MEAN-FIELD LIMIT FOR THE STOCHASTIC VICSEK MODEL

FRANÇOIS BOLLEY, JOSÉ A. CAÑIZO, AND JOSÉ A. CARRILLO

Abstract. We consider the continuous version of the Vicsek model with noise, proposed as a model for collective behavior of individuals with a fixed speed. We rigorously derive the kinetic mean-field partial differential equation satisfied when the number $N$ of particles tends to infinity, quantifying the convergence of the law of one particle to the solution of the PDE. For this we adapt a classical coupling argument to the present case in which both the particle system and the PDE are defined on a surface rather than on the whole space $\mathbb{R}^d$. As part of the study we give existence and uniqueness results for both the particle system and the PDE.

Introduction

The stochastic Vicsek model [13] arises in the study of collective motion of animals and it is receiving lots of attention due to the appearance of a phase transition [2, 9]. A continuum version and variants of this model have been proposed in the recent works [5, 4]. Our objective is to rigorously derive some continuum partial differential equations analysed in [5] from the stochastic Vicsek particle model. This was carried out for a family of collective behaviour models in [1] following the method of [12]. The present models do not fall into this analysis due to the evolution being defined on a surface as we explain next. In the models considered here, individuals are assumed to move with a fixed cruising speed trying to average their orientations with other individuals in the swarm in the presence of noise. This orientation mechanism is modelled by locally averaging in space their relative velocity to other individuals. More precisely, we are interested in the behaviour of $N$ interacting $\mathbb{R}^d$-valued processes $(X_i^t, V_i^t)_{t \geq 0}$ with $1 \leq i \leq N$ with constant speed $|V_i^t|$, say unity. We define them as solutions to the coupled Stratonovich stochastic differential equations

\[
\begin{aligned}
&dX_i^t = V_i^t \, dt, \\
&dV_i^t = \sqrt{2} P(V_i^t) \circ dB_i^t - P(V_i^t) \left( \frac{1}{N} \sum_{j=1}^{N} K(X_i^t - X_j^t)(V_i^t - V_j^t) \right) \, dt.
\end{aligned}
\]

(1)

Here $P(v)$ is the projection operator on the tangent space at $v/|v|$ to the unit sphere $\mathbb{S}^{d-1}$ in $\mathbb{R}^d$, i.e.,

\[ P(v) = I - \frac{v \otimes v}{|v|^2}. \]

This stochastic system is considered with independent and commonly distributed initial data $(X_i^0, V_i^0) \in \mathbb{R}^d \times \mathbb{S}^{d-1}$ with $1 \leq i \leq N$. The $(B_i^t)_{t \geq 0}$ denote $N$ independent standard Brownian motions in $\mathbb{R}^d$. The projection operator ensures that $V_i^t$ keeps constant norm, equal to 1. The second term in the evolution of $V_i^t$ models the tendency of the particle $i$ to have the same orientation as the other particles, in a way weighted by the interaction kernel $K$, as in the model proposed by F. Cucker and S. Smale [3]. Let us observe that $P(V_i^t)V_i^t = 0$, so we can drop the corresponding term when writing (1) to recover the usual formulations as in [5]. We will work with stochastic processes defined on $\mathbb{R}^{2d}$ instead of $\mathbb{R}^d \times \mathbb{S}^{d-1}$. We will check later on that solutions of (1) with initial data in $\mathbb{R}^d \times \mathbb{S}^{d-1}$ remain there for all times. We have written (1) in the Stratonovich sense, since the term involving noise corresponds to Brownian motion on the sphere $\mathbb{S}^{d-1}$ as in [10, Section 1.4] and [11, Section V.31].

Date: February 1, 2011.
By symmetry of the initial configuration and of the evolution, all particles have the same distribution. Even though they are initially independent, correlation builds up in time due to the interaction term. Nevertheless, this interaction term is of order 1/N, and thus, it seems reasonable that two of these interacting particles (or a fixed number k of them) become less and less correlated as N gets large (propagation of chaos).

Following [12] we shall show that the N interacting processes \((X_i^t, V_i^t)_{t \geq 0}\) respectively behave as \(N \to \infty\) like the auxiliary processes \((\overline{X}_i^t, \overline{V}_i^t)_{t \geq 0}\), solutions to

\[
\begin{align*}
\frac{d\overline{X}_i^t}{dt} &= \overline{V}_i^t, \\
\frac{d\overline{V}_i^t}{dt} &= \sqrt{2} \mathcal{P}(\overline{V}_i^t) \circ dB^i_t - \mathcal{P}(\overline{V}_i^t)(H * f_t)(\overline{X}_i^t, \overline{V}_i^t) dt,
\end{align*}
\]

in the Stratonovich sense. Here the Brownian motions \((B^i_t)_{t \geq 0}\) are those governing the evolution of the \((X_i^t, V_i^t)_{t \geq 0}\) and

\[
(H * f)(x, v) = \int_{\mathbb{R}^d} K(x - x') (v - v') f(x', v') \, dx' \, dv', \quad x, v \in \mathbb{R}^d.
\]

Note that (2) consists of \(N\) equations which can be solved independently of each other. Each of them involves the condition that \(f_t\) is the distribution of \((\overline{X}_i^t, \overline{V}_i^t)\), thus making it nonlinear. The processes \((\overline{X}_i^t, \overline{V}_i^t)_{t \geq 0}\) with \(i \geq 1\) are independent since the initial conditions and driving Brownian motions are independent.

We will show that these processes defined on \(\mathbb{R}^{2d}\) are identically distributed, take values in \(\mathbb{R}^d \times S^{d-1}\) if initially so, and their common law \(f_t\) at time \(t\), as a measure on \(\mathbb{R}^d \times S^{d-1}\), evolves according to

\[
\partial_t f_t + \omega \cdot \nabla_x f_t = \Delta_\omega f_t + \nabla_\omega \cdot (f_t(I - \omega \otimes \omega)(H * f_t)), \quad t > 0, x \in \mathbb{R}^d, \omega \in S^{d-1}.
\]

Now the convolution \(H * f\) is over \(\mathbb{R}^d \times S^{d-1}:

\[
(H * f)(x, \omega) = \int_{\mathbb{R}^d \times S^{d-1}} K(x - x') (\omega - \omega') f(x', \omega') \, dx' \, d\omega', \quad x \in \mathbb{R}^d, \omega \in S^{d-1}.
\]

Moreover, \(\nabla_x\) stands for the gradient with respect to the position variable \(x \in \mathbb{R}^d\) whereas \(\nabla_\omega, \Delta_\omega\) respectively stand for the gradient, divergence and Laplace-Beltrami operators with respect to the velocity variable \(\omega \in S^{d-1}\).

This equation is proposed in [4] as a continuous version of the original Vicsek model [13], and one of our purposes is to make this derivation rigorous. The asymptotic behavior and the appearance of a phase transition in the space-homogeneous version of (3) (i.e., without the space variable) has been recently studied in [7]. It is also known as the Doi-Onsager equation, introduced by Doi in [6] as a model for the non-equilibrium Statistical Mechanics of a suspension of polymers in which their spatial orientation (given by the parameter \(\omega \in S^{d-1}\)) is taken into account.

The main result of this paper can be summarized as:

**Theorem 1.** Let \(f_0\) be a probability measure on \(\mathbb{R}^d \times S^{d-1}\) with finite second moment in \(x \in \mathbb{R}^d\) and let \((X_i^0, V_i^0)\) for \(1 \leq i \leq N\) be \(N\) independent variables with law \(f_0\). Let also \(K\) be a Lipschitz and bounded map on \(\mathbb{R}^d\). Then,

i) There exists a pathwise unique global solution to the SDE system (1) with initial data \((X_i^0, V_i^0)\) for \(1 \leq i \leq N\); moreover, the solution is such that all \(V_i^t\) have norm 1.

ii) There exists a pathwise unique global solution to the nonlinear SDE (2) with initial datum \((X_i^0, V_i^0)\); moreover, the solution is such that \(V_i^t\) has norm 1.

iii) There exists a unique global weak solution to the nonlinear PDE (3) with initial datum \(f_0\). Moreover, it is the law of the solution to (2).

Solutions to general SDE’s can be built in submanifolds of \(\mathbb{R}^d\) by means of the Brownian motion of the ambient space as in [11, Theorem V.34.86] for instance; then one can interpret the generator
derivatives of all orders, such that \( \sigma \) since it is a kinetic model.

Here, we give the full construction and derivation of the evolution of the law as it can be done explicitly in the case of the sphere \( \mathbb{S}^{d-1} \). Let us also emphasize that we have only partial diffusion since it is a kinetic model.

We observe that existence of \( L^2 \) and classical solutions for the space-homogeneous version of (3) has also been considered in \([7]\).

As a direct consequence of the classical Sznitman’s theory, we get the following mean-field limit result:

**Theorem 2.** With the assumptions of Theorem 1 and for the respective solutions \((X^i_t, V^i_t)_{t \geq 0}\) and \((\overline{X}_t, \overline{V}_t)_{t \geq 0}\) of (1) and (2), for all \( T > 0 \) there exists a constant \( C \) such that

\[
\mathbb{E} \left[ |X^i_t - \overline{X}_t|^2 + |V^i_t - \overline{V}_t|^2 \right] \leq \frac{C}{N}
\]

for all \( 0 \leq t \leq T, N \geq 1 \) and \( 1 \leq i \leq N \).

This estimate classically ensures quantitative estimates on (see \([12, 1]\) for details)

i) the convergence in \( N \) of the law at time \( t \) of any (by symmetry) of the processes \((X^i_t, V^i_t)\) towards \( f_t \),

ii) the propagation of chaos for the particle system through the convergence of the law at time \( t \) of any \( k \) particles towards the tensor product \( f_t^\otimes k \) (for \( k \) fixed or \( k = o(N) \)),

iii) the convergence of the empirical measure at time \( t \) of the particle system towards \( f_t \).

Of course, the same techniques lead to a corresponding mean-field limit result for the space-homogeneous particle system instead of (1), obtaining the corresponding space-homogeneous PDE.

**Proofs**

Using the standard Itô-Stratonovich calculus, see \([8, p. 99]\) for instance, equations (1) and (2) are respectively equivalent to the Itô stochastic differential equations

\[
\begin{align*}
\frac{dX^i_t}{dt} &= V^i_t dt, \\
\frac{dV^i_t}{dt} &= \sqrt{2} P(V^i_t) dB^i_t - P(V^i_t) \left( \frac{1}{N} \sum_{j=1}^{N} K(X^i_t - X^j_t)(V^i_t - V^j_t) \right) dt - (d - 1) \frac{V^i_t}{|V^i_t|^2} dt.
\end{align*}
\]

and

\[
\begin{align*}
\frac{d\overline{X}_t}{dt} &= \overline{V}_t dt, \\
\frac{d\overline{V}_t}{dt} &= \sqrt{2} P(\overline{V}_t) dB_t - P(\overline{V}_t)(H * f_t)(\overline{X}_t, \overline{V}_t) dt - (d - 1) \frac{\overline{V}_t}{|\overline{V}_t|^2} dt, \\
(\overline{X}_0, \overline{V}_0) &= (X_0, V_0), \quad f_t = \text{law}(\overline{X}_t, \overline{V}_t)
\end{align*}
\]

which we now consider.

We start with the proof of Theorem 1. We use a regularization of the diffusion and drift coefficients. We let \( \sigma_1 \) be a \( d \times d \) matrix valued map on \( \mathbb{R}^d \) with bounded derivatives of all orders such that \( \sigma_1(v) = P(v) \) for all \( v \) with \( |v| \geq 1/2 \), and \( \sigma_2 \) and \( \sigma_3 \) be maps on \( \mathbb{R}^d \), again with bounded derivatives of all orders, such that \( \sigma_2(v) = v/|v|^2 \) if \( |v| \geq 1/2 \) and \( \sigma_3(v) = v \) if \( |v| \leq 2 \).
Existence and uniqueness for the particle system (4). Given such \( \sigma_1, \sigma_2 \), the system of equations

\[
\begin{aligned}
dX^i_t &= V^i_t dt, \\
dV^i_t &= \sqrt{2} \sigma_1(V^i_t) dB^i_t - \sigma_1(V^i_t) \left( \frac{1}{N} \sum_{j=1}^{N} K(X^i_t - X^j_t)(V^i_t - V^j_t) \right) dt - (d-1) \sigma_2(V^i_t) dt 
\end{aligned}
\]  

(6)

starting from \((X^i_0, V^i_0) \in \mathbb{R}^d \times \mathbb{S}^{d-1}\) for \( 1 \leq i \leq N \) has locally Lipschitz coefficients. Moreover, by the Itô formula and as long as \(|V^i_t| \geq 1/2\),

\[
d|V^i|^2 = 2\sqrt{2} V^i \cdot P(V^i) dB^i - 2 V^i \cdot P(V^i) \left( \frac{1}{N} \sum_{j=1}^{N} K(X^i - X^j)(V^i - V^j) \right) dt \\
 \quad - 2(d-1) dt + 2 \sum_{k,l=1}^{d} \delta_{kl} d \left( B^k_t - \frac{d}{|V^i|^2} B^k_t, B^l_t - \frac{d}{|V^i|^2} B^l_t \right) dt
\]

\[
= - 2(d-1) dt + 2 \sum_{k=1}^{d} \left[ 1 - 2 \frac{(V^i_k)^2}{|V^i|^2} + \sum_{p=1}^{d} \frac{(V^i_p)^2}{|V^i|^4} \right] dt = 0.
\]

Here we dropped the time dependence, wrote \(y = (y_1, \ldots, y_d) \in \mathbb{R}^d\) and used the fact that \(P(V^i) y = 0\) for all vectors \(y \in \mathbb{R}^d\). Hence \(|V^i| = 1\) up to explosion time. Since moreover \(dX^i_t = V^i_t dt\), this ensures that the explosion time is infinite, hence global existence and pathwise uniqueness for (6).

Now the solution to (6) for given \( \sigma_1, \sigma_2 \) is a solution to (4) since all velocities have norm 1, which provides global existence of solutions to (4). If we now consider two solutions to (4) for the same initial data and Brownian motions, then they have velocities equal to 1, so that are solutions with law \(\phi\) for all smooth \(x\). Existence and uniqueness for (6).

Existence and uniqueness for the artificial processes (5). Let \( \sigma_1, \sigma_2, \sigma_3 \) be any maps as above and let

\[
H_{\sigma_3}(f)(x) = \int_{\mathbb{R}^d} K(x - y) \sigma_3(v - w) f(y, w) dy dw.
\]

Then, given a distribution \( f_0 \) on \( \mathbb{R}^d \times \mathbb{S}^{d-1} \) with finite second moment in \( x \in \mathbb{R}^d \) and \((\mathbf{X}_0, \mathbf{V}_0)\) with law \(f_0\), the nonlinear equation

\[
\begin{aligned}
d\mathbf{X}_t &= \mathbf{V}_t dt, \\
d\mathbf{V}_t &= \sqrt{2} \sigma_1(\mathbf{V}_t) dB_t - \sigma_1(\mathbf{V}_t)(H_{\sigma_3} \ast f_t)(\mathbf{X}_t, \mathbf{V}_t) dt - (d-1) \sigma_2(\mathbf{V}_t) dt, \\
f_t &= \text{law}(\mathbf{X}_t, \mathbf{V}_t)
\end{aligned}
\]

(7)

has bounded and Lipschitz coefficients on \( \mathbb{R}^{2d} \), so admits a pathwise unique global solution according to [12, Theorem 1.1]. Moreover, as long as \(|\mathbf{V}_t| \geq 1/2\), then we can repeat the argument above to prove that \(d|\mathbf{V}_t|^2 = 0\), so that \(|\mathbf{V}_t| = 1\) for all time. In particular the obtained solution \((\mathbf{X}_t, \mathbf{V}_t)_{t \geq 0}\) is a global solution to the genuine nonlinear equation (5). Pathwise uniqueness of solutions to (5) can be obtained as for (4).

Existence and uniqueness for the PDE (3). Let \( f_0 \) be a distribution on \( \mathbb{R}^d \times \mathbb{S}^{d-1} \) with finite second moment in \( x \in \mathbb{R}^d \), \((\mathbf{X}_0, \mathbf{V}_0)\) with law \(f_0\), and let \((\mathbf{X}_t, \mathbf{V}_t)_{t \geq 0}\) be the solution to (5) with initial datum \((\mathbf{X}_0, \mathbf{V}_0)\). Its law \(f_t\), as a measure on \(\mathbb{R}^{2d}\), satisfies

\[
\frac{d}{dt} \int_{\mathbb{R}^{2d}} \varphi f_t = \int_{\mathbb{R}^{2d}} \left( v \cdot \nabla_x \varphi + \text{Hess}_x \varphi \cdot (I - v \otimes v) + \nabla_v \varphi \cdot (I - v \otimes v) (H \ast f_t) - (d-1) v \cdot \nabla_v \varphi \right) df_t
\]

for all smooth \(\varphi \) on \(\mathbb{R}^{2d}\) by the Itô formula; here \(\nabla_v \) and \(\Delta_v \) are respectively the gradient and Laplace operators with respect to \(v \in \mathbb{R}^d\), and \text{Hess}_x \varphi : M\) is the term by term product of the Hessian with respect to \(v\) matrix of \(\varphi\) with a matrix \(M\).
We have observed that $|\nabla f_i| = 1$ a.s., so $f_i$ is concentrated on $\mathbb{R}^d \times \mathbb{S}^{d-1}$. We now define the restriction $F_i$ of $f_i$ on $\mathbb{R}^d \times \mathbb{S}^{d-1}$ by

$$\int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} \Phi \, df_i = \int_{\mathbb{R}^d} \varphi \, df_i$$

for all continuous maps $\Phi$ on $\mathbb{R}^d \times \mathbb{S}^{d-1}$, where $\varphi$ is any continuous and bounded map on $\mathbb{R}^{2d}$ equal to $\Phi$ on $\mathbb{R}^d \times \mathbb{S}^{d-1}$. Let now $\Phi$ be a $C_c^\infty$ map on $\mathbb{R}^d \times \mathbb{S}^{d-1}$ and $\varphi$ be a $C_c^\infty$ map on $\mathbb{R}^{2d}$ such that $\varphi(x,v) = \Phi(x,v/|v|)$ for all $1/2 \leq |v| \leq 2$. Then $\varphi$ is 0-homogeneous in $v$ in the annulus $1/2 \leq |v| \leq 2$, so that $v \cdot \nabla \varphi = 0$ for all $(x,v)$ in the support of $f_i$. In particular

$$\frac{d}{dt} \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} \Phi \, df_i = \frac{d}{dt} \int_{\mathbb{R}^d} \varphi \, df_i = \int_{\mathbb{R}^{2d}} \left( v \cdot \nabla_x \varphi + \Delta_v \varphi + \nabla_v \cdot (I - v \otimes v)(H \ast f_i) \right) df_i.$$

Then the maps $v \cdot \nabla_x \Phi$ and $v \cdot \nabla_x \varphi$ are equal on $\mathbb{R}^d \times \mathbb{S}^{d-1}$ since $\Phi$ and $\varphi$ have the same $x$-dependence. Moreover, $\nabla_x \Phi = \nabla_v \varphi$ and $\Delta_x \Phi = \Delta_v \varphi$ for $(x, \omega) \in \mathbb{R}^d \times \mathbb{S}^{d-1}$. This last point can be checked by direct computations. Hence

$$\frac{d}{dt} \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} \Phi \, df_i = \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} (\omega \cdot \nabla_x \Phi + \Delta_x \Phi + \nabla_x \omega \cdot (I - \omega \otimes \omega)(H \ast f_i)) df_i.$$

This ensures that $F_i$ is a weak solution to (3).

We now turn to uniqueness of solutions to (3). For that purpose we let $f^1$ and $f^2$ be two solutions with the same initial datum $f_0$, and at each time $t$ we view them as measures on $\mathbb{R}^{2d}$ concentrated on the surface $\mathbb{R}^d \times \mathbb{S}^{d-1}$. We let $(X^i_t, \nabla^i_t)_{t \geq 0}$ and $(\bar{X}^i_t, \bar{\nabla}^i_t)_{t \geq 0}$ be the solutions to (7) with drift given by $H_{\sigma_1} \ast f^i_0$ and $H_{\sigma_2} \ast f^i_0$ respectively, and common initial datum $(\bar{X}_0, \bar{\nabla}_0)$ with law $f_0$. Then their respective laws $g^i_0$ and $g^i_0$, as measures on $\mathbb{R}^{2d}$, are solutions to the linear PDE

$$\partial_t g^i_t + v \cdot \nabla_x g^i_t = \sum_{k,l=1}^d \frac{\partial^2}{\partial y_k \partial y_l} \left( (\sigma_1 \sigma_1^*)_{kl} g^i_t \right) + \nabla_v \cdot \left[ g^i_t \left( \sigma_1 \left( H_{\sigma_2} \ast f^i_0 \right) + (d-1) \sigma_2 \right) \right].$$

Since $f^i_0$ is also a measure solution to this linear PDE on $\mathbb{R}^{2d}$ with bounded and regular coefficients, for which uniqueness classically holds, it follows that $g^i_0 = f^i_0$ ($i = 1, 2$). Consequently, the $(\bar{X}^i_t, \bar{\nabla}^i_t)_{t \geq 0}$ are solutions to the nonlinear SDE (7), for which we have already proved uniqueness. Hence $(\bar{X}^i_t, \bar{\nabla}^i_t)_{t \geq 0}$ and $(\bar{X}^i_t, \bar{\nabla}^i_t)_{t \geq 0}$ are equal, and in particular $f^1_t = f^2_t$.

**Proof of Theorem 2.** Since $|\nabla^i_t| = |\nabla^i_t| = 1$ for all $i$ and $t$, the processes $(X^i_t, \nabla^i_t)_{t \geq 0}$ and $(\bar{X}^i_t, \bar{\nabla}^i_t)_{t \geq 0}$ are solutions of the corresponding equations with bounded and Lipschitz diffusion and drifts coefficients as in (7). Hence we may apply the estimates in [12, Theorem 1.4] to obtain Theorem 2.

**Acknowledgments.** The authors wish to thank Pierre Degond and Vlad Panferov for drawing their attention to this question and for discussion on it. This note was written while the authors were visiting the Isaac Newton Institute, Cambridge; it is a pleasure for them to thank this institution for its kind hospitality.

The last two authors acknowledge support from the project MTM2008-06349-C03-03 DGI-MCI (Spain), the 2009-SGR-345 from AGAUR-Generalitat de Catalunya and the French-Spanish acciones integradas program FR2009-0019. All authors were partially supported by the ANR-08-BLAN-0333-01 Projet CBDif-Fr.

**References**


Université Paris-Dauphine, Ceremade, Place du Maréchal de Lattre de Tassigny, F-75775 Paris cedex 16, France

E-mail address: bolley@ceremade.dauphine.fr

Département de Mathématiques, Universitat Autònoma de Barcelona, E-08193 Bellaterra, Spain

E-mail address: canizo@mat.uab.es

ICREA (Institució Catalana de Recerca i Estudis Avançats) and Département de Matemàtiques, Universitat Autònoma de Barcelona, E-08193 Bellaterra, Spain

E-mail address: carrillo@mat.uab.es