moments associated with (II.1).

Lemma II.1. [10], [22], [24] *Let $\{(\boldsymbol{x}_i, \boldsymbol{v}_i)\}_{i=1}^N$ be a global smooth solution to* (II.1)-(II.2) *with the initial data $\{(\boldsymbol{x}_i^0, \boldsymbol{v}_i^0)\}_{i=1}^N$. Then, the total momentum and energy satisfy the following relations: for $t > 0$,*

(i) $\displaystyle \frac{d}{dt}\sum_{i=1}^N \boldsymbol{v}_i = 0,$

(ii) $\displaystyle \frac{d}{dt}\sum_{i=1}^N \|\boldsymbol{v}_i\|^2 = -\frac{\kappa}{N}\sum_{i,j=1}^N \psi(\|\boldsymbol{x}_j - \boldsymbol{x}_i\|)\|\boldsymbol{v}_j - \boldsymbol{v}_i\|^2 \le 0.$

Proof. The desired estimates are based on the skew-symmetry of $\psi(\|\boldsymbol{x}_j - \boldsymbol{x}_i\|)(\boldsymbol{v}_j - \boldsymbol{v}_i)$ under the exchange symmetry $i \Longleftrightarrow j$. We leave its details. □

In the sequel, we briefly review a Lyapunov functional approach associated with flocking estimate for (II.1) - (II.2). Due to the conservation of momentum in Lemma II.1, we may assume zero sum conditions:
$$\sum_{i=1}^N \boldsymbol{x}_i(t) = 0, \quad \sum_{i=1}^N \boldsymbol{v}_i(t) = 0, \quad t \ge 0.$$
Now, we set
$$\|\boldsymbol{x}\|_\infty := \max_{1\le i\le N}\|\boldsymbol{x}_i\|, \quad \|\boldsymbol{v}\|_\infty := \max_{1\le i\le N}\|\boldsymbol{v}_i\|.$$
Then, $\|\boldsymbol{x}\|_\infty$ and $\|\boldsymbol{v}\|_\infty$ are Lipschitz continuous, and satisfy
$$\begin{cases} \left|\dfrac{d}{dt}\|\boldsymbol{x}\|_\infty\right| \le \|\boldsymbol{v}\|_\infty, \quad \text{a.e. } t > 0, \\ \dfrac{d}{dt}\|\boldsymbol{v}\|_\infty \le -\kappa\psi(\sqrt{2}\|\boldsymbol{x}\|_\infty)\|\boldsymbol{v}\|_\infty. \end{cases} \tag{II.3}$$

Note that once we have a uniform bound for $\|\boldsymbol{x}\|_\infty$, the second inequality in (II.3) yields exponential decay of $\|\boldsymbol{v}\|_\infty$. Thus, we introduce two Lyapunov-type functionals denoted by $\mathcal{L}_\pm(t) \equiv \mathcal{L}_\pm(\boldsymbol{x}(t), \boldsymbol{v}(t))$: for $t \ge 0$,
$$\begin{aligned} \mathcal{L}_+(t) &:= \|\boldsymbol{v}(t)\|_\infty + \frac{\kappa}{\sqrt{2}}\int_0^{\sqrt{2}\|\boldsymbol{x}(t)\|_\infty} \psi(s)ds, \\ \mathcal{L}_-(t) &:= \|\boldsymbol{v}(t)\|_\infty - \frac{\kappa}{\sqrt{2}}\int_0^{\sqrt{2}\|\boldsymbol{x}(t)\|_\infty} \psi(s)ds. \end{aligned}$$

Then, it is easy to see the non-increasing property of $\mathcal{L}_+$ and $\mathcal{L}_-$ using (II.3) (see [22]):
$$\mathcal{L}_+(t) \le \mathcal{L}_+(0) \quad \text{and} \quad \mathcal{L}_-(t) \le \mathcal{L}_-(0), \quad t \ge 0,$$
This leads to the stability estimate of $\mathcal{L}_\pm(t)$:
$$\|\boldsymbol{v}(t)\|_\infty + \frac{\kappa}{\sqrt{2}}\left|\int_{\sqrt{2}\|\boldsymbol{x}_0\|_\infty}^{\sqrt{2}\|\boldsymbol{x}(t)\|_\infty} \psi(s)ds\right| \le \|\boldsymbol{v}^0\|_\infty, \quad t \ge 0.$$
The following theorem is the most relevant result on the flocking estimate.

Theorem II.1. [22] *Let $(\boldsymbol{x}, \boldsymbol{v})$ be a global smooth solution to* (II.1) *with the initial data $(\boldsymbol{x}^0, \boldsymbol{v}^0)$ satisfying the following conditions:*
$$\|\boldsymbol{x}^0\|_\infty > 0, \quad \|\boldsymbol{v}^0\|_\infty < \frac{\kappa}{\sqrt{2}}\int_{\sqrt{2}\|\boldsymbol{x}^0\|_\infty}^\infty \psi(r)dr.$$
Then, there exists a positive number x_M such that
$$\sup_{t\ge0}\|\boldsymbol{x}(t)\| \le x_M, \quad \|\boldsymbol{v}(t)\| \le \|\boldsymbol{v}^0\|e^{-\psi(\sqrt{2}x_M)t}, \quad t \ge 0.$$

Remark II.1. *Related flocking estimates have also been studied in [10], [24]. We refer to [9] for a brief survey of the mathematical results on the CS model* (II.1) - (II.2).

B. The CS model on a Riemannian manifold

In this subsection, we briefly discuss a new CS model (I.2) on a smooth complete connected Riemannian manifold $(M, \langle\cdot,\cdot\rangle_M)$. Let $R \in (0, \infty]$ be the radius of injectivity. This means

> " **(P)**: For any given two points on M, if the geodesic distance between them is less than R, then there is only one distance minimizing geodesic connecting those two points."

Note that for $\mathbb{S}^2$ and Poincaré's half plane $\mathbb{H}$, $R = \pi$ and ∞ respectively. Before we introduce our CS model on $(M, \langle\cdot,\cdot\rangle_M)$, we present a minimal preparation. Let $\mathrm{dist}(\boldsymbol{x}_i, \boldsymbol{x}_k)$ denote the geodesic distance between $\boldsymbol{x}_i$ and $\boldsymbol{x}_k$. If $\mathrm{dist}(\boldsymbol{x}_i(t), \boldsymbol{x}_k(t)) < R$, let $P_{ik}\boldsymbol{v}_k$ represent the parallel transport of the tangent vector $\boldsymbol{v}_k$ at the point $\boldsymbol{x}_k$, along the unique geodesic connecting $\boldsymbol{x}_k$ and $\boldsymbol{x}_i$. Moreover, $\psi : [0, R) \to \mathbb{R}$ is a non-increasing nonnegative $\mathcal{C}^1$ function measuring the degree of communications:
$$\psi(0) = \psi_M \quad \text{and} \quad \psi_m := \inf_{0\le r<R}\psi(r) > 0.$$
We set
$$\psi_{ik}(t) := \psi(\mathrm{dist}(\boldsymbol{x}_i(t), \boldsymbol{x}_k(t))), \quad i, k \in \{1, \cdots, N\}.$$

Consider the Cauchy problem to the CS model (I.2) on $(M, \langle\cdot,\cdot\rangle_M)$:
$$\begin{cases} \dfrac{d\boldsymbol{x}_i}{dt} = \boldsymbol{v}_i, \quad t > 0, \quad 1 \le i \le N, \\ \dfrac{D\boldsymbol{v}_i}{dt} = \dfrac{\kappa}{N}\displaystyle\sum_{k=1}^N \psi_{ik}(t)\big(P_{ik}\boldsymbol{v}_k - \boldsymbol{v}_i\big), \\ \boldsymbol{x}_i(0) = \boldsymbol{x}_i^0, \quad \boldsymbol{v}_i(0) = \boldsymbol{v}_i^0, \end{cases} \tag{II.4}$$

where we assume that for each $\boldsymbol{x}_i^0$ and $\boldsymbol{x}_k^0$, there exists a unique geodesic connecting them with its length less than R.

Next, we give a definition of *velocity alignment* for the model (II.4).

Definition II.2. *Let $\{(\boldsymbol{x}_i, \boldsymbol{v}_i)\}_{i=1}^N$ be a global smooth solution to* (II.4). *Then $\mathcal{G}$ exhibits asymptotic velocity alignment, if $P_{ik}\boldsymbol{v}_k(t) - \boldsymbol{v}_i(t)$ tends to zero asymptotically:*

$$\lim_{t\to\infty} \max_{1\le i,k\le N} \|P_{ik}\boldsymbol{v}_k(t) - \boldsymbol{v}_i(t)\|_{\boldsymbol{x}_i} = 0.$$

In the sequel, we provide a criterion for determining whether asymptotic velocity flocking will emerge or not. For this, we define energy functional $\mathcal{E}$:

$$\mathcal{E} := \sum_{i=1}^N \|\boldsymbol{v}_i\|_{\boldsymbol{x}_i}^2, \quad \boldsymbol{v}_i \in T_{\boldsymbol{x}_i}M.$$

Lemma II.2. *Let $\{(\boldsymbol{x}_i, \boldsymbol{v}_i)\}_{i=1}^N$ be a global smooth solution to* (II.4). *Then, the energy functional $\mathcal{E}$ is monotone decreasing:*

$$\frac{d\mathcal{E}}{dt} = -\frac{\kappa}{N}\sum_{i,k=1}^N \psi(\mathrm{dist}(\boldsymbol{x}_i, \boldsymbol{x}_k))\|P_{ik}\boldsymbol{v}_k - \boldsymbol{v}_i\|_{\boldsymbol{x}_i}^2 \le 0, \quad t > 0. \tag{II.5}$$

Proof. We take inner product of $\boldsymbol{v}_i$ with the second equation in (II.4) and sum the resulting relation over i to obtain

$$\begin{aligned}
\frac{d\mathcal{E}}{dt} &= 2\sum_{i=1}^N \left\langle \boldsymbol{v}_i, \frac{D\boldsymbol{v}_i}{dt}\right\rangle_{\boldsymbol{x}_i} \\
&= 2\sum_{i=1}^N \left\langle \boldsymbol{v}_i, \frac{\kappa}{N}\sum_{k=1}^N \psi_{ik}(P_{ik}\boldsymbol{v}_k - \boldsymbol{v}_i)\right\rangle_{\boldsymbol{x}_i} \\
&= \frac{\kappa}{N}\sum_{i,k=1}^N \psi_{ik}\Big(\langle \boldsymbol{v}_i, P_{ik}\boldsymbol{v}_k - \boldsymbol{v}_i\rangle_{\boldsymbol{x}_i} + \langle v_k, P_{ki}\boldsymbol{v}_i - \boldsymbol{v}_k\rangle_{\boldsymbol{x}_k}\Big) \\
&= \frac{\kappa}{N}\sum_{i,k=1}^N \psi_{ik}\Big(\langle \boldsymbol{v}_i, P_{ik}\boldsymbol{v}_k - \boldsymbol{v}_i\rangle_{\boldsymbol{x}_i} + \langle P_{ik}\boldsymbol{v}_k, \boldsymbol{v}_i - P_{ik}\boldsymbol{v}_k\rangle_{\boldsymbol{x}_i}\Big) \\
&= -\frac{\kappa}{N}\sum_{i,k=1}^N \psi_{ik}\|P_{ik}\boldsymbol{v}_k - \boldsymbol{v}_i\|_{\boldsymbol{x}_i}^2 \le 0,
\end{aligned}$$

where we used $P_{ik}P_{ki} = I$ and parallel transport is an isometry:

$$\langle P_{ik}\boldsymbol{v}_k, P_{ik}\boldsymbol{w}_k\rangle_{\boldsymbol{x}_i} = \langle \boldsymbol{v}_k, \boldsymbol{w}_k\rangle_{\boldsymbol{x}_k}.$$

□

Before we provide a criterion on velocity alignment, we recall Barbalat's lemma.

Lemma II.3. [4] *Suppose that a function $f : [0,\infty) \to \mathbb{R}$ is continuously differentiable, and $\lim_{t\to\infty} f(t) = \alpha \in \mathbb{R}$. If f' is uniformly continuous, then one has*

$$\lim_{t\to\infty} f'(t) = 0.$$

Now, we are ready to provide a sufficient condition for velocity alignment using Lemma II.2 and Lemma II.3.

Proposition II.1. *Let $\{(\boldsymbol{x}_i, \boldsymbol{v}_i)\}_{i=1}^N$ be a global smooth solution to* (II.4) *on a smooth manifold $(M, \langle\cdot,\cdot\rangle_M)$ satisfying*

$$\sup_{0\le t<\infty} |\mathcal{E}''(t)| < \infty. \tag{II.6}$$

Then, velocity alignment emerges:

$$\lim_{t\to+\infty} \|P_{ik}\boldsymbol{v}_k - \boldsymbol{v}_i\|_{\boldsymbol{x}_i}^2 = 0, \quad \forall\, i,k \in \{1,\cdots,N\}.$$

Proof. By Lemma II.2, $\mathcal{E}(t)$ is monotonically decreasing and is bounded below by 0. Hence $\lim_{t\to\infty} \mathcal{E}(t)$ exists. By the a priori condition (II.6), it is easy to see that $\mathcal{E}'$ is Lipschitz continuous. Hence it is uniformly continuous. Thus, it follows from Lemma II.3 and (II.5) that

$$\lim_{t\to\infty} \sum_{i,k=1}^N \psi(\mathrm{dist}(\boldsymbol{x}_i, \boldsymbol{x}_k))\|P_{ik}\boldsymbol{v}_k - \boldsymbol{v}_i\|_{\boldsymbol{x}_i}^2 = 0. \tag{II.7}$$

On the other hand,

$$\begin{aligned}
\psi_m\|P_{ik}\boldsymbol{v}_k - \boldsymbol{v}_i\|_{\boldsymbol{x}_i}^2 &\le \psi(\mathrm{dist}(\boldsymbol{x}_i, \boldsymbol{x}_k))\|P_{ik}\boldsymbol{v}_k - \boldsymbol{v}_i\|_{\boldsymbol{x}_i}^2 \\
&\le \sum_{i,k} \psi(\mathrm{dist}(\boldsymbol{x}_i, \boldsymbol{x}_k))\|P_{ik}\boldsymbol{v}_k - \boldsymbol{v}_i\|_{\boldsymbol{x}_i}^2.
\end{aligned} \tag{II.8}$$

Finally, (II.7) and (II.8) yield the desired estimate:

$$\lim_{t\to+\infty} \|P_{ik}\boldsymbol{v}_k - \boldsymbol{v}_i\|_{\boldsymbol{x}_i}^2 = 0, \quad \forall\, i,k \in \{1,\cdots,N\}.$$

□

III. The Cucker-Smale particles on a sphere

In this section, we present explicit form of the CS model on the sphere $\partial B_r(0)$ from the general form of the CS model introduced in previous section, and present emergent dynamics of the derived model under an a priori assumption. Throughout this section, we set

$$M = \partial B_r(0) \subset \mathbb{R}^3, \quad r > 0.$$

In this case, we can explicitly calculate covariant derivative and parallel transport in terms of state variables of $\mathbb{R}^3$, and we also use $\langle\cdot,\cdot\rangle$ and $\|\cdot\|$ to denote the standard inner product and norm in $\mathbb{R}^3$.

A. The reduction of the CS model to the sphere

In this subsection, we derive the reduction of the CS model (II.4) on a sphere in two steps:

- Step A: A priori, we assume that $\boldsymbol{x}_i$ lies on the sphere $\partial B_r(0)$, i.e., $\|\boldsymbol{x}_i\| = r$, and then we explicitly calculate $\frac{D\boldsymbol{v}_i}{dt}$ and P_{ik} to derive an explicit form of the CS model.
- Step B: We show that the norm of $\boldsymbol{x}_i$ is a constant of motion for the derived model so that $\boldsymbol{x}_i \in \partial B_r(0)$ which justifies our a priori assumption a posteriori.

For a unit vector $\boldsymbol{u} \in \mathbb{R}^3$, we set $[\boldsymbol{u}]_\times$ to be the matrix representation of the linear transformation which map any vector $\boldsymbol{v} \in \mathbb{R}^3$ to $\boldsymbol{u} \times \boldsymbol{v}$. Then, Rodrigues' rotation formula

[14] implies that the matrix for the rotation by the angle θ about the unit vector $\boldsymbol{u}$ is equal to

$$(\cos\theta)I + (\sin\theta)[\boldsymbol{u}]_\times + (1-\cos\theta)\boldsymbol{u}\boldsymbol{u}^\top, \tag{III.1}$$

where I is the 3×3 identity matrix.

Lemma III.1. *Let $\{(\boldsymbol{x}_i,\boldsymbol{v}_i)\}_{i=1}^N$ be a global smooth solution to system* (II.4) *on $\partial B_r(0)$. Then, we have*

$$\frac{D\boldsymbol{v}_i}{dt} = \dot{\boldsymbol{v}}_i - \frac{\langle \dot{\boldsymbol{v}}_i, \boldsymbol{x}_i\rangle}{r^2}\boldsymbol{x}_i, \qquad \text{and}$$
$$P_{ik} = \frac{\langle \boldsymbol{x}_k,\boldsymbol{x}_i\rangle}{r^2} I + \frac{1}{r^2}[\boldsymbol{x}_k\times\boldsymbol{x}_i]_\times + \frac{(\boldsymbol{x}_k\times\boldsymbol{x}_i)(\boldsymbol{x}_k\times\boldsymbol{x}_i)^\top}{r^2(r^2+\langle\boldsymbol{x}_k,\boldsymbol{x}_i\rangle)},$$

where $\langle\cdot,\cdot\rangle$ and $\|\cdot\|$ are Euclidean inner product and Euclidean norm in $\mathbb{R}^3$.

Proof. (i) Let $\boldsymbol{x}_i \in \partial B_r(0)$. Then, the exterior normal vector at $\boldsymbol{x}_i$ is given by

$$\nu_{\boldsymbol{x}_i} := \frac{\boldsymbol{x}_i}{r}.$$

Since $\partial B_r(0)$ is isometrically embedded in $\mathbb{R}^3$, $\frac{D\boldsymbol{v}_i}{dt}$ is the projection of $\dot{\boldsymbol{v}}_i$ on the tangent plane at $\boldsymbol{x}_i$:

$$\frac{D\boldsymbol{v}_i}{dt} = \dot{\boldsymbol{v}}_i - \left\langle \dot{\boldsymbol{v}}_i, \frac{\boldsymbol{x}_i}{r}\right\rangle\frac{\boldsymbol{x}_i}{r} = \dot{\boldsymbol{v}}_i - \frac{\langle \dot{\boldsymbol{v}}_i,\boldsymbol{x}_i\rangle}{r^2}\boldsymbol{x}_i.$$

(ii) The parallel transport P_{ik} moves a tangent vector at $\boldsymbol{x}_k$ to a tangent vector at $\boldsymbol{x}_i$ along the great circle containing these two points, so it is actually the matrix for the rotation by an angle of $\theta = \cos^{-1}\frac{\langle\boldsymbol{x}_k,\boldsymbol{x}_i\rangle}{r^2}$ about an axis in the direction of $\boldsymbol{u} := \frac{\boldsymbol{x}_k\times\boldsymbol{x}_i}{\|\boldsymbol{x}_k\times\boldsymbol{x}_i\|}$. Hence, we apply (III.1) with θ and $\boldsymbol{u}$ using the simple relation $\|\boldsymbol{x}_k\times\boldsymbol{x}_i\| = \sqrt{r^4-\langle\boldsymbol{x}_k,\boldsymbol{x}_i\rangle^2}$ to get

$$\begin{aligned}
P_{ik} &= \frac{\langle\boldsymbol{x}_k,\boldsymbol{x}_i\rangle}{r^2}I + \sqrt{1-\frac{\langle\boldsymbol{x}_k,\boldsymbol{x}_i\rangle^2}{r^4}}\Big[\frac{\boldsymbol{x}_k\times\boldsymbol{x}_i}{\|\boldsymbol{x}_k\times\boldsymbol{x}_i\|}\Big]_\times \\
&\quad + \Big(1-\frac{\langle\boldsymbol{x}_k,\boldsymbol{x}_i\rangle}{r^2}\Big)\frac{(\boldsymbol{x}_k\times\boldsymbol{x}_i)(\boldsymbol{x}_k\times\boldsymbol{x}_i)^\top}{\|\boldsymbol{x}_k\times\boldsymbol{x}_i\|^2} \\
&= \frac{\langle\boldsymbol{x}_k,\boldsymbol{x}_i\rangle}{r^2}I + \frac{1}{r^2}[\boldsymbol{x}_k\times\boldsymbol{x}_i]_\times + \frac{(\boldsymbol{x}_k\times\boldsymbol{x}_i)(\boldsymbol{x}_k\times\boldsymbol{x}_i)^\top}{r^2(r^2+\langle\boldsymbol{x}_k,\boldsymbol{x}_i\rangle)}.
\end{aligned}$$

□

Remark III.1. *We use the result (ii) of Lemma III.1, the vector triple identity*

$$(A\times B)\times C = \langle A,C\rangle B - \langle B,C\rangle A$$

and $\langle\boldsymbol{x}_k,\boldsymbol{v}_k\rangle = 0$ to see

$$\begin{aligned}
P_{ik}\boldsymbol{v}_k - \boldsymbol{v}_i &= \frac{\langle\boldsymbol{x}_k,\boldsymbol{x}_i\rangle}{r^2}\boldsymbol{v}_k + \frac{1}{r^2}(\boldsymbol{x}_k\times\boldsymbol{x}_i)\times\boldsymbol{v}_k \\
&\quad + \frac{\langle\boldsymbol{x}_k\times\boldsymbol{x}_i,\boldsymbol{v}_k\rangle}{r^2(r^2+\langle\boldsymbol{x}_k,\boldsymbol{x}_i\rangle)}(\boldsymbol{x}_k\times\boldsymbol{x}_i) - \boldsymbol{v}_i \\
&= \frac{\langle\boldsymbol{x}_k,\boldsymbol{x}_i\rangle}{r^2}\boldsymbol{v}_k + \frac{\langle\boldsymbol{x}_k,\boldsymbol{v}_k\rangle}{r^2}\boldsymbol{x}_i - \frac{\langle\boldsymbol{x}_i,\boldsymbol{v}_k\rangle}{r^2}\boldsymbol{x}_k \\
&\quad + \frac{\langle\boldsymbol{x}_k\times\boldsymbol{x}_i,\boldsymbol{v}_k\rangle}{r^2(r^2+\langle\boldsymbol{x}_k,\boldsymbol{x}_i\rangle)}(\boldsymbol{x}_k\times\boldsymbol{x}_i) - \boldsymbol{v}_i \\
&= \frac{\langle\boldsymbol{x}_k,\boldsymbol{x}_i\rangle}{r^2}\boldsymbol{v}_k - \frac{\langle\boldsymbol{v}_k,\boldsymbol{x}_i\rangle}{r^2}\boldsymbol{x}_k \\
&\quad + \frac{\langle\boldsymbol{x}_k\times\boldsymbol{x}_i,\boldsymbol{v}_k\rangle}{r^2(r^2+\langle\boldsymbol{x}_k,\boldsymbol{x}_i\rangle)}(\boldsymbol{x}_k\times\boldsymbol{x}_i) - \boldsymbol{v}_i,
\end{aligned} \tag{III.2}$$

where we used $\langle A,B\rangle = A^\top B$.

Now, we are ready to write down the CS model on the sphere $\partial B_r(0)$ explicitly. It follows from (II.4), Lemma III.1 and (III.2) that

$$\begin{cases}
\dot{\boldsymbol{x}}_i = \boldsymbol{v}_i, \quad t>0, \quad 1\le i\le N,\\
\dot{\boldsymbol{v}}_i = -\dfrac{\|\boldsymbol{v}_i\|^2}{r^2}\boldsymbol{x}_i + \dfrac{\kappa}{N}\displaystyle\sum_{k=1}^N \psi_{ik}\Big[\frac{\langle\boldsymbol{x}_k,\boldsymbol{x}_i\rangle}{r^2}\boldsymbol{v}_k \\
\qquad + \dfrac{\langle\boldsymbol{x}_k\times\boldsymbol{x}_i,\boldsymbol{v}_k\rangle}{r^2(r^2+\langle\boldsymbol{x}_k,\boldsymbol{x}_i\rangle)}(\boldsymbol{x}_k\times\boldsymbol{x}_i) - \dfrac{\langle\boldsymbol{v}_k,\boldsymbol{x}_i\rangle}{r^2}\boldsymbol{x}_k - \boldsymbol{v}_i\Big],\\
\boldsymbol{x}_i(0) = \boldsymbol{x}_i^0, \quad \boldsymbol{v}_i(0) = \boldsymbol{v}_i^0,
\end{cases} \tag{III.3}$$

where $\psi_{ik} = \psi\Big(r\arccos\frac{\langle\boldsymbol{x}_i,\boldsymbol{x}_k\rangle}{r^2}\Big)$. This is due to the fact that $r\arccos\frac{\langle\boldsymbol{x}_i,\boldsymbol{x}_k\rangle}{r^2}$ is the geodesic distance between $\boldsymbol{x}_i$ and $\boldsymbol{x}_k$.

Lemma III.2. *Let $\{(\boldsymbol{x}_i,\boldsymbol{v}_i)\}_{i=1}^N$ be a global smooth solution to the Cauchy problem* (III.3) *with the initial data satisfying the constraint:*

$$\langle\boldsymbol{x}_i^0,\boldsymbol{v}_i^0\rangle = 0, \quad i=1,\cdots,N.$$

Then for $i=1,\cdots,N$, we have

$$\langle\boldsymbol{x}_i(t),\boldsymbol{v}_i(t)\rangle = 0 \quad \text{and} \quad \|\boldsymbol{x}_i(t)\| = \|\boldsymbol{x}_i^0\|, \quad t\ge 0.$$

This is clear since the new CS model is defined completely intrinsically on the manifold. However, we give a detailed proof for completeness.

Proof. (i) For the first identity, we take an inner product of $\boldsymbol{x}_i$ with both sides of the second equation in (III.3) to get

$$\langle\dot{\boldsymbol{v}}_i,\boldsymbol{x}_i\rangle = -\|\boldsymbol{v}_i\|^2 - \frac{\kappa}{N}\langle\boldsymbol{x}_i,\boldsymbol{v}_i\rangle\sum_{k=1}^N\psi_{ik}.$$

This yields

$$\begin{aligned}
\frac{d}{dt}\langle\boldsymbol{x}_i,\boldsymbol{v}_i\rangle^2 &= 2\langle\boldsymbol{x}_i,\boldsymbol{v}_i\rangle(\langle\boldsymbol{x}_i,\dot{\boldsymbol{v}}_i\rangle + \|\boldsymbol{v}_i\|^2)\\
&= -\frac{2\kappa}{N}\langle\boldsymbol{x}_i,\boldsymbol{v}_i\rangle^2\sum_{k=1}^N\psi_{ik} \le 0.
\end{aligned}$$

Therefore, we have

$$\langle\boldsymbol{x}_i(t),\boldsymbol{v}_i(t)\rangle^2 \le \langle\boldsymbol{x}_i^0,\boldsymbol{v}_i^0\rangle^2 = 0.$$

This implies

$$\langle\boldsymbol{x}_i(t),\boldsymbol{v}_i(t)\rangle = 0, \quad t\ge 0, \quad 1\le i\le N.$$

(ii) The second relation directly follows from the first relation:

$$\frac{1}{2}\frac{d}{dt}\|\boldsymbol{x}_i(t)\|^2 = \langle\boldsymbol{x}_i(t),\boldsymbol{v}_i(t)\rangle = 0, \quad t>0.$$

□

B. Emergent dynamics

In this subsection, we study the emergent dynamics of system (III.3). Without loss of generality, we assume $r \equiv 1$. Then, system (III.3) becomes

$$\begin{cases} \dot{\boldsymbol{x}}_i = \boldsymbol{v}_i, \quad t > 0, \quad 1 \le i \le N, \\ \dot{\boldsymbol{v}}_i = -\|\boldsymbol{v}_i\|^2 \boldsymbol{x}_i + \dfrac{\kappa}{N} \displaystyle\sum_{k=1}^N \psi_{ik} \Big(\langle \boldsymbol{x}_k, \boldsymbol{x}_i \rangle \boldsymbol{v}_k \\ + \dfrac{\langle \boldsymbol{x}_k \times \boldsymbol{x}_i, \boldsymbol{v}_k \rangle}{1 + \langle \boldsymbol{x}_k, \boldsymbol{x}_i \rangle} (\boldsymbol{x}_k \times \boldsymbol{x}_i) - \langle \boldsymbol{v}_k, \boldsymbol{x}_i \rangle \boldsymbol{x}_k - \boldsymbol{v}_i \Big), \\ \boldsymbol{x}_i(0) = \boldsymbol{x}_i^0, \quad \boldsymbol{v}_i(0) = \boldsymbol{v}_i^0. \end{cases} \tag{III.4}$$

In the sequel, we consider the many-body system with $N \ge 3$ and two-particle system separately. For the former case, we present the emergent dynamics for system (III.4) under the following a priori conditions:

$$\sup_{0 \le t < \infty} \max_{1 \le i,k \le N} \arccos \langle \boldsymbol{x}_k(t), \boldsymbol{x}_i(t) \rangle < \pi \quad \text{and} \quad \psi \equiv 1. \tag{III.5}$$

This condition implies that the geodesic distance between any two particles is always less than π, thereby guaranteeing that the parallel transport operators P_{ik} are well-defined for $t \ge 0$.

1) The many-particle system: In this part, we present an a priori velocity alignment estimate for (III.4):

$$\lim_{t \to \infty} \|P_{ij} \boldsymbol{v}_j(t) - \boldsymbol{v}_i(t)\| = 0, \quad \forall\, i, j \in \{1, \cdots, N\}.$$

For each $i \ne j$, we set

$$\mathbf{p}_{ij} := \frac{\boldsymbol{x}_i \times \boldsymbol{x}_j}{\|\boldsymbol{x}_i \times \boldsymbol{x}_j\|}, \quad \mathbf{q}_{ij} := \frac{\boldsymbol{x}_i + \boldsymbol{x}_j}{\|\boldsymbol{x}_i + \boldsymbol{x}_j\|}, \quad \mathbf{r}_{ij} := \frac{\boldsymbol{x}_i - \boldsymbol{x}_j}{\|\boldsymbol{x}_i - \boldsymbol{x}_j\|}. \tag{III.6}$$

Then, it is easy to check that $\{\mathbf{p}_{ij}, \mathbf{q}_{ij}, \mathbf{r}_{ij}\}$ forms an orthonormal basis for $\mathbb{R}^3$ (whenever $|\langle \boldsymbol{x}_i, \boldsymbol{x}_j \rangle| < 1$) and

$$I = \mathbf{p}_{ij}\mathbf{p}_{ij}^\top + \mathbf{q}_{ij}\mathbf{q}_{ij}^\top + \mathbf{r}_{ij}\mathbf{r}_{ij}^\top.$$

Lemma III.3. *Let $\{(\boldsymbol{x}_i, \boldsymbol{v}_i)\}_{i=1}^N$ be a global solution to system* (III.4) *satisfying*

$$\langle \boldsymbol{x}_i(t), \boldsymbol{v}_i(t) \rangle = 0 \quad t \ge 0, \quad 1 \le i, j \le N.$$

Then the following assertions hold.

1) P_{ij} is given by the following rotation matrix:

$$\begin{aligned} P_{ij} &= \mathbf{p}_{ij}\mathbf{p}_{ij}^\top \\ &+ \left(\langle \boldsymbol{x}_i, \boldsymbol{x}_j \rangle \mathbf{q}_{ij} + \frac{\|\boldsymbol{x}_i + \boldsymbol{x}_j\| \|\boldsymbol{x}_i - \boldsymbol{x}_j\|}{2} \mathbf{r}_{ij} \right) \mathbf{q}_{ij}^\top \\ &+ \left(\langle \boldsymbol{x}_i, \boldsymbol{x}_j \rangle \mathbf{r}_{ij} - \frac{\|\boldsymbol{x}_i + \boldsymbol{x}_j\| \|\boldsymbol{x}_i - \boldsymbol{x}_j\|}{2} \mathbf{q}_{ij} \right) \mathbf{r}_{ij}^\top. \end{aligned}$$

2) P_{ij} satisfies

$$\boldsymbol{x}_i \times P_{ij} \boldsymbol{v}_j = P_{ij} \boldsymbol{x}_j \times P_{ij} \boldsymbol{v}_j = P_{ij}(\boldsymbol{x}_j \times \boldsymbol{v}_j).$$

Proof. We give a proof in Appendix A. □

Lemma III.4. *Let $\{(\boldsymbol{x}_i, \boldsymbol{v}_i)\}_{i=1}^N$ be a global smooth solution to system* (III.4) *satisfying the following relations:*

$$\langle \boldsymbol{x}_i(t), \boldsymbol{v}_i(t) \rangle = 0 \quad \textit{and} \quad |\langle \boldsymbol{x}_i(t), \boldsymbol{x}_j(t) \rangle| < 1,$$

for $\forall\, t \ge 0, \quad 1 \le i, j \le N$. Then, we have

$$\begin{aligned} (i)\ \dot{\mathbf{p}}_{ij} &= \frac{1}{\|\boldsymbol{x}_i \times \boldsymbol{x}_j\|} \Big(\langle \boldsymbol{x}_i \times \boldsymbol{v}_i - \boldsymbol{x}_j \times \boldsymbol{v}_j, \mathbf{q}_{ij} \rangle \mathbf{q}_{ij} \\ &\qquad - \langle \boldsymbol{x}_i \times \boldsymbol{v}_i - \boldsymbol{x}_j \times \boldsymbol{v}_j, \mathbf{r}_{ij} \rangle \mathbf{r}_{ij} \Big). \\ (ii)\ \dot{\mathbf{q}}_{ij} &= \frac{1}{\|\boldsymbol{x}_i \times \boldsymbol{x}_j\|} \Big(- \langle \boldsymbol{x}_i \times \boldsymbol{v}_i - \boldsymbol{x}_j \times \boldsymbol{v}_j, \mathbf{q}_{ij} \rangle \mathbf{p}_{ij} \\ &\qquad + \frac{\langle \boldsymbol{x}_i, \boldsymbol{v}_j \rangle - \langle \boldsymbol{x}_j, \boldsymbol{v}_i \rangle}{2} \mathbf{r}_{ij} \Big). \\ (iii)\ \dot{\mathbf{r}}_{ij} &= \frac{1}{\|\boldsymbol{x}_i \times \boldsymbol{x}_j\|} \Big(\langle \boldsymbol{x}_i \times \boldsymbol{v}_i - \boldsymbol{x}_j \times \boldsymbol{v}_j, \mathbf{r}_{ij} \rangle \mathbf{p}_{ij} \\ &\qquad - \frac{\langle \boldsymbol{x}_i, \boldsymbol{v}_j \rangle - \langle \boldsymbol{x}_j, \boldsymbol{v}_i \rangle}{2} \mathbf{q}_{ij} \Big). \end{aligned}$$

Proof. We give a proof in Appendix B. □

Proposition III.1. *Let $\{(\boldsymbol{x}_i, \boldsymbol{v}_i)\}_{i=1}^N$ be a global smooth solution to system* (III.4) *with* (III.5) *satisfying*

$$|\langle \boldsymbol{x}_i(t), \boldsymbol{x}_j(t) \rangle| < 1.$$

Then, we have

$$\begin{aligned} &\boldsymbol{x}_i \times (P_{ij} \boldsymbol{v}_j - \boldsymbol{v}_i) = -\langle \boldsymbol{x}_i \times \boldsymbol{v}_i - \boldsymbol{x}_j \times \boldsymbol{v}_j, \mathbf{p}_{ij} \rangle \mathbf{p}_{ij} \\ &\quad - \langle \boldsymbol{x}_i \times \boldsymbol{v}_i - \boldsymbol{x}_j \times \boldsymbol{v}_j, \mathbf{r}_{ij} \rangle \Big(- \frac{\|\boldsymbol{x}_i - \boldsymbol{x}_j\|}{\|\boldsymbol{x}_i + \boldsymbol{x}_j\|} \mathbf{q}_{ij} + \mathbf{r}_{ij} \Big). \end{aligned}$$

Proof. We give a proof in Appendix C. □

Remark III.2. *We can use*

$$\mathbf{p}_{ji} = -\mathbf{p}_{ij}, \quad \mathbf{q}_{ji} = \mathbf{q}_{ij} \quad \textit{and} \quad \mathbf{r}_{ji} = -\mathbf{r}_{ij}$$

and Proposition III.1 to obtain

$$\begin{aligned} &\frac{d}{dt} (\boldsymbol{x}_1 \times \boldsymbol{v}_1 - \boldsymbol{x}_2 \times \boldsymbol{v}_2) = \boldsymbol{x}_1 \times \dot{\boldsymbol{v}}_1 - \boldsymbol{x}_2 \times \dot{\boldsymbol{v}}_2 \\ &\quad = \boldsymbol{x}_1 \times \left(-\|\boldsymbol{v}_1\|^2 \boldsymbol{x}_1 + \frac{\kappa}{2} (P_{12} \boldsymbol{v}_2 - \boldsymbol{v}_1) \right) \\ &\quad - \boldsymbol{x}_2 \times \left(-\|\boldsymbol{v}_2\|^2 \boldsymbol{x}_2 + \frac{\kappa}{2} (P_{21} \boldsymbol{v}_1 - \boldsymbol{v}_2) \right) \\ &\quad = -\kappa \big(\langle \boldsymbol{x}_1 \times \boldsymbol{v}_1 - \boldsymbol{x}_2 \times \boldsymbol{v}_2, \mathbf{p}_{12} \rangle \mathbf{p}_{12} \\ &\qquad + \langle \boldsymbol{x}_1 \times \boldsymbol{v}_1 - \boldsymbol{x}_2 \times \boldsymbol{v}_2, \mathbf{r}_{12} \rangle \mathbf{r}_{12} \big) \end{aligned}$$

for a two-particle system.

Now, we are ready to provide the emergent dynamics for the CS model on $\mathbb{S}^2$.

Theorem III.1. *Suppose the initial data satisfy*

$$\langle \boldsymbol{x}_i^0, \boldsymbol{v}_i^0 \rangle = 0, \quad i = 1, \cdots, N,$$

and let $\{(\boldsymbol{x}_i, \boldsymbol{v}_i)\}_{i=1}^N$ be a global smooth solution to system (III.4) - (III.5) *satisfying a priori condition:*

$$\sup_{0 \le t < \infty} \max_{i,k} \arccos \langle \boldsymbol{x}_k(t), \boldsymbol{x}_i(t) \rangle \le \alpha < \pi.$$

Then, one has

$$\lim_{t \to \infty} \max_{1 \le i,k \le N} \|P_{ik} \boldsymbol{v}_k - \boldsymbol{v}_i\| = 0.$$

Proof. By Proposition II.1, it suffices to show that the following quantity is bounded:

$$\frac{d}{dt} \|P_{ik} \boldsymbol{v}_k - \boldsymbol{v}_i\|^2.$$

It follows from (III.2) that

$$\frac{d}{dt}\|P_{ik}\boldsymbol{v}_k-\boldsymbol{v}_i\|^2=\frac{d}{dt}\Big\|\langle \boldsymbol{x}_k,\boldsymbol{x}_i\rangle \boldsymbol{v}_k + \frac{\langle \boldsymbol{x}_k\times \boldsymbol{x}_i,\boldsymbol{v}_k\rangle}{1+\langle \boldsymbol{x}_k,\boldsymbol{x}_i\rangle}(\boldsymbol{x}_k\times\boldsymbol{x}_i)-\langle \boldsymbol{v}_k,\boldsymbol{x}_i\rangle\boldsymbol{x}_k-\boldsymbol{v}_i\Big\|^2. \tag{III.7}$$

Note that the uniform boundedness of $\|\boldsymbol{x}_i\|$, $\|\boldsymbol{v}_i\|$ and $\|\dot{\boldsymbol{v}}_i\|$ implies boundedness of (III.7).

• (Uniform boundedness of $\|\boldsymbol{x}_i\|$ and $\|\boldsymbol{v}_i\|$): It follows from Lemma II.2 that

$$\|\boldsymbol{x}_i(t)\|=1,\quad t\ge 0\quad\text{and}\quad \sup_{0\le t<\infty}\|\boldsymbol{v}_i(t)\|\le\sqrt{\mathcal{E}^0}, \tag{III.8}$$

where we used a simplified notation $\mathcal{E}^0:=\mathcal{E}(0)$.

• (Uniform boundedness of $\|\dot{\boldsymbol{v}}_i\|$): By the a priori assumption, one has

$$1+\langle \boldsymbol{x}_i,\boldsymbol{x}_k\rangle\ge 1+\cos\alpha>0.$$

Note that

$$\dot{\boldsymbol{v}}_i=-\|\boldsymbol{v}_i\|^2\boldsymbol{x}_i+\frac{\kappa}{N}\sum_{k=1}^N\psi_{ik}\Big(\langle \boldsymbol{x}_k,\boldsymbol{x}_i\rangle\boldsymbol{v}_k + \frac{\langle \boldsymbol{x}_k\times\boldsymbol{x}_i,\boldsymbol{v}_k\rangle}{1+\langle \boldsymbol{x}_k,\boldsymbol{x}_i\rangle}(\boldsymbol{x}_k\times\boldsymbol{x}_i)-\langle \boldsymbol{v}_k,\boldsymbol{x}_i\rangle\boldsymbol{x}_k-\boldsymbol{v}_i\Big).$$

This, (III.8) and $\|\boldsymbol{x}_i\times\boldsymbol{x}_j\|\le 1$ imply

$$\begin{aligned}\|\dot{\boldsymbol{v}}_i\|&\le\|\boldsymbol{v}_i\|^2+\frac{\kappa}{N}\sum_{k=1}^N\left(\|\boldsymbol{v}_k\|+\frac{\|\boldsymbol{v}_k\|}{1+\cos\alpha}+\|\boldsymbol{v}_k\|+\|\boldsymbol{v}_i\|\right)\\ &\le\mathcal{E}^0+\kappa\left(3+\frac{1}{1+\cos\alpha}\right)\sqrt{\mathcal{E}^0}.\end{aligned}$$

□

2) The two-particle system: In this part, we consider the two-particle case for the CS model on $\mathbb{S}^2$, (III.4):

$$\begin{cases}\dot{\boldsymbol{x}}_1=\boldsymbol{v}_1,\quad \dot{\boldsymbol{x}}_2=\boldsymbol{v}_2,\quad t>0,\\ \dot{\boldsymbol{v}}_1=-\|\boldsymbol{v}_1\|^2\boldsymbol{x}_1+\frac{\kappa}{2}\Big[-\psi_{11}\langle\boldsymbol{v}_1,\boldsymbol{x}_1\rangle\boldsymbol{x}_1+\psi_{12}\Big(\langle\boldsymbol{x}_2,\boldsymbol{x}_1\rangle\boldsymbol{v}_2\\ \qquad+\frac{\langle\boldsymbol{x}_2\times\boldsymbol{x}_1,\boldsymbol{v}_2\rangle}{1+\langle\boldsymbol{x}_2,\boldsymbol{x}_1\rangle}(\boldsymbol{x}_2\times\boldsymbol{x}_1)-\langle\boldsymbol{v}_2,\boldsymbol{x}_1\rangle\boldsymbol{x}_2-\boldsymbol{v}_1\Big)\Big],\\ \dot{\boldsymbol{v}}_2=-\|\boldsymbol{v}_2\|^2\boldsymbol{x}_2+\frac{\kappa}{2}\Big[-\psi_{22}\langle\boldsymbol{v}_2,\boldsymbol{x}_2\rangle\boldsymbol{x}_2+\psi_{21}\Big(\langle\boldsymbol{x}_1,\boldsymbol{x}_2\rangle\boldsymbol{v}_1\\ \qquad+\frac{\langle\boldsymbol{x}_1\times\boldsymbol{x}_2,\boldsymbol{v}_1\rangle}{1+\langle\boldsymbol{x}_1,\boldsymbol{x}_2\rangle}(\boldsymbol{x}_1\times\boldsymbol{x}_2)-\langle\boldsymbol{v}_1,\boldsymbol{x}_2\rangle\boldsymbol{x}_1-\boldsymbol{v}_2\Big)\Big].\end{cases} \tag{III.9}$$

In this case, we replace the a priori assumption (III.5) by a condition on the initial data and coupling strength (cf. Theorem III.2). Once we have good control of the geodesic distance $\arccos\langle\boldsymbol{x}_1,\boldsymbol{x}_2\rangle$, say

$$\sup_{0\le t<\infty}\arccos\langle\boldsymbol{x}_1(t),\boldsymbol{x}_2(t)\rangle<\pi, \tag{III.10}$$

then the velocity alignment estimate follows from Theorem III.1. In the sequel, we begin to estimate the geodesic distance. Let $(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{v}_1,\boldsymbol{v}_2)$ be a solution to (III.9). We set

$$\mathbf{p}:=\frac{\boldsymbol{x}_1\times\boldsymbol{x}_2}{\|\boldsymbol{x}_1\times\boldsymbol{x}_2\|},\ \mathbf{q}:=\frac{\boldsymbol{x}_1+\boldsymbol{x}_2}{\|\boldsymbol{x}_1+\boldsymbol{x}_2\|},\ \mathbf{r}:=\frac{\boldsymbol{x}_1-\boldsymbol{x}_2}{\|\boldsymbol{x}_1-\boldsymbol{x}_2\|}. \tag{III.11}$$

To derive the estimate (III.10), we differentiate the geodesic distance $\arccos\langle\boldsymbol{x}_1(t),\boldsymbol{x}_2(t)\rangle$ to obtain

$$\begin{aligned}\frac{d}{dt}\arccos\langle\boldsymbol{x}_1,\boldsymbol{x}_2\rangle&=-\frac{\langle\boldsymbol{x}_1,\boldsymbol{v}_2\rangle+\langle\boldsymbol{x}_2,\boldsymbol{v}_1\rangle}{\sqrt{1-\langle\boldsymbol{x}_1,\boldsymbol{x}_2\rangle^2}}\\ &=-\frac{\langle\boldsymbol{x}_1,\boldsymbol{v}_2\rangle+\langle\boldsymbol{x}_2,\boldsymbol{v}_1\rangle}{\|\boldsymbol{x}_1\times\boldsymbol{x}_2\|},\end{aligned} \tag{III.12}$$

where we used that $\|\boldsymbol{x}_1\times\boldsymbol{x}_2\|=\sqrt{1-\langle\boldsymbol{x}_1,\boldsymbol{x}_2\rangle^2}$ on the unit sphere.

On the other hand, we use the definitions in (III.11), and $\langle\boldsymbol{x}_i,\boldsymbol{v}_i\rangle=0$ to get

$$\begin{aligned}&\langle\boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{p}\rangle\\ &=\frac{1}{\|\boldsymbol{x}_1\times\boldsymbol{x}_2\|}\Big(\langle\boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\boldsymbol{x}_1\times\boldsymbol{x}_2\rangle\Big)\\ &=\frac{1}{\|\boldsymbol{x}_1\times\boldsymbol{x}_2\|}\Big(\langle\boldsymbol{x}_1\times\boldsymbol{v}_1,\boldsymbol{x}_1\times\boldsymbol{x}_2\rangle-\langle\boldsymbol{x}_2\times\boldsymbol{v}_2,\boldsymbol{x}_1\times\boldsymbol{x}_2\rangle\Big)\\ &=\frac{1}{\|\boldsymbol{x}_1\times\boldsymbol{x}_2\|}\Big(\langle\boldsymbol{v}_1,\boldsymbol{x}_2\rangle+\langle\boldsymbol{v}_2,\boldsymbol{x}_1\rangle\Big),\end{aligned} \tag{III.13}$$

where we used the vector identities:

$$\begin{aligned}&\langle A\times B,C\times D\rangle=\langle A,B\times(C\times D)\rangle\\ &=\langle A,\langle B,D\rangle C-\langle B,C\rangle D\rangle=\langle A,C\rangle\langle B,D\rangle-\langle A,D\rangle\langle B,C\rangle.\end{aligned}$$

Finally, we combine (III.12) and (III.13) to find

$$\frac{d}{dt}\arccos\langle\boldsymbol{x}_1,\boldsymbol{x}_2\rangle=-\langle\boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{p}\rangle. \tag{III.14}$$

In the next lemma, we estimate the term appearing in the R.H.S. of (III.14).

Lemma III.5. *Let $\{(\boldsymbol{x}_i,\boldsymbol{v}_i)\}_{i=1}^2$ be a solution to system* (III.9) *satisfying*

$$|\langle\boldsymbol{x}_i(t),\boldsymbol{x}_i(t)\rangle|<1,\quad\forall\ t>0.$$

Then, the following estimates hold.

(i)
$$\frac{d}{dt}\langle\boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{p}\rangle=-\kappa\langle\boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{p}\rangle+\frac{\langle\boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{q}\rangle^2-\langle\boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{r}\rangle^2}{\|\boldsymbol{x}_1\times\boldsymbol{x}_2\|}. \tag{III.15}$$

(ii)
$$\begin{aligned}&\frac{d}{dt}\langle\boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{q}\rangle\\ &=-\frac{\langle\boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{p}\rangle\langle\boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{q}\rangle}{\|\boldsymbol{x}_1\times\boldsymbol{x}_2\|}\\ &\quad+\frac{\langle\boldsymbol{x}_1,\boldsymbol{v}_2\rangle-\langle\boldsymbol{x}_2,\boldsymbol{v}_1\rangle}{2\|\boldsymbol{x}_1\times\boldsymbol{x}_2\|}\langle\boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{r}\rangle.\end{aligned}$$

(iii)
$$\begin{aligned}&\frac{d}{dt}\langle\boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{r}\rangle\\ &=\frac{\langle\boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{p}\rangle\langle\boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{r}\rangle}{\|\boldsymbol{x}_1\times\boldsymbol{x}_2\|}\\ &\quad-\frac{\langle\boldsymbol{x}_1,\boldsymbol{v}_2\rangle-\langle\boldsymbol{x}_2,\boldsymbol{v}_1\rangle}{2\|\boldsymbol{x}_1\times\boldsymbol{x}_2\|}\langle\boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{q}\rangle\\ &\quad-\kappa\langle\boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{r}\rangle.\end{aligned}$$

Proof. We give a proof in Appendix D. □

In the next lemma, we estimate the following term appearing in (III.15):

$$\Delta(\boldsymbol{x},\boldsymbol{v}) := \langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{q}\rangle^2-\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{r}\rangle^2.$$

Lemma III.6. *Suppose that parameters δ,κ and initial data satisfy*

$$\delta\in(0,1),\qquad \frac{6\sqrt{\mathcal{E}^0}}{\sqrt{1-\delta^2}}\le\kappa,$$

and let $\{(\boldsymbol{x}_i,\boldsymbol{v}_i)\}_{i=1}^2$ be a smooth solution to system (III.8) *satisfying the a priori condition $\sup\limits_{0\le t<\infty}|\langle\boldsymbol{x}_1(t),\boldsymbol{x}_2(t)\rangle|\le\delta$. Then, we have*

$$\frac{d}{dt}\Delta(\boldsymbol{x},\boldsymbol{v})\ge-\kappa\Delta(\boldsymbol{x},\boldsymbol{v}),\quad t>0. \tag{III.16}$$

Proof. We give a proof in Appendix E. □

Proposition III.2. *Suppose that the parameters δ,κ and initial data satisfy*

$$\delta\in(0,1)\quad\text{and}\quad\frac{6\sqrt{\mathcal{E}^0}}{\sqrt{1-\delta^2}}\le\kappa,$$

and let $\{(\boldsymbol{x}_i,\boldsymbol{v}_i)\}_{i=1}^2$ be a global smooth solution to system (III.8) *satisfying $\sup\limits_{t_0\le t\le T}|\langle\boldsymbol{x}_1(t),\boldsymbol{x}_2(t)\rangle|\le\delta$. Then, for $t\in[t_0,T]$*

$$\begin{aligned}&\arccos\langle\boldsymbol{x}_1(t),\boldsymbol{x}_2(t)\rangle\\&\le\arccos\langle\boldsymbol{x}_1(t_0),\boldsymbol{x}_2(t_0)\rangle+\frac{\sqrt{2\mathcal{E}^0}}{\kappa}+\frac{2\mathcal{E}^0}{\kappa^2\sqrt{1-\delta^2}}.\end{aligned}\tag{III.17}$$

Proof. We give a proof in Appendix F. □

In the following theorem, we remove the a priori assumption $\sup\limits_{t_0\le t\le T}|\langle\boldsymbol{x}_1(t),\boldsymbol{x}_2(t)\rangle|\le\delta$. in Proposition III.2.

Theorem III.2. *Suppose that parameters and initial data satisfy*

$$\delta\in(0,1),\quad\langle\boldsymbol{x}_1(0),\boldsymbol{x}_2(0)\rangle\ge\delta,\quad\frac{6\sqrt{\mathcal{E}^0}}{\sqrt{1-\delta^2}}\le\kappa,$$
$$\pi-2\arccos\delta>\frac{\sqrt{2\mathcal{E}^0}}{\kappa}+\frac{2\mathcal{E}^0}{\kappa^2\sqrt{1-\delta^2}},$$

and let $\{(\boldsymbol{x}_i,\boldsymbol{v}_i)\}_{i=1}^2$ be a global smooth solution to system (III.8). *Then we have*

$$\arccos\langle\boldsymbol{x}_1(t),\boldsymbol{x}_2(t)\rangle<\pi,\quad\textit{for all } t>0.$$

Proof. We split its proof into two steps.

- Step A: We first demonstrate

$$\langle\boldsymbol{x}_1(t),\boldsymbol{x}_2(t)\rangle>-\delta\quad\text{for all } t>0. \tag{III.18}$$

Suppose not. Then, we set

$$T^*:=\inf\{t>0:\langle\boldsymbol{x}_1(t),\boldsymbol{x}_2(t)\rangle\le-\delta\}<\infty,$$
$$t_0:=\sup\{0\le t\le T^*:\langle\boldsymbol{x}_1(t),\boldsymbol{x}_2(t)\rangle\ge\delta\}.$$

Then we have

$$|\langle\boldsymbol{x}_1(t),\boldsymbol{x}_2(t)\rangle|\le\delta\quad\text{for } t\in[t_0,T^*].$$

On the other hand, by the geometric shape of $\arccos\theta$, one has

$$\arccos(-\delta)+\arccos\delta=\pi.$$

Now, it follows from Proposition III.2 that

$$\begin{aligned}\arccos(-\delta)&=\arccos\langle\boldsymbol{x}_1(T^*),\boldsymbol{x}_2(T^*)\rangle\\&\le\arccos\langle\boldsymbol{x}_1(t_0),\boldsymbol{x}_2(t_0)\rangle+\frac{\sqrt{2\mathcal{E}^0}}{\kappa}+\frac{2\mathcal{E}^0}{\kappa^2\sqrt{1-\delta^2}}\\&\le\arccos\delta+\frac{\sqrt{2\mathcal{E}^0}}{\kappa}+\frac{2\mathcal{E}^0}{\kappa^2\sqrt{1-\delta^2}}.\end{aligned}$$

Since $\arccos(-\delta)+\arccos\delta=\pi$, we have

$$\pi-2\arccos\delta\le\frac{\sqrt{2\mathcal{E}^0}}{\kappa}+\frac{2\mathcal{E}^0}{\kappa^2\sqrt{1-\delta^2}}$$

which is contradictory to the condition on δ. Hence we have (III.18).

- Step B: It follows from (III.18), the monotonicity of $\arccos(\cdot)$ and $\delta\in(0,1)$ that

$$\arccos\langle\boldsymbol{x}_1,\boldsymbol{x}_2\rangle<\cos(-\delta)<\arccos(-1)=\pi$$

which verifies the desired estimate. □

IV. Cucker-Smale particles on the unit circle

In this section, we present the Cucker-Smale model on the unit circle $\mathbb{S}^1$ and study its emergent dynamics.

A. The reduction of the CS model to the unit circle

For the reduction of (II.4) to the unit circle $\mathbb{S}^1$, we think of $\mathbb{S}^1$ as standardly embedded in the xy-plane. We introduce the following notation: for $\theta\in\mathbb{R}$,

$$\mathbf{e}(\theta):=(\cos\theta,\sin\theta),\qquad\mathbf{e}^\perp(\theta):=(-\sin\theta,\cos\theta).$$

In those terms, the positions take the following form:

$$\boldsymbol{x}_i(t)=\mathbf{e}(\theta_i(t)),\quad t\ge0,\quad i=1,\cdots,N.$$

Then we have

$$\boldsymbol{v}_i=\dot\theta_i\mathbf{e}^\perp(\theta_i)\quad\text{and}\quad\dot{\boldsymbol{v}}_i=\ddot\theta_i\mathbf{e}^\perp(\theta_i)-\big(\dot\theta_i\big)^2\mathbf{e}(\theta_i).$$

We extend the function $\psi:[0,\pi]\to\mathbb{R}$ to $\widetilde\psi:\mathbb{R}\to\mathbb{R}$ as follows:

$$\widetilde\psi|_{[0,\pi]}=\psi,\quad\widetilde\psi(-\theta)=\widetilde\psi(\theta),\quad\widetilde\psi(\theta+2\pi)=\widetilde\psi(\theta),\quad\theta\in\mathbb{R}.$$

Analogously to the sphere model in Section III, we have $\frac{D}{dt}\boldsymbol{v}_i=\dot{\boldsymbol{v}}_i-\langle\dot{\boldsymbol{v}}_i,\boldsymbol{x}_i\rangle\boldsymbol{x}_i$ and P_{ik} is actually the rotation by an angle of $\theta_i-\theta_k$. By substituting this into the CS model (II.4) and equating the L.H.S. and R.H.S. of the resulting relations, we obtain

$$\begin{aligned}\ddot\theta_i\mathbf{e}^\perp(\theta_i)=\frac{\kappa}{N}\sum_{k=1}^N\widetilde\psi(\theta_k-\theta_i)\Big(&\dot\theta_k\cos(\theta_k-\theta_i)\mathbf{e}^\perp(\theta_k)\\&+\dot\theta_k\sin(\theta_k-\theta_i)\mathbf{e}(\theta_k)-\dot\theta_i\mathbf{e}^\perp(\theta_i)\Big).\end{aligned}\tag{IV.1}$$

Now we rewrite (IV.1) in complex polar coordinates:

$$\begin{aligned} \mathrm{i}\ddot{\theta}_i e^{\mathrm{i}\theta_i} = \frac{\kappa}{N}\sum_{k=1}^N \widetilde{\psi}(\theta_k-\theta_i)\Big(\mathrm{i}\dot{\theta}_k\cos(\theta_k-\theta_i)e^{\mathrm{i}\theta_k} \\ +\dot{\theta}_k\sin(\theta_k-\theta_i)e^{\mathrm{i}\theta_k}-\mathrm{i}\dot{\theta}_i e^{\mathrm{i}\theta_i}\Big). \end{aligned} \tag{IV.2}$$

We multiply $-\mathrm{i}e^{-\mathrm{i}\theta_i}$ to both side of (IV.2) and simplify the resulting relations to obtain

$$\begin{aligned} \ddot{\theta}_i &= \frac{\kappa}{N}\sum_{k=1}^N \widetilde{\psi}(\theta_k-\theta_i)\Big(\dot{\theta}_k\cos(\theta_k-\theta_i)e^{\mathrm{i}(\theta_k-\theta_i)} \\ &\quad -\mathrm{i}\dot{\theta}_k\sin(\theta_k-\theta_i)e^{\mathrm{i}(\theta_k-\theta_i)}-\dot{\theta}_i\Big) \\ &= \frac{\kappa}{N}\sum_{k=1}^N \widetilde{\psi}(\theta_k-\theta_i)(\dot{\theta}_k-\dot{\theta}_i). \end{aligned}$$

Hence, we have the second-order model:

$$\ddot{\theta}_i = \frac{\kappa}{N}\sum_{k=1}^N \widetilde{\psi}(\theta_k-\theta_i)(\dot{\theta}_k-\dot{\theta}_i). \tag{IV.3}$$

Remark IV.1. *Although there are more than one (actually two) length-minimizing geodesics from a point $\boldsymbol{x}_p$ on $\mathbb{S}^1$ to its antipodal point $\boldsymbol{x}_q$, the parallel transports of a tangent vector at $\boldsymbol{x}_p$ along those geodesics to $\boldsymbol{x}_q$ are the same. Hence we can naturally extend the domains of the operators P_{ik} to $\mathbb{S}^1\times\mathbb{S}^1$, which is not the case for $\mathbb{S}^2$.*

Next, we briefly discuss a reduction of (IV.3) to first-order. For this, we integrate (IV.3) to get

$$\dot{\theta}_i(t) = \dot{\theta}_i(0) + \frac{\kappa}{N}\sum_{k=1}^N \int_{\theta_k(0)-\theta_i(0)}^{\theta_k(t)-\theta_i(t)} \widetilde{\psi}(s)ds. \tag{IV.4}$$

We set

$$\alpha := \frac{1}{\pi}\int_0^\pi \tilde{\psi}(s)ds \quad \text{and} \quad f(\tau) := \tilde{\psi}(\tau)-\alpha. \tag{IV.5}$$

Then f is a 2π-periodic even function which monotonically decreases on $[0,\pi]$, with

$$f(\tau)+\alpha \ge 0 \quad (\tau\in\mathbb{R}) \quad \text{and} \quad \int_0^\pi f(\tau)d\tau = 0.$$

We also define

$$F(\tau) := \int_0^\tau f(s)ds. \tag{IV.6}$$

Then, it follows from (IV.4), (IV.5) and (IV.6) that

$$\begin{aligned} \dot{\theta}_i(t) = \dot{\theta}_i(0) &+ \frac{\kappa}{N}\sum_{k=1}^N \Big[\alpha(\theta_k(t)-\theta_i(t)) + F(\theta_k(t)-\theta_i(t))\Big] \\ &- \frac{\kappa}{N}\sum_{k=1}^N \Big[\alpha(\theta_k(0)-\theta_i(0)) + F(\theta_k(0)-\theta_i(0))\Big]. \end{aligned} \tag{IV.7}$$

We set a natural frequency ν_i depending on the initial data:

$$\nu_i := \dot{\theta}_i(0) - \frac{\kappa}{N}\sum_{k=1}^N \Big[\alpha(\theta_k(0)-\theta_i(0)) + F(\theta_k(0)-\theta_i(0))\Big]. \tag{IV.8}$$

Finally, we combine (IV.7) and (IV.8) to get the Kuramoto type model:

$$\dot{\theta}_i(t) = \nu_i + \frac{\kappa}{N}\sum_{k=1}^N \Big[\alpha(\theta_k(t)-\theta_i(t)) + F(\theta_k(t)-\theta_i(t))\Big].$$

Remark IV.2. *If we drop the assumption $\psi \ge 0$ and set $\psi(\theta) := \cos\theta$, then we have*

$$\alpha = 0 \quad \textit{and} \quad F(\theta) \equiv \sin\theta.$$

Thus, we obtain the Kuramoto model [1]:

$$\dot{\theta}_i = \nu_i + \frac{\kappa}{N}\sum_{k=1}^N \sin(\theta_k-\theta_i).$$

B. Emergent dynamics

In this subsection, we consider emergent dynamics of the second-order model:

$$\ddot{\theta}_i = \frac{\kappa}{N}\sum_{k=1}^N \widetilde{\psi}(\theta_k-\theta_i)(\dot{\theta}_k-\dot{\theta}_i).$$

To derive asymptotic behavior of (IV.3), we treat (IV.3) analogously to one-dimensional Euclidean Cucker-Smale model. In the sequel, we assume that

$$\tilde{\psi}(\theta) > 0 \quad \text{on} \quad \theta\in(-\delta,\delta),$$

and introduce phase and frequency diameters:

$$\begin{aligned} \mathcal{D}(\Theta(t)) &:= \max_{1\le i,j\le N} |\theta_i(t)-\theta_j(t)|, \\ \mathcal{D}(\dot{\Theta}(t)) &:= \max_{1\le i,j\le N} |\dot{\theta}_i(t)-\dot{\theta}_j(t)|. \end{aligned}$$

Theorem IV.1. *Suppose δ, θ^∞, initial data and coupling strength satisfy*

$$\begin{aligned} &\delta>0, \quad \mathcal{D}(\Theta^0)<\delta, \quad \theta^\infty\in(\mathcal{D}(\Theta^0),\delta), \\ &\mathcal{D}(\Theta(0)) + \frac{\mathcal{D}(\dot{\Theta}(0))}{\kappa\tilde{\psi}(\theta^\infty)} \le \theta^\infty. \end{aligned}$$

Then, complete synchronization emerges exponentially, i.e.

$$\begin{aligned} &\mathcal{D}(\dot{\Theta}(t)) \le e^{-\kappa\tilde{\psi}(\theta^\infty)t}\mathcal{D}(\dot{\Theta}(0)) \quad \forall\ t\ge 0, \\ &\textit{and} \quad \sup_{0\le t<\infty} \mathcal{D}(\Theta(t)) \le \theta^\infty. \end{aligned}$$

Proof. Since the case $\mathcal{D}(\dot{\Theta}(0))=0$ is trivial, we assume that $\mathcal{D}(\dot{\Theta}(0))>0$. Define the set

$$S := \{T>0 : \mathcal{D}(\Theta(t))<\theta^\infty \quad \text{for} \quad 0\le t\le T\}.$$

By the continuity of $\mathcal{D}(\Theta(t))$, we have $S\neq\emptyset$. By contradiction argument, we will show

$$\sup S = \infty.$$

Suppose $\sup S<\infty$, then we set

$$T^* := \sup S < \infty.$$

For each $0\le t\le T^*$, we choose maximal indices i_t, j_t satisfying

$$\mathcal{D}(\dot{\Theta}(t)) = \dot{\theta}_{i_t}(t) - \dot{\theta}_{j_t}(t),$$

i.e.,

$$\dot\theta_{i_t}(t) = \max_{1\le i\le N} \dot\theta_i(t) \quad \text{and} \quad \dot\theta_{j_t}(t) = \min_{1\le j\le N} \dot\theta_j(t).$$

Note that $\mathcal{D}(\dot\Theta(t))$ is Lipschitz continuous on this interval, since

$$\begin{aligned}\mathcal{D}(\dot\Theta(t)) - \mathcal{D}(\dot\Theta(s)) &= (\dot\theta_{i_t}(t) - \dot\theta_{j_t}(t)) - (\dot\theta_{i_s}(s) - \dot\theta_{j_s}(s)) \\ &= (\dot\theta_{i_t}(t) - \dot\theta_{j_t}(t)) - (\max_{1\le i\le N} \dot\theta_i(s) - \min_{1\le j\le N} \dot\theta_j(s)) \\ &\le (\dot\theta_{i_t}(t) - \dot\theta_{j_t}(t)) - (\dot\theta_{i_t}(s) - \dot\theta_{j_t}(s)) \\ &\le |\dot\theta_{i_t}(t) - \dot\theta_{i_t}(s)| + |\dot\theta_{j_t}(t) - \dot\theta_{j_t}(s)| \\ &\le |t-s| \left(\sup_{0\le r\le T^*} |\ddot\theta_{i_t}(r)| + \sup_{0\le r\le T^*} |\ddot\theta_{i_s}(r)| \right).\end{aligned}$$

Then for a.e. $t \in [0, T^*]$, we have

$$\begin{aligned}&\frac{d}{dt}\mathcal{D}(\dot\Theta(t)) = \ddot\theta_{i_t}(t) - \ddot\theta_{j_t}(t) \\ &= \frac{\kappa}{N}\sum_{k=1}^N \widetilde\psi(\theta_k - \theta_{i_t})(\dot\theta_k - \dot\theta_{i_t}) \\ &- \frac{\kappa}{N}\sum_{k=1}^N \widetilde\psi(\theta_k - \theta_{j_t})(\dot\theta_k - \dot\theta_{j_t}) \\ &\le \frac{\kappa}{N}\sum_{k=1}^N \widetilde\psi(\theta^\infty)(\dot\theta_k - \dot\theta_{i_t}) - \frac{\kappa}{N}\sum_{k=1}^N \widetilde\psi(\theta^\infty)(\dot\theta_k - \dot\theta_{j_t}) \\ &= -\kappa\tilde\psi(\theta^\infty)\mathcal{D}(\dot\Theta(t)).\end{aligned}$$

Hence, for all $0 \le t \le T^*$, we have

$$\mathcal{D}(\dot\Theta(t))) \le \mathcal{D}(\dot\Theta(0))e^{-\kappa\tilde\psi(\theta^\infty)t}. \tag{IV.9}$$

We use the above inequality to deduce the following:

$$\begin{aligned}\mathcal{D}(\Theta(T^*)) &= \mathcal{D}(\Theta(0)) + \int_0^{T^*} \mathcal{D}(\dot\Theta(t))dt \\ &\le \mathcal{D}(\Theta(0)) + \int_0^{T^*} D(\dot\Theta(0))e^{-\kappa\tilde\psi(\theta^\infty)t}dt \\ &< \mathcal{D}(\Theta(0)) + \frac{\mathcal{D}(\dot\Theta(0))}{\kappa\tilde\psi(\theta^\infty)} \le \theta^\infty.\end{aligned}$$

On the other hand, by definition of S, we have

$$\mathcal{D}(\Theta(T^*)) = \theta^\infty.$$

This gives a contradiction. Therefore, we have $\sup S = \infty$, and by (IV.9) the proof is complete. □

V. Cucker-Smale particles on the Poincaré half-plane

In this section, we present the CS model on the Poincaré half plane $\mathbb{H}$ [34] and study its emergent dynamics.

A. The reduction of the CS model to the Poincaré half-plane

We consider the Poincaré half-plane as part of $\mathbb{R}^2$. Using complex cartesian coordinates, we write

$$\boldsymbol{x}_i(t) = z_i(t) = x_i(t) + \mathrm{i}y_i(t), \ \boldsymbol{v}_i(t) = w_i(t) = u_i(t) + \mathrm{i}v_i(t).$$

Then, the CS model in terms of the complex-valued state (z_i, w_i) reduces to

$$\begin{cases} \dfrac{dz_i}{dt} = w_i, \quad t>0, \quad i = 1, \cdots, N, \\ \dfrac{Dw_i}{dt} = \dfrac{\kappa}{N}\displaystyle\sum_{k=1}^N \psi_{ik}(t)(P_{ik}w_k - w_i), \\ z_i(0) = z_i^0, \quad w_i(0) = w_i^0. \end{cases} \tag{V.1}$$

The geodesic distance in terms of z_k and z_i is $2\log\frac{|z_k - z_i| + |z_k - \overline{z}_i|}{2\sqrt{y_i y_k}}$. Therefore, the communication weight function ψ_{ik} is given by

$$\psi_{ik} = \psi(\mathrm{dist}(z_i, z_k)) = \psi\Big(2\log\frac{|z_k - z_i| + |z_k - \overline{z}_i|}{2\sqrt{y_i y_k}}\Big).$$

Lemma V.1. *Let (z_i, w_i) be a global smooth solution to* (V.1). *Then, one has*

(i) $\dfrac{Dw_i}{dt} = \dot w_i + \dfrac{\mathrm{i}}{y_i}w_i^2.$

(ii) $$P_{ik}w_k = -\frac{x_i - x_k + \mathrm{i}y_i + \mathrm{i}y_k}{x_i - x_k - \mathrm{i}y_i - \mathrm{i}y_k}\frac{y_i}{y_k}w_k = -\frac{z_i - \overline{z}_k}{\overline{z}_i - z_k}\frac{y_i}{y_k}w_k. \tag{V.2}$$

Proof. We give a proof in Appendix G. □

Now, we combine (V.1) and (V.2) to get the CS model on the Poincaré half plane $\mathbb{H}$:

$$\begin{cases} \dot z_i(t) = w_i, \quad t>0, \quad 1\le i\le N, \\ \dot w_i = -\dfrac{\mathrm{i}}{y_i}w_i^2 + \dfrac{\kappa}{N}\displaystyle\sum_{k=1}^N \psi_{ik}(t)\Big(-\dfrac{z_i - \overline{z}_k}{\overline{z}_i - z_k}\dfrac{y_i}{y_k}w_k - w_i\Big), \\ z_i(0) = z_i^0, \quad w_i(0) = w_i^0. \end{cases} \tag{V.3}$$

B. Emergent dynamics

Next, we study velocity alignment estimate for the global solutions to (V.3). However, we should show a global existence of a solution. In particular, we show that z_i never hits the x-axis in finite time.

Theorem V.1. *Suppose that the initial data satisfy*

$$y_i^0 > 0 \quad \textit{for } i = 1, \cdots, N,$$

and let (z_i, w_i) be a global smooth solution to the Cauchy problem (V.3). *Then we have the global-in-time existence of the solution to* (V.3).

Proof. Suppose to the contrary, we assume that the maximal interval of existence is $[0, T)$, where $T \in (0, \infty)$. Recall that

$$\|\boldsymbol{v}_i(t)\|^2_{\boldsymbol{x}_i} \le \mathcal{E}^0, \quad i = 1, \cdots, N, \quad t \in [0, T).$$

In other words, we have

$$\left|\frac{w_i(t)}{y_i(t)}\right| \le \sqrt{\mathcal{E}^0}, \quad i = 1, \cdots, N, \quad t \in [0, T). \tag{V.4}$$

By the first equation in (V.3), we have

$$|\dot y_i(t)| = |\operatorname{Im}[w_i(t)]| \le |w_i(t)| \le y_i(t)\sqrt{\mathcal{E}^0}. \tag{V.5}$$

Thus, we have

$$y_i^0 \exp(-\sqrt{\mathcal{E}(0)}t) \le y_i(t) \le y_i^0 \exp(\sqrt{\mathcal{E}^0}t).$$

This yields

$$0 < y_i^0 \exp(-\sqrt{\mathcal{E}(0)}T) \le y_i(t) \le y_i^0 \exp(\sqrt{\mathcal{E}^0}T),\ t \in [0,T). \tag{V.6}$$

By (V.6) we have

$$|\dot{z}_i(t)| = y_i(t) \cdot \left|\frac{w_i(t)}{y_i(t)}\right| \le y_i^0 \exp(\sqrt{\mathcal{E}^0}T)\sqrt{\mathcal{E}^0}.$$

By the mean value theorem, we have

$$|z_i(t)| \le |z_i^0| + Ty_i^0 \exp(\sqrt{\mathcal{E}^0}T)\sqrt{\mathcal{E}^0}, \quad t \in [0,T). \tag{V.7}$$

It follows from (V.3) that

$$\begin{aligned}|\dot{w}_i(t)| &\le \left|\frac{w_i^2}{y_i}\right| + \frac{\kappa}{N}\sum_{k=1}^N \psi_M \Big[\Big|-\frac{z_i - \overline{z}_k}{\overline{z}_i - z_k}\frac{y_i}{y_k}w_k\Big| + |w_i|\Big] \\ &\le y_i\mathcal{E}^0 + \frac{\kappa}{N}\sum_{k=1}^N \psi_M \Big[y_i\sqrt{\mathcal{E}^0} + y_i\sqrt{\mathcal{E}^0}\Big] \\ &\le y_i^0 \exp(\sqrt{\mathcal{E}^0}T)\cdot(\mathcal{E}^0 + 2\kappa\psi_M\sqrt{\mathcal{E}^0}).\end{aligned}$$

This yields

$$\begin{aligned}|w_i(t)| &\le |w_i^0| + Ty_i^0 \exp(\sqrt{\mathcal{E}(0)}T) \\ &\quad \times (\mathcal{E}^0 + 2\kappa\psi_M\sqrt{\mathcal{E}^0}), \quad t \in [0,T).\end{aligned} \tag{V.8}$$

By (V.6), (V.7) and (V.8), we have a contradiction to the fact that T is a life-span. Hence we have a global existence of solutions. □

Theorem V.2. *Suppose the communication weight and the initial data* $\{(z_i^0, w_i^0)\}$ *satisfy*

$$\psi(r) \equiv \psi_M, \quad y_i^0 > 0, \quad i = 1, \cdots, N,$$

and let (z_i, w_i) *be a global smooth solution to the Cauchy problem* (V.3). *Then, we have*

$$\lim_{t\to\infty} \|P_{ik}\boldsymbol{v}_k - \boldsymbol{v}_i\|_{\boldsymbol{x}_i} = 0, \quad \forall i, k.$$

Proof. Note that

$$P_{ik}\boldsymbol{v}_k - \boldsymbol{v}_i = -\frac{z_i - \overline{z}_k}{\overline{z}_i - z_k}\frac{w_k}{y_k} - \frac{w_i}{y_i}.$$

By Lemma II.3, it suffices to show that the following quantity is bounded:

$$\frac{d}{dt}\left| -\frac{z_i - \overline{z}_k}{\overline{z}_i - z_k}\frac{w_k}{y_k} - \frac{w_i}{y_i}\right|^2. \tag{V.9}$$

For this, we need to show the uniform boundedness of the following quantities:

$$\left|\frac{w_i}{y_i}\right|, \quad \left|\frac{d}{dt}\left(\frac{z_i - \overline{z}_k}{\overline{z}_i - z_k}\right)\right|, \quad \left|\frac{d}{dt}\left(\frac{w_i}{y_i}\right)\right|.$$

• Uniform boundedness of $\left|\frac{w_i}{y_i}\right|$: By (V.4), one has

$$\left|\frac{w_i}{y_i}\right| \le \sqrt{\mathcal{E}^0}. \tag{V.10}$$

• Uniform boundedness of $\left|\frac{d}{dt}\left(\frac{z_i - \overline{z}_k}{\overline{z}_i - z_k}\right)\right|$: It follows from (V.4) that

$$\begin{aligned}\left|\frac{d}{dt}\left(\frac{z_i - \overline{z}_k}{\overline{z}_i - z_k}\right)\right| &= \left|\frac{w_i - \overline{w}_k}{\overline{z}_i - z_k} - \frac{(z_i - \overline{z}_k)(\overline{w}_i - w_k)}{(\overline{z}_i - z_k)^2}\right| \\ &= \left|2\mathrm{i}\frac{\mathrm{Im}[w_i(\overline{z}_i - z_k)]}{(\overline{z}_i - z_k)^2} + 2\mathrm{i}\frac{\mathrm{Im}[w_k(z_i - \overline{z}_k)]}{(\overline{z}_i - z_k)^2}\right| \\ &\le 2\left|\frac{w_i}{\overline{z}_i - z_k}\right| + 2\left|\frac{w_k}{\overline{z}_i - z_k}\right| \\ &\le 2\sqrt{\mathcal{E}^0}\left(\left|\frac{y_i}{\overline{z}_i - z_k}\right| + \left|\frac{y_k}{\overline{z}_i - z_k}\right|\right) \\ &= 2\sqrt{\mathcal{E}^0}\cdot\frac{y_i + y_k}{|\overline{z}_i - z_k|} \\ &\le 2\sqrt{\mathcal{E}^0}.\end{aligned} \tag{V.11}$$

• Uniform boundedness of $\left|\frac{d}{dt}\left(\frac{w_i}{y_i}\right)\right|$: It follows from (V.4) that

$$\begin{aligned}\frac{d}{dt}\frac{w_i}{y_i} &= \frac{\dot{w}_i}{y_i} - \frac{w_i\,\mathrm{Im}[w_i]}{y_i^2} \le \left|\frac{\dot{w}_i}{y_i}\right| + \left|\frac{w_i\,\mathrm{Im}[w_i]}{y_i^2}\right| \\ &\le \left|\frac{w_i}{y_i}\right|^2 + \frac{\kappa}{N}\sum_{k=1}^N \psi_M\Big(\left|\frac{w_k}{y_k}\right| + \left|\frac{w_i}{y_i}\right|\Big) + \left|\frac{w_i}{y_i}\right|^2 \\ &\le 2\mathcal{E}^0 + 2\kappa\psi_M\sqrt{\mathcal{E}^0}.\end{aligned} \tag{V.12}$$

Finally, we combine (V.10), (V.11) and (V.12) to see the boundedness of (V.9). □

VI. Conclusion

In this paper, we have introduced the Cucker-Smale flocking model on a Riemannian manifold $(M, \langle\cdot,\cdot\rangle_M)$. For the proposed CS model, we also provided emergent velocity alignment estimates under an a priori assumption. Our proposed model simply becomes the standard CS model in the case of Euclidean space. For definiteness, we considered three explicit Riemannian manifolds such as the unit sphere in $\mathbb{R}^3$, the unit circle and the Poincaré half plane $\mathbb{H}$. For each manifold, we provided explicit reduction model from the proposed CS model by explicitly calculating covariant derivative and parallel transport of the velocity field in terms of Euclidean coordinates of an ambient $\mathbb{R}^d$ in which we embed the manifold. We also studied emergent dynamics for the reduced models. In particular, for the unit circle, our proposed model reduces to the second-order Kuramoto type model with a periodic weight function with mean zero. Our velocity alignment estimate relies on an a priori assumption which can be removed for the two-particle system on the unit sphere and the N-particle system on the Poincaré half-plane. However, we do not have a complete general theory for the emergent dynamics of the CS model on a general Riemannian manifold at present. Thus, one of the remaining interesting issues is to derive emergent dynamics of the general CS model in terms of initial data and system parameters without assuming a priori conditions on the solution itself. We leave this for a future work.

Appendix A
Proof of Lemma III.3

(1) We use Lemma III.1 (ii) to compute P_{ij} as follows:

$$
\begin{aligned}
P_{ij} &= \langle \boldsymbol{x}_j, \boldsymbol{x}_i\rangle I + [\boldsymbol{x}_j \times \boldsymbol{x}_i]_\times + \frac{(\boldsymbol{x}_j \times \boldsymbol{x}_i)(\boldsymbol{x}_j \times \boldsymbol{x}_i)^\top}{1+\langle \boldsymbol{x}_j, \boldsymbol{x}_i\rangle} \\
&= \langle \boldsymbol{x}_j, \boldsymbol{x}_i\rangle(\mathbf{p}_{ij}\mathbf{p}_{ij}^\top + \mathbf{q}_{ij}\mathbf{q}_{ij}^\top + \mathbf{r}_{ij}\mathbf{r}_{ij}^\top) + \boldsymbol{x}_i\boldsymbol{x}_j^\top - \boldsymbol{x}_j\boldsymbol{x}_i^\top \\
&+ \frac{\|\boldsymbol{x}_j \times \boldsymbol{x}_i\|^2}{1+\langle \boldsymbol{x}_j, \boldsymbol{x}_i\rangle}\mathbf{p}_{ij}\mathbf{p}_{ij}^\top \\
&= \langle \boldsymbol{x}_j, \boldsymbol{x}_i\rangle(\mathbf{p}_{ij}\mathbf{p}_{ij}^\top + \mathbf{q}_{ij}\mathbf{q}_{ij}^\top + \mathbf{r}_{ij}\mathbf{r}_{ij}^\top) \\
&+ \Big(\frac{\boldsymbol{x}_i+\boldsymbol{x}_j}{2} + \frac{\boldsymbol{x}_i-\boldsymbol{x}_j}{2}\Big)\Big(\frac{\boldsymbol{x}_i+\boldsymbol{x}_j}{2} - \frac{\boldsymbol{x}_i-\boldsymbol{x}_j}{2}\Big)^\top \\
&- \Big(\frac{\boldsymbol{x}_i+\boldsymbol{x}_j}{2} - \frac{\boldsymbol{x}_i-\boldsymbol{x}_j}{2}\Big)\Big(\frac{\boldsymbol{x}_i+\boldsymbol{x}_j}{2} + \frac{\boldsymbol{x}_i-\boldsymbol{x}_j}{2}\Big)^\top \\
&+ (1-\langle \boldsymbol{x}_j, \boldsymbol{x}_i\rangle)\mathbf{p}_{ij}\mathbf{p}_{ij}^\top \\
&= \langle \boldsymbol{x}_j, \boldsymbol{x}_i\rangle(\mathbf{p}_{ij}\mathbf{p}_{ij}^\top + \mathbf{q}_{ij}\mathbf{q}_{ij}^\top + \mathbf{r}_{ij}\mathbf{r}_{ij}^\top) + (1-\langle \boldsymbol{x}_j, \boldsymbol{x}_i\rangle)\mathbf{p}_{ij}\mathbf{p}_{ij}^\top \\
&+ 2\Big(\frac{\boldsymbol{x}_i-\boldsymbol{x}_j}{2}\Big)\Big(\frac{\boldsymbol{x}_i+\boldsymbol{x}_j}{2}\Big)^\top - 2\Big(\frac{\boldsymbol{x}_i+\boldsymbol{x}_j}{2}\Big)\Big(\frac{\boldsymbol{x}_i-\boldsymbol{x}_j}{2}\Big)^\top \\
&= \mathbf{p}_{ij}\mathbf{p}_{ij}^\top + \langle \boldsymbol{x}_j, \boldsymbol{x}_i\rangle(\mathbf{q}_{ij}\mathbf{q}_{ij}^\top + \mathbf{r}_{ij}\mathbf{r}_{ij}^\top) \\
&+ \frac{\|\boldsymbol{x}_i+\boldsymbol{x}_j\|\|\boldsymbol{x}_i-\boldsymbol{x}_j\|}{2}\mathbf{r}_{ij}\mathbf{q}_{ij}^\top - \frac{\|\boldsymbol{x}_i+\boldsymbol{x}_j\|\|\boldsymbol{x}_i-\boldsymbol{x}_j\|}{2}\mathbf{q}_{ij}\mathbf{r}_{ij}^\top.
\end{aligned}
$$

(2) For the first equality, we use Lemma III.1 (ii) to see that

$$
\begin{aligned}
P_{ij}\boldsymbol{x}_j &= \left(\langle \boldsymbol{x}_j, \boldsymbol{x}_i\rangle I + [\boldsymbol{x}_j \times \boldsymbol{x}_i]_\times + \frac{(\boldsymbol{x}_j \times \boldsymbol{x}_i)(\boldsymbol{x}_j \times \boldsymbol{x}_i)^\top}{1+\langle \boldsymbol{x}_j, \boldsymbol{x}_i\rangle}\right)\boldsymbol{x}_j \\
&= \langle \boldsymbol{x}_j, \boldsymbol{x}_i\rangle\boldsymbol{x}_j + (\boldsymbol{x}_j \times \boldsymbol{x}_i) \times \boldsymbol{x}_j = \boldsymbol{x}_i.
\end{aligned}
$$

The second equality is a result of the following property of an arbitrary rotation matrix $R \in SO(3)$: for $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^3$, we have

$$
R\boldsymbol{x} \times R\boldsymbol{y} = R(\boldsymbol{x} \times \boldsymbol{y}).
$$

Appendix B
Proof of Lemma III.4

(i) Recall that

$$
\mathbf{p}_{ij} := \frac{\boldsymbol{x}_i \times \boldsymbol{x}_j}{\|\boldsymbol{x}_i \times \boldsymbol{x}_j\|}. \tag{B.1}
$$

First, we use the a priori condition $\langle \boldsymbol{x}_j, \boldsymbol{v}_j\rangle = 0$ and the vector triple identity:

$$
A \times (B \times C) = \langle A, C\rangle B - \langle A, B\rangle C
$$

to get

$$
\begin{aligned}
\boldsymbol{x}_i \times \boldsymbol{v}_j &= \boldsymbol{x}_i \times \big(-\boldsymbol{x}_j \times (\boldsymbol{x}_j \times \boldsymbol{v}_j)\big) \\
&= -\langle \boldsymbol{x}_i, \boldsymbol{x}_j \times \boldsymbol{v}_j\rangle\boldsymbol{x}_j + \langle \boldsymbol{x}_i, \boldsymbol{x}_j\rangle(\boldsymbol{x}_j \times \boldsymbol{v}_j), \\
\boldsymbol{x}_j \times \boldsymbol{v}_i &= \boldsymbol{x}_j \times \big(-\boldsymbol{x}_i \times (\boldsymbol{x}_i \times \boldsymbol{v}_i)\big) \\
&= -\langle \boldsymbol{x}_j, \boldsymbol{x}_i \times \boldsymbol{v}_i\rangle\boldsymbol{x}_i + \langle \boldsymbol{x}_j, \boldsymbol{x}_i\rangle(\boldsymbol{x}_i \times \boldsymbol{v}_i).
\end{aligned} \tag{B.2}
$$

Then, we use (B.2) to obtain

$$
\begin{aligned}
&\frac{d}{dt}(\boldsymbol{x}_i \times \boldsymbol{x}_j) = \boldsymbol{x}_i \times \boldsymbol{v}_j + \boldsymbol{v}_i \times \boldsymbol{x}_j \\
&= -\langle \boldsymbol{x}_i, \boldsymbol{x}_j\rangle(\boldsymbol{x}_i \times \boldsymbol{v}_i - \boldsymbol{x}_j \times \boldsymbol{v}_j) \\
&- \langle \boldsymbol{x}_i, \boldsymbol{x}_j \times \boldsymbol{v}_j\rangle\boldsymbol{x}_j + \langle \boldsymbol{x}_j, \boldsymbol{x}_i \times \boldsymbol{v}_i\rangle\boldsymbol{x}_i \\
&= -\langle \boldsymbol{x}_i, \boldsymbol{x}_j\rangle(\boldsymbol{x}_i \times \boldsymbol{v}_i - \boldsymbol{x}_j \times \boldsymbol{v}_j) \\
&+ \langle \boldsymbol{x}_i, \boldsymbol{x}_i \times \boldsymbol{v}_i - \boldsymbol{x}_j \times \boldsymbol{v}_j\rangle\boldsymbol{x}_j + \langle \boldsymbol{x}_j, \boldsymbol{x}_i \times \boldsymbol{v}_i - \boldsymbol{x}_j \times \boldsymbol{v}_j\rangle\boldsymbol{x}_i \\
&= -\langle \boldsymbol{x}_i, \boldsymbol{x}_j\rangle(\boldsymbol{x}_i \times \boldsymbol{v}_i - \boldsymbol{x}_j \times \boldsymbol{v}_j) \\
&+ \langle \boldsymbol{x}_i, \boldsymbol{x}_i \times \boldsymbol{v}_i - \boldsymbol{x}_j \times \boldsymbol{v}_j\rangle\left(\frac{\boldsymbol{x}_i+\boldsymbol{x}_j}{2} - \frac{\boldsymbol{x}_i-\boldsymbol{x}_j}{2}\right) \\
&+ \langle \boldsymbol{x}_j, \boldsymbol{x}_i \times \boldsymbol{v}_i - \boldsymbol{x}_j \times \boldsymbol{v}_j\rangle\left(\frac{\boldsymbol{x}_i+\boldsymbol{x}_j}{2} + \frac{\boldsymbol{x}_i-\boldsymbol{x}_j}{2}\right) \\
&= -\langle \boldsymbol{x}_i, \boldsymbol{x}_j\rangle(\boldsymbol{x}_i \times \boldsymbol{v}_i - \boldsymbol{x}_j \times \boldsymbol{v}_j) \\
&+ \langle \boldsymbol{x}_i \times \boldsymbol{v}_i - \boldsymbol{x}_j \times \boldsymbol{v}_j, \boldsymbol{x}_i + \boldsymbol{x}_j\rangle\frac{\boldsymbol{x}_i+\boldsymbol{x}_j}{2} \\
&- \langle \boldsymbol{x}_i \times \boldsymbol{v}_i - \boldsymbol{x}_j \times \boldsymbol{v}_j, \boldsymbol{x}_i - \boldsymbol{x}_j\rangle\frac{\boldsymbol{x}_i-\boldsymbol{x}_j}{2} \\
&= -\langle \boldsymbol{x}_i, \boldsymbol{x}_j\rangle\langle \boldsymbol{x}_i \times \boldsymbol{v}_i - \boldsymbol{x}_j \times \boldsymbol{v}_j, \mathbf{p}_{ij}\rangle\mathbf{p}_{ij} \\
&- \langle \boldsymbol{x}_i, \boldsymbol{x}_j\rangle\langle \boldsymbol{x}_i \times \boldsymbol{v}_i - \boldsymbol{x}_j \times \boldsymbol{v}_j, \mathbf{q}_{ij}\rangle\mathbf{q}_{ij} \\
&- \langle \boldsymbol{x}_i, \boldsymbol{x}_j\rangle\langle \boldsymbol{x}_i \times \boldsymbol{v}_i - \boldsymbol{x}_j \times \boldsymbol{v}_j, \mathbf{r}_{ij}\rangle\mathbf{r}_{ij} \\
&+ (1+\langle \boldsymbol{x}_i, \boldsymbol{x}_j\rangle)\langle \boldsymbol{x}_i \times \boldsymbol{v}_i - \boldsymbol{x}_j \times \boldsymbol{v}_j, \mathbf{q}_{ij}\rangle\mathbf{q}_{ij} \\
&- (1-\langle \boldsymbol{x}_i, \boldsymbol{x}_j\rangle)\langle \boldsymbol{x}_i \times \boldsymbol{v}_i - \boldsymbol{x}_j \times \boldsymbol{v}_j, \mathbf{r}_{ij}\rangle\mathbf{r}_{ij} \\
&= -\langle \boldsymbol{x}_i, \boldsymbol{x}_j\rangle\langle \boldsymbol{x}_i \times \boldsymbol{v}_i - \boldsymbol{x}_j \times \boldsymbol{v}_j, \mathbf{p}_{ij}\rangle\mathbf{p}_{ij} \\
&+ \langle \boldsymbol{x}_i \times \boldsymbol{v}_i - \boldsymbol{x}_j \times \boldsymbol{v}_j, \mathbf{q}_{ij}\rangle\mathbf{q}_{ij} \\
&- \langle \boldsymbol{x}_i \times \boldsymbol{v}_i - \boldsymbol{x}_j \times \boldsymbol{v}_j, \mathbf{r}_{ij}\rangle\mathbf{r}_{ij}.
\end{aligned} \tag{B.3}
$$

Here we used

$$
\|\boldsymbol{x}_i \pm \boldsymbol{x}_j\|^2 = \|\boldsymbol{x}_i\|^2 + \|\boldsymbol{x}_j\|^2 \pm 2\langle \boldsymbol{x}_i, \boldsymbol{x}_j\rangle = 2(1 \pm \langle \boldsymbol{x}_i, \boldsymbol{x}_j\rangle).
$$

Hence, it follows from (B.1) and (B.3) that

$$
\begin{aligned}
&\frac{d}{dt}\left(\frac{\boldsymbol{x}_i \times \boldsymbol{x}_j}{\|\boldsymbol{x}_i \times \boldsymbol{x}_j\|}\right) \\
&= \frac{1}{\|\boldsymbol{x}_i \times \boldsymbol{x}_j\|}\frac{d}{dt}(\boldsymbol{x}_i \times \boldsymbol{x}_j) \\
&- \left\langle \boldsymbol{x}_i \times \boldsymbol{x}_j, \frac{d}{dt}(\boldsymbol{x}_i \times \boldsymbol{x}_j)\right\rangle \frac{\boldsymbol{x}_i \times \boldsymbol{x}_j}{\|\boldsymbol{x}_i \times \boldsymbol{x}_j\|^3} \\
&= \frac{1}{\|\boldsymbol{x}_i \times \boldsymbol{x}_j\|}\left(\frac{d}{dt}(\boldsymbol{x}_i \times \boldsymbol{x}_j) - \left\langle \mathbf{p}_{ij}, \frac{d}{dt}(\boldsymbol{x}_i \times \boldsymbol{x}_j)\right\rangle \mathbf{p}_{ij}\right) \\
&= \frac{1}{\|\boldsymbol{x}_i \times \boldsymbol{x}_j\|}\big(\langle \boldsymbol{x}_i \times \boldsymbol{v}_i - \boldsymbol{x}_j \times \boldsymbol{v}_j, \mathbf{q}_{ij}\rangle\mathbf{q}_{ij} \\
&- \langle \boldsymbol{x}_i \times \boldsymbol{v}_i - \boldsymbol{x}_j \times \boldsymbol{v}_j, \mathbf{r}_{ij}\rangle\mathbf{r}_{ij}\big),
\end{aligned}
$$

where we used the simplified notation of (III.6).

(ii) Note that

$$
\begin{aligned}
\langle \boldsymbol{x}_i, \boldsymbol{v}_j\rangle &= \big\langle \boldsymbol{x}_i, -\boldsymbol{x}_j \times (\boldsymbol{x}_j \times \boldsymbol{v}_j)\big\rangle = -\big\langle \boldsymbol{x}_i \times \boldsymbol{x}_j, \boldsymbol{x}_j \times \boldsymbol{v}_j\big\rangle, \\
\langle \boldsymbol{x}_j, \boldsymbol{v}_i\rangle &= \big\langle \boldsymbol{x}_j, -\boldsymbol{x}_i \times (\boldsymbol{x}_i \times \boldsymbol{v}_i)\big\rangle = -\big\langle \boldsymbol{x}_j \times \boldsymbol{x}_i, \boldsymbol{x}_i \times \boldsymbol{v}_i\big\rangle.
\end{aligned}
$$

Similar to the estimate in (i), we first express $\boldsymbol{v}_i + \boldsymbol{v}_j$ as a combination of $\mathbf{p}_{ij}$, $\mathbf{q}_{ij}$ and $\mathbf{r}_{ij}$:

$$
\begin{aligned}
&\boldsymbol{v}_i + \boldsymbol{v}_j \\
&= \langle \boldsymbol{v}_i + \boldsymbol{v}_j, \mathbf{p}_{ij}\rangle\mathbf{p}_{ij} + \langle \boldsymbol{v}_i + \boldsymbol{v}_j, \mathbf{q}_{ij}\rangle\mathbf{q}_{ij} + \langle \boldsymbol{v}_i + \boldsymbol{v}_j, \mathbf{r}_{ij}\rangle\mathbf{r}_{ij} \\
&= \frac{\langle \boldsymbol{v}_i, \boldsymbol{x}_i \times \boldsymbol{x}_j\rangle + \langle \boldsymbol{v}_j, \boldsymbol{x}_i \times \boldsymbol{x}_j\rangle}{\|\boldsymbol{x}_i \times \boldsymbol{x}_j\|}\mathbf{p}_{ij} \\
&+ \frac{\langle \boldsymbol{v}_i, \boldsymbol{x}_j\rangle + \langle \boldsymbol{v}_j, \boldsymbol{x}_i\rangle}{\|\boldsymbol{x}_i + \boldsymbol{x}_j\|}\mathbf{q}_{ij} + \frac{\langle \boldsymbol{x}_i, \boldsymbol{v}_j\rangle - \langle \boldsymbol{x}_j, \boldsymbol{v}_i\rangle}{\|\boldsymbol{x}_i - \boldsymbol{x}_j\|}\mathbf{r}_{ij} \\
&= \frac{-\langle \boldsymbol{x}_i \times \boldsymbol{v}_i, \boldsymbol{x}_j\rangle + \langle \boldsymbol{x}_j \times \boldsymbol{v}_j, \boldsymbol{x}_i\rangle}{\|\boldsymbol{x}_i \times \boldsymbol{x}_j\|}\mathbf{p}_{ij}
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{-\langle \boldsymbol{x}_j\times\boldsymbol{x}_i,\boldsymbol{x}_i\times\boldsymbol{v}_i\rangle-\langle \boldsymbol{x}_i\times\boldsymbol{x}_j,\boldsymbol{x}_j\times\boldsymbol{v}_j\rangle}{\|\boldsymbol{x}_i+\boldsymbol{x}_j\|}\mathbf{q}_{ij}\\
&+\frac{\langle \boldsymbol{x}_i,\boldsymbol{v}_j\rangle-\langle \boldsymbol{x}_j,\boldsymbol{v}_i\rangle}{\|\boldsymbol{x}_i-\boldsymbol{x}_j\|}\mathbf{r}_{ij}\\
&=-\frac{\langle \boldsymbol{x}_i\times\boldsymbol{v}_i-\boldsymbol{x}_j\times\boldsymbol{v}_j,\boldsymbol{x}_i+\boldsymbol{x}_j\rangle}{\|\boldsymbol{x}_i\times\boldsymbol{x}_j\|}\mathbf{p}_{ij}\\
&+\frac{\langle \boldsymbol{x}_i\times\boldsymbol{v}_i-\boldsymbol{x}_j\times\boldsymbol{v}_j,\boldsymbol{x}_i\times\boldsymbol{x}_j\rangle}{\|\boldsymbol{x}_i+\boldsymbol{x}_j\|}\mathbf{q}_{ij}+\frac{\langle \boldsymbol{x}_i,\boldsymbol{v}_j\rangle-\langle \boldsymbol{x}_j,\boldsymbol{v}_i\rangle}{\|\boldsymbol{x}_i-\boldsymbol{x}_j\|}\mathbf{r}_{ij}\\
&=-\frac{\|\boldsymbol{x}_i+\boldsymbol{x}_j\|\langle \boldsymbol{x}_i\times\boldsymbol{v}_i-\boldsymbol{x}_j\times\boldsymbol{v}_j,\mathbf{q}_{ij}\rangle}{\|\boldsymbol{x}_i\times\boldsymbol{x}_j\|}\mathbf{p}_{ij}\\
&+\frac{\|\boldsymbol{x}_i\times\boldsymbol{x}_j\|\langle \boldsymbol{x}_i\times\boldsymbol{v}_i-\boldsymbol{x}_j\times\boldsymbol{v}_j,\mathbf{p}_{ij}\rangle}{\|\boldsymbol{x}_i+\boldsymbol{x}_j\|}\mathbf{q}_{ij}\\
&+\frac{\langle \boldsymbol{x}_i,\boldsymbol{v}_j\rangle-\langle \boldsymbol{x}_j,\boldsymbol{v}_i\rangle}{\|\boldsymbol{x}_i-\boldsymbol{x}_j\|}\mathbf{r}_{ij}.
\end{aligned}
$$

Then, it follows from the defining relation of $\mathbf{q}_{ij}$ that

$$
\begin{aligned}
&\frac{d}{dt}\left(\frac{\boldsymbol{x}_i+\boldsymbol{x}_j}{\|\boldsymbol{x}_i+\boldsymbol{x}_j\|}\right)\\
&=\frac{1}{\|\boldsymbol{x}_i+\boldsymbol{x}_j\|}(\boldsymbol{v}_i+\boldsymbol{v}_j)-\langle \boldsymbol{x}_i+\boldsymbol{x}_j,\boldsymbol{v}_i+\boldsymbol{v}_j\rangle\frac{\boldsymbol{x}_i+\boldsymbol{x}_j}{\|\boldsymbol{x}_i+\boldsymbol{x}_j\|^3}\\
&=\frac{1}{\|\boldsymbol{x}_i+\boldsymbol{x}_j\|}\left(\boldsymbol{v}_i+\boldsymbol{v}_j-\langle \mathbf{q}_{ij},\boldsymbol{v}_i+\boldsymbol{v}_j\rangle\mathbf{q}_{ij}\right)\\
&=\frac{1}{\|\boldsymbol{x}_i+\boldsymbol{x}_j\|}\Bigg(-\frac{\|\boldsymbol{x}_i+\boldsymbol{x}_j\|\langle \boldsymbol{x}_i\times\boldsymbol{v}_i-\boldsymbol{x}_j\times\boldsymbol{v}_j,\mathbf{q}_{ij}\rangle}{\|\boldsymbol{x}_i\times\boldsymbol{x}_j\|}\mathbf{p}_{ij}\\
&\qquad+\frac{\langle \boldsymbol{x}_i,\boldsymbol{v}_j\rangle-\langle \boldsymbol{x}_j,\boldsymbol{v}_i\rangle}{\|\boldsymbol{x}_i-\boldsymbol{x}_j\|}\mathbf{r}_{ij}\Bigg)\\
&=-\frac{\langle \boldsymbol{x}_i\times\boldsymbol{v}_i-\boldsymbol{x}_j\times\boldsymbol{v}_j,\mathbf{q}_{ij}\rangle}{\|\boldsymbol{x}_i\times\boldsymbol{x}_j\|}\mathbf{p}_{ij}+\frac{\langle \boldsymbol{x}_i,\boldsymbol{v}_j\rangle-\langle \boldsymbol{x}_j,\boldsymbol{v}_i\rangle}{2\|\boldsymbol{x}_i\times\boldsymbol{x}_j\|}\mathbf{r}_{ij}.
\end{aligned}
$$

(iii) Analogously, we express the relative velocities $\boldsymbol{v}_i-\boldsymbol{v}_j$ in terms of $\mathbf{p}_{ij}$, $\mathbf{q}_{ij}$ and $\mathbf{r}_{ij}$ as follows.

$$
\begin{aligned}
&\boldsymbol{v}_i-\boldsymbol{v}_j\\
&=\langle \boldsymbol{v}_i-\boldsymbol{v}_j,\mathbf{p}_{ij}\rangle\mathbf{p}_{ij}+\langle \boldsymbol{v}_i-\boldsymbol{v}_j,\mathbf{q}_{ij}\rangle\mathbf{q}_{ij}+\langle \boldsymbol{v}_i-\boldsymbol{v}_j,\mathbf{r}_{ij}\rangle\mathbf{r}_{ij}\\
&=\frac{\|\boldsymbol{x}_i-\boldsymbol{x}_j\|\langle \boldsymbol{x}_i\times\boldsymbol{v}_i-\boldsymbol{x}_j\times\boldsymbol{v}_j,\mathbf{r}_{ij}\rangle}{\|\boldsymbol{x}_i\times\boldsymbol{x}_j\|}\mathbf{p}_{ij}\\
&\quad-\frac{\langle \boldsymbol{x}_i,\boldsymbol{v}_j\rangle-\langle \boldsymbol{x}_j,\boldsymbol{v}_i\rangle}{\|\boldsymbol{x}_i+\boldsymbol{x}_j\|}\mathbf{q}_{ij}\\
&\quad-\frac{\|\boldsymbol{x}_i\times\boldsymbol{x}_j\|\langle \boldsymbol{x}_i\times\boldsymbol{v}_i-\boldsymbol{x}_j\times\boldsymbol{v}_j,\mathbf{p}_{ij}\rangle}{\|\boldsymbol{x}_i-\boldsymbol{x}_j\|}\mathbf{r}_{ij}.
\end{aligned}
$$

By the defining relation of $\mathbf{r}_{ij}$, we have

$$
\begin{aligned}
&\frac{d}{dt}\left(\frac{\boldsymbol{x}_i-\boldsymbol{x}_j}{\|\boldsymbol{x}_i-\boldsymbol{x}_j\|}\right)\\
&\quad=\frac{1}{\|\boldsymbol{x}_i-\boldsymbol{x}_j\|}(\boldsymbol{v}_i-\boldsymbol{v}_j)-\langle \boldsymbol{x}_i-\boldsymbol{x}_j,\boldsymbol{v}_i-\boldsymbol{v}_j\rangle\frac{\boldsymbol{x}_i-\boldsymbol{x}_j}{\|\boldsymbol{x}_i-\boldsymbol{x}_j\|^3}\\
&\quad=\frac{1}{\|\boldsymbol{x}_i-\boldsymbol{x}_j\|}\left(\boldsymbol{v}_i-\boldsymbol{v}_j-\langle \mathbf{r}_{ij},\boldsymbol{v}_i-\boldsymbol{v}_j\rangle\mathbf{r}_{ij}\right)\\
&\quad=\frac{1}{\|\boldsymbol{x}_i-\boldsymbol{x}_j\|}\Bigg(\frac{\|\boldsymbol{x}_i-\boldsymbol{x}_j\|\langle \boldsymbol{x}_i\times\boldsymbol{v}_i-\boldsymbol{x}_j\times\boldsymbol{v}_j,\mathbf{r}_{ij}\rangle}{\|\boldsymbol{x}_i\times\boldsymbol{x}_j\|}\mathbf{p}_{ij}\\
&\qquad-\frac{\langle \boldsymbol{x}_i,\boldsymbol{v}_j\rangle-\langle \boldsymbol{x}_j,\boldsymbol{v}_i\rangle}{\|\boldsymbol{x}_i+\boldsymbol{x}_j\|}\mathbf{q}_{ij}\Bigg)\\
&\quad=\frac{\langle \boldsymbol{x}_i\times\boldsymbol{v}_i-\boldsymbol{x}_j\times\boldsymbol{v}_j,\mathbf{r}_{ij}\rangle}{\|\boldsymbol{x}_i\times\boldsymbol{x}_j\|}\mathbf{p}_{ij}-\frac{\langle \boldsymbol{x}_i,\boldsymbol{v}_j\rangle-\langle \boldsymbol{x}_j,\boldsymbol{v}_i\rangle}{2\|\boldsymbol{x}_i\times\boldsymbol{x}_j\|}\mathbf{q}_{ij}.
\end{aligned}
$$

Appendix C
Proof of Proposition III.1

It follows from Lemma III.3 and Lemma III.4 that

$$
\begin{aligned}
&\boldsymbol{x}_i\times(P_{ij}\boldsymbol{v}_j-\boldsymbol{v}_i)=P_{ij}\boldsymbol{x}_j\times P_{ij}\boldsymbol{v}_j-\boldsymbol{x}_i\times\boldsymbol{v}_i\\
&=P_{ij}(\boldsymbol{x}_j\times\boldsymbol{v}_j)-\boldsymbol{x}_i\times\boldsymbol{v}_i\\
&=\langle \boldsymbol{x}_j\times\boldsymbol{v}_j,\mathbf{p}_{ij}\rangle\mathbf{p}_{ij}\\
&+\langle \boldsymbol{x}_j\times\boldsymbol{v}_j,\mathbf{q}_{ij}\rangle\left(\langle \boldsymbol{x}_i,\boldsymbol{x}_j\rangle\mathbf{q}_{ij}+\frac{\|\boldsymbol{x}_i+\boldsymbol{x}_j\|\|\boldsymbol{x}_i-\boldsymbol{x}_j\|}{2}\mathbf{r}_{ij}\right)\\
&+\langle \boldsymbol{x}_j\times\boldsymbol{v}_j,\mathbf{r}_{ij}\rangle\left(\langle \boldsymbol{x}_i,\boldsymbol{x}_j\rangle\mathbf{r}_{ij}-\frac{\|\boldsymbol{x}_i+\boldsymbol{x}_j\|\|\boldsymbol{x}_i-\boldsymbol{x}_j\|}{2}\mathbf{q}_{ij}\right)\\
&-\langle \boldsymbol{x}_i\times\boldsymbol{v}_i,\mathbf{p}_{ij}\rangle\mathbf{p}_{ij}-\langle \boldsymbol{x}_i\times\boldsymbol{v}_i,\mathbf{q}_{ij}\rangle\mathbf{q}_{ij}-\langle \boldsymbol{x}_i\times\boldsymbol{v}_i,\mathbf{r}_{ij}\rangle\mathbf{r}_{ij}\\
&=\langle \boldsymbol{x}_j\times\boldsymbol{v}_j,\mathbf{p}_{ij}\rangle\mathbf{p}_{ij}\\
&+\frac{\langle \boldsymbol{x}_j\times\boldsymbol{v}_j,\boldsymbol{x}_i\rangle}{\|\boldsymbol{x}_i+\boldsymbol{x}_j\|}\left(\langle \boldsymbol{x}_i,\boldsymbol{x}_j\rangle\mathbf{q}_{ij}+\frac{\|\boldsymbol{x}_i+\boldsymbol{x}_j\|\|\boldsymbol{x}_i-\boldsymbol{x}_j\|}{2}\mathbf{r}_{ij}\right)\\
&+\frac{\langle \boldsymbol{x}_j\times\boldsymbol{v}_j,\boldsymbol{x}_i\rangle}{\|\boldsymbol{x}_i-\boldsymbol{x}_j\|}\left(\langle \boldsymbol{x}_i,\boldsymbol{x}_j\rangle\mathbf{r}_{ij}-\frac{\|\boldsymbol{x}_i+\boldsymbol{x}_j\|\|\boldsymbol{x}_i-\boldsymbol{x}_j\|}{2}\mathbf{q}_{ij}\right)\\
&-\langle \boldsymbol{x}_i\times\boldsymbol{v}_i,\mathbf{p}_{ij}\rangle\mathbf{p}_{ij}-\frac{\langle \boldsymbol{x}_i\times\boldsymbol{v}_i,\boldsymbol{x}_j\rangle}{\|\boldsymbol{x}_i+\boldsymbol{x}_j\|}\mathbf{q}_{ij}+\frac{\langle \boldsymbol{x}_i\times\boldsymbol{v}_i,\boldsymbol{x}_j\rangle}{\|\boldsymbol{x}_i-\boldsymbol{x}_j\|}\mathbf{r}_{ij}\\
&=\langle \boldsymbol{x}_j\times\boldsymbol{v}_j,\mathbf{p}_{ij}\rangle\mathbf{p}_{ij}\\
&+\langle \boldsymbol{x}_j\times\boldsymbol{v}_j,\boldsymbol{x}_i\rangle\left(\frac{\langle \boldsymbol{x}_i,\boldsymbol{x}_j\rangle}{\|\boldsymbol{x}_i+\boldsymbol{x}_j\|}\mathbf{q}_{ij}+\frac{\|\boldsymbol{x}_i-\boldsymbol{x}_j\|}{2}\mathbf{r}_{ij}\right)\\
&+\langle \boldsymbol{x}_j\times\boldsymbol{v}_j,\boldsymbol{x}_i\rangle\left(\frac{\langle \boldsymbol{x}_i,\boldsymbol{x}_j\rangle}{\|\boldsymbol{x}_i-\boldsymbol{x}_j\|}\mathbf{r}_{ij}-\frac{\|\boldsymbol{x}_i+\boldsymbol{x}_j\|}{2}\mathbf{q}_{ij}\right)\\
&-\langle \boldsymbol{x}_i\times\boldsymbol{v}_i,\mathbf{p}_{ij}\rangle\mathbf{p}_{ij}-\frac{\langle \boldsymbol{x}_i\times\boldsymbol{v}_i,\boldsymbol{x}_j\rangle}{\|\boldsymbol{x}_i+\boldsymbol{x}_j\|}\mathbf{q}_{ij}+\frac{\langle \boldsymbol{x}_i\times\boldsymbol{v}_i,\boldsymbol{x}_j\rangle}{\|\boldsymbol{x}_i-\boldsymbol{x}_j\|}\mathbf{r}_{ij}\\
&=\langle \boldsymbol{x}_j\times\boldsymbol{v}_j,\mathbf{p}_{ij}\rangle\mathbf{p}_{ij}\\
&+\langle \boldsymbol{x}_j\times\boldsymbol{v}_j,\boldsymbol{x}_i\rangle\left(\frac{\langle \boldsymbol{x}_i,\boldsymbol{x}_j\rangle}{\|\boldsymbol{x}_i+\boldsymbol{x}_j\|}\mathbf{q}_{ij}+\frac{1-\langle \boldsymbol{x}_i,\boldsymbol{x}_j\rangle}{\|\boldsymbol{x}_i-\boldsymbol{x}_j\|}\mathbf{r}_{ij}\right)\\
&+\langle \boldsymbol{x}_j\times\boldsymbol{v}_j,\boldsymbol{x}_i\rangle\left(\frac{\langle \boldsymbol{x}_i,\boldsymbol{x}_j\rangle}{\|\boldsymbol{x}_i-\boldsymbol{x}_j\|}\mathbf{r}_{ij}-\frac{1+\langle \boldsymbol{x}_i,\boldsymbol{x}_j\rangle}{\|\boldsymbol{x}_i+\boldsymbol{x}_j\|}\mathbf{q}_{ij}\right)\\
&-\langle \boldsymbol{x}_i\times\boldsymbol{v}_i,\mathbf{p}_{ij}\rangle\mathbf{p}_{ij}-\frac{\langle \boldsymbol{x}_i\times\boldsymbol{v}_i,\boldsymbol{x}_j\rangle}{\|\boldsymbol{x}_i+\boldsymbol{x}_j\|}\mathbf{q}_{ij}+\frac{\langle \boldsymbol{x}_i\times\boldsymbol{v}_i,\boldsymbol{x}_j\rangle}{\|\boldsymbol{x}_i-\boldsymbol{x}_j\|}\mathbf{r}_{ij}\\
&=\langle \boldsymbol{x}_j\times\boldsymbol{v}_j,\mathbf{p}_{ij}\rangle\mathbf{p}_{ij}-\langle \boldsymbol{x}_i\times\boldsymbol{v}_i,\mathbf{p}_{ij}\rangle\mathbf{p}_{ij}\\
&+\langle \boldsymbol{x}_j\times\boldsymbol{v}_j,\boldsymbol{x}_i\rangle\left(-\frac{1}{\|\boldsymbol{x}_i+\boldsymbol{x}_j\|}\mathbf{q}_{ij}+\frac{1}{\|\boldsymbol{x}_i-\boldsymbol{x}_j\|}\mathbf{r}_{ij}\right)\\
&+\langle \boldsymbol{x}_i\times\boldsymbol{v}_i,\boldsymbol{x}_j\rangle\left(-\frac{1}{\|\boldsymbol{x}_i+\boldsymbol{x}_j\|}\mathbf{q}_{ij}+\frac{1}{\|\boldsymbol{x}_i-\boldsymbol{x}_j\|}\mathbf{r}_{ij}\right)\\
&\quad=-\langle \boldsymbol{x}_i\times\boldsymbol{v}_i-\boldsymbol{x}_j\times\boldsymbol{v}_j,\mathbf{p}_{ij}\rangle\mathbf{p}_{ij}\\
&\quad-\langle \boldsymbol{x}_i\times\boldsymbol{v}_i-\boldsymbol{x}_j\times\boldsymbol{v}_j,\mathbf{r}_{ij}\rangle\left(-\frac{\|\boldsymbol{x}_i-\boldsymbol{x}_j\|}{\|\boldsymbol{x}_i+\boldsymbol{x}_j\|}\mathbf{q}_{ij}+\mathbf{r}_{ij}\right).
\end{aligned}
$$

Appendix D
Proof of Lemma III.5

(i) We use Lemma III.4 (i) and Remark III.2 to get

$$
\begin{aligned}
&\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\dot{\mathbf{p}}\rangle\\
&\quad=\frac{\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{q}\rangle^2-\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{r}\rangle^2}{\|\boldsymbol{x}_1\times\boldsymbol{x}_2\|},
\end{aligned}
$$

and

$$
\left\langle \frac{d}{dt}(\boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2),\mathbf{p}\right\rangle=-\kappa\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{p}\rangle.
$$

We add these two relations to obtain (III.15).
(ii) We use Lemma III.4 (ii) and Remark III.2 to get

$$\begin{aligned}&\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\dot{\mathbf{q}}\rangle\\&\quad=-\frac{\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{p}\rangle\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{q}\rangle}{\|\boldsymbol{x}_1\times\boldsymbol{x}_2\|}\\&\quad+\frac{\langle\boldsymbol{x}_1,\boldsymbol{v}_2\rangle-\langle\boldsymbol{x}_2,\boldsymbol{v}_1\rangle}{2\|\boldsymbol{x}_1\times\boldsymbol{x}_2\|}\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{r}\rangle,\\&\Big\langle\frac{d}{dt}(\boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2),\mathbf{q}\Big\rangle=0\end{aligned}$$

to obtain the desired identity.
(iii) Similarly, we use Lemma III.4 (iii) and Remark III.2 to get

$$\begin{aligned}&\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\dot{\mathbf{r}}\rangle\\&\quad=\frac{\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{p}\rangle\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{r}\rangle}{\|\boldsymbol{x}_1\times\boldsymbol{x}_2\|}\\&\quad-\frac{\langle\boldsymbol{x}_1,\boldsymbol{v}_2\rangle-\langle\boldsymbol{x}_2,\boldsymbol{v}_1\rangle}{2\|\boldsymbol{x}_1\times\boldsymbol{x}_2\|}\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{q}\rangle,\\&\Big\langle\frac{d}{dt}(\boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2),\mathbf{r}\Big\rangle=-\kappa\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{r}\rangle\end{aligned}$$

to finish the proof.

Appendix E
Proof of Lemma III.6

It follows from Lemma III.5 that

$$\begin{aligned}&\frac{d}{dt}\Big(\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{q}\rangle^2-\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{r}\rangle^2\Big)\\&=-\frac{2\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{p}\rangle\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{q}\rangle^2}{\|\boldsymbol{x}_1\times\boldsymbol{x}_2\|}\\&-\frac{2\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{p}\rangle\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{r}\rangle^2}{\|\boldsymbol{x}_1\times\boldsymbol{x}_2\|}\\&+\frac{2\big(\langle\boldsymbol{x}_1,\boldsymbol{v}_2\rangle-\langle\boldsymbol{x}_2,\boldsymbol{v}_1\rangle\big)}{\|\boldsymbol{x}_1\times\boldsymbol{x}_2\|}\\&\quad\times\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{q}\rangle\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{r}\rangle\\&+2\kappa\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{r}\rangle^2\\&\ge-\frac{4\sqrt{\mathcal{E}(0)}}{\sqrt{1-\delta^2}}\Big(\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{q}\rangle^2\\&+\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{r}\rangle^2\\&+|\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{q}\rangle||\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{r}\rangle|\Big)\\&+2\kappa\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{r}\rangle^2\\&\ge-\frac{6\sqrt{\mathcal{E}(0)}}{\sqrt{1-\delta^2}}\Big(\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{q}\rangle^2\\&+\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{r}\rangle^2\Big)+2\kappa\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{r}\rangle^2\\&\ge\kappa\Big(\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{r}\rangle^2-\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{q}\rangle^2\Big).\end{aligned}$$

Appendix F
Proof of Proposition III.2

The estimate (III.17) will be done in three steps.

- Step A: We show

$$\Delta(\boldsymbol{x}(t),\boldsymbol{v}(t))\ge-2\mathcal{E}^0e^{-\kappa(t-t_0)}.\tag{F.4}$$

It follows from (III.16) that

$$\begin{aligned}\Delta(\boldsymbol{x}(t),\boldsymbol{v}(t))&\ge e^{-\kappa(t-t_0)}\Delta(\boldsymbol{x}(t_0),\boldsymbol{v}(t_0))\\&\ge-2\mathcal{E}(t_0)e^{-\kappa(t-t_0)}\ge-2\mathcal{E}^0e^{-\kappa(t-t_0)}.\end{aligned}$$

In the second inequality, the following estimate was used:

$$\begin{aligned}\Delta(\boldsymbol{x},\boldsymbol{v})&\ge-\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{r}\rangle^2\\&\ge-\|\boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2\|^2\\&\ge-2\|\boldsymbol{x}_1\times\boldsymbol{v}_1\|^2-2\|\boldsymbol{x}_2\times\boldsymbol{v}_2\|^2=-2\mathcal{E}.\end{aligned}$$

- Step B: We show

$$\begin{aligned}&\langle \boldsymbol{x}_1(t)\times\boldsymbol{v}_1(t)-\boldsymbol{x}_2(t)\times\boldsymbol{v}_2(t),\mathbf{p}(t)\rangle\\&\ge-\sqrt{2\mathcal{E}^0}e^{-\kappa(t-t_0)}-\frac{2\mathcal{E}^0}{\sqrt{1-\delta^2}}(t-t_0)e^{-\kappa(t-t_0)}.\end{aligned}\tag{F.5}$$

We use Lemma III.5 (i), (F.4) and $\|\boldsymbol{x}_1\times\boldsymbol{x}_2\|\ge\sqrt{1-\delta^2}$ to see

$$\begin{aligned}&\frac{d}{dt}\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{p}\rangle\\&=-\kappa\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{p}\rangle+\frac{\Delta(\boldsymbol{x}(t),\boldsymbol{v}(t))}{\|\boldsymbol{x}_1\times\boldsymbol{x}_2\|}\\&\ge-\kappa\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{p}\rangle-\frac{2\mathcal{E}^0}{\sqrt{1-\delta^2}}e^{-\kappa(t-t_0)}.\end{aligned}\tag{F.6}$$

By Grönwall's lemma, one has

$$\begin{aligned}&\langle \boldsymbol{x}_1(t)\times\boldsymbol{v}_1(t)-\boldsymbol{x}_2(t)\times\boldsymbol{v}_2(t),\mathbf{p}(t)\rangle\\&\ge\langle \boldsymbol{x}_1(t_0)\times\boldsymbol{v}_1(t_0)-\boldsymbol{x}_2(t_0)\times\boldsymbol{v}_2(t_0),\mathbf{p}(t_0)\rangle e^{-\kappa(t-t_0)}\\&-\frac{2\mathcal{E}^0}{\sqrt{1-\delta^2}}(t-t_0)e^{-\kappa(t-t_0)}\\&\ge-\|\boldsymbol{x}_1(t_0)\times\boldsymbol{v}_1(t_0)-\boldsymbol{x}_2(t_0)\times\boldsymbol{v}_2(t_0)\|e^{-\kappa(t-t_0)}\\&-\frac{2\mathcal{E}^0}{\sqrt{1-\delta^2}}(t-t_0)e^{-\kappa(t-t_0)}\\&\ge-\sqrt{2(\|\boldsymbol{x}_1(t_0)\times\boldsymbol{v}_1(t_0)\|^2+\|\boldsymbol{x}_2(t_0)\times\boldsymbol{v}_2(t_0)\|^2)}e^{-\kappa(t-t_0)}\\&-\frac{2\mathcal{E}^0}{\sqrt{1-\delta^2}}(t-t_0)e^{-\kappa(t-t_0)}\\&\ge-\sqrt{2\mathcal{E}^0}e^{-\kappa(t-t_0)}-\frac{2\mathcal{E}^0}{\sqrt{1-\delta^2}}(t-t_0)e^{-\kappa(t-t_0)}.\end{aligned}$$

- Final step: It follows from (III.14) and (F.5) that for $t\ge t_0$,

$$\begin{aligned}&\frac{d}{dt}\arccos\langle\boldsymbol{x}_1,\boldsymbol{x}_2\rangle=-\langle \boldsymbol{x}_1\times\boldsymbol{v}_1-\boldsymbol{x}_2\times\boldsymbol{v}_2,\mathbf{p}\rangle\\&\quad\le\sqrt{2\mathcal{E}^0}e^{-\kappa(t-t_0)}+\frac{2\mathcal{E}^0}{\sqrt{1-\delta^2}}(t-t_0)e^{-\kappa(t-t_0)}.\end{aligned}\tag{F.7}$$

Now, we integrate the differential inequality (F.7) from t_0 to t to obtain the desired estimate:

$$\begin{aligned}&\arccos\langle\boldsymbol{x}_1(t),\boldsymbol{x}_2(t)\rangle\\&\le\arccos\langle\boldsymbol{x}_1(t_0),\boldsymbol{x}_2(t_0)\rangle+\frac{\sqrt{2\mathcal{E}^0}}{\kappa}(1-e^{-\kappa(t-t_0)})\\&+\frac{2\mathcal{E}^0}{\sqrt{1-\delta^2}}\left(\frac{1-e^{-\kappa(t-t_0)}}{\kappa^2}-\frac{(t-t_0)e^{-\kappa(t-t_0)}}{\kappa}\right)\\&\le\arccos\langle\boldsymbol{x}_1(t_0),\boldsymbol{x}_2(t_0)\rangle+\frac{\sqrt{2\mathcal{E}^0}}{\kappa}+\frac{2\mathcal{E}^0}{\kappa^2\sqrt{1-\delta^2}}.\end{aligned}$$

APPENDIX G
PROOF OF LEMMA V.1

(i) Recall that given a frame field (E_1, E_2) and a vector field $\xi = \xi_1 E_1 + \xi_2 E_2$ along a curve $\alpha(t)$ on the manifold, the covariant derivative $\nabla_{\alpha'}\xi$ can be calculated as follows:

$$\nabla_{\alpha'}\xi = (\xi_1' + \xi_2\omega_{21}(\alpha'))E_1 + (\xi_2' + \xi_1\omega_{12}(\alpha'))E_2.$$

where ω_{12} and ω_{21} are connection form. The connection forms ω_{12} and ω_{21} can be computed as

$$\begin{aligned}\omega_{12} = -\omega_{21} &= -\frac{\partial_y\big(\sqrt{\langle \partial_x, \partial_x\rangle_{\mathbb{H}}}\big)}{\sqrt{\langle \partial_y, \partial_y\rangle_{\mathbb{H}}}}dx + \frac{\partial_x\big(\sqrt{\langle \partial_y, \partial_y\rangle_{\mathbb{H}}}\big)}{\sqrt{\langle \partial_x, \partial_x\rangle_{\mathbb{H}}}}dy \\ &= \frac{1}{y}dx \quad \Longrightarrow \quad \omega_{12}(w_i) = -\omega_{21}(w_i) = \frac{u_i}{y_i},\end{aligned}$$

where ∂_x and ∂_y are the coordinate vector fields of the (x, y)-coordinates and $\langle\cdot,\cdot\rangle_{\mathbb{H}}$ denotes the Riemannian metric on $\mathbb{H}$ (see [32]). We set

$$\xi = w_i \quad \text{and} \quad \alpha = z_i$$

to see

$$\alpha' = w_i, \qquad \frac{Dw_i}{dt} = \nabla_{w_i} w_i.$$

Consider the frame field

$$(E_1, E_2) = (y\mathbf{e}_1, y\mathbf{e}_2) \quad \text{on} \quad \mathbb{H} = \{(x, y) \in \mathbb{R}^2 : y > 0\}.$$

In this case, we have

$$w_i = u_i + \mathrm{i}v_i = u_i\mathbf{e}_1 + v_i\mathbf{e}_2 = \Big(\frac{u_i}{y_i}\Big)(y_i\mathbf{e}_1) + \Big(\frac{v_i}{y_i}\Big)(y_i\mathbf{e}_2).$$

Hence, we set

$$\xi_1 = \frac{u_i}{y_i}, \qquad \xi_2 = \frac{v_i}{y_i},$$

$$\xi_1' = \frac{d}{dt}\Big(\frac{u_i}{y_i}\Big) = \frac{\dot{u}_i}{y_i} - \frac{u_i v_i}{y_i^2}, \qquad \xi_2' = \frac{d}{dt}\Big(\frac{v_i}{y_i}\Big) = \frac{\dot{v}_i}{y_i} - \frac{v_i^2}{y_i^2}.$$

By direct calculation, one has

$$\begin{aligned}\frac{Dw_i}{dt} &= \left(\frac{\dot{u}_i}{y_i} - \frac{u_i v_i}{y_i^2} + \frac{v_i}{y_i}\Big(-\frac{u_i}{y_i}\Big)\right) y_i\mathbf{e}_1 \\ &\quad + \left(\frac{\dot{v}_i}{y_i} - \frac{v_i^2}{y_i^2} + \frac{u_i}{y_i}\frac{u_i}{y_i}\right) y_i\mathbf{e}_2 \\ &= (\dot{u}_i\mathbf{e}_1 + \dot{v}_i\mathbf{e}_2) + \frac{-2u_iv_i\mathbf{e}_1 + (u_i^2 - v_i^2)\mathbf{e}_2}{y_i} \\ &= (\dot{u}_i + \mathrm{i}\dot{v}_i) + \frac{-2u_iv_i + \mathrm{i}(u_i^2 - v_i^2)}{y_i} \\ &= \dot{w}_i + \frac{\mathrm{i}}{y_i}w_i^2.\end{aligned}$$

(ii) Let i, k be any indices in $\{1, \cdots, N\}$. For the derivation of (V.2), we consider the following two cases:

$$x_i = x_k \quad \text{and} \quad x_i \neq x_k.$$

• Case A ($x_i = x_k$): The unique geodesic between the two points z_i and z_k is the *straight line segment* connecting them. Since the parallel transport of a tangent vector $\boldsymbol{v}$ along a geodesic γ preserves the angle between $\boldsymbol{v}$ and γ, the angle between w_k and the geodesic should be equal to that between $P_{ik}w_k$ and the geodesic. Hence, one has

$$\frac{P_{ik}w_k}{|P_{ik}w_k|} = \frac{w_k}{|w_k|}. \tag{G.8}$$

On the other hand, note that parallel transport of a tangent vector preserves its length, measured in the Riemannian metric. Since the norm of a tangent vector w at point z in the Poincaré half plane $\mathbb{H}$ is equal to $|w|/\operatorname{Im}[z]$, we have

$$\frac{|P_{ik}w_k|}{y_i} = \|P_{ik}w_k\|_{z_i} = \|w_k\|_{z_k} = \frac{|w_k|}{y_k}. \tag{G.9}$$

By (G.8) and (G.9), we have

$$P_{ik}w_k = \frac{|P_{ik}w_k|}{|w_k|}w_k = \frac{y_i}{y_k}w_k.$$

This coincides with (V.2) because $x_i = x_k$.

• Case B ($x_i \neq x_k$): Here, the geodesic between z_i and z_k is an arc of the circle centered at the real axis which passes through them. Since parallel transport preserves the angle between the vector and the geodesic, the direction of the parallel transported vector is obtained by a rotation along this circle. Let c be the center of the circle. Then we have

$$P_{ik}w_k = \frac{|P_{ik}w_k|}{|w_k|}\frac{z_i - c}{z_k - c}w_k = \frac{x_i + \mathrm{i}y_i - c}{x_k + \mathrm{i}y_k - c}\frac{y_i}{y_k}w_k.$$

We can obtain the value of c in the following way: note that c is on the perpendicular bisector of the line segment connecting z_i and z_k. In other words, there exists some $s \in \mathbb{R}$ such that

$$c = \frac{z_i + z_k}{2} + \mathrm{i}(z_i - z_k)s.$$

On the other hand, c is on the real axis, i.e. $\operatorname{Im}[c] = 0$. Hence, one has

$$\begin{aligned}\operatorname{Im}[c] = \operatorname{Im}\left[\frac{z_i + z_k}{2} + \mathrm{i}(z_i - z_k)s\right] &= \frac{y_i + y_k}{2} + (x_i - x_k)s \\ = 0 \quad \Longrightarrow \quad s &= \frac{y_i + y_k}{2(x_k - x_i)}.\end{aligned}$$

Therefore we have

$$\begin{aligned}c = \operatorname{Re}[c] &= \operatorname{Re}\left[\frac{z_i + z_k}{2} + \mathrm{i}(z_i - z_k)s\right] \\ &= \frac{x_i + x_k}{2} - (y_i - y_k)s = \frac{x_i + x_k}{2} + \frac{y_i^2 - y_k^2}{2(x_i - x_k)}.\end{aligned}$$

Hence we have

$$\begin{aligned}P_{ik}w_k &= \frac{2(x_i - x_k)(x_i + \mathrm{i}y_i - c)}{2(x_i - x_k)(x_k + \mathrm{i}y_k - c)}\frac{y_i}{y_k}w_k \\ &= \frac{(x_i - x_k)^2 + y_k^2 - y_i^2 + 2\mathrm{i}y_i(x_i - x_k)}{-(x_i - x_k)^2 + y_k^2 - y_i^2 + 2\mathrm{i}y_k(x_i - x_k)}\frac{y_i}{y_k}w_k \\ &= \frac{(x_i - x_k + \mathrm{i}y_i)^2 + y_k^2}{-(x_i - x_k - \mathrm{i}y_k)^2 - y_i^2}\frac{y_i}{y_k}w_k \\ &= -\frac{(x_i - x_k + \mathrm{i}y_i + \mathrm{i}y_k)(x_i - x_k + \mathrm{i}y_i - \mathrm{i}y_k)}{(x_i - x_k - \mathrm{i}y_k + \mathrm{i}y_i)(x_i - x_k - \mathrm{i}y_k - \mathrm{i}y_i)}\frac{y_i}{y_k}w_k \\ &= -\frac{x_i - x_k + \mathrm{i}y_i + \mathrm{i}y_k}{x_i - x_k - \mathrm{i}y_k - \mathrm{i}y_i}\frac{y_i}{y_k}w_k.\end{aligned}$$

This yields the desired formula.

Acknowledgment

The work of S.-Y. Ha is supported by NRF Grant (NRF-2017R1A2B2001864) funded by Korean Government. The work of D. Kim is supported by a KIAS Individual Grant (MG073901) at Korea Institute for Advanced Study.

References

[1] Acebron, J. A., Bonilla, L. L., Pérez Vicente, C. J. P., Ritort, F. and Spigler, R.: *The Kuramoto model: A simple paradigm for synchronization phenomena.* Rev. Mod. Phys. **77** (2005), 137-185.

[2] Albi, G., Bellomo, N., Fermo, L., Ha, S.-Y., Pareschi, L., Poyato, D. and Soler, J.: *Vehicular traffic, crowds, and swarms. On the kinetic theory approach towards research perspective.* To appear in *Math. Models Methods Appl. Sci.* **29** (2019), 1901-2005.

[3] Aydoğdu, A., McQuade, S. T. and Duteil, N. P.: *Opinion dynamics on a general compact Riemannian manifold.* Netw. Heterog. Media **12** (2017), 489-523.

[4] Barbălat, I.: *Systèmes d'équations différentielles d'oscillations non Linéaires*, Rev. Math. Pures Appl. **4** (1959), 267-270.

[5] Buck, J. and Buck, E.: *Biology of synchronous flashing of fireflies.* Nature **211** (1966), 562-564.

[6] Cattiaux, P., Delebecque, F. and Pedeches, L.: *Stochastic Cucker-Smale models: old and new.* Ann. Appl. Probab. **28** (2018), 3239-3286.

[7] Carrillo, J. A., Fornasier, M., Rosado, J. and Toscani, G.: *Asymptotic flocking dynamics for the kinetic Cucker-Smale model.* SIAM. J. Math. Anal. **42** (2010), 218-236.

[8] Chi, D., Choi, S.-H. and Ha, S.-Y.: *Emergent behaviors of a holonomic particle system on a sphere.* J. Math. Phys. **55** (2014), 052703.

[9] Choi, Y.-P., Ha, S.-Y. and Li, Z.: *Emergent dynamics of the Cucker-Smale flocking model and its variants.* In N. Bellomo, P. Degond, and E. Tadmor (Eds.), Active Particles Vol.I - Theory, Models, Applications (tentative title), Series: Modeling and Simulation in Science and Technology, Birkhauser, Springer.

[10] Cucker, F. and Smale, S.: *Emergent behavior in flocks.* IEEE Trans. Automat. Control **52** (2007), 852-862.

[11] Cumin, D. and Unsworth, C. P.: *Generalizing the Kuramoto model for the study of neuronal synchronization in the brain.* Physica D **226** (2007), 181-196.

[12] Duan, R., Fornasier, M. and Toscani, G.: *A kinetic flocking model with diffusion.* Comm. Math. Phys. **300** (2010), 95-145.

[13] Degond, P. and Motsch, S.: *Large-scale dynamics of the persistent turing walker model of fish behavior.* J. Stat. Phys. **131** (2008), 989-1022.

[14] Donchev, V. D., Mladenova, C. D., Mladenov, I. M.: *Some alternatives of the Rodrigues axis-angle formula.* C. R. Acad. Bulgare Sci. **69** (2016), 697-706.

[15] Ermentrout. G.B.: *An adaptive model for synchrony in the firefly Pteroptyx malaccae.* J. Math. Biol. **29** (1991), 571-585.

[16] Ferrante, E., Turgut, A. E., Stranieri, A., Pinciroli, C. and Dorigo, M.: *Self-organized flocking with a mobile robot swarm: a novel motion control method.* Adapt. Behav. **20** (2012), 460-477.

[17] Fetecau, R., Patacchini, F. S. and Park, H.: *Well-posedness and asymptotic behaviour of an aggregation model with intrinsic interactions on sphere and other manifolds.* Submitted.

[18] Fetecau, R. C. and Zhang, B.: *Self-organization on Riemannian manifolds.* J. Geom. Mech. **11** (2019), 397-426.

[19] Ha, S.-Y. and Kim, D.: *A second-order particle swarm model on a sphere and emergent dynamics.* SIAM J. Appl. Dyn. Syst. **18** (2019), 80-116.

[20] Ha, S.-Y., Kim, J., Park, J. and Zhang, X.: *Complete cluster predictability of the CuckerSmale flocking model on the real line.* Arch. Ration. Mech. Anal. **231** (2019), 319-365.

[21] Ha, S.-Y., Ko, D. and Ryoo, S. W: *On the relaxation dynamics of Lohe oscillators on some Riemannian manifolds.* J. Stat. Phys. **172** (2018), 1427-1478.

[22] Ha, S.-Y. and Liu, J.-G.: *A simple proof of Cucker-Smale flocking dynamics and mean field limit.* Commun. Math. Sci. **7** (2009), 297-325.

[23] Ha, S.-Y. and Ruggeri, T.: *Emergent dynamics of a thermodynamically consistent particle model.* Arch. Ration. Mech. Anal. **223** (2017), 1397-1425.

[24] Ha, S.-Y. and Tadmor, E.: *From particle to kinetic and hydrodynamic description of flocking.* Kinet. Relat. Models **1** (2008), 415-435.

[25] Lohe, M. A.: *Non-Abelian Kuramoto model and synchronization.* J. Phys. A: Math. Theor. **42** (2009), 395101.

[26] Machida, M., Kano, T., Yamada, S., Okumura, M., Imamura, T. and Koyama, T.: *Quantum synchronization effects in intrinsic Josephson junctions.* Physica C **468** (2008), 689-694.

[27] Markdahl, J.: *Synchronization on Riemannian manifolds: Multiply connected implies multistable.* Available in https://arxiv.org/abs/1906.07452

[28] Markdahl, J., Thunberg, J. and Goncales, J.: *Almost global consensus on the n-sphere.* IEEE Trans. Automat. Contr. **63** (2018), 1664-1675.

[29] Motsch, S. and Tadmor, E.: *A new model for self-organized dynamics and its flocking behavior.* J. Stat. Phys. **144** (2011), 923-947.

[30] Olfati-Saber, R.: *Flocking for multi-agent dynamic systems: algorithms and theory.* IEEE Trans. Automat. Contr. **51** (2006), 401–420.

[31] Olfati-Saber, R.: *Swarms on sphere: A programmable swarm with synchronous behaviors like oscillator networks*. Proc. of the 45th IEEE conference on Decision and Control. (2006), 5060-5066.

[32] O'neil, B.: *Elementary differential geometry.* Academic Press, 1966.

[33] Perea, L., Elosegui, P. and Gomez, G.: *Extension of the Cucker-Smale control law to space flight formations.* J. Guid. Control **32** (2009), 527-537.

[34] Stahl, S.: *The Poincaré half-plane: A gateway to modern geometry.* Johes and Bartlett Publishers, Boston and London, 1993.

[35] Sarlette, A. and Sepulchre, R.: *Consensus optimization on manifolds.* SIAM J. Control Optim. **48** (2009), 56-76.

[36] Tron, R., Afsari, B. and Vidal, R.: *Riemannian consensus for manifolds with bounded curvature.* IEEE Trans. Automat. Contr. **58** (2013), 921–934.

Seung-Yeal Ha received the B.Sc. degree in mathematics from Seoul National University, Seoul, Korea and the Ph.D. degree in mathematics from Stanford University, in 1997 and 2001, respectively. He has worked at University of Wisconsin-Madison, USA. He is currently professor of Department of Mathematical Sciences, Seoul National University, Seoul, Korea. His main research interest lies on the hyperbolic conservation laws, kinetic equations and emergent collective modeling of interacting particle systems.

Doheon Kim received the B.Sc. and Ph.D. degrees, both in mathematics, from Seoul National University, Seoul, Korea, in 2015 and 2019, respectively. He is currently a research fellow of School of Mathematics, Korea Institute for Advanced Study, Seoul, Korea. His current research interest includes dynamical systems and PDE arising from collective dynamics, and the BGK type models.

Franz Wilhelm Schhlöder received a B.Sc degree in physics and an M.Sc degree in mathematics from the University of Heidelberg, Germany, in 2014 and 2016, respectively. Currently, he is a doctoral candidate in mathematics at the University of Milano-Bicocca, Milan, Italy. His main research interests are applications of topology and geometry.