MATHEMATICS OF EMERGENCE: DEVELOPING TRENDS IN ALIGNMENT DYNAMICS

ROMAN SHVYDKOY

Abstract. These notes are based on a series of lectures given at ICMAT, Madrid in November 2018, the Gran Sasso Science Institute, L’Aquila (Italy) in June 2018, and the joint working group at Ecole Normale Superieure and Laboratoire Jacques-Louis Lions at Universite Pierre et Marie Curie (Paris 6), Paris in May 2018, and Charles University in Prague in May 2019.

Contents

1. Background, Basic concepts, Levels of description ........................................ 2
2. Discrete Systems .................................................................................................. 4
   2.1. Momentum, Energy, and Maximum Principle ............................................. 4
   2.2. How much communication is needed? Connectivity and spectral method .. 5
   2.3. Alignment on $\mathbb{R}^n$. Non-degenerate case ....................................... 8
   2.4. Singular kernels and the issue of collisions ............................................... 11
   2.5. Degenerate communication. Corrector Method ....................................... 13
3. Dynamics under potential forces ......................................................................... 17
   3.1. Quadratic confinement .............................................................................. 17
   3.2. Attraction-Alignment models ..................................................................... 19
4. Multi-flocks .......................................................................................................... 25
5. Kinetic models ...................................................................................................... 29
6. Hydrodynamic models .......................................................................................... 29
   6.1. Flocking in hydrodynamic context .............................................................. 31
   6.2. Spectral method. Hydrodynamic connectivity .......................................... 31
   6.3. Topological models. Adaptive diffusion .................................................... 33
7. Regularity I: local well-posedness and continuation criteria ............................... 39
   7.1. Smooth models ........................................................................................... 39
   7.2. Singular models .......................................................................................... 43
8. One-dimensional theory ....................................................................................... 50
   8.1. Threshold condition for global existence with smooth kernels .................. 50
   8.2. Entropy. Csiszár-Kullback inequality. Shape of the limiting Flock .......... 51
   8.3. Alignment on $\mathbb{T}$ with degenerate kernel. The corrector method ......... 55
   8.4. Global existence and strong flocking for singular models ......................... 61
9. Regularity theory of multi-dimensional systems .................................................. 67
   9.1. Oriented flocks ........................................................................................... 67
   9.2. Mikado multi-flocks .................................................................................... 69
   9.3. Spectral dynamics approach ...................................................................... 69
   9.4. Small initial data in Hölder spaces ............................................................. 72

Date: October 22, 2019.

1991 Mathematics Subject Classification. 92D25, 35Q35, 76N10.

Key words and phrases. flocking, alignment, fractional dissipation, Cucker-Smale.

Acknowledgment. Research is supported in part by NSF grants DMS 1515705, DMS-1813351, Simons Foundation, and the College of LAS, UIC.
1. Background, Basic concepts, Levels of description

Agent-based models of collective behavior describe dynamics of a number of objects:

\[ x_i \in \Omega \subset \mathbb{R}^n, \quad i = 1, \ldots, N \]
\[ v_i = \dot{x}_i \]

governed by mutual communication - adjustment of velocity or position to that of nearby neighbors.

Emergence is a phenomenon of self-organization of a system of agents governed by local communication.

The long time dynamics of a self-organized system can be characterized by the following phenomena:

- **alignment**: \( \lim_{t \rightarrow \infty} \max_i |v_i - \bar{v}| = 0 \),
- **flocking**: \( \sup_{i,j} |x_i - x_j| \leq \bar{D} < \infty \),
- **strong flocking**: \( x_i - x_j \rightarrow \bar{x}_{ij}, \text{ as } t \rightarrow \infty \).
• **aggregation:** \( x_i - x_j \to 0, \) as \( t \to \infty \).

Note that alignment implies strong flocking provided it occurs at a sufficiently fast rate

\[
\int_0^\infty \max_{i,j} |v_i - v_j| \, dt < \infty.
\]

Indeed,

\[
x_i(t) - x_j(t) = x_i(0) - x_j(0) + \int_0^t [v_i(s) - v_j(s)] \, ds,
\]

hence

\[
\bar{x}_{ij} = x_i(0) - x_j(0) + \int_0^\infty [v_i(s) - v_j(s)] \, ds.
\]

Most common models that appear in various applications are

- **Vicsek discrete model, 1995:**

\[
\begin{align*}
v_i(k+1) &= v_0 \sum_{j:x_j-x_i < r_0} v_j \frac{v_j}{|\sum_{j:x_j-x_i < r_0} v_j|} + F_i, \\
x_i(k+1) &= x_i(k) + v_i(k+1).
\end{align*}
\]

Here \( F_i \) indicate additional forces.

- **Kuramoto synchronization model, \( \theta \in \mathbb{T}_1 \):**

\[
\dot{\theta}_i = \frac{\lambda}{N} \sum_{j \in N_i} \sin(\theta_j - \theta_i) + \omega_i.
\]

- **1st order environmental averaging:**

\[
\dot{p}_i = \lambda \sum_{j \in N_i} a_{ij}(t)(p_j - p_i) + F_i, \quad \sum_j a_{ij}(t) = 1,
\]

where \( N_i \) is a set of 'active' agents in local proximity, \( p_i \in \mathbb{R}^n \), e.g. \( p = x \) or \( p = v \).

- **2nd order Cucker-Smale type systems:**

\[
\begin{align*}
\dot{x}_i &= v_i, \\
\dot{v}_i &= \lambda \sum_{j=1}^N m_j \phi(x_i, x_j)(v_j - v_i) + F_i,
\end{align*}
\]

where, \( \phi(x, y) \geq 0 \) is a communication kernel and \( \Omega = \mathbb{R}^n \) or \( \mathbb{T}^n \).

The focus of these lectures will be on systems of Cucker-Smale type (2). We distinguish between the following general types of communication:

- **Absolute:** \( \inf_{x,y \in \Omega} \phi(x, y) > 0; \)
- **Global:** \( \phi(x, y) > 0, \) for all \( x, y \in \Omega; \)
- **Local:** there is a \( r_0 > 0 \) so that \( \phi(x, y) = 0 \) if \( |x - y| > r_0; \)
- **Symmetric:** \( \phi(x, y) = \phi(y, x); \)
- **Convolution type:** \( \phi(x, y) = \phi(|x - y|). \)
Sometimes by a local kernel, we mean a lack of global assumptions, for example, when the only information available is

$$\phi(x, y) \geq 1, \text{ for } |x - y| \leq r_0.$$  

A wealth of symmetric examples can be produced by convolution type kernels, where $\phi \in C^\infty(\mathbb{R}^+)$ is a non-negative function. Here are some examples that we will be discussing

$$\text{Cucker-Smale: } \phi(r) = \frac{h(r)}{(1 + r^2)^{\beta/2}}, \quad \beta > 0$$  

$$\text{Singular: } \phi(r) = \frac{h(r)}{r^\beta},$$  

$$\text{Extremely Singular: } \phi(r) = \frac{h(r)}{|r - \delta|^\beta}, \quad r > \delta,$$

where $h$ is a possible cut-off function if we consider local kernels. A measure of strength of communication in long and short range can be expressed by the following more general conditions

$$\text{Long range ("fat tail"): } \int_{r_0}^{\infty} \phi(r) \, dr = \infty$$  

$$\text{Short range: } \int_0^{r_0} \phi(r) \, dr = \infty.$$  

One example of a non-symmetric communication kernel was introduced by Motsch and Tadmor [7]:

$$\psi(x_i, x_j) = \frac{\phi(|x_i - x_j|)}{\sum_k m_k \phi(|x_i - x_k|)}.$$  

Note that MT-model defines an averaging communication protocol: $\sum_j m_j \psi(x_i, x_j) = 1$. Both singular and Motsch-Tadmor kernels are meant to emphasize local interactions over global ones whenever such communication is more realistic. These belong to a class of “geometric” kernels, meaning that the dependencies on other agents $x_k$’s are stated in terms of geometric distances to the center $x_i$.

A new class of “topological” kernels was introduced in [10]. Its construction is based on the following principle: communication between a pair of agents is determined by the density of crowd between them. Effectively, overcrowded areas impede communication, thus information spreads slower. We will present full description of this class of models in Section 2.

2. Discrete Systems

Much of the groundwork on the alignment will be laid out already on the discrete agent-based level of the Cucker-Smale dynamics. We start with basic properties of the system (2) and obtain the classical Cucker-Smale Theorem on emergence with long range interactions on $\mathbb{R}^n$.

2.1. Momentum, Energy, and Maximum Principle. Cucker-Smale systems with symmetric kernels, as opposed to non-symmetric, preserve the average total momentum of the system:

$$\bar{v} = \frac{1}{M} \sum_i m_i v_i, \quad M = \sum_{i=1}^N m_i, \quad \frac{d}{dt} \bar{v} = 0.$$  

Due to conservation of momentum, the center of mass moves with a constant velocity

$$\bar{x} = \frac{1}{M} \sum_i m_i x_i, \quad \frac{d}{dt} \bar{x} = \bar{v}.$$
For convolution type kernels, conservation of momentum can be used to shift the reference frame centered at $\bar{x}$, due to Galilean invariance of the system:

$$x_i \rightarrow x_i - t\bar{v}, \quad v_i \rightarrow v_i - \bar{v}. \tag{12}$$

So, in this case we can assume without loss of generality that $\bar{v} = 0$. In general, such translational invariance is not available. Nonetheless, if alignment occurs, then necessarily all $v_i \rightarrow \bar{v}$. In other words, we can determine the limiting velocity from initial conditions.

Let us consider the following variation and dissipation functions:

$$V_2 = \frac{1}{2} \sum_{i,j} m_i m_j |v_i - v_j|^2 \tag{13}$$

$$I_2 = \sum_{i,j} m_i m_j \phi(x_i, x_j) |v_i - v_j|^2.$$ 

The system has the classical kinetic energy as well defined by

$$\mathcal{E} = \frac{1}{2} \sum_{i=1}^{N} m_i |v_i|^2.$$ 

In case if $\bar{v} = 0$, then $V_2 = 2M\mathcal{E}$, however it is not prudent to use energy as a measure of alignment $\mathcal{E}$ in the non-symmetric case, simply because we do not know if $0$ would remain to be the momentum of the system for all time.

The following energy law is easily verified:

$$\frac{d}{dt} V_2 = -\lambda M I_2. \tag{14}$$

At this point one can obtain an $L^2$-base alignment result under absolute communication: if $\inf \phi = c_0 > 0$, then $I_2 \geq c_0 V_2$, and hence

$$\dot{V}_2 \leq -2c_0 \lambda M V_2.$$ 

Hence,

$$V_2(t) \leq V_2(0)e^{-2c_0 \lambda M t}.$$ 

This result provides exponential alignment "on average", specifically $L^2$-average, which does not translate well into individual information on agents. Indeed, one only obtains

$$|v_i - v_j| \leq \frac{1}{m_i m_j} V_2(0)e^{-\delta t}. \tag{15}$$

This estimate clearly deteriorates in the large crowd limit $N \rightarrow \infty$ when all masses vanish $m_i \rightarrow 0$.

In order to improve upon (15) we must resort to an $\ell^\infty$-based argument and use the maximum principle to be discussed later in Section 2.3.

2.2. How much communication is needed? Connectivity and spectral method. Let us assume for simplicity that all agents have the same mass $m_i = \frac{1}{N}$. It is clear that main "enemy" of alignment is lack of communication, especially at long range. In fact, it is easy to produce an example of initial condition that would diverge and not align if the kernel is local, see the Figure ??.

If global communication is not available, this situation can be remedied by either confined environmental condition, such as periodic boundary conditions to be discussed below in detail, or by an assumption of connectivity. The latter means that for any pair of agents $x_i$ and $x_j$ there exists a chain of agents $x_{k_1}, \ldots, x_{k_l}$ with end-points at $x_i$ and $x_j$ and such that all $|x_{k_p} - x_{k_{p+1}}| < r_0$. In this case one can recover alignment as follows. First, note that the shortest chain connecting
any pair of agents has no repeated agents in the chain. Hence, every chain is limited to length \( N \). Assuming that \( \lambda \phi(r_0) = \varepsilon > 0 \), we can estimate \( V_2 \) as follows:

\[
V_2 = \frac{1}{2N^2} \sum_{i \neq j} |v_i - v_j|^2 \leq \frac{1}{2N^2} \sum_{i \neq j} \sum_{p=1}^{P_{ij}} |v_{kp} - v_{kp+1}|^2 \\
\leq \frac{\lambda}{2\varepsilon N^2} \sum_{i \neq j} \sum_{p=1}^{P_{ij}} \phi(|x_{kp} - x_{kp+1}|)|v_{kp} - v_{kp+1}|^2 \\
\leq \frac{N(N - 1)}{2\varepsilon N^2} \lambda \sum_{k' \neq k''} \phi(|x_{k'} - x_{k''}|)|v_{k'} - v_{k''}|^2 \leq N^2 \mathcal{I}_2.
\]

So, we obtain a desired differential inequality for \( V_2 \): \( \dot{V}_2 \leq -\frac{1}{N^2} V_2 \).

The argument outlined above produces a bad dependence on \( N \), and as such is not suitable in the limit \( N \to \infty \). A continuous analogue of connectivity condition requires more elaboration and is discussed in [6].

The connectivity assumption at all times is hard to verify of course. However, it is guaranteed to hold provided the initial configuration is connected and the communication strength \( \lambda \) is large enough. This is because one can ensure in this case that the system aligns almost instantaneously before any disconnection becomes possible. Indeed, suppose initially the system is \( r_0/2 \) connected, and \( \lambda \phi(r_0) \) is large. Then for some large \( \Lambda > 0 \) we will have

\[
\frac{d}{dt} V_2 \leq -\Lambda V_2,
\]

for as long as the system is \( r_0 \)-connected. So, if a pair \( x_i, x_j \) is initially at most \( r_0/2 \) apart, then

\[
|x_i(t) - x_j(t)| \leq |x_i(0) - x_j(0)| + \frac{C}{\Lambda}.
\]

This shows that the same pair will never get \( r_0 \)-disconnected and hence the connectivity is preserved at all times.

Even if connectivity is intermittent, but reoccurring, one can still achieve the alignment effect if the flock reconnects “more often then not”. There is an elegant algebraic way to make this statement qualitatively precise and in fact it lies in the heart of the original spectral approach introduced by Cucker and Smale in [3], see also Motsch and Tadmor’s implementation to their model in [8], and to topological models in [10] and Section 6.2.

Let us consider the matrix \( A = \{a_{ij}(t)\}_{i,j=1}^{N} \otimes \mathbb{I}_{n \times n} \), where \( a_{ij} = \lambda \phi(x_i(t), x_j(t)), i \neq j \), and \( a_{ij} = 0 \), if \( i = j \). Note that \( a_{ij} \neq 0 \) if and only if the corresponding agents and “connected” in the sense that they communicate through the influence kernel \( \phi \). By analogy with the graph theory we call it the adjacency matrix of the flock. In fact, one can consider the actual adjacency matrix associated with the flock under the above connectivity definition: \( A = \{\hat{a}_{ij}(t)\}_{i,j=1}^{N} \otimes \mathbb{I}_{n \times n} \), where \( \hat{a}_{ij} = 1 \) if \( a_{ij} \neq 0 \), and 0 otherwise. Let \( D = \text{diag}\{b_1, \ldots, b_N\} \otimes \mathbb{I}_{n \times n} \), where \( b_i = \sum_j a_{ij} \), and \( \hat{D} = \text{diag}\{\hat{b}_1, \ldots, \hat{b}_N\} \otimes \mathbb{I}_{n \times n} \), where \( \hat{b}_i = \sum_j \hat{a}_{ij} \). Note that each \( \hat{b}_i \) is precisely the degree of the vertex \( x_i \), i.e. the number of other agents to which it is connected.

In this notation the system (2) can now be written in terms of the grand velocity vector \( \mathbf{V} = (v_1, \ldots, v_N) \), and the Laplacian associated to \( A \), \( L = D - A \), namely,

(16) \[
\frac{d}{dt} \mathbf{V} = -L \mathbf{V}.
\]

The matrix \( L \) is non-negative definite, and hence the spectrum consists of a sequence \( 0 = \kappa_1 \leq \kappa_2 \leq \ldots \leq \kappa_N \). The average grand vector \( \bar{\mathbf{V}} = (\bar{v}_1, \ldots, \bar{v}_N) \) is obviously a member of the kernel of \( L \), hence \( \kappa_1 = 0 \). The next eigenvalue \( \kappa_2 \) is called the Fiedler number, although the classical Fiedler
number is one associated to the Laplacian $\hat{L} = \hat{D} - \hat{A}$ is a similar way, we denote it $\hat{\kappa}_2$. By the min-max theorem $\kappa_2$ is given by

$$\kappa_2 = \min_{\sum_i v_i = 0} \frac{\langle L V, V \rangle}{|V|^2}.$$  

A simple fact to verify is that $\kappa_2 \neq 0$ if and only if the graph is connected, and the relationship to $\hat{\kappa}_2$ is given by (see [3, Proposition 2]):

$$\kappa_2 \geq \hat{\kappa}_2 \min_{i,j: a_{ij} \neq 0} a_{ij}.$$  

For this reason, we can call $\kappa_2$ a weighted Fiedler number, which captures not only algebraic connectivity of the flock as a graph, but also the collective strength of the connection weighted by the kernel $\phi$. With the use of the weighted Fiedler number $\kappa_2 = \kappa_2(t)$, which, let’s recall, depends on time, we can measure alignment in (16) by writing the energy law as

$$\frac{d}{dt} |V - \bar{V}|_2^2 = -2 \langle L(V - \bar{V}), (V - \bar{V}) \rangle \leq -2 \kappa_2(t) |V - \bar{V}|_2^2.$$  

Consequently,

$$|V(t) - \bar{V}|_2 \leq |V_0 - \bar{V}|_2 \exp \left\{ - \int_0^t \kappa_2(s) ds \right\}.$$  

We can see that the divergence of the integral inside leads to alignment, which can be seen as a quantitative measure of connectivity as a function of time. Let us state it precisely.

**Lemma 2.1.** If $\int_0^\infty \kappa_2(s) ds = \infty$, then the flock aligns.

At the center of the original result of Cucker and Smale was a statement that for kernels of type (4) the integral of the weighted Fiedler number is indeed divergent. Here, clearly, the flock remains always algebraically connected, thus $\hat{\kappa}_2 = N$, and one can appeal directly to (18) to restate the problem in terms of control on the decay of the adjacency matrix. The original theorem of Cucker and Smale [2, 3] is the following.

**Theorem 2.2.** Let $\phi(r) = \frac{1}{(1+r^2)^{\beta/2}}$. Then every solution aligns exponentially and flock strongly for $\beta \leq 1$, and conditionally if $\beta > 1$.

In Section 2.3 this result will be proved using a more direct approach due to Ha and Liu, [4], which paves a way to extensions into meso- and macroscopic systems. Let us not, however, underestimate the spectral method as sometimes it is the only one available if no specific structural information is known about the kernel.

To see that the non-integrable decay rate of the kernel is necessary in the Cucker-Smale theorem, let us consider the following example.

**Example 2.3.** Let the kernel be $\phi(r) = \frac{1}{r^2}$ for $r > r_0$, for simplicity, and let $x = x_1 = -x_2 > r_0$ and $v = v_1 = -v_2 > 0$. This symmetry is preserved in time. Then the system (2) becomes

$$\frac{dx}{2^{\beta-1}x^{\beta}} + \frac{dv}{1} = 0.$$
This equation admits a conservation law
\[ J = v + \frac{1}{2^{\beta-1}(1-\beta)x^{\beta-1}}. \]
If \( \beta > 1 \), and the initial velocity is large enough, then \( J(t) = J_0 > 0 \), and hence \( v(t) \geq J_0 \) holds true for all times. This creates permanent misalignment between the two velocities at hand: \( v \) and \(-v\).

2.3. **Alignment on \( \mathbb{R}^n \). Non-degenerate case.** The main goal in this section will be to prove the Cucker-Smale Theorem 2.2 in more general settings of convolution type fat tail kernel using a method based on the maximal principle. We will see that the argument is also easily adaptable to the non-symmetric case of the Motsch-Tadmor kernel. So, we consider the classical Cucker-Smale system

\[
\begin{align*}
\dot{x}_i &= v_i, \\
\dot{v}_i &= \lambda \sum_{k=1}^{N} m_k \phi(x_i - x_k)(v_k - v_i), \quad (x_i, v_i) \in \mathbb{R}^n \times \mathbb{R}^n.
\end{align*}
\]

We assume that \( \phi \) is monotonically decreasing and everywhere positive. It will be useful to use sometimes the following shortcut notation:

\[
\begin{align*}
x_{ij} &= x_i - x_j, \quad v_{ij} = v_i - v_j, \quad \phi_{ij} = \phi(x_i - x_j), \quad \text{etc.}
\end{align*}
\]

Let us also consider the amplitude and flock diameter:

\[
\begin{align*}
D &= \max_{i,j} |x_i - x_j|, \quad A = \max_{i,j} |v_i - v_j|.
\end{align*}
\]

**Theorem 2.4** (Alignment for Cucker-Smale model). Any solution to the system (19) with initial condition \( D_0, A_0 \) satisfying

\[
\int_{D_0}^{\infty} \phi(r) \, dr > \frac{A_0}{\lambda M},
\]

aligns and flocks exponentially fast:

\[
\sup_{t \geq 0} D(t) \leq \bar{D}, \quad A(t) \leq A_0 e^{-t\lambda M \phi(D)}.
\]

In particular, every solution flocks provided the kernel satisfies the fat tail condition (7).

To make the proof perfectly rigorous, let us recall the classical Redemacher Lemma, which we will use throughout.

Suppose \( f(x,t) : X \times \mathbb{R}_+ \to \mathbb{R} \) is a Lipschitz in time function uniformly in \( x \), where \( X \) is an arbitrary index set. So, \( \exists L > 0 \) such that for all \( t, s, x \) we have

\[ |f(x,t) - f(x,s)| \leq L|t - s|. \]

Suppose that at any time \( t \) there is a point \( x(t) \in X \) such that

\[ f(x(t), t) = \sup_{x \in X} f(x,t) := M(t). \]

Note that \( M(t) \) is a Lipschitz function with the same constant \( L \). Indeed, let \( t, s \in \mathbb{R}_+ \) and \( M(t) > M(s) \). Then

\[ M(t) - M(s) = f(x(t), t) - f(x(t), s) + f(x(t), s) - f(x(s), s) \leq f(x(t), t) - f(x(t), s) \leq L|t - s|. \]
Consequently, $M$ is absolutely continuous on any finite interval, i.e.,

$$M(t) - M(s) = \int_s^t m(\tau) d\tau,$$

where $\|m\|_\infty \leq L^\infty_{\text{loc}}$, and hence $M' = m$ a.e.

**Lemma 2.5.** If $f(x, \cdot)$ is differentiable everywhere in $t$ for all $x \in X$, then $M'(t) = \partial_t f(x(t), t)$ holds at any point where $M'$ exists.

Indeed computing one-sided derivative from the right we have

$$M'(t) = \lim_{h \to 0^+} \frac{f(x(t + h), t + h) - f(x(t), t + h) + f(x(t), t + h) - f(x(t), t)}{h}$$

$$\geq \lim_{h \to 0^+} \frac{f(x(t), t + h) - f(x(t), t)}{h} = \partial_t f(x(t), t).$$

Taking $h < 0$ proves the opposite inequality.

**Proof of Theorem 2.4.** Let us represent $A$ as

$$A = \max_{|\ell| = 1, i,j} \ell(v_i - v_j).$$

Note that the maximum is taken over a fixed compact set not changing in time. So, using Rademacher’s lemma we can pick $\ell$ and $i,j$ at each time $t$ for which the maximum is achieved. Using the velocity equation we obtain

$$\frac{d}{dt} \ell(v_{ij}) = \lambda \sum_{k=1}^N m_k \phi_{ik} \ell(v_{ki}) - m_k \phi_{jk} \ell(v_{kj})$$

$$= \lambda \sum_{k=1}^N m_k \phi_{ik} [\ell(v_{kj}) - \ell(v_{ij})] + m_k \phi_{jk} [\ell(v_{ik}) - \ell(v_{ij})]$$

$$\leq \lambda \phi(D) \sum_{k=1}^N m_k [\ell(v_{kj}) - \ell(v_{ij}) + \ell(v_{ik}) - \ell(v_{ij})]$$

$$= -\lambda M \phi(D) \ell(v_{ij}).$$

So, we obtain

$$\begin{cases}
\frac{d}{dt} A \leq -\lambda M \phi(D) A \\
\frac{d}{dt} D \leq A.
\end{cases}$$

This system of ordinary differential inequalities (ODIs) has a decreasing Lyapunov function given by $L = A + \lambda M \int_0^D \phi(r) dr$. This, in particular, implies that

$$\lambda M \int_0^{D(t)} \phi(r) dr \leq A_0 + \lambda M \int_0^{D_0} \phi(r) dr, \quad \forall t > 0.$$

Consequently, $D(t) \leq \bar{D}$, where $\bar{D}$ is obtained from the equation

$$\lambda M \int_{D_0}^D \phi(r) dr = A_0,$$

which is guaranteed to have a finite solution due to (29). Then, $\dot{A} \leq -\lambda M \phi(D) A$ and the theorem follows. \(\square\)
Solving equation (26) allows one to provide explicit decay rates for solutions of (25) for some kernels. In particular, for the classical Cucker-Smale kernel
\[ \phi(r) = \frac{1}{(1 + r^2)^{\frac{\beta}{2}}}, \]
one obtains
\[ \bar{D} \leq \left( \left[ \frac{1 - \beta}{\lambda M} A_0 + (1 + D_0^2)^{\frac{1-\beta}{2}} \right]^{\frac{1}{1-\beta}} - 1 \right)^{\frac{1}{2}}, \quad \beta < 1, \]
rate \[ = \frac{\lambda M}{\left[ \frac{1 - \beta}{\lambda M} A_0 + (1 + D_0^2)^{\frac{1-\beta}{2}} \right]^{\frac{1}{1-\beta}}}. \]
(27)
(here one replaces \( \phi \) with a smaller but explicitly integrable kernel \( \frac{r}{(1+r^2)^{\frac{\beta}{2}+1}} \)), and
\[ \bar{D} \leq \left( e^{\frac{\lambda}{N M} A_0} (1 + D_0^2) - 1 \right)^{\frac{1}{2}}, \quad \text{rate} = \frac{\lambda M}{e^{\frac{\lambda A_0}{N M} (1 + D_0^2)^{\frac{1}{2}}}}, \quad \beta = 1. \]
(28)

Let us note that in the symmetric case we know that the limit of all velocities will be the conserved momentum \( \bar{v} \). However, the argument above does not use the symmetry of the kernel. This allows to adopt it, for example, to the Motsch-Tadmor model we address in the next theorem.

**Theorem 2.6** (Alignment for Motsch-Tadmor model). Consider the forceless system (2) with kernel given by (9). If the initial condition \( D_0, A_0 \) satisfies
\[ \int_{D_0}^{\infty} \phi(r) \, dr > \frac{A_0|\phi|_{\infty}}{\lambda}, \]
then the flock aligns exponentially fast:
\[ \sup_{t \geq 0} D(t) \leq \bar{D}, \quad A(t) \leq A_0 e^{-\lambda|\phi|_{\infty} \bar{D}}. \]
(30)
In particular, every solution flock provides the kernel satisfies the fat tail condition (7).

**Proof.** It suffices to just revisit the computation with the maximal functional:
\[ \frac{d}{dt} \ell(v_{ij}) = \lambda \sum_{k=1}^{N} \sum_{p=1}^{N} \frac{m_k \phi_{ik}}{m_p \phi_{ip}} \ell(v_{kj}) - \frac{m_k \phi_{jk}}{m_p \phi_{jp}} \ell(v_{kj}) \]
\[ = \lambda \sum_{k=1}^{N} \sum_{p=1}^{N} \frac{m_k \phi_{ik}}{m_p \phi_{ip}} \left[ \ell(v_{kj}) - \ell(v_{ij}) \right] + \frac{m_k \phi_{jk}}{m_p \phi_{jp}} \left[ \ell(v_{ik}) - \ell(v_{ij}) \right] \]
\[ \leq \frac{\lambda}{|\phi|_{\infty}} \phi(D) \sum_{k=1}^{N} \frac{m_k}{M} \left[ \ell(v_{kj}) - \ell(v_{ij}) + \ell(v_{ik}) - \ell(v_{ij}) \right] \]
\[ = -\frac{\lambda}{|\phi|_{\infty}} \phi(D) \ell(v_{ij}). \]
(31)
The rest of the proof is similar. \( \square \)

Let us observe two distinctive features of the MT model versus CS model. First, the rate of alignment is independent of the mass. Second, the actual velocity vector to which the system aligns is not determined by initial condition as in the CS case, but rather becomes an emergent quantity in the long time dynamics. Since all the differences \( v_{ij} \) vanish exponentially fast, it implies
that velocities do in fact converge a time independent limit as seen from integrating the velocity equation:

\[
\lim_{t \to \infty} v_i(t) = v_i(0) + \int_0^\infty \lambda \sum_{k=1}^N \frac{m_k \phi_{ik}}{\sum_{p=1}^N m_p \phi_{ip}} v_{ki}(s) \, ds.
\]

It is however hard to predict what that limit will be from the initial conditions only.

2.4. Singular kernels and the issue of collisions. Before we focus solely on alignment, let us introduce into consideration the singular kernels (5), which will play a crucial role in macroscopic description of the system. Clearly, singularity at the origin emphasizes predominantly local communication as is needed in applications. However, with singularity we stumble upon the issue of well-posedness of the system (2) in the case when agents encounter collisions. In fact collisions are common in the bounded kernel case.

Example 2.7. Let us assume that \( \phi = 1 \) in a neighborhood of 0. Let us arrange two agents \( x = x_1 = -x_2 \) with \( 0 < x(0) = \varepsilon \ll 1 \). And let \( v_1 = -v_2 < 0 \) be very large. Clearly \( x(t) \) will remain in the same neighborhood of 0 as where it has started, and so the system reads

\[
\frac{dx}{dt} = v, \quad \frac{dv}{dt} = -2v.
\]

Solving it explicitly we can see that the two agents will collide at the origin.

Heuristically, however, strong singularity should prevent such collisions. The induced alignment forces should become strong enough to correct velocities of converging agents before collision happens in the first place. Exactly how singular the kernel should be can be seen from the following example.

Example 2.8. Let the kernel be given by (5) and let us consider the same setup as previously. Then we obtain the system

\[
\frac{dx}{dt} = v, \quad \frac{dv}{dt} = -2 \frac{v}{x^\beta}.
\]

This system has a conservation law provided \( \beta < 1 \): \( v + \frac{2v^{1-\beta}}{1-\beta} = C_0 \). So, if initially \( C_0 \ll 0 \), then \( v < C_0 \ll 0 \) as well. This means that \( x \) will reach the origin in finite time.

This example demonstrates that the threshold singularity necessary to prevent collisions must be non-integrable. It is indeed true as we prove in the following theorem.

Theorem 2.9. Under the strong singularity condition (8) the flock experiences no collisions between agents for any non-collisional initial datum. Consequently, any non-collisional initial datum gives rise to a unique global solution.

In view of global existence and absence of collisions the content of Theorem 2.4 holds true as stated provided the kernel is singular (8) condition.

Proof. Let us assume that for a given non-collisional initial condition \((x_i,v_i)\) a collision occurs at time \( T^* \) for the first time. Let \( I^* \subset I = \{1, ..., N\} \) be one set of indexes of the agents that collided at one point (note that other groups of agents may collide at the same time at other points of space). Hence, there exists a \( \delta > 0 \) such that \( |x_{ik}(t)| \geq \delta \) for all \( i \in I^* \) and \( k \in \Omega \setminus I^* \). Denote

\[
D^*(t) = \max_{i,j \in I^*} |x_{ij}(t)|, \quad V^*(t) = \sum_{i,j \in I^*} |v_{ij}(t)|^2.
\]

Directly from the characteristic equation we obtain \( |\dot{D}^*| \leq \sqrt{V^*} \), and hence

\[
-\dot{D}^* \leq \sqrt{V^*}.
\]
From the momentum equation we obtain
\[ \frac{d}{dt} V^* = \frac{2}{N} \sum_{k \in I, i,j \in I^*} \phi_{ik} v_{ki} \cdot v_{ij} - \phi_{kj} v_{kj} \cdot v_{ij}. \]

Switching \( i, j \) in the second sum results in the same as first sum. So, we obtain
\[ \frac{d}{dt} V^* \leq \frac{4}{N} \sum_{k \in I^*} \phi_{ki} v_{ki} \cdot v_{ik} \leq C_1 \sqrt{V^*}. \]

Note that in the first sum the agents are separated, and all velocities are bounded by the maximum principle. So, we can estimate it by
\[ \sqrt{V^*} \leq C_1 \sqrt{V^*} + \frac{2}{N} \sum_{k,i,j \in I^*} \phi_{ki} v_{ki} \cdot v_{ij} \leq C_1 \sqrt{V^*}. \]

We thus obtain a system
\[
\begin{cases}
\frac{d}{dt} \sqrt{V^*} \leq -C_2 \phi(D^*) \sqrt{V^*} + C_1 \\
-\frac{d}{dt} D^* \leq \sqrt{V^*}.
\end{cases}
\]

Let us consider the energy functional
\[ E(t) = \sqrt{V^*(t)} + C_2 \int_{D^*(t)} \phi(r) \, dr. \]

We readily find that \( \frac{d}{dt} E \leq C_1 \), hence \( E \) remains bounded up to the critical time. This means that \( D^*(t) \) cannot approach zero value.

The global existence part is now a routine application of the Picard iteration and the standard continuation argument. \( \square \)

A quantitative version on the minimal distance between agents can be obtained in the case of power kernels (5) with \( \beta \geq 2 \). Since, clearly, collision is a local phenomenon, we only need singularity assumption in the local region \( r < r_0 \). So, let us consider the local version of the collision functional introduced in [1] for \( \beta \geq 2 \):

\[
C = \begin{cases} 
\frac{1}{N^2} \sum_{i,j=1}^{N} \frac{1}{(|x_{ij}| \wedge r_0)^{\beta-2}}, & \beta > 2 \\
\frac{1}{N^2} \sum_{i,j=1}^{N} \ln(|x_{ij}| \wedge r_0), & \beta = 2.
\end{cases}
\]

Let us compute the time-derivative of \( C \) for \( \beta > 2 \):
\[
\frac{dC}{dt} = (2 - \beta) \frac{N}{N^2} \sum_{i,j=1}^{N} \frac{d}{dt} \left( \frac{|x_{ij}| \wedge r_0}{(|x_{ij}| \wedge r_0)^{\beta-1}} \right) \leq |\beta - 2| \frac{1}{N^2} \sum_{i,j=1}^{N} \frac{1}{(|x_{ij}| \wedge r_0)^{\beta-1}} |v_{ij}| \mathbf{1}_{|x_{ij}| < r_0}
\]
\[
\leq |\beta - 2| \left( \frac{1}{N^2} \sum_{i,j=1}^{N} v_{ij}^2 \frac{1}{|x_{ij}|^{2-\beta}} \mathbf{1}_{|x_{ij}| < r_0} \right)^{1/2} \left( \frac{1}{N^2} \sum_{i,j=1}^{N} |x_{ij}|^{2-2\beta} \mathbf{1}_{|x_{ij}| < r_0} \right)^{1/2}
\]
\[
\leq C \sqrt{I_2} \sqrt{C}.
\]
This implies
\[(35) \quad \sqrt{C(t)} \leq \sqrt{C(0)} + C \int_0^t \sqrt{\mathcal{I}_2(s)} \, ds,\]
and recalling that \(\mathcal{I}_2\) is integrable on \(\mathbb{R}^+\) we conclude
\[(36) \quad C(t) \lesssim t.\]
For \(\beta = 2\), exact same computation gives \(\frac{d}{dt} C \leq C \sqrt{\mathcal{I}_2}\), hence \(C(t) \lesssim \sqrt{t}\). We thus arrive at the following bounds
\[(37) \quad |x_{ij}(t)| \geq \begin{cases} \frac{c}{t^{(\beta-2)/2}}, & \beta > 2, \\ ce^{-C \sqrt{t}}, & \beta = 2. \end{cases}\]

2.5. Degenerate communication. Corrector Method. Let us consider higher order variations
\[\mathcal{V}_p = \frac{1}{pN^2} \sum_{i,j=1}^N |v_i - v_j|^p, \quad p \geq 1,\]
\[(38) \quad \mathcal{I}_p = \frac{1}{N^2} \sum_{i,j=1}^N \phi(x_i, x_j) |v_i - v_j|^p.\]

We observe that \(\mathcal{V}_p\)'s are non-increasing. Indeed,
\[\frac{d}{dt} \mathcal{V}_p = \frac{1}{N^3} \sum_{i,j,k} |v_{ij}|^{p-2} v_{ij} \cdot (v_{ki} \phi_{ki} - v_{kj} \phi_{kj}) = \frac{2}{N^3} \sum_{i,j,k} |v_{ij}|^{p-2} v_{ij} \cdot v_{ki} \phi_{ki}\]
\[= \frac{1}{N^3} \sum_{i,j,k} (|v_{ij}|^{p-2} v_{ij} - |v_{kj}|^{p-2} v_{kj}) \cdot v_{ki} \phi_{ki}\]
\[= \frac{1}{N^3} \sum_{i,j,k} (|v_{ij}|^{p-2} v_{ij} - |v_{kj}|^{p-2} v_{kj}) \cdot (v_{kj} - v_{ij}) \phi_{ki},\]
with the convention that \(|v_{ij}|^{p-2} v_{ij} = 0\) if \(i = j\). The right hand side is non-positive due to the elementary inequality
\[(|a|^{p-2} a - |b|^{p-2} b) \cdot (a - b) \geq 0.\]
The two special cases, \(i = j\) and \(k = j\), produce the term \(-|v_{ik}|^p - |v_{ij}|^p\). So, in general we have an \(N\)-dependent inequality
\[(39) \quad \frac{d}{dt} \mathcal{V}_p \leq -\frac{1}{N} \mathcal{I}_p, \quad p \geq 1.\]
For the case \(p = 2\) we have the \(N\)-independent energy law:
\[(40) \quad \frac{d}{dt} \mathcal{V}_2 = -\mathcal{I}_2.\]

**Theorem 2.10.** Suppose the kernel \(\phi \geq 0\) is either smooth or satisfying (8). Suppose also that it dominates a monotone fat tail: there exists a non-increasing \(\Phi(r)\), such that for some \(r_0 > 0\)
\[(41) \quad \phi(r) \geq \Phi(r), \forall r > r_0, \quad \text{and} \quad \int_{r_0}^{\infty} \Phi(r) \, dr = \infty.\]
Then
(i) Any solution to the discrete system (2) aligns: \(\mathcal{V}_2(t) \leq \frac{C_N}{t}\), and \(D(t) \leq D_N\), with constants depending on \(N\).
(ii) If \( \phi \) is smooth, then any solution to the discrete system (2) aligns: \( V_4 \to 0 \), with a rate independent of \( N \).

The advantage of (i) over (ii) is that it gives a faster, although \( N \)-dependent, rate as well as flocking. It also holds for singular kernels satisfying non-collision condition (8). However, it is not extendable to the macroscopic or kinetic case, while (ii) is.

Proof of Theorem 2.10 (i). Let us consider the following “distance” – the projection of the displacement \( x_{ij} \) onto the relative velocity direction taken with opposite sign:

\[
d_{ij} = -x_{ij} \cdot \frac{v_{ij}}{|v_{ij}|}.
\]

Next, define two auxiliary functions \( \chi : \mathbb{R}^+ \mapsto \mathbb{R}^+ \) and \( \psi : \mathbb{R} \mapsto \mathbb{R}^+ \) by

\[
\chi(r) = \begin{cases} 
1, & r < r_0, \\
2 - \frac{r}{r_0}, & r_0 \leq r \leq 2r_0, \\
0, & r > 2r_0
\end{cases}
\quad \text{and} \quad
\psi(d) = \begin{cases} 
0, & d < -r_0, \\
|d| - r_0, & |d| \leq r_0, \\
2r_0, & d > r_0.
\end{cases}
\]

We consider the following corrector function with spatial truncation

\[
G = \frac{1}{N^2} \sum_{i,j=1}^{N} |v_{ij}| \psi(d_{ij}) \chi(|x_{ij}|),
\]

and compute its derivative:

\[
\frac{d}{dt} G = -\frac{1}{N^2} \sum_{i,j=1}^{N} |v_{ij}|^2 \mathbb{1}_{|d_{ij}| < r_0} \chi(|x_{ij}|) + \mathcal{R}_1 + \mathcal{R}_2 + \mathcal{R}_3,
\]

where

\[
\mathcal{R}_1 = \frac{2}{N^3} \sum_{i,j,k=1}^{N} \frac{v_{ij}}{|v_{ij}|} \cdot v_{ki} \phi(x_{ik}) \psi(d_{ij}) \chi(|x_{ij}|),
\]

\[
\mathcal{R}_2 = -\frac{2}{N^3} \sum_{i,j,k=1}^{N} x_{ij} \cdot \left( \mathbb{1} - \frac{v_{ij} \otimes v_{ij}}{|v_{ij}|^2} \right) v_{ki} \phi(x_{ik}) \mathbb{1}_{|d_{ij}| < r_0} \chi(|x_{ij}|)
\]

\[
\mathcal{R}_3 = \frac{1}{N^2} \sum_{i,j=1}^{N} |v_{ij}| \psi(d_{ij}) \chi(|x_{ij}|) \frac{x_{ij}}{|x_{ij}|} \cdot v_{ij}.
\]

As seen from the first gain term in (43), the role of \( G \) is compensate for missing communication in the close range. Let us address this term first. Without loss of generality we can assume that \( \Phi \) is a bounded decreasing function on \( \mathbb{R}^+ \). Hence, we have

\[
\mathbb{1}_{r < r_0} + \phi(r) \geq c\Phi(r),
\]

for all \( r > 0 \) and some \( c > 0 \). Using that \( \mathbb{1}_{|d_{ij}| < r_0} \geq \mathbb{1}_{|x_{ij}| < r_0} \) we obtain

\[
-\frac{1}{N^2} \sum_{i,j} |v_{ij}|^2 \mathbb{1}_{|d_{ij}| < r_0} \chi(x_{ij}) \leq -\frac{1}{N^2} \sum_{i,j} |v_{ij}|^2 \mathbb{1}_{|x_{ij}| < r_0} \chi(x_{ij})
\]

\[
= -\frac{1}{N^2} \sum_{i,j} |v_{ij}|^2 (\mathbb{1}_{|x_{ij}| < r_0} + \phi(x_{ij})) + \frac{1}{N^2} \sum_{i,j} |v_{ij}|^2 \phi(x_{ij})
\]

\[
\leq -c\Phi(D) V_2 + I_2,
\]

where \( D = D(t) \) is the diameter of the flock.
Let us proceed to the error terms. By construction of the functions,

\[ |R_1|, |R_2| \lesssim I_1, \]

and

\[ R_3 \leq \frac{1}{N^2} \sum_{i,j} |v_{ij}|^2 |\chi'(x_{ij})| \leq \frac{1}{N^2} \sum_{i,j} |v_{ij}|^2 I_{r_0 < |x_{ij}| < 2r_0} \lesssim \frac{1}{N^2} \sum_{i,j} |v_{ij}|^2 \phi(x_{ij}) = I_2. \]

Hence, we obtain, for constants \( a, b, c \) only depending on \( \phi \), that

\[ \frac{d}{dt} G \leq -c \Phi(D) V_2 + a I_2 + b I_1. \]

Let us form the following functional

\[ L = G + a V_2 + b N V_1. \]

Then

\[ \frac{d}{dt} L \leq -c \Phi(D) V_2. \]

Together with the trivial bound \( \frac{d}{dt} D \leq C_N \sqrt{V_2} \), we find another Lyapunov functional (noting that \( \sqrt{V_2} \) is monotonely decreasing)

\[ \tilde{L} = L + \frac{c}{C_N} \sqrt{V_2} \int_0^r \Phi(r) dr. \]

Hence, \( D(t) \leq D_N \) for all time and for some \( N \)-dependent \( D_N \). Going back to (44) we conclude that

\[ A := \int_0^\infty V_2(t) \, dt < \infty. \]

So, on every time interval \([T, e^A T]\) we find a \( t \) such that \( V_2(t) \leq \frac{1}{r} \). By monotonicity of \( V_2 \) this implies a similar bound for all \( t \). \( \square \)

**Proof of Theorem 2.10 (ii).** Here we consider the third order corrector function

\[ G_3 = \frac{1}{N^2} \sum_{i,j} |v_{ij}|^3 \psi(d_{ij}) \chi(|x_{ij}|), \]

with \( \psi \) and \( \chi \) defined in the previous proof. We find

\[ \frac{d}{dt} G_3 = -\frac{1}{N^2} \sum_{i,j=1}^N |v_{ij}|^4 1_{|d_{ij}| < r_0} \chi_{ij} + R_1 + R_2 + R_3, \]

with

\[ R_1 = \frac{6}{N^3} \sum_{i,j,k=1}^N |v_{ij}| |v_{ij} \cdot v_{jk} \phi_{ik} \psi_{ij} \chi_{ij}, \]

\[ R_2 = -\frac{2}{N^3} \sum_{i,j,k=1}^N |v_{ij}|^2 x_{ij} \cdot \left( I - \frac{v_{ij} \otimes v_{ij}}{|v_{ij}|^2} \right) v_{ki} \phi_{kj} 1_{|d_{ij}| < r_0} \chi_{ij}, \]

\[ R_3 = \frac{1}{N^2} \sum_{i,j=1}^N |v_{ij}|^3 \psi_{ij} \chi(|x_{ij}|) \frac{x_{ij}}{|x_{ij}|} \cdot v_{ij}. \]
The gain term is estimated as before using that we have an a priori uniform bound on all velocities \(|v_i(t)| \leq |v(0)|_\infty\):

\[\frac{1}{N^2} \sum_{i,j=1}^N |v_{ij}|^4 I_{|d_{ij}|<r_0} \chi_{ij} \leq -c \Phi(D) V_4 + |v(0)|_\infty^2 I_2.\]  

(45)

Next, again as before,

\[R_3 \lesssim \frac{1}{N^2} \sum_{i,j} |v_{ij}|^2 \phi(|x_{ij}|) \lesssim I_2.\]

By Young’s inequality we find for \(\varepsilon > 0\)

\[R_2 \lesssim \frac{\varepsilon}{N^2} \sum_{i,j=1}^N |v_{ij}|^4 \chi_{ij}^2 + \frac{1}{\varepsilon N^2} \sum_{i,k=1}^N |v_{ki}|^2 \phi_{ki}.\]

However, \(\chi_{ij}\) is vanishing if \(|x_{ij}| > 2r_0\) so that we find

\[R_2 \lesssim \frac{\varepsilon}{N^2} \sum_{i,j=1}^N |v_{ij}|^4 \chi_{ij} |\leq 2r_0 + \frac{1}{\varepsilon N^2} \sum_{i,k=1}^N |v_{ki}|^2 \phi_{ki} \]

\[\leq \frac{\varepsilon}{N^2} \sum_{i,j=1}^N |v_{ij}|^4 \chi_{ij} |\leq r_0 + \frac{\varepsilon}{N^2} \sum_{i,j=1}^N |v_{ij}|^4 \chi_{ij} |< 2r_0 + \frac{1}{\varepsilon N^2} \sum_{i,k=1}^N |v_{ki}|^2 \phi_{ki}.\]

By choosing \(\varepsilon\) small enough the first sum is absorbed into the gain term (45). The second sum is dominated by \(I_2\), and so is the last third sum.

The term \(R_1\) can be estimated in exact same manner. Thus,

\[
\frac{d}{dt} G_3 \leq -c \Phi(D) V_4 + a I_2.
\]

With the help of the functional \(L = G_3 + a V_2\) we conclude that

\[\int_0^\infty \Phi(D(t)) V_4(t) \, dt < \infty,\]

(46)

with the integral bound being independent of \(N\). Next, let us note that due to uniform bound on velocity, \(D(t) \leq ct + D_0\), so

\[\int_0^\infty \Phi(ct + D_0) V_2(t) \, dt < \infty.\]

(47)

Due to non-integrability of \(\Phi\), \(V_4\) cannot stay bounded away from the origin. Hence, due to its monotonicity, \(V_4 \to 0\). \(\Box\)

**Remark 2.11.** In either of the results above we can actually extract a specific rate of alignment if we make an explicit assumption about the tail of the kernel. Thus, if

\[\phi(r) \sim \frac{1}{r^\beta}, \quad \forall r > r_0,\]

and \(\beta \leq 1\), then from (46),

\[\int_1^\infty \frac{1}{t^\beta} V_4(t) \, dt < \infty,\]

where \(V\) is the corresponding functional. So, if \(\beta = 1\), then there exists an \(A > 0\) such that for any \(T > 1\) there exists a \(t \in [T, T^A]\) such that \(V_4(t) < \frac{1}{t^{1-\beta}}\). Since \(\ln t\) is proportional to \(\ln T\) for all \(t \in [T, T^A]\) this proves the above bound for all large times. If \(\beta < 1\), then we argue that for some large \(A > 0\) and all \(T > 0\) we find \(t \in [T, AT]\) such that \(V_4(t) \leq \frac{1}{t^{1-\beta}}\). But the latter is comparable for all values of \(t \in [T, AT]\). Thus we obtain the power rate as above for all \(t\).
Let us summarize the obtained results.

\[ V_4(t) \lesssim \frac{1}{\ln t}, \quad \beta = 1, \]

\[ V_4(t) \lesssim \frac{1}{t^{1-\beta}}, \quad \beta < 1. \]

Of course the same argument applies to \( V_2 \) under the conditions of part (ii).

3. Dynamics under potential forces

3.1. Quadratic confinement. One of the natural ways to force the Cucker-Smale model is to include a potential confinement to the momentum equation:

\[
\begin{cases}
\dot{x}_i = v_i, \\
\dot{v}_i = \frac{1}{N} \sum_{j=1}^{N} \phi(x_i - x_j)(v_j - v_i) - \nabla U(x_i),
\end{cases}
\]

where \( U \) is a convex radially increasing potential. The macroscopic version of this system was analyzed by Shu and Tadmor in recent work [9]. The case is a good illustration just how dramatically the behavior of the system adapts to present forces and how crucially it relies on the particular structure of potential. While in general one can perform analysis in the case of a strictly convex potential, the results are not as conclusive as they rely on additional assumptions on the strength of communication. We focus here on one particular case of quadratic \( U \):

\[ U(x) = \frac{1}{2} |x|^2. \]

In this case, the system reads

\[
\begin{cases}
\dot{x}_i = v_i, \\
\dot{v}_i = \frac{1}{N} \sum_{j=1}^{N} \phi(x_i - x_j)(v_j - v_i) - x_i,
\end{cases}
\]

The natural limiting state of this system is not, in fact, alignment, but rather a harmonic oscillator. Indeed, denoting as before \( \bar{x}, \bar{v} \) the center of mass and momentum, we obtain

\[ \frac{d}{dt} \bar{x} = \bar{v}, \quad \frac{d}{dt} \bar{v} = -\bar{x}. \]

Also due to the linear nature of the forcing we can shift the system of coordinates to \((\bar{x}, \bar{v})\) and assume that \( \bar{x} = 0, \bar{v} = 0 \). So, the question becomes to show that the solution tends to zero in the new reference frame (still denoted \((x, v)\)).

The full energy of the system is given by

\[ \mathcal{E} = \mathcal{K} + \mathcal{P} \]

\[ \mathcal{K} = \frac{1}{2N} \sum_{i=1}^{N} |v_i|^2, \quad \mathcal{P} = \frac{1}{2N} \sum_{i=1}^{N} |x_i|^2. \]

It will also be important to consider the “particle energy”, i.e. \( L^\infty \)-version of the energy:

\[ \mathcal{E}_\infty = \frac{1}{2} \max_i (|v_i|^2 + |x_i|^2). \]

**Theorem 3.1.** Suppose that the kernel \( \phi \) is bounded, decreasing, and satisfies the weak fat tail condition

\[ \int_0^\infty r \phi(r) \, dr = \infty. \]
Then the system (51) settles on its harmonic oscillator (52) exponentially fast, meaning
\[
\max_i (|v_i(t) - \bar{v}(t)|^2 + |x_i(t) - \bar{x}(t)|^2) \leq Ce^{-\delta t},
\]
for some $\delta > 0$ and $C = C(v_0, x_0, \phi)$ independent of $N$.

**Proof.** The result amounts to establishing an exponential bound on $E_\infty$.

Straight from the equations we obtain the energy law
\[
\frac{d}{dt} E = -I^2 \leq -\phi(D) K.
\]
Note that the dissipation is not coercive even if we knew a lower bound $\phi(D) > c_0$. However, we proceed with $E_\infty$:
\[
\frac{d}{dt} E_\infty \leq \frac{1}{N} \sum_j \phi_{ij} (v_j - v_i) \cdot v_i \leq \frac{1}{2N} \sum_j \phi_{ij} (|v_j|^2 - |v_i|^2) \leq \frac{1}{2N} \sum_j \phi_{ij} |v_j|^2 \leq |\phi|_\infty K.
\]
To combine the two equations into one system, let us note that
\[
D \leq 4\sqrt{E_\infty}. 
\]
Thus, \[
\frac{d}{dt} E \leq -\phi(4\sqrt{E_\infty}) K.
\]
Consider the Lyapunov function
\[
\mathcal{L} = E + \frac{1}{C} \int_0^{E_\infty} \phi(4\sqrt{r}) dr.
\]
From our fat tail assumption it follows that $E_\infty$ remains bounded, from which we conclude that $D(t) \leq D_\infty$. Now the energy law (57) reads
\[
\frac{d}{dt} E \leq -c_0 K.
\]
Next we proceed with a hypocoercivity argument to restore $K$ on the right hand side to the full energy $E$. We consider the corrector (longitudinal momentum)
\[
\mathcal{X} = \frac{1}{N} \sum_{i=1}^N x_i \cdot v_i,
\]
and note that for $\lambda < \frac{1}{4}$, $E + \lambda \mathcal{X} \sim E$. Let us compute the derivative:
\[
\frac{d}{dt} \mathcal{X} = \frac{1}{N^2} \sum_{j,i} \phi_{ij} (v_j - v_i) \cdot x_i + \frac{1}{N} \sum_i (|v_i|^2 - |x_i|^2)
\leq \frac{1}{N^2} \sum_{j,i} (4|v_j|^2 + 4|v_i|^2 + \frac{1}{2}|x_i|^2) + \frac{1}{N} \sum_i (|v_i|^2 - |x_i|^2) \leq C\mathcal{K} - \frac{1}{4}\mathcal{P}.
\]
Choosing $\lambda < c_0/2C$ we obtain
\[
\frac{d}{dt} (E + \lambda \mathcal{X}) \leq -c_1 E \sim -E + \lambda \mathcal{X}.
\]
This establishes exponential decay of $E$.

Now that the $L^2$-energy is decaying exponentially, we can extrapolate to obtain exponential decay for $E_\infty$ as well. The method is similar – we consider an amended version of $E_\infty$:
\[
E^\lambda_\infty = \frac{1}{2} \max_i (|v_i|^2 + |x_i|^2 + \lambda x_i \cdot v_i).
\]
Again, for $\lambda < \frac{1}{4}$, this does not alter the particle energy much, $\mathcal{E}_\infty \sim \mathcal{E}_\infty^\lambda$. Differentiating at a point of maximum we obtain

\[
\frac{d}{dt} \mathcal{E}_\infty^\lambda \leq \frac{1}{N} \sum_j \phi_{ij}(v_j - v_i) \cdot (v_i + \lambda x_i) + \lambda |v_i|^2 - \lambda |x_i|^2
\]

\[
\leq \frac{1}{N} \sum_j \phi_{ij} v_j \cdot v_i + \lambda \frac{1}{N} \sum_j \phi_{ij} v_j \cdot x_i - \frac{1}{N} \sum_j \phi_{ij} |v_i|^2 - \lambda \frac{1}{N} \sum_j \phi_{ij} v_i \cdot x_i + \lambda |v_i|^2 - \lambda |x_i|^2.
\]

In the gain term $-\frac{1}{N} \sum_j \phi_{ij} |v_i|^2$ we replace $\phi_{ij}$ by a lower bound $c_0$:

\[
-\frac{1}{N} \sum_j \phi_{ij} |v_i|^2 \leq -c_0 |v_i|^2.
\]

and in the rest we simply use boundedness of the kernel. Hence, if $\lambda$ is small enough, the term $\lambda |v_i|^2$ gets absorbed. We obtain at this point

\[
\frac{d}{dt} \mathcal{E}_\infty^\lambda \lesssim \frac{1}{N} \sum_j |v_j||v_i| + \lambda \frac{1}{N} \sum_j |v_j||x_i| + \lambda \frac{1}{N} \sum_j |v_i||x_i| - c_1 |v_i|^2 - \lambda |x_i|^2
\]

\[
\leq 4K + \frac{c_1}{4} |v_i|^2 + 4\lambda K + \frac{1}{4} \lambda |x_i|^2 + \lambda |v_i|^2 + \frac{1}{4} \lambda |x_i|^2 - c_1 |v_i|^2 - \lambda |x_i|^2
\]

\[
\lesssim K - |v_i|^2 - |x_i|^2 \lesssim K - \mathcal{E}_\infty^\lambda.
\]

Since we already know that $K$ is exponentially decaying, this establishes a similar bound on $\mathcal{E}_\infty^\lambda$. □

Let us note that for small number of agents, when the dependence on $N$ is not an issue, one can actually obtain a much stronger result: as long as $\phi(r) > 0$ for all $r > 0$, the conclusion of the theorem holds true. Indeed, from the first lines we have established that the total energy is decaying and, hence, bounded. But from the potential part $P$ this immediately implies that the flock is bounded, with a bound depending on $N$. This sets the rest of the argument to go through.

This observation clearly shows that it is impossible to construct a simple example with a few agents, like we did in Example 2.3, to prove sharpness of the fat tail condition (55).

3.2. Attraction-Alignment models. We consider the system with pairwise interactions determined by a radially symmetric smooth potential $U \in C^2(\mathbb{R}_+)$:

\[
\begin{cases}
\dot{x}_i = v_i, \\
\dot{v}_i = \frac{1}{N} \sum_{j=1}^N \phi(x_i - x_j)(v_j - v_i) - \frac{1}{N} \sum_{j=1}^N \nabla U(x_i - x_j).
\end{cases}
\]

Note that in general it may not be possible to achieve a natural aligned steady state as it may not even exist for the system with repulsion/attraction forces. Indeed all such solutions will need to land on the minimum of $U$, which is this case is an annulus. In other words all $|x_{ij}| \in [r_0, r_1]$. This is only possible with a limited number of agents, $N$, where $N$ depends on the width of the annulus. This 3Zone model, however, is very important a mathematical realization of the classical Craig Reynolds’ theory of flock formations, and is an underlying tool in 3D computer animations for picture production. We focus here on two less restrictive cases: repulsion-only or attraction-only potential, the so called 2Zone models.
Notice first that the system (62) preserves momentum and is Galilean invariant. So, we can shift
the center of mass and momentum to zero. Let us define the classical energies
\[ \mathcal{E} = \mathcal{K} + \mathcal{P}, \]
(63)
\[ \mathcal{K} = \frac{1}{2N} \sum_{i=1}^{N} |v_i|^2 = \frac{1}{4N^2} \sum_{i=1}^{N} |v_{ij}|^2, \quad \mathcal{P} = \frac{1}{N^2} \sum_{i,j=1}^{N} U(x_{ij}), \]
the forces
\[ F_i = \frac{1}{N} \sum_{j=1}^{N} \nabla U(x_i - x_j), \quad F = (F_1, \ldots, F_N), \quad |F|^2 = \frac{1}{N} \sum_{j=1}^{N} |F_i|^2. \]
The energy satisfies the classical law:
(64)
\[ \frac{d}{dt} \mathcal{E} = -\frac{1}{N^2} \sum_{i,j=1}^{N} \phi_{ij} |v_{ij}|^2 = -\mathcal{I}. \]

In this section we consider an Attraction-Alignment 2Zone model with non-degenerate communication kernel:
(65)
\[ \phi'(r) \leq 0, \quad \phi(r) \geq \frac{c_0}{(r)^\gamma}, \quad \text{for } r \geq 0. \]

For the potential we assume essentially a power law: for some \( \beta > 1 \) and \( L' > L > 0 \),

\[ \begin{align*}
\text{Support:} & \quad U \in C^2(\mathbb{R}^+), \quad U(r) = 0, \quad \forall r \leq L, \\
\text{Growth:} & \quad U(r) \geq a_0 r^\beta, \quad |U'(r)| \leq a_1 r^{\beta-1}, \quad |U''(r)| \leq a_2 r^{\beta-2}, \quad \forall r > L', \\
\text{Convexity:} & \quad U'(r), U''(r) \geq 0, \quad \forall r > 0.
\end{align*} \]

**Theorem 3.2.** Under the assumptions (65) and (66) on the kernel and potential in the range of parameters given by

(67)
\[ \gamma < \begin{cases} 
1, & 1 < \beta < \frac{4}{3}, \\
\frac{3}{2} \beta - 1, & \frac{4}{3} \leq \beta < 2, \\
2, & \beta \geq 2,
\end{cases} \]

all solutions to the system (62) flock
\[ D(t) \leq D < \infty, \]
and align
(68)
\[ \mathcal{E}(t) \leq \frac{C_\delta}{(t)^{1-\delta}}, \quad \forall \delta > 0. \]

**Proof.** We will operate with the particle energy defined similarly to the confinement case:
(69)
\[ \mathcal{E}_i = \frac{1}{2} |v_i|^2 + \frac{1}{N} \sum_{k=1}^{N} U(x_{ik}), \quad \mathcal{E}_\infty = \max_i \mathcal{E}_i. \]

First, let us observe that the particle energy controls the diameter of the flock. By convexity and
our assumptions on the growth of the potential, we have
(70)
\[ \mathcal{E}_i \geq U(x_i) \geq (|x_i| - L')^\beta. \]

So,
(71)
\[ D \leq \mathcal{E}_\infty^{1/\beta} + L'. \]
Let us now establish a bound on $E_\infty$. For each $i$ we compute

$$\frac{d}{dt}E_i = \frac{1}{N} \sum_{k=1}^{N} \phi_{ik} v_{ki} \cdot v_i - \frac{1}{N} \sum_{k=1}^{N} \nabla U(x_{ik}) \cdot v_k.$$  

For the kinetic part we use the identity

$$v_{ki} \cdot v_i = -\frac{1}{2} |v_{ki}|^2 - \frac{1}{2} |v_i|^2 + \frac{1}{2} |v_k|^2.$$  

Discarding all the negative terms, we bound

$$\frac{1}{N} \sum_{k=1}^{N} \phi_{ik} v_{ki} \cdot v_i \leq |\phi|_\infty K.$$  

Due to the energy law, $K$, of course will remain bounded, but we will keep it for now. As to the potential term, there are several ways we can handle it.

For any $1 \leq \beta \leq \frac{4}{3}$ we can derive a direct estimate from the first derivative:

$$\left| \frac{1}{N} \sum_{k=1}^{N} \nabla U(x_{ik}) \cdot v_k \right| \leq \sqrt{K} \left( \frac{1}{N} \sum_{k=1}^{N} \left| \nabla U(x_{ik}) \right|^2 \right)^{\frac{1}{2}} \leq \sqrt{K} D^{\beta - 1}.$$  

Consequently,

$$\frac{d}{dt} \mathcal{E}^{(i)} \leq c_1 K + c_2 \sqrt{K} D^{\beta - 1} \leq \sqrt{K} (1 + \mathcal{E}_\infty^{\beta - 1}),$$  

and

$$\frac{d}{dt} \mathcal{E}_\infty \leq c_3 \sqrt{K} (1 + \mathcal{E}_\infty^{\beta - 1}) \Rightarrow \mathcal{E}_\infty \lesssim (t)^\beta \Rightarrow \mathcal{D} \lesssim \langle t \rangle.$$  

In the range $\frac{4}{3} \leq \beta \leq 2$ it is better to make use of the second derivative:

$$\left| \frac{1}{N} \sum_{k=1}^{N} \nabla U(x_{ik}) \cdot v_k \right| = \frac{1}{N} \sum_{k=1}^{N} (\nabla U(x_{ik}) - \nabla U(x_i)) \cdot v_k \leq \|D^2 U\|_{\infty} \sqrt{K} \left( \frac{1}{N} \sum_{k=1}^{N} |x_k|^2 \right)^{\frac{1}{2}}$$  

$$\leq c_4 \sqrt{K} \left( \frac{1}{N^2} \sum_{i,j=1}^{N} |x_{ij}|^2 \right)^{\frac{1}{2}}.$$  

\[ \text{FIGURE 1. 2Zone Attraction-Alignment model} \]
The following inequality will be used repeatedly

\[
\frac{1}{N^2} \sum_{i,j=1}^{N} |x_{ij}|^2 \leq (L')^2 + \frac{1}{N^2} \sum_{i,j=1}^{N} |(x_{ij} - L')_+^2 \leq C(1 + D^{(2-\beta)}) \mathcal{P}.
\]

Continuing the above,

\[
\left| \frac{1}{N} \sum_{k=1}^{N} \nabla U(x_{ik}) \cdot v_k \right| \leq c_4 \sqrt{K} (1 + D^{2-\beta})^{1/2} \leq c_5 \sqrt{K} (1 + \mathcal{E}_\infty)^{\frac{2-\beta}{2}}.
\]

In this case,

\[
\frac{d}{dt} \mathcal{E}_\infty \leq c_6 \sqrt{K} (1 + \mathcal{E}_\infty)^{\frac{2-\beta}{2}} \Rightarrow \mathcal{E}_\infty \lesssim \langle t \rangle^{\frac{2-\beta}{2-1}} \Rightarrow D \leq \langle t \rangle^{\frac{2}{2-1}}.
\]

Finally, for \( \beta > 2 \), we argue similarly, using that \( |D^2 U(x_{ik})| \leq D^{\beta-2} \), and (76), to obtain

\[
\left| \frac{1}{N} \sum_{k=1}^{N} \nabla U(x_{ik}) \cdot v_k \right| \leq \sqrt{K} D^{\beta-2},
\]

and hence,

\[
\frac{d}{dt} \mathcal{E}_\infty \leq c_7 \sqrt{K} (1 + \mathcal{E}_\infty)^{\frac{\beta-2}{\beta}} \Rightarrow \mathcal{E}_\infty \lesssim \langle t \rangle^{\frac{\beta}{\beta-2}} \Rightarrow D \leq \langle t \rangle^{\frac{1}{\beta-2}}.
\]

We have proved the following a priori estimate:

\[
\mathcal{D}(t) \lesssim \langle t \rangle^d,
\]

where \( d = \begin{cases} 1, & 1 \leq \beta < \frac{4}{3}, \\ \frac{2}{3\beta - 2}, & \frac{4}{3} \leq \beta < 2, \\ \frac{1}{2}, & \beta \geq 2. \end{cases} \)

Denote

\[
\zeta(t) = \langle t \rangle^{-\gamma d}.
\]

According to the basic energy equation (64) we have

\[
\frac{d}{dt} \mathcal{E} \leq -\frac{1}{2} \mathcal{I} - c\zeta(t) \mathcal{K}.
\]

Considering this as a starting point, just like in the quadratic confinement case, we will build correctors to the energy to achieve full coercivity on the right hand side of (80). We introduce one more auxiliary power function

\[
\eta(t) = \langle t \rangle^{-\alpha}, \quad \gamma d \leq \alpha < 1.
\]

First, we consider the same longitudinal momentum

\[
\mathcal{X} = \frac{1}{N} \sum_{i=1}^{N} x_i \cdot v_i.
\]

It will come with a prefactor \( \varepsilon \eta(t) \), where \( \varepsilon \) is a small parameter. Let us estimate using (76):

\[
\varepsilon \eta(t)|\mathcal{X}| \leq \varepsilon \mathcal{K} + \varepsilon \eta^2(t) \frac{1}{N^2} \sum_{i,j=1}^{N} |x_{ij}|^2 \leq \varepsilon \mathcal{K} + \varepsilon \varepsilon \eta^2(t) + \varepsilon \eta^2(t) D^{(2-\beta)} \mathcal{P}.
\]

The potential term is bounded by \( \varepsilon \mathcal{P} \) as long as

\[
2\alpha \geq d(2 - \beta)_+.
\]
Hence,\
\[(81)\] 
\[\varepsilon \eta(t) |\mathcal{X}| \leq \varepsilon \mathcal{E} + c \eta^2(t).\]

This shows that 
\[\mathcal{E} + \varepsilon \eta(t) \mathcal{X} + 2c \eta^2(t) \sim \mathcal{E} + c \varepsilon \eta^2(t).\]

Let us now consider the derivative 
\[\mathcal{X}' = \frac{1}{N^2} \sum_{i=1}^{N} |v_i|^2 + \frac{1}{N^2} \sum_{i,k=1}^{N} x_{ik} \cdot v_{ki} \phi_{ki} - \frac{1}{N^2} \sum_{i,k=1}^{N} x_{ik} \cdot \nabla U(x_{ik}) = \mathcal{K} + A - B.\]

The gain term \(B\), by convexity dominates the potential energy \(B \geq \mathcal{P}\). This is the main reason why we introduced the \(X\)-corrector. As to \(A\):
\[|A| \leq \frac{1}{2\varepsilon^{1/2} \eta(t)} \mathcal{I} + \frac{\varepsilon^{1/2} \eta(t)}{2} \frac{1}{N^2} \sum_{i,j=1}^{N} |x_{ij}|^2 \leq \frac{1}{\varepsilon^{1/2} \eta(t)} \mathcal{I} + \varepsilon^{1/2} \eta(t) + \varepsilon^{1/2} \eta(t) D^{(2-\beta)} + \mathcal{P}.\]

By requiring a more stringent assumption on parameters \[(82)\]
\[\alpha \geq d(2 - \beta)_+,\]
we can ensure that the potential term is bounded by \(\sim \varepsilon^{1/2} \mathcal{P}\), which can be absorbed by the gain term. So far, we have obtained 
\[(83)\]
\[\frac{d}{dt}(\mathcal{E} + \varepsilon \eta(t) \mathcal{X} + 2c \eta^2(t)) \leq -c_1 \varepsilon \eta(t) \mathcal{E} + c_2 \eta^2(t) + \varepsilon \eta'(t) \mathcal{X}.\]

In view of \((81)\), 
\[|\varepsilon \eta'(t) \mathcal{X}| \leq \varepsilon \frac{1}{\langle t \rangle} \eta(t) |\mathcal{X}| \leq \varepsilon \frac{1}{\langle t \rangle} \mathcal{E} + \varepsilon \frac{\eta^2(t)}{\langle t \rangle}.\]

Since \(\alpha < 1\), the energy term will be absorbed, and the free term is even smaller then \(\eta^2\). Denoting 
\[E = \mathcal{E} + \varepsilon \eta(t) \mathcal{X} + 2c \eta^2(t),\]
we obtain 
\[\frac{d}{dt} E \leq -c_1 \eta(t) E + c_2 \eta^2(t).\]

By Duhamel’s formula,
\[E(t) \lesssim \exp\left\{ -\langle t \rangle^{1-\alpha} \right\} + \exp\left\{ -\langle t \rangle^{1-\alpha} \right\} \int_0^t \frac{e(s)^{1-\alpha}}{\langle s \rangle^{2\alpha}} ds.\]

By an elementary asymptotic analysis,
\[\int_0^t \frac{e(s)^{1-\alpha'}}{\langle s \rangle^{\alpha'-\alpha'}} ds \sim \exp\left\{ \langle t \rangle^{1-\alpha'} \right\} \frac{1}{\langle t \rangle^{\alpha'-\alpha'}}.\]

Thus, we obtain an algebraic decay rate 
\[(84)\] 
\[E(t) \lesssim \frac{1}{\langle t \rangle^\alpha}, \quad \forall \alpha < 1,\]
provided 
\[(85)\] 
\[d\gamma < 1 \quad \text{and} \quad d(2 - \beta)_+ < 1.\]

This translates exactly into the conditions on \(\gamma\) given by \((67)\), and \((84)\) automatically implies \((68)\).

Going back to the estimates \((74)\) and \((77)\), but keeping the kinetic energy with its established decay, we obtain a new decay rate for the diameter 
\[D \leq C_\delta(t)^{\frac{d}{\delta} + \delta}, \quad \forall \delta > 0.\]
At the next stage we prove flocking: \( D(t) < D_\infty \). In order to achieve this we return again to the particle energy estimates. Let us denote

\[
P_i = \frac{1}{N} \sum_{k=1}^{N} U(x_{ik}), \quad I_i = \frac{1}{N} \sum_{k=1}^{N} \phi_{ik} |v_{ki}|^2, \quad X_i = x_i \cdot v_i.
\]

Using (72), (73), (75), (76) and the fact that \( D^{(2-\beta)} + P \) has a negative rate of decrease, we obtain

\[
\frac{d}{dt} \mathcal{E}_i \leq K - \frac{1}{2} \phi(D) |v_i|^2 - I_i + c \sqrt{K} \lesssim -\frac{1}{2} \phi(D) |v_i|^2 - I_i + \frac{1}{(t)^{\frac{1}{2}-\delta}}, \quad \forall \delta > 0.
\]

In view of (85), we can pick \( \alpha \) and \( \delta \) such that

\[
(2-\beta) \alpha + 2 \delta (2-\beta) < 2 \alpha.
\]

We use as before the auxiliary rate function \( \eta(t) = \langle t \rangle^{-\alpha} \). Let us estimate the corrector

\[
|\epsilon \eta(t) \mathcal{X}_i| \leq \epsilon |v_i|^2 + \epsilon \eta^2(t) |x_i|^2 \leq \epsilon |v_i|^2 + \epsilon \eta^2(t) D^2 - \beta \mathcal{P}_i + L^2 \epsilon \eta^2(t) \leq \epsilon |v_i|^2 + c \epsilon \mathcal{P}_i + L^2 \epsilon \eta^2(t).
\]

So,

\[
E_i := \mathcal{E}_i + \epsilon \eta(t) \mathcal{X}_i + L^2 \epsilon \eta^2(t) \sim \mathcal{E}_i + L^2 \epsilon \eta^2(t).
\]

Differentiating,

\[
\mathcal{X}_i' = |v_i|^2 + \frac{1}{N} \sum_{k=1}^{N} x_i \cdot v_{ki} \phi_{ki} - \frac{1}{N} \sum_{k=1}^{N} x_i \cdot \nabla U(x_{ik}) + \frac{1}{N} \sum_{k=1}^{N} x_k \cdot (\nabla U(x_{ik}) - \nabla U(x_i))
\]

\[
\leq |v_i|^2 + \epsilon^{1/2} \eta(t) x_i |x_i|^2 + \frac{1}{\epsilon^{1/2} \eta(t)} I_i - \mathcal{P}_i + \frac{1}{N^2} \sum_{i,k=1}^{N} |x_{ik}|^2
\]

\[
\leq |v_i|^2 + \epsilon^{1/2} L^2 \eta(t) + \epsilon^{1/2} D^{(2-\beta)} + \eta(t) \mathcal{P}_i + \frac{1}{\epsilon^{1/2} \eta(t)} I_i - \mathcal{P}_i + C
\]

in view of (86), \( \epsilon^{1/2} D^{(2-\beta)} + \eta(t) \lesssim \epsilon^{1/2} \), so the potential term is absorbed by \(- \mathcal{P}_i\),

\[
\leq |v_i|^2 + \frac{1}{\eta(t)} I_i - \frac{1}{2} \mathcal{P}_i + C.
\]

Again in view of (86), \( \eta(t) \) decays faster than \( \phi(D) \), so plugging into the energy equation we obtain

\[
\frac{d}{dt} E_i \leq - \epsilon \eta(t) E_i + \eta(t) + \sqrt{K} + \epsilon \eta'(t) \mathcal{X}_i,
\]

and as before \( \epsilon \eta'(t) \mathcal{X} \) is a lower order term which is absorbed in the negative energy term and \( + \eta^2 \).

So,

\[
\frac{d}{dt} E_i \leq - \epsilon \eta(t) E_i + \eta(t) + \sqrt{K}.
\]

By our choice of constants (86), \( \sqrt{K} \) decays faster than \( \eta(t) \), hence,

\[
\frac{d}{dt} E_i \lesssim - \epsilon \eta(t) E_i + \eta(t).
\]

This proves boundedness of \( E_i \), and hence that of \( \mathcal{E}_i + L^2 \epsilon \eta^2(t) \), and hence that of \( \mathcal{E}_i \). In view of (71), this implies flocking:

\[
D(t) < \bar{D}, \quad \forall t > 0.
\]

\[\square\]
It is interesting to note that when the support of the potential spans the entire line, \( L = 0 \), and \( U \) lands at the origin with at least a quadratic touch:

\[
U(r) \geq a_0 r^2, \quad r < L',
\]

then we can establish exponential alignment in terms of the energy \( E \). Indeed, since we already know that the diameter is bounded, the basic energy equation reads

\[
\frac{d}{dt} E \leq -c_0 K - \frac{1}{2} I.
\]

The momentum corrector needs only an \( \varepsilon \)-prefactor to satisfy the bound

\[
|\varepsilon X| \leq \varepsilon K + \varepsilon c P.
\]

This is due to the assumed quadratic order of the potential near the origin and, again, boundedness of the diameter. Hence, \( E + \varepsilon X \sim E \). The rest of the argument is similar to the confinement case. We obtain

\[
X \lesssim K + \varepsilon^{1/2} P + \frac{1}{\varepsilon^{1/2}} I - P \leq K - \frac{1}{2} P \frac{1}{\varepsilon^{1/2}} I.
\]

Thus,

\[
\frac{d}{dt}(E + \varepsilon X) \leq -c_1 E \sim -c_1 (E + \varepsilon X).
\]

This proves exponential decay of \( E \). Going further to consider the individual particle energies, we discover similar decays. Indeed, denoting by \( \text{Exp} \) any quantity that decays exponentially fast, we follow the same scheme:

\[
\frac{d}{dt} E_i \leq -c_1 |v_i|^2 - \frac{1}{2} I_i + \text{Exp}.
\]

In view of \(|x_i|^2 \lesssim P_i\),

\[
\varepsilon |X_i| \leq \varepsilon |v_i|^2 + \varepsilon P_i,
\]

so \( E_i + \varepsilon X_i \sim E_i \). Further following the estimates as in the proof,

\[
X_i' \lesssim |v_i|^2 + \frac{1}{\varepsilon^{1/2}} I_i - \frac{1}{2} P_i.
\]

Thus,

\[
\frac{d}{dt} (E_i + \varepsilon X_i) \leq -c_1 (E_i + \varepsilon X_i) + \text{Exp}.
\]

This establishes exponential decay for \( E_i \), and hence for the individual velocities. This also proves that \( D(t) = \text{Exp} \). So, the alignment outcome here is exponential shrinking a point.

**Theorem 3.3.** Let us assume that the support of the potential spans the entire space and (88). Then the solutions flock and align exponentially fast:

\[
D(t) + |v(t) - \bar{v}|_\infty \leq C e^{-\delta t},
\]

for some \( C, \delta > 0 \).

4. **Multi-flocks**

It is clear from formulas (27) and (28) that the rate of alignment of a flock is proportional to the mass \( M \) if that mass is large. At the same time it is inversely proportional to the diameter of the flock (through the upper bound \( \bar{D} \)). This creates a somewhat unrealistic situation when there are two well separated flocks each is massive are described by the same communication \( \phi \) throughout the environment. In this case the fast alignment which supposed to happen within each flock gets hijacked by the other flock due to long separation. In this situation is better modeled by a system with multi-scaling that takes into account different time scales on which alignment is achieved...
within and between the flocks in a bigger cluster. We can formulate an even more general multi-
scale model where communication between flocks is regulated by a kernel different from internal 
one.

To derive such a model, we assume that positions and velocities \( y_{\alpha i} = y_{\alpha i}(t, \tau), \) \( u_{\alpha i} = u_{\alpha i}(t, \tau), \) where \( \alpha = 1, \ldots, A \) and \( i = 1, \ldots, N_{\alpha}, \) depend on two time parameters, in which \( t \) is a fast time and \( \tau \) is a slow time. We postulate that on the fast scale the \( \alpha \)-flock does not feel the influence of other flocks, and evolves autonomously according to the classical Cucker-Smale system with internal communication law:

\[
\begin{aligned}
\partial_t y_{\alpha i} &= u_{\alpha i}, \\
\partial_t u_{\alpha i} &= \lambda_\alpha \sum_{j=1}^{N_\alpha} m_{\alpha j} \phi_\alpha(y_{\alpha j}, y_{\alpha i})(u_{\alpha j} - u_{\alpha i}).
\end{aligned}
\]  

(89)

On the much slower time scale \( \tau \) the agents of the \( \alpha \)-flock are influenced by other flocks only via their macroscopic parameters, in other words \( \alpha \)-agents adjust to other flocks according to their consensus direction:

\[
\begin{aligned}
Y_\alpha &= \frac{1}{M_\alpha} \sum_{i=1}^{N_\alpha} m_{\alpha i} y_{\alpha i}, \quad U_\alpha = \frac{1}{M_\alpha} \sum_{i=1}^{N_\alpha} m_{\alpha i} u_{\alpha i}, \quad M_\alpha = \sum_{i=1}^{N_\alpha} m_{\alpha i}.
\end{aligned}
\]

Thus,

\[
\begin{aligned}
\partial_\tau y_{\alpha i} &= U_\alpha, \\
\partial_\tau u_{\alpha i} &= \sum_{\beta \neq \alpha} M_\beta \Psi(Y_\alpha, Y_\beta)(U_\beta - u_{\alpha i}),
\end{aligned}
\]  

(90)

where \( \Psi \) is an inter-flock communication kernel. Let us notice that the flock momenta are not moving in the fast scale \( \partial_t U_\alpha = 0 \) as follows from (89). However, in slow scale the set of macroscopic parameters \( (Y_\alpha, U_\alpha) \) satisfies the up-scaled Cucker-Smale system:

\[
\begin{aligned}
\partial_\tau Y_\alpha &= U_\alpha, \\
\partial_\tau U_\alpha &= \sum_{\beta \neq \alpha} M_\beta \Psi(Y_\alpha, Y_\beta)(U_\beta - U_\alpha).
\end{aligned}
\]  

(91)

We next introduce an adimensional ratio of the two time scales,

\[
\varepsilon := \frac{\tau}{t} << \min_\alpha \lambda_\alpha.
\]  

(92)

This assumption allows us to consider pair of parameters for each agent depending only on one time dimension:

\[
\begin{aligned}
x_{\alpha i}(t) &= y_{\alpha i}(t, \varepsilon t), \\
v_{\alpha i}(t) &= u_{\alpha i}(t, \varepsilon t) + \varepsilon U_\alpha(\varepsilon t).
\end{aligned}
\]  

(93)

To derive the system for new variables, notice that \( \dot{x}_{\alpha i} = v_{\alpha i}, \) and

\[
\begin{aligned}
\dot{v}_{\alpha i} &= \partial_t u_{\alpha i}(t, \varepsilon t) + \varepsilon \partial_\tau u_{\alpha i}(t, \varepsilon t) + \varepsilon^2 \partial_\tau U_\alpha(\varepsilon t) \\
&= \lambda_\alpha \sum_{j=1}^{N_\alpha} m_{\alpha j} \phi_\alpha(x_{\alpha i}, x_{\alpha j})(u_{\alpha j} - u_{\alpha i}) + \varepsilon \sum_{\beta \neq \alpha} M_\beta \Psi(X_\alpha, X_\beta)(U_\beta - u_{\alpha i}) \\
&+ \varepsilon^2 \sum_{\beta \neq \alpha} M_\beta \Psi(X_\alpha, X_\beta)(U_\beta - U_\alpha).
\end{aligned}
\]
Noting that \(u_{\alpha j} - u_{\alpha i} = v_{\alpha j} - v_{\alpha i}\) and \(V_\alpha = (1 + \varepsilon)U_\alpha\) we arrive at the following system:

\[
\begin{align*}
\dot{x}_{\alpha i} &= v_{\alpha i}, \\
\dot{v}_{\alpha i} &= \lambda_\alpha \sum_{j=1}^{N_\alpha} m_{\alpha j} \phi_{\alpha}(x_{\alpha i} - x_{\alpha j})(v_{\alpha j} - v_{\alpha i}) + \varepsilon \sum_{\beta=1}^{A} M_\beta \Psi(X_\alpha - X_\beta)(V_\beta - V_\alpha),
\end{align*}
\]

where

\[
X_\alpha = \frac{1}{M_\alpha} \sum_{i \in \Omega_\alpha} m_{\alpha i} x_{\alpha i}, \quad V_\alpha = \frac{1}{M_\alpha} \sum_{i \in \Omega_\alpha} m_{\alpha i} v_{\alpha i}.
\]

Note that the macroscopic variables \(X_\alpha, V_\alpha\) satisfies the upscaled system

\[
\begin{align*}
\dot{X}_\alpha &= V_\alpha, \\
\dot{V}_\alpha &= \varepsilon \sum_{\beta \neq \alpha} M_\beta \Psi(X_\alpha - X_\beta)(V_\beta - V_\alpha).
\end{align*}
\]

which nothing by the classical Cucker-Smale system.

Let us consider the following size metrics of the system (94) and (95):

\[
D_\alpha = \max_{i,j} |x_{\alpha i} - x_{\alpha j}|, \quad \mathcal{D} = \max_{\alpha,\beta} |X_\alpha - X_\beta|,
\]

\[
A_\alpha = \max_{i,j} |v_{\alpha i} - v_{\alpha j}|, \quad \mathcal{A} = \max_{\alpha,\beta} |V_\alpha - V_\beta|.
\]

The alignment of macroscopic quantities follows from the same system of ODEs as we derived in the classical case:

\[
\begin{align*}
\dot{V} &\leq -\varepsilon M \Psi(D) V \\
\dot{D} &\leq V.
\end{align*}
\]

Thus, under fat tail condition on \(\Psi\) the consensus directions \(V_\alpha\) will in fast agree exponentially fast according to Theorem 2.4. To understand the alignment within each flock we need to consider how the cluster dynamics influences internal dynamics inside an \(\alpha\)-flock. First, we observe that any multi-flock system (94) satisfies the global maximum principle – maximum of each coordinate in the total family \(v_{\alpha i}\) is non-increasing, and the minimum is non-decreasing. However, this may not be the case within each individual flock. It is true, however if we consider maxima and mimima relative to the momenta of the flocks. To see this let us pass to the reference frame evolving with the momentum and center of mass of the flock:

\[
w_{\alpha i} = v_{\alpha i} - V_\alpha, \quad y_{\alpha i} = x_{\alpha i} - X_\alpha.
\]

Using (94) and (95) one readily obtains the system

\[
\begin{align*}
\dot{y}_{\alpha i} &= w_{\alpha i}, \\
\dot{w}_{\alpha i} &= \lambda_\alpha \sum_{j=1}^{N_\alpha} m_{\alpha j} \phi_{\alpha ij}(w_{\alpha i} - w_{\alpha j}) - \varepsilon R_\alpha(t)w_{\alpha i},
\end{align*}
\]

where

\[
R_\alpha(t) = \sum_{\beta \neq \alpha} M_\beta \Psi(X_\alpha - X_\beta).
\]

We used a shortcut to denote \(\phi_{\alpha ij} = \phi_{\alpha}(y_{\alpha i} - y_{\alpha j}).\) Clearly, each flock now shrinks towards its momentum.
In order to study flocking behavior within each individual flock we derive an analogue of (96) for $\alpha$-metrics in a way similar to the uni-flock model (19). To this purpose let us observe that

$$V_\alpha = \max_{\ell \in \mathbb{R}^n|\ell| = 1, i, j = 1, ..., N_\alpha} \langle \ell, w_{\alpha i} - w_{\alpha j} \rangle.$$ 

Note that the maximum is always taken over a fixed set not changing in time. Thus, by Rademacher’s lemma, we can evaluate the derivative of $V_\alpha$ by considering $\ell, i, j$ at which that maximum is achieved at any instance of time:

$$\frac{d}{dt} V_\alpha = \langle \ell, \dot{w}_{\alpha i} - \dot{w}_{\alpha j} \rangle = \lambda_\alpha \sum_{k=1}^{N_\alpha} m_{ak} \phi_{\alpha ik} \langle \ell, w_{ak} - w_{\alpha i} \rangle - \lambda_\alpha \sum_{k=1}^{N_\alpha} m_{ak} \phi_{\alpha jk} \langle \ell, w_{ak} - w_{\alpha j} \rangle$$

$$- \varepsilon R_\alpha(t) \langle \ell, w_{\alpha i} - w_{\alpha j} \rangle$$

$$= \lambda_\alpha \sum_{k=1}^{N_\alpha} m_{ak} \phi_{\alpha ik} (\langle \ell, w_{ak} - w_{\alpha j} \rangle - \langle \ell, w_{\alpha i} - w_{\alpha j} \rangle)$$

$$+ \lambda_\alpha \sum_{k=1}^{N_\alpha} m_{ak} \phi_{\alpha jk} (\langle \ell, w_{\alpha i} - w_{ak} \rangle - \langle \ell, w_{\alpha i} - w_{\alpha j} \rangle) - \varepsilon R_\alpha(t) V_\alpha.$$ 

Each difference of the action of $\ell$ is negative due to maximality of $\ell, i, j$. Hence, we replace values of $\phi$ by the minimum, i.e. value at $D_\alpha$:

$$\frac{d}{dt} V_\alpha \leq \lambda_\alpha \phi(D_\alpha) \sum_{k=1}^{N_\alpha} m_{ak} (\langle \ell, w_{ak} - w_{\alpha j} \rangle - \langle \ell, w_{\alpha i} - w_{\alpha j} \rangle + \langle \ell, w_{\alpha i} - w_{ak} \rangle - \langle \ell, w_{\alpha i} - w_{\alpha j} \rangle)$$

$$- \varepsilon R_\alpha(t) V_\alpha = -\lambda_\alpha M_\alpha \phi(D_\alpha) V_\alpha - \varepsilon R_\alpha(t) V_\alpha.$$ 

Noting that $R_\alpha(t) \geq M \Psi(D)$, and combining with (96) we arrive at the following system of ODEs:

$$\begin{align*}
\dot{V}_\alpha &\leq -\lambda_\alpha M_\alpha \phi(D_\alpha) V_\alpha - \varepsilon M \Psi(D) V_\alpha \\
\dot{D}_\alpha &\leq V_\alpha \\
\dot{V} &\leq -\varepsilon M \Psi(D) V \\
\dot{D} &\leq V
\end{align*}$$

(100)

Ignoring the term $-\varepsilon M \Psi(D) V_\alpha$ in the $V_\alpha$ equation we can see that the $\alpha$-flock completely decouples from the rest of the cluster, and one can state a fast internal alignment result by applying Theorem 2.4.

**Theorem 4.1** (Fast local flocking). *Assuming that for a given $\alpha \in \{1, \ldots, A\}$ the kernel $\phi_\alpha$ has a fat tail, the $\alpha$-flock aligns to its momentum exponentially fast at a rate functionally dependent on $\phi_\alpha$, initial data, and $\lambda_\alpha$:

$$\max_i |v_{\alpha i}(t) - V_\alpha(t)| \lesssim e^{-\delta t}.$$ 

It is interesting to note that this alignment process is completely independent from the inter-flock communication. So, long range internal communication leads to local emergence despite potentially destabilizing influence of the outside crowd. On the other hand, if the inter-flock communication $\Psi$ is global, e.g. satisfies the fat tail condition (7), then the global alignment occurs even if internal communications are weak or completely absent. This is evident from (100) where we ignore the term $-\lambda_\alpha M_\alpha \phi(D_\alpha) V_\alpha$ and obtain boundedness of $D$ from the last two equations. Alignment rate in this case is global but slow.
**Theorem 4.2** (Slow global flocking). Suppose $\Psi$ has a fat tail, and $\phi_{\alpha} \geq 0$. Then all solutions to (2) align exponentially fast at a rate functionally dependent on $\Psi$, initial data, and $\epsilon$,

$$\max_{\alpha,i} |v_{\alpha i}(t) - V| \lesssim e^{-\delta t}.$$  

Asymptotic dependence of the implied alignment rates for small $\epsilon$ and large $\lambda_{\alpha}$ for the Cucker-Smale kernel can be readily obtained from formulas (27) and (28). Thus, in the context of fast local alignment we obtain $\delta \sim \lambda_{\alpha}$ for all $\gamma \leq 1$, while in the context of slow alignment, $\delta \sim \epsilon^{1/\gamma}$, for $\gamma < 1$, and $\delta \sim \epsilon^{1}e^{-1/\epsilon}$ for $\gamma = 1$.

5. **Kinetic models**

Kinetic description was derived in [5] via mean-field limit $N \to \infty$:

$$\partial_t f + v \cdot \nabla_x f + \lambda \nabla_v \cdot [fF(f)] = 0,$$

where

$$F(f)(x,v,t) = \int_{\mathbb{R}^{2n}} \phi(x,y)(w-v)f(y,w,t)dwdy.$$  

6. **Hydrodynamic models**

The hydrodynamic model can be derived indirectly from (101):

$$\begin{cases}
\rho_t + \nabla \cdot (\rho u) = 0, \\
u_t + u \cdot \nabla u = \int_{\Omega^n} \phi(x,y)(u(y) - u(x))\rho(y)dy \quad (x,t) \in \Omega^n \times \mathbb{R}_+.
\end{cases}$$

It is sometimes advantageous to write the velocity equation in the conservative form

$$\begin{cases}
\rho u_t + \nabla \cdot (\rho u \otimes u) = \int_{\Omega^n} \phi(x,y)(u(y) - u(x))\rho(y)\rho(x)dy.
\end{cases}$$

The crucial feature of the alignment term in (102) is its commutator representation, which is given by

$$C_{\phi}(u, \rho) = \mathcal{L}_\phi(\rho u) - \mathcal{L}_\phi(\rho)u,$$

where $\mathcal{L}$ can have two different forms

$$\mathcal{L}_\phi(f)(x) = \int_{\Omega^n} \phi(x,y)f(y)dy,$$

more suitable for smooth kernels, and

$$\mathcal{L}_\phi(f)(x) = \int_{\Omega^n} \phi(x,y)(f(y) - f(x))dy,$$

suitable for singular kernels. Note that for smooth kernels of convolution type we have $\mathcal{L}_\phi f = \phi * f$.

Just as in the discrete case, the system (102) conserves mass and momentum. So, we expect the alignment to converge to the average velocity

$$\bar{u} = \frac{1}{M} \int_{\Omega^n} u\rho dx, \quad M = \int_{\Omega^n} \rho(x, t) dx.$$  

For convolution type kernels, $\phi(x,y) = \phi(x-y)$, the system is Galilean invariant, so in this case we can always assume that $\bar{u} = 0$. The kinetic energy is given by

$$\mathcal{E} = \frac{1}{2} \int_{\Omega^n} \rho |u|^2 dx.$$
If \( \bar{u} = 0 \), the energy becomes a measure of alignment due to its relation to the \( L^2 \)-variation given by
\[
\mathcal{E} = \frac{1}{4M} \int_{\Omega^n} |u(x) - u(y)|^2 \rho(x) \rho(y) \, dy \, dx.
\]
However, as in the discrete case we mostly work with the variation functions to make the computations not dependent on Galilean invariance of the models. So, let us introduce the variations:
\begin{equation}
(107) \quad \mathcal{V}_p(t) = \frac{1}{p} \int_{\Omega^n \times \Omega^n} |u(x,t) - u(y,t)|^p \rho(x,t) \rho(y,t) \, dz \, dy.
\end{equation}
The following energy equation is obtained directly by testing the momentum equation:
\begin{equation}
(108) \quad \frac{d}{dt} \mathcal{V}_2 = -MI_2, \quad I_2 = \int_{\Omega^n \times \Omega^n} \phi(x,y) |u(x) - u(y)|^2 \rho(x) \rho(y) \, dy.
\end{equation}

Another fundamental property of the system is the maximum principle for each scalar velocity component \( \ell(u), \ell \in (\mathbb{R}^n)^* \), provided the maxima are achieved, which is course true on periodic domain or if \( \bar{u} = 0 \) and the solution decays at infinity on \( \mathbb{R}^n \).

Since all the macroscopic quantities appear to be measured per mass it is natural to introduce the density measure
\[
dm_t = \rho(x,t) \, dx.
\]
In view of the transport nature of the continuity equation, this measure is transported along the flow of \( u \). Namely, if
\[
\frac{d}{dt}X(\alpha, t, t_0) = u(X(\alpha, t, t_0), t), \quad t > t_0
\]
\[
X(\alpha, t_0, t_0) = \alpha,
\]
then \( dm_t \) is a push-forward of \( dm_{t_0} \) under \( X(\cdot, t, t_0) \):
\begin{equation}
(109) \quad dm_t = X(\cdot, t, t_0)\# dm_{t_0}.
\end{equation}
In other words, for any \( f \),
\begin{equation}
(110) \quad \int_{\mathbb{R}^n} f(X(\alpha, t, t_0)) \, dm_{t_0}(\alpha) = \int_{\mathbb{R}^n} f(x) \, dm_t(x).
\end{equation}
We also denote \( X(\cdot, t, 0) = x(\cdot, t) \). In particular, the density support is transported by the flow:
\[
\text{Supp } \rho(t) = X(t, \text{Supp } \rho_0).
\]
Denoting the Lagrangian velocity by
\begin{equation}
(111) \quad v(\alpha, t) = u(X(\alpha, t), t)
\end{equation}
for short, and denoting
\[v_{\alpha \beta} = v(\alpha, t) - v(\beta, t), \quad \phi_{\alpha \beta} = \phi(X(\alpha, t) - X(\beta, t)),\]
we can rewrite the velocity equation as
\begin{equation}
(112) \quad \frac{d}{dt}v(\alpha, t) = \int_{\Omega^n} \phi_{\alpha \beta} [v(\beta, t) - v(\alpha, t)] \, dm_0(\beta).
\end{equation}
The variation becomes
\[
\mathcal{V}_p(t) = \frac{1}{p} \int_{\Omega^n \times \Omega^n} |v_{\alpha \beta}|^p \, dm_0(\alpha, \beta),
\]
where \( dm_0(\alpha, \beta) = dm_0(\alpha) \times dm_0(\beta) \). Thus, in Lagrangian coordinates some computations become very similar to the discrete case. For example, it is straightforward to see that all \( \mathcal{V}_p \)'s are decaying in time.
6.1. Flocking in hydrodynamic context. The proof of Theorem 2.4 is entirely similar in the hydrodynamic context, where we consider corresponding $L^\infty$-based parameters of the solution

$$
D(t) = \max_{\alpha,\beta \in \text{Supp} \rho_0} |x(\alpha,t) - x(\beta,t)|, \quad A(t) = \max_{\alpha,\beta \in \text{Supp} \rho_0} |v(\alpha,t) - v(\beta,t)|.
$$

Note that $A$ measure amplitude of velocity only over the material domain, not the entire space. Since the domain is fixed for all time, one can apply the Rademacher Lemma 2.5. The proof of the following theorem exactly follows the lines of Theorem 2.4 and will be omitted.

**Theorem 6.1.** The system (102) aligns and flocks exponentially fast provided $\phi$ is a non-increasing, everywhere positive, and satisfying the large tail condition (29):

$$
\sup_{t \geq 0} D(t) = \bar{D}, \quad A(t) \leq A_0 e^{-\lambda M D(t)}.
$$

Theorem 2.6 allows for a similar hydrodynamic formulation, and can be stated ad verbatim. This also applies to the degenerate case of Theorem 2.10(ii), which was proved independent of the number of agents. One makes use of macroscopic quantities throughout:

$$
d_{\alpha\beta} = -x_{\alpha\beta} \cdot \frac{v_{\alpha\beta}}{|v_{\alpha\beta}|}, \quad \mathcal{G}_\beta = \int_{\mathbb{R}^n} |v_{\alpha\beta}|^3 \psi(|x_{\alpha\beta}|) \chi(|x_{\alpha\beta}|) \, dm_0(\alpha, \beta), \quad \text{etc.}
$$

Note that all these results give information about a given global solution. They do not address the issue of well-posedness which will be discussed in later sections.

6.2. Spectral method. Hydrodynamic connectivity. When communication is strictly local, the role of connectivity becomes more prominent and the analysis of flocking behavior is more amenable to the periodic settings. Indeed, like in the discrete case one can easily arrange to flocks heading in opposite directions which will never align. In hydrodynamic context connectivity is encoded in no-vacuum condition: $\rho > 0$. The lower bound on the density quantifies connectivity in a way similar to the weighted Fiedler number. In this section we make this quantification more precise and apply the spectral method to develop a conditional alignment criterion for singular local kernels of rather general nature: $\phi(x,y) = \phi(y,x)$ and satisfying

$$
\lambda \frac{\mathbf{1}_{|x-y|<r_0}}{|x-y|^{n+\alpha}} \leq \phi(x,y,t) \leq \frac{\Lambda(t)}{|x-y|^{n+\alpha}}, \quad 0 < \alpha < 2,
$$

here $\Lambda(t)$ is simply assumed finite for any $t \geq 0$. Since the compactness of embedding $H^s \hookrightarrow L^2$ is crucial to this discussion, as well as a uniform lower bound on the density, which is consistent with finite ness of mass only on a compact environment $\Omega^n$, we restrict ourselves to the periodic domain $\Omega^n = \mathbb{T}^n$.

A quantitative expression of coercivity of the commutator (104) relies on lower bounds on the density. Precisely how large that lower should be in order to ensure alignment is investigated in the following proposition.

**Proposition 6.2.** Let $\phi$ be a symmetric, local, singular kernel satisfying (114) and let $(\rho, u)$ be a global strong solution to (102). Assume that

$$
C \geq \rho(t,x) \geq \frac{c}{\sqrt{1+t}}, \quad C > c > 0.
$$

Then the solution aligns at an algebraic rate. Namely, there exist $\eta > 0$ such that

$$
\int_{\mathbb{T}^d} |u(t,x) - \bar{u}|^2 \rho(t,x) \, dx \leq \frac{1}{2M \eta^n}.
$$

Proof. We consider the family of eigenvalue problems parameterized by time:

$$
\int_{\mathbb{T}^n} \phi(x,y)(u(x) - u(y)) \, dm_t(y) = \kappa(t) u(x), \quad u \in H^2.
$$
We seek the minimal eigenvalue, which is of course 0 corresponding to the constant eigenfunction. To remove this trivial solution we restrict it to the time-dependent 1-codimensional subspace

\begin{equation}
H_0^{\frac{2}{3}} = \left\{ u \in H_0^{\frac{2}{3}} : \int_{\mathbb{T}^n} u \, dm_t = 0 \right\}.
\end{equation}

The key issue here is that $H_0^{\frac{2}{3}}$ depends on time. We will return to this later. So, we seek the second minimal eigenvalue of (117) restricted to $H_0^{\frac{2}{3}}$, as a solution to the variational problem

\begin{equation}
\kappa_2(t) = 2M \inf_{u \in H_0^{\frac{2}{3}}} \frac{\int_{T^{2n}} \phi(x,y)|u(y) - u(x)|^2 \rho(t,y)\rho(t,x) \, dx \, dy}{\int_{T^{2n}} |u(x) - u(y)|^2 \rho(t,x)\rho(t,y) \, dx \, dy}.
\end{equation}

In view of (114), and the assumed bounds on the density (115), the upper norm is equivalent for the $H^{\alpha/2}$, and the lower to $L^2$, so the existence follows classically by compactness. The number $\kappa_2(t)$ bears complete resemblance with the discrete Fiedler number, discussed in Section 2.2. In terms of this Fiedler number the energy equation (108) take form

\begin{equation}
\frac{d}{dt} \mathcal{V}_2 \leq -\kappa_2(t) \mathcal{V}_2.
\end{equation}

Consequently,

\begin{equation}
\mathcal{V}_2(t) \leq \mathcal{V}_2(0) \exp \left\{ -\int_0^t \kappa_2(s) \, ds \right\}.
\end{equation}

We will derive now the lower bound $\kappa_2(t) \geq c/(1 + t)$ which clearly implies the statement of the proposition. Using the bounds on the density (115), the mean-zero condition on $u$, and the lower bound of the kernel (114) we obtain

\begin{align*}
\int_{T^{2n}} |u(x) - u(y)|^2 \rho(t,x)\rho(t,y) \, dx \, dy &= 2M \int_{T^n} |u(x)|^2 \rho(t,x) \, dx \leq C\|u\|_2^2, \\
\int_{T^{2n}} \phi(x,y)|u(y) - u(x)|^2 \rho(t,y)\rho(t,x) \, dx \, dy &\geq \frac{c}{t} \int_{|x-y|<r_0} \frac{|u(x) - u(y)|^2}{|x - y|^{n+\alpha}} \, dx \, dy.
\end{align*}

\begin{equation}
\kappa_2(t) \geq \frac{c}{t} \inf_{u \in H_0^{\frac{2}{3}}} \frac{\int_{|x-y|<r_0} \frac{|u(x) - u(y)|^2}{|x - y|^{n+\alpha}} \, dx \, dy}{\|u\|_2^2}.
\end{equation}

Technically, the infimum still depends on time since the mean-zero condition is. So, the last piece to show is that this infimum stays bounded away from zero. We argue by contradiction. Suppose there is a sequence of times $t_k > 0$, and $u_k \in H^{\alpha/2}$ with $\int u_k \, dm_{t_k} = 0$ such that $\|u_k\|_2 = 1$ and

\begin{equation}
\int_{|x-y|<r_0} \frac{|u_k(x) - u_k(y)|^2}{|x - y|^{n+\alpha}} \, dx \, dy \to 0.
\end{equation}

The latter, in particular, implies compactness of the sequence $\{u_k\}_k$ in $L^2$. Hence, up to a subsequence, $u_k \to u$ strongly in $L^2$ and weakly in $H^{\alpha/2}$. By the lower-weak-semi-continuity, and (123), we conclude that $\|u\|_{H^{\alpha/2}} = 0$, and hence $u$ is a constant field, with $|u| = 1$ due to $\|u_k\|_2 \to \|u\|_2$.

At the same time, since $\int \rho(t_k,x) \, dx = M$, there exists a weak* limit of a further subsequence $m_{t_k} \to m$, where $m$ is a positive Radon measure on $\mathbb{T}^d$ with non-trivial total mass $m(\mathbb{T}^d) = M$. We now reach a contradiction if we prove the limit

\begin{equation}
0 = \int_{\mathbb{T}^n} u_k(x)\rho(t_k,x) \, dx \to M u \neq 0.
\end{equation}
To prove the claimed limit note that the assumed uniform upper-bound of the density implies
\[
\int_{\mathbb{T}^d} u_k(x) \rho(t_k, x) \, dx - M u = \int_{\mathbb{T}^d} (u_k(x) - u) \rho(t_k, x) \, dx,
\]
and the latter is clearly bounded by \(C \|u_k - u\|_2 \to 0\).

We conclude that \(\kappa_2(t) \geq c/t\), and the result follows. \(\square\)

In the course of the proof we essentially established a statement analogous to Lemma 2.1 in the discrete case. Let us state it separately.

**Lemma 6.3.** Let \(\kappa_2(t)\) be the weighted Fiedler number defined by (119), and suppose that \(\int_0^\infty \kappa_2(s) \, ds = \infty\).

Then the solution aligns: \(V_2 \to 0\).

We can see now that unconditional flocking is generally achieved under the lower bound on the density, \(\rho(t, \cdot) \geq (1 + t)^{-1/2}\). The difficulty is that this lower bound is too restrictive and is not given a priori for any strong solution. Situation improves considerably for the topological models which yield unconditional flocking under more accessible assumption on the density. We discuss those next.

6.3. **Topological models. Adaptive diffusion.** Technically, topological models belong to a more general class of systems with kernels depending on all agent positions \(x = (x_1, \ldots, x_N)\). So, we start with the general formulation

\[
\begin{align*}
\dot{x}_i &= v_i, \\
\dot{v}_i &= \lambda \sum_{j:|x_i-x_j|<r_0} m_j \phi_{ij}(x)(v_j - v_i).
\end{align*}
\]

Here, we interpret the mass \(m_j\) as a parameter that quantifies power of the agent \(x_j\) to influence others. Thus, the bigger \(m_j\) is the more direct influence \(x_j\) exerts on others. At the same time, it is natural to assume that the “massive” agents are more resistant to the influence by other agents. This latter property will be encoded into the adjacency matrix \(\{\phi_{ij}(x)\}_{i,j=1}^N\) which reflects a chosen communication protocol. Our particular choice of \(\phi_{ij}\) will involve both geometric and topological features, where agent \(x_i\) senses proximity to others not only according to the Euclidean distance but also by the power of their influence. The model we have in mind is based on the following three principles:

1. Every agent \(x_i\) has a finite influence range, which is a Euclidean ball of radius \(r_0\) centered at \(x_i\), denoted \(B(x_i, r_0)\).
2. Agent \(x_i\) exerts influence on another agent \(x_j\) via a transfer of information through a symmetric communication domain between the two agents, denoted \(\Omega(x_i, x_j)\), where \(x_i, x_j \in \partial\Omega(x_i, x_j)\).
3. The quantity

\[
d_{ij} = \left[ \sum_{k: x_k \in \Omega_{ij}} m_k \right]^{1/3}
\]

is a measure of collective power of the intermediaries. We assume that the strength of communication is inversely proportional to \(d_{ij}\).

Based on the outlined principles, we make the following choice:

\[
\phi_{ij}(x) = \frac{1}{d_{ij}} \psi(|x_i - x_j|),
\]

(125)
where $\psi$ is a non-negative function supported in the ball of radius $r_0$, and $\tau > 0$ is a parameter. The kernel $\psi$ encapsulates the metric component of the kernel and is limited to communication cutoff scale $r_0$, while $\tau$ gauges presence of topological effects.

The choice of the domain $\Omega(x, y)$ can be rather flexible, as long as it satisfies two basic requirements (refer to Figure 2): a) the region is a subset of the ball determined by $[x, y]$ as its diameter chord, and two cones of opening $< \pi$ at vertices $x$ and $y$, b) $\Omega(x, y) = \Omega(y, x)$, and the boundary of the region is smooth everywhere except for $x, y$. For simplicity we also assume that topological communication is homogeneous and isotropic, i.e. $\Omega(x, y)$ is constructed by a shift, rotation, and dilation of a basic domain $\Omega(-e_1, e_1)$. In order to insure symmetry b) we assume that $\Omega(-e_1, e_1) = -\Omega(-e_1, e_1)$.

Note that in the macroscopic limit $N \to \infty$, the natural interpretation of the collective powers $d_{ij}$ is given by

$$d(x, y) = \left[ \int_{\Omega(x, y)} \rho(t, z) \, dz \right]^{\frac{1}{n}}.$$  

Clearly, $d(x, y)$ scales like a distance between $x$ and $y$, that is why we call it a “topological distance”. In fact in 1D, where

$$d(x, y) = \left| \int_x^y \rho(t, z) \, dz \right|,$$

this does define a proper metric. As to the choice of metric component, it will play a more important role when we discuss regularity of the models, but for now it suffices to assume that

$$\psi(r) \geq \lambda I_{r < r_0}.$$  

The corresponding hydrodynamic system is given by

$$\begin{cases} 
\rho_t + \nabla \cdot (\rho u) = 0, \\
u_t + u \cdot \nabla x u = \int_{\Omega} \phi(x, y) (u(y) - u(x)) \rho(y) \, dy, \\
\phi(x, y) = \frac{\psi(|x - y|)}{d^r(x, y)}. 
\end{cases}$$  

We note that a proper care has to be given in order to properly define the singular integral operator $L_\phi f$ and the commutator $C_\phi$ is this case. These issues are discussed in [10].
**Theorem 6.4.** Let \((\rho, u)\) be a global smooth solution of the topological model (127) on the torus \(\mathbb{T}^n\), with \(\tau \geq n\). Assume that the density satisfies the lower bound

\[ \rho(t, x) \geq \frac{c}{1 + t}, \quad \forall t > 0, x \in \mathbb{T}^n. \]

Then the solution aligns with a logarithmic rate given by

\[ |u(t) - \bar{u}|_{\infty} \leq \frac{c}{\sqrt{\ln t}}. \]

**Remark 6.5.** The assumption on the density (128) holds automatically in 1D case, as we will see in Section 8.

**Proof.** Let us fix a coordinate \(i\) and aim to prove (129) for \(u_i\). We denote \(u = u_i\) for notational simplicity. Using the Galilean invariance we can lift \(u\) if necessary and assume that \(u(t) > 0\). Note that the extrema of \(u(t)\), denoted \(u_+(t)\) and \(u_-(t)\), are monotonically decreasing and increasing, respectively.

**Step 1: Alignment near extremes.**

Let \(x_+(t)\) be a point of maximum for \(u(t, \cdot)\) and \(x_-(t)\) a point of minimum. Let us fix a time-dependent \(\delta(t) > 0\) to be specified later, and consider the sets

\[ G_\delta^+(t) = \{ u < u_+(t)(1 - \delta(t)) \}, \quad G_\delta^-(t) = \{ u > u_-(t)(1 + \delta(t)) \}. \]

The effect of flattening is expressed in terms of conditional expectations of the above sets in the balls \(B(x_{\pm}(t), r_0)\) with respect to the mass measure. Let us denote

\[ \mathbb{E}_t[A|B] = \frac{m_t(A \cap B)}{m_t(B)}. \]

We show that

\[ \int_0^\infty \mathbb{E}_t[G_\delta^+(t)|B(x_{\pm}(t), r_0)] \, dt < \infty. \]

To this end, we compute at \((x_+(t), t)\):

\[ \frac{d}{dt} u_+ = \int_{\mathbb{T}^n} \phi(x_+, y)(u(y) - u_+) \rho(y) \, dy. \]

We now use our assumption (126) and the fact that \(\tau \geq n\) to obtain the bound

\[ \frac{\lambda \mathbb{1}_{x < r_0}|x - y|}{d^n(x, y)} \leq \frac{\psi(x - y)d^{\tau - n}(x, y)}{d^\tau(x, y)} \leq \mathcal{M}^{\tau - n} \phi(x, y). \]
Thus, we have
\[-\frac{d}{dt} u_+ = \int \phi(x_+, y)(u_+ - u(y))\rho(y)\,dy\]
\[\geq \frac{1}{m_t(B(x_+, r_0))} \int_{G_+^+(t) \cap B(x_+, r_0)} (u_+ - u(y))\rho(y)\,dy \quad \text{(since } \Omega(x_+, y) \subset B(x_+, 0)\text{)}
\]
\[\geq \frac{\delta(t)u_+}{m_t(B(x_+, r_0))} \int_{G_+^+(t) \cap B(x_+, r_0)} \rho(y)\,dy
\]
\[= \delta(t)u_+ \mathbb{E}_t[G_+^+(t)|B(x_+, r_0)].\]

The result follows by integration:
\[\int_0^\infty \delta(t)\mathbb{E}_t[G_+^+(t)|B(x_+(t), r_0)]\,dt \lesssim \ln \frac{u_+(0)}{\lim_{t \to \infty} u_+(t)} \leq \ln \frac{u_+(0)}{u_-(0)}.
\]

**Step 2: Use of Campanato-Morrey norm.**

On this next step we obtain proper Campanato estimates that measure deviation of \(u\) from its average values in terms of global enstrophy. Denote the averages with respect to mass-measure by
\[u_{x,r} = \frac{1}{m_t(B(x, r))} \int_{B(x,r)} u(t, z)\,dm_t(z).\]

Let us recall that the communication domains \(\Omega(x, x')\) are confined to cones intersected with the diametric ball. This allows us to make the following geometric observation.

**Claim 6.6.** There exists a \(c_0 > 0\) depending only on the opening angle of the cones such that for any \(r > 0\) and any triple of points \((x, x', x^*)\) with \(|x - x^*| < c_0 r\) and \(|x' - x^*| < r\), we have \(\Omega(x, x') \subset B(x^*, r)\).
Hence, for any $T > r_0$ we have obtained

$$\int_{|x-x'| < r_0} |u(x) - u_{x,r}|^2 \rho(x) \, dx \leq \int_{|x-x'| < r_0} \frac{1}{m_t(B(x^*, r))} \rho(x^*) \rho(x') \, dx' \, dx$$

using that $m_t(B(x^*, r)) \geq m_t(\Omega(x,x')) = d^n(x,x')$

$$\leq \int_{|x-x'| < (1+c_0)r} \frac{1}{d^n(x,x')} |u(x) - u(x')|^2 \rho(x) \rho(x') \, dx' \, dx$$

$$\leq C \int_{\mathbb{T}^n} \phi(x,x') |u(x) - u(x')|^2 \rho(x) \rho(x') \, dx' \, dx.$$

From the energy equation (108), the right hand side is globally integrable on $\mathbb{R}_+$. Hence, we conclude with a time bound on the Campanato semi-norm,

$$\int_0^\infty |u|^2_{r_0} \, dt < \infty, \quad [u]|_{r_0} := \sup_{x^* \in \mathbb{T}^n, r < c_0} \int_{|x-x'| < r_0} |u(x) - u_{x,r}|^2 \rho(x) \, dx.$$  

Combined with (130) we have obtained

$$I = \int_0^\infty \left( \delta(t) E_t[G^\pm(t)|B(x_\pm(t), r_0)] + [u(t)]^2_{r_0} \right) \, dt < \infty.$$  

Clearly, for $A = e^{2I}$ we have

$$\int_T^{T+A} \frac{dt}{t \ln t} = 2I \text{ for all } T > 0.$$  

Hence, for any $T > 0$ we can find a $t \in [T, T+A]$ such that

$$[u(t)]^2_{r_0} < \frac{1}{t \ln t}$$

$$\mathbb{E}_t[G^\pm(t)|B(x_\pm(t), r_0)] + \mathbb{E}_t[G^-t(t)|B(x_-(t), r_0)] < \frac{1}{\delta(t) t \ln t}$$

In view of the assumed lower bound on the density this implies in particular that

$$\sup_{x^*, r < c_0} \int_{|x-x'| < r_0} |u(x) - u_{x,r}|^2 \, dx \leq \frac{1}{\ln t}.$$  

**Step 3: Sliding averages.**

Let $t \in [T, T+A]$ be the moment of time fixed above. Let us fix $r = \frac{1}{T+r_0}$, which lies within the reach of the Campanato metric. We will now reconnect the two averages $u_{x^+, r}$ and $u_{x^-, r}$ sliding along the line connecting $x^+$ and $x^-$, and show that the variation of those averages is small.

Denote the direction vector $n = \frac{x^+ - x^-}{|x^+ - x^-|}$ and define a sequence of overlapping balls, $B_k = B(x_k, c_0 r), k = 0, \ldots, K$, with centers given by $x_0 = x_-$ and defined recursively by $x_{k+1} = x_k + c_0 r n$.
up to $k = K - 1$ and ending with $x_K = x_+$, see Figure 5. The point is that the intersection of each consecutive pair is of the measure of order $r_0^2$: $|B_k \cap B_{k+1}| \geq c_1 r_0^2$.

Chebychev inequality, followed by (134) applied to the ball centered at $x_0$, yields
\[
|\{ x \in B_0 \cap B_1 : |u(x) - u_{x_0,r}| > \eta \}| \leq \frac{1}{\eta^2} \int_{B_0} |u(x) - u_{x_0,r}|^2 \, dx \leq \frac{1}{\eta^2 \ln t}.
\]

Let us set $\eta = \frac{2}{\sqrt{c_1 r_0^2 \ln t}}$ so that
\[
|\{ x \in B_0 \cap B_1 : |u(x) - u_{x_0,r}| > \eta \}| \leq \frac{1}{4} |B_0 \cap B_1|.
\]

Applying the same argument to the variation around the averaged value $u_{x_1,r}$, centered at $x_1$, we obtain
\[
|\{ x \in B_0 \cap B_1 : |u(x) - u_{x_1,r}| > \eta \}| \leq \frac{1}{4} |B_0 \cap B_1|.
\]

Consequently the complements of the two sets must have a point in common in $B_0 \cap B_1$:
\[
\{ x \in B_0 \cap B_1 : |u(x) - u_{x_0,r}| \leq \eta \} \cap \{ x \in B_0 \cap B_1 : |u(x) - u_{x_1,r}| \leq \eta \} \neq \emptyset,
\]

which implies that
\[
|u_{x_0,r} - u_{x_1,r}| \leq 2\eta.
\]

Continuing in the same manner we obtain the same bound for all consecutive averages:
\[
|u_{x_k,r} - u_{x_{k+1},r}| \leq 2\eta.
\]

Hence,
\[
(135) \quad |u_{x_-,r} - u_{x_+,r}| \leq 2K\eta \lesssim \frac{1}{\sqrt{\ln t}}.
\]

Note that $K \lesssim 1/r_0$, so it is bounded by an absolute constant. Furthermore, in view of (133), we can estimate
\[
u_{x_+,r} \geq \frac{1}{m_t(B(x_+,r))} \int_{B(x_+,r) \setminus G_+^t} u_+(t)(1 - \delta(t)) \, dm_t
\]
\[
\geq u_+(t)(1 - \delta(t))(1 - \mathbb{E}_t[G_+^t(t)|B(x_+(t), R_0)]) \geq u_+(t)(1 - \delta(t)) \left(1 - \frac{1}{\delta(t)t \ln t}\right).
\]

Hence,
\[
u_+(t) - u_{x_+,r}(t) \lesssim \delta(t) + \frac{1}{\delta(t)t \ln t} \lesssim \frac{1}{\sqrt{t \ln t}}
\]

if we set $\delta(t) = \frac{1}{\sqrt{\ln t}}$. A similar estimate holds for the bottom average. In conjunction with (135) these imply
\[
|u_+(t) - u_-(t)| \lesssim \frac{1}{\sqrt{\ln t}}.
\]

To conclude the proof we note that by the maximum principle
\[
|u_+(T^A) - u_-(T^A)| \lesssim \frac{1}{\sqrt{\ln t}} \sim \frac{1}{\sqrt{\ln(T^A)}}.
\]

Since $T$ is arbitrary this finishes the proof.

\[\square\]
Remark 6.7. Theorem 6.4 allows for an extension which improves upon the rate of alignment under more restrictive bound from below on the density. Specifically, the following statement can be proved along the lines of the above argument: suppose

\[ \rho(t, x) \geq \frac{c}{(1 + t)^\gamma}, \quad 0 \leq \gamma \leq 1, \]

then the solution aligns with the following algebraic rate

\[ \|u(t) - \bar{u}\|_\infty = \frac{o(1)}{t^{\frac{1}{2}(1-\gamma)}}. \]

7. Regularity I: local well-posedness and continuation criteria

The regularity theory for models with smooth communication is understandably quite different from singular models – the former is essentially Burgers’ equation with a dumping mechanism, while the latter is a degenerate fractional parabolic system with dissipation in the momentum equation. In this section we will go through a rather elementary but necessary for future exposition first step – proving local existence of classical solutions. We are not seeking the sharpest spaces to keep exposition simple. However, we do require our local solutions to have certain level of regularity for various phenomena to remain classically verifiable, such as mass conservation, and existence of characteristic flow. So, the results presented below will respect such requirements. We will also limit ourselves to the metric models, where the estimates are not excessively contaminated with density dependent coefficients.

7.1. Smooth models. Let us assume throughout that $\phi$ is sufficiently smooth to take as many derivatives as necessary in the course of our arguments below. We also assume that $\phi = \phi(x - y)$ is of convolution type and that the environment domain is $\mathbb{R}^n$. The exact same results will carry over to $\mathbb{T}^n$ with slight modifications.

Using the commutator structure of the alignment term and (105) we write the system (102) as

\[
\begin{aligned}
\rho_t + \nabla \cdot (\rho u) &= 0, \\
u_t + u \cdot \nabla u &= \phi * (u \rho) - u \phi * \rho.
\end{aligned}
\]

Suppose we would like to prove local existence of solutions in Sobolev class $u \in H^m$, $\rho \in H^k \cap L^1_+$, where $L^1_+$ denotes the set of non-negative functions in $L^1$. Note that the Sobolev embedding does not guarantee that $\rho \in L^1_+$ if it is in a higher Sobolev class, yet $u \in L^\infty$ automatically if $m > \frac{n}{2}$. Both conditions are natural quantities to include in the class as they are controled by dynamics a priori.

One can obtain local existence rather easily for a viscous regularization:

\[
\begin{aligned}
\rho_t + \nabla \cdot (\rho u) &= \varepsilon \Delta \rho, \\
u_t + u \cdot \nabla u &= \phi * (u \rho) - u \phi * \rho + \varepsilon \Delta u.
\end{aligned}
\]

Indeed, we are going to denote the grand quantity $Z = (u, \rho)$ and consider the equivalent mild formulation of (138):

\[
Z(t) = e^{\varepsilon t \Delta} Z_0 + \int_0^t e^{\varepsilon (t-s) \Delta} N(Z(s)) \, ds,
\]

where $N(Z)$ denotes all the non-linear terms in (138). The argument goes by the standard contractivity argument. Let us fix $Z_0 \in H^m \times (H^k \cap L^1_+)$ and consider the map

\[
T[Z](t) = e^{\varepsilon t \Delta} Z_0 + \int_0^t e^{\varepsilon (t-s) \Delta} N(Z(s)) \, ds.
\]
We need to show that for some small $T$ this maps is a contraction on $C([0,T); B_1(Z_0))$, where $B_1$ is understood in metric of $X = H^m \times (H^k \cap L^1)$. Let us invariance, while contractivity following similarly. First, by continuity of the heat semigroup,

$$\|Z_0 - e^{t\Delta} Z_0\|_X \leq \frac{1}{2},$$

for small $t$. To estimate the integral we first recall analyticity property of the heat semigroup:

$$\|\nabla e^{t\Delta} f\|_{L^p} \lesssim \frac{1}{\sqrt{\varepsilon t}}\|f\|_{L^p}, \quad 1 \leq p \leq \infty.$$

So, considering the $\rho$-component we obtain

$$\left\| \int_0^t e^{(t-s)\Delta} \nabla \cdot (\rho u) \, ds \right\|_{L^1} \leq \int_0^t \frac{1}{\sqrt{\varepsilon(t-s)}} \|u\| L^1 \, ds$$

$$\leq t^{1/2} \sup_s \|u(s)\|_\infty \|\rho(s)\|_1 \leq T^{1/2} (\|Z_0\|^2 + 1) < \frac{1}{2}$$

for small $T$. For $H^k$ we argue similarly for $L^2$. So, let $\partial^k$ be a multiindex partial derivative of order $k$. We will use the product estimate

$$\|\partial^k (u \rho)(s)\|_{L^2} \leq \|u\|_\infty \|\rho\|_{H^k} + \|u\|_{H^k} \|\rho\|_\infty.$$

It is clear at this point that in order to close the estimates in $X$ we need to assume that $\frac{n}{2} < k \leq m$, in which case

$$\|\partial^k (u \rho)(s)\|_{L^2} \leq \|Z(s)\|_X^2.$$

So, by the same argument we obtain

$$\left\| \partial^k \int_0^t e^{(t-s)\Delta} \nabla \cdot (\rho u) \, ds \right\|_{L^2} \leq \int_0^t \frac{1}{\sqrt{\varepsilon(t-s)}} \|\partial^k (u \rho)(s)\|_{L^2} \, ds < \frac{1}{2}.$$

On the velocity side there are two terms to handle. For the transport part, the $L^2$-estimate is straightforward, and

$$\left\| \partial^m \int_0^t e^{(t-s)\Delta} u \cdot \nabla u \, ds \right\|_{L^2} \leq \int_0^t \frac{1}{\sqrt{\varepsilon(t-s)}} \|\partial^{m-1} (u \cdot \nabla u)(s)\|_{L^2} \, ds$$

$$\leq \int_0^t \frac{1}{\sqrt{\varepsilon(t-s)}} (\|u\|_{H^{m-1}} \|\nabla u\|_\infty + \|u\|_{H^m} \|u\|_\infty) \, ds < \frac{1}{4},$$

provided $m > \frac{n}{2} + 1$ to ensure embedding of $W^{1,\infty}$ into $H^m$. The commutator term $\phi \ast (u \rho) - u \phi \ast \rho$ is even easier, since the derivatives are absorbed by the kernel except when all fall on $u$, which results in the same estimate.

We have shown that $T : C([0,T); B_1(Z_0)) \to C([0,T); B_1(Z_0))$ is a contraction, and so, we obtain a local solution on a time interval dependent on $\varepsilon$. Denoting $T^*$ the maximal time of existence in $C([0,T); X)$ we show that $T^*$ depends only on the $X$-norm of the initial condition. We do it by establishing a priori estimates that are independent of $\varepsilon$ and which will allow us to pass to the limit of vanishing viscosity. So, the grand quantity we are trying to control is

$$Y_{m,k} = \|u\|_{H^m}^2 + \|\rho\|_{H^k}^2 + |\rho|^2_1.$$

To start, we write the continuity equation as

$$\rho_t + u \cdot \nabla \rho + (\nabla \cdot u)\rho = 0.$$

So, testing with $\partial^k \rho$ we obtain

$$\frac{d}{dt} \|\rho\|_{H^k}^2 = \int (\nabla \cdot u) |\partial^k \rho|^2 \, dx - \int (\partial^k (u \cdot \nabla \rho) - u \cdot \nabla \partial^k \rho) \partial^k \rho \, dx - \int \partial^k ((\nabla \cdot u) \rho) \partial^k \rho \, dx - \varepsilon \|\rho\|_{H^{k+1}}^2.$$
We dismiss the last term. Recalling the classical commutator estimate
\[
\|\partial^k (fg) - f\partial^k g\|_2 \leq |\nabla f|_\infty |g|_{H^{k-1}} + \|f\|_{H^k} |g|_\infty,
\]
we obtain
\[
\frac{d}{dt} \|\rho\|^2_{H^k} \leq |\nabla \rho|_\infty |\rho|_{H^k}^2 + \|u\|_{H^k} |\rho|_{H^k} |\nabla \rho|_\infty + \|u\|_{H^{k+1}} |\rho|_{H^k} |\rho|_\infty \\
\leq C(|\nabla u|_\infty + |\nabla \rho|_\infty + |\rho|_\infty) Y_{m,k},
\]
provided $m \geq k + 1$. The $L^2$ norm of $\rho$ obeys a similar estimate trivially, and the $L^1$-norm is conserved.

For the velocity equation we apply the same commutator estimate for the material derivative part:
\[
\int \partial^m (u \cdot \nabla u) \partial^m u \, dx = - \int \nabla \cdot u |\partial^m u|^2 \, dx + \int [\partial^m (u \cdot \nabla u) - (u \cdot \nabla \partial^m u)] \partial^m u \, dx \\
\lesssim |\nabla u|_\infty \|u\|^2_{H^m}.
\]

For the alignment term, we can put all the derivatives onto the kernel whenever possible and the only term that is left out is $|\partial^m u|^2 \phi \ast \rho$ with $\phi \ast \rho$ clearly bounded by $|\phi|_\infty M$, a priori conserved quantity. So, we obtain
\[
\frac{d}{dt} \|u\|^2_{H^m} \leq (|\nabla u|_\infty + C(|\phi|_{C^m}, M)) \|u\|^2_{H^m}.
\]
The similar bound for $\frac{d}{dt} \|u\|^2_{H^2}$ is derived trivially. So, we obtain
\[
\frac{d}{dt} \|u\|^2_{H^m} \leq (|\nabla u|_\infty + C(|\phi|_{C^m}, M)) \|u\|^2_{H^m}.
\]

It is important to note that this bound is independent of the higher norms of the density. Combining the two equations we obtain
\[
\frac{d}{dt} Y_{m,l} \leq C(|\nabla u|_\infty + |\nabla \rho|_\infty + |\rho|_\infty + C(|\phi|_{C^m}, M)) Y_{m,k}.
\]

Of course $|\nabla u|_\infty + |\nabla \rho|_\infty + |\rho|_\infty \leq Y_{m,l}$ provided $k > \frac{n}{2} + 1$, which adds the last restriction on the exponents for the argument to work. So, if $m \geq k + 1 > \frac{n}{2} + 2$, then
\[
\frac{d}{dt} Y_{m,k} \leq C_1 Y_{m,k} + C_2 Y^2_{m,k}.
\]

Solving the Riccati equation gives a uniform bound on the $X$-norm on a time interval inversely proportional to $\|Z_0\|_X$, but independent of $\varepsilon$. Thus, solutions to (139) with the same initial data exist on a common time interval $[0, T_0]$ where they are uniformly bounded in $C([0, T_0]; X)$.

Let us also note that keeping the dissipative terms in the estimates above also shows that
\[
\varepsilon \int_0^{T_0} \left(\|\rho(s)\|^2_{H^{k+1}} + \|u(s)\|^2_{H^{m+1}}\right) \, ds < C,
\]
where $C$ is independent of $\varepsilon$. Then
\[
\|Z_t\|_{L^2} \leq \|Z\|^2_X + \varepsilon \|Z\|_{H^2} \leq \|Z\|^2_X + \varepsilon \|Z\|_{H^{k+1} \times H^{m+1}}.
\]
So, $Z_t \in L^2([0, T_0]; L^2)$. Passing to a subsequence we find a weak limit $Z_\varepsilon \to Z$ in $L^\infty([0, T_0]; X)$ and $(Z_\varepsilon)_t \to Z_t$ in $L^2([0, T_0]; L^2)$ (technically, a limit in $L^1$ may end up being a measure of bounded variation, however as a member of $H^k$ it is absolutely continuous, hence in $L^1$). Since $Z_t \in L^2([0, T_0]; L^2)$, $Z$ is weakly continuous with values in $L^2$. Since $L^2$ is dense in $H^{-m}$ and $H^{-k}$ this
implies weak continuity $Z \in C_w([0, T_0]; H^m \times H^k)$. Strong continuity of the density follows from the equations itself:

$$\|\rho_t\|_{L^1} \leq \|\rho \nabla u\|_1 + \|u \nabla \rho\|_1 \leq \|Z\|_X^2 < C.$$ 

Further regularity in time $Z_t$ follows from measuring smoothness of the system one level down and performing similar product estimates as above.

Having established local existence in $X$ let us come back to (142) and notice that this solution can in fact be extended beyond $T_0$ if we know that $|\nabla u|_\infty$ is integrable on $[0, T_0]$. Indeed, $|\rho|_\infty$ can be bounded by solving the continuity equation along characteristics

$$\rho(X(t, \alpha), t) = \rho(\alpha, 0) \exp \left\{ - \int_0^t \nabla \cdot u(X(s, \alpha), s) \, ds \right\}. \quad (143)$$

Using (141) we see that $\|u\|_{L^m}$ is also bounded uniformly, and hence so is $|\nabla^2 u|_\infty$ since $m > \frac{n}{2} + 2$. Bootstrapping further by differentiating the continuity equation we bound $|\nabla \rho|_\infty$ in a similar fashion.

Having this continuation criterion at hand we can further improve the local existence result by establishing control over $|\nabla u|_\infty$ directly. First, by the maximum principle, $|u(t)|_\infty \leq |u_0|_\infty$. Writing equation for one component $\partial_i u_j$ we have

$$\partial_i \partial_j u_j + u \cdot \nabla \partial_i u_j + \partial_i u \cdot \nabla u_j = (\partial_i \phi) \ast (u_j \rho) - \partial_i u_j \phi \ast \rho - u_j (\partial_i \phi) \ast \rho.$$ 

Evaluating at the maximum and minimum and adding up over $i, j$ we obtain

$$\frac{d}{dt}|\nabla u|_\infty \leq |\nabla u|_\infty^2 + CM|u|_\infty + M|\nabla u|_\infty.$$ 

Hence, $|\nabla u|_\infty$ is uniformly bounded a priori on a time interval depending only on $|\nabla u_0|_\infty^{-1}$. So, the continuation criterion allows to extend our local solution to that time interval.

Let us record the obtained results in the following theorem.

**Theorem 7.1** (Local existence of classical solutions). Suppose $m \geq k + 1 > \frac{n}{2} + 2$, and $(u_0, \rho_0) \in H^m \times (H^k \cap L^1)$. Then there exists time $T_0 = T_0(|\nabla u_0|_\infty^{-1}, M)$ and a unique solution to (138) on time interval $[0, T_0]$ in the class

$$(u, \rho) \in C_w([0, T_0); H^m \times (H^k \cap L^1)) \cap \text{Lip}([0, T_0); H^{m-1} \times (H^{k-1} \cap L^1)) \quad (144)$$

satisfying the given initial condition. Moreover, any classical local solution on $[0, T_0]$ in class (144) and satisfying

$$\int_0^{T_0} |\nabla u|_\infty \, dt < \infty, \quad (145)$$

can be extended beyond $T_0$.

Theorem 7.1 can be used as a stepping stone to obtain solutions with less smoothness, especially for $\rho$, as long as regularity of $u$ permits to define smooth characteristic flow map with sufficient compactness properties.

So, let us first assume $u_0 \in H^m$, $m > \frac{n}{2} + 1$ and $\rho_0 \in L^1$, the most basic assumption on the density. mollifying the data $((u_0)_\varepsilon, (\rho_0)_\varepsilon)$, due to Theorem 7.1, we obtain a family of local solutions on a common time interval $T_0$, since $H^m \subset W^{1,\infty}$, and $|\nabla (u_0)_\varepsilon|_\infty \leq |\nabla u_0|_\infty$. We also note that the estimate (141) holds for any integer $m$, hence $u_\varepsilon \in L^\infty([0, T_0); H^m)$ uniformly. We also have $\partial_i u_\varepsilon \in L^\infty([0, T_0); H^{m-1}) \subset L^\infty([0, T_0); L^\infty)$. Let us note that $H^m \subset W^{1+\delta,\infty}$ for some $\delta > 0$, and the embedding is compact on any bounded set, and of course $W^{1+\delta,\infty} \subset L^\infty$. So, the Aubin-Lions-Simon Lemma implies that the family is compact in $C([0, T_0); W^{1+\delta,\infty})$ on any bounded set. Passing to a subsequence we obtain a weak limit $u$ in $L^\infty([0, T_0); H^m)$ and strong in $C([0, T_0); W^{1+\delta,\infty})$ on any bounded set. The velocity also belongs to $C_w([0, T_0); H^m)$ as a consequence of the two memberships. Similarly, the family of flows $X_\varepsilon$ belongs to $L^\infty([0, T_0); W^{1,\infty}) \cap \text{Lip}([0, T_0); L^\infty)$
so is compact in $C([0,T_0);C^b)$, $\delta < 1$, on any bounded set. We can thus claim strong uniform convergence of the flow maps as well. Considering the solution to the continuity equation

$$
(146) \quad \rho_\varepsilon(X_\varepsilon(t,\alpha),t) = \rho_\varepsilon(\alpha,0) \exp \left\{ -\int_0^t \nabla \cdot u_\varepsilon(X_\varepsilon(s,\alpha),s) \, ds \right\},
$$

we can clearly pass to the strong limit in $L^1$ and the limit satisfies (143).

Note that the density essentially plays the role of a passive scalar. In particular, if $\rho_0 \in L^1_+ \cap L^\infty$ initially, then by formula (143) it will remain in the same class on the entire time interval.

The argument above can be elevated to any higher smoothness $H^m \times (L^1_+ \cap W^{k,\infty})$ as long as $m > \frac{n}{2} + k + 1$. It goes by differentiating the continuity equation $k$ times and running the same compactness procedure. Weak continuity of the density in $L^1_+ \cap W^{k,\infty}$ follows from the established regularity properties of the velocity field and the corresponding formula for the solution of $\partial^k \rho$.

The continuation criterion (145) remains valid in this setting as well.

**Theorem 7.2** (Local existence of strong solutions). Suppose $m > \frac{n}{2} + k + 1$, $k = 0,1,...$ and $(u_0,\rho_0) \in H^m \times (L^1_+ \cap W^{k,\infty})$. Then there exists time $T_0 = T_0(|\nabla u_0|^{-1},M)$ and a unique solution to (138) on time interval $[0,T_0)$ in the class

$$
(147) \quad (u,\rho) \in C_w([0,T_0);H^m \times (L^1_+ \cap W^{k,\infty})).
$$

Moreover, any such solution satisfying (145) can be extended beyond $T_0$.

The case $k = 1$ will already be interesting in dimension 1, where the given regularity is sufficient to justify pointwise evaluation of $u_\varepsilon$ and to study strong flocking.

### 7.2. Singular models.

For models with singular kernels given by (5), $\beta = n + \alpha$, $0 < \alpha < 2$, the operator $L_\phi$ becomes the fractional Laplacian, although we also consider local kernels. So we set

$$
\phi(r) = \frac{h(r)}{r^{n+\alpha}},
$$

and consider the model on periodic domain

$$
(148) \quad \begin{cases}
\rho_t + \nabla \cdot (\rho u) = 0, \\
u_t + u \cdot \nabla u = \int_{\mathbb{T}^n} \phi(x-y)(u(y) - u(x))\rho(y) \, dy \quad (x,t) \in \mathbb{T}^n \times \mathbb{R}^+.
\end{cases}
$$

The important fact is that we have a coercivity bound

$$
(149) \quad c_1\|f\|_{H^\alpha} - c_2\|f\|_{L^2} \leq \|L_\phi f\|_{L^2} \leq c_1\|f\|_{H^\alpha} + c_3\|f\|_{L^2}.
$$

We will be casting our regularity theory for singular models on the periodic domain $\mathbb{T}^n$ and for non-vacuous solutions only. This is motivated by technical reasons rather than applications, although one can argue that periodic conditions are suitable for studying flocks in the bulk. The primary reason is that we require uniform parabolicity of the commutator (104) for estimates to go through. Such parabolicity depends on the pointwise bound $\rho > c_0 > 0$, which is consistent with finite mass of the flock only on bounded domains.

Furthermore, we can easily construct a blowing up solution with vacuum and with a local kernel. Indeed, consider a local kernel $\text{Supp} \phi \subset B_1(0)$. Let initial density $\text{Supp} \rho_0 \subset B_\varepsilon(0)$, while $u_0 = 1$ on $B_{10}(0)$, $v_0 = 0$ on $B_{10+\varepsilon}(0)$ and smooth in between. Then the density will remain in $B_2(0)$ for a time period of at least $t < 1$, due to $u \leq 1$. During this time the momentum equation will remain pure Burgers, hence the solution will evolve into a shock at a time $t \sim \varepsilon < 1$. A similar argument can be done even for a global singular kernel on the periodic domain and this was formalized in [13].
Performing energy estimates in the same fashion as for smooth models will inevitably create a derivative overload on the density. Instead we consider another "almost conserved" quantity

\[ e = \nabla \cdot u + L_\phi \rho, \]

which satisfies the equation

\[ e_t + \nabla \cdot (ue) = (\nabla \cdot u)^2 - \text{Tr}(\nabla u)^2. \]

Let us derive it in general for the sake of completeness.

We have

\[ \partial_t e + \nabla \cdot [ue] = (\nabla \cdot u)^2 - \text{Tr}(\nabla u)^2. \]

Since \( \phi \) is a convolution kernel, we have that

\[ \partial_t L_\phi + \nabla \cdot L_\phi (\rho u) = 0. \]

Taking the divergence of the velocity equation, we obtain

\[ \partial_t (\nabla \cdot u) + \nabla \cdot [u \cdot \nabla u] = \nabla \cdot L_\phi (\rho u) - \nabla \cdot [u L_\phi \rho] \]

with

\[ \nabla \cdot [u L_\phi \rho] = L_\phi \rho \nabla \cdot u + u \cdot \nabla L_\phi \rho \]

and

\[ \nabla \cdot [u \cdot \nabla u] = \text{Tr}(\nabla u)^2 + u \cdot \nabla (\nabla \cdot u). \]

On one hand, combining (153) and (154), we obtain that

\[ \partial_t e + L_\phi \rho \nabla \cdot u + u \cdot \nabla e + \text{Tr}(\nabla u)^2 = 0. \]

Adding and subtracting now \( (\nabla \cdot u)^2 \) produces (166).

From the order of terms that enter into the formula for \( e \), it is clear that the natural correspondence in regularity for state variables involved is \( (u \in H^{m+1}) \sim (\rho \in H^{m+\alpha}) \).

Note that in 1D the right hand side vanishes and we have a perfect conservation law. This case will be discussed at length in Section 8.

The grand quantity to be estimated is

\[ Y_m = \|u\|^2_{H^{m+1}} + \|e\|^2_{H^m} + |e|_\infty + |\rho|_1 + |\rho^{-1}|_\infty, \]

which is equivalent to \( Y_m \sim \|u\|^2_{H^{m+1}} + \|\rho\|^2_{H^{m+\alpha}} + |\rho^{-1}|_\infty \) in view of (149).

Our strategy will be very similar to the smooth case, where we obtain local solutions via viscous regularization, and prove a continuation criterion via a priori estimates on \( Y_m \). We assume throughout that \( m > \frac{\alpha}{2} + 1 \) and \( 0 < \alpha < 2 \).

So, let us start with (139) and consider the mild formulation

\[ \rho(t) = e^{\varepsilon t \Delta} \rho_0 - \int_0^t e^{\varepsilon (t-s) \Delta} \nabla \cdot (u \rho)(s) \, ds \]

\[ u(t) = e^{\varepsilon t \Delta} u_0 - \int_0^t e^{\varepsilon (t-s) \Delta} u \cdot \nabla u(s) \, ds + \int_0^t e^{\varepsilon (t-s) \Delta} C_\phi (u, \rho)(s) \, ds. \]

Let us denote as before \( Z = (\rho, u) \) and by \( T[Z](t) \) the right hand side of the mild formulation. In order to apply the standard fixed point argument we have to show that \( T \) leaves the set C([0, T_\delta]; B_\delta(Z_0)) invariant, where \( B_\delta(Z_0) \) is the ball of radius \( \varepsilon \) around initial condition \( Z_0 \), and that it is a contraction. We limit ourselves to showing details for invariance as the estimates involved in proving Lipschitzness are similar.

First we assume that \( \rho \) has no vacuum: \( \rho_0(x) \geq c_0 > 0 \). Since the metric we are using for \( \rho \in H^{m+\alpha} \) controls \( L^\infty \) norm, if \( \delta > 0 \) is small enough then for any \( \|\rho - \rho_0\|_{H^{m+\alpha}} < \delta \) one obtains

\[ \rho(x) > \frac{1}{2} c_0. \]
So, let us assume that $Z \in C([0, T); B_\delta(Z_0))$. It is clear that $\|e^{\epsilon t}Z_0 - Z_0\| < \frac{\delta}{2}$ provided time $t$ is short enough. The $Z$ has some bound $\|Z\| \leq C$. Using that let us estimate the norms under the integrals. First, recall that $\|\Lambda_\alpha e^{\epsilon t}Z\|_{L^2} \lesssim \frac{1}{(\epsilon t)^{\alpha/2}}$. In the case $\alpha \geq 1$, we have

$$
\left\| \partial^m \Lambda_\alpha \int_0^t e^{\epsilon (t-s)} \nabla \cdot (u\rho)(s) \, ds \right\|_{L^2} \lesssim \int_0^t \frac{1}{(t-s)^{\alpha/2}} \left\| \partial^{m+1} (u\rho)(s) \right\|_{L^2} \, ds \\
\leq \int_0^t \frac{1}{(t-s)^{\alpha/2}} \left\| u \right\|_{\dot{H}^{m+1}} \left\| \rho \right\|_{\dot{H}^{m+\alpha}} \, ds \leq C^2 t^{1-\alpha/2} < \frac{\delta}{2},
$$

provided $T = T(\delta, \epsilon)$ is small enough. In the case $\alpha < 1$, we combine instead one full derivative with the heat semigroup, and the rest $\partial^{m+\alpha}$ gets applied to $u\rho$, which produces a similar bound.

Moving on to the $u$-equation, we have

$$
\left\| \partial^{m+1} \int_0^t e^{\epsilon (t-s)} \Delta u \cdot \nabla u(s) \, ds \right\|_{L^2} \lesssim \int_0^t \frac{1}{(t-s)^{1/2}} \left\| \partial^{m} (u \cdot \nabla u)(s) \right\|_{L^2} \, ds \\
\leq \int_0^t \frac{1}{(t-s)^{1/2}} \left\| u \right\|_{\dot{H}^{m+1}} \left\| u \right\|_{\dot{H}^{m}} \, ds \leq C^2 t^{1/2} < \frac{\delta}{4}.
$$

As to the commutator form, for $\alpha \leq 1$ the computation is very similar: we combine one derivative with the heat semigroup and for the rest we use (149):

$$
\left\| \partial^m C_\alpha(u, \rho) \right\|_{L^2} \lesssim \left\| u \right\|_{m+\alpha} \left\| \rho \right\|_{m+\alpha} < C^2,
$$

and the rest follows as before. When $\alpha > 1$ we combine $\alpha$ derivatives with the semigroup, and the rest follows as before.

We have proved that $\|T[Z](t) - Z_0\| < \delta$, for a short time and hence, $T$ leaves $C([0, T(\delta, \epsilon)); B_\delta(Z_0))$ invariant.

Now let us make a priori estimates for viscous solutions independent of $\epsilon$. Note that the dissipation terms in all the following computations are negative and as such will be ignored.

First, evaluating the continuity equation at a point of minimum $x_-$ and denoting $\rho_- = \min \rho$ we readily obtain

$$
\frac{d}{dt} \rho_- = \rho_- \Delta u + \epsilon \Delta \rho(x_-) \geq -\rho_- |\nabla u|_\infty.
$$

Hence,

$$
\frac{d}{dt} |\rho|_{\infty} \leq |\rho|_{\infty} |\nabla u|_{\infty} \leq |\nabla u|_{\infty} Y_m.
$$

Furthermore,

$$
(158) \quad \frac{d}{dt} |e|_{\infty} \leq |\nabla u|_{\infty} |e|_{\infty} + |\nabla u|_{\infty}^2 \leq |\nabla u|_{\infty} Y_m.
$$

Let us continue with estimates on the $e$-quantity. We have (dropping integral signs)

$$
\frac{d}{dt} \|e\|_{H^m}^2 \leq \partial^m \nabla u \cdot \nabla \partial^m e + \partial^m e \partial^m (u \cdot \nabla e) - u \cdot \nabla \partial^m e + \partial^m e \partial^m (e \nabla \cdot u) + \partial^m e [(\nabla \cdot u)^2 - \text{Tr}(\nabla u)^2]
$$

In the first term we integrate by part and estimate

$$
|\partial^m e \nabla \cdot \nabla \partial^m e| \leq \|e\|_{H^m}^2 |\nabla u|_{\infty}.
$$

For the next commutator term we use (140)

$$
|\partial^m e \partial^m (u \cdot \nabla e) - u \cdot \nabla \partial^m e| \leq \|e\|_{H^m}^2 \|\nabla u|_{\infty} + \|e\|_{H^m} \|u\|_{H^m} \|\nabla e|_{\infty}.
$$

Using Gagliardo-Nirenberg inequality we estimate the latter term as

$$
\|e\|_{H^m} \|u\|_{H^m} \|\nabla e|_{\infty} \lesssim \|e\|_{H^m}^{\theta_1} \|H^m \nabla u|_{\infty}^{1-\theta_1} \|e\|_{H^m}^{\theta_2} \|H^m \nabla e|_{\infty}^{1-\theta_2},
$$
Thus, and using (140) we obtain
\[ \varepsilon > \frac{2}{2m-n} \]

The two exponents add up to 1, so by the generalized Young inequality,
\[ \leq \left( \|e\|_{H^m}^2 + \|u\|_{H^{m+1}}^2 \right) (\|e\|_{\infty} + \|\nabla u\|_{\infty}) \leq (\|e\|_{\infty} + \|\nabla u\|_{\infty}) Y_m \]

Next term in the e-equation is estimated by the product formula
\[ \|e\|_{H^m}^2 \leq \|f\|_{H^m} \|g\|_{\infty} + \|f\|_{\infty} \|g\|_{H^m}. \]

So, we have
\[ \|e\|_{H^m}^2 \leq \|f\|_{H^m} \|g\|_{\infty} + \|f\|_{\infty} \|g\|_{H^m}. \]

Finally,
\[ \|e\|_{H^m}^2 \leq \|f\|_{H^m} \|g\|_{\infty} + \|f\|_{\infty} \|g\|_{H^m}. \]

Thus,
\[ \|e\|_{H^m}^2 \leq \|f\|_{H^m} \|g\|_{\infty} + \|f\|_{\infty} \|g\|_{H^m}. \]

Next perform the main technical estimate on the velocity equation. We have
\[ \partial_t \|u\|_{H^{m+1}}^2 = -\partial^{m+1} (u \cdot \nabla u) \cdot \partial^{m+1} u + \partial^{m+1} C_\phi (u, \rho) \cdot \partial^{m+1} u. \]

The transport term is estimated using the classical commutator estimate
\[ \partial^{m+1} (u \cdot \nabla u) \cdot \partial^{m+1} u = u \cdot \nabla (\partial^{m+1} u) \cdot \partial^{m+1} u + [\partial^{m+1}, u] \nabla u \cdot \partial^{m+1} u \]

Then
\[ u \cdot \nabla (\partial^{m+1} u) \cdot \partial^{m+1} u = -\frac{1}{2} (\nabla \cdot u) \|\partial^{m+1} u\|_{\infty}^2 \leq \|\nabla u\|_{\infty} \|u\|_{H^{m+1}}^2, \]

and using (140) we obtain
\[ \|\partial^{m+1}, u \|\nabla u \cdot \partial^{m+1} u \| \leq \|\nabla u\|_{\infty} \|u\|_{H^{m+1}}^2. \]

Thus,
\[ \partial_t \|u\|_{H^{m+1}}^2 \leq \|\nabla u\|_{\infty} Y_m + \partial^{m+1} C_\phi (u, \rho) \cdot \partial^{m+1} u. \]

Let us expand the commutator
\[ \partial^{m+1} C_\phi (u, \rho) = \sum_{l=0}^{m+1} \left( \begin{array}{l} m+1 \end{array} \right) C_\phi (\partial^l u, \partial^{m+1-l} \rho). \]

One end-point case, \( l = m + 1 \), gives rise to a dissipative term:
\[ \int_{T^n} C_\phi (\partial^{m+1} u, \rho) \cdot \partial^{m+1} u \, dx = -\frac{1}{2} \int_{T^{m+1}} \phi (z) \|\partial\|_{H^{m+1}}^2 \rho (x + z) \, dz \, dx \]

The first term is bounded by
\[ \rho_- \int_{T^{m+1}} \phi (z) \|\partial\|_{H^{m+1}}^2 \rho (z) \, dz \, dx \sim -\rho_- \|u\|_{H^{m+1} + \frac{1}{2}, \infty}^2, \]

which is the main dissipation term. The second is estimated as follows. Let us pick an \( \varepsilon > 0 \) so small that \( 1 + \frac{2}{\alpha + \varepsilon} > \alpha > \varepsilon \). Then
\[ \int_{T^{m+1}} \phi (z) \|\partial\|_{H^{m+1}}^2 \rho (z) \, dz \, dx \sim -\rho_- \|\rho\|_{H^{m+1} + \frac{1}{2}, \infty}^2 + \rho_- \|\nabla \|_{H^{m+1} + \frac{1}{2}}^2, \]

where \( \theta_1 = \frac{n-2(m+1)}{n-2m} \) and \( \theta_2 = \frac{2}{2m-n} \). The two exponents add up to 1, so by the generalized Young inequality,
\[ \leq \left( \|e\|_{H^m}^2 + \|u\|_{H^{m+1}}^2 \right) (\|e\|_{\infty} + \|\nabla u\|_{\infty}) \leq (\|e\|_{\infty} + \|\nabla u\|_{\infty}) Y_m \]

Next term in the e-equation is estimated by the product formula
\[ \|e\|_{H^m}^2 \leq \|f\|_{H^m} \|g\|_{\infty} + \|f\|_{\infty} \|g\|_{H^m}. \]

So, we have
\[ \|e\|_{H^m}^2 \leq \|f\|_{H^m} \|g\|_{\infty} + \|f\|_{\infty} \|g\|_{H^m}. \]

Finally,
\[ \|e\|_{H^m}^2 \leq \|f\|_{H^m} \|g\|_{\infty} + \|f\|_{\infty} \|g\|_{H^m}. \]

Thus,
\[ \|e\|_{H^m}^2 \leq \|f\|_{H^m} \|g\|_{\infty} + \|f\|_{\infty} \|g\|_{H^m}. \]
where the first term is absorbed into dissipation. So,
\[
\int_{\mathbb{T}^n} C_\phi(\partial^{m+1}\mathbf{u}, \rho) \cdot \partial^{m+1}\mathbf{u} \, dx \lesssim -\rho_- \|\mathbf{u}\|_{H^{m+1+\frac{3}{2}}}^2 + \rho_-^{-1} \|\nabla \rho\|_{C^1}^2 Y_m.
\]
Let us consider first the other end-point case of \(l = 0\). In this case the density suffers a derivative overload. We apply the following “easing” technique:
\[
\int_{\mathbb{T}^n} C_\phi(\mathbf{u}, \partial^{m+1}\rho) \cdot \partial^{m+1}\mathbf{u} \, dx = \int_{\mathbb{T}^{2n}} \phi(z) \delta_z \mathbf{u}(x) \partial^{m+1}\rho(x+z) \partial^{m+1}\mathbf{u}(x) \, dz \, dx.
\]
Observe that
\[
\partial^{m+1}\rho(x+z) = \partial_z \partial_x^m \rho(x+z) = \partial_z (\partial_x^m \rho(x+z) - \partial_x^m \rho(x)) = \partial_z \delta_z \partial_x^m \rho(x).
\]
Let us now integrate by parts in \(z\):
\[
\int_{\mathbb{T}^n} C_\phi(\mathbf{u}, \partial^{m+1}\rho) \cdot \partial^{m+1}\mathbf{u} \, dx = \int_{\mathbb{T}^{2n}} \partial_z \phi(z) \delta_z \mathbf{u}(x) \partial_z \partial_x^m \rho(x) \partial^{m+1}\mathbf{u}(x) \, dz \, dx + \int_{\mathbb{T}^{2n}} \phi(z) \partial_x \mathbf{u}(x+z) \delta_z \partial_x^m \rho(x) \partial^{m+1}\mathbf{u}(x) \, dz \, dx := J_1 + J_2.
\]
Let us start with the \(J_2\) first. By symmetrization,
\[
J_2 = \int_{\mathbb{T}^{2n}} \delta_z \partial_x \mathbf{u}(x) \delta_z \partial_x^m \rho(x) \partial^{m+1}\mathbf{u}(x) \partial(z) \, dz \, dx - \int_{\mathbb{T}^{2n}} \partial_x \mathbf{u}(x) \delta_z \partial_x^m \rho(x) \delta_z \partial^{m+1}\mathbf{u}(x) \phi(z) \, dz \, dx := J_{2,1} + J_{2,2}.
\]
Term \(J_{2,1}\) will appear in a series of similar terms that we will estimate systematically below. The bound for \(J_{2,2}\) is rather elementary:
\[
J_{2,2} \leq \|\nabla \mathbf{u}\|_{C^1} \|\mathbf{u}\|_{H^{m+1+\alpha/2}} + \|\rho\|_{H^{m+\alpha/2}} \leq \epsilon \rho_- \|\mathbf{u}\|_{H^{m+1+\alpha/2}}^2 + \rho_-^{-1} \|\nabla \mathbf{u}\|_{C^1}^2 Y_m.
\]
Similar computation can be made for \(J_1\). Indeed, using that \(\partial_z \phi(z)\) is odd, by symmetrization, we have
\[
J_1 = \frac{1}{2} \int_{\mathbb{T}^{2n}} \partial_z \phi(z) \delta_z \mathbf{u}(x) \partial_z \partial_x^m \rho(x) \delta_z \partial^{m+1}\mathbf{u}(x) \, dz \, dx.
\]
Replacing \(\|\delta_z \mathbf{u}(x)\| \leq \|z\|_{C^1}\), the rest of the term is estimated exactly as \(J_{2,2}\).
To summarize, we have obtained the bound
\[
\int_{\mathbb{T}^n} C_\phi(\mathbf{u}, \partial^{m+1}\rho) \cdot \partial^{m+1}\mathbf{u} \, dx \leq \epsilon \rho_- \|\mathbf{u}\|_{H^{m+1+\alpha/2}}^2 + \rho_-^{-1} \|\nabla \mathbf{u}\|_{C^1}^2 Y_m.
\]
Let us now examine the rest of the commutators \(C_\phi(\partial^l \mathbf{u}, \partial^{m+1-l}\rho)\) for \(l = 1, \ldots, m\). After symmetrization we obtain
\[
\int_{\mathbb{T}^n} C_\phi(\partial^l \mathbf{u}, \partial^{m+1-l}\rho) \cdot \partial^{m+1}\mathbf{u} \, dx = \frac{1}{2} \int_{\mathbb{T}^{2n}} \delta_z \partial^l \mathbf{u}(x) \delta_z \partial^{m+1-l}\rho(x) \partial^{m+1}\mathbf{u}(x) \phi(z) \, dz \, dx + \frac{1}{2} \int_{\mathbb{T}^{2n}} \delta_z \partial^l \mathbf{u}(x) \partial^{m+1-l}\rho(x) \delta_z \partial^{m+1}\mathbf{u}(x) \phi(z) \, dz \, dx := J_1 + J_2.
\]
Estimates on the new terms, \(J_1, J_2\) are a little more sophisticated as we seek to optimize distribution of \(L^p\)-norms inside their components. Notice that the case \(l = 1\) corresponds to the previously appeared term \(J_{2,1}\).
So, let us assume that \(l = 1, \ldots, m\). We will distribute the parameters in \(J_1\) as follows
\[
J_1 = \int_{\mathbb{T}^{2n}} \frac{\delta_z \partial^l \mathbf{u}(x)}{|z|^{\frac{q}{2}+\frac{q}{4}+2\delta}} \frac{\delta_z \partial^{m+1-l}\rho(x)}{|z|^{\frac{q}{2}+\frac{q}{4}}} \frac{\partial^{m+1}\mathbf{u}(x)}{|z|^{\frac{q}{2}-\delta}} \frac{1}{|z|^{\frac{q}{2}-\delta}} \, dz \, dx,
\]
where $\delta > 0$ is a small parameter to be determined later, and $(2, p, q, r)$ is a Hölder quadruple defined by
\[ p = \frac{2m + \frac{\alpha}{2}}{l - 1 + \frac{\alpha}{2}}, \quad q = \frac{2m + \alpha - 1}{m - l + \frac{\alpha}{2}} \quad \frac{1}{r} = 1 - \frac{1}{2} - \frac{1}{p} - \frac{1}{q}. \]
The existence of finite $r$ is warranted by the strict inequality which is verified directly:
\[ \frac{1}{2} + \frac{1}{p} + \frac{1}{q} < 1. \]
By the Hölder inequality,
\[ J_1 \leq \|u\|_{W^{2m+\alpha,2}2,2} \|\rho\|_{W^{m+1-l,2},q} \|u\|_{H^{m+1}}. \]
Let us apply the following Gagliardo-Nirenberg inequalities to all the terms
\[ \|u\|_{H^{m+1}} \leq \|u\|_{H^{m+1,2}}^{\frac{2m}{2m+\alpha}} \|\nabla u\|_{H^{m+1,2}}^{\frac{\alpha}{2m+\alpha}} \leq \|u\|_{H^{m+1,2}}^{\theta_1} \|\nabla u\|_{H^{m+1,2}}^{1-\theta_1}. \]
\[ \|\rho\|_{H^{m+1-l,2}} \leq \|\rho\|_{H^{m+\alpha,1}} \|\nabla \rho\|_{H^{m+\alpha,1}}^{1-\theta_2}, \]
where
\[ \theta_1 = \frac{l - 1 + \frac{\alpha}{2} - \frac{n}{p} + 2\delta}{m + \frac{\alpha}{2} - \frac{n}{p}}, \quad \theta_2 = \frac{m - l + \frac{\alpha}{2} - \frac{n}{q}}{m + \alpha - 1 - \frac{n}{q}}. \]
The exponents satisfy the necessary requirements
\[ 1 \geq \theta_1 \geq \frac{l - 1 + \frac{\alpha}{2} + 2\delta}{m + \frac{\alpha}{2}}, \quad 1 \geq \theta_2 = \frac{m - l + \frac{\alpha}{2}}{m + \alpha - 1}, \]
and in fact,
\[ \theta_1 = \frac{l - 1 + \frac{\alpha}{2}}{m + \frac{\alpha}{2}} + O(\delta). \]
Now, we have
\[ J_1 \leq \|u\|_{H^{m+1,2}}^{2m+\alpha+\theta_1} \|\rho\|_{H^{m+\alpha,1}} \|\nabla u\|_{H^{m+\alpha,1}}^{\alpha + 1-\theta_1} \|\nabla \rho\|_{H^{m+\alpha,1}}^{1-\theta_2}. \]
By generalized Young,
\[ J_1 \leq \varepsilon \rho_- \|u\|_{H^{m+1+l,2}}^2 \rho_0^2 (\|\nabla \rho\|_{H^{m+\alpha,1}}^{\alpha + 1-\theta_1} \|\nabla \rho\|_{H^{m+\alpha,1}}^{1-\theta_2}) Q, \]
where $Q$ is the conjugate to $\frac{2m}{2m+\alpha} + \theta_1$. We have $\theta_2 Q < 2$ as long as
\[ (161) \quad \theta_1 + \theta_2 < 2 - \frac{2m}{2m+\alpha}. \]
We in fact have even stronger inequality, $\theta_1 + \theta_2 < 1$ provided $\delta$ is small enough. So, we arrived at
\[ J_1 \leq \varepsilon \rho_- \|u\|_{H^{m+1+l,2}}^2 + \rho_0^2 p_N (\|\nabla \rho\|_{H^{m+1+l,2}} \|\nabla u\|_{H^{m+1+l,2}} Y_m), \]
for some polynomial $p_N$.
Finally, moving on to $J_2$, we distribute the exponents as follows
\[ J_2 \leq \int_{T^{2n}} \frac{|\delta_x \partial^l u(x)|}{|z|^{\frac{n}{p} + 2\delta + \frac{\alpha}{2}}} \frac{|\delta_x \partial^{m+1} u(x)|}{|z|^{\frac{n}{q} + 2\delta + \frac{\alpha}{2}}} \frac{1}{|z|^{\frac{n}{p} - \delta}} dx \leq \|u\|_{W^{l+\delta,2},q} \|\rho\|_{W^{m+1-l,2}} \|u\|_{H^{m+1+\frac{\alpha}{2}}}. \]
Here we choose $(r, p, q, \delta)$ as follows
\[ q = \frac{2m + \alpha - 1}{m - l}, \quad p = \frac{2m + \frac{\alpha}{2}}{l - 1 + \frac{\alpha}{2}}, \quad \frac{1}{r} = 1 - \frac{1}{2} - \frac{1}{p} - \frac{1}{q}. \]
and $\delta$ is small. With these choices we proceed with the Gagliardo-Nirenberg inequalities

$$
\|u\|_{W^{l,2+2\delta+\frac{\alpha}{2},p}} \leq \|u\|_{F^{m+1+\frac{\alpha}{2},2}} \|\nabla u\|_{\infty}^{1-\theta_1},
$$

$$
\|\rho\|_{W^{m+1-\delta,2}} \leq \|\rho\|_{H^{m+\alpha}} \|\nabla \rho\|_{\infty}^{1-\theta_2},
$$

where

$$
\theta_1 = \frac{l - 1 + \frac{\alpha}{2} + 2\delta}{m + \frac{\alpha}{2} - \frac{n}{2}} = \frac{l - 1 + \frac{\alpha}{2}}{m + \frac{\alpha}{2}} + O(\delta), \quad \theta_2 = \frac{m - l}{m + \alpha - 1}.
$$

Now to achieve the bound

$$
J_2 \leq \varepsilon \rho_+ \|u\|_{H^{m+1+\alpha/2}}^2 + \rho_-^{-1} p_N(\|\nabla \rho\|_{\infty}, |\nabla u|_{\infty}) Y_m,
$$

we have to make sure that $\theta_1 + \theta_2 \leq 1$, which is true for small $\delta$.

We have proved the following a priori bound on $u$:

$$
\partial_t \|u\|_{H^{m+1}}^2 \leq -\frac{1}{2} \rho_- \|u\|_{H^{m+1+\alpha/2}}^2 + \rho_-^{-1} p_N(\|\nabla \rho\|_{\infty}, |\nabla u|_{\infty}) Y_m.
$$

Together with the previously established bounds we obtain

$$
\frac{d}{dt} Y_m \leq -\frac{1}{2} \rho_- \|u\|_{H^{m+1+\alpha/2}}^2 + \rho_-^{-1} p_N(\|\nabla \rho\|_{\infty}, |\nabla u|_{\infty}, |e|_{\infty}) Y_m.
$$

This of course implies a Riccati inequality, provided $m > \frac{n}{2} + 1$:

$$
\frac{d}{dt} Y_m \leq CY_m^N,
$$

and provides a priori bound independent of the viscosity coefficient. Thus, we can extend it to an interval independent of $\varepsilon$ as well. By the compactness argument similar to the smooth kernel case, we obtain a local solution in the same class as initial data and $u \in L^2 H^{m+\alpha/2}$. In addition, we obtain a continuation criterion – as long as $|\nabla \rho|_{\infty}, |\nabla u|_{\infty}, |e|_{\infty}$ remain bounded on $[0, T_0)$ the solutions can be extended beyond $T_0$. However everything is reduced to a control over the first two quantities, because $|e|_{\infty}$ remains bounded as long as $|\nabla u|_{\infty}$ in view of (158).

**Theorem 7.3** (Local existence of classical solutions). Suppose $m > \frac{n}{2} + 1$, and

$$(u_0, \rho_0) \in H^{m+1}(\mathbb{T}^n) \times H^{m+\alpha}(\mathbb{T}^n),$$

and $\rho_0(x) > 0$ for all $x \in \mathbb{T}^n$. Then there exists time $T_0 > 0$ and a unique non-vacuous solution to (148) on time interval $[0, T_0)$ in the class

$$(162) \quad u \in C_w([0, T_0); H^{m+1}) \cap L^2([0, T_0); H^{m+1+\alpha/2}), \quad \rho \in C_w([0, T_0); H^{m+\alpha}).$$

Moreover, any such solution satisfying

$$(163) \quad \sup_{t \in [0, T_0)} (|\nabla \rho(t)|_{\infty} + |\nabla u(t)|_{\infty}) < \infty$$

can be extended beyond $T_0$.

It is clear from the proof that (163) can be replaced with an integrability condition with some high power depending on $m, n, \alpha$. 

8. One-dimensional theory

In dimension 1, most complete regularity theory of alignment models is available. In this chapter we discuss some of its highlights.

In 1D the system is given by

\[
\begin{aligned}
\rho_t + (\rho u)_x &= 0, \\
u_t + uu_x &= \int_\Omega \phi(x-y)(u(y) - u(x))\rho(y) \, dy \\
(x,t) &\in \Omega \times \mathbb{R}_+.
\end{aligned}
\]  

(164)

Here \( \Omega \) is either \( \mathbb{R} \) or periodic circle \( \mathbb{T} \). The underlying theme in this chapter will be to relate regularity and flocking behavior of the system to the new conserved quantity which is available in dimension 1 and a few other exceptional multi-D cases:

\[
e = u_x + L_\phi \rho,
\]

(165)

where \( L_\phi \) takes form of either of the integral representations (105) or (106) depending on the context.

8.1. Threshold condition for global existence with smooth kernels. The system (164) possesses an extra conservation law provided \( \phi \) is a convolution kernel, \( \phi = \phi(x-y) \). To see this we note that the alignment term in the velocity equation is given by the commutator

\[
C_\phi(u,\rho) = \phi * (u\rho) - u \phi * \rho.
\]

Using this commutator structure and by elementary manipulation with the equations we obtain

\[
ea_t + (ue)_x = 0, \quad e = u_x + \phi * \rho.
\]

(166)

This quantity, whose physical role will be illuminated later in Section 8.2 plays a crucial role in regularity theory. Note that the Burgers’ part of the velocity equation tends to create shocks, while the alignment part would counterforce it to smooth out the solution. One would expect then that a threshold condition would separate singular behavior from regular. For pure Burgers equation such condition is provided by positive initial slope \( u_x \geq 0 \). In view of the regularizing additional effect of alignment such condition can be relaxed to \( e_0 \geq 0 \). To see that let us rewrite the \( e \)-equation as a non-autonomous logistic ODE along characteristics:

\[
\dot{e} = e(\phi * \rho - e)
\]

(167)

It is clear that the sign of \( e \) will be preserved pointwise. So, if \( e_0(x_0) < 0 \) at some point \( x_0 \), then \( \dot{e} < -e^2 \), and consequently the solution blows up in finite time. On the other hand, supposing \( e_0 \geq 0 \) we have \( e(t) \geq 0 \) for all times, and \( e \) remains a priori bounded since \( \phi * \rho \leq CM \). This in particular implies that \( u_x \in L^\infty_t \). In view of Theorem 7.1, this ensures global existence of solutions. Moreover, we can also obtain a hydrodynamic version of strong flocking. Let us make this precise.

**Theorem 8.1** (Threshold for global existence). Consider the system (164) on \( \mathbb{R} \) or \( \mathbb{T} \) with smooth kernel. For any initial condition \( (u_0,\rho_0) \in H^m \times (L^1_+ \cap W^{1,\infty}) \), \( m > \frac{n}{2} + 2 \), which satisfies the threshold condition \( e_0 \geq 0 \) there exists a unique global solution \( (u,\rho) \in C([0,\infty);H^m \times (L^1_+ \cap W^{1,\infty}) \). If \( e_0(x) < 0 \) at some point, then the solution blows up in finite time.

Furthermore, suppose that \( \phi \) has a fat tail. Let the initial flock has compact support \( \text{Supp} \rho_0 \) and \( e_0 \geq 0 \). Then the solution flock strongly in the following sense: there exists \( C,\delta > 0 \) depending on \( \phi \), and initial data, such that the velocity satisfies

\[
\sup_{x \in \text{Supp} \rho(t)} |u(x,t) - \bar{u}| + |u_x(x,t)| + |u_{xx}(x,t)| \leq Ce^{-\delta t}.
\]

(168)

and the density \( \rho \) converges to a traveling wave \( \bar{\rho} \) in the metric of \( C^\gamma \) for any \( 0 < \gamma < 1 \):

\[
\|\rho(\cdot,t) - \bar{\rho}(\cdot - \bar{u}t)\|_{C^\gamma} \leq Ce^{-\delta t}, \quad t > 0.
\]

(169)
Proof. The global existence has already been proved in the preceding remarks.

Let us now assume that we have a global solution with \( e_0 \geq 0 \) and \( \phi \) has a fat tail. We know from Theorem 6.1 that the diameter of the flock will remain finite, \( \bar{D} \). Then we can estimate the convolution from below: for any \( x \in \text{Supp} \rho(t) \):

\[
\phi \ast \rho(t, x) \geq \phi(D)M := c_0.
\]

Solving the logistic ODE: \( \dot{e} \geq e(c_0 - e) \) we find that starting from some time \( t_0 \) for all \( t > t_0 \) and \( x \in \text{Supp} \rho(t) \) we have \( e(x, t) \geq c_0/2 \). With this in mind let us write the equation for \( u_x \) as follows

\[
\frac{D}{Dt}u_x = \int_{\mathbb{R}} \phi'(x-y)(u(y) - u(x))\rho(y) \, dy - e(x)u_x(x).
\]

We already know from Theorem 6.1 that the velocity fluctuations are decaying with exponential rate. Hence, the integral above will be bounded by \( |\phi'|_\infty ME(t) \), where we denote by \( E \) any quantity that shows exponential decay. Thus, multiplying (170) with \( u_x \) and evaluating at the maximum over \( \text{Supp} \rho_0 \) we obtain

\[
\frac{d}{dt}\|u_x\|_{L^\infty(\text{Supp} \rho(t))} = E(t) - \frac{1}{2}c_0\|u_x\|_{L^\infty(\text{Supp} \rho(t))}.
\]

This implies the desired result by integration. In view of (143) the density enjoys a pointwise global bound

\[
\sup_{t>0} \|\rho(\cdot, t)\|_\infty = R < \infty.
\]

For the second derivative, we have

\[
\frac{D}{Dt}u_{xx} + 2u_x u_{xx} = \int_{\mathbb{R}} \phi''(x-y)(u(y) - u(x))\rho(y) \, dy - 2u_x \phi' \ast \rho - \epsilon u_{xx}.
\]

We see that the integral term as well as \( u_x \) is of type \( E(t) \) in view of the previously established bounds. Note also that \( |\phi_x \ast \rho| \leq |\phi_x|_\infty M \). So, we obtain

\[
\frac{d}{dt}\|u_{xx}\|_{L^\infty(\text{Supp} \rho(t))} = E(t) - \frac{1}{2}c_0\|u_{xx}\|_{L^\infty(\text{Supp} \rho(t))}.
\]

This implies exponential decay. Moving to the density, we have

\[
\partial_t \rho_x + u_x \rho_{xx} = -2u_x \rho_x - u_{xx} \rho = E\rho_x + E.
\]

This shows that \( \rho_x \) remains uniformly bounded. Now, to establish strong flocking we pass to the moving frame \( x - \bar{u}t \) and write the continuity equation in new coordinates

\[
\frac{D}{Dt}\rho = -(u - \bar{u})\rho_x - u_x \rho = E.
\]

This shows that \( \rho(t) \) is Cauchy in \( t \) in the metric of \( L^\infty \). Hence, there exists \( \bar{\rho} \in L^\infty \) such that \( \|\rho(t) - \bar{\rho}\|_\infty = E(t) \). Since \( \rho' \) is uniformly bounded, this also shows that \( \bar{\rho} \) is Lipschitz. Convergence in \( C^\gamma \), \( \gamma < 1 \), follows by interpolation. \( \square \)

8.2. Entropy. Csiszár-Kullback inequality. Shape of the limiting Flock. Although in the case of symmetric communication the limiting velocity is determined from the initial condition by use of momentum conservation, the limiting shape of the density profile \( \bar{\rho} \) is an emergent quantity which is not known a priori. Yet it is easy to notice that the \( e \)-quantity is somehow involved in determining \( \bar{\rho} \). Let us assume in this section that \( e = u_x + \mathcal{L}_\phi \rho \), where \( \mathcal{L}_\phi \) is defined by (106) in either smooth or singular case. Let us write the continuity equation with the use of \( e \):

\[
\rho_t + u \rho_x + e \rho = \rho \mathcal{L}_\phi \rho.
\]
Let us assume for a moment that $\phi$ is a smooth absolute kernel on $\mathbb{T}$. Suppose that $e_0 = 0$, and hence $e(t) = 0$ for all time. Note that in this case $u_x + \phi \ast \rho \geq \rho(x)|\phi|_1 \geq 0$, so the global solution exists. Then

$$\rho_t + u_x \rho = \rho \mathcal{L}_\phi \rho,$$

and so, $\rho$ obeys the maximum principle. Let us assume that there is no vacuum $\rho_-(0) > 0$. Let us then write the equation for the new quantity $\ln \rho$:

$$(\ln \rho)_t + u(\ln \rho)_x = \mathcal{L}_\phi \rho.$$

Evaluating at a point of minimum we obtain

$$(\ln \rho_-)_t = \int \phi(x,y)(\rho(y) - \rho_-) \, dy \geq c_0(M - 2\pi \rho_-),$$

and at the maximum

$$(\ln \rho_+)_t = \int \phi(x,y)(\rho(y) - \rho_+) \, dy \leq c_0(M - 2\pi \rho_+).$$

Subtracting the two we obtain

$$\frac{d}{dt} \ln \frac{\rho_+}{\rho_-} \leq -2\pi c_0(\rho_+ - \rho_-) \leq -2\pi c_0 \rho_-(0) \left( \frac{\rho_+}{\rho_-} - 1 \right) \leq -2\pi c_0 \rho_-(0) \ln \frac{\rho_+}{\rho_-}.$$

We conclude that

$$\ln \frac{\rho_+}{\rho_-} \leq c_1 e^{-c_2 t}.$$

Since the maximum also stays bounded we have inequality

$$\ln \frac{\rho_+}{\rho_-} \geq c \left( \frac{\rho_+}{\rho_-} - 1 \right),$$

So, we conclude that the density flattens out exponentially fast to a uniformly distributed state $\bar{\rho} = \frac{1}{2\pi} M$.

In view of this computation we can see that $e$ is directly responsible for the flattening of the density. It turns out that, first, a similar result is true for local kernels and even vacuous solutions. And second, in general the size of $e$ per mass, i.e. the quotient $q = \frac{e}{\rho}$, measures how far $\bar{\rho}$ is from the uniform distribution. Thus, $e$ plays the role of a topological entropy of the flock – a measure of disorder. We will address this interpretation in the next theorem.

**Theorem 8.2.** Let $(\rho, u)$ be a smooth solution to the system (164) on the 1D torus $\mathbb{T}$, and $\phi$ is a smooth local kernel:

$$\phi(r) \geq \lambda_1 r < r_0.$$

If $e_0 = 0$, then

$$\|\rho(t) - \bar{\rho}\|_1 \leq c_1(\|\rho_0\|_2) e^{-c_2(\lambda, r_0, M, \|\rho_0\|_\infty)t},$$

where $\bar{\rho} = \frac{1}{2\pi} M$.

In general, provided $\|q_0\|_\infty < \|\phi\|_1$, one has

$$\limsup_{t \to \infty} \|\rho(\cdot, t) - \bar{\rho}\|_1 \leq \frac{M\|q_0\|_\infty\|\phi\|_\infty}{\lambda c(r_0)(\|\phi\|_1 - \|q_0\|_\infty)}.$$

Let us note that the dependence on $\|q_0\|_\infty$ is linear for small values. At the same time, the bound is inversely proportional to the strength $\lambda$, which shows the stabilizing effect of communication on the structure of the flock. Let us note that $q$ satisfies the transport equation

$$\frac{d}{dt} q + u q_x = 0, \quad q = \frac{e}{\rho},$$

and hence the value of $\|q\|_{L^\infty}$ is preserved for all time.
The proof will be split in the following few subsections. First, let us recall a very useful tool – the Csiszár-Kullback inequality.

8.2.1. The Csiszár-Kullback inequality. The main object to study will actually be not $L^1$-norm, but rather the relative entropy defined by

$$ H = \int_\mathbb{T} \rho \log \frac{\rho}{\bar{\rho}} \, dx = \int_\mathbb{T} \rho \log \rho \, dx - M \log \bar{\rho}, $$

(180)

The classical Csiszár-Kullback inequality states

$$ \|\rho - \bar{\rho}\|^2_{L^1} \leq 4\pi \bar{\rho} H. $$

(181)

Furthermore, by the elementary inequality $\log x \leq x - 1$, we also have

$$ H \leq \int_\mathbb{T} \rho \left( \frac{\rho}{\bar{\rho}} - 1 \right) \, dx = \rho^{-1} \|\rho - \bar{\rho}\|^2_{L^2}. $$

(182)

Thus we obtain the two sided bounds

$$ \frac{1}{4\pi} \|\rho - \bar{\rho}\|^2_{L^1} \leq \rho H \leq \|\rho - \bar{\rho}\|^2_{L^2}. $$

(183)

8.2.2. Evolution of the entropy. At the heart of the argument is the equation for the entropy (180) which one obtains testing the continuity equation with $\log \rho + 1$:

$$ (\rho \log \rho)_t = \rho_t (\log \rho + 1) = -(\rho u)' (\log \rho + 1) $$

$$ = -\rho' (\log \rho + 1) u - \rho u' (\log \rho + 1) $$

$$ = -\rho \rho' + (\rho \log \rho)' u - \rho u' $$

$$ = -[u(\rho \log \rho)]' - \rho u' = -[u(\rho \log \rho)]' - \rho^2 q + \mathcal{L}_\psi \rho. $$

Therefore,

$$ \frac{dH}{dt} = \frac{d}{dt} \int_\mathbb{T} \rho \log \rho \, dx = - \int_\mathbb{T} \rho^2 q \, dx - \int_{\mathbb{T}^2} \phi(x-y)(\rho(x) - \rho(y)) \rho(x) \, dx \, dy. $$

(184)

Noting that $\int_\mathbb{T} \rho q \, dx = \int_\mathbb{T} e \, dx = 0$, we can subtract $\bar{\rho}$ from one density in the first integral on the left hand side. After additionally symmetrizing the last integral we obtain

$$ \frac{dH}{dt} = - \int_\mathbb{T} (\rho - \bar{\rho}) \rho q \, dx - \frac{1}{2} \int_{\mathbb{T}^2} \phi(x-y) |\rho(x) - \rho(y)|^2 \, dx \, dy. $$

(185)

8.2.3. Bounds on the dissipation. If our kernel was absolute, it would be easy to get a positive lower bound on the dissipation term:

$$ \int_{\mathbb{T}^2} \phi(x-y)|\rho(x) - \rho(y)|^2 \, dx \, dy \geq (\inf \phi) \int_{\mathbb{T}^2} |\rho(x) - \rho(y)|^2 \, dx \, dy = 2(\inf \phi) \|\rho - \bar{\rho}\|^2_{L^2}. $$

(186)

Since we have a non-trivial lower bound on the kernel only near the diagonal $\{ (x, y) \in \mathbb{T}^2 : |x-y| < r_0 \}$, we need a substitute for (186) stated in the following Lemma.

**Lemma 8.3.** The following inequality holds:

$$ \int_{|x-y| < r_0} |\rho(x) - \rho(y)|^2 \, dy \, dx \geq c(r_0) \|\rho - \bar{\rho}\|^2_{L^2}. $$

(187)

**Proof.** Let $\chi$ be a nonnegative cutoff function on $\mathbb{T}$ with support in $B_{r_0}(0)$, which is constant on $B_{r_0/2}$ and has integral 1. Then $\chi(0) = 1$ and $|\hat{\chi}(k)| < 1$ for all $k \in \mathbb{Z} \setminus \{0\}$. On the other hand, we have by the Riemann-Lebesgue Lemma that $\hat{\chi}(k) \to 0$ as $k \to \infty$; therefore we have in fact that $|\hat{\chi}(k)| \leq 1 - \varepsilon$ for some $\varepsilon > 0$ depending only on $r_0$ ($k \neq 0$). Define $\rho_0(x) = \chi * \rho(x)$, so that

$$ (\rho - \bar{\rho}_0)(k) = (1 - \hat{\chi}(k))\bar{\rho}(k). $$
consequently \(|(\rho - \hat{\rho}_0)(k)| \geq \varepsilon|\hat{\rho}(k)|, \quad k \in \mathbb{Z}, \ k \neq 0\), and \(\hat{\rho}(0) = \hat{\rho}_0(0)\). Thus
\[
\|\rho - \hat{\rho}\|^2 = \sum_{k \in \mathbb{Z} \setminus \{0\}} |\hat{\rho}(k)|^2 \leq \varepsilon^{-2} \sum_{k \in \mathbb{Z}} |(\rho - \hat{\rho}_0)(k)|^2 = \varepsilon^{-2}\|\rho - \hat{\rho}_0\|^2_{L^2}.
\]
By the fact that \(\int_T \chi = 1\) and Minkowski, we have
\[
\|\rho - \hat{\rho}_0\|^2_{L^2} = \left\| \int_T \chi(y)(\rho(\cdot) - \rho(\cdot - y))\,dy \right\|_{L^2} \leq \int_{|y|<r_0} \|\rho(\cdot) - \rho(\cdot - y)\|^2_{L^2} \,dy
= \int_T \int_{|z|<r_0} |\rho(x) - \rho(x + z)|^2 \,dz\,dx.
\]
Combining the inequalities above yields
\[
\|\rho - \hat{\rho}\|^2 \leq \varepsilon^{-2} \int_T \int_{|z|<r_0} |\rho(x) - \rho(x + z)|^2 \,dz\,dx;
\]
taking \(c(r_0) = \varepsilon^2\) finishes the proof.

Estimating the dissipation term now becomes trivial:
\[(188) \quad \frac{1}{2} \int_{\mathbb{R}^2} \phi(x-y)|\rho(x) - \rho(y)|^2 \,dx \,dy \geq c\|\rho - \hat{\rho}\|^2_{L^2}.
\]

8.2.4. The entropy equation revisited. We now return to the evolution equation (185) and make estimates; in doing so we will prove Theorem 8.2. We have
\[
\dot{\mathcal{H}}(t) \leq \|\rho(t)\|_{L^\infty} \|\dot{q}_0\|_{L^\infty} \|\rho(\cdot, t) - \hat{\rho}\|_{L^1} - c\|\rho(\cdot, t) - \hat{\rho}\|^2_{L^2}
\leq \|\rho(t)\|_{L^\infty} \|\dot{q}_0\|_{L^\infty} \sqrt{4\pi \hat{\rho}\mathcal{H}(t)} - c\hat{\rho}\mathcal{H}(t).
\]
Setting \(Y = \sqrt{\mathcal{H}}\), we find
\[
\dot{Y}(t) \leq \|\rho(t)\|_{L^\infty} \|\dot{q}_0\|_{L^\infty} \sqrt{4\pi \hat{\rho}} - c\hat{\rho}Y(t).
\]
Using Grönwall’s lemma we obtain
\[(189) \quad Y(t) \leq Y_0 e^{-c\hat{\rho}t} + \sqrt{\pi \hat{\rho}} \|\dot{q}_0\|_{L^\infty} \int_0^t \|\rho(s)\|_{L^\infty} e^{-c\hat{\rho}(t-s)} \,ds.
\]
From here, it is easy to prove Theorem 8.2. Indeed, if \(e_0 \equiv 0\), then the second term in (189) drops out completely. Using (183) proves this part of the theorem.

For the general \(\varepsilon\), we have from the above
\[(190) \quad \lim_{t \to \infty} \sup_{t \to \infty} \|\rho(\cdot, t) - \hat{\rho}\|_{L^1} \leq M\|\dot{q}_0\|_{L^\infty} \lim_{t \to \infty} \int_0^t \|\rho(s)\|_{L^\infty} e^{-c\hat{\rho}(t-s)} \,ds
\]
\[
\leq \frac{\|\dot{q}_0\|_{L^\infty}}{\lambda c(r_0)} \lim_{t \to \infty} \sup_{t \to \infty} \|\rho(t)\|_{L^\infty}.
\]
Thus the proof of Theorem 8.2 is reduced to estimating the density amplitude.
8.2.5. Bounds on the Density Amplitude. Throughout our discussion of bounds on the density amplitude, we will make use of the following differential inequality: if $\dot{X}(t) \leq A X(t) [B - X(t)]$, where $A$ and $B$ are positive constants and $X(t)$ is a positive function, then

$$X(t) \leq \frac{BX(0)}{X(0) + (B - X(0)) \exp(-ABt)}.$$  

In particular, $\limsup_{t \to \infty} X(t) \leq B$.

Let $\rho_+(t)$ denote the maximum value of $\rho$ at time $t$, and let $x_+$ denote the $x$-value where the maximum is achieved. Then if $\|q_0\|_{L^\infty} < \|\phi\|_{L^1}$, one can get an upper bound on $\|\rho(t)\|_{L^\infty}$ by integrating the differential inequality derived below.

$$\frac{d}{dt} \rho_+(t) = -\rho_+(t) u'(x_+, t) = -\rho_+(t)^2 q(x_+, t) + \rho_+(t) \int_T \phi(x_+ - y)(\rho(y, t) - \rho_+(t)) dy,$$

$$\leq (\|q_0\|_{L^\infty} - \|\phi\|_{L^1}) \rho_+(t)^2 + M \|\phi\|_{L^\infty} \rho_+(t),$$

$$= (\|\phi\|_{L^1} - \|q_0\|_{L^\infty}) \rho_+(t) \left[ \frac{M \|\phi\|_{L^\infty}}{\|\phi\|_{L^1} - \|q_0\|_{L^\infty}} - \rho_+(t) \right].$$

In view of (191) we obtain

$$\limsup_{t \to \infty} \|\rho(t)\|_{L^\infty} \leq \frac{M \|\phi\|_{L^\infty}}{\|\phi\|_{L^1} - \|q_0\|_{L^\infty}}.$$

Plugging into (190) we conclude

$$\limsup_{t \to \infty} \|\rho(\cdot, t) - \bar{\rho}\|_{L^1} \leq \frac{M \|q_0\|_{L^\infty} \|\phi\|_{L^\infty}}{\lambda c(r_0)(\|\phi\|_{L^1} - \|q_0\|_{L^\infty})}.$$

8.3. Alignment on $\mathbb{T}$ with degenerate kernel. The corrector method. On the periodic circle, there is a mechanism for alignment with local communication even for vacuous solutions. It is clear already on the level of particle dynamics – if two agents have not yet aligned and are past their communication range, they would meet at the opposite end of the circle and reestablish communication. The alignment can be established in this case with an adaptation of the corrector method we introduced in Section 2.5.

**Theorem 8.4.** For any solution of either the discrete or hydrodynamic system on $\mathbb{T}$ the following holds

(i) For sub-quadratic communication

$$\lambda 1_{r < r_0} \leq \phi(r) \leq \frac{\Lambda}{r^2},$$

one has

$$V_2(t) \leq C \frac{\ln t}{t},$$

as $t \to \infty$, where $C$ depends only on the initial condition.

(ii) If the kernel satisfies the more singular assumption

$$1_{r < r_0} \frac{\lambda}{r^\beta} \leq \phi(r) \leq \frac{\Lambda}{r^\beta}, \quad \beta > 2$$

then we can only conclude $V_2(t) \to 0$, $t \to \infty$.

Let us note that this result does not improve the rate in the topological models since the variations of the density preclude us from making assumption (192).
Proof. The proof will be carried out in the discrete case only since the continuous version is entirely similar. We will limit our selves to providing necessary modifications in that case.

We go back to the basic energy law (40) and construct a corrector functional $G$ which serves to compensate for the missing interactions. To do that we first define a periodic analogue of the directed distance:

$$d_{ij}(t) = -x_{ij} \text{sgn}(v_{ij}) \mod 2\pi,$$

where $x_i, x_j \in [0, 2\pi)$ are viewed on the same coordinate chart. The distance picks up the length of the arch between $x_i$ and $x_j$ which contracts under the evolution of the agents. The distance undergoes jump discontinuities at $x_i = x_j$ and $v_i = v_j$. At any other point, we have

$$d_{ij}(t) = -|v_{ij}|.$$

Next, we define a slope kernel $\psi \geq 0$ as follows (see Figure 6)

$$\psi(x) = \begin{cases} 
-x + r_0, & -r_0 \leq x \leq r_0 \\
\frac{r_0}{\pi - r_0}x - \frac{r_0^2}{\pi - r_0}, & r_0 < x < 2\pi - r_0,
\end{cases}$$

extended periodically on $\mathbb{R}$. Finally, we define the corrector

$$G(t) = \frac{1}{N^2} \sum_{i,j=1}^{N} |v_{ij}| \psi(d_{ij}).$$

Let us look into differentiability of $G$:

$$\frac{d}{dt}G = -\frac{1}{N^2} \sum_{i,j=1}^{N} |v_{ij}|^2 \psi'(d_{ij}) + \frac{2}{N^3} \sum_{i,j=1}^{N} \psi(d_{ij}) \text{sgn}(v_{ij}) \sum_{k=1}^{N} v_{ki} \phi_{ki}.$$

The formula can be justified classically, at those times when there is no jump, i.e. $x_i \neq x_j$ and $v_i \neq v_j$, due to (195). When two agents pass each other $x_i = x_j$ we use periodicity of $\psi$, and when $v_{ij} = 0$, the factor $|v_{ij}|$ vanishes.

We continue

$$\frac{d}{dt}G(t) = \frac{1}{N^2} \sum_{i,j=1}^{N} |v_{ij}|^2 I_{|x_{ij}| \leq r_0} - \frac{r_0}{\pi - r_0} \frac{1}{N^2} \sum_{i,j=1}^{N} |v_{ij}|^2 I_{|x_{ij}| \geq r_0} + \mathcal{R},$$

where

$$\mathcal{R} = \frac{2}{N^3} \sum_{i,j,k=1}^{N} \psi(d_{ij}) \text{sgn}(v_{ij}) v_{ki} \phi_{ki}.$$
Symmetrizing over \(i, k\), we obtain
\[
\mathcal{R} = \frac{1}{N^3} \sum_{i,j,k=1}^{N} (\psi(d_{ij}) \text{sgn}(v_{ij}) - \psi(d_{kj}) \text{sgn}(v_{kj})) v_{ki} \phi_{ki}.
\]

In the case \(v_i \geq v_j \geq v_k\) or \(v_i \leq v_j \leq v_k\), the summand is negative, and so we can neglect it. Continuing,
\[
\mathcal{R} \leq \frac{1}{N^3} \sum_{i,j,k=1}^{N} \left( \psi(d_{ij}) \text{sgn}(v_{ij}) - \psi(d_{kj}) \text{sgn}(v_{kj}) \right) v_{ki} \phi_{ki} \mathbb{1}_{v_j > \max(v_i,v_k)}
\]
\[
+ \frac{1}{N^3} \sum_{i,j,k=1}^{N} \left( \psi(d_{ij}) - \psi(d_{kj}) \right) v_{ki} \phi_{ki} \mathbb{1}_{v_j < \min(v_i,v_k)}
\]
\[
= \frac{1}{N^3} \sum_{i,j,k=1}^{N} \left( \psi(d_{kj}) - \psi(d_{ij}) \right) v_{ki} \phi_{ki} \mathbb{1}_{v_j > \max(v_i,v_k)}
\]
\[
+ \frac{1}{N^3} \sum_{i,j,k=1}^{N} \left( \psi(d_{ij}) - \psi(d_{kj}) \right) v_{ki} \phi_{ki} \mathbb{1}_{v_j < \min(v_i,v_k)}.
\]

In the cases \(v_j > \max(v_i,v_k)\) and \(v_j < \min(v_i,v_k)\), we see that \(v_i - v_j\) and \(v_k - v_j\) have the same sign so \(d_{ij}\) and \(d_{kj}\) are computed in the same direction. So, by the Lipschitz continuity of \(\psi\) and by the triangle inequality we find that \(|\psi(d_{ij}) - \psi(d_{kj})| \leq C|x_i - x_k|\). Therefore,
\[
\mathcal{R} \leq \frac{C}{N^3} \sum_{i,j,k=1}^{N} |x_{ik}| v_{ik} \phi_{ki} = \frac{C}{N^2} \sum_{i,k=1}^{N} |x_{ik}| v_{ik} \phi_{ki} \leq \frac{1}{t} \frac{C}{bN^2} \sum_{i,k=1}^{N} |x_{ik}|^2 \phi_{ki} + t b \frac{N^2}{N^2} \sum_{i,k=1}^{N} v_{ik}^2 \phi_{ki}.
\]

Let us proceed now under the assumption of (i). Here we obtain
\[
\mathcal{R} \leq \frac{c}{t} + bt \mathcal{I}_2.
\]

Then the corrector equation becomes
\[
\frac{d}{dt} G(t) \leq a \mathcal{I}_2 + b(t \mathcal{I}_2 - \mathcal{V}_2) + \frac{c}{t}.
\]

Let us form another functional:
\[
\mathcal{L} = G + bt \mathcal{V}_2 + a \mathcal{V}_2.
\]

It satisfies inequality \(\frac{d}{dt} \mathcal{L} \leq \frac{c}{t}\). Thus, \(\mathcal{L}(t) \leq \ln t\), and the resulting bound follows.

For part (ii) we use the collision potential (34) with precomputed bound in (35),
\[
\mathcal{R} \leq C \left( \frac{1}{N^2} \sum_{i,k=1}^{N} |x_{ik}|^2 - \beta |x_{ik}| \beta \phi_{ki} \right)^{1/2} \sqrt{\mathcal{I}_2} \leq \sqrt{\mathcal{I}_2} \sqrt{\mathcal{C}} \leq c_1 \sqrt{\mathcal{I}_2(t)} + c_2 \sqrt{\mathcal{I}_2(t)} \int_0^t \sqrt{\mathcal{I}_2(s)} \, ds.
\]

We can replace by the generalized Young inequality,
\[
c_1 \sqrt{\mathcal{I}_2(t)} \leq \frac{c_3}{t} + bt \mathcal{I}_2(t),
\]

and obtain
\[
\frac{d}{dt} G(t) \leq a \mathcal{I}_2 + b(t \mathcal{I}_2 - \mathcal{V}_2) + \frac{c_3}{t} + c_2 \sqrt{\mathcal{I}_2(t)} \int_0^t \sqrt{\mathcal{I}_2(s)} \, ds.
\]
With \( L \) being defined as before, we continue
\[
\frac{d}{dt} L \lesssim \frac{c_3}{t} + c_2 \sqrt{I_2(t)} \int_0^t \sqrt{I_2(s)} \, ds.
\]
Integrating over \([0, T]\),
\[
L(T) \lesssim L(0) + \ln T + \left( \int_0^T \sqrt{I_2(s)} \, ds \right)^2.
\]
Thus,
\[
\mathcal{V}_2(T) \leq \frac{1}{T} L(T) \lesssim \frac{\ln T}{T} + \frac{1}{T} \left( \int_0^T \sqrt{I_2(s)} \, ds \right)^2.
\]
The right hand side tends to zero which can be readily seen by splitting the integral into \((0, T')\) and \((T', T)\), where \( T' \) is large.

We already noted that the hydrodynamic version of the result is identical. Let us make some remarks about the proof. We work with the Lagrangian formulation (112). As in the discrete case we define the directed distance
\[
d_{\alpha \beta}(t) = (x(\alpha, t) - x(\beta, t)) \, \text{sgn}(v(\beta, t) - v(\alpha, t)) \mod 2\pi,
\]
and the corrector with \( \psi \) as before:
\[
\mathcal{G} = \int_{T^2} |u_{\alpha \beta}| \psi(d_{\alpha \beta}) \, dm_0(\alpha, \beta).
\]
We calculate the derivative of \( \mathcal{G} \):
\[
\frac{d}{dt} \mathcal{G} = -\int_{T^2} |u_{\alpha \beta}|^2 \psi'(d_{\alpha \beta}) \, dm_0(\alpha, \beta) + \int_{T^3} \text{sgn}(u_{\alpha \beta}) \psi(d_{\alpha \beta}) \phi_{\alpha \gamma} u_{\gamma \alpha} \, dm_0(\alpha, \beta, \gamma) 
\leq a I_2 - b \mathcal{V}_2 + \mathcal{R}.
\]
Here,
\[
\mathcal{R} = \int_{T^3} \text{sgn}(u_{\alpha \beta}) \psi(d_{\alpha \beta}) \phi_{\alpha \gamma} u_{\gamma \alpha} \, dm_0(\alpha, \beta, \gamma)
= \frac{1}{2} \int_{T^3} \left[ \text{sgn}(u_{\alpha \beta}) \psi(d_{\alpha \beta}) - \text{sgn}(u_{\gamma \beta}) \psi(d_{\gamma \beta}) \right] \phi_{\alpha \gamma} u_{\gamma \alpha} \, dm_0(\alpha, \beta, \gamma)
\leq \int_{T^3} (\psi(d_{\alpha \beta}) - \psi(d_{\gamma \beta})) \phi_{\alpha \gamma} u_{\gamma \alpha} \mathbb{I}_{u(\beta) < \min\{u(\alpha), u(\gamma)\}} \, dm_0(\alpha, \beta, \gamma)
+ \int_{T^3} (\psi(d_{\gamma \beta}) - \psi(d_{\alpha \beta})) \phi_{\alpha \gamma} u_{\gamma \alpha} \mathbb{I}_{u(\beta) > \max\{u(\alpha), u(\gamma)\}} \, dm_0(\alpha, \beta, \gamma)
\leq \int_{T^2} |x_{\alpha \gamma}| |u_{\gamma \alpha}| \phi_{\alpha \gamma} \, dm_0(\alpha, \gamma).
\]
In case (i) we obtain
\[
\mathcal{R} \leq c \frac{t}{T} + bt I_2,
\]
and the proof concludes as in the agent-based settings. In case (ii) we consider the collisional potential
\[
\mathcal{C} = \int_{T^2} \frac{dm_0(\alpha, \beta)}{|x_{\alpha \beta} \wedge r_0|^{\beta-2}}.
\]
It is well-posed for \( \beta < 3 \) (in view of also the fact that the density is bounded for regular solutions). A similar computation establishes (35), and from this point on the proof is ad verbatim. \( \square \)
In hydrodynamic settings the $L^2$-based alignment result does not provide sufficient information for pointwise behavior. So, it is desirable to obtain $L^\infty$-based alignment statement in this context. The mechanism for such alignment comes from considering regions where the density is non-negligible – here the alignment term works faster than the transport to avoid agent collisions. At the same time if density is thin the equation acts as classical Burgers’ equation. So, in order to avoid a blowup it must have low velocity fluctuations. In other words, it has to be aligned sufficiently well.

**Theorem 8.5.** Consider the system (164) on $\mathbb{T}$ with a smooth non-trivial non-negative kernel. Then any global classical solution aligns:

\[
\sup_x |u(x,t) - \bar{u}| \leq C \left( \frac{\ln t}{t} \right)^{\frac{1}{3}}.
\]

**Proof.** By the Galilean invariance we can assume throughout that $\bar{u} = 0$. As a consequence of the energy equality and (193) we obtain

\[
\int_T^\infty \int_{\mathbb{T}^2} \phi_{\alpha\beta} |v(\alpha,t) - v(\beta,t)|^2 \, dm_0(\alpha,\beta) \, dt \leq C \frac{\ln T}{T} := \varepsilon.
\]

Here we passed to Lagrangian coordinates $v(\alpha,t) = u(x(\alpha,t),t)$. Denote

\[
F(\alpha,T) = \int_T^\infty \int_{\mathbb{T}} \phi_{\alpha\beta} |v(\alpha,t) - v(\beta,t)|^2 \, dm_0(\beta) \, dt.
\]

So, we have

\[
\int_T^\infty F(\alpha, T) \, dm_0(\alpha) \leq \varepsilon.
\]

Let us fix another small parameter $\delta > 0$ and define the “good set”:

\[
G_\delta(T) = \{ \alpha : F(\alpha, T) \leq \delta \}.
\]

We denote by $G^c_\delta$ the complement of $G_\delta$, so that $m_0(G^c_\delta) = M - m_0(G_\delta)$ (recall that $M$ is the total mass of the flock). By the Chebychev inequality,

\[
m_0(G^c_\delta) < \frac{\varepsilon}{\delta}.
\]

Thus, the good set occupies almost all of the domain provided $\varepsilon \ll \delta$. We now proceed by proving that alignment occurs first on the good set identified above, and then on the rest of the torus later in time within a controlled time scale.

**Lemma 8.6** (Alignment on the good set). We have

\[
\sup_{\alpha_1, \alpha_2 \in G_\delta(t), t \geq T} |v(\alpha_1,t) - v(\alpha_2,t)| \lesssim \delta^{2/3}.
\]

**Proof.** Note that it suffices to establish alignment at time $T$ only. This simply follows from monotonicity of the our $F$-function:

\[
F(\alpha, t) \leq F(\alpha, T), \quad t > T,
\]

which implies that the good sets are increasing in time, $G_\delta(T) \subset G_\delta(t)$.

Integrating the equation

\[
\frac{d}{dt} v(\alpha,t) = \int_{\mathbb{T}} \phi_{\alpha\beta} v_{\alpha\beta} \, dm_0(\beta)
\]

over $[T,t]$ for any $\alpha \in G_\delta$ we obtain

\[
|v(\alpha,t) - v(\alpha,T)| \lesssim \int_T^t \int_{\mathbb{T}} |\phi_{\alpha\beta} v_{\alpha\beta}| \, dm_0(\beta) \lesssim \delta \sqrt{t - T}.
\]
Suppose that for some \( \alpha_1, \alpha_2 \in G_\delta \) we have
\[
v(\alpha_1, T) - v(\alpha_2, T) > U,
\]
where \( U \) to be determined later. Then in view of (199),
\[
v(\alpha_1, t) - v(\alpha_2, t) > \frac{U}{2},
\]
as long as
\[
t - T \lesssim \frac{U^2}{\delta^2}.
\]
During this time the corresponding characteristics will undergo a significant relative displacement
\[
x(\alpha_1, t) - x(\alpha_2, t) \geq x(\alpha_1, T) - x(\alpha_2, T) + \frac{1}{2}U(t - T) \mod 2\pi,
\]
where \( \frac{1}{2}U(t - T) > 4\pi \) as long as \( t - T \gtrsim \frac{1}{T} \). If this is allowed to happen, then the characteristics will find themselves at the separation distance equal \( 2\pi = 0 \), so they collapse. We necessarily obtain
\[
\frac{1}{U} \gtrsim \frac{U^2}{\delta^2},
\]
which gives \( U \lesssim \delta^{2/3} \) as claimed. \( \square \)

Next step is to show that the solution aligns completely at a not too distant later time \( t > T \).

**Lemma 8.7** (Alignment outside the good set). For all \( t \gtrsim T + \frac{1}{\delta^{2/3} + (\varepsilon/\delta)^{1/2}} \) we have
\[
\sup_{\alpha \in \mathbb{T}, \gamma \in G_\delta(T)} |v(\alpha, t) - v(\gamma, t)| \lesssim \delta^{1/3} + (\varepsilon/\delta)^{1/2}.
\]

**Proof.** Let \( \alpha \in \mathbb{T} \) and \( \gamma \in G_\delta(T) \). Let us write the momentum equation in Lagrangian coordinates as follows
\[
\frac{d}{dt} v(\alpha, t) = \int_\mathbb{T} v_{\beta\alpha} \phi_{\alpha\beta} \, dm_0(\beta) = \int_\mathbb{T} (v_{\beta\gamma} + v_{\gamma\alpha}) \phi_{\alpha\beta} \, dm_0(\beta)
\]
\[
= (\phi * \rho)(x(\alpha, t), t) v_{\gamma\alpha} + \int_\mathbb{T} v_{\beta\gamma} \phi_{\alpha\beta} \, dm_0(\beta).
\]
The integral term will remain small for all \( t \geq T \), in view Lemma 8.6 and (198). Indeed,
\[
\left| \int_\mathbb{T} v_{\beta\gamma}(t) \phi_{\alpha\beta} \, dm_0(\beta) \right| = \left| \int_{G_\delta(T)} v_{\beta\gamma}(t) \phi_{\alpha\beta} \, dm_0(\beta) \right| + \left| \int_{G_\delta(T)} v_{\beta\gamma}(t) \phi_{\alpha\beta} \, dm_0(\beta) \right|
\]
\[
\lesssim \delta^{2/3} + \frac{\varepsilon}{\delta}.
\]
So,
\[
(\phi * \rho)v_{\gamma\alpha} - \delta^{2/3} - \frac{\varepsilon}{\delta} \leq \frac{d}{dt} v(\alpha, t) \leq (\phi * \rho)v_{\gamma\alpha} + \delta^{2/3} + \frac{\varepsilon}{\delta}.
\]
Let us fix a time \( t \gtrsim T + \frac{1}{\delta^{2/3} + (\varepsilon/\delta)^{1/2}} \), and assume that \( v_{\alpha\gamma}(t) = U > 0 \), where \( U \) is to be determined later. Let us drive the dynamics backwards in time from the moment \( t \). For a time period \([s, t] \), where \( T < s < t \), the difference will remain positive \( v_{\alpha\gamma}(s) > 0 \). On that time period, the right hand side of (202) implies
\[
\frac{d}{dt} v \leq \delta^{2/3} + \frac{\varepsilon}{\delta}
\]
and hence,
\[
v(\alpha, t) - \left( \delta^{2/3} + \frac{\varepsilon}{\delta} \right) (t - s) \leq v(\alpha, s).
\]
At the same time, by \((199)\) applied to \(\gamma \in G_\delta\), we have
\[
|v(\gamma, t) - v(\gamma, s)| \leq \delta(t - s)^{1/2}.
\]
So, combined with the previous,
\[
U - \left( \frac{\delta^{2/3} + \frac{\varepsilon}{\delta}}{2} \right) (t - s) - \delta(t - s)^{1/2} = v_{\alpha\gamma}(t) - \left( \frac{\delta^{2/3} + \frac{\varepsilon}{\delta}}{2} \right) (t - s) - \delta(t - s)^{1/2} \leq v_{\alpha\gamma}(s).
\]
We will have
\[
v_{\alpha\gamma}(s) \geq \frac{U}{2},
\]
as long as \((t - s) \lesssim \frac{U}{\delta^{1/3} + \frac{\varepsilon}{\delta}}\) and \((t - s) \lesssim \frac{T^2}{\delta^2}\). The former is more restrictive, unless \(U \lesssim \delta^{4/3}\), in which case we have achieved our objective. Arguing as in the proof of Lemma 8.6 we obtain collision backwards in time, provided \((t - s) \sim 1/U\). This is possible when \(U \gtrsim \delta^{1/3} + (\varepsilon/\delta)^{1/2}\) on the time interval \(t - T \gtrsim 1/U\), which is true under the assumption.

Arguing from the opposite end, \(v_{\alpha\gamma}(t) = -U < 0\), we obtain the bound from below as well. \(\square\)

Lemma 8.7 implies the global alignment: for all \(t \gtrsim T + \frac{1}{\delta^{1/3} + (\varepsilon/\delta)^{1/2}}\)
\[
\sup_{\alpha, \gamma \in T} |v(\alpha, t) - v(\gamma, t)| \lesssim \delta^{1/3} + (\varepsilon/\delta)^{1/2}.
\]
Optimizing over \(\delta\), we pick \(\delta = \varepsilon^{3/5}\), and recalling that \(\varepsilon = \ln T/T\), we obtain
\[
\sup_{\alpha, \gamma \in T} |v(\alpha, t) - v(\gamma, t)| \lesssim \left( \frac{\ln T}{T} \right)^{1/5},
\]
for \(t \sim T + (\frac{T}{\ln T})^{1/5} \sim T\). This proves the result. \(\square\)

8.4. Global existence and strong flocking for singular models. In this section we establish global well-posedness of solutions to the Euler alignment system \((164)\) on the torus \(\mathbb{T}\) for the case of singular kernel \(\phi\). More specifically, we assume that \(\phi\) is the kernel of the classical fractional Laplacian \(\Lambda_\alpha\):
\[
\phi(z) = \sum_{k \in \mathbb{Z}} \frac{1}{|z + 2\pi k|^{1+\alpha}}, \quad 0 < \alpha < 2.
\]

Very often we refer to \(\phi\) as the kernel over the line \(\phi(z) = \frac{1}{|z|^{1+\alpha}}\) which is justified by extending corresponding functions periodically to the whole line:
\[
\Lambda_\alpha f(x) = p.v. \int_{\mathbb{T}} \phi(z) \delta_z f(x) \, dz = p.v. \int_{\mathbb{R}} \delta_z f(x) \frac{dz}{|z|^{1+\alpha}}.
\]
The analysis can be carried out for local metric kernels as well in a similar fashion.

As always in 1D the corresponding entropy will play a key role in establishing regularity of solutions to \((164)\):
\[
e = u_x + \Lambda_\alpha \rho.
\]
The main result is the following.

**Theorem 8.8.** Suppose \(m \geq 3\) and \(0 < \alpha < 2\). Let \((u_0, \rho_0) \in H^{m+1}(\mathbb{T}) \times H^{m+\alpha}(\mathbb{T})\), and \(\rho_0(x) > 0\) for all \(x \in \mathbb{T}\). Then there exists a unique non-vacuous global in time solution to \((164)\) in the class
\[
u \in C_w([0, \infty); H^{m+1}) \cap L^2([0, \infty); H^{m+1+\alpha}/2), \quad \rho \in C_w([0, \infty); H^{m+\alpha}).
\]
Moreover, the solution obeys uniform bounds on the density
\[
c_0 \leq \rho(x, t) \leq C_0, \quad t \geq 0,
\]
and strong flocking: \( \tilde{\rho} \in H^{m+\alpha} \) such that
\[
\|u(t) - \tilde{u}\|_{W^{2,\infty}} + \|\rho(\cdot, t) - \tilde{\rho}(\cdot - \tilde{u}t)\|_{C^\gamma} \leq Ce^{-\delta t} \quad t > 0, \ 0 < \gamma < 1.
\]

According to our local well-posedness Theorem 7.3 we already have a local solution \((u, \rho)\) on time interval \([0, T_0)\). We proceed in several steps. First, we establish uniform bounds (205) on the density which depend only on the initial conditions. So, such bounds hold uniformly on the available time interval \([0, T_0)\). Next, we invoke results from the theory fractional parabolic equations to conclude that our solution gains Hölder regularity after a short period of time, and the Hölder exponent and well as the bound on the Hölder norm depend on \(L^\infty\) metrics of the solution. Finally, we establish a continuation criterion much weaker than that of Theorem 7.3 – claiming that any uniform modulus of continuity for the density propels higher order norms beyond \(T_0\).

Paired with the mass equation we find that the ratio \(q = e/\rho\) satisfies the transport equation
\[
\frac{D}{Dt} q = q_t + u q_x = 0.
\]
Starting from sufficiently smooth initial condition with \(\rho_0\) away from vacuum we can assume that
\[
Q = |q(t)|_\infty = |q_0|_\infty < \infty.
\]

**Step 1: Bounds on the density.** We start by establishing (205) on the given time interval.

First, recall that \(q = \frac{e}{\rho}\) is transported, see (179), and hence is bounded for all time with its initial value \(|q_0|_\infty\). So, we can write the continuity equation as
\[
\rho_t + u \rho_x = -q\rho^2 + \rho \Lambda_\alpha(\rho).
\]
Let us evaluate at a point \(x_+\) where the maximum of \(\rho\), denoted \(\rho_+\), is reached. We obtain
\[
\frac{d}{dt} \rho_+ = -q(x_+, t) \rho_+^2 + \rho_+ \int \phi(|z|)(\rho(x_+ + z, t) - \rho_+) \, dz
\leq |q_0|_\infty \rho_+^2 + \rho_+ \int_{|z| < r} \phi(|z|)(\rho(x_+ + z, t) - \rho_+) \, dz
\leq |q_0|_\infty \rho_+^2 + \frac{1}{r^{1+\alpha}} \rho_+(M - 2r \rho_+) = |q_0|_\infty \rho_+^2 + \frac{1}{r^{1+\alpha}} M \rho_+ - \frac{2}{r^\alpha} \rho_+^2.
\]

Let us pick \(r\) large enough so that \(\frac{2}{r^\alpha} > |q_0|_\infty + 1\). Then
\[
\frac{d}{dt} \rho_+ \leq -\rho_+^2 + C(M, r) \rho_+,
\]
which establishes the upper bound by integration.

As to the lower bound we argue similarly. Let \(\rho_-\) be the minimum value of \(\rho\) and \(x_-\) a point where such value is achieved. We have
\[
\frac{d}{dt} \rho_- \geq -|q_0|_\infty \rho_-^2 + \rho_- \int \phi(|z|)(\rho(x_- + z, t) - \rho_-) \, dz
\geq -|q_0|_\infty \rho_-^2 + \phi_- \rho_- (M - 2\pi \rho_-) = -c_1 \rho_-^2 + c_2 \rho_-.
\]
This readily implies the bound from below. Note that at this point the global communication of the model is crucial: \(\phi_- > 0\).

As a consequence of the lower bound on the density we have a global bound on the entropy:
\[
\sup_{t \in [0,T_0]} |e(t)|_\infty < \infty.
\]

**Step 2: Hölder regularization.** The representation of continuity equation in the form (209) puts it into the class of forced fractional parabolic equations with bounded drift and force:
\[
\partial_t v + u \cdot \nabla v = L[v] + f,
\]
where $L$ has kernel
\[ K(x, z, t) = \rho(x) \frac{1}{|z|^{1+\alpha}}, \]
which is even with respect to $z$. The bounds on the density provide uniform ellipticity bounds on the kernel $|z|^{1+\alpha} \lesssim K(x, z, t) \lesssim |z|^{1+\alpha}$.

With these ingredients at hand, the case $\alpha = 1$ falls under the assumptions of Silverstre’s results [12] which provides Hölder regularization bound given by
\[
|\rho|_{C^\gamma([t_0, T_0] \times \mathbb{R}^3)} \leq C(|\rho|_{L^\infty([t_0, T_0] \times \mathbb{R}^3)} + |e\rho|_{L^\infty([t_0, T_0] \times \mathbb{R}^3)}),
\]
for some $\gamma > 0$.

The case $\alpha < 1$ falls under the same result provided $u \in L^\infty([0, T_0); C^{1-\alpha})$. This is indeed the case as follows from
\[
\Lambda_\alpha^{-1} \partial_x u = \Lambda_\alpha^{-1} e - \rho \in L_t^\infty.
\]
Note that $\Lambda_\alpha^{-1} \partial_x$ a $(1 - \alpha)$-order differential operator.

Finally, for $\alpha > 1$ the Hölder continuity follows from a similar identity for $\rho$:
\[
\Lambda_\alpha^{-1} \rho = \Lambda_1^{-1} e - \mathcal{H} u,
\]
where $\mathcal{H}$ is the Hilbert transform. Note that it sends functions in $L^\infty$ to $B^0_{\infty, \infty}$. Hence, $\rho \in B^0_{\infty, \infty} = C^{\alpha-1}$.

**Step 3: Continuation Criterion.** Last step is to show that if the density is bounded in $C^\gamma$ on time interval $[T_0/2, T_0)$ then the solution remains uniformly in $W^{1, \infty}$, and hence the continuation criterion of Theorem 7.3 applies. While doing that we will keep track of estimates on the $W^{1, \infty}$ with the purpose to obtaining long time asymptotics.

**Step 3a: Control over $\rho'$.** So, let us start with $\rho'$:
\[
\partial_t \rho' + u \rho'' + u' \rho' + e' \rho + e \rho' = \rho' \Lambda_\alpha \rho + \rho \Lambda_\alpha \rho'.
\]
Using again $u' = e - \Lambda_\alpha \rho$ we rewrite
\[
\partial_t \rho' + u \rho'' + u' \rho + 2e \rho' = 2\rho' \Lambda_\alpha \rho + \rho \Lambda_\alpha \rho'.
\]
Evaluating at the maximum of $|\rho'|$ and multiplying by $\rho'$ we obtain
\[
\partial_t |\rho'|^2 + e \rho' \rho' + 2e |\rho'|^2 = 2|\rho'|^2 \Lambda_\alpha \rho + \rho \Lambda_\alpha \rho'.
\]
Let us note that $q'$ satisfies the continuity equation, and consequently, $\frac{q'}{\rho}$ is transported. So, $|q'| \leq C \rho$ pointwise. For the e-quantity itself this implies pointwise bound
\[
|e'(x, t)| \leq C(|\rho'(x, t)| + \rho(x, t)).
\]
Let us note that in order to make pointwise evaluation possible in (214) one has to assume regularity $e' \in C^1(\mathbb{T})$ which guaranteed provided $m \geq 2$. With this at hand, and in view of (205) and (210) we can bound
\[
|e' \rho \rho' + 2e |\rho'|^2 | \leq C(|\rho'|^2 + |\rho'|).
\]
Thus,
\[
\partial_t |\rho'|^2 = C(|\rho'|^2 + |\rho'|) + 2|\rho'|^2 \Lambda_\alpha \rho + \rho \Lambda_\alpha \rho'.
\]
Due to the bound from below on $\rho$, we estimate
\[
\rho \rho' \Lambda_\alpha \rho' \leq c_1 \int_{\mathbb{R}} \frac{(\rho'(x + z) - \rho'(x))\rho'(x + z)}{|z|^{1+\alpha}} \, dz \leq -c_2 D_\alpha \rho'(x).
\]
where
\[
D_\alpha \rho'(x) = \int_{\mathbb{R}} \frac{|\rho'(x) - \rho'(x + z)|^2}{|z|^{1+\alpha}} \, dz.
\]
Lemma 8.9 (Non-local maximum principle). The following pointwise bound holds
\[(217) \quad D_{\alpha}\rho'(x) \geq c \frac{\rho'(x)^{2+\alpha}}{\rho|_{\infty}}.\]

\textbf{Proof.} Fix an \(r > 0\) to be determined later. We write
\[
D_{\alpha}\rho'(x) \geq \int_{|z| > r} \frac{|\rho'(x) - \rho'(x + z)|^2}{|z|^{1+\alpha}} \, dz \geq \int_{|z| > r} \frac{|\rho'(x)|^2 - 2\rho'(x + z)\rho'(x)}{|z|^{1+\alpha}} \, dz
\]
\[
= \frac{|\rho'(x)|^2}{r^{\alpha}} - 2\rho'(x) \int_{|z| > r} \frac{\rho'(x + z)}{|z|^{1+\alpha}} \, dz.
\]
Integrating by parts in the last integral we further estimate
\[
D_{\alpha}\rho'(x) \geq \frac{|\rho'(x)|^2}{r^{\alpha}} - c_{\alpha} \rho'(x) \|\rho\|_{\infty} \frac{1}{r^{1+\alpha}}.
\]
Choosing \(r = C |\rho|_{\infty}/|\rho'(x)|\), where \(C\) is large proves the estimate. \(\square\)

In view of density bounds we have a priori (205), the non-local maximum principle yields the following non-linear bound
\[
D_{\alpha}\rho'(x) \geq c |\rho'(x)|^{2+\alpha}.
\]

We arrive at
\[(218) \quad \partial_t |\rho|^2 = C(|\rho|^2 + |\rho'|) + 2|\rho'|^2 \Lambda_{\alpha}\rho - c|\rho|^{2+\alpha} - \frac{1}{2} D_{\alpha}\rho'(x).
\]
The lower order terms \(|\rho|^2 + |\rho'|\) can be absorbed into dissipation by the generalized Young inequality:
\[
|\rho|^2 + |\rho'| \leq c_{\varepsilon} + \varepsilon |\rho|^{2+\alpha},
\]
for \(\varepsilon > 0\) small. So, it remains to obtain estimate on the remaining term \(|\rho|^2 \Lambda_{\alpha}\rho\).

To do that we fix a scale parameter \(1 > r > 0\) to be determined later, and split the integral representation of the fractional Laplacian into three parts: short-range, mid-range, and long-range
\[
\Lambda_{\alpha}\rho(x) = \int_{|z| < r} \delta_z \rho(x) - \rho'(x)z \, \frac{dz}{|z|^{1+\alpha}} + \int_{r < |z| < 1} \delta_z \rho(x) \, \frac{dz}{|z|^{1+\alpha}} + \int_{|z| > 1} \delta_z \rho(x) \, \frac{dz}{|z|^{1+\alpha}}
\]
\[
:= S + M + L.
\]
For the short-range we use the dissipation directly:
\[(219) \quad |\rho(x + z) - \rho(x) - \rho'(x)z| = \left| \int_0^z (\rho'(x + w) - \rho'(x)) \, dw \right| \leq \sqrt{D_{\alpha}\rho'(x)} |z|^{1 + \frac{\alpha}{2}},
\]
so,
\[
|S| \leq r^{1-\alpha/2} \sqrt{D_{\alpha}\rho'(x)}.
\]
In the mid-range we use the available Hölder continuity (here we can assume without loss of generality that \(\gamma < \alpha\)):
\[
|M| \leq |\rho|_{C^{1,\gamma}} r^{\gamma - \alpha} \lesssim r^{\gamma - \alpha}.
\]
And finally, for the long-range we simply use the boundedness of \(\rho\):
\[
|L| \lesssim |\rho|_{\infty}.
\]

The competition occurs only between the short- and mid-range terms. Optimizing over \(r\) we set \(r = (D_{\alpha}\rho'(x))^{-\frac{\gamma}{\gamma + \alpha + 2\gamma}}\) unless such expression is \(> 1\), in which case we have an absolute bound on the dissipation and the proof proceeds trivially. With the established bounds we obtain the following pointwise estimate
\[
|\Lambda_{\alpha}\rho(x)| \lesssim c_1 + c_2 (D_{\alpha}\rho'(x))^{\frac{\alpha - \gamma}{\gamma + \alpha + 2\gamma}}.
\]
Note that $\frac{\alpha - \gamma}{2 + \alpha + 2\gamma} < \frac{\alpha}{2 + \alpha}$. So, we can use the generalized Young inequality we obtain

$$|\rho'|^2|\Lambda_\alpha \rho| \lesssim c_\varepsilon + \varepsilon|\rho'|^{2+\alpha} + \varepsilon D_\alpha \rho'(x).$$

Plugging this into (218) we arrive at

(220) \[ \partial_t |\rho'|^2 \leq c_1 - c_2 |\rho'|^{2+\alpha}. \]

This concludes the proof of uniform bound $\rho \in L^\infty([0, T_0]; W^{1, \infty})$.

**Step 3b: Control over $u'$.** Note that for the case $0 < \alpha < 1$ this bound is straightforward from the $\varepsilon$-quantity. Indeed, the $\varepsilon$-quantity is uniformly bounded by (210), while $\Lambda_\alpha \rho \in L^\infty$ simply by $|\Lambda_\alpha \rho|_\infty \leq |\rho'|_\infty$. However, we will seek more precise estimates with the view towards long time behavior. So, we will revisit this case also in the course of completing this step.

So, let us write the equation for $u'$, evaluated at maximum of $|u'|$ and multiplied by $u'$:

$$\frac{d}{dt} |u'|^2 \leq |u'|^3 + u'(x) \int_R \delta_z u'(x) \rho(x + z) \frac{dz}{|z|^{1+\alpha}} + u'(x) \int_R \delta_z u(x) \rho'(x + z) \frac{dz}{|z|^{1+\alpha}}.$$

Note that the last term is bounded by $C |u'(x)| \mathcal{A}(t)$, and we know that a priori $\mathcal{A}(t)$ is an exponentially decaying quantity. The dissipation term is bounded, as before by

$$u'(x) \int_R \delta_z u'(x) \rho(x + z) \frac{dz}{|z|^{1+\alpha}} \leq -c D_\alpha u'(x).$$

The dissipation term obeys another non-local maximum principle similar to (217) where instead of $u'$ we replace it with $(u - \bar{u})'$, thus the denominator contain the amplitude $\mathcal{A}$ rather than $|u|_\infty$:

(221) \[ D_\alpha u'(x) \geq c \frac{|u'(x)|^{2+\alpha}}{\mathcal{A}^\alpha(t)}. \]

We continue

(222) \[ \frac{d}{dt} |u'|^2 \leq |u'|^3 - c \frac{|u'(x)|^{2+\alpha}}{\mathcal{A}^\alpha(t)} + C |u'(x)| \mathcal{A}(t). \]

If $\alpha > 1$, we absorb the cubic and linear terms by

$$|u'|^3 + C |u'(x)| \mathcal{A}(t) \leq \varepsilon \frac{|u'(x)|^{2+\alpha}}{\mathcal{A}^\alpha(t)} + c_\varepsilon \mathcal{A}^{\frac{3\alpha}{\alpha-1}}(t) + c_\varepsilon \mathcal{A}^2(t).$$

So,

(223) \[ \frac{d}{dt} |u'|^2 \leq E(t) - c \frac{|u'(x)|^{2+\alpha}}{\mathcal{A}^\alpha(t)}, \]

where $E$ denotes an exponentially decaying quantity. In particular this establishes uniform control over $|u'|_\infty$.

If $0 < \alpha < 1$, we already know from the remark in the beginning of this step that $|u'|$ is uniformly bounded. So, we estimate $|u'|^3 \lesssim |u'|^2$, and the rest goes as before to obtain (223).

Let us investigate the remaining critical case $\alpha = 1$. This is when the dissipation is not strong enough to control nonlinearity, and at the same time $\Lambda_\alpha \rho$ is not automatically bounded either. We will do it on the next step.

**Step 3c: Control over $\Lambda_1 \rho$.** Let us assume that $\alpha = 1$ and denote $\Lambda_1 = \Lambda$. We obtain an estimate on $\Lambda \rho$ indirectly, by establishing an energy-type bound on $\rho'''$. So, assuming we have proved that $|\rho'''| \in L^\infty([0, T_0])$, control over $\Lambda \rho$ goes as follows:

$$\Lambda \rho(x) = \int_{|z| < 1} [\delta_z \rho(x) - \rho'(x)z] \frac{dz}{|z|^2} + \int_{1 < |z|} \delta_z \rho(x) \frac{dz}{|z|^2}.$$
The second integral is clearly bounded uniformly. Next,

\[ |\delta_z \rho(x) - \rho'(x)z| = \left| \int_0^z \int_0^y \rho''(x+y) \, dy \right| \leq |\rho''|_2|z|^{3/2}. \]

So, the first integral is bounded by a constant multiple of $|\rho''|_2$. This shows that $\Lambda \rho \in L^\infty([0,T_0); L^\infty)$.

So, let us write the second derivative of density:

\[ \partial_t \rho'' + u \rho''' + u' \rho'' + e'' \rho + 3e' \rho' + 2e \rho'' = 2\rho'' \mathcal{L}_\rho \rho + 3\rho' \mathcal{L}_\rho \rho' + \rho \mathcal{L}_\rho \rho'' \tag{224} \]

Let us apply the test-function $\rho''/\rho$. Via routine computation with the use of the density equation, one can observe that

\[ \left\langle \partial_t \rho'' + u \rho''' + u' \rho'' + \frac{\rho''}{\rho}, \frac{\rho''}{\rho} \right\rangle = \frac{1}{2} \partial_t \int \frac{1}{\rho} |\rho''|^2 \, dx. \]

In view of the bounds on the density we note that $\int \frac{1}{\rho} |\rho''|^2 \, dx \sim |\rho''|^2_2$. So, it is sufficient to bound the rest of the terms in terms of $|\rho''|_2^2$. Considering the last three terms on the left hand side, let us make one observation: since $q'/\rho$ is transported, then $(q'/\rho)'$ satisfied the continuity equation, and hence $(q'/\rho)/\rho$ is transported again. Solving for $e''$ in this expression results in poinwise bound

\[ |e''(x,t)| \leq C(|\rho''(x,t)| + |\rho'(x,t)| + \rho(x,t)). \tag{225} \]

In order for this bound to make sense we require $m \geq 3$. With the use of a priori estimates established so far,

\[ \left\langle e'' \rho + 3e' \rho' + 2e \rho'' \right\rangle \lesssim 1 + |\rho''|^2_2. \]

At this point we have (dropping $\rho, \rho'$ that are already bounded)

\[ \partial_t \int \frac{1}{\rho} |\rho''|^2 \, dx \lesssim 1 + |\rho''|^2_2 + \int |\rho''|^2 |\Lambda \rho| \, dx + \int |\rho''| |\Lambda \rho'| \, dx + \int \rho'' \Lambda \rho'' \, dx \]

\[ = 1 + |\rho''|^2_2 + I_1 + I_2 + I_3. \tag{226} \]

Clearly, the last term $I_3$ is dissipative:

\[ I_3 \lesssim - \int \mathcal{D}_\alpha \rho''(x) \, dx - \frac{1}{|\rho''|_\infty} \int |\rho''|^3 \, dx = -\|\rho''\|_{L^2_H}^2 - |\rho''|_3^3, \]

where in the latter we dropped $\frac{1}{|\rho''|_\infty}$ from inside the integral since this term is bounded from below.

To tackle $I_1$ we fix $\varepsilon > 0$ small and split the fractional Laplacian:

\[ \Lambda \rho(x) = \int_{|z|<\varepsilon} \delta_z \rho(x) - \rho'(x)z \frac{dz}{|z|^2} + \int_{\varepsilon<|z|} \delta_z \rho(x) \frac{dz}{|z|^2} \lesssim \varepsilon^{1/2} |\rho''|_2 + c_\varepsilon. \]

So,

\[ I_1 \lesssim \varepsilon^{1/2} |\rho''|_2^2 + c_\varepsilon |\rho''|_2^2 \lesssim \varepsilon^{1/2} |\rho''|^3_3 + c_\varepsilon |\rho''|_2^2. \]

The cubic term gets absorbed by dissipation for small $\varepsilon$.

For $I_2$ we simply use the H"older inequality:

\[ |I_2| \leq |\rho''|_2 |\Lambda \rho'|_2 \lesssim |\rho''|^2_2. \]

So,

\[ \partial_t \int \frac{1}{\rho} |\rho''|^2 \, dx \lesssim 1 + |\rho''|^2_2 - |\rho''|^3_3. \tag{227} \]

This finishes the step.
Coming back to Step 3b, we conclude that \( u' = \Lambda - \rho \) remains bounded. This concludes the proof of global existence.

**STEP 4: FLOCKING.** In the cases when \( \alpha \neq 1 \), the estimates above come with a good a priori bound (223), which shows that \( |u'|_\infty \leq E(t) \), an exponentially decaying quantity. In the exceptional case \( \alpha = 1 \) we go back to (222) to obtain

\[
\frac{d}{dt} |u'|^2 \leq |u'|^3 - c |u'(x)|^3 \Lambda(t) + \Lambda^2(t).
\]

Clearly, in the long run the dissipation term overtakes \( |u'|^3 \) and we arrive at the same conclusion.

Since on Step 3a we showed that \( \rho \) is uniformly bounded \( W^{1,\infty} \), the proof of strong flocking for the density, \( \rho \to \bar{\rho}(-\bar{u}t) \), follows along the lines of Theorem 8.1.

Lastly, showing exponential decay of \( |u''|_\infty \) follows similar estimates on the evolution of the norm \( |u''|^2 \), and will not be presented here for the sake of brevity. We refer to [11] for full details.

### 9. Regularity theory of multi-dimensional systems

#### 9.1. Oriented flocks.

One class of solutions that behaves like 1D is the class of unidirectional oriented flows. These are given by

\[
\mathbf{u}(x,t) = u(x,t) \mathbf{d}, \quad \mathbf{d} = \mathbb{S}^{n-1}, \quad u : \mathbb{R}^n \times \mathbb{R}^+ \to \mathbb{R}.
\]

The same conservation law holds for the entropy

\[
e = \mathbf{d} \cdot \nabla u + \phi * \rho, \quad \partial_t e + \nabla \cdot (e \mathbf{u}) = 0,
\]

although in this case the entropy does not control the full gradient of the velocity field. Nonetheless, one can develop a theory fully analogous to the 1D in the smooth communication case, which is what we will cover in this section.

First of all by the maximum principle applied in any direction perpendicular to \( \mathbf{d} \) one can see that the ansatz (228) is preserved in time. Second, in view of rotational invariance of the Euler Alignment System, we can assume that \( \mathbf{d} \) points in the direction of the \( x_{11} \)-axis. So, we can assume

\[
\mathbf{u}(x,t) = \langle u(x,t), 0, \ldots, 0 \rangle \quad \text{for} \quad u : \mathbb{R}^n \times \mathbb{R}^+ \to \mathbb{R}.
\]

The full system (102) takes for of a system of scalar conservation laws:

\[
(x,t) \in \mathbb{R}^n \times \mathbb{R}^+ \quad \begin{cases} \partial_t \rho + \partial_1 (\rho u) = 0, \\ \partial_t u + \frac{1}{2} \partial_1 (u^2) = \phi * (\rho u) - u \phi * \rho. \end{cases}
\]

The entropy takes form

\[
e := \partial_1 u + \phi * \rho, \quad \partial_t e + \partial_1 (ue) = 0.
\]

We prove the analogue of our one dimensional Theorem 8.1, which reads exactly the same if \( u \) is understood to be the first coordinate of the velocity field, and the alignment part involved full gradients:

\[
\mathcal{A}(t) + \|\nabla u(t)\|_{L^\infty(\text{Supp}(\rho(t)))} + \|\nabla^2 u(t)\|_{L^\infty(\text{Supp}(\rho(t)))} \leq C e^{-\delta t}.
\]

According to our continuation criterion Theorem 7.2 we only have to establish a uniform control over the full gradient \( |\nabla u|_{\infty} \) on a local interval of existence \( [0,T_0) \). Due to entropy law we can resort to the same argument initially as in 1D case. We have the logistic equation (167) as before. So, for the supercritical case \( e_0(x_0) < 0 \) the argument is the same to show blow up in finite time.
Above the threshold, \( e_0 \geq 0 \), the \( e \) will remain non-negative and bounded. As a result we obtain an a priori bound
\[
|\partial_1 u(t)|_{\infty} < C_1, \quad \forall t \in [0, T_0).
\]
In order to apply the continuation criterion we need to show that the full gradient of \( u \) remains bounded.

So, let us fix \( i \neq 1 \) and establish control over \( \partial_i u \). We write the equation of \( \partial_i u \) along characteristics:
\[
(234) \quad \frac{D}{Dt} \partial_i u + \partial_i u \partial_1 u = \partial_i \phi * (\rho u) - (\partial_i \phi * \rho) u - (\phi * \rho) \partial_i u
\]
Here and below, \( \frac{D}{Dt} \) denotes differentiation along generic particle path \( X(t, x_0) \) initiated at \( x_0 \). Let us evaluate (234) at the maximum of \( \partial_1 u \):
\[
\partial_i |\partial_1 u|_{\infty} \leq (|\partial_1 u|_{\infty} + |\phi|_{\infty}) |\partial_1 u|_{\infty} + 2|\partial_1 \phi|_{\infty} u|_{L^\infty}.
\]
In view of the \( L^\infty \)-bound of \( u \) and \( \partial_1 u \), we have that
\[
\partial_t \|\partial_1 u\|_{L^\infty} \lesssim \|\partial_1 u\|_{L^\infty} + 1.
\]
This readily implies the desired result by integration.

For flocking estimates, we start as before again to conclude that since the diameter of the flock \( D(t) \) will remain finite there exists a time \( t^* > 0 \) such that \( e(x, t) \geq \frac{1}{2} \phi(D)M \) for all \( x \in \text{Supp} \rho(\cdot, t) \) and \( t > t^* \). With this in mind we write another equation for \( \partial_i u \) in the following form
\[
\partial_i \partial_j u + \partial_i u \partial_1 u + u \partial_i \partial_1 u = \partial_i \phi * (\rho u) - \partial_i u \phi * (\rho) - u \partial_i \phi * (\rho),
\]
or along characteristics
\[
\frac{D\partial_i u}{Dt} = \partial_i \phi * (\rho u) - u \partial_i \phi * (\rho) - (\partial_i u + \phi * (\rho)) \partial_i u
\]
\[
= \int \partial_i \phi(|x - y|) (u(y) - u(x)) \rho(y) dy - e \partial_i u.
\]
We already know that the velocity fluctuations \( A(t) \) are exponentially decaying. Hence, the integral above will be bounded by \( |\partial_1 \phi|_{\infty} ME(t) \) where \( E(t) \) is a generic exponentially decaying quantity. Evaluating (235) at the maximum over \( \text{Supp} \rho(\cdot, t) \) we obtain
\[
\partial_i \|\partial_1 u\|_{L^\infty(\text{Supp} \rho(\cdot, t))} \leq E(t) - \frac{1}{2} \phi(D)M \|\partial_1 u\|_{L^\infty(\text{Supp} \rho(\cdot, t))}.
\]
This readily implies the exponential bound on \( \|\partial_1 u\|_{L^\infty(\text{Supp} \rho(\cdot, t))} \) for \( i = 1, \ldots, n \).

Moving on to the second order derivatives, we write the equations along characteristics
\[
\frac{D\partial_j \partial_i u}{Dt} = \int \partial_j \partial_i \phi(|x - y|) (u(y) - u(x)) \rho(y) dy - e \partial_j \partial_i u - \partial_i \partial_j u - \partial_i e \partial_j u,
\]
where
\[
\partial_i e = \partial_j \partial_i u + \partial_i \phi * \rho \quad \text{and} \quad \partial_i e = \partial_i \partial_1 u + \partial_i \phi * \rho.
\]
We prove exponential decay by bootstrapping information from the partial \( \partial_1 \partial_1 \), then \( \partial_1 \partial_j \), and then general \( \partial_i \partial_j \). So, first, we consider the case \( i = j = 1 \). Note that for this particular case, we have
\[
\frac{D\partial_1^2 u}{Dt} = \int \partial_1^2 \phi(|x - y|) (u(y) - u(x)) \rho(y) dy - e \partial_1^2 u - 2 \partial_1 e \partial_1 u.
\]
Using that \( \partial_1 e = \partial_1^2 u + \partial_1 \phi * \rho \), we arrive at
\[
\frac{D\partial_1^2 u}{Dt} \leq E(t) - \partial_1^2 u(e - E(t)).
\]
Now, as \( e(x, t) \geq \frac{1}{2} \phi(D)M > 0 \) for all \( x \in \text{Supp} \rho(\cdot, t) \) and \( t > t^* \) we have that
\[
\frac{D\partial_t^2 u}{Dt} \leq E(t) - \partial_t^2 u \left( \frac{1}{2} \phi(D)M - E(t) \right) \quad \text{for } t > t^*.
\]
As \( E(t) \) decays exponentially fast, there must exist \( t^{**} > t^* \) such that
\[
\frac{D\partial_t^2 u}{Dt} \leq E(t) - \frac{1}{4} \phi(D)M \partial_t^2 u \quad \text{for } t > t^{**}.
\]
Then, evaluating the previous inequality at the maximum over \( \text{Supp} \rho(\cdot, t) \) we have obtained the desired result by integration.

Second, we consider the case \( i = 1 \) and \( j \neq 1 \). In this case, we have
\[
\frac{D\partial_j \partial_1 u}{Dt} = \int_{\Omega} \partial_j \partial_1 \phi(|x - y|) (u(y) - u(x)) \rho(y) \, dy - e \partial_j \partial_1 u - \partial_1 e \partial_j u,
\]
where
\[\partial_j e = \partial_j \partial_1 u + \partial_1 \phi \ast (\rho) \quad \text{and} \quad \partial_1 e = \partial_1^2 u + \partial_1 \phi \ast (\rho).\]
Using that \( \|\nabla u\|_{L^\infty(\text{Supp} \rho(\cdot, t))} \leq E(t) \) and the fact that \( \|\partial_t^2 u\|_{L^\infty(\text{Supp} \rho(\cdot, t))} \leq E(t) \) we get
\[
\frac{D\partial_j \partial_1 u}{Dt} \leq E(t) - \partial_j \partial_1 u (e - E(t))
\]
and doing the same as before we obtain that \( \|\partial_j \partial_1 u\|_{L^\infty(\text{Supp} \rho(\cdot, t))} \leq E(t) \) for \( j \neq 1 \).

Finally, the case \( i, j \neq 1 \) relies on the previous in a similar manner. We get
\[
\frac{D\partial_j \partial_j u}{Dt} \leq E(t) - e \partial_j \partial_j u
\]
and hence \( \|\partial_j \partial_j u\|_{L^\infty} \leq E(t) \).

The same argument as in 1D shows that \( \|\nabla \rho\|_{L^\infty} \) remains uniformly bounded, and with the exponential decay of velocity this implies strong flocking as in 1D again.


9.3. Spectral dynamics approach. In dimension 2 one can obtain an alternative threshold condition based on spectral dynamics approach. Let us assume that the kernel \( \phi \) is smooth, of convolution type, and satisfies the usual fat tail condition (7). Recall the entropy
\[
e = \nabla \cdot u + \phi \ast \rho,
\]
satisfying the equation
\[
e_t + \nabla \cdot (ue) = (\nabla \cdot u)^2 - \text{Tr}(\nabla u)^2.
\]
In 2D the right hand side is equal exactly to \( 2 \det(\nabla u) \). So, if we attempt to appeal as in 1D to the logistic nature of the equation we write
\[
\frac{d}{dt} e = e(\phi \ast \rho - e) + 2 \det(\nabla u).
\]
So, the residual term \( \det(\nabla u) \) gets in the way of controlling the growth or sign of \( e \). It is difficult however to track down dynamics of \( \det(\nabla u) \) since \( \nabla u \) is non-symmetric. Instead, one can track the dynamics of the symmetric part of \( \nabla u \), and in particular the eigenvalues of \( S = \frac{1}{2}(\nabla u + \nabla^\perp u) \).

In order to see exactly what we are aiming for, let us note that
\[
\det(\nabla u) = \det S + \omega^2,
\]
where \( \omega = \frac{1}{2}(\partial_1 u_2 - \partial_2 u_1) \) is the scalar vorticity of the field. Denote by \( \mu_1, \mu_2 \) the eigenvalues of \( S \). Then \( \det S = \mu_1 \mu_2 \). At the same time
\[
\mu_1 \mu_2 = \frac{1}{4}(\mu_1 + \mu_2)^2 - \frac{1}{4}(\mu_1 - \mu_2)^2.
\]
The first term is exactly $\frac{1}{4}(\nabla \cdot \mathbf{u})^2$, and the second involves spectral gap denoted $\eta = \mu_1 - \mu_2$. So,

$$2 \det(\nabla \mathbf{u}) = \frac{1}{2}(\nabla \cdot \mathbf{u})^2 - \frac{1}{2} \eta^2 + 2 \omega^2.$$ 

Expanding $\nabla \cdot \mathbf{u} = e - \phi * \rho$ the e-equation can now be rewritten as follows

$$2 \frac{d}{dt} e = (\phi * \rho)^2 + 4 \omega^2 - \eta^2 - e^2.$$ 

The issue now reduces to whether we can control the spectral gap $\eta$ and vorticity $\omega$. It turns out that evolution of both quantities can be read off easily from the equation for $\nabla \mathbf{u}$. Indeed, let us write the full matrix equation first:

$$\partial_t \nabla \mathbf{u} + \mathbf{u} \cdot \nabla (\nabla \mathbf{u}) + (\nabla \mathbf{u})^2 = - (\phi * \rho) \nabla \mathbf{u} + \mathbf{E},$$

where $\mathbf{E}$ is an exponentially decaying quantity. To be precise,

$$\mathbf{E} = C \nabla \phi \mathbf{u},$$

and according to (23),

$$|\mathbf{E}| \leq A_0 e^{-\lambda M \phi(\bar{D})} |\nabla \phi|_{\infty} M,$$

where $\bar{D}$ is determined solely from initial condition by equation (26). We decompose $\nabla \mathbf{u}$ into symmetric and skew-symmetric parts

$$\nabla \mathbf{u} = \mathbf{S} + \Omega, \quad \mathbf{S} = \frac{1}{2}(\nabla \mathbf{u} + (\nabla \mathbf{u})^\top), \quad \Omega = \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix}.$$ 

And we have decomposition of the square matrix

$$\begin{pmatrix} \nabla \mathbf{u} \end{pmatrix}^2 = \frac{\mathbf{S}^2 - \omega^2 \mathbf{1}_{2 \times 2}}{\text{sym}} + \frac{\mathbf{S} \Omega + \Omega \mathbf{S}}{\text{skew-sym}} = \mathbf{S}^2 - \omega^2 \mathbf{1}_{2 \times 2} + \Omega \nabla \cdot \mathbf{u}.$$ 

So, reading off the equation for skew-symmetric part we obtain (in Lagrangian coordinates)

$$\frac{d}{dt} \omega + e \omega = \mathbf{E}.$$ 

For the symmetric part we have

$$\frac{d}{dt} \mathbf{S} + \mathbf{S}^2 = \omega^2 \mathbf{1}_{2 \times 2} - (\phi * \rho) \mathbf{S} + \mathbf{E}.$$ 

Now the advantage of considering symmetric $\mathbf{S}$ come into light at this point. Denote $(\mathbf{s}_1(x, t), \mathbf{s}_2(x, t))$ the orthonormal basis of eigenvectors corresponding to $\mu_1$ and $\mu_2$, respectively. Then $\mu_i = \mathbf{s}_i \mathbf{S} \mathbf{s}_i$, and note that

$$\frac{d}{dt} \mu_i = \mathbf{s}_i \left[ \frac{d}{dt} \mathbf{S} \right] \mathbf{s}_i,$$

due to orthogonality $\mathbf{s}_i \cdot \frac{d}{dt} \mathbf{s}_i = 0$. So, multiplying the $\mathbf{S}$-equation by $\mathbf{s}_i$ from both sides, we obtain a system for spectral dynamics of the eigenvalues

$$\frac{d}{dt} \mu_i + \mu_i^2 = \omega^2 - (\phi * \rho) \mu_i + \mathbf{E}.$$ 

Lastly, taking the difference and using that $\mu_1^2 - \mu_2^2 = \eta (\nabla \cdot \mathbf{u})$ we obtain

$$\frac{d}{dt} \eta + e \eta = \mathbf{E}.
Collecting together the equations we have obtained the system

\[
\begin{align*}
2\dot{e} + e^2 &= (\phi \ast \rho)^2 + 4\omega^2 - \eta^2 \\
\dot{\omega} + e\omega &= E \\
\dot{\eta} + e\eta &= E.
\end{align*}
\]

(242)

Let us note in passing that the bound on $E$'s in (240) is still valid up to an absolute constant due to algebraic manipulations above.

So, let us now fix an initial condition $(u_0, \rho_0) \in H^m \times (L^1_+ \cap W^{k, \infty})$ and assume that $e_0(x) > 0$ for every $x \in \mathbb{R}^2$. According to Theorem 7.2 we have a local solution on a maximal time interval $[0, T_0)$. By continuity, $e(X(t, x), t) > 0$ for some short time $t < T(x)$. On that time interval the spectral gap solution reads

\[
\eta(t) = \eta_0 \exp \left\{ -\int_0^t e(s) \, ds \right\} + \int_0^t \exp \left\{ -\int_s^t e(\tau) \, d\tau \right\} E(s) \, ds.
\]

So,

\[
|\eta(t)| \leq |\eta_0| + cA_0 \frac{\|\nabla \phi\|_{\infty}}{\phi(\bar{D})}.
\]

Using this we obtain from the $e$-equation

\[
2\dot{e} \geq \phi^2(\bar{D})M^2 - \left( |\eta_0| + cA_0 \frac{\|\nabla \phi\|_{\infty}}{\phi(\bar{D})} \right)^2 - e^2.
\]

Assuming now the small gap and small amplitude condition

\[
|\eta_0| < \frac{1}{4} \phi(\bar{D})M, \quad A_0 < \frac{1}{4} \frac{\phi^2(\bar{D})M}{c\|\nabla \phi\|_{\infty}}
\]

we find that

\[
2\dot{e} \geq \frac{1}{2} \phi^2(\bar{D})M^2 - e^2.
\]

This shows that $e$ will remain positive on the entire interval $[0, T_0)$, which in turn implies that $\eta$ will remain bounded on $[0, T_0)$. Solving the analogous vorticity equation we conclude that $\omega$ also remains bounded on $[0, T_0)$. Now the $e$-equation again can be written as

\[
2\dot{e} \leq C - e^2,
\]

which shows that $e$ itself remains bounded on $[0, T_0)$, which implies that $\nabla \cdot u$ remains bounded. The three parameters $\eta, \nabla \cdot u, \omega$ control the full gradient of $u$. Hence, $|\nabla u|_{\infty}$ is bounded. The continuation criterion of Theorem 7.2 implies global existence.

**Theorem 9.1.** Suppose $(u_0, \rho_0) \in H^m \times (L^1_+ \cap W^{k, \infty})$ and assume that $e_0(x) > 0$ for every $x \in \mathbb{R}^2$. Assume also the smallness conditions

\[
|\eta_0| < \frac{1}{4} \phi(\bar{D})M, \quad A_0 < \frac{1}{4} \frac{\phi^2(\bar{D})M}{c\|\nabla \phi\|_{\infty}}.
\]

Then there exists a global solution with this initial condition.

It is interesting to note that the small amplitude condition alone would guarantee global existence for singular models due to additional dissipation enhancement effect, see Section 9.4.

Let observe that in the case of oriented flocks described in the previous section, $|\eta_0| = |\nabla u_0|$. So, the spectral gap not always applies to those flows. Yet, we know they are globally well-posed. It would be interesting to bridge the gap between 1D theory and the spectral gap condition.
9.4. Small initial data in Hölder spaces. Lack of control on $e$ in multi dimensional case is part of the reason why the model has no well developed regularity theory. In the singular kernel case the dissipation provided by the alignment commutator may in some cases be strong enough to overcome possible growth of solution. The non-local maximum principle expressed in the nonlinear bound (221) gives a good indication that the alignment amplifies diffusion in the velocity equation, which if initially small remains so for all time. This gives a mechanism to control the entropy and subsequently obtain a global existence of small amplitude solutions. We carry out this idea in the case of $n$-dimensional torus, naturally, for non-vacuous data.

To fix the notation, we denote by $(\cdot,\cdot)$ the metric of the homogeneous Hölder class $C^\alpha(\mathbb{T}^n)$. We consider the global singular communication given by kernel of the classical fractional Laplacian $\Lambda_\alpha$:

$$\phi(z) = \sum_{k \in \mathbb{Z}^n} \frac{1}{|z + 2\pi k|^{n+\alpha}}, \quad 0 < \alpha < 2.$$  (243)

Like in 1D it is sometimes convenient to use both open space and periodic representations of $\Lambda_\alpha$:

$$\Lambda_\alpha f(x) = p.v. \int_{\mathbb{T}^n} \phi(z) \delta_z f(x) \, dz = p.v. \int_{\mathbb{R}^n} \delta_z f(x) \frac{dz}{|z|^{n+\alpha}}.$$  (244)

Let us state the result.

**Theorem 9.2.** Consider the Euler Alignment System (148) on the torus $\mathbb{T}^n$ with kernel given by (243). There exists an $N \in \mathbb{N}$ such that for any sufficiently large $R > 0$, depending only on $\alpha$ and dimension $n$, any initial condition $(u_0, \rho_0) \in H^{2+\alpha}(\mathbb{T}^n) \times H^{2+\alpha}(\mathbb{T}^n)$, $m > \frac{2}{\gamma} + 3$, satisfying

$$|\rho_0|_\infty, |\rho_0^{-1}|_\infty, [u_0]_3, [\rho_0]_3 \leq R,$$  (245)

$$\mathcal{A}_0 \leq \frac{1}{R^N},$$

gives rise to a unique global solution in class $C_w([0,\infty); H^{2+\alpha} \times H^{2+\alpha})$. Moreover, the solution converges to a flocking state exponentially fast:

$$\mathcal{A}(t) + [u(t)]_1 + [u(t)]_2 + ||\rho(t) - \bar{\rho}(t)||_{C^1} < Ce^{-\delta t}.$$  

The proof of Theorem 9.2 establishes a uniform control on $C^2$-norm of $u$ and the distance between the initial density $\rho_0$ and its final profile $\bar{\rho}$. As a consequence, we obtain a stability result for flocking states.

**Theorem 9.3.** Let $(\bar{u}, \bar{\rho})$ be a traveling wave, where $\bar{\rho}(x,t) = \bar{\rho}(x - t\bar{u})$, and let $(v_0, r_0)$ be an initial data satisfying the conditions of Theorem 9.2. Suppose $|v_0 - \bar{u}|_\infty + |r_0 - \bar{\rho}|_\infty < \varepsilon$. Then the solution will converge to another flock $\bar{r}$ with $|\bar{r} - \bar{\rho}_0|_\infty < \varepsilon^\theta$, where $\theta \in (0,1)$ depends only on $\alpha$.

The idea of the proof is to establish control over a higher Hölder norm $|u|_{2+\gamma}$. This serves multiple purposes. First, it automatically shows boundedness of the gradients $|\nabla u|_\infty$ and $|\nabla \rho|_\infty$, where for the latter we need to control $|\nabla^2 u|_\infty$. So, we fulfill the continuation criterion of Theorem 7.3 and conclude global existence as stated. Second, with $C^{2+\gamma}$ norm uniformly bounded, we obtain exponential decay of $|u(t)|_1 + |u(t)|_2$ simply by interpolation with $\mathcal{A}$, which readily implies strong flocking as in Theorem 8.1. From the technical point of view we need $C^{2+\gamma}$ to overcome the singularity of the kernel for $\alpha > 1$ where $C^{1+\gamma}$ is insufficient.

From now on we will fix an exponent $0 < \gamma < 1$ to be identified later but dependent only on $\alpha$. The following notation will be used:

$$\delta_h f(x) = f(x + h) - f(x), \quad \tau_z f(x) = f(x + z)$$

$$\delta_h^2 f = \delta_h(\delta_h f), \quad \delta_h^3 f = \delta_h(\delta_h(\delta_h f)).$$
We adopt the finite difference definition of the Hölder metric:

\[ [f]_{2+\gamma} = \sup_{x,h \in \mathbb{T}^n} \frac{|\delta_h^3 f(x)|}{|h|^{2+\gamma}}. \]  

(246)

The equivalence of (246) to the classical norm \( [\nabla^2 f]_{\gamma} \) is a well known result in approximation theory, see \[?]\ For integer values of smoothness parameter \( k \in \mathbb{N} \) we use classical homogeneous metric \( [f]_k = [\nabla^k f]_{\infty} \).

The proof will be structured in several steps.

**Step 1: Breakthrough scenario.** According to Theorem 7.3 we have a local solution \((u, \rho) \in C_w((0, T_0) : H^{m+1} \times H^{m+\alpha})\) satisfying the assumptions of Theorem 9.2. Note that in view of the smallness assumption on \( \mathcal{A}_0 \), the norm \([u(t)]_{2+\gamma}\) will remain smaller than 1 at least for a short period of time. We thus consider a possible critical time \( t^* < T_0 \) at which the solution reaches size \( R \) for the first time:

\[ [u(t^*)]_{2+\gamma} = R, \quad [u(t)]_{2+\gamma} < R, \quad t < t^*. \]  

(247)

A contradiction is be achieved when we show that \( \partial_t [u(t^*)]_{2+\gamma} < 0 \). This establishes the bound \([u(t)]_{2+\gamma} < R\) on the entire interval \([0, T_0]\), and hence, the solution can be extended by Theorem 7.3.

In the course of the argument we pick \( \gamma \) based on several occurring restrictions, but ultimately depending only on \( \alpha \).

**Step 2: Preliminary estimates on \([0, t^*]\).** First we observe two simple bounds:

\[ [u(t)]_1, [u(t)]_2 < R^{1/6} e^{-c_0 t / R}, \quad \text{for all } t \leq t^*, \]  

(248)

provided \( R \) and \( N \) are sufficiently large, and \( N \) depending only on \( \alpha \). Indeed, in view of (245),

\[ [u]_1 \leq \mathcal{A}_0^{1+\gamma} [u]_{2+\gamma}^{1/2} \leq R^{1-N/2} e^{-c_0 t / R} < R^{1/6} e^{-c_0 t / R}, \]

and similarly,

\[ [u]_2 \leq \mathcal{A}_0^{2+\gamma} [u]_{2+\gamma}^{2} \leq R^{1-N} e^{-c_0 t / R} \leq R^{1/6} e^{-c_0 t / R}. \]

Next, we provide preliminary bounds on the density. Let us denote \( \underline{\rho} \) and \( \overline{\rho} \) the minimum and maximum of \( \rho \), respectively. Denote \( d = \nabla \cdot u \). Solving the continuity equation along characteristics, we estimate

\[ \rho_0 \exp \left\{ -\int_0^t |d(s)|_\infty \, ds \right\} \leq \underline{\rho}(t), \quad \overline{\rho}(t) \leq \rho_0 \exp \left\{ \int_0^t |d(s)|_\infty \, ds \right\}. \]

By (248), \( |d(s)|_\infty \leq R^{-3} e^{-c_0 s / R} \). Consequently,

\[ \int_0^t |d(s)|_\infty \, ds \leq c R^{-2} \leq \ln 2, \]

We have obtained the estimates

\[ \frac{1}{2R} \leq \underline{\rho}(t), \quad \overline{\rho}(t) \leq 2R. \]  

(249)

To get similar bounds for higher order derivatives of \( \rho \) we resort to the help of the \( e \)-quantity in the local well-posedness case. Note that the right hand side of the \( e \)-equation (166) is bounded by

\[ |(\nabla \cdot u)^2 - \text{Tr}(\nabla u)^2| \leq c |u|_2^2 \lesssim R^{-6} e^{-c_0 t / R}. \]

From (166) we obtain directly the estimate

\[ \frac{d}{dt} |e|_\infty \leq R^{-3} e^{-c_0 t / R} |e|_\infty + R^{-6} e^{-c_0 t / R}. \]

Again, by the Grönwall inequality, and using that \( |e_0|_\infty < R \) with \( R \) large enough,

\[ |e(t)|_\infty \leq 2R, \quad t < t^*. \]  

(250)
The estimate for $\nabla e$ follows a similar computation:

$$\frac{d}{dt}[e]_1 \lesssim [u]_1[e]_1 + [u]_2[e]_\infty + c[u]_1[u]_2.$$  

Using (248) we obtain

$$\frac{d}{dt}[e]_1 \lesssim R^{-3}e^{-\epsilon_0 t/R}[e]_1 + 2R^{-1}e^{-\epsilon_0 t/R} + cR^{-3}e^{-\epsilon_0 t/R}.$$  

Recall that initially $[e_0]_1 \leq [u_0]_2 + [\rho_0]_3 < 2R$. So, by the Grönwall inequality, we have

$$[e]_1 \leq 4R,$$

and hence, for $\alpha \neq 1$, we obtain

$$[\rho]_{1+\alpha} \leq 5R,$$

while for $\alpha = 1$,

$$[\Lambda \rho]_1 \leq 5R.$$

The latter does not guarantee a bound in $W^{2,\infty}$, however it implies bounds in other border-line classes such as Besov class $B^2_{\infty,\infty}$. It will be sufficient for what follows to reduce the exponent 2 by $\gamma > 0$, which will ultimately depend on $\alpha$ only, and cumulatively quote the case $\alpha \geq 1$ as

(252) $$[\rho]_{1+\alpha-\gamma} \leq C_\alpha R.$$  

**STEP 3: HIGHER ORDER NON-LOCAL MAXIMUM PRINCIPLE.** We establish another non-linear bound on the dissipation term – an adaptation of Lemma 8.9 to higher variations. Denote as before the dissipation term by

(253) $$D_\alpha f(x) = \int_{\mathbb{R}^n} |f(x + z) - f(x)|^2 \frac{dz}{|z|^{n+\alpha}}.$$  

**Lemma 9.4.** There is an absolute constant $c_0 > 0$ such that

(254) $$D_\alpha [\delta_3^h f](x) \geq c_0 \frac{|\delta_{\alpha}^3 f(x)|^{2+\alpha}}{|f|_2^\alpha |h|^{3\alpha}}.$$  

**Proof.** Let us fix a smooth cut-off function $\psi$, and fix an $r > 0$. We obtain

$$D_\alpha [\delta_{\alpha}^3 f](x) \geq \int_{\mathbb{R}^n} |\delta_{\alpha}^3 f(x)|^2 \frac{1 - \psi(z/r)}{|z|^{n+\alpha}} \frac{dz}{r^\alpha}.$$  

Notice that

$$\delta_3^h f(x + z) = \int_0^1 \int_0^1 \int_0^1 \nabla z \left( \frac{1 - \psi(z/r)}{|z|^{n+\alpha}} \right) \frac{dz}{r^\alpha}.$$  

Integrating by parts in $z$ once, and using the bound

$$|\nabla z \left( \frac{1 - \psi(z/r)}{|z|^{n+\alpha}} \right)| \leq \frac{c}{|z|^{n+\alpha+1}} \chi_{|z| > r},$$  

we obtain

$$\int_{\mathbb{R}^n} \delta_3^h f(x + z) \frac{1 - \psi(z/r)}{|z|^{n+\alpha}} \frac{dz}{r^\alpha} \leq C[u]_2 \frac{|h|^{3}}{r^{\alpha+1}}.$$  

We continue with the estimate:

$$D_\alpha [\delta_3^h f](x) \geq |\delta_3^h f(x)|^2 \frac{1}{r^\alpha} - C[f]_2 |\delta_3^h f(x)| \frac{|h|^{3}}{r^{\alpha+1}}.$$
Optimizing in \( r \) we pick \( r \sim \frac{\|f\|_{L^1}^2}{\|f\|_{L^1}^2} \) and this produces the desired estimate.

In view of the preliminary estimates and the assumptions of the breakthrough scenario, we consequently obtain

\[
(255) \quad \frac{1}{|h|^{4+2\gamma}} D_\alpha \delta_h^3 u(x) \geq \frac{R^{8+\alpha}}{|h|^{\alpha(1-\gamma)}}.
\]

**Step 4: Main Estimates.** With all preliminaries at hand we are now ready to use the equation to make estimate on the evolution of \([u]_{2+\gamma}\). Let \((x, h) \in \mathbb{T}^n\) be a pair for which the supremum is attained. We write the equation for the third finite difference:

\[
(256) \quad \partial_t \delta_h^3 u + \delta_h^3 (u \cdot \nabla u) = \int_{\mathbb{R}^n} \delta_h^3 [\rho(x + z) \delta_z u(x)] \frac{dz}{|z|^{n+\alpha}}.
\]

Denote

\[
B = \delta_h^3 (u \cdot \nabla u),
\]

\[
(257) \quad I = \int_{\mathbb{R}^n} \delta_h^3 [\rho(x + z) \delta_z u(x)] \frac{dz}{|z|^{n+\alpha}}.
\]

We will use the test-function \(\delta_h^3(u(x)/|h|^{4+2\gamma})\). Let us note the product formula

\[
\delta_h^3(fg) = \delta_h^3 f \tau_{3h} g + 3\delta_h^2 f \delta_h \tau_{2h} g + 3\delta_h f \delta_h^2 \tau_{h} g + f \delta_h^3 g.
\]

So, for the B-term we obtain

\[
B = \delta_h^3 u \cdot \tau_{3h} \nabla u + 3\delta_h^2 u \cdot \delta_h \tau_{2h} \nabla u + 3\delta_h u \cdot \delta_h^2 \tau_{h} \nabla u + u \cdot \nabla \delta_h^3 u.
\]

Note that the last term vanishes due to criticality. Thus, we can estimate

\[
\frac{1}{|h|^{2+\gamma}} |B| \leq [u]_{2+\gamma} [u]_1 + 3|h|^{1-\gamma} |u|_2 + 3[u]_1 [u]_{2+\gamma} \lesssim [u]_{2+\gamma} [u]_1 + |h|^{1-\gamma} |u|_2.
\]

Multiplying by another \([u]_{2+\gamma} \leq R\) and using (248) we obtain

\[
(258) \quad \frac{|u|_{2+\gamma}}{|h|^{2+\gamma}} |B| \lesssim R^{-1} + R^{-5} < 1.
\]

We now turn to the I-term which contains dissipation. The integrand is given by \(\delta_h^3[\tau_\rho \delta_z u]\). So, we expand similarly using commutativity \(\partial_t \partial_z = \partial_z \partial_t\):

\[
\delta_h^3[\tau_\rho \delta_z u] = \delta_h^3 \tau_\rho \tau_{3h} \delta_z u + 3\delta_h^2 \tau_\rho \tau_{2h} \delta_z u + 3\delta_h \tau_\rho \tau_h \delta_z u + \tau_\rho \delta_z \delta_h^3 u.
\]

Multiplying by \(\delta_h^3 u\) the last term provides dissipation:

\[
\tau_\rho \delta_z \delta_h^3 u \delta_h^3 u \leq -\frac{1}{2} \rho |\delta_z \delta_h^3 u|^2.
\]

Dividing by \(|h|^{4+2\gamma}\) and using (255) we obtain the lower bound

\[
(260) \quad \frac{1}{2} |h|^{4+2\gamma} \rho D_\alpha \delta_h^3 u(x) \geq \frac{R^8}{|h|^{\alpha(1-\gamma)}}.
\]

It is clear that the transport term treated in (258) is entirely absorbed into dissipation at the time \(t^*\):

\[
\partial_t [u]_{2+\gamma} \leq -\frac{R^7}{|h|^{\alpha(1-\gamma)}} \frac{|h|^{4+2\gamma} J}{|h|^{\alpha(1-\gamma)}} J,
\]

where \(J\) contains all the remaining three terms of \(I\):

\[
J = \int_{\mathbb{R}^n} \delta_h^3 \tau_\rho \tau_{3h} \delta_z u + 3\delta_h^2 \tau_\rho \tau_{2h} \delta_z u + 3\delta_h \tau_\rho \tau_h \delta_z u \frac{dz}{|z|^{n+\alpha}} = J_1 + 3J_2 + J_3.
\]
It remains to estimate each of the remaining $J_i$ terms. Specifically, we will aim to obtain bounds of the form

$$\frac{1}{|h|^{2+\gamma}} |J_i| \lesssim \frac{|h|^\varepsilon}{|h|^{\alpha(1-\gamma)}},$$

for some $\varepsilon > 0$ and $\gamma$ is sufficiently small. This makes the dissipation absorb all the remaining $J$-terms in the equation.

We begin with $J_2$. For $\alpha < 1$, let us use (248) and (251) to infer

$$\delta_h^2 \tau_z \rho \leq \rho |1+\alpha| |h|^{1+\alpha} \lesssim R |h|^{1+\alpha}$$

and

$$\tau_2 h \delta_h \delta_z u \leq |u|_2 |h| \min\{|z|, 1\} \lesssim R^{-1} |h| \min\{|z|, 1\}.$$ Thus, the singularity is removed and we get

$$\frac{1}{|h|^{2+\gamma}} |J_2| \lesssim |h|^\alpha,$$

which clearly implies (261) for sufficiently small $\gamma$. In the case $\alpha \geq 1$ we first perform symmetrization

$$J_2 = \frac{1}{2} \int_{\mathbb{R}^n} \left[ \delta_h^2 (\tau_z - \tau_z) \rho \tau_2 h \delta_h \delta_z u + \delta_h^2 \tau_z \rho \tau_2 h (\delta_z + \delta_z u) \right] \frac{dz}{|z|^{n+\alpha}}.$$ For the first summand we use (252):

$$\delta_h^2 (\tau_z - \tau_z) \rho \leq R \min\{|h|^{1+\alpha-\gamma}, |h|^{\alpha-\gamma} |z|\}$$

$$\tau_2 h \delta_h \delta_z u \leq R^{-1} |h| \min\{|z|, 1\}.$$ with this we can estimate

$$\frac{1}{|h|^{2+\gamma}} \int_{\mathbb{R}^n} |\delta_h^2 (\tau_z - \tau_z) \rho \tau_2 h \delta_h \delta_z u| \frac{dz}{|z|^{n+\alpha}} \leq \frac{|h|^{1+\alpha-\gamma}}{|h|^{2+\gamma}} \leq \frac{|h|^{\alpha-2\gamma - 1+\alpha(1-\gamma)}}{|h|^{\alpha(1-\gamma)}}.$$ Clearly, $\alpha - 2\gamma - 1 + \alpha(1-\gamma) > 0$. In the second summand, using that $(\delta_z + \delta_z u)$ is the second difference,

$$\delta_h^2 \tau_z \rho \tau_2 h (\delta_z + \delta_z u) \leq |h|^{2-\gamma} \min\{|z|^2, 1\}$$

we obtain

$$\frac{1}{|h|^{2+\gamma}} \int_{\mathbb{R}^n} |\delta_h^2 \tau_z \rho \tau_2 h (\delta_z + \delta_z u)| \frac{dz}{|z|^{n+\alpha}} \leq \frac{|h|^{2-\gamma}}{|h|^{2+\gamma}} \leq \frac{|h|^{\alpha(1-\gamma)-2\gamma}}{|h|^{\alpha(1-\gamma)}}.$$ This finishes the bound on $J_2$.

As to $J_3$ we proceed similarly. For $\alpha < 1$, we use

$$|\delta_h \tau_z \rho \tau_h \delta_h^2 \delta_z u| \leq |h|^2 \min\{|z|, 1\}.$$ Hence, we have

$$\frac{1}{|h|^{2+\gamma}} |J_3| \lesssim \frac{|h|^2}{|h|^{2+\gamma}} \leq \frac{h^{\alpha(1-\gamma)-\gamma}}{|h|^{\alpha(1-\gamma)}}.$$ The power on the bottom is positive for sufficiently small $\gamma$. For $\alpha \geq 1$, we again perform symmetrization first

$$J_3 = \frac{1}{2} \int_{\mathbb{R}^n} \left[ \delta_h (\tau_z - \tau_z) \rho \tau_h \delta_h^2 \delta_z u + \delta_h \tau_z \rho \tau_h \delta_h^2 (\delta_z + \delta_z) u \right] \frac{dz}{|z|^{n+\alpha}}.$$ Thus,

$$|\delta_h (\tau_z - \tau_z) \rho \tau_h \delta_h^2 \delta_z u| \leq \min\{|h|^{3+\alpha-\gamma}, |h|^{1+\alpha} |z|^2\},$$

$$|\delta_h \tau_z \rho \tau_h \delta_h^2 (\delta_z + \delta_z) u| \leq \min\{|h|^3, |h| |z|^2\}.$$ The first term leads to an estimate as before. For the second, we split the integration into subregions $|z| < r$ and $|z| > r$ to obtain the bound $|h|^{2-\alpha} + |h|^3 r^{-\alpha}$. Setting $r = |h|$ leads to further estimate by $|h|^{3-\alpha}$ which implies (261).
Estimate on $J_1$ is slightly more involved. Let us start with the case $\alpha \geq 1$. We symmetrize first as usual

$$ J_1 = \frac{1}{2} \int_{\mathbb{R}^n} [\delta^3_h(x) - \tau_x \rho - \tau_z \rho \tau_z \delta_z u + \delta^3_h \tau_z \rho \tau_z \delta_z u] \frac{dz}{|z|^{n+\alpha}}. $$

For the first summand we use

$$ |\delta^3_h(x) - \tau_x \rho - \tau_z \rho \tau_z \delta_z u| \leq |h|^{\alpha-\gamma} \min\{|z|^2, |h|\}. $$

For the second summand,

$$ |\delta^3_h \tau_z \rho \tau_z \delta_z u + \delta_z \delta_z u| \leq |h|^{1+\alpha-\gamma} \min\{|z|^2, 1\}. $$

So, this is estimated as earlier. In fact one can observe that the estimates above extend to the range $\alpha > \frac{1}{2}$, but not all the way to zero. The problem is that the density receives all the variations in $h$ and not fully utilizes them, while $u$ can no longer directly contribute. So, we switch one $h$-difference back onto $u$. We do it by starting from the original formula

$$ J_1 = \int_{\mathbb{R}^n} \delta^3_h \tau_x \rho(x) \tau_z \delta_z u(x) \frac{dz}{|z|^{n+\alpha}}. $$

Over the domain $|z| < 10|h|$ we estimate using the same cut-off function $\psi$ as before,

$$ \int_{\mathbb{R}^n} |\delta^3_h \tau_x \rho(x) \tau_z \delta_z u(x)| \psi \left( \frac{z}{10|h|} \right) \frac{dz}{|z|^{n+\alpha}} \leq \int_{|z| < 10|h|} |h|^{1+\alpha} \frac{dz}{|z|^{n+\alpha-1}} \leq |h|^2. $$

This culminates into (261). For the remaining part, denote for clarity $f = \delta^3_h \rho$. So, $\delta^3_h \tau_x \rho(x) = f(x + h + z) - f(x + z)$. Let us write

$$ \int_{\mathbb{R}^n} (f(x + h + z) - f(x + z)) \tau_z \delta_z u(x) \frac{(1 - \psi(\frac{z}{10|h|}))}{|z|^{n+\alpha}} dz $$

$$ = \int_{\mathbb{R}^n} f(x + z) \tau_z \delta_z - h u(x) \frac{(1 - \psi(\frac{z-h}{10|h|}))}{|z-h|^{n+\alpha}} - \tau_z \delta_z u(x) \frac{(1 - \psi(\frac{z}{10|h|}))}{|z|^{n+\alpha}} dz $$

$$ = \int_{\mathbb{R}^n} f(x + z) \tau_z \delta_z - h u(x) \frac{(1 - \psi(\frac{z-h}{10|h|}))}{|z-h|^{n+\alpha}} - \tau_z \delta_z u(x) \frac{(1 - \psi(\frac{z}{10|h|}))}{|z|^{n+\alpha}} dz $$

$$ - \int_{\mathbb{R}^n} f(x + z) \tau_z \delta_z u(x) \frac{(1 - \psi(\frac{z-h}{10|h|}))}{|z-h|^{n+\alpha}} - \frac{(1 - \psi(\frac{z}{10|h|}))}{|z|^{n+\alpha}} dz $$

The integrals are still supported on $|z| > 9|h|$, where $|z - h| \sim |z|$. Estimating the first component we use bounds

$$ |\delta_z - h u(x) - \delta_z u(x)| = |u(x + z - h) - u(x + z)| \leq |h| $$

$$ |f(x + z)| \leq |h|^{1+\alpha}. $$

thus,

$$ \left| \int_{\mathbb{R}^n} f(x + z) \tau_z \delta_z - h u(x) \frac{(1 - \psi(\frac{z-h}{10|h|}))}{|z-h|^{n+\alpha}} \frac{dz}{|z|^{n+\alpha}} \right| \leq |h|^{2+\alpha} \int_{|z| > 9|h|} \frac{dz}{|z|^{n+\alpha}} \leq |h|^2, $$

which implies (261). Finally, for the second component we use

$$ \frac{(1 - \psi(\frac{z-h}{10|h|}))}{|z-h|^{n+\alpha}} - \frac{(1 - \psi(\frac{z}{10|h|}))}{|z|^{n+\alpha}} \leq \frac{|h|}{|z-h|^{n+\alpha+1}} \leq \frac{|h|}{|z|^{n+\alpha+1}} |I_{|z| > 9|h|} \leq \frac{|h|}{|z|^{n+\alpha+1}} |I_{|z| > 9|h|}, $$

and

$$ |f(x + z) \tau_z \delta_z u(x)| \leq |h|^{1+\alpha} |z|. $$

Integration reproduces the same bound as for the first part.
We have established that $\partial_t [u(t^*)]_2^{2+\gamma} < 0$ at the critical time. This means that such time $t^*$ does not exist and which finishes the proof of existence part.

**Step 5: Flocking and Stability.** As we noted in the beginning, exponential decay of $[u]_2^2$ implies uniform control over $|\nabla \rho|_\infty$.

Arguing as in the proof of Theorem 8.1 we can slightly improve the space in which string flocking occurs. This is due to (251) - (252) bounds, which imply that $\bar{\rho} \in W^{1+\alpha-\gamma, \infty}$ by compactness. Using again (252) and by interpolation we have convergence in the $W^{1,\infty}$-metric as well:

$$[\rho(\cdot, t) - \bar{\rho}(\cdot, t)]_1 < C_2 e^{-\delta t}.$$  

As far stability is concerned, the computation above shows that in fact the limiting profile $\bar{\rho}$ differs little from initial density $\rho_0$ under the conditions of Theorem 9.3. Indeed, setting $R$ such that $\varepsilon = 1/R^N$ (here $\varepsilon > 0$ is small), we obtain via (248),

$$|\partial_t r|_\infty \leq C R^{-2} e^{-c_0 t/R}.$$  

Hence, $|\bar{r} - r_0|_\infty \leq \frac{C}{c_0 R} = \varepsilon^\theta$. Since $|r_0 - \bar{\rho}_0|_\infty < \varepsilon$, this finishes the result.

**9.5. Notes and remarks.** The two notable exceptions are the result of Ha, et al [?] demonstrating global existence in the case of smooth communication kernel $\phi$ with small initial data in higher order Sobolev spaces, $\|u\|_{H^{s+1}} < \varepsilon_0$, where $\varepsilon_0$ depends on $\|\rho_0\|_{H^s}$;

**References**


