Alignment of self-propelled rigid bodies: from particle systems to macroscopic equations

Pierre Degond, Amic Frouvelle, Sara Merino-Aceituno, Ariane Trescases

Alignment of self-propelled rigid bodies: from particle systems to macroscopic equations

Pierre Degond[1], Amic Frouvelle[2], Sara Merino-Aceituno[3], and Ariane Trescases[4]

[1] Department of Mathematics, Imperial College London, South Kensington Campus, London, SW7 2AZ, UK,
pdegond@imperial.ac.uk
[2] CEREMADE, CNRS, Université Paris-Dauphine, Université PSL, 75016 PARIS, FRANCE,
frouvelle@ceremade.dauphine.fr
[3] School of Mathematical and Physical Sciences, University of Sussex, Falmer BN1 9RH, UK,
s.merino-aceituno@sussex.ac.uk
Faculty of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Wien, Austria
sara.merino@univie.ac.at
Department of Mathematics, Imperial College London, South Kensington Campus, London, SW7 2AZ, UK,
[4] IMT; UMR5219, Université de Toulouse; CNRS, F-31400 Toulouse, France,
ariane.trescases@math.univ-toulouse.fr

Abstract. The goal of these lecture notes is to present in a unified way various models for the dynamics of aligning self-propelled rigid bodies at different scales and the links between them. The models and methods are inspired from [12,13], but, in addition, we introduce a new model and apply on it the same methods. While the new model has its own interest, our aim is also to emphasize the methods by demonstrating their adaptability and by presenting them in a unified and simplified way. Furthermore, from the various microscopic models we derive the same macroscopic model, which is a good indicator of its universality.

Keywords: Self-propelled particles, rotation matrix, quaternions, alignment, velocity jumps, generalized collisional invariants, self-organized hydrodynamics.

1 Introduction

Collective behavior arises ubiquitously in nature: fish schools, flocks of birds, herds, colonies of bacteria, pedestrian dynamics, opinion formation, are just some examples. One of the main challenges in the investigation of collective behavior is to explain its emergent properties, that is, how from the local interactions between a large number of agents, large-scale structures and self-organization arise at a much larger scale than the agents' sizes. Kinetic theory provides a

mathematical framework for the study of emergent phenomena with the rigorous derivation of equations for the large-scale dynamics (called macroscopic equations) from particle or individual-based models. The derivation of macroscopic equations establishes a rigorous link between the particle dynamics and the large-scale dynamics. Moreover, the simulation of macroscopic equations have the advantage of being, generally, computationally far more efficient than particle simulations, especially as the number of agents grows large.

Tools for the derivation of macroscopic equations were first developed in Mathematical Physics, particularly, in the framework of the Boltzmann equation for rarefied gases [7,9,26]. However, compared to the case of classical equations in Mathematical Physics, an additional difficulty arises here in the study of living systems: the lack of conservation laws. In classical physical systems, each macroscopic quantity corresponds to a conservation law (like the conservation of the total mass, momentum and energy). However, in the models that we will consider here, the number of conserved quantities is less than the number of macroscopic quantities to be determined. To overcome this difficulty we will use the methodological breakthrough presented in [16]: the Generalized Collision Invariant (GCI). This new concept relaxes the condition of being a conserved quantity, and has then been used in a lot of works related to alignment of self-propelled particles [4,10,11,12,13,14,15,17,18,19,22]. The goal of this exposition is precisely to clearly illustrate the application of this methodology to models for collective dynamics based on alignment of the body position.

Specifically, in the models that will be considered in this exposition each agent is described by its location in the three-dimensional space and the orientation of its body, represented by a three-dimensional frame. Each agent perceives (directly or indirectly) the orientations of the bodies of the neighboring agents and tends to align with them. This type of collective motion can be found, e.g., in sperm dynamics and animals (birds, fish), and it is a stepping stone to modeling more complex agents composed of articulated bodies (corpora [8]). For more examples and applications based on body attitude coordination see [25] and references therein.

Our models are inspired by time-continuous versions of the Vicsek model, introduced in the 90's [28]. The Vicsek model is now a classic in the field of collective motion: self-propelled particles move at constant speed while trying to align their direction of movement with their neighbors up to some noise. We consider time-continuous versions of the Vicsek model since they are more prone to mathematical studies, as pointed out in [16]. However, there is no obvious unique way of writing a time-continuous version. In [16] and then in [19], two different continuous versions have been proposed that differ by the way agents approach the aligned state: in the first one the particles' velocities align gradually over time towards an aligned state, and in the second one the velocities make discontinuous jumps at discrete times towards an aligned state. Interestingly, both models in [16] and in [19] give rise to the same hydrodynamic/macroscopic limit (with different values for the constants in the equations). Inspired by this, here we will present two models for alignment of rigid bodies, one given by

a time-continuous gradual alignment (taken from the references [12,13]), and another one for alignment based on a jump process on the velocities, that we present here for the first time.

The reason for considering here these two types of models is the following. The main difficulty in applying the Generalized Collision Invariant method to obtain the macroscopic equations lays, precisely, on finding the explicit form of the Generalized Collision Invariants. Indeed, in [12,13] that was the main mathematical difficulty. However, we will see that in the jump model it is straightforward to obtain the GCI but, at the same time, the computation of the macroscopic limit keeps the same structure as in the previous results [12,13]. Particularly, we will obtain the same macroscopic equations (though with different values for the coefficients). The jump model constitutes, therefore, an excellent framework for a didactic exposition of the GCI methodology. With this, the proofs in [12,13] will become more accessible to the reader.

Here, to model alignment of the orientations of the agents seen as rigid bodies (and not only the alignment of their velocities as in the original Vicsek model), we represent the body orientation of an agent as a three-dimensional frame, obtained by the rotation of a fixed frame. Therefore, we will represent the orientations of the agents as rotations. But, as we will see in Section 2, in the three-dimensional space rotations can be equivalently represented by rotation matrices and unitary quaternions. Using rotation matrices, the modeling at the individual-based level is more natural and intuitive. However, in terms of numerical efficiency, quaternions require less memory usage (it only requires storing 4 entries rather than 9 entries for matrices) and are less costly to renormalize (while obtaining a rotation matrix from an approximate matrix typically requires a polar decomposition, obtaining a unit quaternion from an approximate quaternion only requires dividing by the norm). We will also see that working with quaternions can give rise to a better presentation of the macroscopic equations.

We conclude by noting that the study of collective behavior based on the Vicsek model and its variations is a fertile field. Among many of the existing mathematical works, we highlight [16] where the hydrodynamic limit has been computed as well as [11], where the emergence of phase transitions is investigated. Many refinements have been proposed to incorporate additional mechanism, such as, to cite only a few of them, volume exclusion [18], presence of leaders [21] or polarization of the group [6]. We refer the interested reader to [12,13] and the references therein for more on this topic.

The structure of the document is as follows. We start by introducing some notations and recalling some useful properties on matrices, rotations and quaternions in Sec. 2. We present the individual-based models in Sec. 3. From there we derive the mesoscopic models in Sec. 4. Finally we compute the hydrodynamic limit in Sec. 5 and discuss the results in Sec. 6.

2 Preliminaries: matrices, rotations and quaternions

We present in this section some notations and useful properties on matrices, rotations, and quaternions. We first introduce some notations and present some properties on matrices and rotation matrices. In a second subsection we detail the link between rotations in $\mathbb{R}^3$ and unit quaternions.

The main drawback in using quaternion to represent a rotation is that two opposite quaternions represent the same rotation. In analogy with the theory of rodlike polymers [20], where two opposite unit vectors represent the same orientation, we also present in this first subsection the formalism of Q-tensors, which will be helpful for the modeling. Most of the results will be given without detailed proofs, which can be found in Section 5 of [13].

2.1 Matrices, rotation matrices and $\mathbb{R}^3$

We start by introducing a few notations. We will use the following matrix spaces:

- $\mathcal{M}$ is the set of three-by-three matrices,
- $\mathcal{S}$ is the set of symmetric three-by-three matrices,
- $\mathcal{A}$ is the set of antisymmetric three-by-three matrices,
- $O_3(\mathbb{R})$ is the orthogonal group in dimension three,
- $SO_3(\mathbb{R})$ is the special orthogonal group in dimension three.

For a matrix $A \in \mathcal{M}$, we denote by A^{T} its transpose, and we write $\operatorname{Tr} A$ its trace, $\operatorname{Tr} A = \sum_i A_{ii}$. The matrix I is the identity matrix in $\mathcal{M}$. We use the following definition of the dot product on $\mathcal{M}$: for A, $B \in \mathcal{M}$,

$$A \cdot B := \frac{1}{2} \sum_{i,j=1}^{3} A_{ij} B_{ij}.$$

The choice of this dot product (note in particular the factor $\frac{1}{2}$) is motivated by the following property: for any $u = (u_1, u_2, u_3) \in \mathbb{R}^3$, define the antisymmetric matrix $[u]_\times$ such that

$$[u]_\times := \begin{pmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{pmatrix},$$

(or equivalently such that for any $v \in \mathbb{R}^3$, we have $[u]_\times v = u \times v$). Then we have for any u, $v \in \mathbb{R}^3$:

$$[u]_\times \cdot [v]_\times = u \cdot v.$$

The following properties will be useful in the sequel. We state them without proof but the interested reader can find them in Ref. [12].

Proposition 1 (Space decomposition in symmetric and antisymmetric matrices). *We have*

$$\mathcal{S} \oplus \mathcal{A} = \mathcal{M} \text{ and } \mathcal{A} \perp \mathcal{S}.$$

Proposition 2 (Tangent space to $SO_3(\mathbb{R})$, and projection). *For a matrix $A \in SO_3(\mathbb{R})$, denote by T_A the tangent space to $SO_3(\mathbb{R})$ at A. Then*

$$M \in T_A \text{ if and only if there exists } P \in \mathcal{A} \text{ s.t } M = AP,$$

or equivalently the same statement with $M = PA$. Consequently, the orthogonal projection of a matrix M on T_A is given by

$$P_{T_A}(M) = \frac{1}{2}(M - AM^{\mathrm{T}}A), \tag{1}$$

and we have that $M \in T_A^{\perp}$ if and only if $M = AS$ (or equivalently $M = SA$), for some $S \in \mathcal{S}$.

We end up by recalling the polar decomposition of a matrix.

Proposition 3. *Let $M \in \mathcal{M}$. There exist $A \in O_3(\mathbb{R})$ and $S \in \mathcal{S}$ such that*

$$M = AS.$$

Furthermore, if $\det M \neq 0$*, then A and S are unique. In this case, we write*

$$PD(M) := A.$$

2.2 Quaternions, rotations and Q-tensors

Besides rotation matrices, another common representation of rotations in $\mathbb{R}^3$ is done through the unit quaternions, which will be denoted by $\mathbb{H}_1$. Recall that any quaternion q can be written as $q = a + b\mathrm{i} + c\mathrm{j} + d\mathrm{k}$ with $a, b, c, d \in \mathbb{R}$. Quaternions form a four dimensional (non commutative) division algebra, by the rules $\mathrm{i}^2 = \mathrm{j}^2 = \mathrm{k}^2 = \mathrm{ijk} = -1$. The real part $\mathrm{Re}(q)$ of the quaternion q is a and its imaginary part, denoted $\mathrm{Im}(q)$ is $b\mathrm{i}+c\mathrm{j}+d\mathrm{k}$. The three-dimensional space of purely imaginary quaternions is then identified with $\mathbb{R}^3$, therefore whenever in the paper we have a vector in $\mathbb{R}^3$ which is used as a quaternion, it should be understood that it is a purely imaginary quaternion thanks to this identification. For instance, the vector $e_1 \in \mathbb{R}^3$ (resp. e_2, e_3) is identified with the quaternion i (resp. j, k). The conjugate of the quaternion q is given by $q^* = \mathrm{Re}(q) - \mathrm{Im}(q)$, therefore we get $qq^* = |q|^2 = a^2 + b^2 + c^2 + d^2 \geqslant 0$.

We now explain how the group $\mathbb{H}_1$ (the unit quaternions q, such that $|q| = 1$) provides a representation of rotations. Any unit quaternion $q \in \mathbb{H}_1$ can be written in a polar form as $q = \cos(\theta/2)+\sin(\theta/2)n$, where $\theta \in [0, 2\pi)$ and $n \in \mathbb{S}^2$ (a purely imaginary quaternion with the previous identification). With this notation, the unit quaternion q represents the rotation of angle θ around the axis given by the direction n, anti-clockwise. More specifically, for any vector $u \in \mathbb{R}^3$, the vector quq^* (which is indeed a pure imaginary quaternion whenever $q \in \mathbb{H}_1$ and u is a pure imaginary quaternion, so it can be seen as a vector in $\mathbb{R}^3$) is the rotation of u of angle θ around the axis given by the direction n (note that θ and n are uniquely defined except when $q = \pm 1$: in this case the associated rotation is the identity, and any direction $n \in \mathbb{S}^2$ is suitable).

The underlying map from the group of unit quaternions to the group of rotation matrices is then given by

$$\Phi : \begin{array}{c} \mathbb{H}_1 \to SO_3(\mathbb{R}) \\ q \mapsto \Phi(q) : \begin{array}{c} \mathbb{R}^3 \to \mathbb{R}^3 \\ u \mapsto quq^*. \end{array} \end{array} \tag{2}$$

It is then straightforward to get that Φ is a morphism of groups: for any q and $\widetilde{q}$ in $\mathbb{H}_1$, we have $\Phi(q\widetilde{q}) = \Phi(q)\Phi(\widetilde{q})$ and $\Phi(q^*) = \Phi(q)^{\mathrm{T}}$.

An important remark is that two opposite unit quaternions represent the same rotation:

$$\forall q \in \mathbb{H}_1, \quad \Phi(q) = \Phi(-q). \tag{3}$$

More precisely, the kernel of Φ is given by $\{\pm 1\}$, so that Φ induces an isomorphism between $\mathbb{H}_1/\{\pm 1\}$ and $SO_3(\mathbb{R})$.

We finally briefly introduce the notion of Q-tensors. Indeed, since a unitary quaternion and its opposite correspond to the same rotation matrix, we can see an analogy with the theory of suspensions of rodlike polymers [20]. Those polymers are also modeled using unit vectors (in this case, in $\mathbb{R}^3$), and two opposite vectors are describing the same orientation. Their alignment is called nematic. One relevant object in this theory is the so-called Q-tensor associated with the unit quaternion q, given by the matrix $Q = q \otimes q - \frac{1}{4}\mathrm{I}_4$, where q is seen as a unit vector in $\mathbb{R}^4$, and I_4 is the identity matrix of size four. This object is a symmetric and trace free four by four matrix, which is invariant under the transformation $q \mapsto -q$. We denote by $\mathcal{S}_4^0$ the space of symmetric trace free 4×4 matrices (a vector space of dimension 9), and endow it with the dot product known as "contraction of tensors"; more precisely if $Q, \widetilde{Q}$ are in $\mathcal{S}_4^0$, their contraction $Q : \widetilde{Q} = \sum_{i,j} Q_{ij}\widetilde{Q}_{ij}$ is the trace of $Q\widetilde{Q}^{\mathrm{T}}$. We then get a map

$$\Psi : \begin{array}{c} \mathbb{H}_1 \to \mathcal{S}_4^0 \\ q \mapsto q \otimes q - \frac{1}{4}\mathrm{I}_4, \end{array}$$

whose image can also be identified with $\mathbb{H}_1/\{\pm 1\}$. Indeed, the preimage of $\Psi(q)$ is always equal to $\{q, -q\}$. We therefore have two ways to see $\mathbb{H}_1/\{\pm 1\}$ as a submanifold of a nine-dimensional vector space : either as the image of Φ (in $\mathcal{M}$), which is exactly $SO_3(\mathbb{R})$, or as the image of Ψ (in $\mathcal{S}_4^0$). It appears that the dot products on these spaces behave remarkably well, regarding the maps Φ and Ψ, as stated in the following proposition, from which we can also see that the images are submanifolds of the spheres of radii $\sqrt{\frac{3}{2}}$ (in $\mathcal{M}$) and $\frac{\sqrt{3}}{2}$ (in $\mathcal{S}_4^0$).

Proposition 4. *For any unit quaternions q and $\widetilde{q}$, we have*

$$\frac{1}{2}\Phi(q) \cdot \Phi(\widetilde{q}) = (q \cdot \widetilde{q})^2 - \frac{1}{4} = \Psi(q) : \Psi(\widetilde{q}).$$

Proof. For the second equality, recall that for any quaternions q and $\widetilde{q}$ we have by definition $(q \otimes \widetilde{q})_{ii} = q_i\widetilde{q}_i$, therefore $\mathrm{Tr}(q \otimes \widetilde{q}) = q \cdot \widetilde{q}$ (this justifies the

fact that $\mathrm{Tr}(q \otimes q - \frac{1}{4}\mathrm{I}_4) = 0$ when q is a unit quaternion). Using the fact that $(q \otimes q)(\widetilde{q} \otimes \widetilde{q}) = (q \cdot \widetilde{q})\, q \otimes \widetilde{q}$, we get, when q and $\widetilde{q}$ are unit quaternions:

$$\begin{aligned}\Psi(q) : \Psi(\widetilde{q}) &= \mathrm{Tr}((q \otimes q - \tfrac{1}{4}\mathrm{I}_4)(\widetilde{q} \otimes \widetilde{q} - \tfrac{1}{4}\mathrm{I}_4)) \\ &= \mathrm{Tr}((q \otimes q)(\widetilde{q} \otimes \widetilde{q} - \tfrac{1}{4}\mathrm{I}_4)) = (q \cdot \widetilde{q})^2 - \frac{1}{4}.\end{aligned}$$

For the first equality, we first prove that $\mathrm{Tr}(\Phi(q)) = 4\mathrm{Re}(q)^2 - |q|^2$ for any quaternion q. Indeed we first have that $\mathrm{Tr}(\Phi(q)) = \sum_{i=1}^{3} e_i \cdot \Phi(q) e_i$. Writing q as $a + b\mathrm{i} + c\mathrm{j} + d\mathrm{k}$ and using the identifications between $\mathbb{R}^3$ and the purely imaginary quaternions, we get for instance that $e_1 \cdot \Phi(q) e_1 = \mathrm{Re}(\mathrm{i}(q\mathrm{i}q^*)^*)$, which after computations is $a^2 + b^2 - c^2 - d^2$. At the end, with similar computations for e_2 and e_3, we get $\mathrm{Tr}(\Phi(q)) = 3a^2 - b^2 - c^2 - d^2 = 4\mathrm{Re}(q)^2 - |q|^2$. Therefore, if q and $\widetilde{q}$ are unit quaternions, we get

$$\begin{aligned}2\Phi(q) \cdot \Phi(\widetilde{q}) &= \mathrm{Tr}(\Phi(q)\Phi(\widetilde{q})^T) \\ &= \mathrm{Tr}(\Phi(q\widetilde{q}^*)) = 4\mathrm{Re}(q\widetilde{q}^*)^2 - |q\widetilde{q}^*|^2 = 4(q \cdot \widetilde{q})^2 - 1. \square\end{aligned}$$

We will also make use of the two following properties regarding the differentiability of the map Φ and the volume forms in $SO_3(\mathbb{R})$ and $\mathbb{H}_1$.

Proposition 5. *The map Φ is continuously differentiable on $\mathbb{H}_1$. Denoting its differential at $q \in \mathbb{H}_1$ by $\mathrm{D}_q\Phi : q^\perp \longrightarrow T_{\Phi(q)}$, we have that for any $p \in q^\perp$,*

$$\mathrm{D}_q\Phi(p) = 2\,[pq^*]_\times\, \Phi(q).$$

Here, we wrote T_A the tangent space of $SO_3(\mathbb{R})$ at $A = \Phi(q)$, and $q^\perp$ the orthogonal of q.

Proposition 6. *Consider a function $g : SO_3(\mathbb{R}) \to \mathbb{R}$, then*

$$\int_{SO_3(\mathbb{R})} g(A)\,\mathrm{d}A = \int_{\mathbb{H}_1} g(\Phi(q))\,\mathrm{d}q,$$

where $\mathrm{d}q$ and $\mathrm{d}A$ are the normalized Lebesgue measures on the hypersphere $\mathbb{H}_1$ and on $SO_3(\mathbb{R})$, respectively. Furthermore, if $B \in SO_3(\mathbb{R})$, then

$$\int_{SO_3(\mathbb{R})} g(A)\,\mathrm{d}A = \int_{SO_3(\mathbb{R})} g(AB)\,\mathrm{d}A = \int_{SO_3(\mathbb{R})} g(BA)\,\mathrm{d}A.$$

3 Individual Based Modeling: alignment of self-propelled rigid bodies

Our goal is to model a large number N of particles described, for $n = 1, \ldots, N$, by their positions $X_n \in \mathbb{R}^3$ and their orientations as rigid bodies. The most natural way to describe such an orientation is to give three orthogonal unit vectors u_n, v_n, and w_n. For instance, one way to describe the full orientation of

a bird, would be to set the first vector u_n as the direction of its movement (from the center to the beak), the second vector v_n as the direction of its left wing (from the center to the wing), and the last one w_n as the direction of the back, so that u_n, v_n and w_n form a direct orthogonal frame. Therefore the matrix A_n whose three columns are exactly u_n, v_n and w_n is a special orthogonal matrix. This rotation matrix A_n represents the rotation that has to be done between a reference particle the orthogonal vectors of which are exactly the canonical basis of $\mathbb{R}^3$ (denoted by $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, and $e_3 = (0, 0, 1)$), and the particle number n. A particle can then be described by a pair $(X_n, A_n) \in \mathbb{R}^3 \times SO_3(\mathbb{R})$.

In the spirit of the Vicsek model [28], we want to include in the modeling the three following rules :

- particles move at constant speed,
- particles try to align with their neighbors,
- this alignment is subject to some noise.

Up to changing the time units, we will consider that all particles move at speed one. The first rule requires a direction of movement for each particle. Therefore, in the following, we will suppose that the first vector $u_n = A_n e_1$ of the matrix A_n represents the velocity of the particle number n. Then, the evolution of the position X_n will simply be given by

$$\frac{\mathrm{d}X_n}{\mathrm{d}t} = A_n e_1. \tag{4}$$

In the quaternion framework, if the quaternion q_n represents the orientation of particle number n (meaning that $\Phi(q_n) = A_n$) then the equation corresponding to (4) reads:

$$\frac{\mathrm{d}X_n}{\mathrm{d}t} = q_n e_1 q_n^*. \tag{5}$$

We now want to describe the evolution of q_n (or of the rotation matrix A_n), taking into account the two remaining rules.

3.1 Defining the target for the alignment mechanism

To implement the second rule in the modeling, in the spirit of the Vicsek model, we need to provide for each particle a way to compute the "average orientation" of the neighbors. In the Vicsek model, the idea was to take the sum of all the velocities of the neighbors and to normalize it in order to have a unit target velocity.

In our framework of rotation matrices, to apply the same procedure, if we want the target orientation $\bar{A}_n$ (viewed from the particle number n) to be a rotation matrix, we need a procedure of normalization which from any matrix gives a matrix of $SO_3(\mathbb{R})$. Indeed, the sum of all rotations matrix A_m of the neighbors need not be a rotation matrix (nor a multiple of such a matrix). The

choice that had been done in [12] was to take the polar decomposition : we denote

$$\bar{J}_n = \frac{1}{N}\sum_{m=1}^{N} K(X_m - X_n)A_m \tag{6}$$

$$\bar{A}_n = \mathrm{PD}(\bar{J}_n), \tag{7}$$

where $\mathrm{PD}(J)$ (when $\det(J) \neq 0$) denotes the orthogonal matrix in the polar decomposition (see Prop. 3) of a matrix J, and K is an observation kernel, which weights the orientations of neighbors. A simple example is $K(x) = 1$ if $0 \leqslant |x| < R$ and $K(x) = 0$ if $|x| \geqslant R$. In that case all the neighbors located in the ball of radius R and center X_n have the same influence on the computation of the average matrix $\bar{A}_n$, and all individuals located outside this ball have no influence on this computation.

A first difficulty arises here when the polar decomposition is not a rotation matrix, that is to say $\det(\bar{J}_n) \leqslant 0$. Indeed, due to random effects, we can expect that the polar decomposition is almost surely defined (that is to say $\det(\bar{J}_n) \neq 0$, which happens on a negligible set), but we cannot expect that $\det(\bar{J}_n) > 0$ almost surely.

In the framework of unitary quaternions, things are slightly more complicated. Indeed, since a unitary quaternion and its opposite correspond to the same rotation matrix (see eq. (3)), the expression used to compute an "average orientation" needs to be invariant by the change of sign of any of the quaternions appearing in the formula. Using the Q-tensors as in the theory of suspensions of rodlike polymers [20] is a good option. We are then led to averaging objects of the form $q \otimes q - \frac{1}{4}\mathrm{I}_4$ which are invariant under the transformation $q \mapsto -q$. The average Q-tensor of the neighbors would then take the form

$$\bar{Q}_n = \frac{1}{N}\sum_{m=1}^{N} K(X_m - X_n)(q_m \otimes q_m - \tfrac{1}{4}\mathrm{I}_4). \tag{8}$$

To define now an "average" quaternion from this Q-tensor $\bar{Q}_n$, we need a procedure which provides a unit vector. We expect that if all quaternions q_n are all equal to a given q (or to $-q$), the procedures returns q or $-q$. Therefore, from the form $\alpha(q \otimes q - \frac{1}{4}\mathrm{I}_4)$, with $\alpha > 0$, it should return q or $-q$. These two vectors are precisely the unit eigenvectors associated to the maximal eigenvalue (which is equal to $\frac{3}{4}$, the other eigenvectors, orthogonal to q, being associated to the eigenvalue $\frac{-1}{4}$). Therefore, in [13], we defined

$$\bar{q}_n = \text{one of the unit eigenvectors of } \bar{Q}_n \text{ of maximal eigenvalue.} \tag{9}$$

Note that the direction of $\bar{q}_n$ is uniquely defined when the maximal eigenvalue is simple. Since symmetric matrices with multiple maximal eigenvalues are negligible, we can expect this definition to be well-posed almost surely.

The first unexpected link found in [13] between this framework of quaternions and the previous framework of average matrices, is that these two averaging

procedures are actually equivalent (when the polar decomposition in formula (7) actually returns a rotation matrix). This is due to the following observations [13]:

Proposition 7.

(i) If $M \in \mathcal{M}$ is such that $\det(M) > 0$, then the polar decomposition of M is a rotation matrix, and it is the unique maximizer of the function $A \mapsto A \cdot M$ among all matrices A in $SO_3(\mathbb{R})$.

(ii) A unit eigenvector corresponding to the maximal eigenvalue of a symmetric matrix Q maximizes the function $q \mapsto q \cdot Qq$ among all unit vectors q of $\mathbb{H}_1$.

(iii) If for all n we have $\Phi(q_n) = A_n$, and $\det(\bar{J}_n) > 0$, then $\Phi(\bar{q}_n) = \bar{A}_n$, where $\bar{J}_n$, $\bar{A}_n$ and $\bar{q}_n$ are given by (6)-(7), and (8)-(9).

We therefore have now a good procedure to compute $\bar{A}_n$ thanks to the following maximization problem (instead of polar decomposition):

$$\bar{A}_n = \operatorname{argmax} \{A \in SO_3(\mathbb{R}) \mapsto A \cdot M_n\}, \tag{10}$$

where $\bar{J}_n$ is defined in Eq. (6), and from now on we use this definition of $\bar{A}_n$, which ensures that it corresponds to the definition (9) of $\bar{q}_n$ in the world of quaternions.

Since the next part of the modeling will include some random effects, we can expect that the configurations for which the average is not well-defined will be of negligible probability.

We now need to have evolution equations for the orientations (either the rotation matrices A_n or the unit quaternions q_n). In the spirit of the time discrete Vicsek model, it would correspond to saying $A_n(t + \Delta t) = \bar{A}_n(t)+$ "noise", or $q_n(t + \Delta t) = \bar{q}_n(t)+$ "noise". However, as was pointed out in [16], in this procedure Δt is actually a parameter of the model, and not a time discretization of an underlying process : indeed, this parameter controls the frequency at which particles change their orientation, and changing the value of this frequency leads to drastic changes in the behavior of the model. Regarding the mathematical study of this type of time-discrete models, it is far from being clear how to go beyond observations of numerical simulations. However, it is possible to build models in the same spirit as the Vicsek model which will be much more prone to mathematical study, in particular if we want to derive a kinetic description (when the number of particles is large) and macroscopic limit (when the scale of observation is large). In the next two subsections we present two ways of building such models. The first one corresponds to a time-continuous alignment mechanism as was proposed in [16], in which the orientation of one particle continuously tries to align with its target orientation, up to some noise. This leads to the models presented in [12] (in the framework of rotation matrices) and in [13] (in the framework of unit quaternions). The second one, as in the Vicsek model, corresponds to a process in which orientations undergo jumps as time evolves, but where the jumps are not synchronous: instead of taking place every time step for all particles, they all have independent times at which they change from their orientations to their target orientations, up to some

noise. This leads to a new model, which is different at the particle and kinetic levels, but for which the derivation of the macroscopic model gives the same system of evolution equations (up to the values of the constant parameters of the model). This procedure was studied in [19] for the alignment mechanism of the Vicsek model, and they also found that the macroscopic model corresponds to the Self-Organized Hydrodynamic system of [16] (derived from the time continuous alignment process).

3.2 Gradual alignment model

We first consider a time-continuous alignment mechanism. We have to take into account the last two rules (particles try to align with their neighbors, and this alignment is subject to some noise). For the sake of simplicity, we first present the alignment dynamics without noise. We will add noise at the end of this subsection.

The alignment is modeled by a gradual alignment of an agent's body orientation towards its local average defined in the previous subsection. We express the evolution towards the average as the gradient of a polar distance between the agent and the average. It takes the form, in the world of rotation matrices

$$\frac{\mathrm{d}A_n}{\mathrm{d}t} = \nabla_{A_n}\left[A_n \cdot \bar{A}_n\right],$$

and in the world of unit quaternions

$$\frac{\mathrm{d}q_n}{\mathrm{d}t} = \nabla_{q_n}\left[\frac{1}{2}(q_n \cdot \bar{q}_n)^2\right],$$

where the strength of alignment (or equivalently the relaxation frequency) has been taken to be one (which can be done without loss of generality by changing time units), and ∇_{A_n} and ∇_{q_n} represent the gradients on $SO_3(\mathbb{R})$ and $\mathbb{H}_1$ respectively. For the quaternions, we took the square of the norm to account for the fact that only the directions of the vectors q_n and $\bar{q}_n$, and not their sign, should influence the alignment dynamics (this is called nematic alignment, and it is analogous to the case of rodlike polymers, as described in Subsec. 3.1).

The alignment forces can be rewritten respectively as

$$\nabla_{A_n}\left[A_n \cdot \bar{A}_n\right] = P_{T_{A_n}}\bar{A}_n,$$

for the matrices, where $P_{T_{A_n}}$ is the orthogonal projection on the tangent space of $SO_3(\mathbb{R})$ at A_n, given by Eq. (1) (see Prop. 2), and

$$\nabla_{q_n}\left[\frac{1}{2}(q_n \cdot \bar{q}_n)^2\right] = P_{q_n^\perp}\left[\left(\bar{q}_n \otimes \bar{q}_n - \tfrac{1}{4}\mathrm{I}_4\right) q_n\right].$$

for the quaternions, where $P_{q_n^\perp} = \mathrm{I}_4 - q_n \otimes q_n$ is the projection on the orthogonal of q_n.

The second link found in [13] between the frameworks of quaternions and rotation matrices, is that these alignment mechanisms are also equivalent.

Proposition 8. *Consider the system, for all $n = 1..N$,*

$$\frac{\mathrm{d}A_n}{\mathrm{d}t} = \nabla_{A_n}\left[A_n \cdot \bar{A}_n\right], \tag{11}$$

$$A(t=0) = A_n^0 \in SO_3(\mathbb{R}), \tag{12}$$

with $\bar{A}_n$ defined in (10)*, and the system, for all $n = 1..N$,*

$$\frac{\mathrm{d}q_n}{\mathrm{d}t} = \nabla_{q_n}\left[\frac{1}{2}(q_n \cdot \bar{q}_n)^2\right], \tag{13}$$

$$q(t=0) = q_n^0 \in \mathbb{H}_1, \tag{14}$$

with $\bar{q}_n$ defined in (9)*. If $A_n^0 = \Phi(q_n^0)$ for $n = 1..N$, then, for any solution $(q_n)_n$ of the Cauchy problem* (13)–(14)*, the N-tuple $(A_n)_n := (\Phi(q_n))_n$ is a solution of the Cauchy problem* (11)–(12)*.*

The proof of this proposition relies on two main properties: the equivalence of the averaging procedures of Prop. 7 on one hand, and, on the other hand, the computation of the differential of Φ in Prop. 5, which allows us to write a link between the gradient operators on $SO_3(\mathbb{R})$ and on $\mathbb{H}_1$.

We finally describe the complete model by adding the third rule (the fact that the alignment is subject to some noise). The natural way to introduce it is to transform the ordinary differential equations (4)-(11) and (5)-(13), into stochastic differential equations, which take the form of the two following systems:

$$\begin{cases} \mathrm{d}X_n = A_n e_1 \,\mathrm{d}t, \\ \mathrm{d}A_n = P_{T_{A_n}} \circ \left[\bar{A}_n \mathrm{d}t + 2\sqrt{D}\,\mathrm{d}B_t^{9,n}\right], \end{cases} \tag{15}$$

and

$$\begin{cases} \mathrm{d}X_n = q_n e_1 q_n^* \,\mathrm{d}t, \\ \mathrm{d}q_n = P_{q_n^\perp} \circ \left[\left(\bar{q}_n \otimes \bar{q}_n - \frac{1}{4}\mathrm{I}_4\right) q_n \mathrm{d}t + \sqrt{D/2}\,\mathrm{d}B_t^{4,n}\right], \end{cases} \tag{16}$$

where $(B_t^{9,n})_n$ are matrices of $\mathcal{M}$ with coefficients given by standard independent Brownian motions, and $(B_t^{4,n})_n$ are independent standard Brownian motions on $\mathbb{R}^4$, $D > 0$ representing the noise intensity. The stochastic differential equations have to be understood in the Stratonovich sense, which is well adapted to write stochastic processes on manifolds [24].

Theorem 1 (Equivalence in law [13]). *The processes* (15) *and* (16) *are equivalent in law.*

This theorem relies on the properties of the map Φ defined in (2), in the same way as they are used to prove the equivalence of the alignment dynamics alone in Proposition 8. However in that case the trajectories were exactly the same due to the uniqueness of the solution of the Cauchy problem. Here, since the driving Brownian motions do not belong to the same space (one is on a nine-dimensional space, the other one in a four-dimensional one) we cannot easily give a sense to

some pathwise equivalence. However, the projection of these driving Brownian motions on the tangent space of the manifold we consider produce process which are actually three-dimensional, in the sense that their trajectories are contained in a three-dimensional manifold. This is why the equivalence is at the level of the law of the trajectories. Working on the partial differential equations satisfied by the densities of the laws of the processes, and relying on the equivalence of measures in Prop. 6, we can make further use of the differential properties of Φ to write a link between the divergence and Laplacian operators on $SO_3(\mathbb{R})$ and on $\mathbb{H}_1$. We obtain that these partial differential equation are equivalent, which give the equivalence in law. More precise details on the law of such a stochastic differential equation is given in Subsection 4.1 for the case of a single individual evolving in a given orientation field.

3.3 Alignment model with orientation jumps

In this section we describe an alternative alignment mechanism where the orientations of the particles make jumps at random times. For this, we attach a Poisson point process with parameter 1 to each particle n (for $n = 1, \dots, N$), which corresponds to the times at which this particle updates its orientation. The increasing sequence of positive times will be denoted by $(t_{n,m})_{m \geqslant 1}$, and can be constructed by independent increments between two consecutive times, given by exponential variables of parameter 1. This means that the unit of time has been chosen in order that it corresponds to the average of the time between two jumps of a given particle.

Next we need to define how the orientation (A_n or q_n) of a particle changes when there is a jump. Recall the definition of the averages $\bar{A}_n$ and $\bar{q}_n$ in (10) and (9) respectively. We want the new orientation to be drawn according to a probability "centered" around $\bar{A}_n$ (resp. $\pm\bar{q}_n$) and radially symmetric, that is, it should have a density of the form $A \mapsto M_{\bar{A}_n}(A)$ (resp. $q \mapsto \widetilde{M}_{\bar{q}_n}(q)$), which only depends on the distance between A and $\bar{A}_n$ (resp. the distance between $\pm q$ and $\pm\bar{q}_n$). In the matrix world, the square of the norm of an orthogonal matrix is $\frac{1}{2}\operatorname{Tr}(A^{\mathrm{T}}A) = \frac{3}{2}$, therefore we have $\|A - \bar{A}_n\|^2 = 3 - 2A \cdot \bar{A}_n$, we are thus looking at a probability density only depending on $A \cdot \bar{A}_n$. Thanks to Prop. (4), in the world of quaternions, it corresponds to a function only depending on $(q \cdot \bar{q}_n)^2$.

To fix the ideas, and to see analogies with the gradual alignment model, we will take for $M_{\bar{A}_n}$ the von-Mises distribution centered around $\bar{A}_n$ and with concentration parameter $\frac{1}{D}$. We will indeed see that in this case, the results of the computations for the macroscopic limits that were done in [12,13] can directly be reused. Of course, the method that we present here still applies for a generic smooth function of $A \cdot \bar{A}_n$.

The von-Mises distribution centered in $\Lambda \in SO_3(\mathbb{R})$ and with concentration parameter $\frac{1}{D}$ is defined, for A in $SO_3(\mathbb{R})$, by

$$M_\Lambda(A) = \frac{1}{Z} \exp\left(\frac{1}{D} A \cdot \Lambda\right), \quad \int_{SO_3(\mathbb{R})} M_\Lambda(A)\, dA = 1, \tag{17}$$

where $Z = Z_D < \infty$ is a normalizing constant such that this function is a probability density on $SO_3(\mathbb{R})$.

Analogously, we define the von-Mises distribution on $\mathbb{H}_1$ as

$$M_{\bar{q}}(q) = \frac{1}{Z'} \exp\left(\frac{2}{D}\left((\bar{q}\cdot q)^2 - \frac{1}{4}\right)\right), \qquad \int_{\mathbb{H}_1} M_{\bar{q}}(q)\ dq = 1, \tag{18}$$

where $Z' = Z'_D$ is a normalizing constant. Thanks to Prop. 4 and Prop. 6, if q is a random variable on $\mathbb{H}_1$ distributed according to $M_{\bar{q}}$, then $A = \Phi(q)$ is a random variable on $SO_3(\mathbb{R})$ distributed according to M_Λ, where $\Lambda = \Phi(\bar{q})$.

A useful property of $SO_3(\mathbb{R})$ (or $\mathbb{H}_1$) is that the dot product is invariant by multiplication : we have that $M_{\bar{A}_n}(A) = M_{\mathrm{I}_3}(\bar{A}_n^{\mathrm{T}} A)$, since $\mathrm{I}_3 \cdot (\bar{A}_n^{\mathrm{T}} A) = \bar{A}_n \cdot A$. Furthermore, the measure on $SO_3(\mathbb{R})$ is also left-invariant. We therefore only need to be able to draw random variable according to M_{I_3}, thanks to the following proposition.

Proposition 9. *If $B \in SO_3(\mathbb{R})$ is a random variable distributed according to the density M_{I_3}, then $\bar{A}_n B$ is a random variable distributed according to the density $M_{\bar{A}_n}$.*

Analogously, if $r \in \mathbb{H}_1$ is a random variable distributed according to the density M_1, then $\bar{q}_n r$ is a random variable distributed according to the density $M_{\bar{q}_n}$.

Proof. If U is a measurable set of $SO_3(\mathbb{R})$, then, by left invariance of the measure

$$\begin{aligned}\mathbb{P}(\bar{A}_n B \in U) &= \mathbb{P}(B \in \bar{A}_n^{\mathrm{T}} U) \\ &= \int_{\bar{A}_n^{\mathrm{T}} U} M_{\mathrm{I}_3}(A)\mathrm{d}A = \int_U M_{\mathrm{I}_3}(\bar{A}_n^{\mathrm{T}} A)\mathrm{d}A = \int_U M_{\bar{A}_n}(A)\mathrm{d}A.\end{aligned}$$

Notice that this proof does not rely on the particular expression of the von-Mises distribution, and still applies if $M_{\bar{A}_n}(A)$ is a generic function of $\bar{A}_n \cdot A$. The proof is analogous for the quaternion version. □

We are now ready to construct the stochastic process corresponding to the evolution of positions and orientations of the particles.

Definition 1. *We are given:*

- *a probability density M_{I_3} on $SO_3(\mathbb{R})$, with the property that $M_{\mathrm{I}_3}(A)$ only depends on $\mathrm{I}_3 \cdot A = \frac{1}{2}Tr(A)$ (we will take the von-Mises distribution defined in (17) in the following of the paper),*
- *some independent random variables $S_{n,m} > 0$ and $\eta_{n,m} \in SO_3(\mathbb{R})$, such that for $1 \leqslant n \leqslant N$ and $m \in \mathbb{N}$, $S_{n,m}$ is distributed according to an exponential law of parameter 1 and $\eta_{n,m}$ is distributed according to M_{I_3},*
- *some initial positions $X_{n,0} \in \mathbb{R}^3$ and initial body orientations $A_{n,0} \in SO_3(\mathbb{R})$ for $1 \leqslant n \leqslant N$.*

The variables $(S_{n,m})_{m\in\mathbb{N}}$ represent the intervals of time between consecutive jumps for particle number n. Therefore we define $t_{n,m} = \sum_{0\leqslant \ell < m} S_{n,\ell}$, which

corresponds to the time at which particle number n changes its orientation for the m-th time. The positions and orientations are then defined inductively (almost surely, all times $t_{n,m}$ are distinct) by

$$\begin{cases} X_n(0) = X_{n,0}, \\ X_n(t) = X_n(t_{n,m}) + (t - t_{n,m}) A_n(t) e_1, & \text{if } t \in [t_{n,m}, t_{n,m+1}), \\ A_n(t) = A_{n,0}, & \text{if } t \in [0, t_{n,1}), \\ A_n(t) = \bar{A}_n(t_{n,m}^-)\eta_{n,m}, & \text{if } t \in [t_{n,m}, t_{n,m+1}), m \geqslant 1, \end{cases} \tag{19}$$

where $\bar{A}_n$ is the maximizer of the function $A \mapsto A \cdot \frac{1}{N} \sum_{l=1}^N K(X_l - X_n) A_l$.

Since all the independent random variables $\eta_{n,m}$ are distributed according to a law which has a density with respect to the Lebesgue measure on SO_3, and the set of configurations for which this maximizer $\bar{A}_n$ is not well defined are included in low dimensional manifolds (compared to the configuration space), we expect that this process is almost surely well defined. We do not give a detailed proof of this fact here since we are interested in derivation of kinetic models which will share the same issues, therefore we will focus on the formal derivation of these model in the case where this maximizer is well defined everywhere. The rigorous treatment of this issue is outside the scope of these lecture notes. It is even far from being well understood, even in the case of the Vicsek model, for which the only bad configurations are those with a zero average velocity. At the kinetic level, the only known global existence of solutions requires very strong assumptions of non-vanishing average velocity (which are not only assumptions on the initial conditions) [23].

Analogously, we can define this process in the world of quaternions.

Definition 2. *We are given:*

- *a probability density M_1 on $\mathbb{H}_1$, with the property that $M_1(q)$ only depends on $(1 \cdot q)^2 = \mathrm{Re}(q)^2$ (we will take the von-Mises distribution defined in* (18) *in the following of the paper),*
- *some independent random variables $S_{n,m} > 0$ and $\eta_{n,m} \in \mathbb{H}_1$, such that for $1 \leqslant n \leqslant N$ and $m \in \mathbb{N}$, $S_{n,m}$ is distributed according to an exponential law of parameter 1 and $\eta_{n,m}$ is distributed according to M_1,*
- *some initial positions $X_{n,0} \in \mathbb{R}^3$ and initial body orientations $q_{n,0} \in \mathbb{H}_1$ for $1 \leqslant n \leqslant N$.*

Again, we define $t_{n,m} = \sum_{0 \leqslant \ell < m} S_{n,\ell}$, which corresponds to the time at which particle number n changes its orientation for the m-th time. The positions and orientations are then defined inductively (almost surely, all times $t_{n,m}$ are distinct) by

$$\begin{cases} X_n(0) = X_{n,0}, \\ X_n(t) = X_n(t_{n,m}) + (t - t_{n,m})\, q_n e_1 q_n^*, & \text{if } t \in [t_{n,m}, t_{n,m+1}), \\ q_n(t) = q_{n,0}, & \text{if } t \in [0, t_{n,1}), \\ q_n(t) = \bar{q}_n(t_{n,m}^-)\eta_{n,m}, & \text{if } t \in [t_{n,m}, t_{n,m+1}), m \geqslant 1, \end{cases} \tag{20}$$

where $\bar{q}_n$ is defined in (8)-(9).

Once again, we expect this process to be defined almost surely, and as we remarked, thanks to Prop. 4, these two definitions give rise to processes which are equivalent in law, through the map Φ. A last remark is that these processes are a particular case of Piecewise Deterministic Markov Processes (PDMP's): between two jumps, the configuration follows an Ordinary Differential Equation (which in our case is nothing else than free transport). More comments on PDMP's will be made in Subsection 4.2.

4 Derivation of kinetic models

The aim of this section is to present a heuristic derivation of kinetic models corresponding to the limit of the particle systems when the number of particles is large. We present this derivation in the framework of rotation matrices, and we will give the corresponding kinetic models in the framework of quaternions at the end of this section.

To this aim, we introduce the so-called empirical distribution f^N of the particles as the measure

$$f^N(x, A, t) = \frac{1}{N} \sum_{i=1}^{N} \delta_{X_i(t)}(x) \otimes \delta_{A_i(t)}(A),$$

that is to say that if φ is a continuous and bounded function from $\mathbb{R}^3 \times SO_3(\mathbb{R})$ to $\mathbb{R}$, the integral of φ with respect to this measure (at time t) is given by

$$\int_{\mathbb{R}^3 \times SO_3(\mathbb{R})} \varphi(x, A) f^N(x, A, t) \mathrm{d}x \, \mathrm{d}A = \frac{1}{N} \sum_{i=1}^{N} \varphi(X_i(t), A_i(t)). \tag{21}$$

This function is independent of the change of numbering of particles, we say that particles are indistinguishable.

Notice that the average orientation $\bar{A}_n$ defined in (10) can be constructed through the empirical distribution : if we define, for a given probability density f, the functions J_f^K and Λ_f^K by

$$J_f^K(x) = \int_{\mathbb{R}^3 \times SO_3(\mathbb{R})} K(x - y) A \, f(y, A) \, \mathrm{d}y \, \mathrm{d}A, \tag{22}$$

$$\Lambda_f^K(x) \text{ is a maximizer on } SO_3(\mathbb{R}) \text{ of } A \mapsto A \cdot J_f^K, \tag{23}$$

we get that the definition (6) can be written as $\bar{J}_n = J_{f^N}^K(X_n)$. And therefore we get $\bar{A}_n = \Lambda_{f^N}^K(X_n)$. Therefore we obtain that the interaction between particles (which is only due to this target orientation $\bar{A}_n$) corresponds to an interaction, for each particle, with the field generated by the empirical distribution f^N. The type of limit we want to understand is called mean-field limit: when the number of particles is large, correlations between finite numbers of particles tend to

vanish, and a kind of law of large numbers gives that the empirical distribution is well approached by the law of one single particle. This phenomenon is linked to the notion of propagation of chaos, and we refer to [27] for an introduction. This type of limit has been rigorously shown to be valid in various models of collective behavior, such as [3] in a regularized Vicsek model, and [2,5] in cases with less regularity. In our model, it is not straightforward to apply this strategy (due to the regularity issues for the definition of Λ_f^K), therefore we only present a heuristic derivation of the mean-field limit one would obtain if the empirical distribution f^N converges to the law of one single particle when N is large.

Let us now focus for the moment on a single particle model aligning with a given "target field" $\Lambda(x,t) \in SO_3(\mathbb{R})$, and subject to some noise, as in the models given in the previous section. This corresponds to replacing $\bar{A}_n(t)$ by $\Lambda(X_n,t)$ in the models given by (15) (for the gradual alignment model) and (19) (for the alignment model by orientation jumps).

4.1 Gradual alignment of a single individual in an orientation field

We then consider the following stochastic differential equation, for the evolution of a particle at position X_t and body orientation A_t, in an orientation field $\Lambda(x,t) \in SO_3(\mathbb{R})$:

$$\begin{cases} \mathrm{d}X_t = A_t e_1 \,\mathrm{d}t, \\ \mathrm{d}A_t = P_{T_{A_t}} \Lambda(X_t,t)\,\mathrm{d}t + 2\sqrt{D}\, P_{T_{A_t}} \circ \mathrm{d}B_t^9, \end{cases} \tag{24}$$

where B_t^9 is a matrix with independent coefficients given by 9 standard Brownian motions on $\mathbb{R}$, and the $\circ$ indicates that this has to be understood in the Stratonovich sense. Let us see how this last fact ensures that the orientation $A(t)$ stays on $SO_3(\mathbb{R})$. Thanks to the classical chain rule satisfied by Stratonovich SDE's [24], for a smooth function $\varphi(x,A)$, we have

$$\begin{aligned} \varphi(X_t,A_t) =& \varphi(X_0,A_0) \\ &+ \int_0^t \left(\mathrm{D}_x\varphi(X_s,A_s)[A_s e_1] + \mathrm{D}_A\varphi(X_s,A_s)[P_{T_{A_s}}\Lambda(X_s,s)]\right)\mathrm{d}s \\ &+ 2\sqrt{D}\int_0^t \left(\mathrm{D}_A\varphi(X_s,A_s)[P_{T_{A_s}}(\cdot)]\right)\circ \mathrm{d}B_s^9, \end{aligned} \tag{25}$$

where D_x and D_A are the differentials with respect to $x \in \mathbb{R}^3$ and $A \in \mathcal{M}$. Now, if we take $\varphi(x,A) = A^{\mathrm{T}}A - \mathrm{I}_3$, we get that $\mathrm{D}_A\varphi(x,A)[H] = A^{\mathrm{T}}H + H^{\mathrm{T}}A$. Thanks to the formula (1), we then get that the linear operator $\mathrm{D}_A\varphi(x,A)[P_{T_A}(\cdot)]$ (from $\mathcal{M}$ to $\mathcal{M}$) is given by

$$\begin{aligned} \mathrm{D}_A\varphi(X,A)[P_{T_A}H] &= \frac{1}{2}A^{\mathrm{T}}(H - AH^{\mathrm{T}}A) + \frac{1}{2}(H^{\mathrm{T}} - A^{\mathrm{T}}HA^{\mathrm{T}})A \\ &= \frac{1}{2}(A^{\mathrm{T}}H\varphi(x,A) - \varphi(x,A)H^{\mathrm{T}}A). \end{aligned}$$

Therefore, if we define the linear operator $L(Y,t) : H \mapsto \frac{1}{2}(A_t^{\mathrm{T}} H Y - Y H^{\mathrm{T}} A_t)$ and the process $Y_t = \varphi(X_t, A_t) = A_t^{\mathrm{T}} A_t - \mathrm{I}_3$, Eq. (25) becomes

$$Y_t = Y_0 + \int_0^t L(Y_s, s)[\Lambda(X_s, s)]\mathrm{d}s + 2\sqrt{D}\int_0^t L(Y_s, s) \circ \mathrm{d}B_s^9,$$

hence the process Y_t satisfies a linear SDE with initial condition 0, therefore it is 0 for all time, which means that A_t stays in $SO_3(\mathbb{R})$ for all time.

Moreover, it is shown in Chapter 3 of [24] that for a manifold $\mathcal{N}$ embedded in the euclidean space $\mathbb{R}^d$, the generator of the SDE equation $\mathrm{d}Z_t = \sigma P_{T_{Z_t}} \circ \mathrm{d}B_t^d$ (where P_{T_y} is the orthogonal projection on the tangent space T_y of $\mathcal{N}$ at y) is $\frac{\sigma^2}{2}\Delta_{\mathcal{N}}$, where $\Delta_{\mathcal{N}}$ is the Laplace-Beltrami operator on $\mathcal{N}$ (the solution of this SDE is called Brownian motion on $\mathcal{N}$). This means that for a smooth function φ on $\mathcal{N}$,

$$\mathbb{E}[\varphi(Z_t)] = \mathbb{E}[\varphi(Z_0)] + \frac{\sigma^2}{2}\mathbb{E}\left[\int_0^t \Delta_{\mathcal{N}}\varphi(Z_s)\,\mathrm{d}s\right].$$

Using the Stratonovich chain rule, this means that

$$\mathbb{E}\left[\sigma\int_0^t \mathrm{D}_Z\varphi(Z_t)[P_{T_{Z_t}}(\cdot)] \circ \mathrm{d}B_t^d\right] = \frac{\sigma^2}{2}\mathbb{E}\left[\int_0^t \Delta_{\mathcal{N}}\varphi(Z_s)\,\mathrm{d}s\right].$$

In our case, if we write $\mathcal{N} = SO_3(\mathbb{R})$, with the metric induced by the euclidean metric in $\mathbb{R}^9$, this would mean that the expectation of the last term of (25) is $2D\int_0^t \Delta_{\mathcal{N}}\varphi(X_s, A_s)\mathrm{d}s$. However, the metric we used for $SO_3(\mathbb{R})$ is induced by the dot product $(A, B) \mapsto \frac{1}{2}\operatorname{Tr}(A^T B)$, which is half of what corresponds to the euclidean dot product in $\mathbb{R}^9$. The Riemannian metric is then divided by 2, and the formula for the Laplace-Beltrami operator gives that it is then multiplied by 2 (recall the condensed form $\Delta_g\varphi = \frac{1}{\sqrt{|g|}}\partial_i(\sqrt{|g|}g^{ij}\partial^i\varphi)$, where g^{ij} are the coefficients of the inverse of the metric tensor $(g_{ij})_{i,j}$). Therefore we get $\Delta_A\varphi = 2\Delta_{\mathcal{N}}\varphi$. Finally, for an arbitrary test function φ with values in $\mathbb{R}$, taking the expectation in Eq. (25), and using the gradient formulation instead of the differential, we get

$$\begin{aligned}\mathbb{E}[\varphi(X_t, A_t)] &= \mathbb{E}[\varphi(X_0, A_0)] \\ &+\mathbb{E}\Bigg[\int_0^t [\nabla_x\varphi(X_s, A_s)\cdot A_s e_1 + \nabla_A\varphi(X_s, A_s)\cdot P_{T_{A_s}}\Lambda(X_s, s) \\ &\qquad + D\Delta_A\varphi(X_s, A_s)]\mathrm{d}s\Bigg].\end{aligned} \tag{26}$$

Finally, we denote by $f(x, A, t)$ the law of such a particle at time t, which is defined by the formula $\mathbb{E}[\varphi(X_t, A_t)] = \int_{\mathbb{R}\times SO_3(\mathbb{R})}\varphi(x, a)f(x, A, t)\mathrm{d}A\mathrm{d}x$. Then, the fact that Eq. (26) holds for any test function φ, corresponds exactly to the fact that f is a weak solution of the following linear evolution equation:

$$\partial_t f + (Ae_1)\cdot\nabla_x f = -\nabla_A\cdot(P_{T_A}\Lambda f) + D\Delta_A f. \tag{27}$$

4.2 Alignment by orientation jumps, for a single individual in a field

We now turn to the model of alignment by orientation jumps. We then consider the following process: given some independent random variables $S_m > 0$ and $\eta_m \in SO_3(\mathbb{R})$, such that for $m \in \mathbb{N}$, S_m is distributed according to an exponential law of parameter 1 and η_m is distributed according to M_{I_3}, an initial position $X_0 \in \mathbb{R}^3$ and orientation $A_0 \in SO_3(\mathbb{R})$, we define $t_m = \sum_{0 \leqslant \ell < m} S_\ell$, and the position and orientation at time t are then defined inductively (almost surely, all times t_m are distinct) by

$$\begin{cases} X_t = X_{t_m} + (t - t_m) A_t e_1, & \text{if } t \in [t_m, t_{m+1}), \\ A_t = A_0, & \text{if } t \in [0, t_1), \\ A_t = \Lambda(X_{t_m}, t_m^-)\eta_m, & \text{if } t \in [t_m, t_{m+1}), \text{ with } m \geqslant 1, \end{cases} \tag{28}$$

Another way to describe this process (X_t, A_t) is to say that it is a (non autonomous) Piecewise Deterministic Markov Process (PDMP) with jump rate 1, with flow ϕ given by $\phi((X, A), t) = (X + tAe_1, A)$ and with transition measure $Q_t((X, A), \cdot) = \delta_X \otimes M_{\Lambda(X,t)}$. The only difference with classical description of PDMP's (see for instance [1] for a review of recent results), except from the fact that we work on a manifold rather than an open set of $\mathbb{R}^d$, is that the transition measure depends on time.

Let us explain how to derive the evolution equation for the law of the process (X_t, A_t). We take once again a smooth test function $\varphi(x, A)$, we fix a small time interval δt, and we evaluate the expectation of $\varphi(X_{t+\delta t}, A_{t+\delta t})$. With probability $1 - \delta t + o(\delta t)$, there is no jump in $(t, t + \delta t)$ and therefore $A_{t+\delta t} = A_t$, and $X_{t+\delta t} = X_t + \delta t A_t e_1$. With probability $\delta t + o(\delta t)$, there is exactly one jump at time s in $(t, o(\delta t))$, and therefore $A_{t+\delta t} = A_s$ which follows the distribution $M_{\Lambda(X_s,s)}$. Of course we have $(t - s) = o(1)$. Finally, there are two or more jumps in $(t, t + \delta t)$ with probability $o(\delta t)$. We therefore get

$$\begin{aligned} \mathbb{E}[\varphi(X_{t+\delta t}, A_{t+\delta t})] &= (1 - \delta t + o(\delta t))\mathbb{E}[\varphi(X_t + \delta t A_t e_1, A_t)] \\ &\quad + \delta t\, \mathbb{E}\left[\int_{SO_3(\mathbb{R})} \varphi(X_t + o(1), A') M_{\Lambda(X_t + o(1), t + o(1))}(A') \mathrm{d}A'\right] + o(\delta t), \end{aligned}$$

which gives, by smoothness of φ, and if we assume that Λ and $\Lambda \mapsto M_\Lambda$ are smooth enough, that

$$\begin{aligned} \frac{1}{\delta t}\Big(\mathbb{E}[\varphi(X_{t+\delta t}, A_{t+\delta t})] - \mathbb{E}[\varphi(X_t, A_t)]\Big) &= \mathbb{E}[\nabla_x \varphi(X_t, A_t) \cdot A_t e_1] - \mathbb{E}[\varphi(X_t, A_t)] \\ &\quad + \mathbb{E}\left[\int_{SO_3(\mathbb{R})} \varphi(X_t, A') M_{\Lambda(X_t, t)}(A') \mathrm{d}A'\right] + o(1), \end{aligned}$$

that is to say

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[\varphi(X_t,A_t)] = \mathbb{E}\Bigg[\nabla_x\varphi(X_t,A_t)\cdot A_t e_1 - \varphi(X_t,A_t) + \int_{SO_3(\mathbb{R})}\varphi(X_t,A')M_{\Lambda(X_t,t)}(A')\mathrm{d}A'\Bigg]. \tag{29}$$

Finally, as in the previous subsection, we denote by $f(x,A,t)$ the law of such a particle at time t, defined by $\mathbb{E}[\varphi(X_t,A_t)] = \int_{\mathbb{R}\times SO_3(\mathbb{R})}\varphi(x,a)f(x,A,t)\mathrm{d}A\mathrm{d}x$. Now, the fact that Eq. (29) holds for any test function φ corresponds exactly to the fact that f is a weak solution of the following linear evolution equation:

$$\partial_t f + (Ae_1)\cdot\nabla_x f = \rho_f M_\Lambda - f, \tag{30}$$

where

$$\rho_f(x,t) = \int_{SO_3(\mathbb{R})} f(x,A,t)\mathrm{d}A. \tag{31}$$

4.3 Kinetic mean-field models of alignment

Let us summarize the results of the two previous subsections: for the evolution of a particle in an orientation field $\Lambda(x,t)$ according to one of the models (24) or (28), the law f of the particle is evolving according to one of the (linear) kinetic equations (27) or (30) which is of the form:

$$\partial_t f + (Ae_1)\cdot\nabla_x f = \Gamma_\Lambda(f), \tag{32}$$

with

$$\Gamma_\Lambda(f) = \begin{cases} -\nabla_A\cdot(P_{T_A}\Lambda f) + D\Delta_A f & \text{in the gradual alignment model,} \\ (\rho_f M_\Lambda - f) & \text{in the jump model.} \end{cases} \tag{33}$$

We are now ready to provide a formal derivation of the equation satisfied by the law of one particle in the limit of a large number of particles. The heuristic is as follows: if we consider that the empirical distribution f^N of the particles converges to a deterministic law f, either for the gradual alignment process (15) or for the model with orientation jumps (19), then each particle will evolve in the limit $N\to\infty$ as a single particle in a orientation field $\Lambda(x,t)$ corresponding to $\Lambda^K_{f(t,\cdot)}(x)$, given by the formulas (22)-(23). Therefore the evolution of the law of one particle, in the limit $N\to\infty$, is governed by the evolution equation (32) where Λ is replaced by Λ^K_f. This gives the following (now non-linear and non-local) evolution equation:

$$\partial_t f + (Ae_1)\cdot\nabla_x f = \Gamma_{\Lambda^K_f}(f), \tag{34}$$

where $\Gamma_\Lambda(f)$ is defined above in (33), and with

$$J^K_f(x) = \int_{\mathbb{R}^3\times SO_3(\mathbb{R})} K(x-y)A\, f(y,A)\,\mathrm{d}y\,\mathrm{d}A,$$
$$\Lambda^K_f(x) \text{ maximizes } A\mapsto A\cdot J^K_f(x) \text{ on } SO_3(\mathbb{R}).$$

This heuristic can be made rigorous if the map $f \mapsto \Lambda_f$ is regular (which is not the case here, as there are configurations for which it is not even well defined). Let us quickly present the coupling method (see for instance [27]) to understand how we can indeed use the law of large numbers for independent processes. The idea is first to construct a nonlinear process (for one single particle) which is the natural limit of the evolution of one particle in the particle system corresponding to (15) or (19), and for which the law is following the evolution equation (34). For the gradual alignment process, given a Brownian motion B_t^9 as in (24), it would be defined as follows

$$\begin{cases} \mathrm{d}\overline{X}_t = \overline{A}_t e_1 \,\mathrm{d}t, \\ \mathrm{d}\overline{A}_t = P_{T_{\overline{A}_t}} \Lambda_{f(t,\cdot)}(\overline{X}_t)\,\mathrm{d}t + 2\sqrt{D}\, P_{T_{\overline{A}_t}} \circ \mathrm{d}B_t^9, \\ f(t,\cdot) \text{ is the law of } (\overline{X}_t, \overline{A}_t). \end{cases} \tag{35}$$

For the orientation by jumps, given random variables t_m and η_m as in (28), it would be given by

$$\begin{cases} \overline{X}_t = \overline{X}_{t_m} + (t - t_m)\overline{A}_t e_1, & \text{if } t \in [t_m, t_{m+1}), \\ \overline{A}_t = A_0, & \text{if } t \in [0, t_1), \\ \overline{A}_t = \Lambda_{f(t_m,\cdot)}(\overline{X}_{t_m})\eta_m, & \text{if } t \in [t_m, t_{m+1}), \text{ with } m \geqslant 1, \\ f(t,\cdot) \text{ is the law of } (\overline{X}_t, \overline{A}_t). \end{cases} \tag{36}$$

These constructions can be seen as fixed point problems for the laws of the trajectories, and this is where the regularity of $f \mapsto \Lambda_f$ can be used to prove a contraction property in an appropriate space. Once these processes are well defined, the second idea of the method of couplings is to introduce N of these processes (35) or (36), for which the initial conditions and random variables (Brownian motions B_t^9, n or jump times $t_{n,m}$ and rotations $\eta_{n,m}$) are the same as for the particle systems (15) and (19). By construction, these auxiliary nonlinear processes $(\overline{X}_n(t), \overline{A}_n(t))$ are independent and identically distributed according to the law f, solution of the kinetic equation (34). Therefore the last step of the coupling method is to perform estimates of the differences between the trajectories of the particle system and of the auxiliary process in order to let appear quantities reminiscent of (21), but of the form

$$\frac{1}{N}\sum_{i=1}^{N} \varphi(\overline{X}_i(t), \overline{A}_i(t)), \tag{37}$$

for which the law of large numbers applies.

Let us finish this subsection by presenting the kinetic equation we obtain (exactly in the same manner) when working with unit quaternions instead of rotation matrices. The formal mean-field limit of the particle system (16) or (20) is given by the following evolution equation, for the density $f(t, x, q)$ of finding a particle at position x with orientation given by the unit quaternion q at time t:

$$\partial_t f + \Phi(q)(e_1) \cdot \nabla_x f = \Gamma_{\overline{q}_f^K} f, \tag{38}$$

with

$$\Gamma_{\overline{q}}(f) = \begin{cases} -\nabla_q \cdot (P_{q^\perp}(\overline{q} \otimes \overline{q}) q\, f) + \frac{D}{4} \Delta_q f & \text{in the gradual alignment model,} \\ (\rho_f M_{\overline{q}} - f) & \text{in the jump model,} \end{cases} \tag{39}$$

where

$$\rho_f = \int_{\mathbb{H}_1} f(q) \mathrm{d}q, \tag{40}$$

$$Q_f^K(x) = \int_{\mathbb{R}^3 \times \mathbb{H}_1} K(x-y)(q \otimes q - \tfrac{1}{4}\mathrm{I}_4) f(y,q) \mathrm{d}y\, \mathrm{d}q, \tag{41}$$

$$\overline{q}_f^K(x) \text{ is an eigenvector of } Q_f^K(x) \text{ of maximal eigenvalue.} \tag{42}$$

5 Macroscopic limit

In this section we derive the macroscopic dynamics for the kinetic equations (34) and (38). This means that we are interested in the dynamics in the large time as well as large-space scale. For this we first introduce a scaling with respect to a small parameter ε. We then determine the local equilibria of the collision operator, which depend on two macroscopic quantities, a density ρ and a local orientation Λ. The final step is then the derivation of the evolution equations of these macroscopic functions ρ and Λ. The first one comes from the conservation of mass (a collision invariant), and the second one needs more work, and can be derived using the concept of Generalized Collisional Invariants introduced in [16]. The subsection presenting this concept and how to use it to obtain the evolution equation for Λ is the main part of this section.

5.1 Scaling

We introduce the macroscopic temporal and spatial variables (t', x') given by

$$t' = \varepsilon t, \quad x' = \varepsilon x,$$

where $0 < \varepsilon \ll 1$ is a scale parameter. We also consider the following rescaling for the interaction kernel:

$$K_\varepsilon(x) = \frac{1}{\varepsilon^3} K\left(\frac{x}{\varepsilon}\right).$$

This corresponds to localized interactions as $\varepsilon \to 0$ (see Rem. 1 below). Notice that

$$\int_{\mathbb{R}^3} K_\varepsilon(x)\, dx = \int_{\mathbb{R}^3} K(x)\, dx = 1.$$

Define the function f_ε in the macroscopic variables as

$$f_\varepsilon(t', x', A) = f(t, x, A).$$

Our goal is to determine the dynamics for this function as $\varepsilon \to 0$. Firstly, one can check that the evolution equation for f_ε is given by

$$\varepsilon(\partial_t f_\varepsilon + (Ae_1)\cdot\nabla_x f_\varepsilon) = \Gamma_{\Lambda_{f_\varepsilon}^{K_\varepsilon}}(f_\varepsilon), \qquad \text{(matrix formulation)}, \tag{43}$$

$$\varepsilon(\partial_t f_\varepsilon + \Phi(q)(e_1)\cdot\nabla_x f_\varepsilon) = \Gamma_{\bar{q}_{f_\varepsilon}^{K_\varepsilon}}(f_\varepsilon), \qquad \text{(quaternion formulation)}, \tag{44}$$

where the primes have been skipped. We recall that the definition of the operators Γ_Λ and $\Gamma_{\bar{q}}$ are given in (33) and (39) respectively, and the definitions of the average orientations $\Lambda_{f_\varepsilon}^{K_\varepsilon}$ and $\bar{q}_{f_\varepsilon}^{K_\varepsilon}$ are given in (22)–(23) and (41)–(42) respectively.

Next, we expand the collision operators in the parameter ε.

Lemma 1 (Expansion for localized interactions). *The following expansion holds:*

$$J_f^{K_\varepsilon} = K_\varepsilon *_x J_f = J_f + \mathcal{O}(\varepsilon^2),$$

where $J_f(x)$ takes in account the dependence of f on the variable A only:

$$J_f(x) = \int_{SO_3(\mathbb{R})} A\, f(x,A)\,\mathrm{d}A. \tag{45}$$

Consequently, we can recast Eq. (43) *as*

$$\varepsilon(\partial_t f_\varepsilon + (Ae_1)\cdot\nabla_x f_\varepsilon) = \Gamma_{\Lambda_{f_\varepsilon}}(f_\varepsilon) + \mathcal{O}(\varepsilon^2), \tag{46}$$

where

$$\Lambda_f \text{ maximizes } A \mapsto A\cdot J_f \text{ on } SO_3(\mathbb{R}). \tag{47}$$

Analogously, we have that

$$Q_f^{K_\varepsilon} = K_\varepsilon *_x Q_f = Q_f + \mathcal{O}(\varepsilon^2),$$

where

$$Q_f(x) = \int_{\mathbb{R}^3\times\mathbb{H}_1} (q\otimes q - \tfrac{1}{4}\mathrm{I}_4) f(x,q)\,\mathrm{d}q, \tag{48}$$

and Eq. (44) *is recast as*

$$\varepsilon(\partial_t f_\varepsilon + \Phi(q)(e_1)\cdot\nabla_x f_\varepsilon) = \Gamma_{\bar{q}_{f_\varepsilon}} f_\varepsilon + \mathcal{O}(\varepsilon^2), \tag{49}$$

where

$$\bar{q}_f \text{ is an eigenvector of } Q_f \text{ of maximal eigenvalue.} \tag{50}$$

Remark 1 (Localized interactions). Notice that in the leading order of the expansion of K_ε, we obtain a delta distribution in x. This is why we say that this kind of rescaling corresponds to localized interactions in the limit $\varepsilon \to 0$.

The proof of the expansions in Lem. 1 is straightforward using the Taylor expansion and the fact that

$$\int_{\mathbb{R}^3} xK(x)\ dx = 0,$$

for more details on this and the fact that Λ_f and $\bar{q}_f$ are indeed defined as in the lemma, the reader is referred to [12, Lem. 4.1] and [13, Lem. 4.2, Prop. 4.3], respectively.

Our goal is to investigate the limit of f_ε as $\varepsilon \to 0$. Firstly, notice that, formally, from Eq. (46) we have that

$$\text{if } f_\varepsilon(t,x,\cdot) \to f_0(t,x,\cdot) \text{ as } \varepsilon \to 0 \text{ then } f_0(t,x,\cdot) \in \ker(\Gamma_{\Lambda_{f_0(t,x,\cdot)}}), \tag{51}$$

in the matrix formulation, or

$$\text{if } f_\varepsilon(t,x,\cdot) \to f_0(t,x,\cdot) \text{ as } \varepsilon \to 0 \text{ then } f_0(t,x,\cdot) \in \ker(\Gamma_{\bar{q}_{f_0(t,x,\cdot)}}), \tag{52}$$

in the quaternion formulation. For this reason, we study next the kernel of the operator Γ_Λ (which is an operator acting on functions of $A \in SO_3(\mathbb{R})$ only) for a fixed $\Lambda \in SO_3(\mathbb{R})$ (analogously $\Gamma_{\bar{q}}$ for fixed $\bar{q} \in \mathbb{H}_1$) in the following section.

5.2 Study of the collision operator Γ

The goal of this subsection is to show that both the jump model and the gradual alignment model have the same type of equilibria. More precisely, we show the following proposition.

Proposition 10 (Equilibria, matrix formulation). *Recall the definition of the operator Γ_Λ in Eq.* (33)*, and the definition of the von Mises distribution M_Λ in Eq.* (17)*. Then, for any $f \geq 0$, we have*

$$\Gamma_{\Lambda_f}(f) = 0 \iff f = \rho M_\Lambda, \text{ for some } \rho \geq 0, \Lambda \in SO_3(\mathbb{R}). \tag{53}$$

Furthermore, any element f of the form $f = \rho M_\Lambda$, with $\rho \geq 0$ and $\Lambda \in SO_3(\mathbb{R})$ satisfies the consistency relations

$$\rho_f = \rho, \quad J_f = \rho c_1 \Lambda, \quad \Lambda_f = \Lambda, \tag{54}$$

where $c_1 \in (0,1)$ is an explicit constant, and ρ_f, J_f and Λ_f are defined in Eq. (31)*, Eq.* (45)*, and Eq.* (47)*, respectively.*

As a consequence, both the gradual alignment model and the jump model have the same equilibria, and therefore, the same type of (formal) limit as $\varepsilon \to 0$: we can write

$$f_0(t,x,A) = \rho(t,x) M_{\Lambda(t,x)}(A),$$

for some $\rho(t,x) \geq 0$ and $\Lambda \in SO_3(\mathbb{R})$ satisfying furthermore the consistency relations

$$\rho_{f_0}(t,x) = \rho(t,x), \quad J_{f_0}(t,x) = \rho(t,x) c_1 \Lambda(t,x), \quad \Lambda_{f_0}(t,x) = \Lambda(t,x). \tag{55}$$

Remark 2 (Variants in the jump-based model). In the jump-based model, the result holds true formally if we replace (in the collision operator and in the result) the von-Mises distribution by any distribution. Therefore, the jump-based model can reproduce a great variety of behaviors in terms of equilibria.

The proof of Prop. 10 is done in [12] in the case of the gradual alignment model. We summarize here the main ideas.

We first prove the consistency relations. Let us take a function f of the form

$$f = \rho M_\Lambda, \tag{56}$$

for some $\rho \geq 0$ and $\Lambda \in SO_3(\mathbb{R})$. Now, one can check by direct computation that the following consistency relation holds for the average of M_Λ (see proof in [12, Lem. 4.4]):

$$\int_{SO_3(\mathbb{R})} A\, M_\Lambda(A)\, dA = c_1 \Lambda, \text{ for some } c_1 \in (0,1) \text{ explicit.}$$

With this, integrating expression (56) against 1 and A in $SO_3(\mathbb{R})$ we obtain the two first equalities of Eq. (54). To conclude the proof of Eq. (54), the last equality is a consequence of the second one.

Now, in the gradual alignment model, it is proved in [12] that the operator Γ_{Λ_f} can be recast as

$$\Gamma_{\Lambda_f}(f) = D\nabla_A \cdot \left(M_{\Lambda_f} \nabla_A \left(\frac{f}{M_{\Lambda_f}} \right) \right). \tag{57}$$

Using expression (57), one can obtain that

$$\ker(\Gamma_{\Lambda_f}) = \{\rho M_{\Lambda_f} \text{ for any } \rho = \rho(t,x)\}$$

(see detailed proof of this statement in [12, Lem. 4.3]), which, thanks to the consistency relations proved before, is equivalent to Eq. (53).

In the case of the jump model, from the definition of the operator Γ_{Λ_f} it is straightforward that its kernel is given by the functions f such that

$$f = \rho_f M_{\Lambda_f},$$

that is, since we have taken M_Λ to be the von-Mises distribution, and using again the consistency relations, exactly Eq. (53).

Therefore, for both models, we can use Eq. (51) to see that the limit f_0 must be of the form

$$f_0(t,x,A) = \rho(t,x) M_{\Lambda(t,x)}(A), \tag{58}$$

for some $\rho = \rho(t,x)$ and $\Lambda = \Lambda(t,x) \in SO_3(\mathbb{R})$ to be determined, and which satisfy the consistency relations.

Analogously, we obtain the same kind of results for the formulation with quaternions. We write only the result on the limiting function.

Lemma 2 (Equilibria, quaternion formulation). *Recall the definition of the operator $\Gamma_{\bar{q}}$ in Eq.* (39) *and of the von Mises distribution $M_{\bar{q}}$ in Eq.* (18)*. Then, both the gradual alignment model and the jump model have the same equilibria, and therefore, the same type of limit as $\varepsilon \to 0$: we can write*

$$f_0(t,x,q) = \rho(t,x) M_{\bar{q}(t,x)}(q),$$

for some $\rho(t,x) \geq 0$ and $\bar{q}(t,x) \in \mathbb{H}_1$ satisfying furthermore the consistency relations

$$\rho_{f_0}(t,x) = \rho(t,x), \quad \bar{q}_{f_0}(t,x) = \bar{q}(t,x), \tag{59}$$

where $c_1 \in (0,1)$ is an explicit constant, and ρ_f and $\bar{q}_f$ are defined in Eq. (40) *and Eq.* (50)*, respectively.*

The proof of this proposition is done in [13] in the case of the gradual alignment model. We only recall the main ideas here. First, the consistency relations rely on the consistency relation satisfied by the von-Mises distribution on $\mathbb{H}_1$, which is (see [13, Prop 4.4]):

$$\text{the leading eigenvector of } \int_{\mathbb{H}_1} (q \otimes q - \frac{1}{4} I_4) M_{\bar{q}} \, dq \text{ corresponds to } \bar{q}. \tag{60}$$

Therefore, if we take any f of the form

$$f = \rho M_{\bar{q}},$$

multiplying this expression by 1 and $(q \otimes q - \frac{1}{4}\mathrm{I}_4)$ and integrating on $\mathbb{H}_1$ we have that

$$\rho_f = \rho, \qquad Q_f = \rho_f \int_{\mathbb{H}_1} (q \otimes q - \frac{1}{4} I_4) M_{\bar{q}} \, dq,$$

where Q_f is defined in Eq. (48). As a consequence of the last equality,

$$\bar{q}_f = \bar{q}.$$

To compute the kernel of the collision operator, in the gradual alignment model we use that the collision operator can be recast as

$$\Gamma_{\bar{q}_f}(f) = \frac{D}{4} \nabla_q \cdot \left(M_{\bar{q}_f} \nabla_q \left(\frac{f}{M_{\bar{q}_f}} \right) \right),$$

(proved in [13]). In the jump-based model the computation of the kernel is straightforward.

We then use Eq. (52) to conclude the proposition.

In summary, we have seen that, formally, the limit of f_ε will be of the form ρM_Λ (or $\bar{\rho} M_{\bar{q}}$ for the quaternion case). We are left with determining the dynamics of the functions $\rho = \rho(t,x)$, $\bar{\rho} = \bar{\rho}(t,x)$, $\Lambda = \Lambda(t,x)$ and $\bar{q} = \bar{q}(t,x)$ (macroscopic quantities). This is done in the following section.

5.3 The equation for the density ρ

We first compute the evolution for the density $\rho = \rho(t, x)$. We integrate the rescaled kinetic equation (46) over $SO_3(\mathbb{R})$ and divide by ε to obtain

$$\partial_t \left(\int_{SO_3(\mathbb{R})} f_\varepsilon \, dA \right) + \nabla_x \cdot \left(\int_{SO_3(\mathbb{R})} Ae_1 \, f_\varepsilon \, dA \right) = 0.$$

Importantly, the right-hand side has vanished in the integration. This cancellation reflects the fact that the total mass is conserved, i.e., the number of particles is preserved through the interactions. Now we can take formally the limit $\varepsilon \to 0$ and since we know the limit of f_ε in Eq. (58) we have that

$$\partial_t \left(\int_{SO_3(\mathbb{R})} \rho M_\Lambda(A) \, dA \right) + \nabla_x \cdot \left(\int_{SO_3(\mathbb{R})} (Ae_1) \, \rho M_\Lambda(A) \, dA \right) = 0,$$

which corresponds to

$$\partial_t \rho + c_1 \nabla_x \cdot (\rho \Lambda e_1) = 0, \tag{61}$$

given the consistency relations (55). The equation for the density ρ in (61) corresponds to the continuity equation: the density of particles is transported with a velocity equal to $c_1 \Lambda e_1$.

In the formulation with quaternions, analogous computations give the same equation for $\bar{\rho}$ with $\Phi(\bar{q})(e_1)$ instead of Λe_1, that is,

$$\partial_t \rho + c_1 \nabla_x \cdot (\rho \, \Phi(\bar{q})(e_1)) = 0.$$

We are left with computing the evolution for $\Lambda = \Lambda(t, x)$ and $\bar{q} = \bar{q}(t, x)$. This is done in the following section.

5.4 The equation for the body orientation Λ

The natural path to obtain an equation for $\Lambda = \Lambda(t, x)$ is to multiply the rescaled kinetic equation (46) by A; integrate this expression in $SO_3(\mathbb{R})$; and use the consistency relations (55) at the limit $\varepsilon = 0$. First multiplying by A and integrating we obtain

$$\partial_t \int_{SO_3(\mathbb{R})} A f_\varepsilon \, dA + \int_{SO_3(\mathbb{R})} [A \, (Ae_1 \cdot \nabla_x) f_\varepsilon] \, dA = \frac{1}{\varepsilon} \int_{SO_3(\mathbb{R})} A \Gamma_{\Lambda_{f_\varepsilon}}(f_\varepsilon) \, dA + \mathcal{O}(\varepsilon),$$

after dividing by ε on both sides. Notice that the limit of the first term indeed will correspond to $c_1 \partial_t(\rho \Lambda)$ thanks to the second consistency relation in Eq. (55). However, it is unclear how to deal with the ε^{-1} term on the right hand side as we do not have enough information on the asymptotics of the integral (and the same difficulty arises in the quaternion framework). In classical kinetic theory (in Mathematical Physics), this difficulty does not arise: typically every macroscopic

quantity corresponds to what is called a conserved quantity or collision invariant. We say that a function ψ is a collision invariant if for all f (in a reasonable class of functions)

$$\int_{SO_3(\mathbb{R})} \psi \Gamma_{\Lambda_f} \, dA = 0.$$

We have already seen that $\psi = 1$ is a collision invariant corresponding to the total mass being conserved. However, here the body orientation Λ is not a conserved quantity. The same kind of non-conservative property arises in the Vicsek model for the momentum of the particles. To sort out this problem we will relax the condition of being a collision invariant taking into account the constraint given by the second equality in Eq. (55) for the limiting function. This gives rise to the concept coined as the Generalized Collision Invariant in Ref. [16] and that we explain in the following.

The Generalized Collision Invariant. Consider the following definition:

Definition 3 (Generalised Collision Invariant). *A function ψ_{Λ_0} is a Generalized Collision Invariant (GCI) associated with $\Lambda_0 \in SO_3(\mathbb{R})$ of the operator Γ if it holds that*

$$\int_{SO_3(\mathbb{R})} \Gamma_{\Lambda_0}(f)\psi_{\Lambda_0} = 0, \quad \text{for all } f \text{ such that } P_{T_{\Lambda_0}}\left(\int_{SO_3(\mathbb{R})} A \; f \; dA\right) = 0.$$

We denote by $\mathrm{GCI}(\Lambda_0)$ *this set of Generalized Collision Invariants associated with Λ_0.*

In the quaternion formulation, we say that a function ψ_{q_0} is a Generalized Collision Invariant associated to $q_0 \in \mathbb{H}_1$ of the operator Γ if it holds that

$$\int_{\mathbb{H}_1} \Gamma_{q_0}(f)\psi_{q_0} = 0, \text{ for all } f \text{ such that } P_{q_0^\perp}\left(\int_{\mathbb{H}_1} (q \otimes q - \tfrac{1}{4}\mathrm{I}_4) f(q) \; dq \; q_0\right) = 0.$$

We also denote by $\mathrm{GCI}(q_0)$ *this set of Generalized Collision Invariants associated with q_0.*

Remark 3 (On the constraints on the test functions). In the matrix formulation, one can notice that the condition on the test functions f is equivalent to saying that $J_f \in T_{\Lambda_0}^\perp$, which is equivalent to $J_f = \Lambda_0 S$ for some symmetric matrix S (see Prop. 2). Taking $S = \rho c_1 \mathrm{I}_3$ and $\Lambda_0 = \Lambda_{f_0}$, we recover the second equality in (55). That is, the limiting function f_0 is an admissible test function in the definition of the GCI (associated with Λ_{f_0}). Something similar happens in the case of the quaternions: the conditions on the test functions f is equivalent to asking that $P_{q_0^\perp}(Q_f \, q_0) = 0$, which will hold true if q_0 is an eigenvector of Q_f, which is what happens, in particular, for $f = f_0$ and $q_0 = \bar{q}_{f_0}$ by the second equality in (59). Therefore, the limiting function f_0 is an admissible test function in the definition of the GCI (associated with $\bar{q}_{f_0}$).

Remark 4. It is straightforward to see that this notion extends the notion of collision invariant. In particular, the mass $\psi = 1$, which is a collision invariant, is also a GCI (associated with Λ_0 for any $\Lambda_0 \in SO_3(\mathbb{R})$ if we see ψ as a function on $SO_3(\mathbb{R})$, and associated with q_0 for any $q_0 \in \mathbb{H}_1$ if we see ψ as a function on $\mathbb{H}_1$). Note in particular that the definition of the GCI is non-empty.

We explain next how the GCI is useful. Multiplying Eq. (46) by a GCI associated with Λ_{f_ε} and integrating with respect to A we obtain

$$\varepsilon \int \Big(\partial_t f_\varepsilon + (Ae_1 \cdot \nabla_x)(f_\varepsilon)\Big)\ \psi_{\Lambda_{f_\varepsilon}}\ dA = 0.$$

Notice that, indeed, the right hand side vanishes since

$$\int_{SO_3(\mathbb{R})} \Gamma_{\Lambda_{f_\varepsilon}}(f_\varepsilon)\ \psi_{\Lambda_{f_\varepsilon}} = 0,$$

given that f_ε satisfies the condition

$$P_{T_{\Lambda_{f_\varepsilon}}}\left(\int_{SO_3(\mathbb{R})} Af_\varepsilon\ dA\right) = P_{T_{\Lambda_{f_\varepsilon}}} J_{f_\varepsilon} = P_{T_{\Lambda_{f_\varepsilon}}}(S_\varepsilon \Lambda_{f_\varepsilon}) = 0,$$

where $J_{f_\varepsilon} = S_\varepsilon \Lambda_{f_\varepsilon}$ is the Polar Decomposition of J_{f_ε} and S_ε is a symmetric matrix, therefore $J_{f_\varepsilon} \in T^{\perp}_{\Lambda_{f_\varepsilon}}$ (see Prop. 2).

Now, dividing by ε and then making $\varepsilon \to 0$, using (58) we obtain

$$\int_{SO_3(\mathbb{R})} \Big(\partial_t(\rho M_{\Lambda_{f_0}}) + (Ae_1) \cdot \nabla_x(\rho M_{\Lambda_{f_0}})\Big)\ \psi_{\Lambda_{f_0}}\ dA = 0. \tag{62}$$

Consequently, if we can compute the Generalized Collision Invariants in an explicit form, then we will be able to make explicit the limit given in Eq. (62) and we will be done. This is done in the following.

Description of the GCI. The explicit description of the GCI is given in the following proposition. For this, we need to introduce $h = h(r)$ the unique solution (see Ref. [13]) of the following differential equation on $(-1, 1)$:

$$\begin{aligned}(1-r^2)^{3/2} \exp\left(\frac{2r^2}{d}\right)\left(\frac{-4}{d}r^2 - 3\right) h(r) + \frac{d}{dr}\left[(1-r^2)^{5/2} \exp\left(\frac{2r^2}{d}\right) h'(r)\right] \\ = r\,(1-r^2)^{3/2} \exp\left(\frac{2r^2}{d}\right).\end{aligned} \tag{63}$$

The function h is *odd*: $h(-r) = -h(r)$, and it satisfies for all $r \geq 0$, $h(r) \leq 0$ (by maximum principle).

Proposition 11 (Description of the GCI). *Let $\Lambda_0 \in SO_3(\mathbb{R})$ and $q_0 \in \mathbb{H}_1$. Then, it holds that*

$$\begin{aligned} &\mathrm{GCI}(\Lambda_0) = \mathrm{span}\left\{1, \cup_{P\in\mathcal{A}} \psi^P_{\Lambda_0}\right\}, && \textit{(matrix formulation),}\\ &\mathrm{GCI}(q_0) = \mathrm{span}\left\{1, \cup_{\beta\in q_0{}^\perp} \psi^\beta_{q_0}\right\}, && \textit{(quaternion formulation),} \end{aligned}$$

where, for $P \in \mathcal{A}$ and $\beta \in q_0^\perp$,

$$\begin{aligned} &\psi^P_{\Lambda_0}(A) = P\cdot(\Lambda_0^T A)\,\bar{k}(\Lambda_0\cdot A), && \textit{(matrix formulation),}\\ &\psi^\beta_{q_0}(q) := (\beta\cdot q)\,\bar{h}(q\cdot q_0), && \textit{(quaternion formulation)} \end{aligned} \tag{64}$$

with $\bar{h}$ given by, for $r \in (-1,1)$,

$$\bar{h}(r) = \begin{cases} h(r) & \textit{in the gradual alignment model,}\\ r & \textit{in the jump model,} \end{cases}$$

where h is the unique solution of the differential equation (63)*, and $\bar{k}$ given by, for $r \in (-1/2, 3/2)$,*

$$\bar{k}(s) = \frac{\bar{h}(\frac{1}{2}\sqrt{2s+1})}{\frac{1}{2}\sqrt{2s+1}}. \tag{65}$$

The function $\bar{k}$ is designed so that $\bar{k}(\frac{1}{2}+\cos\theta) = \frac{\bar{h}(\cos\frac{\theta}{2})}{\cos\frac{\theta}{2}}$. It is negative in the gradual alignment model and a constant equal to $\bar{k} = 1$ in the jump model.

Remark 5. The relation between the functions $\bar{k}$ and $\bar{h}$ in Eq. (65) is related to the relation between dot products, see Prop. 4.

The first step to prove this proposition is a characterization of the GCI in terms of the adjoint of the collision operator.

Lemma 3 (Characterization of the GCI). *A function $\psi_{\Lambda_0} : SO_3(\mathbb{R}) \to \mathbb{R}$ (resp., $\psi_{q_0} : \mathbb{H}_1 \to \mathbb{R}$) is a GCI associated with $\Lambda_0 \in SO_3(\mathbb{R})$ (resp., associated with $q_0 \in \mathbb{H}_1$) if and only if there exists $P \in \mathcal{A}$ (resp., $\beta \in q_0^\perp$) such that ψ_{Λ_0} (resp., ψ_{q_0}) is solution of*

$$\begin{aligned} &\Gamma^*_{\Lambda_0}\psi_{\Lambda_0}(A) = P\cdot\Lambda_0^T A, \ \textit{for all } A\in SO_3(\mathbb{R}) \ \textit{(matrix formulation),}\\ &\Gamma^*_{q_0}\psi_{q_0}(q) = (\beta\cdot q)(q\cdot q_0), \ \textit{for all } q\in\mathbb{H}_1 \ \textit{(quaternion formulation),} \end{aligned} \tag{66}$$

where $\Gamma^*_{\Lambda_0}$ denotes the adjoint in $L^2(SO_3(\mathbb{R}))$ of the operator Γ_{Λ_0} (resp., and $\Gamma^*_{q_0}$ denotes the adjoint in $L^2(\mathbb{H}_1)$ of Γ_{q_0}).

Proof. We show the proof here for the matrix formulation. For the quaternion formulation it is done analogously. Given $f : SO_3(\mathbb{R}) \to \mathbb{R}$ and $\Lambda_0 \in SO_3(\mathbb{R})$,

we have the following equivalences (in the second equivalence we use Prop. 2):

$$\begin{aligned} P_{T_{\Lambda_0}}\left(\int_{SO_3(\mathbb{R})} A\ f\ dA\right) = 0 &\Leftrightarrow \int_{SO_3(\mathbb{R})} A\ f\ dA \in T^{\perp}_{\Lambda_0}, \\ &\Leftrightarrow (\Lambda_0 P)\cdot \int_{SO_3(\mathbb{R})} A\ f\ dA = 0 \ \text{ for all } P\in\mathcal{A}, \\ &\Leftrightarrow \int_{SO_3(\mathbb{R})} P\cdot(\Lambda_0^T A)\ f\ dA = 0 \ \text{ for all } P\in\mathcal{A}, \\ &\Leftrightarrow f\in G^{\perp}, \end{aligned}$$

where

$$G = \{g\in L^2(SO_3(\mathbb{R})) \mid g(A) = P\cdot\Lambda_0^T A, \text{ for some } P\in\mathcal{A}\}.$$

Starting from Def. 3, we then get, for $\psi_{\Lambda_0} : SO_3(\mathbb{R})\to\mathbb{R}$:

$$\begin{aligned} \psi_{\Lambda_0}\in \mathrm{GCI}(\Lambda_0) &\Leftrightarrow \int_{SO_3(\mathbb{R})} \Gamma_{\Lambda_0}(f)\psi_{\Lambda_0} = 0 \quad \text{for all } f \text{ such that } f\in G^{\perp}, \\ &\Leftrightarrow \int_{SO_3(\mathbb{R})} f\Gamma^*_{\Lambda_0}(\psi_{\Lambda_0}) = 0 \quad \text{for all } f \text{ such that } f\in G^{\perp}, \\ &\Leftrightarrow \Gamma^*_{\Lambda_0}(\psi_{\Lambda_0})\in (G^{\perp})^{\perp} = G, \end{aligned}$$

where $\Gamma^*_{\Lambda_0}$ is the adjoint of $\Gamma^*_{\Lambda_0}$ in $L^2(SO_3(\mathbb{R}))$. The last equality comes from the fact that the space G is a finite-dimensional subspace of L^2. The last equivalence therefore implies that ψ_{Λ_0} is a GCI if and only if there exists $P\in\mathcal{P}$ such that ψ_{Λ_0} is solution of (66). □

One can check that, in the matrix formulation, for $\psi : SO_3(\mathbb{R})\to\mathbb{R}$, the adjoint is given by

$$\Gamma^*_{\Lambda_0}(\psi) = \begin{cases} D\nabla_A\cdot(M_{\Lambda_0}\nabla_A\psi) & \text{(gradual alignment model)}, \\ \int_{SO_3(\mathbb{R})}\psi(A)\,M_{\Lambda_0}(A)\ dA - \psi & \text{(jump model)}, \end{cases} \tag{67}$$

and in the quaternion formulation we have, for $\bar{\psi}:\mathbb{H}_1\to\mathbb{R}$:

$$\Gamma^*_{q_0}(\bar{\psi}) = \begin{cases} D\nabla_q\cdot(M_{q_0}\nabla_q\bar{\psi}) & \text{(gradual alignment model)}, \\ \int_{\mathbb{H}_1}\bar{\psi}(q)M_{q_0}(q)\ dq - \bar{\psi} & \text{(jump model)}. \end{cases}$$

The end of the proof of Prop. 11 for the gradual alignment model relies on the application of Lax-Milgram theorem. It is done in references [12] (for the matrix formulation) and [13] (for the quaternion formulation). We do not repeat it here.

In the case of the jump model, for the matrix formulation it is a direct check that for any $P'\in\mathcal{A}$, the function $\psi^{P'}_{\Lambda_0}$ defined by

$$\psi^{P'}_{\Lambda_0}(A) = P'\cdot\Lambda_0^T A,$$

satisfies Eq. (66) with $P = -P'$ and is, therefore, a GCI. As noticed in Rem. 4, the constant function $\psi = 1$ is also a GCI. Conversely, using the explicit form (67) of the adjoint operator $\Gamma^*_{\Lambda_0}$, it is also direct to see that any solution ψ_{Λ_0} of Eq. (66) for some $P \in \mathcal{A}$ satisfies

$$\psi_{\Lambda_0}(A) = -P \cdot \Lambda_0^T A + \int_{SO_3(\mathbb{R})} \psi_{\Lambda_0}(A')\ M_{\Lambda_0}(A')\ dA' \ \in\ \operatorname{span}\left\{1,\ \psi_{\Lambda_0}^{-P}\right\}.$$

Analogously, one can check that for any $\beta' \in q_0^\perp$, the function $\psi = \psi_{q_0}^{\beta'}$ given in Eq. (64) is indeed a GCI using Eq. (66) with $\beta = -\beta'$ and the consistency relation (60). Conversely, one can check that any solution ψ of Eq. (66) for some $\beta \in q_0^\perp$ belongs to $\operatorname{span}\left\{1,\ \psi_{q_0}^{-\beta}\right\}$.

Limiting equation. Now that we have an explicit form for the GCI, we can go back to the limiting equation (62) (in the matrix formulation) and substitute its value. This way we have that for all $P \in \mathcal{A}$ it holds:

$$\int_{SO_3(\mathbb{R})} \Big(\partial_t(\rho M_\Lambda) + (Ae_1 \cdot \nabla_x)(\rho M_\Lambda)\Big)\ (P \cdot \Lambda^T A)\ dA = 0.$$

This is equivalent to:

$$P \cdot \left[\int_{SO_3(\mathbb{R})} \Big(\partial_t(\rho M_\Lambda) + (Ae_1 \cdot \nabla_x)(\rho M_\Lambda)\Big)\ \Lambda^T A\ dA\right] = 0 \quad \text{for all } P \in \mathcal{A},$$

which implies thanks to Prop. 1 that

$$\int_{SO_3(\mathbb{R})} \Big(\partial_t(\rho M_\Lambda) + (Ae_1 \cdot \nabla_x)(\rho M_\Lambda)\Big)\ \Lambda^T A\ dA \in \mathcal{S},$$

or, in other words,

$$\int_{SO_3(\mathbb{R})} \Big(\partial_t(\rho M_\Lambda) + (Ae_1 \cdot \nabla_x)(\rho M_\Lambda)\Big)(\Lambda^T A - A^T \Lambda)\ dA = 0.$$

It remains only to compute this expression. This expression is exactly the same is in Ref. [12, Equation (4.25)] with the function $\bar{\psi}_0$ appearing in this reference to be taken equal to one. Therefore, here we do not repeat again the computation for this expression and put directly the result in Th. 2 (Eq. (70)) in the following section.

5.5 Main results

To introduce the results on the matrix formulation we need to introduce first some notation: For a smooth function Λ from $\mathbb{R}^3$ to $SO_3(\mathbb{R})$, and for $x \in \mathbb{R}^3$, we define the matrix $\mathcal{D}_x(\Lambda)$ such that

$$(w \cdot \nabla_x)\Lambda = [\mathcal{D}_x(\Lambda)w]_\times \Lambda, \qquad \text{for any } w \in \mathbb{R}^3.$$

This matrix is well defined (see [12, Sec. 4.5]). With this, we define the following first-order operators

$$\delta_x(\Lambda) = \mathrm{Tr}(\mathcal{D}_x(\Lambda)), \qquad [\mathrm{r}_x(\Lambda)]_\times = \mathcal{D}_x(\Lambda) - \mathcal{D}_x(\Lambda)^T.$$

In order to present the results on the quaternion formulation, we first introduce the (right) relative differential operator on $\mathbb{H}_1$: for a function $q = q(t,x)$ where $q(t,x) \in \mathbb{H}_1$ and for $\partial \in \{\partial_t, \partial_{x_1}, \partial_{x_2}, \partial_{x_3}\}$, let

$$\partial_{\mathrm{rel}} q := (\partial q) q^*, \ \Big(= \mathrm{Im}((\partial q) q^*) \Big), \tag{68}$$

where ∂q belongs to the orthogonal space of q, and the product has to be understood in the sense of quaternions. Notice that, effectively, $\partial_{\mathrm{rel}} q$ is a purely imaginary quaternion, since $\mathrm{Re}((\partial q) q^*) = q \cdot \partial q = 0$ (by the fact that q is a unit quaternion), and it can be identified with a vector in $\mathbb{R}^3$. With this, we define the (right) relative space differential operators

$$\nabla_{x,\mathrm{rel}} q = (\partial_{x_i,\mathrm{rel}} q)_{i=1,2,3} = ((\partial_{x_i} q) q^*)_{i=1,2,3} \in (\mathbb{R}^3)^3 \subset \mathbb{H}^3,$$
$$\nabla_{x,\mathrm{rel}} \cdot q = \sum_{i=1,2,3} (\partial_{x_i,\mathrm{rel}} q)_i = \sum_{i=1,2,3} ((\partial_{x_i} q) q^*)_i \in \mathbb{R},$$

where $((\partial_{x_i} q) q^*)_i$ indicates the i-th component of $(\partial_{x_i} q) q^*$.

With these notations, we can state the main result:

Theorem 2 ((Formal) macroscopic limit). *The following results hold true for both the jump model and the gradual alignment model. When $\varepsilon \to 0$ in the kinetic equations* (46) *(matrix representation) and* (49) *(quaternion representation) it holds (formally) that*

$$f_\varepsilon \to f = f(t,x,A) = \rho M_\Lambda(A), \quad \text{with } \Lambda = \Lambda(t,x) \in SO_3(\mathbb{R}),\ \rho = \rho(t,x) \geq 0,$$
$$f_\varepsilon \to f = f(t,x,q) = \bar\rho M_{\bar q}(q), \quad \text{with } \bar q = \bar q(t,x) \in \mathbb{H}_1,\ \bar\rho = \bar\rho(t,x) \geq 0,$$

for the matrix representation and the quaternion representation, respectively. Moreover, if the convergence is strong enough and the pair functions (ρ, Λ), $(\bar\rho, \bar q)$ are regular enough, then they satisfy the following systems, respectively:

$$\partial_t \rho + \nabla_x \cdot (c_1 \rho \Lambda e_1) = 0, \tag{69}$$
$$\begin{aligned} &\rho(\partial_t \Lambda + c_2((\Lambda e_1) \cdot \nabla_x)\Lambda) \\ &\quad + [(\Lambda e_1) \times (2c_3 \nabla_x \rho + c_4 \rho \mathrm{r}_x(\Lambda)) + c_4 \rho \delta_x(\Lambda) \Lambda e_1]_\times \Lambda = 0, \end{aligned} \tag{70}$$

and

$$\partial_t \bar\rho + \nabla_x \cdot (c_1 e_1(\bar q) \bar\rho) = 0, \tag{71}$$
$$\begin{aligned} &\bar\rho(\partial_t \bar q + c_2'(e_1(\bar q) \cdot \nabla_x)\bar q) \\ &\quad + c_3 [e_1(\bar q) \times \nabla_x \bar\rho]\, \bar q + c_4 \bar\rho\, [\nabla_{x,\mathrm{rel}} \bar q\, e_1(\bar q) + (\nabla_{x,\mathrm{rel}} \cdot \bar q) e_1(\bar q)]\, \bar q = 0, \end{aligned} \tag{72}$$

where the (right) relative differential operator $\nabla_{x,\mathrm{rel}}$ *is defined in Eq.* (68)*; and*

$$e_1(\bar{q}) = \mathrm{Im}(\bar{q} e_1 \bar{q}^*),$$

and where c_i*,* $i = 1, \dots, 4$ *are explicit constants. To define them we use the following notation: for two real functions* g*,* w *consider*

$$\langle g \rangle_w := \int_0^\pi g(\theta) \frac{w(\theta)}{\int_0^\pi w(\theta') d\theta'} \, d\theta.$$

Then the constants are given by

$$\begin{aligned}
c_1 &= \frac{2}{3} \langle 1/2 + \cos\theta \rangle_{m(\theta)\, \sin^2(\theta/2)}, \\
c_2 &= \frac{1}{5} \langle 2 + 3\cos\theta \rangle_{m(\theta)\, \sin^4(\theta/2) \bar{h}(\cos(\theta/2)) \cos(\theta/2)}, \\
c_2' &= \frac{1}{5} \langle 1 + 4\cos\theta \rangle_{m(\theta)\, \sin^4(\theta/2) \bar{h}(\cos(\theta/2)) \cos(\theta/2)}, \\
c_3 &= \frac{D}{2}, \\
c_4 &= \frac{1}{5} \langle 1 - \cos\theta \rangle_{m(\theta)\, \sin^4(\theta/2) \bar{h}(\cos(\theta/2)) \cos(\theta/2)},
\end{aligned}$$

where

$$m(\theta) := \exp\left(\frac{1}{D} \left(\frac{1}{2} + \cos\theta \right) \right),$$

with $\bar{h}$ *given by, for* $r \in (-1, 1)$*,*

$$\bar{h}(r) = \begin{cases} h(r) & \text{in the gradual alignment model,} \\ r & \text{in the jump model,} \end{cases}$$

where h *is the unique solution of the differential equation* (63)*.*

Note that the matrix product in the fourth term of Eq. (72) has to be understood as a matrix product, giving rise to a scalar product in $\mathbb{H}$:

$$\nabla_{x,\mathrm{rel}} \bar{q} \, e_1(\bar{q}) = ((\partial_{x_i,\mathrm{rel}} \bar{q}) \cdot e_1(\bar{q}))_{i=1,2,3} \, .$$

We now state the equivalence of the matrix formulation and the quaternion formulation:

Theorem 3 (Equivalences of the equations [13]). *Let* $\rho_0 = \rho_0(x) \geq 0$*. Let* $\bar{q}_0 = \bar{q}_0(x) \in \mathbb{H}_1$ *and* $\Lambda_0 = \Lambda_0(x) \in SO_3(\mathbb{R})$ *represent the same rotation, i.e.,* $\Lambda_0(x) = \Phi(\bar{q}_0(x))$ *for all* $x \in \mathbb{R}^3$*. Then the system* (69)–(70) *and the system* (71)–(72) *are equivalent (in the sense that any solution* $(\rho, \Lambda = \Phi(\bar{q}))$ *of the system* (69)–(70) *is a solution* $(\bar{\rho}, \bar{q})$ *of* (71)–(72)*).*

Therefore the equations in the matrix formulation and in the quaternion formulation are equivalent. For an explicit term-by-term equivalence, the reader is referred to [13, Sec. 5.3.3]. Moreover, we have the following corollary:

Corollary 1. *The jump model and the gradual alignment model give rise to the same macroscopic equations with different constants when the equilibria in the jump model is given by a von-Mises distribution.*

We conclude this section by giving a short interpretation of the macroscopic equations obtained in Th. 2. For a full description and justification we refer the reader to [12,13]. Since by Th. 3 we know that the systems (69)–(70) and (71)–(72) are equivalent, we will restrict ourselves to interpreting the matrix formulation (for more details on the quaternion formulation the reader is referred to [13]).

Eq. (69) is the continuity equation for ρ and ensures mass conservation. The convection velocity is given by $c_1 \Lambda e_1$ and Λe_1 gives the direction of motion. Eq. (70) gives the evolution of the mean orientation Λ. We remark that every term in Eq. (70) belongs to the tangent space at Λ in $SO(3)$; this is true for the first term since $(\partial_t + c_2(\Lambda e_1) \cdot \nabla_x)$ is a differential operator and it also holds for the second term because it is the product of an antisymmetric matrix with Λ (see Prop. 2).

The term corresponding to c_3 in (70) gives the influence of $\nabla_x \rho$ (pressure gradient) on the body attitude Λ. It has the effect of rotating the body around the vector directed by $(\Lambda e_1) \times \nabla_x \rho$ at an angular speed given by $\frac{c_3}{\rho} \|(\Lambda e_1) \times \nabla_x \rho\|$, so as to align Λe_1 with $-\nabla_x \rho$ (for more details on this, see [12]). Therefore, the $\nabla_x \rho$ term has the same effect as a pressure gradient in classical hydrodynamics. In this case the pressure gradient has the effect of rotating the whole body frame.

If we had that $c_3 = c_4 = 0$, then we would recover the Self-Organized Hydrodynamic (SOH) model. The SOH model corresponds to the macroscopic equations of the Vicsek model [16]. The SOH model bears analogies with the compressible Euler equations, where (69) is obviously the mass conservation equation and (70) is akin to the momentum conservation equation. There are however major differences. The first one is that we preserve the constraint $\Lambda(t) \in SO_3(\mathbb{R})$ for all times and so the mass convection speed is $|c_1 \Lambda(t) e_1| = c_1$ for all times, while the velocity in the Euler equations is an arbitrary vector. The second one is that the convection speed c_2 is a priori different from the mass convection speed c_1. This difference is a signature of the lack of Galilean invariance of the system, which is a common feature of all dry active matter models.

The major novelty of the present model are the terms with constants c_3 and c_4. They influence the transport of the direction of motion Λe_1. The overall dynamics tends to align the velocity orientation Λe_1, not opposite to the density gradient $\nabla_x \rho$ but opposite to a composite vector $(c_3 \nabla_x \rho + c_4 \rho \, \mathrm{r}_x)$. The vector r_x gives rise to an effective pressure force which adds up to the usual pressure gradient. In addition to this effective force, spatial inhomogeneities of the body attitude also have the effect of inducing a proper rotation of the frame about the direction of motion. This proper rotation is proportional to δ_x. For an interpretation of r_x, δ_x, see [12].

Finally, we add the following interpretation based on the quaternion formulation. First, note that considering $\partial = \partial_t$ the time derivative, for a func-

tion $q = q(t,x)$ with values in $\mathbb{H}_1$ the vector $\partial_{t,\mathrm{rel}} q = \partial_t q\, q^{-1}$ is half of the angular velocity of a solid of orientation represented by q. By analogy, the vector $\partial_{x_i,\mathrm{rel}} q = \partial_{x_i} q\, q^{-1}$ for $i = 1, 2, 3$ is half of the angular variation in space of a solid of orientation represented by q. Now, in the quaternion formulation the evolution equation for the body attitude can be rewritten as

$$\bar\rho(\partial_{t,\mathrm{rel}}\bar q + c_2(e_1(\bar q)\cdot\nabla_{x,\mathrm{rel}})\bar q) \\ + c_3 e_1(\bar q)\times\nabla_x\bar\rho + c_4\bar\rho\,[\nabla_{x,\mathrm{rel}}\bar q\, e_1(\bar q) + (\nabla_{x,\mathrm{rel}}\cdot\bar q)e_1(\bar q)] = 0,$$

simply by multiplying Eq. (72) by $\bar q^{-1}$ on the right. This equation lives in $\mathbb{R}^3$ (since $\partial_{\mathrm{rel}}\bar q$ lives in $\mathbb{R}^3$), and it only involves the following physical quantities: the macroscopic density ρ (and its space gradient), the macroscopic direction of movement $e_1(\bar q)$, and the macroscopic angular time/space variations of the body attitude $2\partial_{\mathrm{rel}}\bar q$.

6 Conclusion

In these notes, we have formally derived macroscopic models, starting from the description of particle systems, and using an intermediate kinetic model to link the two scales. The two limits ($N \to \infty$ for the particle system, and $\varepsilon \to 0$ for the rescaled kinetic equations) are formal derivations, but some steps towards a rigorous limit can be done. A way to recover a rigorous mean-field limit is to change the model in such a way that the singular behavior of the alignment is removed, as in [3], but it introduces a phenomenon of phase transition as in [10,11]. The study of this phase transition is an ongoing work. Another issue to have a better understanding of the limit $\varepsilon \to 0$ is to have well-posedness of the macroscopic system (69)-(70), so we need to study its hyperbolicity. This is also an ongoing work.

7 Acknowledgments

P.D. acknowledges support from the Royal Society and the Wolfson foundation through a Royal Society Wolfson Research Merit Award ref WM130048; the British "Engineering and Physical Research Council" under grants ref: EP/M006883/1 and EP/P013651/1; the National Science Foundation under NSF Grant RNMS11-07444 (KI-Net). P.D. is on leave from CNRS, Institut de Mathématiques de Toulouse, France.
A.F. acknowledges support from the EFI project ANR-17-CE40-0030 and the Kibord project ANR-13-BS01-0004 of the French National Research Agency (ANR), as well as from the project Défi S2C3 POSBIO of the interdisciplinary mission of CNRS, and the project SMS co-funded by CNRS and the Royal Society.
A.T. acknowledges support from the Kibord project ANR-13-BS01-0004 of the French National Research Agency (ANR).

References

1. Azaïs, R., Bardet, J.B., Génadot, A., Krell, N., Zitt, P.A.: Piecewise deterministic Markov process—recent results. In: Journées MAS 2012, *ESAIM Proc.*, vol. 44, pp. 276–290. EDP Sci., Les Ulis (2014)
2. Bolley, F., Cañizo, J.A., Carrillo, J.A.: Stochastic mean-field limit: non-Lipschitz forces & swarming. Math. Models Methods Appl. Sci. **21**(11), 2179–2210 (2011)
3. Bolley, F., Cañizo, J.A., Carrillo, J.A.: Mean-field limit for the stochastic Vicsek model. Appl. Math. Lett. **3**(25), 339–343 (2012)
4. Bostan, M., Carrillo, J.A.: Reduced fluid models for self-propelled particles interacting through alignment. Math. Models Methods Appl. Sci. **27**(7), 1255–1299 (2017)
5. Carrillo, J.A., Choi, Y., Hauray, M., Salem, S.: Mean-field limit for collective behavior models with sharp sensitivity regions. J. Europ. Math. Soc. (2018). To appear
6. Cavagna, A., Del Castello, L., Giardina, I., Grigera, T., Jelic, A., Melillo, S., Mora, T., Parisi, L., Silvestri, E., Viale, M., et al.: Flocking and turning: a new model for self-organized collective motion. J. Stat. Phys. **158**(3), 601–627 (2014)
7. Cercignani, C., Illner, R., Pulvirenti, M.: The mathematical theory of dilute gases, vol. 106. Springer Science & Business Media (2013)
8. Constantin, P.: The onsager equation for corpora. Journal of Computational and Theoretical Nanoscience **7**(4), 675–682 (2010)
9. Degond, P.: Macroscopic limits of the Boltzmann equation: a review. In: Modeling and Computational Methods for Kinetic Equations, pp. 3–57. Springer (2004)
10. Degond, P., Frouvelle, A., Liu, J.G.: Macroscopic limits and phase transition in a system of self-propelled particles. J. Nonlinear Sci. **23**(3), 427–456 (2013)
11. Degond, P., Frouvelle, A., Liu, J.G.: Phase transitions, hysteresis, and hyperbolicity for self-organized alignment dynamics. Arch. Ration. Mech. Anal. **216**(1), 63–115 (2015)
12. Degond, P., Frouvelle, A., Merino-Aceituno, S.: A new flocking model through body attitude coordination. Math. Models Methods Appl. Sci. **27**(06), 1005–1049 (2017)
13. Degond, P., Frouvelle, A., Merino-Aceituno, S., Trescases, A.: Quaternions in collective dynamics. Multiscale Mod. Simul. **16**(1), 28–77 (2018)
14. Degond, P., Liu, J.G., Motsch, S., Panferov, V.: Hydrodynamic models of self-organized dynamics: derivation and existence theory. Methods Appl. Anal. **20**(2), 89–114 (2013)
15. Degond, P., Manhart, A., Yu, H.: A continuum model for nematic alignment of self-propelled particles. Discrete Contin. Dyn. Syst. Ser. B **22**(4), 1295–1327 (2017)
16. Degond, P., Motsch, S.: Continuum limit of self-driven particles with orientation interaction. Math. Models Methods Appl. Sci. **18**, 1193–1215 (2008)
17. Degond, P., Motsch, S.: A macroscopic model for a system of swarming agents using curvature control. J. Stat. Phys. **143**(4), 685–714 (2011)
18. Degond, P., Navoret, L.: A multi-layer model for self-propelled disks interacting through alignment and volume exclusion. Math. Models Methods Appl. Sci. **25**(13), 2439–2475 (2015)
19. Dimarco, G., Motsch, S.: Self-alignment driven by jump processes: Macroscopic limit and numerical investigation. Math. Models Methods Appl. Sci. **26**(07), 1385–1410 (2016)

20. Doi, M., Edwards, S.F.: The Theory of Polymer Dynamics, *International Series of Monographs on Physics*, vol. 73. Oxford University Press (1999)
21. Ferdinandy, B., Ozogány, K., Vicsek, T.: Collective motion of groups of self-propelled particles following interacting leaders. Phys. A **479**, 467–477 (2017)
22. Frouvelle, A.: A continuum model for alignment of self-propelled particles with anisotropy and density-dependent parameters. Math. Models Methods Appl. Sci. **22**(7), 1250,011, 40 (2012)
23. Gamba, I.M., Kang, M.J.: Global weak solutions for Kolmogorov-Vicsek type equations with orientational interactions. Arch. Ration. Mech. Anal. **222**(1), 317–342 (2016)
24. Hsu, E.P.: Stochastic Analysis on Manifolds, *Graduate Series in Mathematics*, vol. 38. American Mathematical Society, Providence, Rhode Island (2002)
25. Sarlette, A., Sepulchre, R., Leonard, N.E.: Autonomous rigid body attitude synchronization. Automatica **45**(2), 572–577 (2009)
26. Sone, Y.: Kinetic theory and fluid dynamics. Springer Science & Business Media (2012)
27. Sznitman, A.S.: Topics in propagation of chaos. In: École d'Été de Probabilités de Saint-Flour XIX — 1989, *Lecture Notes in Mathematics*, vol. 1464, pp. 165–251. Springer, Berlin (1991)
28. Vicsek, T., Czirók, A., Ben-Jacob, E., Cohen, I., Shochet, O.: Novel type of phase transition in a system of self-driven particles. Phys. Rev. Lett. **75**(6), 1226–1229 (1995)