ABSTRACT

Title of dissertation: MULTISCALE PROBLEMS ON COLLECTIVE DYNAMICS AND IMAGE PROCESSING: THEORY, ANALYSIS AND NUMERICS

Changhui Tan, Doctor of Philosophy, 2014

Dissertation directed by: Professor Eitan Tadmor
Department of Mathematics
Institute for Physical Science & Technology
Center for Scientific Computation and Mathematical Modeling

Multi-scale problems appear in many contexts. In this thesis, we study two different subjects involving multi-scale problems: (i) collective dynamics, and (ii) image processing.

For *collective dynamics*, we concentrate on flocking models, in particular, Cucker-Smale and Motsch-Tadmor systems. These models characterize the emergent behaviors of self-organized dynamics. We study flocking systems in three different scales, from microscopic agent-based models, through mesoscopic kineitc discriptions, to macroscopic fluid systems. Global existence theories are developed for all three scales, with the proof of asymptotic flocking behaviors. In the macroscopic level, a critical threhold phenomenon is addressed to obtain global regularity. Similar idea is implemented to other fluid systems as well, like Euler-Poisson equations. In the kinetic level, a discontinuous Galerkin method is introduced to overcome the numerical difficulty due to the precence of δ-singularity.

For *image processing*, we apply the idea of multi-scale image representation to construct uniformly bounded solutions for $div\ U = F$. Despite the fact that the equation is simple and linear, it is suprisingly true that its bounded solution can not be constructed through a linear procedure. In particular, the Holmholtz solution $U = \nabla\Delta^{-1}F$ is *not* always bounded. A hierarchical construction of the bounded solution of the equation is proposed, borrowing the idea from image processing. We also present a numerical implementation to deal with the highly nonlinear construction procedure. Solid numerical result verifies that the constructed solution is indeed uniformly bounded.

MULTISCALE PROBLEMS ON COLLECTIVE DYNAMICS AND IMAGE PROCESSING: THEORY, ANALYSIS AND NUMERICS

by

Changhui Tan

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2014

Advisory Committee:

Professor Eitan Tadmor, Chair/Advisor
Professor Howard Elman
Professor Pierre-Emmanuel Jabin
Professor Dave Levermore
Professor Antoine Mellet

To my wonderful parents, and my beloved wife.
Thanks for your consistent support and love!

Acknowledgments

I owe my gratitude to all the people who have made this thesis possible and because of whom my graduate experience has been one that I will cherish forever.

First and foremost I would like to express my deepest appreciation to my advisor, Professor Eitan Tadmor. Without his advice and support over the years, this thesis would never exist. Thank you!

I want to thank my committees, Professor Howard Elman and Professor Dave Levermore, for consistent support from my preliminary to my defense, as well as Professor Pierre-Emmanuel Jabin and Professor Antoine Mellet for valuable helps and discussions. I am grateful to Professors Laurent Desvillettes, Alex Kiselev, Hailiang Liu, Jian-Guo Liu, Ronghua Pan and Konstantina Trivisa, who offer me great help and encouragement.

I would also like to thank CSCAMM for providing me a comfortable working environment. Many thanks to my colleagues at CSCAMM, Dr. Hantaek Bae, Dr. Trygve Karper, Prof. Sebastien Motsch, and in particular, my collaborator Dr. Thomas Rey. I want to acknowledge the NSF Ki-Net research network for providing lots of activities and financial support.

Thanks to Chunting, Feng, Jinxin, Jong-Jun, Miao, Minyu, Qin, Ran, Rongrong, Tong, Xuan, Yongyong and Yujie. It is great to have you in these years.

Finally, I thank my parents, Hua and Xiaodong, who have always stood by me and given me endless love and support. Words cannot express the gratitude I owe them. And most importantly, I would like to thank my lovely wife Wei. I am so lucky to have you, and your love means everything to me.

Table of Contents

List of Tables

List of Figures

Part I

Collective Dynamics

Collective behaviors are common in nature and human societies. It refers to social processes and events which emerge to global structures in a spontaneous way. Many biological examples reflect collective behaviors, including the dynamics of a flock of birds, a school of fish, a colony of ants, a swarm of bacteria, even opinions for a group of people. In these systems, Each individual only accesses limited environmental information to determine their motions. Yet the whole systems self-organize into global patterns: flock, mill, concentration, consensus, clusters, etc.

Modeling collective behaviors brings challenges to scientists and mathematicians. The goal is to impose simple interaction rules between individuals, which lead to desired global structures. A celebrated 3-zone interaction framework is widely accepted in the models of collective dynamics, consisting long range attraction, short range repulsion and mid range alignment. We shall start this part with a survey on models of collective dynamics, under this framework (section 1.1).

In particular, we are interested in flocking models, where only mid range zone is highlighted. The *flocking* phenomenon characterizes the fact that self-propelled individuals organize into an ordered motion. The pioneering Cucker-Smale model and a normalized Motsch-Tadmor model are introduced in section 1.2, alongside with other inspiring models. With simple interaction rules, these two models enjoy unconditional flocking property: the flocking phenomenon is detected with all initial configurations.

We study flocking models in different scales. As the number of individuals becomes larger and larger, it is costly to trace the dynamics of each individual. Instead, we study the mean-field phase diagram of the systems, governed by Vlasov-type kinetic equations. The elegant kinetic theory, originally developed for gas dynamics and other physics sys-

tems, is well adapted to flocking dynamics. Under this *mesoscopic* scale, global existence theory and unconditional flocking property are addressed in section 1.3.1-1.3.2. From the numerical prospective, flocking property implies a concentration in velocity as time approaches infinity. Such δ-singularity need to be treated carefully to maintain high accuracy and stability, as the solution becomes more and more singular. Chapter 3 is devoted to design numerical schemes to attack the δ-singularity, and achieve asymptotic flocking behavior for the numerical solution.

A hydrodynamic representation of flocking models can be derived at least formally (see section 1.3.3) by taking moments of the kinetic systems. In this *macroscopic* scale, flocking systems can be viewed as compressible Eulerian dynamics, with nonlocal alignment forcing. With the presence of nonlinear convection, global existence theory is way more complicated, comparing with kinetic systems. Nevertheless, we prove in section 1.3.4 that *"strong solution must flock"*: thus, a global existence theory directly implies flocking behavior.

We study global existence of strong solutions for macroscopic flocking models in chapter 2. For general compressible Eulerian dynamics, there is finite time shock formations, due to nonlinear convection. To achieve global regularity, one should hope the external forcing regularizes the system and competes with the convection. In the case of macroscopic flocking systems, the alignment forcing is rather weak. Moreover, the nonlocal nature of the forcing adds more difficulty to the analysis. We prove a *critical threshold* result for the macroscopic system: subcritical initial configurations guarantee global existence of strong solutions, while supercritical initial configurations lead to finite time blowup. It is remarkable that the theory can be extended to two space dimension

(section 2.4), as well as vacuum stage (section 2.5), while most of the celebrating results on Eulerian dynamics are presented in 1D or scaler equation, and vacuum is often avoided.

A prototype problem for critical thresholds phenomenon is studied as well: Euler-Poisson equations, sharing the similar strategy applied on macroscopic flocking models. While the global regularity for 2D Euler-Poisson system is still open, it can be achieved with a slight restriction.

Main results presented in this part follow from the work in [95] for the macroscopic scale, and [97] for the mesoscopic scale. We shall also list key references [33, 45, 64, 72].

Chapter 1: Modeling Collective Dynamics

In this chapter, we discuss various models on collective dynamics. The philosophy of modeling collective dynamics is to use simple "local" interaction rules to characterize "global" structures.

Many systems can be represented under different scales. The starting point is the *microscopic* representation, or agent-based models, where the dynamics of each individual is traced. Though agent-based models are often intuitive, it is not efficient to keep track of each agent if the group size is large. *Macroscopic* representation is introduced to capture the mean-field dynamics of the system. A much smaller in size PDE system is used to replace the huge ODE system, which preserves various of conservations. For second order systems, there is an intermediate *mesoscopic* representation to link the microscopic world to the macroscopic one. Powerful kinetic theory enters in this level. All three scales have their own mathematical interests and challenges.

Among different models on collective dynamics, we are interested in systems reflecting an intriguing *flocking* phenomenon. Two models, Cucker-Smale and Motsch-Tadmor, are studied in all three scales, where flocking phenomenon is detected and proved, under suitable assumptions.

We organize this chapter as follows. We start with an introduction of modeling

collective dynamics in section 1.1. Then several flocking models are presented in 1.2. We continue with multi-scale representations for flocking models in 1.3, featuring on the flocking phenomenon. More discussions on macroscopic and kinetic models are given in chapter 2 and 3, respectively.

1.1 Background and motivation

Everyone must observe some beautiful swinging movements of animals like birds, fishes, insects and many more. They often emerge into surprised and astonished structures, without a clear indication of any role of a leader. There is a constant rising of interests in understanding these remarkable phenomena to biologists, physicists, ecologists, as well as mathematicians. Scientists are motivated to study various of collective behaviors, by proposing models consisting simple rules, yet capturing the global structures. Successful models are not limited to biological systems. Many other intriguing sociological behaviors have been studied as well.

1.1.1 Collective dynamics in biology and beyond

There are many successful mathematical models on collective dynamics, in the biological contexts. Examples include a flock of birds flying towards the same direction [5, 19, 26, 72], a school of fishes forming a specific pattern [2, 8, 30, 47, 77], a group of locusts rolling across the farmland [14, 99], a colony of bacteria and social insects swarming into some structures [10, 13, 52, 56, 100], etc.

Besides biological systems, models for collective dynamics have appeared in a large

variety of other contexts, including ecological models [42], multi-agent robot networks [49, 50], opinion dynamics [17, 57, 73], traffic and pedestrian networks [3, 46, 80], and more.

All these systems share relatively simple mathematical structures, while gracefully capture phenomena observed in the nature. Powerful mathematics kicks in to provide a clearer picture of the dynamics.

1.1.2 The three-zone framework on modeling collective dynamics

One simple elegant way to model collective dynamics is to propose rules on interactions between each pair of individuals in the group. It is natural to assume that the interaction depends on the physical distance between the individuals. A widely accepted three-zone framework proposed by Reynolds in [82] is used on modeling the interactions.

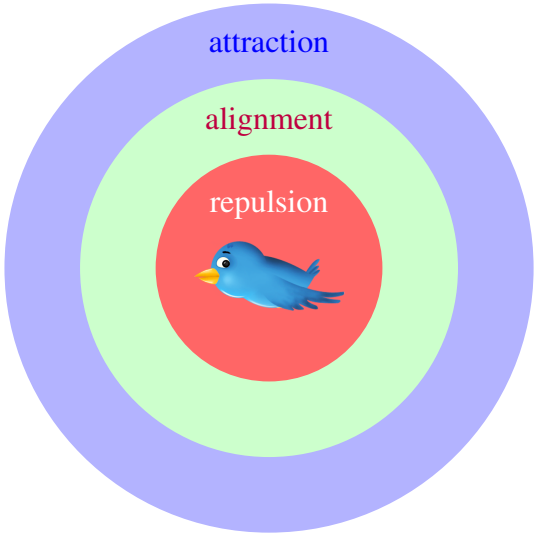

Figure 1.1: The three-zone framework: short range repulsion, mid range alignment and long range attraction.

As illustrated in figure 1.1, the interactions are divided into three zones, under dif-

ferent rules. In the short range, individuals have the tendency to move apart from others to avoid collision, thus the interaction is *repulsive*. In the mid range, repulsion is substituted by *alignment*, where individuals try to align with the direction of its neighbors. If the individual is far part from the group, it will be *attracted* by others relatively far away and try to stay closer to the group.

The attractive and repulsive interactions can be modeled through an interaction potential $\Psi = \Psi(r)$, which decays when r is small (representing repulsion), and increases when r is large (representing attraction). The agent-based dynamics reads

$$\dot{\mathbf{x}}_i = \frac{1}{N}\sum_{i=1}^{N} \nabla_{\mathbf{x}_j}\Psi(|\mathbf{x}_i - \mathbf{x}_j|). \tag{1.1}$$

The attraction-repulsion system above has been well-studied, through its mean field hydrodynamic representation [4, 18, 31, 100, 101]. For more discussions on this type of models, we refer to a review [87] and references therein.

1.2 Models on flocking and alignment

We turn our main focus onto modeling mid range alignment interactions, starting from this section. The key phenomenon to capture in the modeling prospectus is velocity matching, *i.e.*, flocking.

1.2.1 Cucker-Smale model

We start with the pioneering work of Cucker and Smale [26, 27]. The model reads

$$\dot{\mathbf{x}}_i = \mathbf{v}_i, \quad \dot{\mathbf{v}}_i = \frac{1}{N}\sum_{j=1}^{N} \phi_{ij}(\mathbf{v}_j - \mathbf{v}_i). \tag{1.2}$$

Here, ϕ_{ij} indicates the influence from individual j to individual i. It is natural to measure the strength of influence by the physical distance between the two individuals,

$$\phi_{ij} = \phi(|\mathbf{x}_i - \mathbf{x}_j|). \tag{1.3}$$

$\phi = \phi(r)$ is called the *influence function*, which is bounded, non-negative and decreasing in r. Without loss of generosity, we can assume $\|\phi\|_{L^\infty(\mathbb{R}^+)} = \phi(0) = 1$.

System (1.2) celebrates the conservation of total momentum, due to symmetry of interaction ϕ_{ij},

$$\frac{d}{dt}\left(\sum_{i=1}^N \mathbf{v}_i\right) = \frac{1}{N}\sum_{i=1}^N\sum_{j=1}^N \phi_{ij}(\mathbf{v}_j - \mathbf{v}_i) = 0.$$

It implies the average velocity $\bar{\mathbf{v}} = \dfrac{1}{N}\displaystyle\sum_{i=1}^N \mathbf{v}_i$ is conserved in time.

Moreover, symmetry also implies dissipation in l^2 variation of velocity

$$\frac{d}{dt}\left(\sum_{i=1}^N |\mathbf{v}_i - \bar{\mathbf{v}}|^2\right) = -\frac{1}{N}\sum_{i=1}^N\sum_{j=1}^N \phi_{ij}|\mathbf{v}_j - \mathbf{v}_i|^2 \le 0. \tag{1.4}$$

Therefore, the equilibrium state for Cucker-Smale system $(\mathbf{x}_i^*, \mathbf{v}_i^*)$ satisfies either or the following

$$\phi_{ij} = 0, \quad \text{or} \quad \mathbf{v}_i = \mathbf{v}_j, \quad \forall\, i, j = 1, \cdots, N.$$

The large time behavior of the system would be *cluster formation*. Individuals in the same cluster will align with each other, $\mathbf{v}_i - \mathbf{v}_j \to 0$ as $t \to +\infty$. Individuals from different clusters will have no influence with each other asymptotically, $\phi_{ij} \to 0$.

In particular, if the influence function is strong enough in the far field, namely it has a diverging tail (c.f. [44, 73]),

$$\int^\infty \phi(r)dr = +\infty,$$

then, the number of clusters is guaranteed to be 1, where all individuals align with others,

$$\mathbf{v}_i(t) \xrightarrow{t\to\infty} \bar{\mathbf{v}}, \quad \forall\, i = 1, \cdots, N,$$

and stay within a finite distance with respect to the center $\bar{\mathbf{x}}(t) = \bar{\mathbf{v}}t$,

$$|\mathbf{x}_i(t) - \bar{\mathbf{x}}(t)| \leq C, \quad \forall\, t \geq 0.$$

It yields the so-called *flocking phenomenon*, where (i) the distance between individuals are bounded for all time, and (ii) all individuals have the same asymptotic velocity. Figure 1.2 is a simulation of 1D Cucker-Smale system (1.2). The flocking properties can be observed: (i) in left graph, and (ii) in right graph. We postpond more discussions to section 1.2.3.

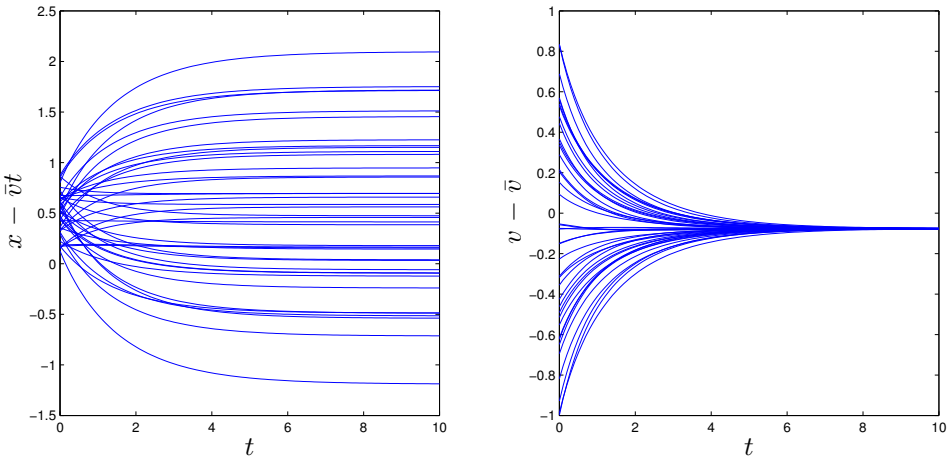

Figure 1.2: A 1D simulation to illustration the flocking phenomenon. Left graph shows $\{\mathbf{x}_i(t) - \bar{\mathbf{x}}(t)\}_i$ are bounded in all time. Right graph shows $\{\mathbf{v}_i(t) - \bar{\mathbf{v}}\}_i$ converge to 0 as time approaches infinity.

1.2.2 Motsch-Tadmor model

Cucker-Smale model beautifully captures the flocking phenomenon. However, there is a drawback of this model which is worth mentioning, consulting [72].

Figure 1.3 states an example, discussed in [72, Section 2.1], where Cucker-Smale model provides "unreasonable" dynamics. Suppose there is a small group G_1 with N_1 individuals and a large group G_2 with N_2 individuals which is far from G_1. The interactions between individuals in G_1 and G_2 are weak. In particular, if the distance is larger than the length of the support of ϕ, there is no interaction between G_1 and G_2. As the total number of individuals is $N = N_1 + N_2$, the Cucker-Smale model (1.2) has the following dynamics of $\mathbf{v}_i$ for $i \in G_1$.

$$\dot{\mathbf{v}}_i = \frac{1}{N_1 + N_2} \sum_{j=1}^{N_1} \phi_{ij}(\mathbf{v}_j - \mathbf{v}_i), \quad \forall\, i \in G_1.$$

As $N_2 \gg N_1$, $\dot{\mathbf{v}}_i \approx 0$. Hence, the invisible large group far away almost halts the internal interactions within G_1.

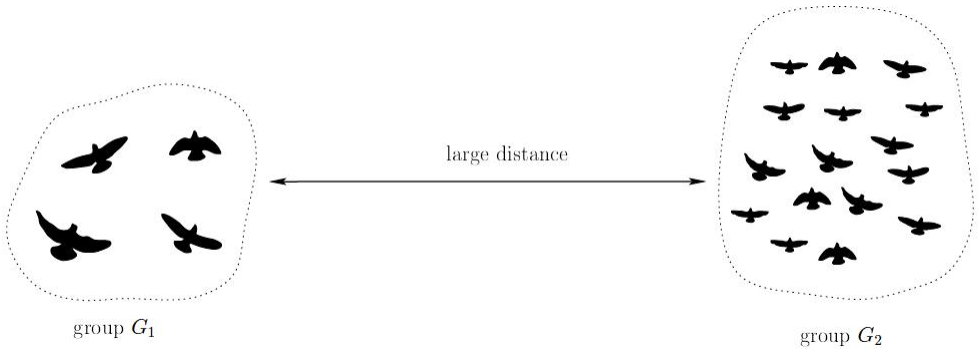

Figure 1.3: A drawback of Cucker-Smale model when the configuration is far from equilibrium [72]: the interactions between individuals within the small group G_1 becomes almost zero due to a large group G_2 far away.

Motsch and Tadmor in [72] proposed the following new model to overcome such

drawback of Cucker-Smale model.

$$\dot{\mathbf{x}}_i = \mathbf{v}_i, \quad \dot{\mathbf{v}}_i = \frac{1}{\Phi_i}\sum_{j=1}^{N}\phi_{ij}(\mathbf{v}_j - \mathbf{v}_i), \quad \Phi_i = \sum_{j=1}^{N}\phi_{ij}. \tag{1.5}$$

The total interaction Φ_i for an individual i is used to normalize the system instead of the total number of individuals N.

Consider the same example illustrated in figure 1.3. The dynamics of $\mathbf{v}_i$ under (1.5) for $i \in G_1$ reads

$$\dot{\mathbf{v}}_i = \frac{1}{\Phi_i}\sum_{j=1}^{N_1}\phi_{ij}(\mathbf{v}_j - \mathbf{v}_i) \approx \frac{1}{N_1}\sum_{j=1}^{N_1}\phi_{ij}(\mathbf{v}_j - \mathbf{v}_i), \quad \forall\, i \in G_1,$$

where $\Phi_i \approx N_1$ as $\phi_{ij} \approx 1$ for $j \in G_1$ and $\phi_{ij} = 0$ for $j \in G_2$. This yields a much more reasonable dynamics as the invisible group G_2 should have no effect on the dynamics of individuals in G_1.

The dynamics of $\mathbf{v}_i$ in (1.5) can be also written as $\dot{\mathbf{v}}_i = \tilde{\mathbf{v}}_i - \mathbf{v}_i$, where $\tilde{\mathbf{v}}_i = \sum_{j=1}^{N}\frac{\phi_{ij}}{\Phi_i}\mathbf{v}_j$ is the *environmental average velocity* corresponding to individual i. It is a convex combination of $\{\mathbf{v}_j\}$. Every individual tends to align with its environment average velocity.

The normalization Φ_i breaks the symmetry of the system. Hence, conservation of momentum is not valid for (1.5). It is more realistic but less mathematical friendly. For instance, the l^2 variation of velocity is no longer dissipative. Nevertheless, flocking phenomenon is still observed from simulations, with the result similar to what is illustrated in figure 1.2. And it could be proved (e.g. [72, 73]) under l^∞ framework.

1.2.3 Unconditional flocking

In this section, we discuss flocking properties for Cucker-Smale system (1.2) and Motsch-Tadmor system (1.5).

We start with a concrete mathematical definition of flocking, for an agent-based system.

Definition 1.2.1 (Flocking)**.** We say an multi-agent dynamical system $\{(\mathbf{x}_i, \mathbf{v}_i)\}$ converges to a flock if

(i) Variation in position $S(t)$ stays bounded for all time.

$$S(t) := \|\{|\mathbf{x}_i(t) - \bar{\mathbf{x}}(t)|\}_i\| \leq D \quad \forall\, t \geq 0.$$

(ii) Variation in velocity $V(t)$ decays to 0 as time approaches infinity.

$$V(t) := \|\{|\mathbf{v}_i(t) - \bar{\mathbf{v}}(t)|\}_i\| \xrightarrow{t\to\infty} 0.$$

The norm $\|\cdot\|$ on $\mathbb{R}^N$ has not been specified in the definition. Indeed, flocking properties can be derived under l^1, l^2 or l^∞ frameworks, consulting [26, 44, 45, 72, 73].

Here, we define l^p norm in the usual fashion.

$$\|\mathbf{z}\|_{l^p} = \begin{cases} \left(\dfrac{1}{N}\displaystyle\sum_{i=1}^{N} |z_i|^p\right)^{\frac{1}{p}} & p < \infty, \\ \max\limits_{1\leq i\leq N} |z_i| & p = \infty. \end{cases}$$

The starting point is the Cucker-Smale system (1.2), which enjoys the dissipation property (1.4). It implies the following estimate

$$\frac{d}{dt}\sum_{i=1}^{N} |\mathbf{v}_i(t) - \bar{\mathbf{v}}|^2 \leq -\min_{ij} \phi_{ij}(t) \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N} |\mathbf{v}_j(t) - \mathbf{v}_i(t)|^2 = -2\min_{ij} \phi_{ij}(t) \sum_{i=1}^{N} |\mathbf{v}_i(t) - \bar{\mathbf{v}}|^2.$$

Hence, under l^2 framework, $\dot{V}(t) \leq -\min_{ij} \phi_{ij}(t) V(t)$. In particular, if ϕ has a global lower bound $\gamma > 0$, then $\dot{V} \leq -\gamma V$, which implies exponential decay of V. Moreover, $|\dot{S}| \leq V$ regardless of the norm picked. It clearly implies boundedness of S in all time.

Remark 1.2.1. The flocking property stated above is satisfied regardless of the choice of initial configurations. We call this *unconditional flocking*, as suggested in [72]. In contrast, *conditional flocking* occurs when definition 1.2.1 is satisfied only for some choices of initial configurations. If ϕ has a global lower bound, Cucker-Smale system (1.2) has unconditional flocking property under any l^p frameworks.

The global lower bound assumption for the influence function ϕ is too strong and unrealistic, as the influence is usually "local". We would like to find a weaker condition to guarantee unconditional flocking, where ϕ is allowed to vanish at infinity. To proceed, we follow the discussion in [44]. Estimate

$$\min_{ij}\phi_{ij}(t)=\phi(\max_{ij}|\mathbf{x}_i(t)-\mathbf{x}_j(t)|)\geq\phi(2\max_{i}|\mathbf{x}_i(t)-\bar{\mathbf{x}}(t)|)\geq\phi\left(2\left(\sum_{i=1}^{N}|\mathbf{x}_i(t)-\bar{\mathbf{x}}(t)|^2\right)^{1/2}\right).$$

It implies the following estimates on (S,V), under l^2 framework:

$$|\dot{S}|\leq V,\quad \dot{V}\leq-\phi(2\sqrt{N}S)V.$$

The following proposition provides a sufficient condition of the influence function ϕ to ensure flocking, making use of the above decay estimates.

Proposition 1.2.1 ([44, Theorem 3.2])**.** *Suppose $(S(t),V(t))$ satisfying decay estimates*

$$|\dot{S}|\leq V,\quad \dot{V}\leq-k\phi(\alpha S)V,$$

where k,α are positive constants. If ϕ decays sufficiently slow, namely

$$k\int_{S_0}^{\infty}\phi(\alpha r)dr>V_0, \tag{1.6}$$

there exists a finite number D, defined in (1.7)*, such that*

$$\sup_{t\geq0}S(t)\leq D,\quad V(t)\leq V_0e^{-k\phi(\alpha D)t}.$$

Proof. Consider free energy $\mathscr{E}(t) := V(t) + k\int_0^{S(t)} \phi(\alpha s)ds$. From condition (1.11), it is easy to check $\frac{d}{dt}\mathscr{E} \leq 0$. We then have

$$V(t) - V_0 \leq -k\int_{S_0}^{S(t)} \phi(\alpha s)ds.$$

Assume (1.6) holds, then there exists a finite number D, depending only on $\phi, \rho_0, \mathbf{u}_0$, defined as

$$D = \psi^{-1}\big(V_0 + \psi(S_0)\big), \quad \text{where} \quad \psi(t) = k\int_0^t \phi(\alpha s)ds. \tag{1.7}$$

such that $V_0 = k\int_{S_0}^{D} \phi(\alpha s)ds$. Hence, we have

$$0 \leq V(t) \leq k\int_{S(t)}^{D} \phi(\alpha s)ds.$$

It yields that $S(t) \leq D < \infty$. Moreover, from (1.11b) and the monotone decreasing property of ϕ, we obtain

$$\frac{d}{dt}V(t) \leq -k\phi(\alpha D)V(t) \quad \Rightarrow \quad V(t) \leq V_0 e^{-k\phi(\alpha D)t} \to 0, \ \text{ as } t \to +\infty.$$

□

Remark 1.2.2. The proposition states that condition (1.6) implies flocking. Moreover, the decay on V is exponential. We name this *flocking with fast alignment*. We will take advantage of the fast alignment property later in section section 2.3.2 for the macroscopic system.

We introduce a new assumption on ϕ:

$$\phi(r) \text{ is bounded, positive, Lipschitz, non-increasing in } r, \text{ and } \int^{\infty} \phi(r)dr = \infty. \tag{1.8}$$

An prototype choice of ϕ is $\phi(r) = (1+r)^{-\alpha}$, with $\alpha < 1$. The new assumption allows ϕ to vanish at infinity, thus it is weaker than the uniform lower boundedness assumption.

We observe that with the assumption $\int^{\infty} \phi(r)dr = \infty$, condition (1.6) is satisfied regardless of choices of initial configuration. Therefore, we conclude that Cucker-Smale system (1.2) has unconditional flocking property.

Similar type of estimates on (S,V) under l^∞ framework can be derived for both Cucker-Smale (1.2) and Motsch-Tadmor (1.5) systems, consulting [72,73]. The estimates read

$$|\dot{S}| \leq V, \quad \dot{V} \leq -\phi(S)V.$$

A direct consequence of proposition 1.2.1 yields unconditional flocking for (1.2) and (1.5), under assumption (1.6).

The major difference between the l^2 approach and l^∞ approach is, estimates of (S,V) under l^∞ framework is independent of the total number of individuals N. Therefore, when $N \to \infty$, the estimate is valid, only under l^∞ framework.

1.2.4 Other models

We briefly discuss other alignment models in this section.

One celebrating model on self-organized dynamics is proposed by Vicsek in [102]. A continuous in time description reads

$$\dot{\mathbf{x}}_i = c\omega_i, \quad \dot{\omega}_i = (Id - \omega_i \otimes \omega_i)\tilde{\omega}_k.$$

Here, constant speed c is assumed. ω_i denotes the orientation of the i-th individual, where $\mathbf{v}_i = c\omega_i$. $\tilde{\omega}_i$ is the environmental average, given by

$$\tilde{\omega}_i = \frac{J_i}{|J_i|}, \quad J_i = \sum_{\{j\,:\,|\mathbf{x}_i - \mathbf{x}_j| \leq R|\}} \omega_j.$$

This model discribes the dynamics of a group of animals like a school of fish, where each individual change their orientation based on the orientations of its neighbors of distance less than R. Interesting alignment and phase transition behaviors are observed. We refer to [29, 30, 35] for details.

Another popular model on synchronization is proposed by Kuramoto in [58]. It was motivated by chemical and biological oscillators, and has wide applications in neuroscience. The governing equation reads:

$$\frac{d\theta_i}{dt} = \omega_i + \frac{K}{N}\sum_{j=1}^{N}\sin(\theta_j - \theta_i), \quad i = 1, \cdots, N,$$

where $\theta_i = \theta_i(t)$ discribes the frequency of ith oscillator, and ω_i is its intrinsic natural frequency, which is independence in time. The synchronization behavior is reflected by alignment of the asymptotic frequency θ_i, as $t \to \infty$. Consult the review paper [1] and references therein for various discussions on Kuramoto model.

Finally, we would like to mention Krause's consensus model [57] on opinion dynamics. A continuous version reads

$$\dot{\mathbf{x}}_i = \frac{1}{N}\sum_{j=1}^{N}\phi(|\mathbf{x}_i - \mathbf{x}_j|)(\mathbf{x}_j - \mathbf{x}_i).$$

The model has a similar flavor as Cucker-Smale model (1.2). It is a first-order model where the influence is realized by "velocity" rather than "acceleration". Consensus is mathematically defined as the alignment on $\{\mathbf{x}_i\}$, namely, $\lim_{t\to\infty} S(t) = 0$. A sufficient condition to ensure consensus is $S_0 < R$, where $supp(\phi) = [0, R)$. The proof is left to readers.

It is worth noting that Krause's model is a special case of the attraction-repulsion

model (1.1), with potential $\Psi(r) = \int_0^r s\phi(s)ds$. In fact,

$$\dot{\mathbf{x}}_i = \frac{1}{N}\sum_{j=1}^{N} \Psi'(|\mathbf{x}_i-\mathbf{x}_j|)\partial_{\mathbf{x}_j}|\mathbf{x}_i-\mathbf{x}_j| = \frac{1}{N}\sum_{j=1}^{N} \frac{\Psi'(|\mathbf{x}_i-\mathbf{x}_j|)}{|\mathbf{x}_i-\mathbf{x}_j|}(\mathbf{x}_j-\mathbf{x}_i) = \frac{1}{N}\sum_{j=1}^{N} \phi(|\mathbf{x}_i-\mathbf{x}_j|)(\mathbf{x}_j-\mathbf{x}_i).$$

As the potential Ψ above is increasing for all $r > 0$, the system reflects an aggregation model, which leads to concentration [7, 16], or in our words, consensus.

Krause's model also has a normalized version of Motsch-Tadmor type,

$$\dot{\mathbf{x}}_i = \frac{1}{\Phi_i}\sum_{j=1}^{N} \phi(|\mathbf{x}_i-\mathbf{x}_j|)(\mathbf{x}_j-\mathbf{x}_i), \quad \Phi_i = \sum_{j=1}^{N} \phi(|\mathbf{x}_i-\mathbf{x}_j|).$$

We refer to [73] for a comprehensive review on both models. In particular, it is observed that *heterophilous dynamics* might enhance consensus, namely, if ϕ is increasing inside its support, it is more likely to achieve consensus, rather than a usual homophilous decreasing influence ϕ. The full explanation of such phenomenon is still open.

1.3 Multi-scale representation of Flocking

In this section, we derive the kinetic and hydrodynamic representations of flocking systems (1.2) and (1.5), and discuss flocking behaviors under different scales.

1.3.1 Kinetic representation

For Cucker-Smale system (1.2) and Motsch-Tadmor system (1.5) with large number of agents, we can formally derive the asymptotic behaviors of the system when number of agents goes to infinity. The mean field limit yields the following Vlasov-type kinetic equation

$$\partial_t f + \mathbf{v}\cdot\nabla_{\mathbf{x}} f + \nabla_{\mathbf{v}}\cdot(fL(f)) = 0. \tag{1.9}$$

Here, $f = f(t, \mathbf{x}, \mathbf{v})$ denotes the number density of agents at postion $\mathbf{x}$, velocity $\mathbf{v}$ and time t. The operator L is different for the two systems, see [45, 72] for derivation. In detail, for Cucker-Smale system (1.2), the corresponding operator is

$$L(f)(t, \mathbf{x}, \mathbf{v}) = \iint \phi(|\mathbf{x} - \mathbf{y}|)(\mathbf{v}^* - \mathbf{v}) f(t, \mathbf{y}, \mathbf{v}^*) d\mathbf{y} d\mathbf{v}^*. \tag{C-S}$$

For Motsch-Tadmor system (1.5), the corresponding operator is

$$L(f)(t, \mathbf{x}, \mathbf{v}) = \iint \frac{\phi(|\mathbf{x} - \mathbf{y}|)}{\Phi(t, \mathbf{x})} (\mathbf{v}^* - \mathbf{v}) f(t, \mathbf{y}, \mathbf{v}^*) d\mathbf{y} d\mathbf{v}^*, \tag{M-T}$$

where

$$\Phi(t, \mathbf{x}) = \iint \phi(|\mathbf{x} - \mathbf{y}|) f(t, \mathbf{y}, \mathbf{v}^*) d\mathbf{y} d\mathbf{v}^*.$$

We provide a brief formal derivation of (1.9) in appendix section 1.A for completeness.

The global existence theory for Vlasov-type equation has been studied in many contexts. We refer readers to [36, 61, 79, 86] for Vlasov-Poisson equations and [37, 38] for Vlasov-Maxwell equations, as well as a nice technical review paper [75]. For kinetic flocking system (1.9), it is significantly easier to obtain existence of global C^1 solution, as the alignment forcing has no singularity. We state the following theorem.

Theorem 1.3.1 (Global strong solution)**.** *Consider* (1.9) *with* (C-S) *or* (M-T) *setup. Suppose the initial profile $f_0 \in C^1 \cap W^{1,\infty}(\mathbb{R}^{2n})$ is compactly supported and C^1-regular. Then, for any $T \in (0, \infty)$, there exists a unique strong solution $f \in C^1([0, T) \times \mathbb{R}^{2n})$.*

The proof for (C-S) setup is given in [45, Theorem 3.3]. We will give a similar (but stronger) proof in the appendix section 1.B, for both setups.

1.3.2 Unconditional flocking for kinetic models

In this section, we discuss the flocking property for the solution of kinetic system (1.9), under (C-S) and (M-T) setups.

To illustrate flocking in kinetic level, we first define the total variation in position $\mathbf{x}$ and velocity $\mathbf{v}$:

$$S(t) := \sup_{(\mathbf{x},\mathbf{v}),(\mathbf{y},\mathbf{v}^*)\in \mathrm{supp} f(t)} |\mathbf{x}-\mathbf{y}|, \quad V(t) := \sup_{(\mathbf{x},\mathbf{v}),(\mathbf{y},\mathbf{v}^*)\in \mathrm{supp} f(t)} |\mathbf{v}-\mathbf{v}^*|. \tag{1.10}$$

As mentioned in section 1.2.3, the definition of (S,V) are under L^∞ framework.

Definition 1.3.1 (Kinetic flocking)**.** We say a solution $f(t,\mathbf{x},\mathbf{v})$ converges to a flock in the kinetic level, if $S(t)$ remains bounded in all time, and $V(t)$ decays to 0 asymptotically:

$$S(t) \le D, \ \ \forall t \ge 0; \quad V(t) \to 0 \text{ as } t \to \infty.$$

We claim that (1.9) has unconditional flocking property if ϕ satisfies (1.8).

Theorem 1.3.2 (Unconditional flocking for kinetic systems)**.** *Suppose f is the solution of the system* (1.9) *with* (C-S) *or* (M-T) *setup, with regular initial profile $f_0 \in C^1 \cap W^{1,\infty}$. Then f converges to a flock in the sense of definition* 1.3.1.

The theorem has been proved for Cucker-Smale system in [19]. Here, we give a new proof for both setups. The idea is addressed in [72], and [97] with more details.

The heart of the matter is to prove the following decay estimates for (S,V). It immediately implies unconditional flocking, thanks to proposition 1.2.1.

Proposition 1.3.3 (L^∞-type Decay estimates)**.** (S,V) *are defined in* (1.10)*, where* f *is a solution of* (1.9)*. Then,*

$$\frac{d}{dt}S(t) \le V(t), \tag{1.11a}$$

$$\frac{d}{dt}V(t) \le -k\phi(S(t))V(t)., \tag{1.11b}$$

where $k = m$ *for* (C-S) *and* $k = 1$ *for* (M-T)*.*

Here, m is the total mass $m = \iint f dxdv$, which is conserved in time.

Proof. The characteristic curve of the system reads $(\mathbf{x}(t), \mathbf{v}(t))$ where

$$\frac{d}{dt}\mathbf{x}(t) = \mathbf{v}(t), \quad \frac{d}{dt}\mathbf{v}(t) = L(f)(t, \mathbf{x}(t), \mathbf{v}(t)).$$

We consider two characteristics $(\mathbf{x}_1(t), \mathbf{v}_1(t))$ and $(\mathbf{x}_2(t), \mathbf{v}_2(t))$, both starting inside the support of f_0. It is clear that f is constant along the characteristic curve. Therefore, $(\mathbf{x}_i(t), \mathbf{v}_i(t))$ lies inside the support of $f(t)$.

To simplify the notations, we omit the time variable throughout the proof.

Compute

$$\frac{d}{dt}|\mathbf{x}_1 - \mathbf{x}_2|^2 = 2\langle \mathbf{x}_1 - \mathbf{x}_2, \mathbf{v}_1 - \mathbf{v}_2 \rangle \le 2SV.$$

By taking the supreme of the left hand side for all $\mathbf{x}_1, \mathbf{x}_2$, the inequality yields (1.11a).

Similarly, we have

$$\frac{d}{dt}|\mathbf{v}_1 - \mathbf{v}_2|^2 = 2\big\langle \mathbf{v}_1 - \mathbf{v}_2, L(f)(\mathbf{x}_1, \mathbf{v}_1) - L(f)(\mathbf{x}_2, \mathbf{v}_2) \big\rangle.$$

The following key estimates can be proved later in lemma 1.3.5 for (C-S) and lemma 1.3.4 for (M-T)

$$L(f)(\mathbf{x}_1, \mathbf{v}_1) - L(f)(\mathbf{x}_2, \mathbf{v}_2) \le k(1 - \phi(S))V - k(\mathbf{v}_1 - \mathbf{v}_2), \tag{1.12}$$

for all $(\mathbf{x}_1, \mathbf{v}_1), (\mathbf{x}_2, \mathbf{v}_2)$ in the support of f. It yields

$$\frac{d}{dt}|\mathbf{v}_1 - \mathbf{v}_2|^2 \leq 2k(1-\phi(S))|\mathbf{v}_1 - \mathbf{v}_2|V - 2k|\mathbf{v}_1 - \mathbf{v}_2|^2.$$

Take $\mathbf{v}_1, \mathbf{v}_2$ where $|\mathbf{v}_1 - \mathbf{v}_2| \to V$, we end up with (1.11b). □

To prove the remaining key estimate, we start with (M-T) configuration.

Lemma 1.3.4. *Inequality* (1.12) *holds for* (M-T) *with* $k = 1$.

Proof. Given any pairs $(\mathbf{x}, \mathbf{v})$ and $(\mathbf{y}, \mathbf{v}^*)$ inside the support of f, define

$$b(t, \mathbf{x}, \mathbf{v}, \mathbf{y}, \mathbf{v}^*) := \frac{\phi(|\mathbf{x}-\mathbf{y}|)f(t, \mathbf{y}, \mathbf{v}^*)}{\Phi(t, \mathbf{x})}.$$

Such function b enjoys the following properties

(P1) $\iint b(t, \mathbf{x}, \mathbf{v}, \mathbf{y}, \mathbf{v}^*) d\mathbf{y} d\mathbf{v}^* = 1$, for all t,

(P2) $k \iint b(t, \mathbf{x}, \mathbf{v}, \mathbf{y}, \mathbf{v}^*)(\mathbf{v}^* - \mathbf{v}) d\mathbf{y} d\mathbf{v}^* = L(f)(t, \mathbf{x}, \mathbf{v})$,

(P3) There exists $\eta(t, \mathbf{v}^*)$ such that $\int b(t, \mathbf{x}, \mathbf{v}, \mathbf{y}, \mathbf{v}^*) d\mathbf{y} \geq \eta(\mathbf{v}^*)$ for all $t, \mathbf{x}$ and $\mathbf{v}$. Also, $\int \eta(t, \mathbf{v}^*) d\mathbf{v}^* = \phi(S) > 0$ for all t.

The first two properties are easy to check. The last one provides some positivity property of the uniform lower bound of b with respect to $(\mathbf{x}, \mathbf{v})$, which is essential to the estimate. Here, we have

$$\int b(t, \mathbf{x}, \mathbf{v}, \mathbf{y}, \mathbf{v}^*) dy = \frac{\int \phi(|\mathbf{x}-\mathbf{y}|)f(t, \mathbf{y}, \mathbf{v}^*) d\mathbf{y}}{\Phi(t, \mathbf{x})} \geq \frac{\phi(S) \int f(t, \mathbf{y}, \mathbf{v}^*) d\mathbf{y}}{m},$$

thanks to the decreasing property of ϕ and the universal assumption of $\phi(0) = 1$. The right hand side is independent with respect to $\mathbf{x}$, and could be defined as the η. Clearly, (P3) is satisfied.

We continue to compute, omitting the t variable.

$$
\begin{aligned}
& L(f)(\mathbf{x}_1,\mathbf{v}_1) - L(f)(\mathbf{x}_2,\mathbf{v}_2) \\
\overset{(P2)}{=}\ & k \iint \left[b(\mathbf{x}_1,\mathbf{v}_1,\mathbf{y},\mathbf{v}^*)(\mathbf{v}^*-\mathbf{v}_1) - b(\mathbf{x}_2,\mathbf{v}_2,\mathbf{y},\mathbf{v}^*)(\mathbf{v}^*-\mathbf{v}_2) \right] d\mathbf{y} d\mathbf{v}^* \\
\overset{(P1)}{=}\ & k \iint \left(b(\mathbf{x}_1,\mathbf{v}_1,\mathbf{y},\mathbf{v}^*) - b(\mathbf{x}_2,\mathbf{v}_2,\mathbf{y},\mathbf{v}^*) \right) \mathbf{v}^* d\mathbf{y} d\mathbf{v}^* - k(\mathbf{v}_1-\mathbf{v}_2) \\
=\ & k \left[\int \left(\int b(\mathbf{x}_1,\mathbf{v}_1,\mathbf{y},\mathbf{v}^*) d\mathbf{y} - \eta(\mathbf{v}^*) \right) d\mathbf{v}^* - \int \left(\int b(\mathbf{x}_2,\mathbf{v}_2,\mathbf{y},\mathbf{v}^*) d\mathbf{y} - \eta(\mathbf{v}^*) \right) d\mathbf{v}^* \right] \\
& - k(\mathbf{v}_1-\mathbf{v}_2) \\
=\ & k(1-\phi(S)) \left[\int \hat{b}(\mathbf{x}_1,\mathbf{v}_1,\mathbf{v}^*)\mathbf{v}^* d\mathbf{v}^* - \int \hat{b}(\mathbf{x}_2,\mathbf{v}_2,\mathbf{v}^*) d\mathbf{v}^* \right] - k(\mathbf{v}_1-\mathbf{v}_2).
\end{aligned}
$$

Here, $\hat{b}$ is defined as $\hat{b}(\mathbf{x},\mathbf{v},\mathbf{v}^*) = \dfrac{\int b(\mathbf{x},\mathbf{v},\mathbf{y},\mathbf{v}^*) d\mathbf{y} - \eta(\mathbf{v}^*)}{1-\phi(S)}$. From (P1) and (P3), we know $\hat{b}$ is positive, supported inside the support of f in $\mathbf{v}$, and $\displaystyle\int \hat{b}(\mathbf{x},\mathbf{v},\mathbf{v}^*) d\mathbf{v}^* = 1$ for all $(\mathbf{x},\mathbf{v})$. Therefore, $\displaystyle\int \hat{b}(\mathbf{x},\mathbf{v},\mathbf{v}^*)\mathbf{v}^* d\mathbf{v}^*$ lies inside the convex envelope of the support of f in $\mathbf{v}$. Hence,

$$
\left| \int \hat{b}(\mathbf{x}_1,\mathbf{v}_1,\mathbf{v}^*)\mathbf{v}^* d\mathbf{v}^* - \int \hat{b}(\mathbf{x}_2,\mathbf{v}_2,\mathbf{v}^*) d\mathbf{v}^* \right| \leq V,
$$

and (1.12) holds. □

We now apply the same idea to (C-S) configuration.

Lemma 1.3.5. *Inequality* (1.12) *holds for* (C-S) *with* $k = m$.

Proof. Our goal is to find function b satisfying (P1)-(P3). After then, we can proceed using lemma 1.3.4. We pick

$$
b(t,\mathbf{x},\mathbf{v},\mathbf{y},\mathbf{v}^*) = \frac{1}{m}\phi(|\mathbf{x}-\mathbf{y}|) f(t,\mathbf{y},\mathbf{v}^*) + \left(1 - \frac{\Phi(t,\mathbf{x})}{m}\right) \delta_0(\mathbf{x}-\mathbf{y})\delta_0(\mathbf{v}-\mathbf{v}^*),
$$

where δ_0 is the Dirac delta at the origin. With this setup, (P1) and (P2) are easily satisfied. For (P3), the same choice of η in lemma 1.3.4 can be used. □

1.3.3 Macroscopic representation

In this section, we discuss the macroscopic representation for flocking models, which is formally obtained by taking moments of the kinetic system (1.9).

Multiply (1.9) against $\{1, \mathbf{v}\}$ and integrate over velocity space $\mathbb{R}^n$. We get the hydrodynamic description of the system

$$\rho_t + div(\rho \mathbf{u}) = 0, \tag{1.13a}$$

$$(\rho \mathbf{u})_t + div(\rho \mathbf{u} \otimes \mathbf{u} + P) = \rho \mathsf{F}. \tag{1.13b}$$

Here, $(\rho, \mathbf{u})$ are hydrodynamic variables

$$\rho(\mathbf{x}) = \int_{\mathbb{R}^n} f(\mathbf{x}, \mathbf{v}) d\mathbf{v}, \quad \rho(\mathbf{x})\mathbf{u}(\mathbf{x}) = \int_{\mathbb{R}^n} \mathbf{v} f(\mathbf{x}, \mathbf{v}) d\mathbf{v}.$$

F is the non-local alignment forcing. For Cucker-Smale system,

$$\mathsf{F}(\mathbf{x}) = \int_{\mathbb{R}^n} \phi(|\mathbf{x}-\mathbf{y}|)(\mathbf{u}(\mathbf{y}) - \mathbf{u}(\mathbf{x}))\rho(\mathbf{y}) d\mathbf{y}. \tag{C-S}$$

For Motsch-Tadmor system,

$$\mathsf{F}(\mathbf{x}) = \int_{\mathbb{R}^n} \frac{\phi(|\mathbf{x}-\mathbf{y}|)}{\Phi(\mathbf{x})}(\mathbf{u}(\mathbf{y}) - \mathbf{u}(\mathbf{x}))\rho(\mathbf{y}) d\mathbf{y}, \tag{M-T}$$

where

$$\Phi(\mathbf{x}) = \int_{\mathbb{R}^n} \phi(|\mathbf{x}-\mathbf{y}|)\rho(\mathbf{y}) d\mathbf{y}. \tag{1.14}$$

P is the stress tensor (or pressure) which is defined by

$$P(\mathbf{x}) = \int_{\mathbb{R}^n} (\mathbf{v}-\mathbf{u}) \otimes (\mathbf{v}-\mathbf{u}) f(\mathbf{x}, \mathbf{v}) d\mathbf{v}. \tag{1.15}$$

To close the system with $(\rho, \mathbf{u})$, we apply different ansatz, which lead to different types of pressure.

Pressure-less closure: $P = 0$**.**

Consider kinetic flocking system (1.9) with strong local alignment

$$f_t + \mathbf{v}\cdot\nabla_{\mathbf{x}} f + \nabla_v \cdot (fL(f)) = \frac{1}{\varepsilon}\nabla_v \cdot ((\mathbf{v}-\mathbf{u})f).$$

Take $\varepsilon \to 0$, formally the right hand side should converges to zero. The limit must have the form of mono kinetic type:

$$f(\mathbf{x},\mathbf{v}) = \rho(\mathbf{x})\delta(\mathbf{v}-\mathbf{u}(\mathbf{x})).$$

Plug in this ansatz to (1.15). We get

$$P(\mathbf{x}) = \rho(\mathbf{x})\int_{\mathbb{R}^n} (\mathbf{v}-\mathbf{u})\otimes(\mathbf{v}-\mathbf{u})\delta(\mathbf{v}-\mathbf{u})d\mathbf{v} = 0.$$

Isothermal pressure closure: $P = \rho\mathbb{I}_{n\times n}$**.**

Consider kinetic flocking system (1.9) with strong local alignment and strong noise

$$f_t + \mathbf{v}\cdot\nabla_{\mathbf{x}} f + \nabla_v \cdot (fL(f)) = \frac{1}{\varepsilon}\nabla_v \cdot ((\mathbf{v}-\mathbf{u})f) + \frac{1}{\varepsilon}\Delta_{\mathbf{v}} f.$$

Again, take $\varepsilon \to 0$, formally the right hand side should converges to zero. The limit must have the form

$$f(\mathbf{x},\mathbf{v}) = \rho(\mathbf{x})\frac{1}{(2\pi)^{n/2}} e^{-\frac{|\mathbf{v}-\mathbf{u}(\mathbf{x})|^2}{2}},$$

where f is a Maxwellian with constant temperature $\mathscr{M}(\rho,\mathbf{u},1)$. Plug in this ansatz to (1.15). We get

$$P_{ij}(\mathbf{x}) = \rho(\mathbf{x})\int_{\mathbb{R}^n} (v_i-u_i(\mathbf{x}))(v_j-u_j(\mathbf{x}))\frac{1}{(2\pi)^{n/2}} e^{-\frac{|\mathbf{v}-\mathbf{u}(\mathbf{x})|^2}{2}} d\mathbf{v} = \frac{\rho(\mathbf{x})}{(2\pi)^{n/2}}\int_{\mathbb{R}^n} v_i v_j e^{-\frac{|\mathbf{v}|^2}{2}} d\mathbf{v}.$$

If $i \neq j$, then the integral at the right hand side equals 0, due to anti-symmetry. If $i = j$,

$$P_{ii}(\mathbf{x}) = \frac{\rho(\mathbf{x})}{(2\pi)^{n/2}}\int_{\mathbb{R}^n} v_i^2 e^{-\frac{|\mathbf{v}|^2}{2}} d\mathbf{v} = \frac{\rho(\mathbf{x})}{(2\pi)^{1/2}}\int_{-\infty}^{\infty} v_i^2 e^{-\frac{v_i^2}{2}} dv_i = \rho.$$

The rigorous derivation of hydrodynamic limiting system with pressure has been studied in [51]. The authors show that

$$f^{\varepsilon}(t,\mathbf{x},\mathbf{v}) \to \rho(\mathbf{x},t)\frac{1}{(2\pi)^{n/2}}e^{-\frac{|\mathbf{v}-\mathbf{u}(\mathbf{x},t)|^2}{2}},$$

as $\varepsilon \to 0$, for any time t when strong solution $(\rho,\mathbf{u})$ exists and solves (1.13) with isothermal pressure.

There are other closures on pressure P. For instance, the isentropic pressure is given by $P = P(\rho) = \rho^{\gamma}\mathbb{I}_{n\times n}$ for $\gamma > 1$.

We shall concentrate on the pressure-less setup as it is relatively easier to adapt our techniques without taking into account pressure. It reads

$$\rho_t + div(\rho\mathbf{u}) = 0,$$

$$(\rho\mathbf{u})_t + div(\rho\mathbf{u}\otimes\mathbf{u}) = \rho\mathsf{F}.$$

It has be also expressed by the following non-conservative form.

$$\rho_t + div(\rho\mathbf{u}) = 0, \tag{1.16a}$$

$$\mathbf{u}_t + \mathbf{u}\cdot\nabla\mathbf{u} = \mathsf{F}. \tag{1.16b}$$

It is a pressure-less compressible Eulerian dynamics, with alignment forcing.

The main question we address to system (1.16) is: *does this system reflect flocking behaviors, as observed and proved in microscopic and kinetic systems.*

The question will be answered in two steps. In the next section, we prove the statement: strong solution must flock. And the whole chapter 2 is devoted to study global existence of strong solution for macroscopic flocking systems.

1.3.4 Strong solution must flock

In this section, we discuss the flocking behavior of the macroscopic system (1.16). Similar than the kinetic system, we start with the definition of flocking in hydrodynamic level. Define

$$S(t) := \sup_{\mathbf{x},\mathbf{y}\in supp\rho(t)} |\mathbf{x}-\mathbf{y}|, \quad V(t) := \sup_{\mathbf{x},\mathbf{y}\in supp\rho(t)} |\mathbf{u}(\mathbf{x})-\mathbf{u}(\mathbf{y})|. \tag{1.17}$$

Definition 1.3.2 (Hydrodynamic flocking)**.** We say a solution $(\rho,\mathbf{u})$ of (1.16) converges to a flock in the hydrodynamic level, if $S(t)$ remains bounded in all time, and $V(t)$ decays to 0 asymptotically:

$$\exists\, D < +\infty, \text{ such that } S(t) \le D, \ \ \forall t \ge 0; \quad V(t) \to 0 \text{ as } t \to \infty.$$

The following theorem has the same flavor as theorem 1.3.2, which claims any strong solution of the kinetic flocking system must flock. Macroscopic system (1.16) enjoys the same property, assuming the existence of strong solution globally in time.

Theorem 1.3.6 (Strong solution must flock)**.** *Let* $(\rho,\mathbf{u})$ *be a global strong solution of system* (1.16)*, where in particular* $\mathbf{u} \in C(\mathbb{R}^+, W^{1,\infty}(\mathbb{R}^n))$*. Then,* $(\rho,\mathbf{u})$ *converges to a flock, in the sense of definition* 1.3.2*.*

Remark 1.3.1. The assumption of existence of strong solution is necessary for the theorem. It ensures that there is no intersection among characteristics and no shock formation. For kinetic system (1.9), existence of strong solution is easily proved for all regular initial profiles, as the free transport is linear, and the forcing is not harmful. However, it is not trivial to establish an existence theory for the macroscopic system (1.16). The main

difference is that nonlinear convection replaces the linear transport in the kinetic system, which may lead to finite time loss of regularity. We will study the existence theory of strong solutions for macroscopic flocking models (1.16) in chapter 2. As long as strong solution exists globally in time, flocking property is granted, thanks to theorem 1.3.6.

We first show the following decay estimates. Then, the theorem is a consequence of proposition 1.2.1. The idea of the proof is similar to the kinetic system (proposition 1.3.3), see [95].

Proposition 1.3.7 (Decay estimates)**.** *If* $(\rho, \mathbf{u})$ *are strong solutions of* (1.16)*, then the decay estimates* (1.11) *are satisfies, with* (S,V) *defined in* (1.17)*.*

Proof. Consider two characteristics $\dot{X}(t) = \mathbf{u}(X,t)$, $\dot{Y}(t) = \mathbf{u}(Y,t)$ subject to initial conditions $X(0) = \mathbf{x}$, $Y(0) = \mathbf{y}$, where $\mathbf{x}, \mathbf{y} \in supp(\rho_0)$. Along the characteristics, the evolution of $\rho(t)$ reads

$$\frac{d}{dt}\rho(X(t),t) = \rho_t + \mathbf{u}\cdot\nabla\rho = -\rho div\,\mathbf{u}.$$

As $div\,\mathbf{u}$ is assumed to be bounded, it is clear that $\rho(t,X(t)) > 0$ if and only if $\rho_0(\mathbf{x}) > 0$. Therefore, $X(t) \in supp(\rho(t))$ as long as $\mathbf{x} \in supp(\rho_0)$.

To simplify the notations, we omit the time variable throughout the proof.

Compute

$$\frac{d}{dt}|X-Y|^2 = 2\langle X-Y, \mathbf{u}(X)-\mathbf{u}(Y)\rangle \le 2SV.$$

By taking the supreme of the left hand side for all $\mathbf{x}, \mathbf{y} \in supp(\rho_0)$, the inequality yields (1.11a).

Similarly, we have

$$\frac{d}{dt}|\mathbf{u}(X)-\mathbf{u}(Y)|^2 = 2\left\langle \mathbf{u}(X)-\mathbf{u}(Y), \mathsf{F}(X)-\mathsf{F}(Y)\right\rangle.$$

Moreover, we claim the following key inequality. The proof is postponed to lemma 1.3.8 for (C-S) and lemma 1.3.9 for (M-T).

$$\mathsf{F}(X) - \mathsf{F}(Y) \leq k(1-\phi(S))V - k(\mathbf{u}(X) - \mathbf{u}(Y)), \tag{1.18}$$

for all X, Y starting at $\mathbf{x}, \mathbf{y} \in supp(\rho_0)$. Hence,

$$\frac{1}{2}\frac{d}{dt}|\mathbf{u}(X) - \mathbf{u}(Y)|^2 \leq k(1-\phi(S))V|\mathbf{u}(X) - \mathbf{u}(Y)| - k|\mathbf{u}(X) - \mathbf{u}(Y)|^2.$$

Take X, Y where $|\mathbf{u}(X) - \mathbf{u}(Y)| \to V$, we end up with (1.11b). □

Lemma 1.3.8. *Inequality* (1.18) *holds for* (C-S) *setup with* $k = m$.

Proof. Define

$$b(\mathbf{x}, \mathbf{y}) := \frac{1}{m}\phi(|\mathbf{x}-\mathbf{y}|)\rho(\mathbf{y}) + \left(1 - \frac{\Phi(\mathbf{x})}{m}\right)\delta_0(\mathbf{x}-\mathbf{y}),$$

where Φ is defined in (1.14) and δ_0 is the Dirac delta.

b has the following properties:

(P1) $b(\mathbf{x}, \mathbf{y}) \geq 0$, for all $\mathbf{x}, \mathbf{y} \in supp(\rho)$.

(P2) $\displaystyle\int_{\mathbb{R}^n} b(\mathbf{x}, \mathbf{y}) d\mathbf{y} = 1$, for all $\mathbf{x} \in supp(\rho)$.

(P3) $\displaystyle\int_{\mathbb{R}^n} b(\mathbf{x}, \mathbf{y})(\mathbf{u}(\mathbf{y}) - \mathbf{u}(\mathbf{x})) d\mathbf{y} = \frac{1}{k}\int_{\mathbb{R}^n} \phi(|\mathbf{x}-\mathbf{y}|)(\mathbf{u}(\mathbf{y}) - \mathbf{u}(\mathbf{x}))\rho(\mathbf{y}) d\mathbf{y}.$

Here, we add the second part to the definition of b to make sure (P2) holds. As we only add a delta mass, the equality of (P3) remains true as well.

To prove (1.18), we need to improve (P1). A minimum assumption on b introduced in [73] is

(P1') $\Psi(\mathbf{x},\mathbf{y}) > 0$, for all $\mathbf{x},\mathbf{y} \in supp(\rho)$, where

$$\Psi(\mathbf{x},\mathbf{y}) := \int_{\mathbb{R}^n} \eta_{\mathbf{x},\mathbf{y}}(\mathbf{z})d\mathbf{z} \quad \text{and} \quad \eta_{\mathbf{x},\mathbf{y}}(\mathbf{z}) \leq \min\{b(\mathbf{x},\mathbf{z}), b(\mathbf{y},\mathbf{z})\}.$$

In our case, We pick $\eta_{\mathbf{x},\mathbf{y}}(\mathbf{z}) = \eta(\mathbf{z}) = \frac{1}{m}\phi(S)\rho(\mathbf{z})$ and $\Psi(\mathbf{x},\mathbf{y}) = \Psi = \phi(S) > 0$.

Next, compute

$$\begin{aligned}
\mathsf{F}(X) - \mathsf{F}(Y) \;&=\; \int_{\mathbb{R}^n} [\phi(|X-\mathbf{z}|)(\mathbf{u}(\mathbf{z}) - \mathbf{u}(X)) - \phi(|Y-\mathbf{z}|)(\mathbf{u}(\mathbf{z}) - \mathbf{u}(Y))]\,\rho(\mathbf{z})d\mathbf{z} \\
&\overset{(P3)}{=}\; k\int_{\mathbb{R}^n} [b(X,\mathbf{z})(\mathbf{u}(\mathbf{z}) - \mathbf{u}(X)) - b(Y,\mathbf{z})(\mathbf{u}(\mathbf{z}) - \mathbf{u}(Y))]\,d\mathbf{z} \\
&\overset{(P2)}{=}\; k\int_{\mathbb{R}^n} (b(X,\mathbf{z}) - b(Y,\mathbf{z}))\,\mathbf{u}(\mathbf{z})d\mathbf{z} - k(\mathbf{u}(X) - \mathbf{u}(Y)) \\
&=\; k\int_{\mathbb{R}^n} (b(X,\mathbf{z}) - \eta(\mathbf{z}))\mathbf{u}(\mathbf{z})d\mathbf{z} - k\int_{\mathbb{R}^n} (b(Y,\mathbf{z}) - \eta(\mathbf{z}))\mathbf{u}(\mathbf{z})d\mathbf{z} - k(\mathbf{u}(X) - \mathbf{u}(Y)) \\
&\overset{(P1')}{=}\; k(1-\Psi)\int_{\mathbb{R}^n} \hat{b}(X,\mathbf{z})\mathbf{u}(\mathbf{z})d\mathbf{z} - k(1-\Psi)\int_{\mathbb{R}^n} \hat{b}(Y,\mathbf{z})\mathbf{u}(\mathbf{z})d\mathbf{z} - k(\mathbf{u}(X) - \mathbf{u}(Y)),
\end{aligned}$$

where $\hat{b}$ is defined as

$$\hat{b}(\mathbf{x},\mathbf{y}) := \frac{b(\mathbf{x},\mathbf{y}) - \eta(\mathbf{y})}{1-\Psi}.$$

Condition (P1') guarantees $\hat{b}(\mathbf{x},\mathbf{y}) \geq 0$ for all $\mathbf{x},\mathbf{y} \in supp(\rho)$, and $\int_{supp(\rho)} \hat{b}(\mathbf{x},\mathbf{y})d\mathbf{y} = 1$, for all $\mathbf{x} \in supp(\rho)$. Therefore, define

$$\bar{\mathbf{u}}(X) := \int_{\mathbb{R}^n} \hat{b}(X,\mathbf{z})\mathbf{u}(\mathbf{z})d\mathbf{z},$$

then $\bar{\mathbf{u}}(X)$ is included in the convex envelope of the set $\{\mathbf{u}(\mathbf{x}) \mid \mathbf{x} \in supp(\rho)\}$ as long as $X \in supp(\rho)$. It yields that $|\bar{\mathbf{u}}(X) - \bar{\mathbf{u}}(Y)| \leq V$. Therefore,

$$\begin{aligned}
\mathsf{F}(X) - \mathsf{F}(Y) &= k(1-\Psi)(\bar{\mathbf{u}}(X) - \bar{\mathbf{u}}(Y)) - k(\mathbf{u}(X) - \mathbf{u}(Y)) \\
&\leq k(1-\Psi)V - k(\mathbf{u}(X) - \mathbf{u}(Y)).
\end{aligned}$$

□

Lemma 1.3.9. *Inequality* (1.18) *holds for* (M-T) *setup with* $k = 1$.

Proof. In this case, we define

$$b(\mathbf{x}, \mathbf{y}) = \frac{\phi(|\mathbf{x} - \mathbf{y}|)\rho(\mathbf{y})}{\Phi(\mathbf{x})}.$$

It is easy to check that b satisfies (P1')(P2) and (P3) with $k = 1$.

A same argument as lemma 1.3.8 yields (1.18). □

Appendix

1.A Formal derivation of the kinetic system

We provide a formal derivation of kinetic flocking systems. More details and discussions are referred to [45, 72].

We represent the number density f as an empirical distribution with respect to the N-particle system:

$$f(t,\mathbf{x},\mathbf{v}) = \frac{1}{N}\sum_{i=1}^{N} \delta_{\mathbf{x}_i(t)}\delta_{\mathbf{v}_i(t)}.$$

Take smooth test function $\varphi(\mathbf{x},\mathbf{v})$ and compute

$$\iint \partial_t f(t,\mathbf{x},\mathbf{v})\varphi(\mathbf{x},\mathbf{v})d\mathbf{x}d\mathbf{v} = \frac{1}{N}\sum_{i=1}^{N} \partial_t \varphi(\mathbf{x}_i(t),\mathbf{v}_i(t)) = \frac{1}{N}\sum_{i=1}^{N} \left[\dot{\mathbf{x}}_i \cdot \nabla_{\mathbf{x}}\varphi(\mathbf{x}_i,\mathbf{v}_i) + \dot{\mathbf{v}}_i \cdot \nabla_{\mathbf{v}}\varphi(\mathbf{x}_i,\mathbf{v}_i)\right].$$

We now apply the microscopic model $\dot{\mathbf{x}}_i = \mathbf{v}_i, \quad \dot{\mathbf{v}}_i = \mathsf{F}_i$ to the right hand side. F_i is the forcing for particle i. For Cucker-Smale system, we have

$$\mathsf{F}_i = \frac{1}{N}\sum_{j=1}^{N} \phi(|\mathbf{x}_i - \mathbf{x}_j|)(\mathbf{v}_j - \mathbf{v}_i) = \iint \phi(|\mathbf{x}_i - \mathbf{y}|(\mathbf{v}^* - \mathbf{v}_i)f(\mathbf{y},\mathbf{v}^*)d\mathbf{y}d\mathbf{v}^* = L(f)(\mathbf{x}_i,\mathbf{v}_i),$$

where $L(f)$ is defined in section 1.3.1. Similar argument can be applied to Motsch-Tadmor setup as well.

Let's continue with our formal calculation.

$$\begin{aligned}\iint \partial_t f(\mathbf{x},\mathbf{v})\varphi(\mathbf{x},\mathbf{v})d\mathbf{x}d\mathbf{v} &= \frac{1}{N}\sum_{i=1}^{N}\left[\mathbf{v}_i\cdot\nabla_{\mathbf{x}}\varphi(\mathbf{x}_i,\mathbf{v}_i)+L(f)(\mathbf{x}_i,\mathbf{v}_i)\cdot\nabla_{\mathbf{v}}\varphi(\mathbf{x}_i,\mathbf{v}_i)\right]\\ &= \iint\left[\mathbf{v}\cdot\nabla_{\mathbf{x}}\varphi(\mathbf{x},\mathbf{v})+L(f)(\mathbf{x},\mathbf{v})\cdot\nabla_{\mathbf{v}}\varphi(\mathbf{x},\mathbf{v})\right]f(\mathbf{x},\mathbf{v})d\mathbf{x}d\mathbf{v}\\ &= -\iint\left[\mathbf{v}\cdot\nabla_{\mathbf{x}}f(\mathbf{x},\mathbf{v})+\nabla_{\mathbf{v}}\cdot(L(f)(\mathbf{x},\mathbf{v})f(\mathbf{x},\mathbf{v}))\right]\varphi(\mathbf{x},\mathbf{v})d\mathbf{x}d\mathbf{v}.\end{aligned}$$

We end up with the weak formulation of the kinetic flocking system (1.9).

1.B Global existence theory for kinetic flocking equations

In this section, we prove theorem 1.3.1: a global existence result for kinetic flocking system (1.9), under both (C-S) and (M-T).

We proceed with standard estimates for Vlasov-type kinetic equations. See [6] for similar argument for kinetic system for granular gas.

Take characteristic curve (X,V) starting at point $(\mathbf{x},\mathbf{v})$.

$$\dot{X}(t,\mathbf{x},\mathbf{v}) = V(t,\mathbf{x},\mathbf{v}), \tag{1.19a}$$

$$\dot{V}(t,\mathbf{x},\mathbf{v}) = L(f)(t,X(t,\mathbf{x},\mathbf{v}),V(t,\mathbf{x},\mathbf{v})). \tag{1.19b}$$

Note that there is a slight change of notation. We use (X,V) to denote the characteristic curve. As for variation in velocity defined in definition 1.2.1, we use $[V]$ instead, only in this section. Clearly $[V](t)$ is bounded in all time. We denote the uniform bound as $[V]$. In the case where we have unconditional flocking, $[V] = [V](0)$.

Define the Jacobian J as the transformation matrix from Eulerian coordinates $(\mathbf{x},\mathbf{v})$

to Lagrangian coordinates (X,V).

$$J(t,\mathbf{x},\mathbf{v}) = \begin{bmatrix} \partial_{\mathbf{x}} X & \partial_{\mathbf{v}} X \\ \partial_{\mathbf{x}} V & \partial_{\mathbf{v}} V \end{bmatrix}, \quad A(t,\mathbf{x},\mathbf{v}) = \begin{bmatrix} 0 & 1 \\ \partial_{\mathbf{x}} L(f) & \partial_{\mathbf{v}} L(f) \end{bmatrix}.$$

It is easy to check from (1.19) that

$$\dot{J}(t,\mathbf{x},\mathbf{v}) = A(t,X,V)J(t,\mathbf{x},\mathbf{v}), \quad J(\mathbf{x},\mathbf{v},0) \equiv \mathbb{I}_{2n\times 2n}, \tag{1.20}$$

$$\dot{J}^{-1}(t,\mathbf{x},\mathbf{v}) = -J^{-1}(t,\mathbf{x},\mathbf{v})A(t,X,V), \quad J^{-1}(\mathbf{x},\mathbf{v},0) \equiv \mathbb{I}_{2n\times 2n}, \tag{1.21}$$

$$det\ J(t,\mathbf{x},\mathbf{v}) = \exp\left(\int_0^t \mathrm{tr}A(s,X(s,\mathbf{x},\mathbf{v}),V(s,\mathbf{x},\mathbf{v}))ds\right). \tag{1.22}$$

Along the characteristics, we have

$$f(t,X(t,\mathbf{x},\mathbf{v}),V(t,\mathbf{x},\mathbf{v})) = f_0(\mathbf{x},\mathbf{v})(det\ J(t,\mathbf{x},\mathbf{v}))^{-1}.$$

Plug in (1.22) to the equation. It is sufficient to prove that $f(t,\cdot,\cdot) \in L^\infty_{\mathbf{x},\mathbf{v}}$ in any finite time as long as $\|A\|_{L^\infty}$ is finite. And we claim that it is true for kinetic flocking systems (1.9).

In fact, under (C-S) setup, we obtain

$$|\partial_{\mathbf{x}} L(f)| = \left| \iint \partial_{\mathbf{x}} \phi(|\mathbf{x}-\mathbf{y}|)(\mathbf{v}^*-\mathbf{v}) f(\mathbf{y},\mathbf{v}^*) d\mathbf{y} d\mathbf{v}^* \right| \le \|\phi\|_{\dot{W}^{1,\infty}}[V]m,$$
$$|\partial_{\mathbf{v}} L(f)| = \left| -\iint \phi(|\mathbf{x}-\mathbf{y}|) f(\mathbf{y},\mathbf{v}^*) d\mathbf{y} d\mathbf{v}^* \right| = \Phi(\mathbf{x}) \le m.$$

Here, m is the total mass. As ϕ satisfies (1.8), in particular ϕ is Lipschitz, hence A is

bounded in all time. Similarly, under (M-T) setup,

$$
\begin{aligned}
|\partial_{\mathbf{x}} L(f)| &= \left| \iint \partial_{\mathbf{x}} \left(\frac{\phi(|\mathbf{x}-\mathbf{y}|)}{\Phi(\mathbf{x})} \right) (\mathbf{v}^* - \mathbf{v}) f(\mathbf{y},\mathbf{v}^*) d\mathbf{y} d\mathbf{v}^* \right| \\
&\leq [V] \iint \frac{|\partial_{\mathbf{x}}\phi(|\mathbf{x}-\mathbf{y}|) f(\mathbf{y},\mathbf{v}^*)\Phi(\mathbf{x}) - \phi(|\mathbf{x}-\mathbf{y}|)\partial_{\mathbf{x}} f(y,v^*)\Phi(\mathbf{x})|}{\Phi^2(\mathbf{x})} d\mathbf{y} d\mathbf{v}^* \\
&\leq 2[V] \iint \frac{|\partial_{\mathbf{x}}\phi(|\mathbf{x}-\mathbf{y}|) f(\mathbf{y},\mathbf{v}^*)|}{\Phi(\mathbf{x})} d\mathbf{y} d\mathbf{v}^* \leq \frac{2\|\phi\|_{\dot{W}^{1,\infty}}[V]m}{\phi(D)}, \\
\partial_{\mathbf{v}} L(f) &= -\iint \frac{\phi(|\mathbf{x}-\mathbf{y}|)}{\Phi(\mathbf{x})} f(\mathbf{y},\mathbf{v}^*) d\mathbf{y} d\mathbf{v}^* = -1.
\end{aligned}
$$

Note that a normalization factor $\Phi(\mathbf{x})$ makes the estimate on $\partial_{\mathbf{x}} L(f)$ more delicate. Thanks to unconditional flocking property, we are able to obtain a positive lower bound on $\phi(\mathbf{x})$. As long as the denominator is bounded away from zero, we conclude with the desired estimate. Here, we recall D the uniform bound on the total variation in position $S(t)$, defined in (1.7).

For existence of classical solutions, we need to bound $\nabla_{(\mathbf{x},\mathbf{v})} f$.

$$
\begin{aligned}
(\nabla f)(t,X(t),V(t)) &= J^{-1}(t)\nabla f_0(\mathbf{x},\mathbf{v}) \exp\left(-\int_0^t \mathrm{tr}A(s,X(s),V(s))ds \right) \\
&\quad + f_0(\mathbf{x},\mathbf{v}) \exp\left(-\int_0^t \mathrm{tr}A(s,X(s),V(s))ds \right) \int_0^t J(s)\ (\nabla \mathrm{tr}A)(s,X(s),V(s))ds.
\end{aligned}
$$

As $\|A\|_{L^\infty}$ is bounded, it is clear that J and J^{-1} are bounded point-wise by e^{Ct}, thanks to (1.20) and (1.21). To obtain boundedness of ∇f, we are left to estimate $\nabla \mathrm{tr}A = \nabla \partial_{\mathbf{v}} L(f)$. Notice that $L(f)$ is linear in v for both setups. Hence, $\partial_{\mathbf{v}}^2 L(f) = 0$.

Compute $\partial_{\mathbf{x}}\partial_{\mathbf{v}} L(f)$ for (C-S) setup:

$$
|\partial_{\mathbf{x}}\partial_{\mathbf{v}} L(f)| = \left| -\iint \partial_{\mathbf{x}}\phi(|\mathbf{x}-\mathbf{y}|) f(\mathbf{y},\mathbf{v}^*) d\mathbf{y} d\mathbf{v}^* \right| \leq \|\phi\|_{\dot{W}^{1,\infty}} m.
$$

For (M-T) setup, as $\partial_{\mathbf{v}} L(f) = -1$, it directly implies $\partial_{\mathbf{x}}\partial_{\mathbf{v}} L(f) = 0$.

We end up with global existence of classical solutions with

$$\|f(t,\cdot,\cdot)\|_{W^{1,\infty}} \leq \|f_0\|_{W^{1,\infty}} e^{Ct}.$$

Chapter 2: Macroscopic Flocking Models

In this chapter, we develop a global existence theory for macroscopic flocking models.

Macroscopic flocking models can be viewed as compressible Eulerian dynamics coupled with non-local alignment forcing, as stated in section 1.3.3. It is well known that compressible Euler system will lose C^1 regularity in finite time, due to formation of shock discontinuities caused by nonlinear convection. On the other hand, the forcing term regularizes the system and prevents finite time break down. This brings mathematical challenges to quantify and balance the competition between convection and forcing.

One prototype problem is the Euler-Poisson system which models plasma with ions and electrons interactions. The competition between convection and Poisson forcing is difficult to compare, and whether the system has global existence of smooth solution is still open, except for the 1D case, where a *critical threshold phenomenon* has been shown in [33]: the initial configurations are divided into two parts through a critical threshold, where subcritical initials lead to global smooth solution, while supercritical initials lead to finite time break down of the system. The critical threshold beautifully expresses the competition, where forcing dominates in the subcritical case, and convection dominates in the supercritical case. The major difficulty to extend the result to higher dimension is

due to the fact that Poisson forcing becomes non-local for dimension 2 or higher.

We study the critical threshold phenomenon for macroscopic flocking models. The alignment forcing is always non-local. We use L^∞ estimate to handle the non-locality and derive a subcritical region for the initial configuration where alignment forcing dominates and global strong solution exists. Meanwhile, there is a supercritical region where the system has a finite time break down. There is a gap between the two regions due to non-locality of the alignment forcing.

As discussed in section 1.3.4, *strong solution must flock*. It implies that the system converges to a flock if starting with subcritical initial data. Moreover, we take advantage of the flocking property and achieve a much larger subcritical region, where the existence of global strong solution is guaranteed.

This chapter is organized as follows. We start in section 2.1 with a review of known results for systems related to macroscopic Cucker-Smale system (2.1). Section 2.2 is devoted to introduce the key tool to study macroscopic flocking models: critical thresholds in Eulerian dynamics. The prototype Euler-Poisson system, interesting from its own sake, is introduced in section 2.2.3-2.2.4, where we address the difficulty for analyzing the 2D system due to non-locality. More discussions regarding 2D Euler-Poisson system are presented in section 2.8. In section 2.3-2.7, we focus on global existence theory for macroscopic flocking systems. A 1D critical threshold result for Cucker-Smale system is presented in section 2.3, where we make full use of the flocking property and derive a large subcritical region. If the initial configuration lies inside the region, there is global existence of strong solution, and flocking behavior follows due to the fact that strong solution must flock. The result can be extended to the more complicated 2D system (section

2.4) and vacuum area (section 2.5). Higher regularity and integrability are also discussed in section 2.6. Finally, we claim in section 2.7 that a similar theory could be established for macroscopic Motsch-Tadmor system as well.

2.1 Macroscopic Cucker-Smale and related systems

In this section, we briefly discuss macroscopic Cucker-Smale system

$$\rho_t + div(\rho \mathbf{u}) = 0, \qquad \mathbf{x} \in \mathbb{R}^n, t \geq 0, \tag{2.1a}$$

$$(\rho \mathbf{u})_t + div(\rho \mathbf{u} \otimes \mathbf{u}) + \nabla P = \int_{\mathbb{R}^n} \phi(|\mathbf{x}-\mathbf{y}|)(\mathbf{u}(\mathbf{y}) - \mathbf{u}(\mathbf{x}))\rho(\mathbf{x})\rho(\mathbf{y})d\mathbf{y}, \tag{2.1b}$$

derived in section 1.3.3. With different choice of the influence function ϕ, (2.1) is related to some systems which has been studied from different prospectus.

We proceed with a survey on known results from the literature for related systems. Techniques and details are not going to be addressed in this section.

2.1.1 Local dissipation

Assume ϕ is bounded and decay sufficiently fast at infinity, such that $\int_0^\infty \phi(r) r^{n+1} dr$ is finite. We process a hyperbolic scaling $(\mathbf{x},t) \to \left(\frac{\mathbf{x}}{\varepsilon}, \frac{t}{\varepsilon}\right)$ with $(\rho, \mathbf{u}, P) \to \left(\frac{\rho}{\varepsilon}, \mathbf{u}, \frac{P}{\varepsilon}\right)$ and get the following scaled system

$$\rho_t + div(\rho \mathbf{u}) = 0,$$

$$(\rho \mathbf{u})_t + div(\rho \mathbf{u} \otimes \mathbf{u}) + \nabla P = \int_{\mathbb{R}^n} \phi_\varepsilon(|\mathbf{x}-\mathbf{y}|)(\mathbf{u}(\mathbf{y}) - \mathbf{u}(\mathbf{x}))\rho(\mathbf{x})\rho(\mathbf{y})d\mathbf{y}.$$

where

$$\phi_\varepsilon := \frac{1}{\varepsilon^{n+2}} \phi\left(\frac{|\mathbf{x}|}{\varepsilon}\right).$$

When $\varepsilon \to 0$, the interaction becomes local. Through a formal calculation, consulting appendix section 2.A, we derive the limiting local system

$$\rho_t + div\,(\rho \mathbf{u}) = 0, \tag{2.2a}$$

$$(\rho \mathbf{u})_t + div\,(\rho \mathbf{u} \otimes \mathbf{u}) + \nabla P = C div\,(\mu(\rho)\nabla \mathbf{u})\,. \tag{2.2b}$$

where the viscosity coefficient $\mu(\rho) = \rho^2$, and the constant $C = \dfrac{\omega_n}{2n}\displaystyle\int_0^\infty \phi(r) r^{n+1} dr$, with ω_n denoting the surface area of a unit sphere in $\mathbb{R}^n$. System (2.2) belongs to the class of compressible Navier-Stokes equations with degenerate viscosity $\mu = \rho^\theta$, $\theta > 0$, which vanishes at the vacuum.

The study of such equations is mostly limited to one dimension. For existence and uniqueness of the weak solution with "moderate degeneracy", $\theta < 1/2$, we refer to [67, 104, 107]. Mellet and Vasseur in [74] proved global existence and uniqueness of the strong solution. They assume $\rho_0 > 0$ and show that $\rho(t) > 0$ for all $t \geq 0$, namely, vacuum never appears and the system is in fact not degenerate in all time.

However, with the presence of vacuum, global regularity of the system is not known. In fact, for the 1D pressure-less case, (2.2b) acts like inviscid Burgers equation at vacuum. It is well known that initial C^1 regularity will lose in finite time for general non-increasing initial data.

2.1.2 Fractional dissipation

One common way to introduce non-locality to the viscous term is to consider fractional dissipation, with

$$\phi(r) = r^{-n-2\alpha}.$$

Here, ϕ has a singularity at the origin. This setup is widely discussed in the incompressible case, where the density ρ is formally set to be 1 and system (2.1b) reads

$$\mathbf{u}_t + \mathbf{u}\cdot\nabla\mathbf{u} + \nabla p = -(-\Delta)^{\alpha}\mathbf{u},$$

$$div\,\mathbf{u} = 0.$$

L^2-energy bound implies that global smooth solutions exist for $\alpha > \frac{1}{2} + \frac{n}{4}$, consulting *e.g.*, [55, 103].

If we enforce $\rho \equiv 1$ and $p \equiv 0$, (2.1b) in one-dimension becomes fractal Burgers equation

$$u_t + uu_x = -(-\Delta)^{\alpha}u.$$

With additional pointwise bounds, this system admits global solutions for $\alpha > 1/2$. The critical case, $\alpha = 1/2$ was the subject of extensive recent studies [15, 25, 53], using different approaches.

For the compressible system, global regularity is still open, especially for the case when vacuum arises.

2.1.3 Nonlocal alignment

We are interested in models with flocking property. The discussion of unconditional flocking for the agent-based systems in section 1.2.3 suggests that the influence function should satisfy condition (1.8). This condition is violated in both local dissipation and fractional dissipation setups. We shall introduce a third prototype of influence function

$$\phi(r) = (1+r)^{-\alpha} \quad \alpha < 1,$$

where condition (1.8) is satisfied. In contrast with the other two setups, the influence function is *not singular* at the origin, which causes different behaviors of the system.

If we enforce $\rho \equiv 1$, the 1D pressure-less system (2.1b) reads

$$u_t + uu_x = \int_{-\infty}^{\infty} \phi(|x-y|)(u(y)-u(x))dy.$$

It is realized as Burgers equation with a nonlocal source term. Liu and Tadmor in [62] proved critical thresholds phenomenon for the system, where global smooth solution exists with subcritical initial configuration. We postpone the details in section 2.2.4.

2.1.4 Common difficulties

In all three prototypes of influence functions discussed above, the known results are far from complete, and share some common difficulties.

- *Restrictions in dimension.* Most of the results are restricted in 1D or scaler equations. For instance, the beautiful De Giorgi-Nash-Moser theory, which is elegantly applied in [15], is limited to scaler equations. Very few tools can be applied to multi-dimensional systems.

- *Presence of vacuum.* The vacuum is usually difficult to treat with. Take the local system (2.2) as an example. Due to degeneracy, the viscosity term vanishes at vacuum. The system behaves as a compressible Euler equation, where shock occurs in finite time for general initial data.

The heart of the chapter is the study of the full compressible system (2.1) with nonlocal alignment setup. The simplest 1D pressure-less system is discussed in section

2.3. Several extensions will be discussed as well. In particular, we study the 2D system in section 2.4, and the vacuum situation in section 2.5. The common difficulties above could be overcome thanks to the choice of global influence function.

2.2 Eulerian dynamics and critical threshold

In this section, we develop the main tools to study global existence theory for macroscopic flocking models. We shall explain the main ideas in a general framework of Eulerian dynamics.

We introduce the *critical threshold phenomenon* for general Eulerian dynamics in section 2.2.2. It discribes that global regularity depends on initial configurations. A prototype Euler-Poisson system is discussed in section 2.2.3, where the 1D system is perfectly characterized by a critical threshold. On the other hand, the multi-D system can not be resolved in the same fashion, as non-locality enters when $n > 1$. Note that flocking models has non-local forcing as well. We will focus on techniques on Eulerian dynamics with non-local forcing in section 2.2.4.

2.2.1 General setup

A general description of Eulerian dynamics with forcing is expressed in the form

$$\mathbf{u}_t + \mathbf{u} \cdot \nabla \mathbf{u} = \mathbf{F}, \tag{2.3}$$

where $\mathbf{u}$ is the velocity and $\mathbf{F}$ is the external force.

Theorem 2.2.1 (Local well-posedness)**.** *Suppose the initial condition* $\mathbf{u}_0 \in H^s(\mathbb{R}^n)$ *with*

$s > \frac{n}{2} + 1$, and the external force $\mathbf{F} \in L^1([0,T];H^s(\mathbb{R}^n))$. Then there exists $T_0 > 0$ and $\mathbf{u} \in C([0,T_0];H^s(\mathbb{R}^n))$, such that $\mathbf{u}$ solves (2.3) with initial profile $\mathbf{u}_0$, up to time T_0.

Proof. We start with a standard energy estimate. Denote pseudo-differential operator $\Lambda := (I-\Delta)^{1/2}$. Act Λ^s on equation (2.3) and integrate by parts against $\Lambda^s \mathbf{u}$.

$$\frac{1}{2}\frac{d}{dt}\|\mathbf{u}(\cdot,t)\|_{H^s}^2 = -\sum_j \left(\left[\Lambda^s, u_j\right] \partial_j \mathbf{u}, \Lambda^s \mathbf{u}\right) + \frac{1}{2}\left(\Lambda^s \mathbf{u}, (div\,\mathbf{u})\Lambda^s \mathbf{u}\right) + \left(\Lambda^s \mathbf{u}, \Lambda^s \mathbf{F}\right),$$

where $(\cdot,\cdot)$ is the L^2 inner product. A commutator estimate (c.f. [54]) implies the following inequality

$$\| \left[\Lambda^s, u_j\right] \partial_j \mathbf{u}\|_{L^2} \lesssim \|\nabla \mathbf{u}\|_{L^\infty} \|\mathbf{u}\|_{H^s}.$$

Therefore, we get

$$\frac{d}{dt}\|\mathbf{u}(\cdot,t)\|_{H^s} \lesssim \left(\|\nabla \mathbf{u}\|_{L^\infty} + \|\mathbf{F}\|_{H^s}\right)\|\mathbf{u}\|_{H^s}$$

Applying Gronwall's inequality, we end up with

$$\|\mathbf{u}(\cdot,t)\|_{H^s} \lesssim \|\mathbf{u}_0\|_{H^s} \exp\left\{ \int_0^t \left(\|\nabla \mathbf{u}(\cdot,\tau)\|_{L^\infty} + \|\mathbf{F}(\cdot,\tau)\|_{H^s}\right) d\tau \right\}. \tag{2.4}$$

When $s > \frac{n}{2} + 1$, Sobolev imbedding implies $H^s(\mathbb{R}^n) \subset W^{1,\infty}(\mathbb{R}^n)$, so $\|\nabla \mathbf{u}\|_{L^\infty} \lesssim \|\mathbf{u}\|_{H^s}$. The inequality above implies local well-posedness for the system (2.3). □

The proof of theorem 2.2.1 implies more than local well-posedness of the system. In particular, due to (2.4), if $\nabla \mathbf{u}$ is bounded in $\mathbb{R}^n \times [0,T]$, then $\|\mathbf{u}(\cdot,t)\|_{H^s}$ is bounded up to time T. On the other hand, the blowup in $\nabla \mathbf{u}$ implies blowup in $\|\mathbf{u}\|_{H^s}$ by the the same Sobolev imbedding. Hence, global existence of smooth solution of (2.3) is equivalent to global in time boundedness of $\nabla \mathbf{u}$. We conclude with the following proposition.

Proposition 2.2.2 (Global regularity)**.** *Consider* (2.3) *subject to initial condition* $\mathbf{u}_0 \in H^s(\mathbb{R}^n)$ *with* $s > \frac{n}{2} + 1$*, and external force* $\mathbf{F} \in L^1(\mathbb{R}^+; H^s(\mathbb{R}^n))$*. Then, for any finite time* T*, there exists a solution* $\mathbf{u} \in C([0,T]; H^s(\mathbb{R}^n))$*, if and only if,* $\|\nabla \mathbf{u}(\cdot,t)\|_{L^\infty}$ *is bounded for all* $t \in [0,T]$*.*

2.2.2 Critical threshold phenomenon in Eulerian dynamics

Critical threshold phenomenon describes that the boundedness of $\nabla \mathbf{u}$ for Eulerian dynamics depends on the choice of initial profile $\mathbf{u}_0$. We use a simple example to address this phenomenon.

Consider the following 1D Burgers equation with a damping force

$$u_t + uu_x = -\kappa u, \quad \kappa > 0. \tag{2.5}$$

We denote $d = u_x$. Differentiating (2.5) with respect to x yields the following dynamics of d

$$d' = -d^2 - \kappa d,$$

where $' = \partial_t + u\partial_x$ is the material derivative. This Riccati-type ODE has explicit solution along the characteristic curve

$$d(t) = \frac{\kappa}{e^{\kappa t}[1 + \kappa/d(0)] - 1}.$$

The solution has different behaviors given different initial profile $d(0)$:

- If $d(0) \geq -\kappa$, then $d(t)$ exists in all time;
- If $d(0) < -\kappa$, then $d(t) \to -\infty$ when $t \to \frac{1}{\kappa} \ln \left[\frac{d(0)}{d(0) + \kappa} \right]$.

In fact, we observe that the contribution from nonlinear convection is the $-d^2$ term, which always drives d to be more negative. The contribution from damping force the $-\kappa d$ term. When d is negative, it prevents d from getting more negative. There is a competition between the two terms. In the subcritical case where $d(0) \geq -\kappa$, the forcing term dominates, and d is bounded in all time. In the supercritical case where $d(0) < -\kappa$, the convection term dominates, and d goes to $-\infty$ in finite time.

Combining across the fan of all characteristics, we conclude with the critical threshold phenomenon for (2.5). The initial configurations are divided into two parts:

- *Subcritical:* if initially $\inf_x u_0(x) \geq -\kappa$, then u_x is bounded in all time;
- *Supercritical:* if initially $\inf_x u_0(x) < -\kappa$, then $\|u_x(\cdot,t)\|_{L^\infty}$ blows up in finite time.

This result can be used immediately to develop a global existence theory of system (2.5), thanks to proposition 2.2.2. Hence, the understanding of the critical threshold phenomenon is very important and powerful for Eulerian dynamics.

For general system (2.3), $M := \nabla \mathbf{u}$ is an $n \times n$ matrix, with the following dynamics

$$M' + M^2 = \nabla \mathbf{F}, \tag{2.6}$$

where $' = \partial_t + \mathbf{u} \cdot \nabla_{\mathbf{x}}$ is again the material derivative. It requires more efforts to study the boundedness of M for higher dimensions systems.

It is natural to use $d := div\,\mathbf{u}$ to play the role of u_x in one-dimension. As $d = \mathrm{tr}(M)$, we apply trace operator on (2.6) and get the dynamics of d:

$$d' + \mathrm{tr}(M^2) = div\,\mathbf{F}.$$

Due to the fact that $\mathrm{tr}(M^2) \neq (\mathrm{tr}(M))^2$ for dimension $n > 1$, we can not simply close the system with the dynamics of d on its own.

Suppose $(\lambda_1, \cdots, \lambda_n)$ are eigenvalues of M, we compute

$$\mathrm{tr}(M^2) = \sum_{i=1}^{n} \lambda_i^2 = \frac{1}{n}\left(\sum_{i=1}^{n} \lambda_i\right)^2 + \frac{1}{2n}\sum_{i=1}^{n}\sum_{j=1}^{n}(\lambda_i - \lambda_j)^2.$$

The dynamics of d can be rewritten as

$$d' + \frac{d^2 + \eta^2}{n} = div\,\mathbf{F}.$$

Here, η is called *spectral gap*, defined through

$$\eta^2 = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(\lambda_i - \lambda_j)^2. \tag{2.7}$$

It measures the difference among eigenvalues of M. η can be either real or purely imaginary if M has real entries.

It is essential to bound the spectral gap in order to derive critical threshold for systems with dimension bigger than 1. Different techniques are used to treat with specific problems.

2.2.3 Pressure-less Euler-Poisson system

One prototype example of studying critical threshold phenomenon is the Pressure-less Euler-Poisson system

$$\rho_t + div\,(\rho\mathbf{u}) = 0, \tag{2.8a}$$

$$\mathbf{u}_t + \mathbf{u}\cdot\nabla\mathbf{u} = -k\nabla\phi, \tag{2.8b}$$

$$-\Delta\phi = \rho. \tag{2.8c}$$

The compressible Euler system (2.8a)-(2.8b) is coupled with a Poisson forcing (2.8c). It models plasma with ion-electron interactions (c.f. [43]).

Critical threshold phenomenon for Euler-Poisson system is first discussed in 1D in [33]. A series of studies follows in order to extend the series into high dimensions and more complicated setups [22, 63, 64, 96]. However, despite the success in 1D, it is extremely difficult to find the precise critical threshold in 2D and higher dimensions.

We briefly go over the known results in this section, and address the main difficulties on solving the full 2D system.

Consider (2.8) in 1D. We apply the procedure discussed in section 2.2.2 to (2.8b) and get

$$d' + d^2 = -k\phi_{xx}.$$

The right hand side is equal to $k\rho$, from the Poisson equation (2.8c). Rewrite the continuity equation (2.8a) along the characteristic curve, we get the following coupled system of ODEs

$$\rho' + \rho d = 0,$$

$$d' + d^2 = k\rho.$$

A study on the dynamics of (ρ, d) along the characteristics will provide the critical threshold. For detailed expressions, consult [33].

Now, consider multi-dimensional Euler-Poisson system (2.8). Again, the analysis

in section 2.2.2 yields the following coupled system of ODEs

$$\rho' + \rho d = 0,$$
$$d' + \frac{d^2 + \eta^2}{n} = k\rho.$$

The major difference compared to 1D system is the presence of spectral gap. In general, it is very difficult to express the dynamics of η.

For the 2D case, $M = \nabla \mathbf{u}$ is a 2×2 matrix. We can write the dynamics of M (2.6) explicitly

$$\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}' + \begin{bmatrix} M_{11}^2 + M_{12}M_{21} & M_{12}(M_{11} + M_{12}) \\ M_{21}(M_{11} + M_{12}) & M_{22}^2 + M_{12}M_{21} \end{bmatrix} = -k \begin{bmatrix} \phi_{x_1x_1} & \phi_{x_1x_2} \\ \phi_{x_1x_2} & \phi_{x_2x_2} \end{bmatrix},$$

Note that $d = div\mathbf{u} = M_{11} + M_{12}$. Moreover, we define $q := M_{11} - M_{22}, r := M_{12} + M_{21}, \omega := \nabla \times \mathbf{u} = M_{12} - M_{21}$. The spectral gap can be express in terms of (p, r, ω) as the following

$$\eta^2 = q^2 + r^2 - \omega^2.$$

Meanwhile, the dynamics of (q, r, ω) along the characteristic curves are relatively easier to trace. We derive the following system

$$\rho' + \rho d = 0, \tag{2.9a}$$
$$d' + \frac{d^2 + \eta^2}{2} = k\rho, \tag{2.9b}$$
$$q' + qd = k(\phi_{x_2x_2} - \phi_{x_1x_1}), \tag{2.9c}$$
$$r' + rd = -2k\phi_{x_1x_2}, \tag{2.9d}$$
$$\omega' + \omega d = 0. \tag{2.9e}$$

In this way, we can express and propagate the spectral gap η along the characteristics in dimension 2.

Let $\beta_q = q/\rho$. From (2.9a) and (2.9c), we get

$$\beta_q' = \frac{q'\rho - \rho' q}{\rho^2} = k\frac{\phi_{x_2x_2} - \phi_{x_1x_1}}{\rho} = k\frac{(R_1R_1 - R_2R_2)[\rho]}{\rho}.$$

Here, R_j is the Reisz transform, which is a Fourier multiplier with symbol $i\xi_j/|\xi|$. It follows from the fact $\phi_{x_{j_1}x_{j_2}} = -(\Delta^{-1}\rho)_{x_{j_1}x_{j_2}}$ has symbol $\xi_{j_1}\xi_{j_2}/|\xi|^2$. Thus $\phi_{x_{j_1}x_{j_2}} = -R_{j_1}R_{j_2}[\rho]$.

Similarly, we define $\beta_r = r/\rho$ and $\beta_\omega = \omega/\rho$. System (2.9) is equivalent to

$$\rho' = -\rho d, \tag{2.10a}$$

$$d' = -\frac{d^2 + \rho^2(\beta_q^2 + \beta_r^2 - \beta_\omega^2)}{2} + k\rho, \tag{2.10b}$$

$$\beta_q' = k\frac{(R_1R_1 - R_2R_2)[\rho]}{\rho}, \tag{2.10c}$$

$$\beta_r' = -2k\frac{R_1R_2[\rho]}{\rho}, \tag{2.10d}$$

$$\beta_\omega' = 0. \tag{2.10e}$$

Now, we address the major difficulty of solving system (2.10): *non-locality*. Different from 1D system, the Reisz transform is a non-local operator in 2D. Hence, information along a characteristic curve is not enough to propagate the system. We will discuss techniques on solving Eulerian dynamics with non-local forcing in section 2.2.4.

To avoid non-locality, Liu and Tadmor in [64] introduce a restricted Euler-Poisson model, where $D^2\phi$ is substitute by a matrix with local entries:

$$\begin{bmatrix} \rho/2 & 0 \\ 0 & \rho/2 \end{bmatrix}.$$

The new matrix preserves the trace and the curl of $D^2\phi$.

For the restricted model, it is easy to observe

$$\beta_q' = \beta_r' = \beta_\omega' = 0.$$

Therefore,

$$\beta := \beta_q^2 + \beta_r^2 - \beta_\omega^2$$

is a constant in time along every characteristic curve.

Critical threshold can be derived by analyzing the dynamics of the following system of ODEs along the characteristics:

$$\rho' = -\rho d, \tag{2.11a}$$

$$d' = -\frac{d^2 + \beta\rho^2}{2} + k\rho. \tag{2.11b}$$

We refer the details to [64].

For the full 2D Euler-Poisson system, the problem is still open, due to the presence of non-local Reisz operator. Recent progress to treat with non-locality will be discussed at the end of the chapter (section 2.8).

2.2.4 Eulerian dynamics with non-local forcing

Many systems of Eulerian dynamics involve non-local forcing. One example is Euler-Poisson equations in 2D or higher dimensions, as discussed in section 2.2.3. Another example is macroscopic flocking systems (*e.g.* (2.1)) where non-local alignment forcing is coupled with compressible Eulerian dynamics.

As stated in section 2.2.3, non-local forcing requires information in the whole space, thus the propagation of the system along different characteristic curves becomes tangled.

In order to localize the system on each characteristic curve, the following two steps should be accomplished.

- A uniform estimate for the non-local term.
- A valid comparison principle.

We use an elementary example [64] to illustrate how to handle the non-locality through the two steps above. Consider 1D Burgers equation with source term of convolution type. This system has been introduced earlier in section 2.1.3. Recall the equation

$$u_t + uu_x = \int_{-\infty}^{\infty} \phi(x-y)(u(y)-u(x))dy. \tag{2.12}$$

For simplicity, we assume $\int_{-\infty}^{\infty} \phi(x)dx = 1$.

Differentiate the equation with respect to x. We get the following dynamics of $d = u_x$ along the characteristics.

$$d' + d^2 = -d + \int_{-\infty}^{\infty} \phi_x(x-y)(u(y)-u(x))dy.$$

The last term is non-local. We perform with the following uniform estimate

$$\left| \int_{-\infty}^{\infty} \phi_x(x-y)(u(y)-u(x))dy \right| \leq C := \|\phi\|_{\dot{W}^{1,\infty}} \left(\sup_x u_0(x) - \inf_x u_0(x) \right).$$

Note that the inequality follows from a maximum (and minimum) principle on u. See [64] for details.

The bound C only depend on initial quantities and the choice of ϕ, which are given apriori. It is easy to study the following local system by substituting the non-local term with its uniform bound

$$e' + e^2 = -e + C.$$

Indeed, we state the following critical threshold phenomenon:

- If $e(0) \geq \frac{1-\sqrt{1+4C}}{2}$, then $e(t)$ exists in all time;
- If $e(0) < \frac{1-\sqrt{1+4C}}{2}$, then $e(t) \to -\infty$ in finite time.

To extend the critical threshold result to the non-local dynamics, we link the two systems with a comparison principle.

Lemma 2.2.3 (Comparison principle)**.** *If $d(0) \leq e(0)$, then $d(t) \leq e(t)$.*

Proof. Subtracting the dynamics of e from the dynamics of d, we get

$$(d-e)' = -(d^2-e^2) - (d-e) + \left(\int_{-\infty}^{\infty} \phi_x(x-y)(u(y)-u(x))dy - C \right) \leq -(d-e)(d+e+1).$$

Suppose by contradiction $d(t) > e(t)$ for some $t > 0$. As d, e are continuous in t, there exists $\tau \in [0,t)$ such that $d(\tau) - e(\tau) = 0$ and $(d(\tau) - e(\tau))' > 0$. This violates the inequality above. Therefore, $d(t) \leq e(t)$ for all t. □

A supercritical threshold on d follows directly from the comparison principle, by comparing with a local dynamics starting at $e(0) \in \left(d(0), \frac{1-\sqrt{1+4C}}{2} \right)$:

- *Supercritical:* if $d(0) < \frac{1-\sqrt{1+4C}}{2}$, then $d(t) \to -\infty$ in finite time.

Similarly, a subcritical threshold can be generated by comparing with

$$e' + e^2 = -e - C.$$

Remark 2.2.1. A drawback of this approach is that there is a gap between the subcritical region and supercritical region, due to the loss from the uniform estimate. However, this is the only way to deal with non-locality if we want to study the system along the characteristics.

2.3 Critical threshold for 1D macroscopic pressure-less Cucker-Smale system

This section is devoted to study critical threshold phenomenon and develop global existence theory for macroscopic flocking systems. The first system to start with is the 1D macroscopic pressure-less Cucker-Smale systems

$$\rho_t + (\rho u)_x = 0, \quad x \in \mathbb{R}, t \geq 0, \tag{2.13a}$$

$$u_t + u u_x = \int_{-\infty}^{\infty} \phi(|x-y|)(u(y)-u(x))\rho(y)dy. \tag{2.13b}$$

subject to initial condition

$$\rho(x,0) = \rho_0(x), \quad u(x,0) = u_0(x). \tag{2.13c}$$

Recall that ϕ is the influence function, which satisfies (1.8). The initial density ρ_0 is compactly supported, and the initial velocity u_0 is bounded.

We state the main theorem before getting into details. The following two initial quantities play an important role in the theorem:

$$d_0 := \inf_{x \in supp(\rho_0)} u_{0x}(x), \quad V_0 := \sup_{x,y \in supp(\rho_0)} |u_0(x) - u_0(y)|. \tag{2.14}$$

Here, d_0 represents the smallest slope of u_0, and V_0 represents the maximal variation of the initial velocity.

Theorem 2.3.1 (Critical Thresholds for 1D Cucker-Smale system)**.** *Consider initial value problem of* (2.13)*. There exists threshold functions $\sigma_+ > \sigma_-$, such that*

- *Subcritical: If the initial condition satisfies*

$$d_0 > \sigma_+(V_0), \tag{2.15}$$

then $u_x(x,t)$ remains bounded for all $(x,t) \in supp(\rho)$.

Moreover, strong solution $\rho(\cdot,t) \in L^1(\mathbb{R})$, $u(\cdot,t) \in W^{1,\infty}(supp(\rho(t)))$ exists for all $t \geq 0$, and converges to a flock in the sense of definition 1.3.2.

- *Supercritical: If the initial condition satisfies*

$$d_0 < \sigma_-(V_0),$$

then there exists a finite time T and a position $x^* \in supp(\rho(\cdot,T))$, $u_x(x^*,T) \to -\infty$.

Remark 2.3.1. We use figure 2.1 to illustrate the two thresholds. Detailed expressions of threshold functions σ_+ and σ_- are given in (2.23). To ensure boundedness of u_x, there are two requirements for the initial configurations:

– Initial slop of velocity u_{0x} is not too negative,

– Initial variation of velocity V_0 is not too large.

Note that one steady state of the system is flocking, when $u = \bar{u}_0$, *i.e.*, the velocity is constant. The subcritical threshold condition says that if initial configuration is not far away from equilibrium, then strong solution exists globally and converges to the steady state.

Remark 2.3.2. For both subcritical and supercritical thresholds, there is a darker area and a much larger lighter area shown in figure 2.1. The darker areas represent the thresholds stated in theorem 2.3.4. It is an extension of critical threshold result for 1D Burgers equation with nonlocal source (2.12), as discussed in section 2.2.4. Taking advantage of the fast alignment property of the system, we improve the result to the much lighter area. See section 2.3.2 for details.

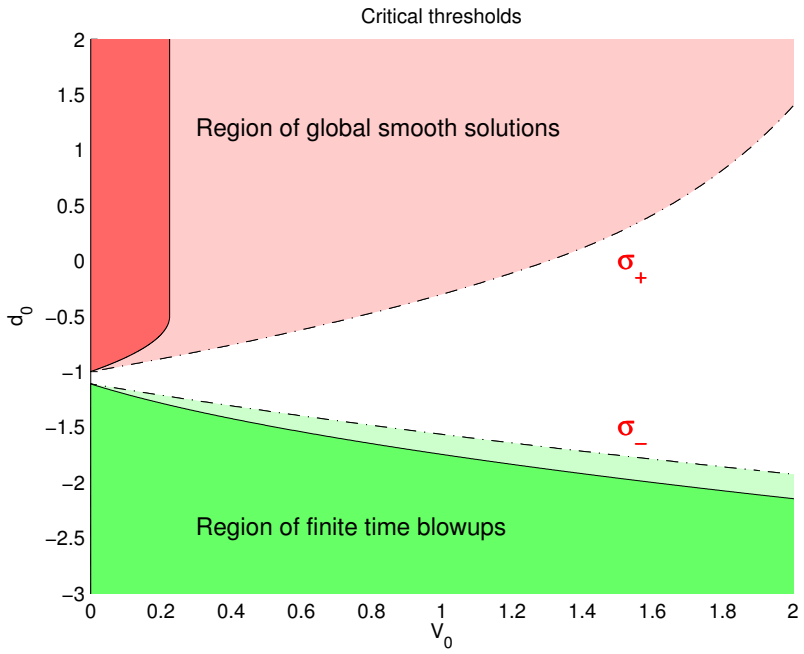

Figure 2.1: Illustration of the critical thresholds in one dimension.

2.3.1 A first result on the thresholds

To study the critical threshold phenomenon for (2.13), we follow the technique introduced in section 2.2.4, as the alignment forcing is non-local.

Differentiate (2.13b) with respect to x. We get the dynamics of $d = u_x$ along the characteristic curves:

$$d' + d^2 = \int_{-\infty}^{\infty} \phi_x(|x-y|)(u(y)-u(x))\rho(y)dy - d\int_{-\infty}^{\infty} \phi(|x-y|)\rho(y)dy. \tag{2.16}$$

Again, $' = \partial_t + u\partial_x$ is the material derivative. The time variable is omitted for simplicity, unless necessary.

There are two non-local terms in this dynamics. We denote

$$p(x) := \int_{-\infty}^{\infty} \phi(|x-y|)\rho(y)dy,$$
$$Q(x) := \int_{-\infty}^{\infty} \phi_x(|x-y|)(u(y)-u(x))\rho(y)dy.$$

Remark 2.3.3. We use lower-case p and upper-case Q in order to be compatible with the 2D case, where p is a scaler and Q is a 2×2 matrix.

As discussed in section 2.2.4, the idea of treating non-local terms is to establish uniform bounds and proceed with a comparison principle. We prove that (2.16) is an example of the following prototype problem:

$$d' = -d^2 - pd + Q, \quad \text{where } 0 < \gamma \le p \le \Gamma \text{ and } |Q| \le c, \tag{2.17}$$

where γ, Γ, c are positive constant coefficients.

In particular, these coefficients are given below for (2.16), shown in proposition 2.3.2.

$$\gamma = \phi(D)m, \quad \Gamma = m, \quad c = V_0\|\phi\|_{\dot{W}^{1,\infty}}m. \tag{2.18}$$

As a reminder, $m := \int_{-\infty}^{\infty} \rho(x)dx$ is the total mass, which is bounded and conserved in time.

Proposition 2.3.2. *Suppose (ρ, u) is a strong solution of system* (2.13)*. Then, for any $x \in supp(\rho(t))$,*

$$\phi(D)m \le \int_{-\infty}^{\infty} \phi(|x-y|)\rho(y)dy \le m.$$
$$\left| \int_{-\infty}^{\infty} \phi_x(|x-y|)(u(y)-u(x))\rho(y)dy \right| \le V_0\|\phi\|_{\dot{W}^{1,\infty}}m,$$

Proof. For the first inequality,

$$\int_{-\infty}^{\infty} \phi(|x-y|)\rho(y)dy \leq \|\phi\|_{L^\infty} m = m.$$

On the other hand, as (ρ, u) is a strong solution, theorem 1.3.6 implies that it converges to a flock. In particular, $S(t)$ is uniformly bounded by D, defined in (1.7). Therefore,

$$\int_{-\infty}^{\infty} \phi(|x-y|)\rho(y)d\mathbf{y} = \int_{supp(\rho(t))} \phi(|x-y|)\rho(y,t)dy \geq \phi(D)\int_{-\infty}^{\infty} \rho(y,t)d\mathbf{y} = \phi(D)m,$$

for all $x \in supp(\rho(t))$.

For the second inequality,

$$\begin{aligned}\left|\int_{-\infty}^{\infty} \phi_x(|x-y|)(u(y)-u(x))\rho(y)dy\right| &\leq \int_{-\infty}^{\infty} |u(y)-u(x)|\,|\phi_x(|x-y|)|\,\rho(y)dy \\ &\leq V(t)\|\phi\|_{\dot{W}^{1,\infty}} \int_{-\infty}^{\infty} \rho(y)dy \leq V_0\|\phi\|_{\dot{W}^{1,\infty}} m = V_0\|\phi\|_{\dot{W}^{1,\infty}} m.\end{aligned}$$

□

Remark 2.3.4. In the proof of the second inequality, recall

$$V(t) = \max_{x\in supp(\rho(t))} u(x,t) - \min_{x\in supp(\rho(t))} u(x,t)$$

the total variance of u at time t. It is easy to prove $V(t) \leq V(0)$ through a maximum principle of u, see lemma 2.6.2. As a matter of fact, theorem 1.3.6 provides a better estimate $V(t) \leq V_0 e^{-m\phi(D)t}$, which will be used to improve the result in section 2.3.2.

We proceed to discuss the evolution of the initial value problem of (2.17). The same comparison principle in section 2.2.4 can be applied and it yields the following threshold result.

Proposition 2.3.3 (1D dynamics)**.** *Consider initial value problem of* (2.17)*. We have the following*

- *Subcritical: If* $\gamma^2 - 4c \geq 0$ *and* $d(0) \geq -(\gamma + \sqrt{\gamma^2 - 4c})/2$*, then* $d(t)$ *is bounded for all time* $t \geq 0$*.*

- *Supercritical: If* $d(0) < -(\Gamma + \sqrt{\Gamma^2 + 4c})/2$*, then* $d(t) \to -\infty$ *in finite time.*

For Cucker-Smale system (2.13), the constants γ, Γ and c are given in (2.18). Applying proposition 2.3.3 to all characteristic paths, we derive critical thresholds for boundedness of u_x in the inside the support of ρ.

Theorem 2.3.4 (1D critical thresholds)**.** *Consider initial value problem of* (2.13)

- *Subcritical: If the initial configuration satisfies*

$$V_0 \leq \frac{\phi(D)^2 m}{4\|\phi\|_{\dot{W}^{1,\infty}}} \quad \text{and} \quad d_0 \geq -\frac{1}{2}\left(\phi(D)m + \sqrt{\phi(D)^2 m^2 - 4V_0\|\phi\|_{\dot{W}^{1,\infty}} m}\right),$$

  then $u_x(x,t)$ *is bounded for all* $(x,t) \in supp(\rho)$*.*

- *Supercritical: If the initial configuration satisfies*

$$d_0 < -\frac{1}{2}\left(m + \sqrt{m^2 + 4V_0\|\phi\|_{\dot{W}^{1,\infty}} m}\right)$$

  then there exists a finite time T *and a position* $x \in supp(\rho(\cdot,T))$*,* $u_x(x,T) \to -\infty$*.*

Remark 2.3.5. The thresholds in theorem 2.3.4 correspond to darker areas in figure 2.1. The result will be improved in section 2.3.2, taking into account of the additional fast alignment property.

2.3.2 Enhanced dynamics with fast alignment

In this section, we improve theorem 2.3.4, by substituting the estimate $V(t) \leq V_0$ in proposition 2.3.2 with a stronger one $V(t) \leq V_0 e^{-m\phi(D)t}$. The improved estimate on $V(t)$ comes from the fast alignment property of flocking systems discussed in section 1.3.4.

We write a new prototype problem

$$d' = -d^2 - pd + Q, \quad p \in [\gamma, \Gamma], \quad Q \in [-cV, cV]. \tag{2.19a}$$

$$\tfrac{d}{dt} V \leq -GV. \tag{2.19b}$$

It is clear that the dynamics (2.16) is an example of (2.19) with coefficients

$$\gamma = \phi(D)m, \quad \Gamma = m, \quad c = \|\phi\|_{\dot{W}^{1,\infty}} m, \quad G = -m\phi(D). \tag{2.20}$$

Remark 2.3.6. The new dynamics of d is coupled with the dynamics of V, where $V(t)$ vanishes to 0 exponentially fast. As time goes by, the influence of Q (which is a "bad term") on the dynamics of d becomes weaker and weaker. Therefore, we expect a wider set of initial configurations which ensures the boundedness of d. Better critical thresholds could be derived just as illustrated in figure 2.1.

The following theorem characterizes the dynamics of (d, V). The proof will be provided in the next section.

Theorem 2.3.5. *Consider initial value problem of* (2.19)*. We have the following*

- *There exists a continuous function $\sigma_+ : \mathbb{R}^+ \to [-\gamma, +\infty)$, defining implicitly as*

$$\sigma_+(0) = -\gamma, \ \sigma_+'(x) = \begin{cases} \frac{c}{\gamma + G}, & x \to 0+ \\ \frac{-\sigma_+(x)^2 - \gamma\sigma_+(x) - cx}{-Gx} & \text{if } \sigma_+(x) < 0 \\ \frac{-\sigma_+(x)^2 - \Gamma\sigma_+(x) - cx}{-Gx} & \text{if } \sigma_+(x) \geq 0 \end{cases} \tag{2.21}$$

  such that, if $d_0 > \sigma_+(V_0)$, i.e. $(V_0, d(0))$ lies above σ_+, then (V, d) are bounded all time, and $d(t) \to 0, V(t) \to 0$ as $t \to \infty$.

- *There exists a function* $\sigma_- : \mathbb{R}^+ \to (-\infty, -\Gamma]$*, defining implicitly as*

$$\sigma_-(0) = -\Gamma,\ \sigma_-'(x) = \begin{cases} -\frac{c}{\Gamma+G}, & x \to 0+ \\ \frac{-\sigma_-(x)^2 - \Gamma\sigma_-(x) + cx}{-cx} & x > 0. \end{cases} \tag{2.22}$$

such that, if $d_0 < \sigma_-(V_0)$*, i.e.* $(V_0, d(0))$ *lies below* σ_-*, then* $d(t) \to -\infty$ *in finite time.*

Apply theorem 2.3.5 to Cucker-Smale system (2.13) by plugging in the values of the constants given in (2.20) and combine all characteristic paths. We conclude with theorem 2.3.1 with the following threshold functions.

$$\sigma_+(0) = -\phi(D)m,\ \sigma_+'(x) = \begin{cases} \frac{\|\phi\|_{\dot{W}^{1,\infty}}}{2\phi(D)} & x \to 0+ \\ \frac{-\sigma_+(x)^2 - \phi(D)m\sigma_+(x) - \|\phi\|_{\dot{W}^{1,\infty}}mx}{-\phi(D)mx} & \text{if } \sigma_+(x) < 0\,, \\ \frac{-\sigma_+(x)^2 - m\sigma_+(x) - \|\phi\|_{\dot{W}^{1,\infty}}mx}{-\phi(D)mx} & \text{if } \sigma_+(x) \geq 0 \end{cases} \tag{2.23a}$$

$$\sigma_-(0) = -m,\ \sigma_-'(x) = \begin{cases} -\frac{\|\phi\|_{\dot{W}^{1,\infty}}}{1+\phi(D)} & x \to 0+ \\ \frac{-\sigma_-(x)^2 - m\sigma_-(x) + \|\phi\|_{\dot{W}^{1,\infty}}mx}{-\phi(D)mx} & x > 0 \end{cases}. \tag{2.23b}$$

Remark 2.3.7. The additional fast alignment property enables us to establish a much larger area of (V_0, d_0) such that u_x is bounded in the non-vacuum area. In particular, the crucial upper bound of V_0 is not any more required, see figure 2.1.

Remark 2.3.8. To further reduce the gap between σ_+ and σ_-, one can trace the dynamics of $d + \phi \star \rho$ along the characteristics, namely $\sigma_+ = \sigma_-$. It provides a perfect threshold for 1D Cucker-Smale system. This is an ongoing work joint with Carrillo, Choi and Tadmor.

2.3.3 Proof of the key theorem

The proof of the key theorem 2.3.5 can be separate into two parts. First, we discuss the evolution of the equality system

$$\frac{d}{dt}\omega = -\omega^2 - E\omega + F\eta, \tag{2.24a}$$

$$\frac{d}{dt}\eta = -G\eta, \tag{2.24b}$$

where $E > 0, F \in \mathbb{R}, G > 0$ are constant coefficients. Then, we state a comparison principle to compare (d,V) with (ω,η) and therefore derive the evolution of the inequality system (2.19).

The evolution of system (2.24) is summarized in the following proposition.

Proposition 2.3.6 (Critical threshold for the equality system)**.** *Suppose* $(\eta(t),\omega(t))$ *satisfy* (2.24) *where* $\eta(t) \geq 0$*, with initial condition* $\omega(0) = \omega_0$*,* $\eta(0) = \eta_0 > 0$*. Then,*

- *If* $\omega_0 > f(\eta_0)$*, i.e.* (η_0,ω_0) *lies above* f*, we have* $\omega(t) \to 0, \eta(t) \to 0$ *as* $t \to \infty$*,*
- *If* $\omega_0 = f(\eta_0)$*, i.e.* (η_0,ω_0) *lies on* f*, we have* $\omega(t) \to -E, \eta(t) \to 0$ *as* $t \to \infty$*,*
- *If* $\omega_0 < f(\eta_0)$*, i.e.* (η_0,ω_0) *lies below* f*, we have* $\omega(t) \to -\infty$ *as* $t \to \infty$*,*

where f *is a function defined on* $\mathbb{R}^+$*, such that*

$$f(0) = -E,\ f'(0) = -\frac{F}{E+G},\ f'(x) = \frac{-f(x)^2 - Ef(x) + Fx}{-Gx} \ \text{ for } \ x \in (0,+\infty). \tag{2.25}$$

Proof. The system has two stationary points, $O(0,0)$ and $A(0,-E)$. To study the stability property of these points, we consider the linearization of the coefficient matrix

$$L(\eta,\omega) = \begin{pmatrix} -G & 0 \\ F & -2\omega - E \end{pmatrix},$$

where the two eigenvalues are G and $-2\omega - E$.

At $O(0,0)$, both eigenvalues $-G$ and $-E$ are negative. So it is stable. At $A(0,-E)$, we have a positive eigenvalue E and a negative one $-G$. Hence, it is a saddle.

We use the above facts to construct the critical threshold via the phase place analysis.

Figure 2.2 is the phase plane of (η,ω). $\frac{d}{dt}\omega = 0$ is a parabola and $\frac{d}{dt}\eta = 0$ is a line. There exists a critical curve f, starting from A and travel along the vector field, which divided the plane $\mathbb{R}^+ \times \mathbb{R}$ into two parts. Flows starting above f converge to the stable point O, while flows starting below f will diverge.

Along the line, clearly we have

$$f'(x) = \frac{d\omega}{d\eta} = \frac{\frac{d}{dt}\omega}{\frac{d}{dt}\eta} = \frac{-\omega^2 - E\omega + F\eta}{-G\eta} = \frac{-f(x)^2 - Ef(x) + Fx}{-Gx}.$$

When $x \to 0$, we have

$$f'(0) = \lim_{x\to 0} \frac{-f(x)^2 - Ef(x) + Fx}{-Gx} = \frac{1}{G}\left[f(0)\lim_{x\to 0}\frac{f(x)+E}{x} - F\right] = \frac{-Ef'(0) - F}{G}.$$

It yields $f'(0) = -\dfrac{F}{E+G}$. □

The following lemma states the relationship between the solution of the equality system (2.24) and the inequality system (2.19). It allows us to extend the critical thresholds result to the inequality system.

Lemma 2.3.7 (Comparison principles)**.** *Let* (d,V) *satisfy* (2.19)*, and* (ω,η) *satisfy* (2.24) *with* E,F *defined as below.* $t_0 \geq 0$ *and* $T \in (t_0,+\infty]$.

1. *Suppose* $\omega(t) \geq 0$*, for* $t \in [t_0,T]$.

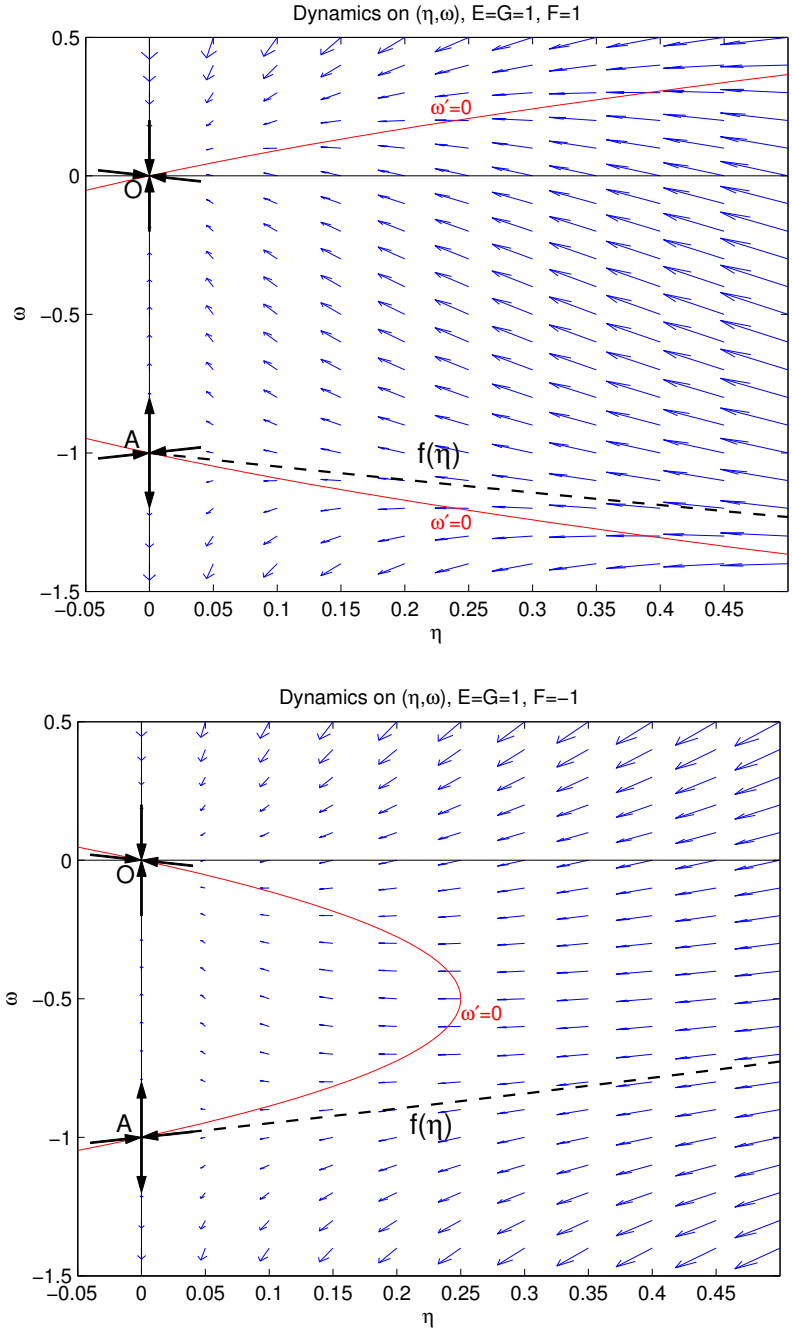

Figure 2.2: Phase plane of the equality system (η, ω). Critical threshold is represented as the dashed curve. The graph above represents the case when F is positive, while the graph below represents the case when F is negative.

(1a) Let $E=\gamma, F=C$. *If* $\begin{cases} d(t_0)\le \omega(t_0) \\ V(t_0)\le \eta(t_0) \end{cases}$ *, then* $\begin{cases} d(t)\le \omega(t) \\ V(t)\le \eta(t) \end{cases}$ *for* $t\in[t_0,T]$.

(1b) Let $E=\Gamma, F=-C$. *If* $\begin{cases} d(t_0)\ge \omega(t_0) \\ V(t_0)\le \eta(t_0) \end{cases}$ *, then* $\begin{cases} d(t)\ge \omega(t) \\ V(t)\le \eta(t) \end{cases}$ *for* $t\in[t_0,T]$.

2. *Suppose* $\omega(t)\le 0$, *for* $t\in[t_0,T]$.

(2a) Let $E=\Gamma, F=C$. *If* $\begin{cases} d(t_0)\le \omega(t_0) \\ V(t_0)\le \eta(t_0) \end{cases}$ *, then* $\begin{cases} d(t)\le \omega(t) \\ V(t)\le \eta(t) \end{cases}$ *for* $t\in[t_0,T]$.

(2b) Let $E=\gamma, F=-C$. *If* $\begin{cases} d(t_0)\ge \omega(t_0) \\ V(t_0)\le \eta(t_0) \end{cases}$ *, then* $\begin{cases} d(t)\ge \omega(t) \\ V(t)\le \eta(t) \end{cases}$ *for* $t\in[t_0,T]$.

Proof. We only prove (1a) as the proofs of the others are the same.

Subtracting (2.19) with (2.24), we get

$$\frac{d}{dt}(\omega-d)\ge -(\omega+d)(\omega-d)-p(\omega-d)+C(\eta-V),$$

$$\frac{d}{dt}(\eta-V)\ge -G(\eta-V).$$

Suppose by contradiction $V(t)>\eta(t)$ for some $t\in(t_0,T)$. As V,η are continuous, there exists $\tau\in(t_0,t)$ such that $\eta(\tau)-V(\tau)=0$ and $\frac{d}{dt}(\eta(\tau)-V(\tau))<0$. This violates the second inequality. So, $V(t)\le\eta(t)$ for all $t\in[t_0,T]$.

Similarly, suppose by contradiction $d(t)>\omega(t)$ for some $t\in(t_0,T)$. As V,η are continuous, there exists $\tau\in(t_0,t)$ such that $d(\tau)-\omega(\tau)=0$ and $\frac{d}{dt}(d(\tau)-\omega(\tau))<0$. Meanwhile, $V(\tau)\le\eta(\tau)$. This violates the first inequality. So, $d(t)\le\omega(t)$ for all $t\in[t_0,T]$. □

We use the comparison principles to prove theorem 2.3.5.

Claim 1: $d(t)$ could not diverge to $+\infty$.

Suppose by contradiction $d(t) \to +\infty$ as $t \to T$. Then, there exists a $t_0 \in [0,T)$ such that $d(t_0) > 0$. Construct (ω, η) by (2.24) with $E = \gamma$, $F = C$ and with initial values $\omega(t_0) = d(t_0) > 0, \eta(t_0) = V(t_0)$. From proposition 2.3.6, $\omega(t)$ is either bounded all time or diverges to $-\infty$ in finite time. Therefore, $\omega(t)$ is upper bounded. Comparison principle (1a) implies that $d(t) \leq \omega(t)$ is also upper bounded. Moreover, $d(T) \leq \omega(T) \leq 0$, where T is either infinity or the blowup time.

Claim 2: $d(t) \to -\infty$ in finite time if $d(0) < \sigma_-(V(0))$, where σ_- is defined in (2.22).

Clearly, $d(0) < 0$. Again, construct (ω, η) by (2.24) with $E = \Gamma$, $F = C$ and with initial values $\omega(0) = d(0) < 0, \eta(0) = V(0)$. Note that under this setup, σ_- is the same as f defined in (2.25). (Actually this is how σ_- is determined.) From proposition 2.3.6, $\omega(0) < \sigma_-(\eta(0))$ implies $\omega(t) \to -\infty$ in finite time. Comparison principle (2a) implies that $d(t) \leq \omega(t) \to -\infty$ in finite time.

Claim 3: $d(t), V(t)$ are bounded all time if $d(0) > \sigma_+(V(0))$, where σ_+ is defined in (2.21). Moreover, $d(t) \to 0, V(0) \to 0$ as $t \to \infty$.

First, we assume $d(0) \leq 0$. Similarly, construct (ω, η) by (2.24) with $E = \gamma$, $F = -C$ and with initial values $\omega(0) = d(0) \leq 0, \eta(0) = V(0)$. Define f_l to be the same as f in (2.25) as long as $f_l \leq 0$, with the choice of E, F. Then, using proposition 2.3.6, $\omega(0) > \sigma_+(\eta(0))$ implies $\omega(t)$ is lower bounded all time, and $\omega(t) \to 0$ as $t \to \infty$. Comparison principle (2b) implies that $d(t) \geq \omega(t)$ is also lower bounded in all time, and $\lim_{t\to\infty} d(t) \geq 0$. Combine with the result in claim 1, we conclude that $d(t)$ is bounded in all time, and $\lim_{t\to\infty} d(t) = 0$.

Next, we discuss the case when $d(0) > 0$. To apply comparison principle (1b), we have to construct (ω, η) by (2.24) with $E = \Gamma$, $F = -C$ and with initial values $\omega(0) \leq d(0), \eta(0) \geq V(0)$.

Extend σ_+ continuously such that

$$\sigma'_+(x) = \frac{-\sigma_+(x)^2 - \Gamma\sigma_+(x) - Cx}{-Gx},$$

where the range of the extended part is in $\mathbb{R}^+$. It represents a path of the evolution of (ω, η). Take $\eta(0) = V(0)$ and $\omega(0) = f'_r(\eta(0))$. Clearly, $\omega(0) \leq d(0)$. Following the path, there exists a time t_0 such that $\omega(t_0) = 0$ and $\eta(t_0) = f_r(0)$. Comparison principle (1b) implies $d(t_0) \geq \omega(t_0)$ and $V(t_0) \leq \eta(t_0)$. Starting with time t_0, we proceed with the first step and it yields the same result.

2.4 Extension to 2D Cucker-Smale system

This section is devoted to extend the critical thresholds result in section 2.3 for macroscopic Cucker-Smale system to 2D.

Consider 2D pressure-less compressible Euler equations with nonlocal alignment of Cucker-Smale type

$$\rho_t + div(\rho \mathbf{u}) = 0, \qquad \mathbf{x} \in \mathbb{R}^2, t \geq 0, \tag{2.26a}$$

$$\mathbf{u}_t + \mathbf{u} \cdot \nabla \mathbf{u} = \int_{\mathbb{R}^2} \phi(|\mathbf{x} - \mathbf{y}|)(\mathbf{u}(\mathbf{y}) - \mathbf{u}(\mathbf{x}))\rho(\mathbf{y}) d\mathbf{y}, \tag{2.26b}$$

subject to compactly supported initial density ρ_0 and bounded initial velocity $\mathbf{u}_0$,

$$\rho(\mathbf{x}, 0) = \rho_0(\mathbf{x}) \in L^1_+(\mathbb{R}^n), \qquad \mathbf{u}(\mathbf{x}, 0) = \mathbf{u}_0(\mathbf{x}) \in W^{1,\infty}(\mathbb{R}^n), \tag{2.26c}$$

with influence function ϕ satisfying (1.8).

The goal is to understand the behavior of $\|\nabla \mathbf{u}(\cdot,t)\|_{L^\infty}$ in time. If it is bounded in all time, $(\rho, \mathbf{u})$ is a strong solution of the system (2.26), and flocking property follows.

We first state the main theorem. Recall the two initial quantities which are essential in 1D system

$$d_0 := \inf_{\mathbf{x}\in supp(\rho_0)} div\, \mathbf{u}_0(\mathbf{x}), \quad V_0 := \sup_{\mathbf{x},\mathbf{y}\in supp(\rho_0)} |\mathbf{u}_0(\mathbf{x}) - \mathbf{u}_0(\mathbf{y})|.$$

Here, $div\, \mathbf{u}_0$ is playing the role as u_{0x} in 1D. Introduce another initial quantity

$$B_0 := \sup_{x\in supp(\rho_0)} \max\left\{2|\partial_{x_1} u_{02}|, 2|\partial_{x_2} u_{01}|, |\partial_{x_1} u_{01} - \partial_{x_2} u_{02}|\right\}.$$

It characterizes information in $\nabla \mathbf{u}_0$ other than its trace. The critical thresholds are expressed in terms of (d_0, V_0, B_0).

Theorem 2.4.1 (2D Critical Thresholds for Cucker-Smale)**.** *Consider system* (2.26) *in two-dimension. Then,*

- *Subcritical: There exists threshold functions σ_+, ζ such that, if initially*

$$d_0 > \sigma_+(V_0) \quad and \quad B_0 < \zeta(V_0),$$

  then $\nabla \mathbf{u}(x,t)$ is bounded for all $(x,t) \in supp(\rho)$.

- *Supercritical: There exists a threshold function σ_- such that, if initially*

$$d_0 < \sigma_-(V_0), \quad and \quad |\partial_{x_1} u_{02}|, |\partial_{x_2} u_{01}| \text{ are big enough},$$

  then there exists a finite time T and a position $x \in supp(\rho(\cdot,T))$, $div\, \mathbf{u}(x,T) \to -\infty$.

Remark 2.4.1. B_0 reflects the initial spectral gap. The additional information encoded in theorem 2.4.1 says: if the spectral gap is initially small, it remains small for all time.

Thanks to this bound, we are able to reduce the 2D system into the 1D case, where the divergence direction dominates the spectral gap in the subcritical region. For subcritical region, the threshold functions σ_+ and ζ are illustrated in figure 2.4 and figure 2.3, respectively.

2.4.1 Boundedness of spectral gap implies critical thresholds

We start with applying $\nabla_{\mathbf{x}}$ on (2.26b). It yields the dynamics of the gradient velocity matrix $M = \nabla_{\mathbf{x}}\mathbf{u}$ along the characteristic curves

$$M' + M^2 = \int_{\mathbb{R}^2} \nabla_{\mathbf{x}}\phi(|\mathbf{x}-\mathbf{y}|) \otimes (\mathbf{u}(\mathbf{y},t) - \mathbf{u}(\mathbf{x},t))\rho(\mathbf{y},t)d\mathbf{y} - M\int_{\mathbb{R}^2} \phi(|\mathbf{x}-\mathbf{y}|)\rho(\mathbf{y},t)d\mathbf{y}. \tag{2.27}$$

Proposition 2.3.2 stays true for 2D (and higher dimensional) system, where the second inequality is satisfied for each entry of the matrix, namely

$$\left| \int_{\mathbb{R}^2} \partial_{x_j}\phi(|\mathbf{x}-\mathbf{y}|)(u_i(\mathbf{y}) - u_i(\mathbf{x}))\rho(\mathbf{y})d\mathbf{y} \right| \leq V_0\|\phi\|_{\dot{W}^{1,\infty}} m, \quad \forall\, i,j = 1,2.$$

Therefore, (2.27) is an example of the following prototype problem

$$M' = -M^2 - pM + Q, \quad \text{where } 0 < \gamma \leq p \leq \Gamma \text{ and } |Q_{ij}| \leq c, i,j = 1,2. \tag{2.28}$$

Here, p is bounded above and below by positive constants γ and Γ uniformly in time and paths. Q is an 2×2 matrix, with all entries bounded by $\pm c$ uniformly in time and paths. For (2.27), the constants are the same as 1D system, given in (2.18).

To study (2.28), we first take the trace of the system. It provides the dynamics of $d = \text{tr}M = div\,\mathbf{u}$ along the characteristics.

$$d' = -\frac{d^2+\eta^2}{2} - pd + (Q_{11}+Q_{22}).$$

Here, η is the spectral gap defined in (2.7). Recall $\eta^2 = (\lambda_1 - \lambda_2)^2$ where λ_1 and λ_2 are eigenvalues of M.

If η is bounded by a constant uniformly in time, i.e. $\eta(t) \leq \tilde{c}$ for all $t \geq 0$, then

$$d' = -\frac{d^2}{2} - pd + \tilde{Q},$$

where $p \in [\gamma, \Gamma]$ and $\tilde{Q} = Q_{11} + Q_{22} - \eta^2/2 \in [-2c - \tilde{c}/2, 2c + \tilde{c}/2]$. The system is reduced to one dimensional under prototype (2.28). Therefore, to bound the spectral gap η is the main idea when extending a one-dimension model into two dimensions.

We proceed with the same procedure as for Euler-Poisson equations. Let $q := M_{11} - M_{22}, r := M_{12} + M_{21}, s := M_{12} - M_{21} = \nabla \times M$. Here, we change the notation ω to s since ω has been used elsewhere in this section.

System (2.28) can be expressed in terms of (d, q, r, s).

$$d' + \frac{d^2 + \eta^2}{2} + pd = Q_{11} + Q_{22}, \tag{2.29a}$$

$$q' + q(d + p) = Q_{11} - Q_{22}, \tag{2.29b}$$

$$r' + r(d + p) = Q_{12} + Q_{21}, \tag{2.29c}$$

$$s' + s(d + p) = Q_{12} - Q_{21}, \tag{2.29d}$$

where the spectral gap $\eta^2 = q^2 + r^2 - s^2$. Hence, the boundedness of (q, r, s) implies the boundedness of η. Note that equations (2.29b)-(2.29d) have the same type. The following lemma states the uniform boundedness property of (q, r, s) (and hence η).

Lemma 2.4.2 (Uniform bound for the spectral gap)**.** *Suppose* (q, r, s) *are bounded initially by*

$$\max\{|q(0)|, |r(0)|, |s(0)|\} \leq B. \tag{2.30}$$

If $d(t) \geq -\gamma + 2cB^{-1}$ for $t \in [0,T]$, then the boundedness of (q,r,s) is preserved, i.e.

$$\max\{|q(t)|,|r(t)|,|s(t)|\} \leq B, \ \ \textit{for } t \in [0,T].$$

Moreover, the spectral gap $|\eta(t)| \leq \sqrt{2}B$ is also bounded for $t \in [0,T]$.

Proof. We prove the result for q by contradiction. Suppose there exists a (smallest) $t_0 \in [0,T]$ such that $|q(t)| > B$ for $t \in (t_0, t_0+\delta)$. By continuity, $|q(t_0)| = B$. There are two cases.

- $q(t_0) = B, q'(t_0) > 0$. Then $q'(t_0) + q(t_0)(d(t_0)+p) > 0 + B(2cB^{-1}) = 2c$. This contradicts with (2.29b) as $Q_{11} - Q_{22} \leq 2c$.

- $q(t_0) = -B, q'(t_0) < 0$. Then $q'(t_0) + q(t_0)(d(t_0)+p) < 0 - B(2cB^{-1}) = -2c$. This also contradicts with (2.29b) as $Q_{11} - Q_{22} \geq -2c$.

Therefore, $|q(t)| \leq B$ for $t \in [0,T]$. Same argument yields the boundedness of r and s.

Finally, $|\eta(t)| = \sqrt{|q(t)^2 + r(t)^2 - s(t)^2|} \leq \sqrt{2}B$, for $t \in [0,T]$. □

Lemma 2.4.2 says that the spectral gap is bounded as long as d is not too negative. Under this assumption, (2.29a) can be written as

$$d' = -\frac{d^2}{2} - pd + \tilde{Q},$$

where $p \in [\gamma, \Gamma]$ and $\tilde{Q} \in [-2c - B^2, 2c + B^2]$. The system is the same as (2.17) after a simple scaling. Proposition 2.3.3 implies the following result.

Proposition 2.4.3. *Assume* (2.30) *and $d(0) \geq -\gamma + \sqrt{\gamma^2 - 4c - 2B^2} \geq -\gamma + 2cB^{-1}$. Then M is bounded in all time.*

Proof. We claim that

$$d(t) \geq -\gamma + \sqrt{\gamma^2 - 4c - 2B^2} \geq -\gamma + 2cB^{-1} \quad \text{and} \quad \max\{|q(t)|, |r(t)|, |s(t)|\} \leq B.$$

Violation of the first condition contradicts proposition 2.3.3. Violation of the second condition contradicts lemma 2.4.2. □

Remark 2.4.2. We can rewrite the assumption in proposition 2.4.3 as follows:

$$d(0) \geq -\gamma + \sqrt{\gamma^2 - 4c - 2B^2}, \quad B \leq \frac{1}{2}\sqrt{(\gamma^2 - 4c) + \sqrt{(\gamma^2 - 4c)^2 - 32c^2}}.$$

Therefore, to ensure boundedness of M in all time, we need $d(0)$ not too negative, and $q(0), r(0), s(0)$ small.

Adding up all characteristic paths, we conclude with the following theorem.

Theorem 2.4.4 (2D critical thresholds)**.** *Consider system* (2.26) *in two-dimension.*

- *If the initial configuration satisfies*

$$V_0 \leq \frac{(\sqrt{2}-1)\phi(D)^2 m}{4\|\phi\|_{\dot{W}^{1,\infty}}}, \quad d_0 \geq -\frac{1}{2}\left(\phi(D)m + \sqrt{\phi(D)^2 m^2 - 4V_0\|\phi\|_{\dot{W}^{1,\infty}} m - 2B_0^2}\right),$$

$$B_0 \leq \frac{1}{2}\sqrt{\phi(D)^2 m^2 - 4V_0\|\phi\|_{\dot{W}^{1,\infty}} m + \sqrt{\left(\phi(D)^2 m^2 - 4V_0\|\phi\|_{\dot{W}^{1,\infty}} m\right)^2 - 32V_0^2[\phi]_{Lip}^2 m^2}},$$

then $\nabla_{\mathbf{x}}\mathbf{u}$ *is bounded for all* $(\mathbf{x}, t) \in supp(\rho)$.

- *If the initial configuration satisfies*

$$d_0 < -\frac{1}{2}\left(m + \sqrt{m^2 + 4V_0\|\phi\|_{\dot{W}^{1,\infty}} m}\right),$$

$$|\partial_{x_2} u_{01}|, |\partial_{x_1} u_{02}| \geq \frac{V_0\|\phi\|_{\dot{W}^{1,\infty}} m}{\sqrt{m^2 + 4V_0\|\phi\|_{\dot{W}^{1,\infty}} m}}, \text{ and } \partial_{x_2} u_{01} \cdot \partial_{x_1} u_{02} > 0, \text{ for every } \mathbf{x},$$

then there exists a finite time T *and a position* $x \in supp(\rho(\cdot, T))$, $div\,\mathbf{u}(x, T) \to -\infty$.

Remark 2.4.3. If $B_0 = 0$, the result reduces to the one dimensional case. In general, to bound B in all time, we need to assume more strict initial condition for $d(0)$.

Remark 2.4.4. For the second part of the theorem, we provide a critical threshold of the initial profile which leads to a finite time break down. The idea and the result are similar to the one-dimensional case. The extra assumption on $\partial_{x_2} u_{01}$ and $\partial_{x_1} u_{02}$ is to make sure the spectral gap is real for all time. So it does not help prevent d from blowup to $-\infty$ in finite time. As it is not our main concern, we omit the proof.

2.4.2 2D Enhanced dynamics with fast alignment

Theorem 2.4.4 can be improved by taking into account the fast alignment property: $V(t)$ decays exponentially. Similar to 1D system, we shall study the following prototype of problems

$$M' = -M^2 - pM + Q, \quad p \in [\gamma, \Gamma], \;\; Q_{ij} \in [-cV, cV], \;\; i, j = 1, 2, \tag{2.31a}$$

$$\tfrac{d}{dt} V \leq -GV, \tag{2.31b}$$

where γ, Γ, c, G are positive constants. For Cucker-Smale system (2.26), the coefficients are given in (2.20).

Rewrite system (2.29) coupled with fast decay property (2.31b).

$$d' + \frac{d^2+q^2}{2} + 2rs + pd = (\hat{Q}_{11} + \hat{Q}_{22})V, \tag{2.32a}$$

$$q' + q(d+p) = (\hat{Q}_{11} - \hat{Q}_{22})V, \tag{2.32b}$$

$$r' + r(d+p) = (\hat{Q}_{12} + \hat{Q}_{21})V, \tag{2.32c}$$

$$s' + s(d+p) = (\hat{Q}_{12} - \hat{Q}_{21})V, \tag{2.32d}$$

$$\frac{d}{dt}V = -GV, \tag{2.32e}$$

where $p \in [\gamma, \Gamma]$ and $\hat{Q} = Q/V$ with $|\hat{Q}_{ij}| \leq c$ for $i, j = 1, 2$.

Now, we state the uniform boundedness result for the spectral gap.

Lemma 2.4.5. *Let $b_0 = \max\{|q(0)|, |r(0)|, |s(0)|\}$. Suppose there exists a positive constant δ such that $d(t) \geq -\gamma + \delta$ for all $t \geq 0$.*

If $b_0 \leq \zeta(V_0; \delta, B)$, then (q, r, s) are uniformly bounded,

$$\max\{|q(0)|, |r(0)|, |s(0)|\} \leq B.$$

The function ζ is defined as below

$$\zeta(x; \delta, B) = \begin{cases} B & x \in [0, \frac{\delta B}{2C}] \\ \frac{B}{\delta - G}\left[-G\left(\frac{2C}{\delta B}x\right)^{\delta/G} + \frac{2C}{\delta B}x\right] & x \in [\frac{\delta B}{2C}, \left(\frac{\delta}{G}\right)^{\frac{G}{\delta-G}}\frac{\delta B}{2C}], \ \ \delta \neq G \cdot \\ \frac{2C}{\delta}\left(1 - \log\left(\frac{2C}{\delta B}x\right)\right)x & x \in [\frac{\delta B}{2C}, \frac{\delta B e}{2C}], \ \ \delta = G \end{cases} \tag{2.33}$$

Lemma 2.4.5 provides a region of the initial (b_0, V_0) such that the spectral gap is uniformly bounded in all time. From the definition of ζ, we observe that, to guarantee a uniform upper bound B, V_0 can not be too large. Given δ and B, the upper bound of V_0 is $\left(\frac{\delta}{G}\right)^{\frac{G}{\delta-G}}\frac{\delta B}{2C}$,independent of the choice of b_0.

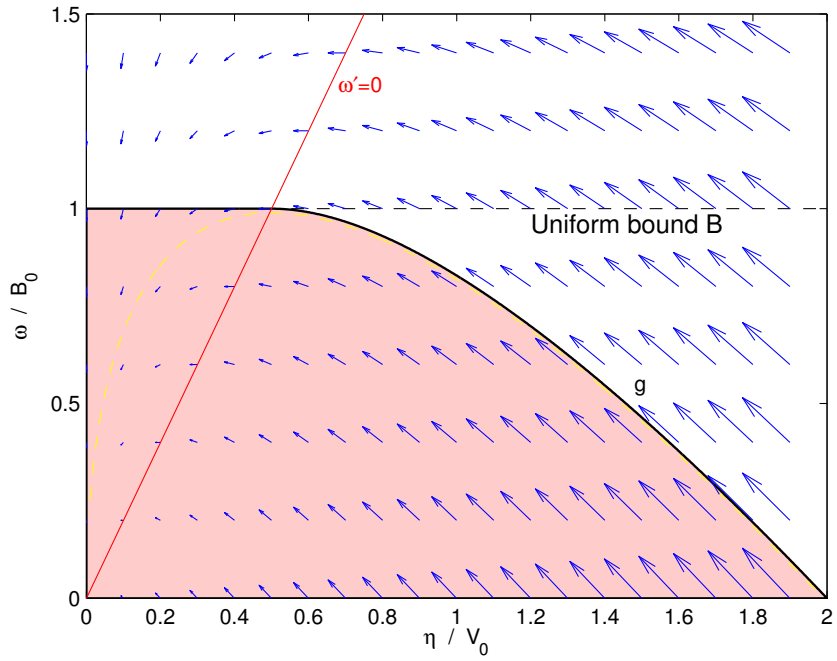

Figure 2.3: Subcritical threshold conditions for 2D system: $B_0 \leq \zeta(V_0)$. If (V_0, B_0) lies in the shaded area, B_0 will remain uniformly bounded. This is the phase diagram of the equality system (η, ω). It represents (V_0, B_0) through a comparison principle.

Proof. We prove the result for q. Consider the coupled system (2.32b) and (2.32e). The corresponding equality system reads

$$\omega' = -\delta\omega + 2C\eta,$$

$$\eta' = -G\eta.$$

This system can be easily solved. Figure 2.3 shows the dynamics of (η, ω). The filled area includes all initial conditions such that $\omega(t) \leq B$ for all $t \geq 0$. The area is governed by a function g. A simple computation yields an explicit expression of g, which is stated in (2.33).

A comparison argument enable us to connect the equality system with the inequality

system, which says

$$\text{If } \begin{cases} |q(0)| \le \omega(0) \\ V_0 \le \eta(0) \end{cases}, \quad \text{then } \begin{cases} |q(t)| \le \omega(t) \\ V(t) \le \eta(t) \end{cases}, \quad \text{for all } t \ge 0.$$

Therefore, $|q|$ is bounded by B uniformly in time as long as $(V_0, |q(0)|)$ lies inside the area, i.e. $|q(0)| \le g(V_0)$. Similarly, we prove for r and s which ends the proof. □

Next, for given δ and B, we consider the coupled system (2.32a) and (2.32e) and find the region of $(V_0, d(0))$ such that $d(t) \ge -\gamma + \delta$.

Proposition 2.4.6. *Suppose there exists a B such that $|\eta(t)| \le B$ for $t \ge 0$, and $B \le \gamma/\sqrt{2}$. Also, suppose $\delta \in (0, \sqrt{\gamma^2 - 2B^2}]$.*

If $d(0) \ge \sigma_+(V_0; \delta, B)$, then d is bounded all time, and $d(t) \ge -\gamma + \delta$ for $t \ge 0$.

The function σ_+ is continuous and defined implicitly as below

$$\sigma_+(x; \delta, B) = -\gamma + \delta, \qquad x \in [0, \frac{\gamma^2 - \delta^2 - 2B^2}{4C}) \tag{2.34a}$$

$$\sigma'_+(x; \delta, B) = \begin{cases} \frac{\sigma_+(x)^2 + 2\gamma\sigma_+(x) + 4Cx + 2B^2}{2Gx} & \text{if } \sigma_+(x) < 0 \\ \frac{\sigma_+(x)^2 + 2\Gamma\sigma_+(x) + 4Cx + 2B^2}{2Gx} & \text{if } \sigma_+(x) \ge 0 \end{cases}, \quad x \in [\frac{\gamma^2 - \delta^2 - 2B^2}{4C}, +\infty). \tag{2.34b}$$

Similar to the one-dimensional case, proposition 2.4.6 can be easily proved by analyze on the equality system and a comparison principle. Figure 2.4 shows the area of $(V_0, d(0))$ such that $d(t)$ is lower bounded by $-\gamma + \delta$ for all time. The area is governed by h defined in (2.34). We omit the detail of the proof.

Theorem 2.4.7 (2D Critical Thresholds)**.** *Consider initial value problem of* (2.26) *in two-dimension. Constants (γ, Γ, C, G) are defined in* (2.20)*. Then,*

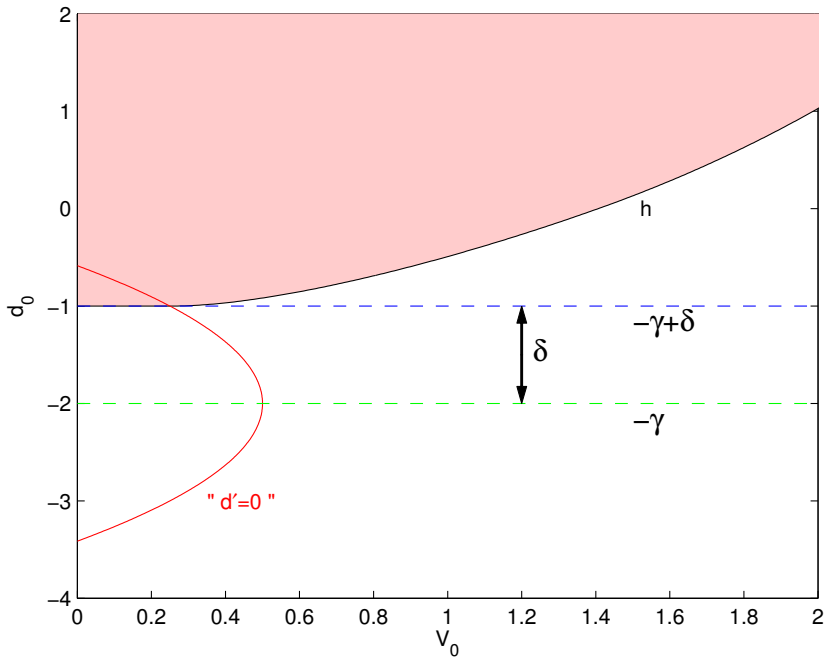

Figure 2.4: Subcritical threshold conditions for 2D system: $d_0 \geq \sigma_+(V_0)$. The shaded area represents the upper threshold. δ characterizes the effect of spectral gap, which makes the subcritical area smaller.

- *Subcritical: If there exists* (δ, B) *such that* $\delta^2 + 2B^2 \leq \gamma^2$*, and the initial profiles* (V_0, d_0, B_0) *satisfies*

  1. $B_0 \leq \zeta(V_0; \delta, B)$*, where* ζ *is defined in* (2.33)*,*

  2. $d_0 \geq \sigma_+(V_0; \delta, B)$*, where* σ_+ *is defined in* (2.34)*.*

  Then, $|\nabla \mathbf{u}(x,t)|$ *is bounded all time for all* $(x,t) \in supp(\rho)$*.*

Remark 2.4.5. The theorem says, two conditions need to be satisfied to guarantee the boundedness of $\nabla \mathbf{u}$:

First, B_0 should be small. If B_0 is too big, we can not control the spectral gap, then the two-dimensional system will collapse.

Second, d_0 can not be too negative. This condition is needed for both one and two dimensions, in order to prevent $div\,\mathbf{u}$ from blowing up to $-\infty$ in finite time.

2.5 Extension to vacuum area

In this section, we discuss the boundedness of $\nabla\mathbf{u}$ when $(\mathbf{x},t) \notin supp(\rho)$.

As discussed in section 2.1.4, treating with vacuum is difficult in general. For the local system (2.2), there is no external forcing inside vacuum as the viscosity is degenerate. The system acts like compressible Euler equations, and the solution will form a shock in finite time. There is no hope to bound $\nabla\mathbf{u}$ global in time.

However, the alignment forcing is non-local, which helps smoothing the equation even in vacuum area. It can compete with convection if non-locality is strong enough.

The study of critical thresholds in vacuum area enables us to study the system in the whole space, without worrying about the free boundary. It also extend the global existence result to initial density whose support is not connected.

For simplicity, we focus on 1D systems. Similar result can be established for 2D systems, with no additional difficulty.

The next theorem shows the upper threshold to ensure boundedness of u_x outside the support of ρ.

Theorem 2.5.1 (1D Upper Threshold for vacuum area)**.** *Consider initial value problem of* (2.13)*. Let V_0^λ denote the variation of the initial velocity between a point in the non-vacuum area and a point at most λ away from the non-vacuum area,*

$$V_0^\lambda = \sup\left\{|u_0(x)-u_0(y)| \;:\; dist(x,supp(\rho_0)) \le \lambda,\;\; y \in supp(\rho_0),\right\}.$$

Subcritical: If the initial configuration satisfies

$$V_0^\lambda \leq \frac{m\phi^2(\lambda+D)}{4|\phi_x(\lambda)|+2|\phi_x(\lambda+D)|} \quad \textit{for all } \lambda \geq 0, \tag{2.35a}$$

$$u_{0x}(x) \geq -\frac{m}{2}\phi(dist(x, supp(\rho_0))+D). \tag{2.35b}$$

Then $u_x(x,t)$ is bounded for all $(x,t) \notin supp(\rho)$.

Remark 2.5.1. Condition (2.35) has the same flavor as (2.15) for the non-vacuum area: variation of initial velocity is not too big and slop of velocity is not too negative. For (2.35a), when λ approaches zero, the condition is equivalent to the non-vacuum case. On the other hand, when λ approaches infinity, if $\phi(r) \sim r^{-\alpha}$, the right hand side is proportional to $r^{1-\alpha}$. Thanks to the slow decay assumption on ϕ, *i.e.* $\alpha < 1$, (2.35a) provides no restrictions on V_0^∞. Note that if $\alpha > 1$, the condition requires $V_0^\infty = 0$ which can not be achieved unless u is a constant.

If we combine theorem 2.3.1 for non-vacuum area and theorem 2.5.1 for vacuum area, we conclude that the 1D system (2.13) has global strong solutions provided suitable subcritical initial conditions.

Theorem 2.5.2 (1D global strong solution)**.** *Consider initial value problem of* (2.13)*.*

- *Subcritical: If initial configuration satisfies both* (2.15) *and* (2.35)*, then there exists a strong solution $\rho \in L^\infty([0,+\infty), L^1(\mathbb{R}))$ and $u \in L^\infty([0,+\infty), W^{1,\infty}(\mathbb{R}))$. Moreover, the solution converges to a flock in the sense of definition* 1.3.2*.*

- *Supercritical: If initial configuration satisfies $d_0 < \sigma_-(V_0)$, then solution (ρ,u) will blow up in finite time.*

We prove theorem 2.5.1 in the rest of the section.

2.5.1 Dynamics inside the vacuum

Consider the dynamics of $d = u_x$ (2.28) for $x \notin supp(\rho_0)$. Define maximum variation of the velocity field between a point in the whole space and a point in the non-vacuum area

$$V^{\infty}(t) := \sup\{|u(x,t) - u(y,t)|,\ x \in \mathbb{R},\ y \in supp(\rho(\cdot,t))\}.$$

and the distance between x and the non-vacuum area at time t

$$L(x,t) := dist(x, supp(\rho(\cdot,t)).$$

We have the following bounds in contrast with proposition 2.3.2.

Proposition 2.5.3 (Bounds inside the vacuum)**.** *Suppose* (ρ,u) *is a strong solution of system* (2.13)*. Then, for any* $x \notin supp(\rho(t))$*,*

$$\left|\int_{-\infty}^{\infty} \phi_x(|x-y|)(u(y,t)-u(x,t))\rho(y,t)dy\right| \leq V^{\infty}(0)\,|\phi_x(L(x,t))|\,m,$$
$$\phi(L(x,t)+D)m \leq \int_{-\infty}^{\infty} \phi(|x-y|)\rho(y,t)dy \leq m.$$

Proof. For the first inequality,

$$\left|\int_{-\infty}^{\infty} \phi_x(|x-y|)(u(y,t)-u(x,t))\rho(y,t)dy\right| \leq \int_{supp(\rho(t))} |u(y,t)-u(x,t)|\,|\phi_x(|x-y|)|\,\rho(y,t)d\mathbf{y}$$
$$\leq V^{\infty}(t)|\phi_x(L(x,t))| \int_{supp(\rho(t))} \rho(y,t)dy \leq V^{\infty}(0)|\phi_x(L(x,t))|m.$$

The last inequality is valid due to maximum principle.

For the second inequality,

$$\int_{-\infty}^{\infty} \phi(|x-y|)\rho(y,t)dy \leq \|\phi\|_{L^{\infty}} m = m.$$

On the other hand, as (ρ,u) converges to a flock, $S(t)$ is uniformly bounded by D, defined in (1.7). Hence, $\max_{y\in supp(\rho(t))} dist(x,y) \le L(x,t)+D$. It yields

$$\int_{-\infty}^{\infty} \phi(|x-y|)\rho(y,t)dy = \int_{supp(\rho(t))} \phi(|x-y|)\rho(y,t)dy \ge \phi(L(x,t)+D)m.$$

□

Remark 2.5.2. The key estimate in proposition 2.5.3 is the positive lower bound of $\phi\star\rho$. This lower bound helps preventing shocks. For local system, we are not able to find such lower bound for $x\notin supp(\rho(\cdot,t))$, where regularity can not be preserved in general.

Now, we discuss the criterion to guarantee the boundedness of $\|u_x(\cdot,t)\|_{L^\infty}$ in whole space, using similar technique as Section 2.3.

Lemma 2.5.4. *Suppose V^∞ satisfies the condition*

$$V^\infty(0) \le \inf_{r\ge 0}\left[\frac{m\phi(r+D)^2}{4|\phi_x(r)|+2|\phi_x(r+D)|}\right], \tag{2.36}$$

and u_0 satisfies the following criterion

$$u_{0x}(x) \ge -\frac{m}{2}\phi(L(x,0)+D), \quad for\ x\notin supp(\rho_0). \tag{2.37}$$

Then, $u_x(x,t)$ is bounded for all time for $(x,t)\notin supp(\rho)$.

There are several remarks about lemma 2.5.4.

Remark 2.5.3. Condition (2.36) carries two aspects.

1. Slow decay when $r\to\infty$. Suppose $\phi(r)\approx r^{-\alpha}$ as $r\to\infty$. The right hand side is proportional to $r^{1-\alpha}$. If ϕ decays fast with $\alpha>1$, *i.e.* (1.8) is violated, the right hand side goes to 0. The condition can not be achieved unless u_0 is a constant. A slow decay assumption on ϕ is needed to make sure the condition is meaningful.

2. Condition for V^∞ when $r \to 0$. Take $r = 0$, the condition reads

$$V^\infty(0) \le \frac{m\phi(D)^2}{4\|\phi\|_{\dot{W}^{1,\infty}} + 2|\phi_x(D)|}.$$

This is equivalent to the thresholds of V_0 in proposition 2.3.3, assuming $V^\infty(0) \lesssim V_0$.

Remark 2.5.4. Criterion (2.37) is satisfied automatically for large $|x|$. As the matter of fact, when $|x| \to \infty$, (2.37) says that $u_{0x}(x) \gtrsim -|x|^{-\alpha}$. This is a consequence of $u_0 \in L^\infty(\mathbb{R})$ and the fact $\alpha \le 1$.

Proof of Lemma 2.5.4. Consider $(x,t) \notin supp(\rho)$. It belongs to a characteristic starting from $(x_0, 0)$ where $x_0 \notin supp(\rho_0)$, as long as u_x is bounded. At this point, we have

$$d' = -d^2 - pd + Q,$$

where $p \in [\phi(L(x,t)+D)m, m]$ and $|Q| \le V^\infty(0)|\phi_x(L(x,t))|m$.

It is sufficient to discuss the following equality system and use the comparison principle to draw conclusion on d.

$$\omega' = -\omega^2 - \phi(L(x,t)+D)m\omega - V^\infty(0)|\phi_x(L(x,t))|m.$$

Condition (2.36) ensures that the right hand side has two distinguished solutions. Especially, if we pick $\omega = -\frac{1}{2}\phi(L(x,t)+D)m$, then

$$\omega' = \frac{1}{4}\phi(L(x,t)+D)^2m^2 - V^\infty(0)|\phi_x(L(x,t))|m \overset{(2.36)}{\ge} \frac{1}{2}|\phi_x(L(x,t)+D)|V^\infty(0)m > 0.$$

Let $A(x_0)$ denote the area where $\omega \ge -\frac{1}{2}\phi(L(x,t)+D)m$, and $(x,t) = (X(t),t)$ is a point on the characteristic starting from $(x_0, 0)$, namely

$$A(x_0) := \left\{ (z,t) \,\middle|\, z \ge -\frac{1}{2}(\phi(L(X(t),t)+D)m), t \ge 0 \right\}.$$

Its boundary $\partial A(x_0)$ reads

$$\partial A(x_0) = \{(\gamma(t),t)|t \geq 0\} \quad \text{where} \quad \gamma(t) = -\frac{1}{2}\phi(L(X(t),t)+D)m.$$

Criterion (2.37) implies $(\omega(0),0) \in A(x_0)$. We are left to show that $(\omega(t),t)$ stays in $A(x_0)$ for all $t \geq 0$. As $A(x_0)$ is uniformly bounded from below in z, it implies ω is lower bounded in all time.

Finally, we prove that $(\omega(t),t) \in A(x_0)$ for $t \geq 0$ by contradiction.

Suppose there exist $t > 0$ such that $(\omega(t),t) \in \partial A(x_0)$, and $(\omega(t+\delta),t+\delta) \notin A(x_0)$. It means that $\omega(t) = \gamma(t)$ and $\omega'(t) < \gamma'(t)$. On the other hand, we compute

$$\gamma'(t) = \frac{m}{2}\phi_x(L(X(t),t)+D)\frac{d}{dt}L(X(t),t) \leq \frac{1}{2}|\phi_x(L(x,t)+D)|V^{\infty}(0)m.$$

The last inequality is true as both $\partial(supp(\rho))$ and X are travelling with the speed between u_{min} and u_{max}. It yields $\frac{d}{dt}L(X(t),t) \leq V^{\infty}(t) \leq V^{\infty}(0)$.

Combine with the estimate on ω', we conclude that $\omega'(t) \geq \gamma'(t)$ which leads to a contradiction. □

2.5.2 Fast alignment property inside the vacuum

We showed in Section 2.3.2 a much wider critical threshold for $(x,t) \in supp(\rho)$, assuming fast alignment property. In this subsection, we extend the enhanced result to the vacuum area.

We start with showing a fast alignment property where vacuum is involved. As the strength of the alignment forcing at point (x,t) is determined by $L(x,t)$, it is natural to introduce the following definitions.

Definition 2.5.1 (Level Sets)**.** For any level $\lambda \geq 0$, define

$$\Omega^{\lambda}(t) = \left\{ X(t) \;\middle|\; \begin{cases} \dot{X}(t) = u(X,t) \\ \\ X(0) = x \end{cases} , \; L(x,0) \leq \lambda \right\},$$

$$S^{\lambda}(t) = \sup\left\{ |x-y|, \;\; x \in \Omega^{\lambda}(t), \; y \in \Omega^{0}(t) \right\}.$$

$$V^{\lambda}(t) = \sup\left\{ |u(x,t)-u(y,t)|, \;\; x \in \Omega^{\lambda}(t), \; y \in \Omega^{0}(t) \right\}.$$

If $\lambda = 0$, $\Omega^0(t) = supp(\rho(t))$, and $S^0(t), V^0(t)$ coincides with $S(t), V(t)$ respectively. Moreover, $S^{\lambda}(0) = S_0 + \lambda$. If $\lambda = \infty$, $V^{\infty}(t)$ coincide with the definition before.

Theorem 2.5.5 (Fast alignment on Ω^{λ})**.** *Let (ρ, u) be a global strong solution of system* (2.13)*. Suppose the influence function ϕ satisfies*

$$m \int_{S_0}^{\infty} \phi(r) dr > V^{\lambda}(0).$$

Then, there exists a finite number D^{λ}, such that

$$\sup_{t \geq 0} S^{\lambda}(t) \leq D^{\lambda}, \quad V^{\lambda}(t) \leq V^{\lambda}(0) e^{-m\phi(D^{\lambda})t}.$$

Moreover, D^{λ} has the following expression

$$D^{\lambda} = \psi^{-1}(V^{\lambda}(0) + \psi(S_0 + \lambda)), \text{ where } \psi(t) = m \int_0^t \phi(s) ds.$$

Remark 2.5.5. The proof of theorem 2.5.5 follows the same idea in proposition 1.3.7 and theorem 1.3.6 by considering X is a characteristic starting from $x \in \Omega^{\lambda}(0)$. We observe that $V^{\lambda}(t)$ still has an exponential decay in time, with rate $m\phi(D^{\lambda})$. When λ becomes larger, the rate becomes smaller. However, as long as λ is finite, we always have fast alignment.

Now, we are ready to prove theorem 2.5.1. It is an improvement of lemma 2.5.4 by using fast alignment property.

Proof of theorem 2.5.1. We repeat the proof of lemma 2.5.4 using a better bound on the term Q which reads $|Q| \leq V^{L(x_0,0)}(t)|\phi_x(L(x,t))|m$. Also, we use a better bound on $\frac{d}{dt}L(X(t),t) \leq V^{L(x_0,0)}(t)$. It yields the following modified condition

$$V^{L(x_0,0)}(t) \leq \frac{m\phi(L(x,t)+D)^2}{4|\phi_x(L(x,t))|+2\phi_x(L(x,t)+D)|},$$

for all x_0 and t, with $(x,t) = (X(t),t)$ being a point on the characteristics starting from $(x_0,0)$.

When $t = 0$, let $\lambda = L(x_0,0)$, we get the condition (2.35a) stated in the theorem, i.e.

$$V^{\lambda}(0) \leq \frac{m\phi(\lambda+D)^2}{4|\phi_x(\lambda)|+2\phi_x(\lambda+D)|}.$$

Finally, we prove that if (2.35a) holds, the modified condition automatically holds for all $t > 0$.

Take $\lambda = L(x,t)$, It suffies to prove that $V^{L(x_0,0)}(t) \leq V^{\lambda}(0)$. Apply theorem 2.5.5, we are left to prove $V^{L(x_0,0)}(0)e^{-m\phi(D^{L(x_0,0)})t} \leq V^{L(x,t)}(0)$. This is true if we assume that V^{λ} grows slower than exponential in λ. The assumption is valid as V^{∞} is finite. □

2.6 Higher integrability and regularity

In this section, we discuss the smoothness of the strong solution of macroscopic Cucker-Smale system, assuming the initial configuration lies in the subcritical region.

Recall the system:

$$\rho_t + div(\rho \mathbf{u}) = 0, \tag{2.38a}$$

$$\mathbf{u}_t + \mathbf{u}\cdot\nabla\mathbf{u} = \int_{\mathbb{R}^n} \phi(|\mathbf{x}-\mathbf{y}|)(\mathbf{u}(\mathbf{y})-\mathbf{u}(\mathbf{x}))\rho(\mathbf{y})d\mathbf{y}. \tag{2.38b}$$

The main question is: *does the strong solution preserves initial integrability and regularity?*

We first show that $\|\rho(\cdot,t)\|_{L^p}$ is bounded in all finite time if ρ_0 is L^p integrable.

Lemma 2.6.1 (L^p boundedness on ρ)**.** *Suppose* $(\rho,\mathbf{u})$ *are strong solution of* (2.38)*. Suppose* $\rho_0 \in L^p_+(\mathbb{R}^n)$ *compactly supported. Moreover,* $div\,\mathbf{u} \in L^1(0,T,L^\infty(\mathbb{R}^n))$*. Then,* $\rho(\cdot,t) \in L^p_+(\mathbb{R}^n)$ *for all* $t \in [0,T]$*. Here,* $1 \le p \le \infty$*.*

Proof. Multiply (2.38a) by $p\rho^{p-1}$ and integrate in space. We get

$$\begin{aligned}\frac{d}{dt}\|\rho(\cdot,t)\|^p_{L^p} &= -p\int_{\mathbb{R}^n}\rho^{p-1}div(\rho\mathbf{u})d\mathbf{x} = p(p-1)\int_{\mathbb{R}^n}\rho^{p-1}\nabla\rho\cdot\mathbf{u}d\mathbf{x}\\ &= (p-1)\int_{\mathbb{R}^n}\nabla(\rho^p)\cdot\mathbf{u}d\mathbf{x} \le (p-1)\|div\,\mathbf{u}\|_{L^\infty}\|\rho\|^p_{L^p}.\\ \Rightarrow \frac{d}{dt}\|\rho(\cdot,t)\|_{L^p} &\le \frac{p-1}{p}\|div\,\mathbf{u}\|_{L^\infty}\|\rho\|_{L^p}.\end{aligned}$$

Using Gronwall's inequality, we conclude

$$\|\rho(\cdot,T)\|_{L^p} \le \|\rho_0\|_{L^p}\exp\left(\frac{p-1}{p}\int_0^T\|div\,\mathbf{u}(\cdot,t)\|_{L^\infty}dt\right).$$

The argument is also true for $p=\infty$ with

$$\|\rho(\cdot,T)\|_{L^\infty} \le \|\rho_0\|_{L^\infty}\exp\left(\int_0^T\|div\,\mathbf{u}(\cdot,t)\|_{L^\infty}dt\right).$$

□

As we know $div\,\mathbf{u}$ is bounded in all time with subcritical initial data, we conclude that L^p integrability on ρ is preserved.

For velocity $\mathbf{u}$, maximum principle clearly holds.

Lemma 2.6.2 (Maximum principle)**.** *Suppose* $(\rho,\mathbf{u})$ *are smooth solution of* (2.38)*. Suppose* $\rho_0 \geq 0$*, then* $\rho \geq 0$*. Also, if* $\mathbf{u}_0 \in L^\infty(\mathbb{R}^n)$*, then*

$$\min_{\mathbf{x}\in\mathbb{R}^n} u_{0i}(\mathbf{x}) \leq u_i(\cdot,t) \leq \max_{\mathbf{x}\in\mathbb{R}^n} u_{0i}(\mathbf{x},0), \;\; for \;\; i=1,\cdots,n.$$

Next, we focus on initial data with higher regularity. The following theorem shows that the initial regularity is preserved if the nonlocal term satisfies the following condition:

$$\left\| \int_{\mathbb{R}^n} \phi(|\mathbf{x}-\mathbf{y}|)(\mathbf{u}(\mathbf{y})-\mathbf{u}(\mathbf{x}))\rho(\mathbf{y})d\mathbf{y} \right\|_{H^{s+1}} \lesssim \|\rho\|_{H^s}\|\nabla\mathbf{u}\|_{L^\infty} + \|\rho\|_{L^\infty}\|\mathbf{u}\|_{H^{s+1}}. \quad (2.39)$$

Theorem 2.6.3 (Global regularity)**.** *Consider system* (2.38) *subject to smooth initial conditions* $\rho_0 \in H^s_+(\mathbb{R}^n)$ *and* $\mathbf{u}_0 \in H^{s+1}(\mathbb{R}^n)$*, where* $s > n/2$*. Suppose* (2.39) *stays true.*

Then for any $T > 0$*, there exists a unique solution* $(\rho,\mathbf{u}) \in C([0,T];H^s_+(\mathbb{R}^n)) \times C([0,T];H^{s+1}(\mathbb{R}^n))$*, if and only if,* $\|\nabla\mathbf{u}(\cdot,t)\|_{L^\infty(\mathbb{R}^n)}$ *is bounded for all* $t \in [0,T]$*.*

Proof. We start with acting operator Λ^s on equation (2.38a) and integrate by parts against $\Lambda^s\rho$. Here $\Lambda := (I-\Delta)^{1/2}$ is the pseudo-differential operator.

The evolution of the H^s norm reads [60]

$$\frac{d}{dt}\|\rho(\cdot,t)\|^2_{H^s} = -\left([\Lambda^s div,\mathbf{u}]\rho,\Lambda^s\rho\right) + \frac{1}{2}\left(\Lambda^s\rho,(div\,\mathbf{u})\Lambda^s\rho\right).$$

By commutator estimates [54], we have

$$\|\,[\Lambda^s div,\mathbf{u}]\,\rho\|_{L^2} \lesssim \|\nabla\mathbf{u}\|_{L^\infty}\|\rho\|_{H^s} + \|\mathbf{u}\|_{H^{s+1}}\|\rho\|_{L^\infty}.$$

Use lemma 2.6.1, we get the following estimate

$$\frac{d}{dt}\|\rho(\cdot,t)\|_{H^s}^2 \lesssim \left[\|\nabla \mathbf{u}\|_{L^\infty} + \exp\left(\int_0^t \|div\,\mathbf{u}(\cdot,s)\|_{L^\infty} ds\right)\right]\left(\|\rho\|_{H^s}^2 + \|\mathbf{u}\|_{H^{s+1}}^2\right). \quad (2.40)$$

Similarly, for equation (2.38b), we have

$$\begin{aligned}\frac{d}{dt}\|\mathbf{u}(\cdot,t)\|_{H^{s+1}}^2 = &-\sum_j \left(\left[\Lambda^{s+1}, u_j\right]\partial_j \mathbf{u}, \Lambda^{s+1}\mathbf{u}\right) + \frac{1}{2}\left(\Lambda^{s+1}\mathbf{u}, (div\,\mathbf{u})\Lambda^{s+1}\mathbf{u}\right) \\ &+ \left(\Lambda^{s+1}\mathbf{u}, \Lambda^{s+1}\int_{\mathbb{R}^n} \phi(|\mathbf{x}-\mathbf{y}|)(\mathbf{u}(\mathbf{y})-\mathbf{u}(\mathbf{x}))\rho(\mathbf{y})d\mathbf{y}\right).\end{aligned}$$

For the commutator, we have the same estimates

$$\|\left[\Lambda^{s+1}, u_j\right]\partial_j \mathbf{u}\|_{L^2} \lesssim \|\nabla \mathbf{u}\|_{L^\infty}\|\mathbf{u}\|_{H^{s+1}}.$$

For the nonlocal term, condition (2.39) implies

$$\left(\Lambda^{s+1}\mathbf{u}, \Lambda^{s+1}\int_{\mathbb{R}^n} \phi(|\mathbf{x}-\mathbf{y}|)(\mathbf{u}(\mathbf{y})-\mathbf{u}(\mathbf{x}))\rho(\mathbf{y})d\mathbf{y}\right) \lesssim \|\mathbf{u}\|_{H^{s+1}}\left(\|\rho\|_{H^s}\|\nabla \mathbf{u}\|_{L^\infty} + \|\rho\|_{L^\infty}\|\mathbf{u}\|_{H^{s+1}}\right).$$

Using lemma 2.6.1 and 2.6.2, it yields the following estimate

$$\frac{d}{dt}\|\mathbf{u}(\cdot,t)\|_{H^{s+1}}^2 \lesssim \left[1 + \|\nabla \mathbf{u}\|_{L^\infty} + \exp\left(\int_0^t \|div\,\mathbf{u}(\cdot,s)\|_{L^\infty} ds\right)\right]\left(\|\rho\|_{H^s}^2 + \|\mathbf{u}\|_{H^{s+1}}^2\right). \quad (2.41)$$

Let $Y(t) := \|\mathbf{u}(\cdot,t)\|_{H^{s+1}}^2 + \|\rho(\cdot,t)\|_{H^s}^2$. Sum up (2.40) and (2.41) and use Gronwall's inequality. We get

$$Y(T) \lesssim Y(0)\exp\left\{\int_0^T \left[1 + \|\nabla \mathbf{u}(\cdot,t)\|_{L^\infty} + \exp\left(\int_0^t \|div\,\mathbf{u}(\cdot,s)\|_{L^\infty} ds\right)\right] dt\right\}. \quad (2.42)$$

Therefore, if $\|\nabla \mathbf{u}(\cdot,t)\|_{L^\infty}$ is bounded for all $t \in [0,T]$, then $Y(t)$ is bounded as well in $t \in [0,T]$.

The other direction is obviously true, as $H^s(\mathbb{R}^n) \subset L^\infty(\mathbb{R}^n)$ for $s > n/2$. □

To understand when (2.39) is satisfied, we can do the following formal Fourier transform

$$\begin{aligned}
&\left\| \int_{\mathbb{R}^n} \phi(|\mathbf{x}-\mathbf{y}|)(\mathbf{u}(\mathbf{y})-\mathbf{u}(\mathbf{x}))\rho(\mathbf{y})d\mathbf{y} \right\|_{H^{s+1}} \\
&= \left\| (1+|\xi|^2)^{(s+1)/2} \int_{\mathbb{R}^n} (\hat{\phi}(\xi)-\hat{\phi}(\eta))\hat{\rho}(\eta)\hat{\mathbf{u}}(\xi-\eta)d\eta \right\|_{L^2} \\
&\leq \left\| (1+|\xi|^2)^{(s+1)/2} |D_\xi \hat{\phi}(\xi)| \int_{\mathbb{R}^n} \hat{\rho}(\eta)|\xi-\eta||\hat{\mathbf{u}}(\xi-\eta)|d\eta \right\|_{L^2} + l.o.t \\
&\lesssim \left\| (1+|\xi|^2)^{1/2} |D_\xi \hat{\phi}(\xi)| \right\|_{L^\infty} \|\rho\nabla\mathbf{u}\|_{H^s}.
\end{aligned}$$

Here, $\|\rho\nabla\mathbf{u}\|_{H^s} \lesssim \|\rho\|_{H^s}\|\nabla\mathbf{u}\|_{L^\infty} + \|\rho\|_{L^\infty}\|\mathbf{u}\|_{H^{s+1}}$. Therefore, we need ϕ to satisfy the following property:

$$(1+|\xi|^2)^{1/2}|D_\xi\hat{\phi}(\xi)| \in L^\infty, \quad \text{i.e.} \quad \Lambda(\mathbf{x}\phi(|\mathbf{x}|)) \in L^1.$$

It requires that ϕ decays sufficiently fast at infinity. In fact, take for instance the prototype $\phi(\mathbf{x}) = (1+\mathbf{x})^{-\alpha}$, we need $\alpha > n$ to satisfy the property. This violates the slow decay assumption $\alpha < 1$, which ensures flocking.

Therefore, the preservation of higher initial regularity in $\mathbb{R}^n$ is not guaranteed for the main system (2.38).

2.7 On macroscopic Motsch-Tadmor system

In this section, we briefly discuss the critical thresholds phenomenon for macroscopic Motsch-Tadmor system

$$\rho_t + div(\rho\mathbf{u}) = 0, \tag{2.43a}$$

$$\mathbf{u}_t + \mathbf{u}\cdot\nabla\mathbf{u} = \int_{\mathbb{R}^n} \frac{\phi(|\mathbf{x}-\mathbf{y}|)}{\Phi(\mathbf{x})}(\mathbf{u}(\mathbf{y})-\mathbf{u}(\mathbf{x}))\rho(\mathbf{y})d\mathbf{y}. \tag{2.43b}$$

The procedure to bound $M = \nabla \mathbf{u}$ is the same as Cucker-Smale system.

The evolution of M along the characteristic curve is derived by take derivative on (2.43b) with respect to $\mathbf{x}$:

$$M' + M^2 = \int_{\mathbb{R}^n} \nabla_{\mathbf{x}} \left(\frac{\phi(|\mathbf{x}-\mathbf{y}|)}{\Phi(\mathbf{x},t)} \right) \otimes (\mathbf{u}(\mathbf{y},t) - \mathbf{u}(\mathbf{x},t)) \rho(\mathbf{y},t) d\mathbf{y} - M. \tag{2.44}$$

The non-local term has a uniform bounded, given by the following proposition.

Proposition 2.7.1. *Suppose* $(\rho, \mathbf{u})$ *is a strong solution of system* (2.43). *Then, for any* $\mathbf{x} \in supp(\rho(t))$,

$$\left| \int_{\mathbb{R}^n} \partial_{x_j} \left(\frac{\phi(|\mathbf{x}-\mathbf{y}|)}{\Phi(\mathbf{x},t)} \right) (u_i(\mathbf{y},t) - u_i(\mathbf{x},t)) \rho(\mathbf{y},t) d\mathbf{y} \right| \leq \frac{2\|\phi\|_{\dot{W}^{1,\infty}}}{\phi(D)} V(t), \quad i,j = 1, \cdots, n.$$

Proof.

$$\begin{aligned}
&\left| \int_{\mathbb{R}^n} \partial_{x_j} \left(\frac{\phi(|\mathbf{x}-\mathbf{y}|)}{\Phi(\mathbf{x},t)} \right) (u_i(\mathbf{y},t) - u_i(\mathbf{x},t)) \rho(\mathbf{y},t) d\mathbf{y} \right| \\
&\leq \int_{\mathbb{R}^n} |u_i(\mathbf{y},t) - u_i(\mathbf{x},t)| \left| \partial_{x_j} \left(\frac{\phi(|\mathbf{x}-\mathbf{y}|)}{\Phi(\mathbf{x},t)} \right) \rho(\mathbf{y},t) \right| d\mathbf{y} \\
&\leq V(t) \int_{\mathbb{R}^n} \frac{|\partial_{x_j} \phi(|\mathbf{x}-\mathbf{y}|) \rho(\mathbf{y},t) \Phi(\mathbf{x},t) - \phi(|\mathbf{x}-\mathbf{y}|) \rho(\mathbf{y},t) \partial_{x_j} \Phi(\mathbf{x},t)|}{\Phi^2(\mathbf{x},t)} d\mathbf{y} \\
&\leq V(t) \left[\frac{\int_{\mathbb{R}^n} |\partial_{x_j} \phi(|\mathbf{x}-\mathbf{y}|)| \rho(\mathbf{y}) d\mathbf{y} + |\partial_{x_j} \Phi(\mathbf{x},t)|}{\Phi(\mathbf{x},t)} \right] \\
&\leq \frac{2V(t) \int_{\mathbb{R}^n} |\partial_{x_j} \phi(|\mathbf{x}-\mathbf{y}|)| \rho(\mathbf{y}) d\mathbf{y}}{\Phi(\mathbf{x},t)}.
\end{aligned}$$

From theorem 1.3.6 we know that $S(t) \leq D$. Therefore,

$$\begin{aligned}
\Phi(\mathbf{x},t) &= \int_{\mathbb{R}^n} \phi(|\mathbf{x}-\mathbf{y}|) \rho(\mathbf{y},t) d\mathbf{y} = \int_{supp(\rho(t))} \phi(|\mathbf{x}-\mathbf{y}|) \rho(\mathbf{y},t) d\mathbf{y} \\
&\geq \phi(D) \int_{\mathbb{R}^n} \rho(\mathbf{y},t) d\mathbf{y} = \phi(D) m,
\end{aligned}$$

for all $\mathbf{x} \in supp(\rho(t))$. Hence, we conclude

$$\left| \int_{\mathbb{R}^n} \partial_{x_j} \left(\frac{\phi(|\mathbf{x}-\mathbf{y}|)}{\Phi(\mathbf{x},t)} \right) (u_i(\mathbf{y},t) - u_i(\mathbf{x},t)) \rho(\mathbf{y},t) d\mathbf{y} \right| \leq \frac{2\|\phi\|_{\dot{W}^{1,\infty}} m}{\phi(D) m} V(t).$$

□

The proposition shows that (2.44) is a special case of the prototype (2.19) for 1D and (2.31) for 2D, with coefficients

$$\gamma = 1,\ \Gamma = 1,\ c = \frac{2\|\phi\|_{\dot{W}^{1,\infty}}}{\phi(D)},\ G = \phi(D).$$

It yields the same result as Cucker-Smale. For instance, in 1D , theorem 2.3.1 holds for Motsch-Tadmor system with the following thresholds functions:

$$\sigma_+(0) = -1,\ \sigma_+'(x) = \begin{cases} \frac{2\|\phi\|_{\dot{W}^{1,\infty}}}{(1+\phi(D))\phi(D)}, & x \to 0+ \\ \frac{-\phi(D)\sigma_+(x)^2 - \phi(D)\sigma_+(x) - 2\|\phi\|_{\dot{W}^{1,\infty}} x}{-\phi(D)^2 x} & x > 0 \end{cases}, \tag{2.45a}$$

$$\sigma_-(0) = -1,\ \sigma_-'(x) = \begin{cases} -\frac{2\|\phi\|_{\dot{W}^{1,\infty}}}{(1+\phi(D))\phi(D)}, & x \to 0+ \\ \frac{-\phi(D)\sigma_-(x)^2 - \phi(D)\sigma_-(x) + 2\|\phi\|_{\dot{W}^{1,\infty}} x}{-\phi(D)^2 x} & x > 0 \end{cases}. \tag{2.45b}$$

2.8 2D Euler-Poisson equation revisit

In this section, we come back to 2D Euler-Poisson equation. As discussed in section 2.2.3, critical thresholds are difficult to achieve because of the presence of the non-local Reisz transform. We introduce some new existence results which is stronger than what's been proved in the restricted models.

Recall the dynamics (2.11),

$$\begin{aligned} \rho' &= -\rho d, \\ d' &= -\frac{d^2 + \beta\rho^2}{2} + k\rho. \end{aligned}$$

where $\beta = \beta_q^2 + \beta_r^2 - \beta_\omega^2$. The dynamics of β_q, β_r and β_q are given in (2.10). Here, we repeat the expression.

$$\begin{aligned}
\beta_q' &= k\frac{(R_1R_1 - R_2R_2)[\rho]}{\rho}, \\
\beta_r' &= -2k\frac{R_1R_2[\rho]}{\rho}, \\
\beta_\omega' &= 0.
\end{aligned}$$

For restricted model, we assume that β is a constant in time. But for the full model, β is allowed to grow in time.

2.8.1 Global existence for a modified system

We will prove a global existence result for a modified system with the following hypotheses:

$$|\beta_q'| \leq C, \quad |\beta_r'| \leq C. \tag{2.46}$$

Remark 2.8.1. The Reisz transform R_iR_j "almost" maps L^∞ to L^∞, with possible violation of a logarithmic blowup. In fact, it maps L^∞ to BMO. Therefore, the hypothesis is "almost" true. But, "almost" true can still be wrong. See discussions in section 2.8.2.

Under this hypotheses, β is allow to grow at most t^2 along the characteristics.

$$\begin{aligned}
\beta =& \beta_p^2 + \beta_q^2 - \beta_\omega^2 \leq (|\beta_p(0)| + Ct)^2 + (|\beta_q(0)| + Ct)^2 - \beta_\omega(0)^2 \\
=& 2C^2\left(t + \frac{|\beta_p(0)| + |\beta_q(0)|}{2C}\right)^2 - \left[\beta_\omega(0)^2 - \frac{(|\beta_p(0)| - |\beta_q(0)|)^2}{2}\right].
\end{aligned}$$

Assume $|\beta_\omega(0)| \geq (|\beta_p(0)| - |\beta_q(0)|)/\sqrt{2}$. Define $D := (|\beta_p(0)| + |\beta_q(0)|)/(2C)$. We get

$$\begin{aligned}
\rho' &= -\rho d, \\
d' &\geq -\frac{d^2}{2} - C^2(t+D)^2\rho^2 + k\rho.
\end{aligned}$$

Apply the following scaling to the system:

$$\tilde{d}(t) = (t+D)d(t), \quad \tilde{\rho}(t) = (t+D)^2\rho(t). \tag{2.47}$$

The dynamics on $(\tilde{d}, \tilde{\rho})$ reads

$$\begin{aligned}
(t+D)\tilde{d}' &\geq \tilde{d} - \frac{\tilde{d}^2}{2} - k^2C^2\tilde{\rho}^2 + k\tilde{\rho}, \\
(t+D)\tilde{\rho}' &= (2-\tilde{d})\tilde{\rho}.
\end{aligned}$$

Finally, we take $\tau = \log(t+D) - \log D$. The dynamics of $(\tilde{d}, \tilde{\rho})$ do not depend explicitly on τ.

$$\begin{aligned}
\tilde{d}' &\geq \tilde{d} - \frac{\tilde{d}^2}{2} - k^2C^2\tilde{\rho}^2 + k\tilde{\rho}, \\
\tilde{\rho}' &= (2-\tilde{d})\tilde{\rho}.
\end{aligned}$$

We first analyze the equality system

$$\begin{aligned}
e' &= e - \frac{e^2}{2} - k^2C^2\eta^2 + k\eta, \\
\eta' &= (2-e)\eta.
\end{aligned}$$

We draw the phase plane of the equality system in figure 2.5. The system has a source $O(0,0)$, a sink $B(0,2)$ and a saddle point $A((kC^2)^{-1}, 2)$. If (η_0, e_0) lies in the shaded area, (η, e) will converges to point B as time goes to infinity. The procedure to derive the critical threshold (dashed line) is the same as section 2.3.3. Hence, details are omitted.

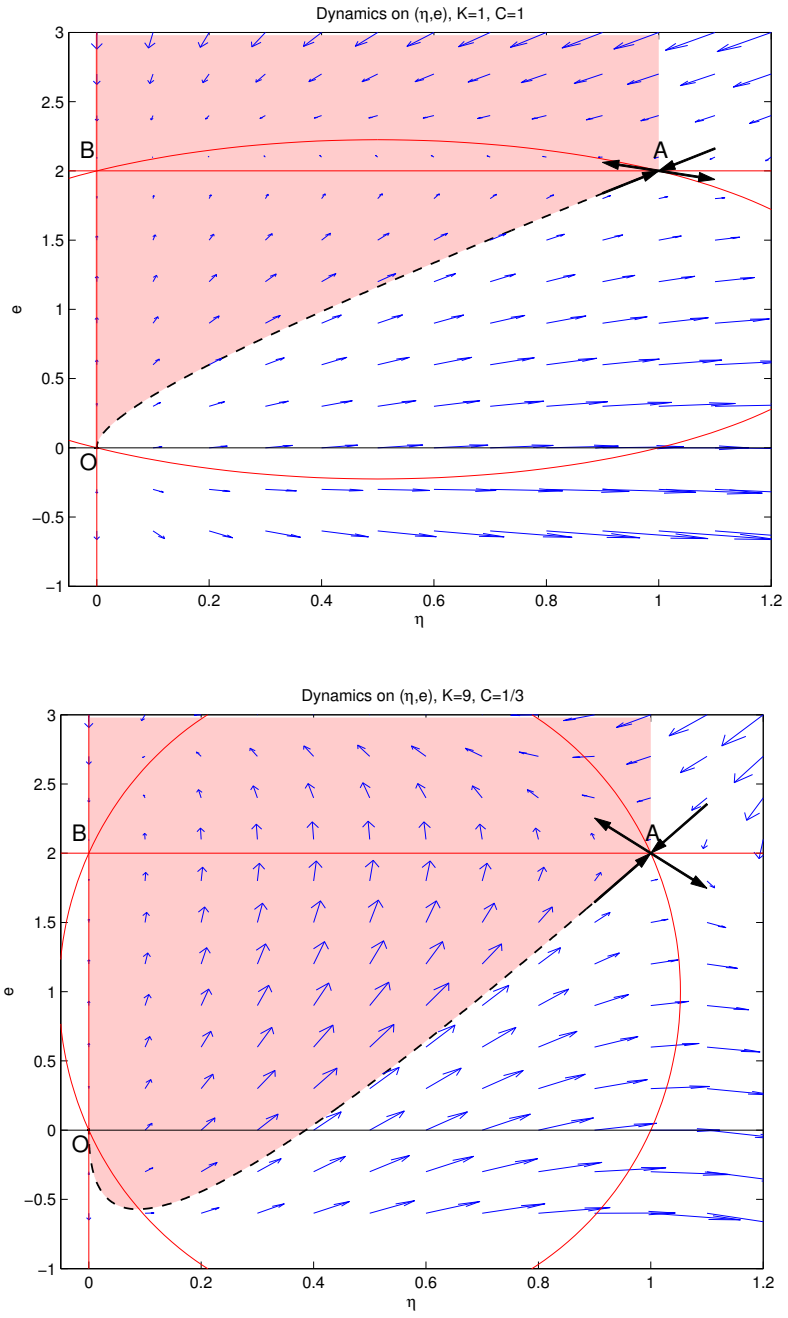

Figure 2.5: Phase plane and critical threshold of the scaled system (η, e). It also represents the threshold for (ρ, d) as well, through a scaling and a comparison principle. The graph above is the case where k is relatively small, where subcritical region only contains increasing profiles. The graph below is the case where k is relatively large, where subcritical region allows more general initial configurations.

Remark 2.8.2. There are two constants in the dynamics. k prescribes the strength of the force. Larger k helps to achieve global regularity. C is a measurement of spectral gap. Smaller C gives better bounds of the spectral gap, which yields better estimates. In particular, if $C = 0$, we recover the restricted model (2.11), where the critical threshold is precisely given. Figure 2.5 provides thresholds for two sets of choices of (k, C), where kC^2 is fixed in order to have the same saddle point A. In the left graph, k is relatively smaller and C is relatively larger. We observe that the shaded area has no intersection with $e < 0$. In this case, convection dominates, and general initial profile lead to blowup, just like Euler equation. In contrast, the right graph has larger k and smaller C. The forcing is stronger. It is possible to have initial profile with $e_0 < 0$ which leads to global regularity.

Next, we use the following comparison principle to link $(\tilde{\rho}, \tilde{d})$ with (η, e).

Lemma 2.8.1. *If* $\begin{cases} 0 \le \tilde{\rho}(0) \le \eta(0) \\ \tilde{d}(0) \ge e(0) \end{cases}$ *and* $\eta(t) \le (kC^2)^{-1}$, *then* $\begin{cases} 0 \le \tilde{\rho}(t) \le \eta(t) \\ \tilde{d}(t) \ge e(t) \end{cases}$.

Proof. If the first inequality is violated starting at time t_0, then $\tilde{\rho}(t_0) = \eta(t_0)$ and $\frac{d}{dt}(\tilde{\rho} - \eta)(t_0) > 0$. This leads to contradiction

$$\frac{d}{dt}(\tilde{\rho} - \eta)(t_0) = (2 - \tilde{d}(t_0))\tilde{\rho}(t_0) - (2 - e(t_0))\eta(t_0) = -(\tilde{d}(t_0) - e(t_0))\tilde{\rho}(t_0) \le 0.$$

Similarly, if the second inequality is violated starting at time t_0, then $\tilde{d}(t_0) = e(t_0)$ and $\frac{d}{dt}(\tilde{d} - e)(t_0) > 0$. We get the following contradiction

$$\begin{aligned} \frac{d}{dt}(\tilde{d} - e)(t_0) &= -k^2C^2\tilde{\rho}(t_0)\left((kC^2)^{-1} - \tilde{\rho}(t_0)\right) + k^2C^2\eta(t_0)\left((kC^2)^{-1} - \eta(t_0)\right) \\ &\le -k^2C^2(\tilde{\rho}(t_0) - \eta(t_0))((kC^2)^{-1} - \eta(t_0)) \le 0. \end{aligned}$$

□

Using the comparison principle, we conclude that if $(\tilde{\rho}(0), \tilde{d}(0))$ lies in the shaded area, then $(\tilde{\rho}(t), \tilde{d}(t))$ exists in all time.

Note that $\tilde{d}(0) = d(0)$ and $\tilde{\rho}(0) = \rho(0)$. Therefore, we achieve global existence for the modified system with hypothesis (2.46), provided that the initial data lies in the subcritical region, namely $(\rho(\mathbf{x}), div\,\mathbf{u}_0(\mathbf{x}))$ belongs to the shaded area in figure 2.5, for all $\mathbf{x}$.

2.8.2 Discussion on the full system

In this section, we briefly discuss how realistic the hypotheses (2.46), for the full 2D Euler-Poisson equation (2.8).

The rescaling argument (2.47) implies that long time behaviors of (ρ, d) for subcritical initial data are

$$\rho(t) \sim \frac{1}{t^2}, \quad d(t) \sim \frac{1}{t}.$$

Thus, if $R_i R_j[\rho](t) \sim \frac{1}{t^2}$, hypotheses (2.46) is valid.

However, the Riesz transform fails to map L^∞ to L^∞. One needs a bit more regularity and integrability to control $R_i R_j[\rho]$ point-wise. The following lemma provides an estimate.

Lemma 2.8.2. *$\|R_1 R_2[\rho]\|_{L^\infty}$ and $\|(R_1 R_1 - R_2 R_2)[\rho]\|_{L^\infty}$ are both bounded (up to a constant) by $\|\rho\|_{L^\infty}\left[\frac{1}{\gamma}\log_+\|\rho\|_{C^\gamma} + \frac{p-1}{2}\log_+\|\rho\|_{L^p}\right]$.*

Proof. We postpond the proof to the appendix section 2.B. □

Heuristically speaking, $\|d\|_{L^\infty} \sim t^{-1}$, $\|\rho\|_{L^\infty} \sim t^{-2}$. From the continuity equation, we get

$$\|\rho(\cdot,t)\|_{C^\gamma} \lesssim \|\rho_0\|_{C^\gamma} \exp\left(\int_0^t \|d(\cdot,s)\|_{L^\infty} ds\right) \sim \exp\left(\int^t \frac{1}{s} ds\right) \sim t,$$
$$\|\rho(\cdot,t)\|_{L^p} \lesssim \|\rho_0\|_{L^p} \exp\left(\int_0^t \|d(\cdot,s)\|_{L^\infty} ds\right) \sim \exp\left(\int^t \frac{1}{s} ds\right) \sim t.$$

It implies

$$\frac{R_i R_j[\rho]}{\rho} \lesssim \frac{t^{-2}\log t}{t^{-2}} = \log t.$$

We are not able to close the system as hypothesis (2.46) is not valid with an extra logarithmic growth.

Appendix

2.A Local dissipation system

In this section, we formally derive the local dissipation system (2.2). Recall the rescaled system

$$\rho_t + div(\rho \mathbf{u}) = 0,$$
$$(\rho \mathbf{u})_t + div(\rho \mathbf{u} \otimes \mathbf{u}) + \nabla P = \int_{\mathbb{R}^n} \phi_\varepsilon(|\mathbf{x}-\mathbf{y}|)(\mathbf{u}(\mathbf{y})-\mathbf{u}(\mathbf{x}))\rho(\mathbf{x})\rho(\mathbf{y})d\mathbf{y}.$$

where

$$\phi_\varepsilon := \frac{1}{\varepsilon^{n+2}}\phi\left(\frac{|\mathbf{x}|}{\varepsilon}\right).$$

The alignment becomes local when ε goes to zero. Formally, assuming that $(\rho, \mathbf{u})$ are sufficiently smooth, we apply Taylor expansion on $(\rho, \mathbf{u})$ around point $\mathbf{x}$.

$$\begin{aligned}
&\int_{\mathbb{R}^n} \phi_\varepsilon(|\mathbf{x}-\mathbf{y}|)(\mathbf{u}(\mathbf{y})-\mathbf{u}(\mathbf{x}))\rho(\mathbf{y})d\mathbf{y} \\
=&\int_{\mathbb{R}^n} \phi_\varepsilon(|\mathbf{y}|)(\mathbf{u}(\mathbf{x}-\mathbf{y})-\mathbf{u}(\mathbf{x}))\rho(\mathbf{x}-\mathbf{y})d\mathbf{y} \\
=&\int_{\mathbb{R}^n} \phi_\varepsilon(|\mathbf{y}|)\left(-D\mathbf{u}(\mathbf{x})\mathbf{y} + \frac{1}{2}\mathbf{y}\cdot D^2\mathbf{u}(\mathbf{x})\mathbf{y} + o(|\mathbf{y}^2|)\right)(\rho(\mathbf{x}) - D\rho(\mathbf{x})\mathbf{y} + o(|\mathbf{y}|))\,d\mathbf{y} \\
=&\frac{1}{2}\rho(\mathbf{x})\int_{\mathbb{R}^n} \phi_\varepsilon(|\mathbf{y}|)\mathbf{y}\cdot D^2\mathbf{u}(\mathbf{x})\mathbf{y}d\mathbf{y} \;+\; \int_{\mathbb{R}^n} \phi_\varepsilon(|\mathbf{y}|)D\rho(\mathbf{x})\mathbf{y}\; D\mathbf{u}(\mathbf{x})\mathbf{y}d\mathbf{y} \;+\; \int_{\mathbb{R}^n} \phi_\varepsilon(|\mathbf{y}|)o(|\mathbf{y}|^2)d\mathbf{y} \\
=&\,\mathrm{I} \,+\, \mathrm{II} \,+\, \mathrm{III}.
\end{aligned}$$

We simplify each term

$$\begin{aligned} \mathrm{I} &= \frac{1}{2}\rho(\mathbf{x})\int_{\mathbb{R}^n}\phi_\varepsilon(|\mathbf{y}|)\sum_{i,j}\partial_{ij}^2\mathbf{u}(\mathbf{x})y_iy_jd\mathbf{y} = \frac{1}{2}\rho(\mathbf{x})\int_{\mathbb{R}^n}\phi_\varepsilon(|\mathbf{y}|)\sum_{i}\partial_{ii}^2\mathbf{u}(\mathbf{x})y_i^2d\mathbf{y} \\ &= \frac{1}{2}\rho(\mathbf{x})\Delta\mathbf{u}(\mathbf{x})\left[\frac{1}{n}\int_{\mathbb{R}^n}\phi_\varepsilon(|\mathbf{y}|)|\mathbf{y}|^2d\mathbf{y}\right]. \end{aligned}$$

Here we use the fact that $\int_{\mathbb{R}^n}\phi_\varepsilon(\mathbf{y})y_iy_jd\mathbf{y}=0$ for $i\neq j$ by anti-symmetry, and $\int_{\mathbb{R}^n}\phi_\varepsilon(\mathbf{y})y_i^2d\mathbf{y}$ is the same quantity for all i's. The coefficient is independent with respect to ε,

$$\int_{\mathbb{R}^n}\phi_\varepsilon(|\mathbf{y}|)|\mathbf{y}|^2d\mathbf{y}=\omega_n\int_{\mathbb{R}}\phi_\varepsilon(s)s^2s^{n-1}ds\overset{r=s/\varepsilon}{=}\omega_n\int_{\mathbb{R}}\phi(r)r^{n+1}dr,$$

where ω_n is the surface area of a unit sphere in $\mathbb{R}^n$. Similarly,

$$\mathrm{II} = \int_{\mathbb{R}^n}\phi_\varepsilon(|\mathbf{y}|)\sum_{i,j}\partial_i\rho(\mathbf{x})\partial_j\mathbf{u}(\mathbf{x})y_iy_jd\mathbf{y}=(\nabla\rho\cdot\nabla)\mathbf{u}\left[\frac{1}{n}\int_{\mathbb{R}^n}\phi_\varepsilon(|\mathbf{y}|)|\mathbf{y}|^2d\mathbf{y}\right],$$

and $\mathrm{III}=o(1)$ vanishes as $\varepsilon\to 0$.

Put the three parts together, we conclude with the limiting system (2.2), as

$$\begin{aligned} &\int_{\mathbb{R}^n}\phi_\varepsilon(|\mathbf{x}-\mathbf{y}|)(\mathbf{u}(\mathbf{y})-\mathbf{u}(\mathbf{x}))\rho(\mathbf{x})\rho(\mathbf{y})d\mathbf{y}=\rho(\mathbf{x})(\mathrm{I}+\mathrm{II}+\mathrm{III}) \\ &\quad= C\rho(\mathbf{x})\big(\rho(\mathbf{x})\Delta\mathbf{u}(\mathbf{x})+2(\nabla\rho\cdot\nabla)\mathbf{u}\big)+o(1)\rho(\mathbf{x})\overset{\varepsilon\to0}{\longrightarrow}Cdiv\,(\rho^2\nabla\mathbf{u}), \end{aligned}$$

where $C=\dfrac{\omega_n}{2n}\displaystyle\int_{\mathbb{R}}\phi(r)r^{n+1}dr$.

2.B L^∞ estimate for Reisz transform

The Reisz transform $R_iR_j[\rho]=\nabla\otimes\nabla\Delta^{-1}\rho$ does not map L^∞ to L^∞. Here is an estimate which requires a little more than L^∞. We proceed with $\mathbb{R}^2$ for simplicity.

$\Delta^{-1}\rho=\mathscr{N}\star\rho$, where $\mathscr{N}$ is the Newtonian potential

$$\mathscr{N}(\mathbf{x})=\frac{1}{2\pi}\log|\mathbf{x}|.$$

Take the first derivative

$$\begin{aligned}
\partial_{x_i}\Delta^{-1}\rho(\mathbf{x}) &= p.v.\int_{\mathbb{R}^2}\mathscr{N}(\mathbf{y})\partial_{x_i}\rho(\mathbf{x}-\mathbf{y})d\mathbf{y} \\
&= -p.v.\int_{\mathbb{R}^2}\mathscr{N}(\mathbf{y})\partial_{y_i}\rho(\mathbf{x}-\mathbf{y})d\mathbf{y} \\
&= \lim_{\delta\to 0}\left[\int_{|\mathbf{y}|\geq\delta}\partial_{y_i}\mathscr{N}(\mathbf{y})\rho(\mathbf{x}-\mathbf{y})d\mathbf{y} - \int_{|\mathbf{y}|=\delta}\mathscr{N}(\mathbf{y})\rho(\mathbf{x}-\mathbf{y})\frac{-y_i}{\delta}d\mathbf{y}\right] \\
&= p.v.\int_{\mathbb{R}^2}\partial_{x_i}\mathscr{N}(\mathbf{y})\rho(\mathbf{x}-\mathbf{y})d\mathbf{y} + \lim_{\delta\to 0}\int_{|\mathbf{z}|=1}\delta\mathscr{N}(\delta\mathbf{z})\rho(\mathbf{x}-\delta\mathbf{z})z_i d\mathbf{z}
\end{aligned}$$

The second term vanishes as

$$\lim_{\delta\to 0}\int_{|\mathbf{z}|=1}\mathscr{N}(\delta\mathbf{z})\rho(\mathbf{x}-\delta\mathbf{z})z_i d\mathbf{z} = \rho(\mathbf{x})\lim_{\delta\to 0}\int_{|\mathbf{z}|=1}\delta\mathscr{N}(\delta\mathbf{z})z_i d\mathbf{z} = 0.$$

Take the second derivative. The same procedure yields

$$\partial^2_{x_ix_j}\Delta^{-1}\rho(\mathbf{x}) = p.v.\int_{\mathbb{R}^2}\partial^2_{x_ix_j}\mathscr{N}(\mathbf{y})\rho(\mathbf{x}-\mathbf{y})d\mathbf{y} + \rho(\mathbf{x})\lim_{\delta\to 0}\int_{|\mathbf{z}|=1}\delta\partial_{x_i}\mathscr{N}(\delta\mathbf{z})z_j d\mathbf{z}.$$

This time, the second term does not vanish as $\partial_{x_i}\mathscr{N}$ is homogeneous of degree -1. Hence $\delta\partial_{x_i}\mathscr{N}(\delta\mathbf{z}) = \partial_{x_i}\mathscr{N}(\mathbf{z})$ and the second term equals to $\int_{|\mathbf{z}|=1}\partial_{x_i}\mathscr{N}(\mathbf{z})z_j d\mathbf{z}$.

Remark 2.B.1. Take $(\partial^2_{x_1x_1}+\partial^2_{x_2x_2})\Delta^{-1}\rho$. The first term equals to 0 as $\Delta\mathscr{N}=0$. The second term equals to ρ. It clearly implies that if we take the trace, we recover ρ.

Remark 2.B.2. Take $(\partial^2_{x_1x_1}-\partial^2_{x_2x_2})\Delta^{-1}\rho$ or $\partial^2_{x_1x_2}\Delta^{-1}\rho$. It is easy to see that the second term equals to 0 due to symmetry. It means that local information are completely stored in the divergence part. All other parts don't "see" the local information.

To bound the full matrix $\nabla\otimes\nabla\Delta^{-1}\rho$, we just need to bound the three terms in the above remarks. The divergence part is clearly bounded. We try to bound the rest.

Take in general K a singular integral homogeneous of degree -2. Note for the two terms in remark 2, K can be expressed as $\frac{x_2^2-x_1^2}{\pi|\mathbf{x}|^4}$ and $\frac{x_1x_2}{\pi|\mathbf{x}|^4}$, respectively. The goal is to bound $p.v.\int_{\mathbb{R}^2} K(\mathbf{y})\rho(\mathbf{x}-\mathbf{y})d\mathbf{y}$.

Decompose $\int_{|\mathbf{y}|\geq\delta} K(\mathbf{y})\rho(\mathbf{x}-\mathbf{y})d\mathbf{y}$ to three parts: near origin, mid-range and near infinity.

$$\int_{|\mathbf{y}|\geq\delta} = \int_{\delta\leq|\mathbf{y}|<\varepsilon} + \int_{\varepsilon\leq|\mathbf{y}|<R} + \int_{|\mathbf{y}|\geq R} = \mathrm{I}+\mathrm{II}+\mathrm{III}.$$

For I, we assume some Hölder continuity C^γ on ρ at point $\mathbf{x}$:

$$\sup_{\mathbf{y}} \frac{|\rho(\mathbf{x})-\rho(\mathbf{y})|}{|\mathbf{x}-\mathbf{y}|^\gamma} = C(\mathbf{x}) < \infty.$$

Note that by anti-symmetry, $\int_{a<|\mathbf{x}|<b} K(\mathbf{x})d\mathbf{x} = 0$ on any ring. Therefore,

$$\mathrm{I} = \int_{\delta\leq|\mathbf{y}|<\varepsilon} K(\mathbf{y})(\rho(\mathbf{x}-\mathbf{y})d\mathbf{y}-\rho(\mathbf{x})) \leq C(\mathbf{x})\int_{\delta\leq|\mathbf{y}|<\varepsilon} K(\mathbf{y})|\mathbf{y}|^\gamma d\mathbf{y}.$$

Thanks to the additional γ regularity, the integral at the right hand side converges as $\delta\to 0$. Moreover,

$$\int_{\delta\leq|\mathbf{y}|<\varepsilon} K(\mathbf{y})|\mathbf{y}|^\gamma d\mathbf{y} \sim \int_\delta^\varepsilon \frac{1}{r^2} r^\gamma r dr \sim \varepsilon^\gamma.$$

Remark 2.B.3. It is clear that continuity is not enough to control the origin. When $\gamma = 0$, $\int \frac{1}{r}dr$ blows up at zero.

We pick $\varepsilon = \min\{C(\mathbf{x})^{-1/\gamma}, 1\}$ to guarantee boundedness of I.

For III, we observe that K blow up at infinity as well. Therefore, we can not use $\|\rho\|_{L^\infty}$ to control. Any L^p norm which carries decay information at infinity will be enough to control the far field.

$$\mathrm{III} \leq \|\rho\|_{L^p}\left(\int_{|\mathbf{y}|\geq R}|K(\mathbf{y})|^{p'}d\mathbf{y}\right)^{1/p'} \sim \|\rho\|_{L^p}\int_R^\infty \frac{1}{r^{2p'}} r dr \sim \|\rho\|_{L^p} R^{-\frac{2}{p-1}}.$$

Similarly, we pick $R = \max\{\|\rho\|_{L^p}^{\frac{p-1}{2}}, 1\}$.

For II, there is no singularity, so we simply use $\|\rho\|_{L^\infty}$ to control.

$$\begin{aligned}\text{II} \leq & \|\rho\|_{L^\infty} \int_{\varepsilon \leq |\mathbf{y}| < 1} K(\mathbf{y}) d\mathbf{y} \sim \|\rho\|_{L^\infty} \int_{\varepsilon}^{R} \frac{1}{r^2} r dr \sim \|\rho\|_{L^\infty} \left(\log \frac{1}{\varepsilon} + \log R \right) \\ \sim & \|\rho\|_{L^\infty} \left[\frac{1}{\gamma} \log_+ C(\mathbf{x}) + \frac{p-1}{2} \log_+ \|\rho\|_{L^p} \right].\end{aligned}$$

We end up with the lemma 2.8.2.

Chapter 3: Kinetic Flocking Models

In this chapter, we turn to study flocking models in the kinetic level.

The existence theory for kinetic flocking systems has been stated in theorem 1.3.1. The Vlasov-type equation is easier to analyze, comparing to the macroscopic system, as the nonlinear convection term is replaced by a linear free transport. Moreover, the large time behavior of the system is also discussed in section 1.3.2. If the influence function ϕ decays slow enough at infinity, namely (1.8) is satisfied, then there is *unconditional flocking*. For more realistic influence function which is compactly supported, there is cluster formation in large time. One big open problem is to identify or estimate number of clusters when time goes to infinity.

Both flocking and clustering effects require concentration of velocity. In the kinetic level, the solution f tends to be singular in v variable as time approaches infinity, despite the fact of been smooth in any finite time. Such concentration appears in many kinetic systems with energy dissipation [48]. The generation of $\delta-$singularity brings challenges for the numerical implementation. Many techniques use smooth approximations for the singularity. They suffer large errors as the solution becomes more and more singular.

We propose a discontinuous Galerkin (DG) method to solve kinetic flocking systems numerically. Discontinuous Galerkin method is first introduced by Reed and Hill

in [81] and has many successful applications in hyperbolic conservation laws. The idea is to use piecewise polynomials to approximate the solution in the weak sense. The use of weak formulation of the solution overcomes the inaccuracy of the scheme. The efficiency of DG method on δ-singularity has been studied in [105] and more applications are discussed in [106].

We organize this chapter as follows. We start with a splitting argument in section 3.1 to separate the free transport and the alignment. The main focus thereafter is the alignment part (3.3), which is the main reason for generation of δ-singularities. A preliminary introduction on DG method for kinetic Cucker-Smale system is given in section 3.2, followed by a first order scheme. Higher order DG schemes are established and discussed as well, where the imperative positivity preserving property is proved to ensure stability of the numerical scheme. The detailed algorithm is listed in appendix section 3.A. In section 3.3, we briefly discuss the kinetic Motsch-Tadmor system, with all arguments similar as the Cucker-Smale case. Some examples are provided in section 3.4 to demonstrate the good performance of the high order DG scheme.

3.1 General setup and splitting

Recall kinetic flocking system

$$\partial_t f + \mathbf{v}\cdot\nabla_{\mathbf{x}} f + \nabla_{\mathbf{v}}\cdot(fL(f)) = 0. \tag{3.1}$$

Here, $f = f(t,\mathbf{x},\mathbf{v})$ denotes the number density at postion $\mathbf{x}$, velocity $\mathbf{v}$ and time t. The operator L is defined as

$$L(f)(t,\mathbf{x},\mathbf{v}) = \iint \phi(|\mathbf{x}-\mathbf{y}|)(\mathbf{v}^*-\mathbf{v})f(t,\mathbf{y},\mathbf{v}^*)d\mathbf{y}d\mathbf{v}^*, \tag{C-S}$$

for Cucker-Smale model, and as

$$L(f)(t,\mathbf{x},\mathbf{v}) = \iint \frac{\phi(|\mathbf{x}-\mathbf{y}|)}{\Phi(t,\mathbf{x})}(\mathbf{v}^*-\mathbf{v})f(t,\mathbf{y},\mathbf{v}^*)d\mathbf{y}d\mathbf{v}^*, \tag{M-T}$$

for Motsch-Tadmor model, where

$$\Phi(t,\mathbf{x}) = \iint \phi(|\mathbf{x}-\mathbf{y}|)f(t,\mathbf{y},\mathbf{v}^*)d\mathbf{y}d\mathbf{v}^*.$$

Our goal is to design an accurate and stable numerical scheme which successfully captures the asymptotic behavior of velocity concentration.

The main system (3.1) can be split into two components: the free transport part

$$\partial_t f(t,\mathbf{x},\mathbf{v}) = -\mathbf{v}\cdot\nabla_{\mathbf{x}} f(t,\mathbf{x},\mathbf{v}), \tag{3.2}$$

and the alignment part

$$\partial_t f(t,\mathbf{x},\mathbf{v}) = -\nabla_{\mathbf{v}}\cdot(fL(f)). \tag{3.3}$$

The free transport part does not lead to concentrations or singularities. Many standard methods can be used to solve (3.2). For instance, to get first order accuracy, one can follow the characteristics, which are parallel straight lines travelling along the direction v. A simple linear interpolation can be used to obtain a solution with first order accuracy. For higher accuracy, many stable and effective methods can be used. For instance, we apply WENO scheme (c.f. [88]) to solve (3.2).

The alignment part is our main concern. Through a similar proof as in proposition 1.3.3, we argue that (3.3) converges to a flock provided that $\phi(S_0) > 0$. In fact, S won't change in time due to no presence of transport, and V decays exponentially with rate $\phi(S_0)$. If the support of ϕ is small, clusters may generate in large time.

Both flocking and clustering imply generations of δ-singularities as time approaches infinity. We derive a DG scheme in section 3.2 and 3.3, for (C-S) setup and (M-T) setup, respectively.

Provided solvers (3.2) and (3.3), we solve the full system (3.1), using Strang splitting method [90] or other higher order splitting method. We refer the reader to [70] for a review of splitting methods.

We summarize the full algorithm in appendix 3.A.

3.2 A discontinuous Galerkin method on kinetic flocking systems

In this section, we discuss the numerical implementation of the alignment part of the main system (3.3). We shall first focus on the (C-S) setup. For (M-T) case, similar argument can be made. See section 3.3 for details. Also, we take 1D as an easy illustration. See remark 3.2.5 for discussions on multi-dimensional system.

Rewrite (3.3) as follows

$$\partial_t f(t,x,v) = -\nabla_v \cdot (fL(f)) = -\nabla_v \cdot \left(f(t,x,v) \int (v^* - v) G(t,x,v^*) dv^* \right), \tag{3.4}$$

where $G(t,x,v) = \int \phi(|x-y|) f(t,y,v) dy$, which can be efficiently solved by numerical solvers for convolutions.

As (3.4) is homogeneous in x, we drop the x dependency for simplicity from now on.

3.2.1 The DG framework

The idea of the discontinuous Galerkin method is to use piecewise polynomial to approximate the solution.

We partition the computational domain $\Omega = [a,b]$ on v into N cells $\{I_j\}_{j=1}^N$

$$I_j = \left(v_{j-1/2}, v_{j+1/2}\right), \quad v_j = a + (j-1/2)\Delta v, \quad \Delta v = \frac{b-a}{N},$$

with uniform mesh size $h := \Delta v$ for simplicity. The space we are working with is

$$V_h := \left\{ f \ : \ \text{For all } j = 1, \cdots, N, \ f|_{I_j} \in \mathscr{P}_k \right\},$$

where $\mathscr{P}_k$ denotes polynomial of degree at most k. The weak formulation of (3.4) reads

$$\frac{d}{dt}\int_{I_j} f(v)p(v)dv = -pfL(f)\Big|_{v_{j-1/2}}^{v_{j+1/2}} + \int_{I_j} fL(f)\phi' dv, \quad \forall p = p(v) \in V_h. \tag{3.5}$$

The DG scheme is to find $f \in V_h$ which satisfies (3.5).

If we apply test function $p(v) = 1$ on (3.5), we get

$$\frac{d}{dt}\bar{f}_j = -\frac{1}{h} fL(f)\Big|_{v_{j-1/2}}^{v_{j+1/2}},$$

where $\bar{f}_j$ is the cell average of I_j. With a forward Euler scheme in time, this becomes the classical finite volume method, namely

$$\bar{f}_j(t+\Delta t) = \bar{f}_j(t) + \frac{\Delta t}{h}\left[f(v_{j-1/2}^+)\cdot L(f)(v_{j-1/2}) - f(v_{j+1/2}^-)\cdot L(f)(v_{j+1/2})\right].$$

The key part is to approximate the flux at the cell interfaces. To ensure the conservation law, we modify the scheme using a numerical flux

$$\bar{f}_j(t+\Delta t) = \bar{f}_j(t) + \frac{\Delta t}{h}\left[\hat{f}(v_{j-1/2})\cdot L(f)(v_{j-1/2}) - \hat{f}(v_{j+1/2})\cdot L(f)(v_{j+1/2})\right] \tag{3.6a}$$

so that the outflux and influx at the same interface add up to zero. Note that L is a global operator on f, and $L(f)$ is a continuous at the interface, we need to compute $L(f)$ using information from all cells. Then, with fixed $L(f)(v_{j+1/2})$, the flux is linear in f. We use upwind fluxes where

$$\hat{f}_{j+1/2} := \hat{f}(v_{j+1/2}) = \begin{cases} f(v^-_{j+1/2}) & \text{if } L(f)(v_{j+1/2}) \geq 0 \\ f(v^+_{j+1/2}) & \text{if } L(f)(v_{j+1/2}) < 0 \end{cases}. \tag{3.6b}$$

Remark 3.2.1. We usually use monotone numerical flux for DG scheme. In our simple case when the flux is linear, some widely used flux such as Godunov flux, Lax-Friedrich flux are coincide with the upwind flux.

3.2.2 A first order scheme

Let us consider the simple case when $k = 0$. A piecewise constant approximation yields first order accuracy. To obtain $\bar{f}_j(t + \Delta t)$, we apply scheme (3.6) with

$$f(v^+_{j+1/2}) = \bar{f}_{j+1}, \quad f(v^-_{j-1/2}) = \bar{f}_j,$$

as v is a constant in each cell. We are left with computing $L(f)$. As f is piecewise constant in v for all x, clearly G is also a piecewise constant in v. Hence,

$$\begin{aligned} L(f)(v_{j+1/2}) &= \int_\Omega (v^* - v_{j+1/2}) G(v^*) dv^* \\ &= \sum_{l=1}^{N} \bar{G}_l \int_{I_l} (v^* - v_{j+1/2}) dv^* = h^2 \sum_{l=1}^{N} (l - j - 1/2) \bar{G}_l, \end{aligned}$$

where $\bar{G}_l$ is the value of G in I_l. We can use any first order numerical integration on x to compute $\bar{G}_l$ from $\bar{f}_l$.

We prove the positivity preserving property of the first order scheme, which ensures L^1 stability of the numerical solution.

Proposition 3.2.1. *Suppose $\bar{f}_j(t) > 0$ for all j. Applying the first order scheme, we have $\bar{f}_j(t+\Delta t) > 0$ under CFL condition*

$$\frac{\Delta t}{h} \max_j \left| L(f)(v_{j+1/2}) \right| < \frac{1}{2}. \tag{3.7}$$

Proof. Rewrite (3.6a) as following

$$\bar{f}_j(t+\Delta t) = \frac{1}{2}\left[\bar{f}_j(t) + \frac{2\Delta t}{h}\hat{f}(v_{j-1/2}) \cdot L(f)(v_{j-1/2})\right] + \frac{1}{2}\left[\bar{f}_j(t) - \frac{2\Delta t}{h}\hat{f}(v_{j+1/2}) \cdot L(f)(v_{j+1/2})\right].$$

We will show that both terms are positive under CFL condition.

For the first term, if $L(f)(v_{j-1/2}) \geq 0$, clearly

$$\bar{f}_j(t) + \frac{2\Delta t}{h}\hat{f}(v_{j-1/2}) \cdot L(f)(v_{j-1/2}) = \bar{f}_j(t) + \frac{2\Delta t}{h}\bar{f}_{j-1}(t) \cdot L(f)(v_{j-1/2}) > 0.$$

if $L(f)(v_{j-1/2} < 0$, then under CFL condition, we have

$$\bar{f}_j(t) + \frac{2\Delta t}{h}\hat{f}(v_{j-1/2}) \cdot L(f)(v_{j-1/2}) = \left[1 - \frac{2\Delta t}{h}\left|L(f)(v_{j-1/2})\right|\right]\bar{f}_j(t) > 0.$$

Similarly, the second term is positive under the same CFL condition. Therefore, $\bar{f}_j(t+\Delta t) > 0$, for all j. □

Remark 3.2.2. The CFL condition (3.7) depends on time t. We can derive a sufficient CFL condition where the choice of Δt is independent of t.

$$\left|L(f)(v_{j+1/2})\right| = h^2 \left|\sum_{l=1}^{N}(l-j-1/2)\bar{G}_l\right| \leq h^2(N-1/2)\sum_{l=1}^{N}\bar{G}_l,$$

for any $j = 0, \cdots, N-1$. Here, $M_0 = \iint f(t,x,v)dvdx$ is the total mass, which is conserved in time. Recall the assumption on ϕ: $\|\phi\|_{L^\infty} = \phi(0) = 1$. Hence

$$\sum_{l=1}^{N}\bar{G}_l = \sum_{l=1}^{N}\int \phi(|x-y|)\bar{f}_l(y)dy \leq \sum_{l=1}^{N}\int \bar{f}_l(y)dy = M_0,$$

for f which is assume to be piecewise constant in v.

This implies $\left|L(f)(v_{j+1/2})\right| \leq h(b-a)M_0$ and thus a sufficient CFL condition

$$\frac{\Delta t}{h} < \frac{1}{2M_0(b-a)}.$$

3.2.3 High order DG schemes

In order to obtain high order accuracy, we apply (3.5) with test functions with high orders. Choose Legendre polynomials on I_j

$$p_j^{(0)}(v) = 1, \quad p_j^{(1)}(v) = v - v_j, \quad p_j^{(2)}(v) = (v-v_j)^2 - \frac{1}{12}h^2, \quad \cdots.$$

Denote $f_j^{(l)} = \frac{1}{h^{l+1}} \int_{I_j} f(v) p_j^{(l)} dv$. Clearly, all $f \in \mathscr{P}_k$ can be determined by $f_j^{(l)}$ for $j = 1, \cdots, N$, $l = 0, \cdots, k$. As a matter of fact, we can write $f(v) = \sum_{l=0}^{k} a_l f_j^{(l)} p_j^{(l)}(v)$ for $v \in I_j$, with $a_0 = 1, a_1 = 12/h, a_2 = 180/h^2$, etc. (Consulting [23].)

From (3.5), we obtain the evolution of $f_j^{(l)}$.

$$\begin{aligned}
\frac{d}{dt} f_j^{(0)} &= \frac{1}{h}(\hat{f}_{j-1/2} L_{j-1/2} - \hat{f}_{j+1/2} L_{j+1/2}), \\
\frac{d}{dt} f_j^{(1)} &= -\frac{1}{2h}(\hat{f}_{j-1/2} L_{j-1/2} + \hat{f}_{j+1/2} L_{j+1/2}) + \frac{1}{h^2} \int_{I_j} f L(f) dv, \\
\frac{d}{dt} f_j^{(2)} &= \frac{1}{6h}(\hat{f}_{j-1/2} L_{j-1/2} - \hat{f}_{j+1/2} L_{j+1/2}) + \frac{2}{h^3} \int_{I_j} f L(f)(v - v_j) dv,
\end{aligned} \tag{3.8}$$

etc. Here, we denote $L_{j\pm1/2} = L(f)(v_{j\pm1/2})$ for simplicity.

Next, we compute $L_{j+1/2}$ and the two integrals in the dynamics above, given $f \in V_h$.

For $k = 0$, $L_{j+1/2}$ is given in section 3.2.2. $f_j^{(0)}$ coincide with $\bar{f}_j$.

For $k \geq 1$, we use L^2-orthogonality property of Legendre polynomial to compute

$$\begin{aligned} L(f)(v) &= \int (v^* - v) G(v^*) dv^* \\ &= \sum_{l=1}^{N} \int_{I_l} \left[(v_l - v) p_l^{(0)}(v^*) + p_l^{(1)}(v^*) \right] \cdot \left[G_l^{(0)} p_l^{(0)}(v^*) + \frac{12}{h} G_l^{(1)} p_l^{(1)}(v^*) + \cdots \right] dv^* \\ &= h \sum_{l=1}^{N} (v_l - v) G_l^{(0)} + h^2 \sum_{l=1}^{N} G_l^{(1)}. \end{aligned}$$

All other terms of $G(v^*)$ is L^2-orthogonal to $v^* - v$ and have no contribution to $L(f)(v)$. This implies

$$L_{j+1/2} = h \sum_{l=1}^{N} (v_l - v_{j+1/2}) G_l^{(0)} + h^2 \sum_{l=1}^{N} G_l^{(1)} = h^2 \sum_{l=1}^{N} \left[(l - j - 1/2) G_l^{(0)} + G_l^{(1)} \right].$$

Moreover, $L(f)(v)$ is linear in terms of v. Again, by orthogonality, we get

$$\begin{aligned} &\frac{1}{h^2} \int_{I_j} f L(f) dv \\ =\ &\frac{1}{h^2} \int_{I_j} f(v) \left[\left(h \sum_{l=1}^{N} (v_l - v_j) G_l^{(0)} + h^2 \sum_{l=1}^{N} G_l^{(1)} \right) p_j^{(0)}(v) + \left(-h \sum_{l=1}^{N} G_l^{(0)} \right) p_j^{(1)}(v) \right] dv \\ =\ &h \left\{ f_j^{(0)} \sum_{l=1}^{N} [(l - j) G_l^{(0)} + G_l^{(1)}] - f_j^{(1)} \sum_{l=1}^{N} G_l^{(0)} \right\}. \end{aligned}$$

Finally, for $k \geq 2$,

$$\frac{2}{h^3} \int_{I_j} f L(f)(v - v_j) dv = 2h \left\{ f_j^{(1)} \sum_{l=1}^{N} [(l - j) G_l^{(0)} + G_l^{(1)}] - \left(\frac{1}{12} f_j^{(0)} + f_j^{(2)} \right) \sum_{l=1}^{N} G_l^{(0)} \right\}.$$

Remark 3.2.3. As shown above, to compute the right hand side of (3.8), we need to calculate the following sums:

$$\sum_{l=1}^{N} G_l^{(0)}, \quad \sum_{l=1}^{N} G_l^{(1)} \quad \text{and} \quad \sum_{l=1}^{N} (l - j) G_l^{(0)}.$$

The first two sums are independent of j. The third sum has a convolution structure. Fast convolution solvers could be used to compute the sum.

3.2.4 Positivity preserving

One major difficulty of high order schemes is that the reconstructed solution is not necessarily positive. A negative computational solution will quickly become unstable. Suitable limiters are needed to preserve positivity of the numerical solution. We proceed with the limiter introduced in [108].

First, we extend proposition 3.2.1 to high order schemes and prove positivity for $\bar{f}_j$. To proceed, we use Gauss-Lobatto quadrature points on I_j, denoting $\{v_j^i\}_{i=1}^n$. In particular, $v_j^1 = v_{j-1/2}$ and $v_j^n = v_{j+1/2}$. For f_j a polynomial of degree up to $2n-3$,

$$\bar{f}_j = \frac{1}{h}\int_{I_j} f_j(v)dv = \frac{1}{h}\sum_{i=1}^{n} \alpha_i f_j(v_j^i),$$

where α_i are Gauss-Lobatto weights. For example, when $n=2$, $\alpha_1 = \alpha_2 = 1/2$; when $n=3$, $\alpha_1 = \alpha_3 = 1/6$ and $\alpha_2 = 2/3$. Note that α_i's are all positive, summing up to 1, and symmetric $\alpha_i = \alpha_{n+1-i}$.

Proposition 3.2.2. *Suppose $f_j(t, v_j^i) > 0$ for all Gauss-Lobatto quadrature points v_j^i. Then, for any scheme with forward Euler in time and DG in space with order $k \le 2n-3$, we have $\bar{f}_j(t+\Delta t) > 0$, under CFL condition*

$$\frac{\Delta t}{h}\max_j \left|L_{j+1/2}\right| < \alpha_1. \tag{3.9}$$

Proof. The dynamic of $\bar{f}_j = f_j^{(0)}$ reads

$$\begin{aligned}
\bar{f}_j(t+\Delta t) =& \bar{f}_j(t) + \frac{\Delta t}{h}\left[\hat{f}(v_{j-1/2})\cdot L_{j-1/2} - \hat{f}(v_{j+1/2})\cdot L_{j+1/2}\right] \\
=& \frac{1}{h}\sum_{i=2}^{n-1}\alpha_i f_j(v_j^i) + \alpha_1\left(f_j(v_{j-1/2}) + \frac{\Delta t}{\alpha_1 h}\hat{f}(v_{j-1/2})\cdot L_{j-1/2}\right) \\
&+ \alpha_n\left(f_j(v_{j+1/2}) - \frac{\Delta t}{\alpha_n h}\hat{f}(v_{j+1/2})\cdot L_{j+1/2}\right).
\end{aligned}$$

We check positivity for the last two terms. For the second term, if $L_{j-1/2} \geq 0$, clearly

$$f_j(v_{j-1/2}) + \frac{\Delta t}{\alpha_1 h}\hat{f}(v_{j-1/2}) \cdot L_{j-1/2} = f_j(v_{j-1/2}) + \frac{\Delta t}{\alpha_1 h} f_{j-1}(v_{j-1/2}) \cdot L_{j-1/2} > 0.$$

If $L_{j-1/2} < 0$, then under CFL condition, we have

$$f_j(v_{j-1/2}) + \frac{\Delta t}{\alpha_1 h}\hat{f}(v_{j-1/2}) \cdot L_{j-1/2} = \left[1 - \frac{\Delta t}{\alpha_1 h}|L_{j-1/2}|\right] f_j(v_{j-1/2}) > 0.$$

Similarly, the third term is positive under the same CFL condition, as $\alpha_n = \alpha_1$. Therefore, $\bar{f}_j(t+\Delta t) > 0$, for all j. □

Remark 3.2.4. Similar to remark 3.2.2, there is a sufficient CFL condition independent of t for the high order DG scheme. We estimate the additional part of $L^{j+1/2}$ as below.

$$\left| h^2 \sum_{l=1}^{N} G_l^{(1)} \right| = \sum_{l=1}^{N} \int\int_{I_l} f(y,v)|v - v_l|\phi(|x-y|)dvdy \leq \frac{h}{2} M_0.$$

Together with the estimate for the first part (shown in remark 3.7), we get

$$\left|L_{j+1/2}\right| \leq h\left(N - \frac{1}{2}\right) M_0 + \frac{h}{2} M_0 = hNM_0 = M_0(b-a).$$

With the correction term, we have the same bound on $L_{j+1/2}$. It yields the following sufficient CFL condition

$$\frac{\Delta t}{h} < \frac{\alpha_1}{M_0(b-a)}. \tag{3.10}$$

It is consistent with the condition derived for first order scheme, provided that f is positive at all Gauss-Lobatto quadrature points.

To make sure f_j is positive at Gauss-Lobatto quadrature points, we modify $f(t)$ using an interpolation between the current f and the positive constant $\bar{f} = f^{(0)}$, namely,

in I_j at time $t+\Delta t$,

$$\tilde{f}_j(v) = \theta_j f_j(v) + (1-\theta_j)\bar{f}_j,$$

where $\theta_j \in [0,1]$ to be chosen. When $\theta_j = 1$, there is no modification and high accuracy is preserved. When, $\theta_j = 0$, the modified solution coincides with the first order scheme. Hence, for higher accuracy, θ_j should be as large as possible. On the other hand, we need positivity of $\tilde{f}_j(v_j^i)$, i.e.

$$(\bar{f}_j - f_j(v_j^i))\theta_j < \bar{f}_j,$$

for all i. Therefore, we shall choose θ_j as the following

$$\theta_j = \begin{cases} \frac{\bar{f}_j - \varepsilon}{\bar{f}_j - m_j}, & \text{if } m_j < \varepsilon \\ 1 & \text{if } m_j \geq \varepsilon \end{cases},$$

where $m_j := \min_i f_j(v_j^i), \varepsilon = \min\{10^{-13}, \bar{f}_j\}$.

The modified solution $\tilde{f}_j$ preserves the total mass as well. It implies L^1 stability of the scheme.

We can write the modification in terms of $f_j^{(l)}$ where

$$\tilde{f}_j^{(0)} = f_j^{(0)}, \quad \tilde{f}_j^{(l)} = \theta_j f_j^{(l)}, l \geq 1. \tag{3.11}$$

Indeed, the modification weakens the high order correction at several cells to enforce positivity.

Theorem 3.2.3 (Positivity preserving)**.** *Consider* (3.3) *with operator L defined as* (C-S). *Suppose the initial density f_0 is positive. Consider the above scheme: forward Euler in time, DG in space, with limiter* (3.11). *The solution is positive in all time, under CFL*

condition

$$\lambda < \frac{\alpha_1}{M_0(b-a)},$$

where $\lambda = \Delta t/\Delta v$, α_1 *is a constant depending on order of accuracy,* M_0 *is the total mass and* $b-a$ *is the length of the computational domain in* v.

Remark 3.2.5. The positivity preserving property remains true for multi-dimensional system as well. To construct the limiter, it is necessary to find Gauss-Lobatto quadrature points in multi-D. We refer to [108] for related discussions.

3.2.5 High order time discretization

In this subsection, we discuss time discretization for the ODE systems with respect to $f_j^{(l)}$. We already show positivity preserving and L^1 stability for forward Euler time discretization, under CFL condition (3.9). To get high order accuracy in time, we use strong stability preserving (SSP) Runge-Kutta method [40, 89]. For instance, a second order SSP scheme reads

$$\begin{aligned} f_{[1]} &= \mathtt{FE}(f(t),\Delta t) \\ f(t+\Delta t) &= \frac{1}{2}f(t)+\frac{1}{2}\mathtt{FE}(f_{[1]},\Delta t), \end{aligned}$$

and a third order SSP scheme reads

$$\begin{aligned} f_{[1]} &= \mathtt{FE}(f(t),\Delta t) \\ f_{[2]} &= \frac{3}{4}f(t)+\frac{1}{4}\mathtt{FE}(f_{[1]},\Delta t) \\ f(t+\Delta t) &= \frac{1}{3}f(t)+\frac{2}{3}\mathtt{FE}(f_{[2]},\Delta t). \end{aligned}$$

Here, $\mathtt{FE}(f,\Delta t)$ represents a forward Euler step with size Δt.

As an SSP time discretization is a convex combination of forward Euler, positivity preserving property is granted automatically.

3.3 On Motsch-Tadmor system

In this section, we apply the DG method to solve the kinetic flocking system (3.3) with (M-T) setup. The scheme works the same as the (C-S) setup. As an extra, we compute the normalization factor $\Phi(x)$ as

$$\Phi(x) = \iint \phi(|x-y|) f(t,y,v) dy dv = \sum_{j=1}^{N} \int_{I_j} G(x,v) dv = \frac{1}{h} \sum_{j=1}^{N} G_j^{(0)}.$$

Similar CFL condition can be derived for positivity preserving and L^1 stability.

Theorem 3.3.1 (Positivity preserving)**.** *Consider* (3.3) *with operator L defined as* (M-T). *Suppose the initial density f_0 is positive. Then, the solution generated by the DG scheme is positive in all time, under CFL condition*

$$\lambda < \frac{\alpha_1}{\phi(D)(b-a)},$$

where $\lambda = \Delta t/\Delta v$, α_1 is a constant depending on order of accuracy, $b-a$ is the length of the computational domain in v, and D is a constant defined in (1.7).

Proof. A similar argument as in proposition 3.2.2 implies the following CFL condition:

$$\frac{\Delta t}{h} \max_j \left| \frac{L_{j+1/2}}{\Phi(x)} \right| < \alpha_1.$$

To get a time space independent sufficient CFL condition, we obtain a lower bound on $\Phi(x)$.

$$\Phi(x) = \iint \phi(|x-y|) f(t,y,v) dy dv \geq \phi(D) \iint f(t,y,v) dy dv = \phi(D) M_0,$$

for all x in the support of f. Together with the estimate $|L^{j+1/2}| \leq M_0(b-a)$ shown in remark 3.2.2, we conclude with the desired sufficient condition. □

3.4 Numerical examples

In this section, we present some numerical examples to demonstrate the good performance of the DG scheme applied to kinetic flocking models.

3.4.1 Global influence and unconditional flocking

We consider 1D kinetic Cucker-Smale model with influence function $\phi(r) = (1+r)^{-1/2}$, and the initial density $f_0(x,v) = \chi_{|x|<1}\chi_{|v|<1}$, where χ is the indicator function. As ϕ satisfies (1.8), the solution should converge to a flock.

We set the computational domain as follows. In x direction, we compute D from (1.7) and get $D \approx 2.93$. By symmetry, the support of the solution in x direction lies in $(-1.5, 1.5)$. We set the computational domain on x to be $[-2,2]$ for safety. In v direction, the variation becomes smaller as time increases. Therefore, $[-1,1]$ is an appropriate domain for v. We start the test with mesh size 40×40.

For the time step, the CFL condition (3.10) suggests $\Delta t < \alpha_1/160$. So, for first and second order schemes, we take $\Delta t = 0.003$. For third order scheme, we take $\Delta t = 0.001$.

Figure 3.1 shows the dynamics of density f under DG schemes using piecewise polynomials of degree $k = 0,1,2$. We observe that all three schemes converge to flock. On the other hand, high order schemes converge faster than the low order scheme, which is an indicator of better performance. In figure 3.2, we plot $\int f(t,x,v)dx$ against v at

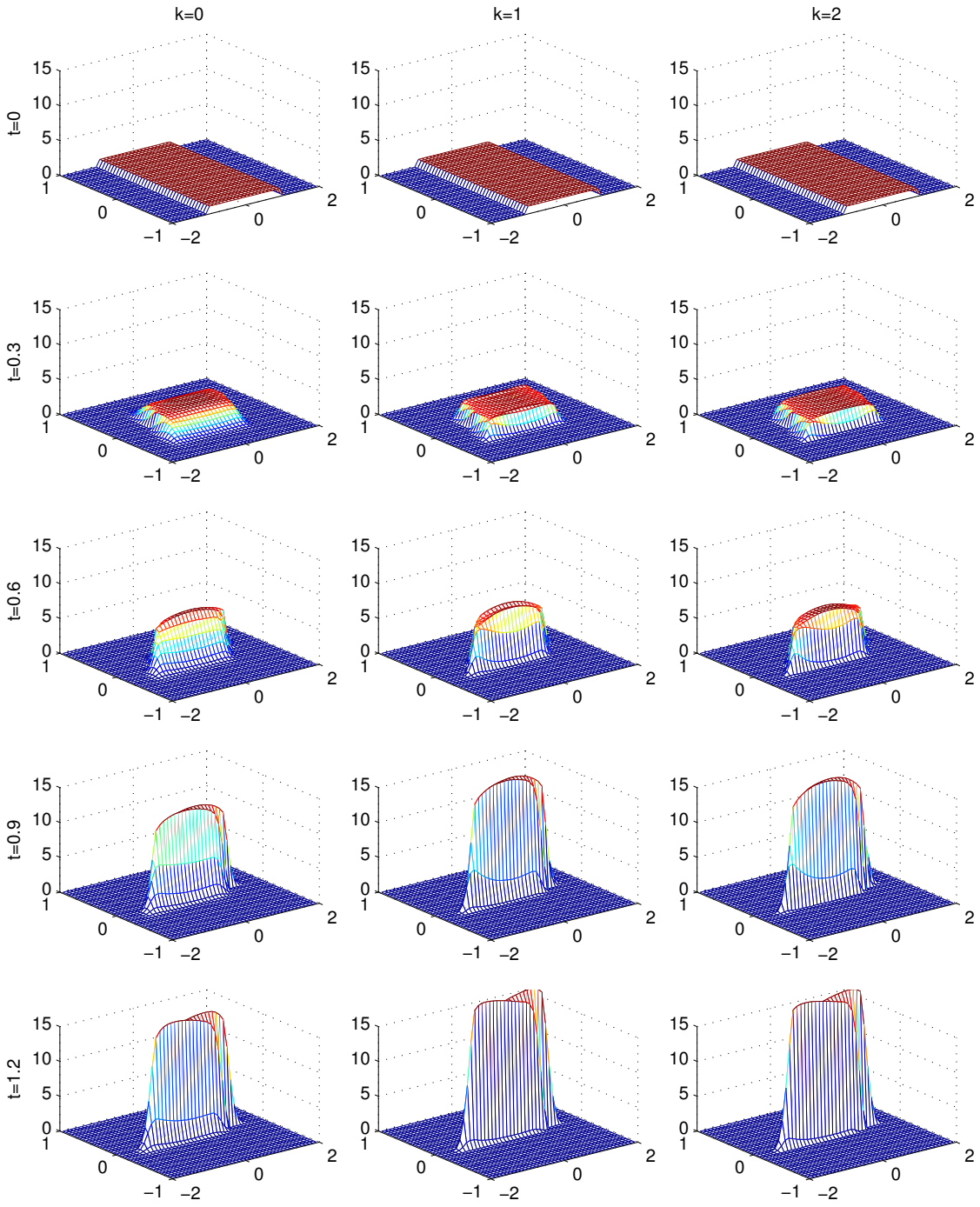

Figure 3.1: Density f at time $t = 0, .3, .6, .9, 1.2$ for DG schemes with $k = 0, 1, 2$. The numerical solution has a concentration in v direction in large time.

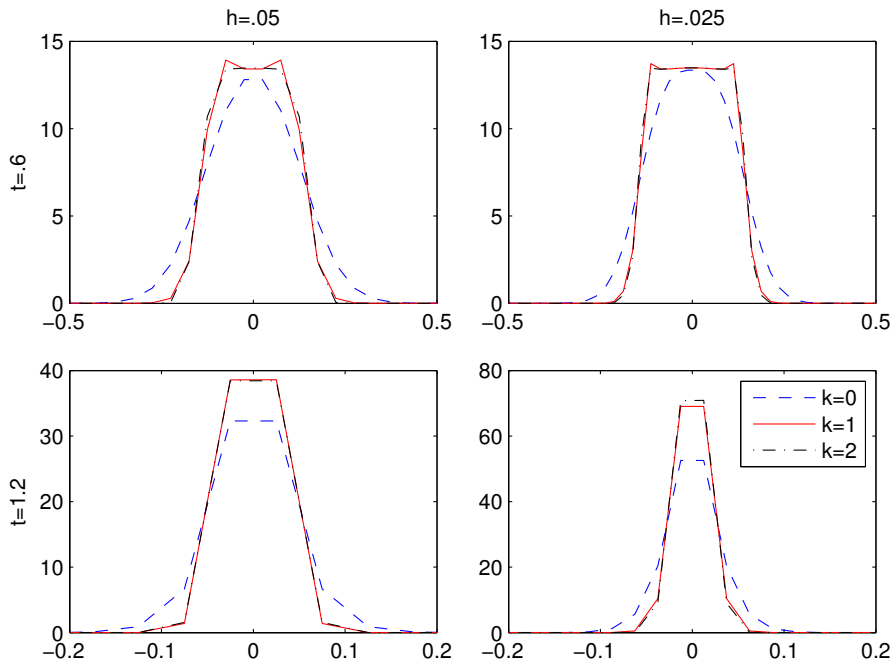

Figure 3.2: $\int f(t,x,v)dx$ at time $t = .6, 1.2$ and mesh size(in v) $h = .05, .025$ for DG schemes. Higher order scheme has better perference in terms of less diffusion and faster concentration.

different time. Indeed, high order schemes concentrate faster. If we refine the mesh size in v, we observe from the lower right graph that third order scheme performs better than the second order scheme.

3.4.2 "Local influence" and cluster formation

It is known that flocking is *not* guaranteed if the influence function is compactly supported, especially when (1.6) does not hold. Multiple clusters might form as time goes. This example is designed to compare the two asymptotic behaviors. In fact, our DG scheme captures both flocking and clusters very well. Let

$$f_0(x,v) = 5 * [\chi_{-5<x<-4.5} \cdot \chi_{.4<v<.5} + \chi_{4.5<x<5} \cdot \chi_{-.5<v<-.4}].$$

It represents two groups, where the left group is travelling to the right and the right group is travelling to the left. We consider two different influence functions:

$$\phi_1(r) = \left(1 - \frac{r}{5}\right)^2 \chi_{r<5}, \quad \phi_2(r) = \left(1 - \frac{r}{.5}\right)^2 \chi_{r<.5}.$$

Both functions are compactly supported. Yet ϕ_1 is much stronger than ϕ_2. In particular, $\phi_1(r) \geq \phi_2(r)$.

Figure 3.3 shows the evolution of the (C-S) model under two influence functions. We observe that with strong influence ϕ_1, the system converges to a flock. In contrast, with relatively weak influence ϕ_2, the interaction is not strong enough and multiple clusters are forming in large time.

3.4.3 Numerical rate of convergence

In this example, we test the rate of convergence of our DG method. We concentrate on the flocking part, *i.e.* equation (3.3). We set the same influence function $\phi(r) = (1+r)^{-1/2}$, and the following smooth initial density

$$f_0(x,v) = \begin{cases} \exp\left(-\frac{1}{.9-x^2-v^2}\right) & \text{if } x^2+v^2 < .9 \\ 0 & \text{otherwise.} \end{cases}$$

As there is no free transport, we set the computational domain $[-1,1]\times[-1,1]$. Fix the number of partitions on x to be 10. For v, we test on 10×2^s partitions, with $s = 0,1,\cdots,6$. To satisfy the CFL condition (3.10), we pick $\Delta t = .1\times 2^{-s}$ for second order scheme, and $\Delta t = .05\times 2^{-s}$ for third order scheme. Denote the corresponding numerical solution be $f^{[s]}$.

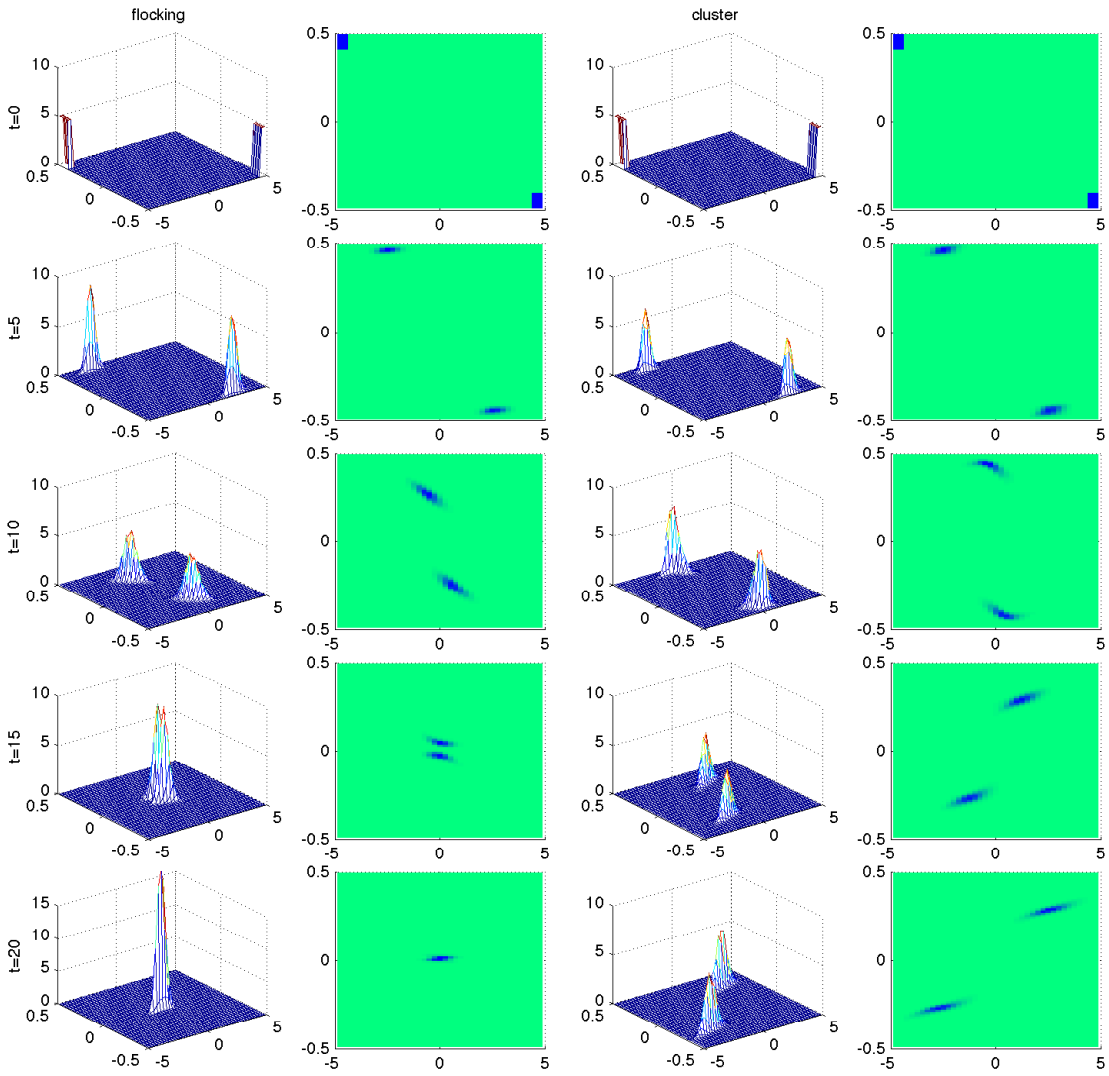

Figure 3.3: Flocking vesus cluster formation. With the same initial profile, strong inflence (left graphs) leads to flocking, and weak inflence (right graphs) leads to more clusters.

As this type of equations have no explicit solutions, we use $f^{[6]}$ as the reference solution. To concentrate on v variable, we integrate x and set $F^{[s]}(t,v) = \int_{-1}^{1} f^{[s]}(t,x,v)dx$. The L^1 error is computed as

$$e_s(t) = \left\| F^{[s]}(t,\cdot) - F^{[6]}(t,\cdot) \right\|_{L^1_v[-1,1]}, \quad s = 0,\cdots,5.$$

The following table shows the computational convergence rates

$$r_s = -\log_2(e_s/e_{s-1}), \quad s = 1,\cdots,5$$

for $t = 0,1,\cdots,8$. The numerical results validate the desired order of convergence of the corresponding schemes. Note that at time $t = 8$, the solution is already very concentrated in v. The DG schemes still have good performances.

3.4.4 Motsch-Tadmor vesus Cucker-Smale

The last example is designed to compare the two models (C-S) and (M-T).

A drawback of Cucker-Smale model has been pointed out in section 1.2.2, mentioning its poor performance in modeling the dynamics under far-from-equilibrium condition, as illustrated in figure 1.3. On the other hand, Motsch-Tadmor model overcomes the drawback and provides more realistic dynamics even when the data is far from equilibrium.

We setup up a numerical example to verify the argument. Our DG schemes have good performances on both setups. It captures the difference of the two models in kinetic level, which agrees with the discussion in section 1.2.2.

Consider the initial configuration as a combination of a small group (with mass .02) and a large flock (with mass .98) far away

$$f_0(x,v) = \chi_{|x|<.1}\chi_{|v|<.05} + .98\delta(x-5)\delta(v-1),$$

Second order scheme

t	0	1	2	3	4	5	6	7	8
r_1	1.8465	2.0820	2.0554	2.2063	1.9445	1.8082	1.4335	0.9248	0.6094
r_2	2.0595	2.3851	2.1841	1.9734	1.9677	1.9551	1.8305	1.8867	1.5480
r_3	2.0172	2.3027	2.4948	2.3888	2.3025	2.0338	2.0378	1.7847	1.7644
r_4	1.9983	2.0971	2.2556	2.4189	2.5530	2.3960	2.2283	2.2273	2.1043
r_5	2.0448	2.0639	2.0979	2.2186	2.2794	2.5724	2.5779	2.3025	2.2066

Third order scheme

t	0	1	2	3	4	5	6	7	8
r_1	2.6925	4.1556	2.7288	3.4275	2.3059	1.7800	2.6563	2.2884	1.1553
r_2	3.0736	3.0650	3.4810	3.5644	3.1477	3.3098	2.1924	1.7520	2.4830
r_3	3.0794	2.9025	2.8867	2.9380	3.3538	3.7456	3.4532	3.0015	2.0851
r_4	2.9998	3.1586	3.1191	2.9967	2.9232	2.8592	3.1041	3.5266	3.6594
r_5	2.9809	3.0599	3.0873	3.1422	3.1529	2.9013	2.9281	3.0585	3.4617

Table 3.1: Computational convergence rates for second and third order DG schemes at different times.

with compact supported influence function $\phi(r) = (1-r)^2\chi_{r<1}$. It is easy to check that the large flock never interact with the small group.

Figure 3.4 shows numerical results of the evolutions of the small group in both (C-S) and (M-T) setups. We observe that under (C-S) setup, the faraway large flock eliminates the interactions inside the small group. The evolution is almost like a pure transport. In contrast, (M-T) setup yields the reasonable flocking behavior for the small group.

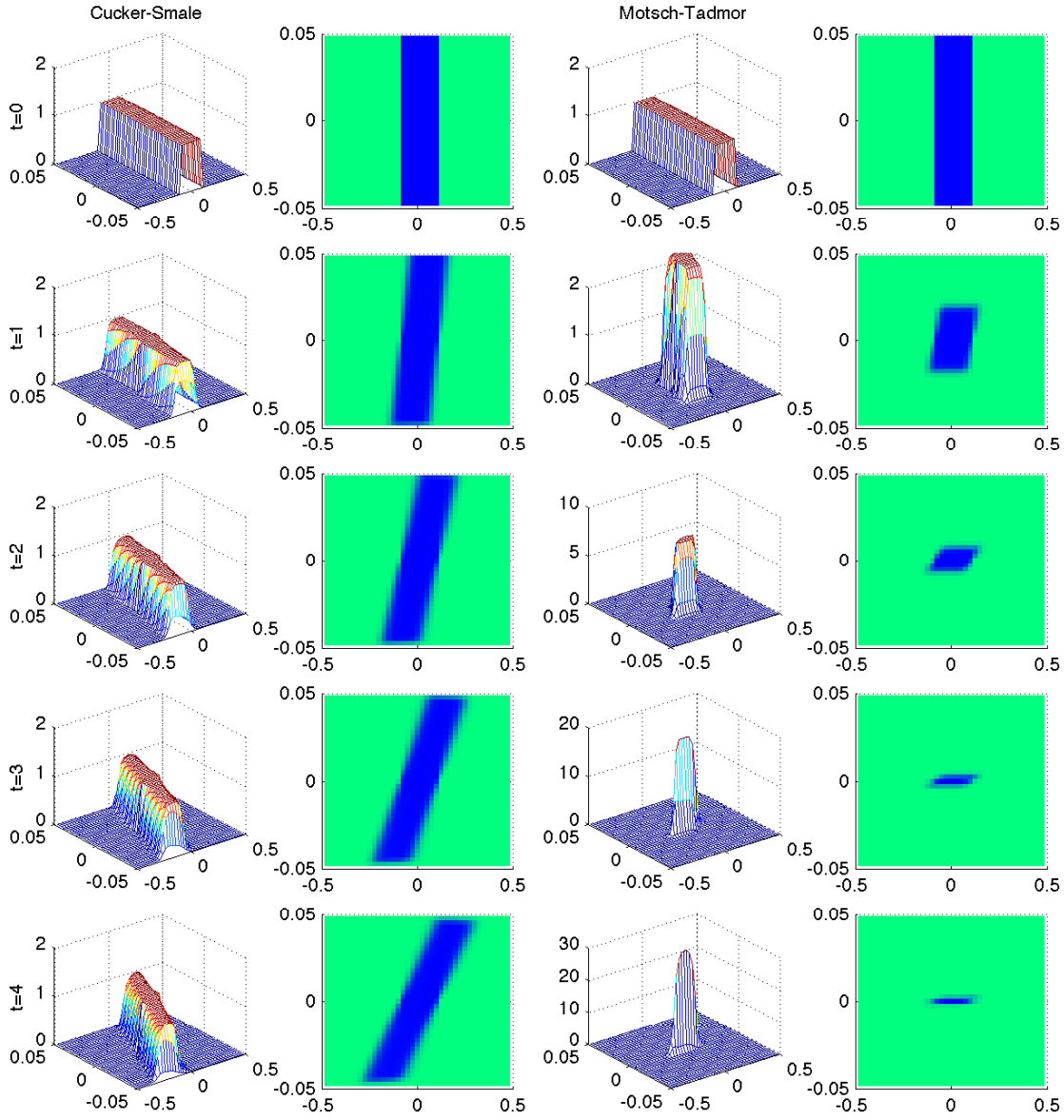

Figure 3.4: Evolution of the small group under 2 models, when there is a large group very far away. Under Cucker-Smale setup, the dynamics looks like pure transport. Under Motsch-Tadmor setup, flocking behavior is correctly captured.

Appendix

3.A Complete algorithm

In this section, we summarize the full procedure of our proposed SSP Runge-Kutta discontinuous Galerkin (RKDG) method on kinetic flocking models.

Initialization

Given an initial profile f_0, we need to use piecewise polynomials in v to approximate f_0. In particular, we compute

$$f_0{}_j^{(l)}(x) = \frac{1}{h^{l+1}} \int_{I_j} f_0(x,v) p_j^{(l)}(v) dv$$

for all x on the mesh grid, using numerical integration with desired accuracy.

Splitting and free transport

We split the full system into the free transport part (3.2) and the alignment part (3.3), using standard splitting method.

As for the free transport part, we apply WENO scheme on $f_j^{(l)}$ separately for all $l = 0, \cdots, k$. It is worth noting that we approximate f by piecewise polynomials in v direction. After a free transport in x direction, the resulting f is still a piecewise polynomial in v.

Flocking

The flocking part (3.3) is solved by RKDG method. We evolve $f_j^{(l)}$ using SSP scheme (discussed in section 3.2.5) to achieve time accuracy. In each step, we solve a forward Euler step FE. In detail, we first apply limiter (3.11) to $f_j^{(l)}$. Then, evolve $f_j^{(l)}$ by (3.8), with forward Euler time discretization, where Δt should satisfy the CFL condition (3.9), or simply (3.10).

Part II

Image Processing

Image processing is one of the fast growing technology worldwide, where mathematics enters beautifully in many subjects: denoising, deblurring, segmentation, inpainting, etc.

The starting point of this part is the famous Rudin-Osher-Fatemi (ROF) model for image denoising (section 4.1). The main idea is to decompose the image to two scales, where noise is extracted and distinguished from other textures of the image, *e.g.* edges. The two-scale decomposition of image can be extended to a multi-scale representation, through a hierarchical procedure (section 4.2).

The main problem we are dealing with has no direct relation to image processing. The goal is to find uniformly bounded solution to $divU = F$ for $F \in L^n(\mathbb{T}^n)$. A surprising fact is that such solution can not be constructed linearly, even if the equation is linear (section 5.1).

In chapter 5, we introduce a hierarchical construction of a bounded solution for the equation, which is motivated from multi-scale image representation. The solution, constructed through a highly nonlinear procedure, is proved to be uniformly bounded (section 5.2). In section 5.3, we propose a numerical implementation for the minimization problem in each hierarchical step. A duality argument is used to avoid difficulties in the minimization which involves both L^2 norm and L^∞ norm. The dual problem has an ROF substructure where a modified ROF solver could be used to solve the problem.

Main references for this part are: ROF model [84], multi-scale image representation [92] and hierarchical construction [91]. Chapter 5 is presented along the storyline of the work [94].

Chapter 4: Multi-scale image representation

Image processing has an explosive impact on modern sciences. With the help of advanced computer technologies, man can acquire images far beyond their eye visions, *e.g.*, satellite imaging and medical imaging like CT, MRI. As the acquisition procedures have some limitations and are not always perfect, images might be noisy, blurry or even incomplete. It promotes the study of reconstructing images for a degraded version in an accurate and stable way, which is the main goal of image processing.

Mathematically speaking, image processing can be considered as an inverse problem. The main challenge is the ill-posedness of the problem, where direct inversion leads to non-uniqueness or instability. Many elegant image processors are designed, depending on the type of errors in the acquisition procedure, and different prospectus of the reconstruction: denoising, deblurring, inpainting, segmentation, etc. We refer to classical image processing books [21, 39, 59] for various of examples and discussions.

In this chapter, we go over variational approaches for images processing and corresponding PDE interpretations. In section 4.1, we discuss the celebrated Rudin-Osher-Fatemi (ROF) model [84], which plays an important role in image denoising. In section 4.2, we introduce a multi-scale representation of images using hierarchical decomposition, followed from [92,93]. The decomposition successfully captures the structure of the

image under different levels, and it is constructive.

4.1 Variational decomposition of images

A grey-scaled digital image is represented by a distribution f, defined on a bounded regular image domain Ω. In general, f lies in the distribution space $\mathscr{D}'(\Omega)$, where $\mathscr{D}(\Omega)$ is the space of C^∞ test functions, and $\mathscr{D}'(\Omega)$ is the space which consists all linear functional on $\mathscr{D}(\Omega)$.

The space $\mathscr{D}'(\Omega)$ is very board. It contains many diversified images with various types of textures. Hence, most realistic images are contained in this space. However, it is hard to derive any exciting image features from this board space. To realize more structures, we shall go to smaller spaces.

4.1.1 L^p images vesus BV images

Most images lie in the space $L^p(\Omega)$, which is a subspace of $\mathscr{D}'(\Omega)$. Equipped with the norms

$$\|f\|_{L^p} = \left(\int_\Omega |f|^p dx\right)^{1/p},$$

L^p images permit many image structures, including

1. Edges: It corresponds to discontinuity (or jump) of f,

2. Noises: It usually corresponds to oscillations in f.

If we add some regularity, the resulting $W^{1,p}$ images are more specific: both edges and noises are forbidden, as they both drive $W^{1,p}$ norm to infinity.

The mission of image denoising is to remove the oscillations from the image, while maintain other key structures, including edges and other textures. Thus, we need to find a space between L^p and $W^{1,p}$ which permits edges but not noises.

The novel space of BV images, introduced in [83, 84], achieve the good balance between penalizing noises and respecting edges. The norm is defined as

$$\|f\|_{BV} := \|f\|_{L^1} + \mathsf{TV}[f],$$

where $\mathsf{TV}[f]$ is the total variation of f

$$\mathsf{TV}[f] := \sup\left\{ \int_\Omega f(x) div\, \varphi(x) dx \ : \ \varphi \in C_c^1(\Omega),\ \|\varphi\|_{L^\infty(\Omega)} \le 1 \right\}. \tag{4.1}$$

It is easy to check the piecewise smooth image belongs to $BV(\Omega)$. Hence edge is permitted. On the other hand, highly oscillatory f drives the total variation to be infinity.

4.1.2 Rudin-Osher-Fatemi model

The essence of image denoising is to find a way to extract the noise from the image. As we argued in the previous section, noise is captured in the space $L^p \backslash BV$. Hence, the total variation is a good measurement of the noise.

The Rudin-Osher-Fatemi (ROF) denoising model [84] carries the idea above. Given a noisy image f, one can decompose f into two parts: the recovered (de-noised) image u and the noise v, where v captures most of the total variation and u is close enough to f in terms of edges, textures, etc. The model can be expressed as the following minimization problem.

$$J(f,\lambda) = \min_{u+v=f} \left(\mathsf{TV}[u] + \lambda \|v\|_{L^2}^2 \right). \tag{4.2}$$

Here, small $\mathsf{TV}[u]$ indicates few noises for the recovered image, and small $\|v\|_{L^2}$ indicates the similarity between f and u. λ is a parameter which characterizes the balance between the two terms. In the two extreme cases:

1. If $\lambda = 0$, then $u = 0, v = f$. With no restriction on the similarity between f and u, the recovered image is simply flat.

2. If $\lambda = +\infty$, then $u = f, v = 0$. No noise is sensed without the measure of total variation. Thus, the recovered image is the same as the original.

Different choices of λ give decompositions under different scales. The recovered image depends on λ. So we denote u_λ and v_λ as the solution of the (4.2) in level λ.

$$u_\lambda = \underset{u \in BV}{arg\,min} \int_\Omega \left(|\nabla u| + \lambda |f - u|^2 \right), \quad v_\lambda = f - u_\lambda.$$

Theorem 4.1.1 (Existence and uniqueness)**.** *Assume that $f \in L^2(\Omega)$. Then for any $\lambda \geq 0$, there exists a unique pair (u_λ, v_λ) which minimizes* (4.2)*.*

Proof. See for instance [26, Theorem 4.14]. □

Proposition 4.1.2. *Suppose $f \in L^2(\Omega)$. (u_λ, v_λ) are minimizers of ROF model* (4.2)*.*

(a). u_λ is bounded in $L^2(\Omega) \cap BV(\Omega)$: $\|u_\lambda\|_{L^2} \leq 2\|f\|_{L^2}$ and $\mathsf{TV}[u_\lambda] \leq \lambda \|f\|_{L^2}^2$,

(b). v_λ has zero average: $\int_\Omega v_\lambda = 0$,

(c). (Meyer [71, Theorem 4]) Define $\|f\|_* := \inf\{\|g\|_{L^\infty} : f = div\, g\}$. If $\lambda \leq \dfrac{1}{2\|f\|_*}$, then $u_\lambda = 0$ and $v_\lambda = f$. If $\lambda \geq \dfrac{1}{2\|f\|_*}$, then $\|v_\lambda\|_* = \dfrac{1}{2\lambda}$ and $\displaystyle\int_\Omega u_\lambda v_\lambda = \frac{1}{2\lambda}\mathsf{TV}[u_\lambda]$.

Remark 4.1.1. Meyer's result addresses that if λ is small, then the recovered image is flat. To get a non-trivial u_λ, we have to pick λ large enough. Moreover, (u_λ, v_λ)

forms an extremal pair, in the sense of achieving an equality in the duality inequality $\left|\int_\Omega u_\lambda v_\lambda\right| \leq \mathsf{TV}[u_\lambda]\|v_\lambda\|_*$.

4.1.3 Euler-Lagrange equations

To solve ROF model (4.2), we derive the formal Euler-Lagrange equations, using the standard variational approach.

Suppose (u,v) is the minimizer of (4.2). We compute the energy $\tilde{J}$ for the perturbed pair $(u+\varepsilon h, v-\varepsilon h)$.

$$\begin{aligned}\tilde{J} =& \int_\Omega \left(|\nabla(u+\varepsilon h)| + \lambda(v-\varepsilon h)^2\right) = J + \varepsilon\int_\Omega \left(\frac{\nabla u}{|\nabla u|}\cdot\nabla h - 2\lambda v h\right) + o(\varepsilon)\\ =& J + \varepsilon\left[\int_\Omega -h\left(div\left(\frac{\nabla u}{|\nabla u|}\right) + 2\lambda v\right) + \int_{\partial\Omega}\frac{h}{|\nabla u|}\frac{\partial u}{\partial \mathbf{n}}\right] + o(\varepsilon).\end{aligned}$$

Note that the lower order term $o(\varepsilon) \geq 0$. To ensure $\tilde{J} \geq J$ for all $\varepsilon \in \mathbb{R}$, the minimizer u must satisfy the Euler-Lagrange equation

$$div\left(\frac{\nabla u}{|\nabla u|}\right) + 2\lambda(f-u) = 0, \quad \frac{\partial u}{\partial \mathbf{n}}|_{\partial\Omega} = 0.$$

The main difficulty in the computational prospectus is the degeneracy caused by the gradient term in the denominator of the elliptic operator. One way to overcome such degeneracy is to *regularize* the diffusivity coefficient $|\nabla u|^{-1}$ to $|\varepsilon + \nabla u|^{-1}$ with some small number $\varepsilon > 0$. We refer to [20] for more discussions about the regularization.

Figure 4.1 provides an example of ROF decomposition for a finger print. A Gauss-Seidel iteration is used to get the numerical result. With larger λ, more textures are resolved. Meanwhile, noises are inserted into u_λ as well for large λ. Therefore, an appropriate scale should be picked to balance the two effects.

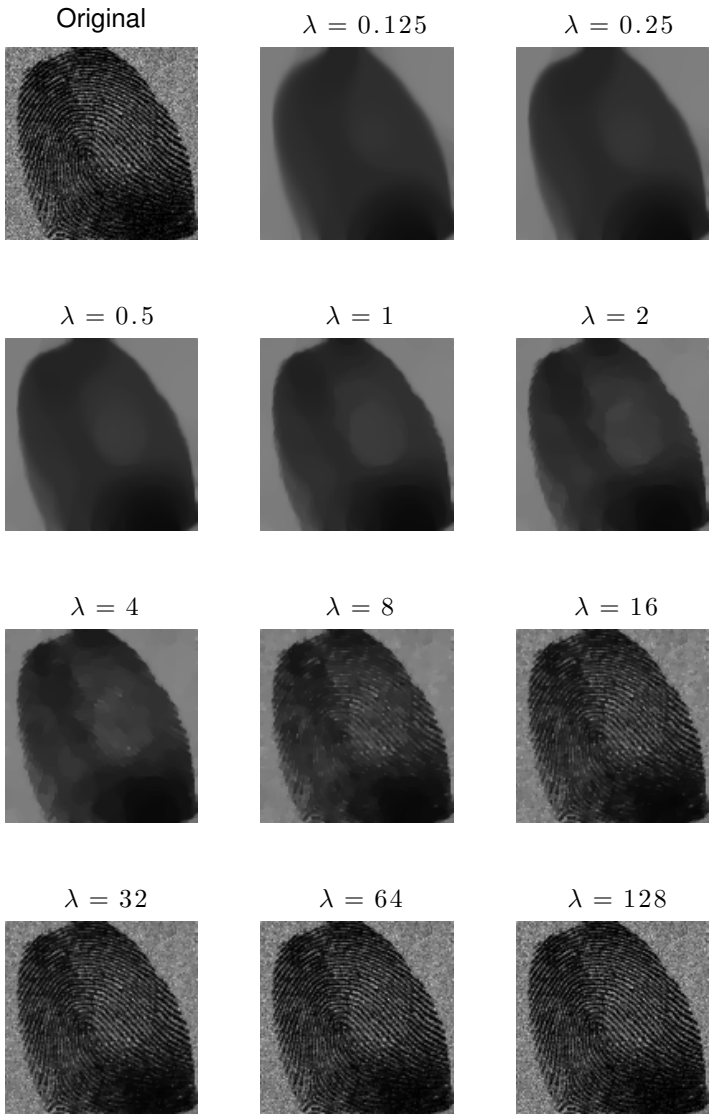

Figure 4.1: The recovered image u_λ using ROF denoising model under different λ

4.2 Multi-scale image representation using hierarchical decompositions

Tadmor, Nezzar and Vese in [92] proposed a hierarchical decomposition projecture for image denoising. They extend the ROF model from two-scale to multi-scale, where the noise is represented in different levels.

Recall ROF model (4.2), f is decomposed to u_λ and v_λ. The idea is to decompose v_λ using ROF model with a larger scale, *e.g.* 2λ,

$$v_\lambda = u_{2\lambda} + v_{2\lambda}, \quad \{u_{2\lambda}, v_{2\lambda}\} = \underset{u+v=v_\lambda}{arg\,min} \left(\mathsf{TV}[u] + 2\lambda \|v\|_{L^2}^2\right).$$

$u_{2\lambda}$ resolves more textures up to scale 2λ. Repeating the procedure, f can be decomposed into pieces consisting multi-scale representations. A hierarchical decomposition of f under dyadic refinement reads

$$\begin{aligned} f =& u_\lambda + v_\lambda = u_\lambda + u_{2\lambda} + v_{2\lambda} = \cdots = \sum_{j=1}^{k} u_{\lambda_j} + v_{\lambda_k} \qquad \lambda_j = 2^{j-1}\lambda, \qquad (4.3) \\ v_{\lambda_j} =& u_{\lambda_{j+1}} + v_{\lambda_{j+1}}, \quad \{u_{\lambda_{j+1}}, v_{\lambda_{j+1}}\} = \underset{u+v=v_{\lambda_j}}{arg\,min} \left(\mathsf{TV}[u] + \lambda_{j+1}\|v\|_{L^2}^2\right). \end{aligned}$$

Due to proposition 4.1.2(c), we take $\lambda > \dfrac{1}{2\|f\|_*}$ to avoid trivial solutions.

The following theorem provides the convergence of the hierarchical decomposition $f = \sum_{j=1}^{\infty} u_{\lambda_j}$. See [92, Theorem 2.1, 2.2] for proofs.

Theorem 4.2.1 ([92]). *For $f \in L^2$, $\|v_{\lambda_k}\|_* = \dfrac{1}{2\lambda_k} \downarrow 0$, as $k \to \infty$. Moreover, if $f \in BV$, then $\|v_{\lambda_k}\|_{L^2} \downarrow 0$ as $k \to \infty$.*

Remark 4.2.1. The infinite sum converges strongly in $\|\cdot\|_*$ for general L^2 image. Note that the norm captures the strength of oscillations. Therefore, u_{λ_k} represents oscillations

scaled in $\left[\frac{1}{2\lambda_k}, \frac{1}{\lambda_k}\right]$. To get strong L^2 convergence, we need to assume f to be more regular. The main reason is that the construction of the hierarchical decomposition is highly *nonlinear*. Hence, additional energy estimate is required to pass to the limit, where higher regularity (*e.g.* BV) is necessary. The nonlinear nature of the decomposition motivates the construction of a bounded solution for $div\,\mathbf{u} = F$, as no linear construction is available for such solution. See chapter 5 for related discussions.

Figure 4.2 shows an example of hierarchical decomposition of a fingerprint. In each hierarchical step, more textures are resolved. We pick $\lambda_1 = 0.125$. u_{λ_1} is blurry as oscillations has scale smaller than $\lambda_1^{-1} = 8$ are all considered as noises, leaving only the rough structures and main edges. In each successive step, we resolve more textures by including more scales of the original image f to the recovered image $\sum_{j=1}^{k} u_{\lambda_j}$. For k large, we know from theorem 4.2.1 that the recovered image is close to the original image, as some "noise" are interpreted as textures due to the smallness of λ_k^{-1}. Hence, for the purpose of denoising, one has to stop at some stage. The stopping criterion is subject to the type of noise as well as textures.

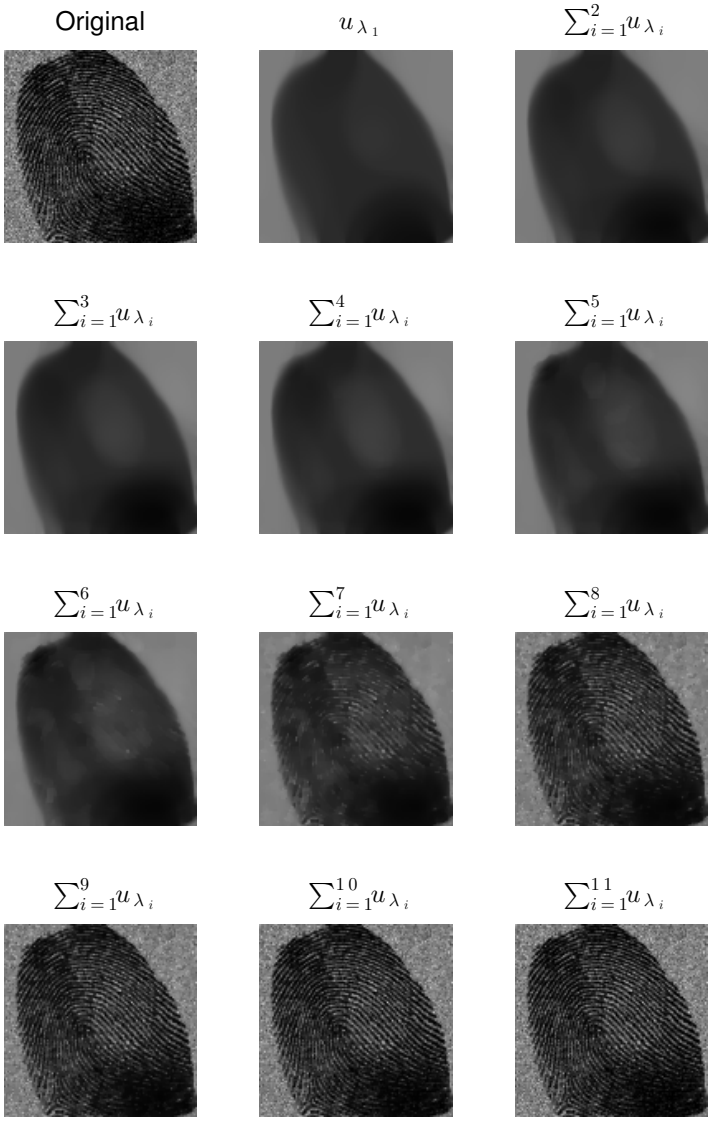

Figure 4.2: Multi-scale images under hierarchical decomposition. Here, $\lambda_j = 2^{j-4}$. More textures (as well as noises) are resolved with more hierarchical steps

Chapter 5: Hierarchical construction of bounded solutions of $div\, U = F$

In this chapter, we discuss the construction of bounded solution of the linear equation $div\, U = F$, where F lies in the critical space $L^n(\mathbb{T}^n)$.

The existence of uniformly bounded solutions follows from the closed range theorem together with Gagliardo-Nirenberg inequality, [11]. Moreover, Bourgain and Brezis [11] proved that any mapping, $F \in L^n \mapsto U \in L^\infty(\mathbb{T}^n)$, must be *nonlinear*: thus, the intriguing aspect here is that although the equation is linear, the construction of its uniformly bounded solutions for L^n-data is not.

In particular, the favorite linear Helmholtz solution $U_{\text{Hel}} = \nabla\Delta^{-1}F \in W^{1,n}$ is not guaranteed to be uniformly bounded, as the Sobolev imbedding $W^{1,n}(\mathbb{T}^n) \not\subset L^\infty(\mathbb{T}^n)$ fails in the critical borderline case, see *e.g.* [34, pp.280].

Inspired by the hierarchical decompositions in image processing, Tadmor [91] utilized such decompositions to construct a bounded solution for the equation.

The constructive procedure is highly nonlinear. It brings challenges for the numerical implementation. We propose a highly non-trivial numerical approach in [94] for the 2D case, through a duality argument. It successfully captures the key structure of the

constructed solution: uniform boundedness.

We organize the chapter as follows. First, we discuss the existence theory of bounded solution in section 5.1. In particular, the solution can not be constructed linearly. In section 5.2, we propose a hierarchical construction based on the idea from image processing. A duality argument is presented in section 5.3, which simply the minimization problem in order to implement numerically. The numerical algorithm is discussed in section 5.4, followed by an example, showing that the hierarchical solution is bounded while Helmholtz solution is not (section 5.5). We end up this chapter with a brief discussion on potential applications, which leads to future research.

5.1 Bounded solution for $divU = F$

We are concerned with the uniformly bounded solutions, $U \in L^\infty(\mathbb{T}^n, \mathbb{R}^n)$ of the linear equation

$$divU = F, \quad F \in L^n_\#(\mathbb{T}^n), \tag{5.1}$$

where $L^n_\#(\mathbb{T}^n)$ is the space of L^n integrable functions over the n-dimensional torus $\mathbb{T}^n$ with zero mean.

This system has been studied in [11, 12, 28, 69, 91, 94] and many more literatures. We refer a recent review [85] for a full discusion.

5.1.1 Existence of bounded solution

The existence of bounded solutions for (5.1) has been proved by Bourgain and Brezis in [11]. Here, we give some details for completeness.

Consider operator $A:\ D(A)\subset L^\infty \to L^n_\#$ be a densely defined and closed linear operator, such that

$$AU = divU, \quad \text{where} \quad D(A) = \{U \in L^\infty \ :\ divU \in L^n_\#\}.$$

Then, the adjoint operator $A^* : D(A^*) \subset L^{\frac{n}{n-1}}_\# \to \mathscr{M}$, where $\mathscr{M}$ is the space of measures, which is the dual space of L^∞. By duality,

$$A^*u = -\nabla u, \quad \text{and then} \quad D(A^*) = \{u \in L^{\frac{n}{n-1}}_\# \ :\ A^*u \in \mathscr{M}\} = L^{\frac{n}{n-1}}_\# \cap BV.$$

An Gagliardo-Nirenberg-Sobolev imbedding states $BV \subset L^{\frac{n}{n-1}}_\#$, namely

$$\|u - \fint u\| \lesssim \|\nabla u\|_{\mathscr{M}}, \quad \forall\ u \in BV.$$

It implies that the null space of A^* $\mathsf{Null}(A) = 0$.

Applying closed range theorem, we get $\mathsf{Range}(A) = \mathsf{Null}(A^*)^\perp = L^n_\#$. Therefore, for all $F \in L^n_\#$, there exists a $U \in D(A) \subset L^\infty$ which solves (5.1).

Remark 5.1.1. The proof of existence of bounded solutions is not constructive. In fact, the construction of such solution is non-trivial, despite the easy linear expression of the equation. We will address this difficulty in section 5.1.3.

5.1.2 Helmholtz solution is not bounded

One classical solution of (5.1) is the Helmholtz solution

$$U_{\text{Hel}} = \nabla\Delta^{-1}F. \tag{5.2}$$

For $F \in L^n_\#(\mathbb{T}^n)$, it is clear that $U_{\text{Hel}} \in W^{1,n}(\mathbb{T}^n)$. But since $W^{1,n}(\mathbb{T}^n)$ is not a subset of $L^\infty(\mathbb{T}^n)$, Helmholtz solution need not to be uniformly bounded.

The following concrete counterexample due to L. Nirenberg, [11, Remark 7], demonstrates this type of unboundedness: let

$$F = \Delta v, \qquad v(\mathbf{x}) := x_1 |\log r|^\theta \zeta(r), \quad r = |\mathbf{x}|, \tag{5.3}$$

where $\theta \in (0, \frac{n-1}{n})$, and $\zeta(r)$ is a smooth cut-off function supported near the origin with $\zeta(r) = 1$ for $r \in [0, \varepsilon]$ and $\zeta(r) = 0$ for $r \geq 2\varepsilon$, where ε is a fixed constant.

In this case, $F \in L^n_\#(\mathbb{T}^n)$, but the Helmholtz solution, $U_{\text{Hel}} = \nabla \Delta^{-1} F = \nabla v$, has a fractional logarithmic growth at the origin. Hence, Helmholtz solution (5.2) is not good candidate of bounded solution.

We first check $F \in L^n_\#$. Take $g(r) = |\log r|^\theta \zeta(r)$. Compute

$$\begin{aligned}
\partial_{x_j} v &= x_1 g'(r) \frac{x_j}{r} + g(r) \delta_{1j}, \\
\partial^2_{x_j x_j} v &= x_1 \left(\frac{g'(r)}{r} \right)' \frac{x_j^2}{r} + x_1 \frac{g'(r)}{r} + 2 g'(r) \frac{x_1}{r} \delta_{1j}, \\
F = \Delta v &= x_1 \left[g''(r) + (n+1) \frac{g'(r)}{r} \right].
\end{aligned}$$

Note that x_1 is bounded by 2ε as ζ is supported near origin. And

$$\left| g''(r) + (n+1) \frac{g'(r)}{r} \right| \lesssim \frac{|\log r|^{\theta - 1}}{r}, \quad r \leq \varepsilon.$$

It is clearly bounded for $r > \varepsilon$. The only singularity is at 0, and it is L^n-integrable. In fact,

$$\int_{\mathbb{R}^n} |F(\mathbf{x})|^n d\mathbf{x} \lesssim \int_0^\varepsilon \left(\frac{|\log r|^{\theta-1}}{r} \right)^n r^{n-1} dr + C \lesssim \lim_{r \to 0} |\log r|^{(\theta-1)n+1} + C'.$$

As we pick $\theta < \frac{n-1}{n}$, the power $(\theta - 1)n + 1 < 0$, and the limit goes to zero, which implies boundedness of $\|F\|_{L^n}$. The mean zero property is true due to anti-symmetry with respect to the first component.

On the other hand, $U_{\text{Hel}} = \nabla v$ in not in L^∞. To see this, we compute the first component at origin

$$u_1(0) = \lim_{\mathbf{x}\to 0}\left[g'(r)\frac{x_1^2}{r} + g(r)\right].$$

It is clear that the limit of the first part is zero. However, the second part has fractional logarithmic growth as $g(r) = |\log r|^\theta$ near origin as long as $\theta > 0$. Therefore, u_1 diverges at the origin and $U_{\text{Hel}} \notin L^\infty$.

5.1.3 No linear construction of bounded solution

In this section, we state a surprising argument: *the construction of bounded solution for* (5.1) *can not be linear*. It implies the fact that all solutions via a linear operator, including the Helmholtz solution, are not bounded in $L^\infty \to L^n_\#$. We have to construct the uniformly bounded solution through a nonlinear procedure. This statement is also proved in [11].

Suppose by contradiction, there exists a bounded linear operator $\mathscr{L} : L^n_\# \to L^\infty$ such that $div\,\mathscr{L} = I$. We define $\bar{\mathscr{L}}$ be the average among all translations

$$\bar{\mathscr{L}} = \int_{\mathbb{R}^n} \tau_{-\mathbf{x}}\mathscr{L}\tau_{\mathbf{x}}d\mathbf{x},$$

where $\tau_{\mathbf{x}}$ is the usual translation. Then, $\bar{\mathscr{L}}$ belongs to $\mathscr{M}^{n,\infty}(\mathbb{T}^n)$, where the space $\mathscr{M}^{p,q}(\mathbb{T}^n)$ denotes the set of all bounded linear operators from $L^p(\mathbb{T}^n)$ to $L^q(\mathbb{T}^n)$ that commute with translations.

Note that $\mathscr{M}^{n,\infty}(\mathbb{T}^n) = \mathscr{M}^{1,\frac{n}{n-1}}(\mathbb{T}^n)$, see *e.g.* [41, Theorem 2.5.7]. Therefore, if we write $\tilde{\Gamma}$ as a Fourier multiplier $\widehat{(\bar{\mathscr{L}}f)}_{\mathbf{k}} = \mathbf{m}(\mathbf{k})\hat{f}_{\mathbf{k}}$ for $\mathbf{k} \in \mathbb{Z}^n$, the symbol $\mathbf{m} = (m_1, \cdots, m_n)$

satisfies

$$\sum_{\mathbf{k}\in\mathbb{Z}^n\setminus\{0\}} \frac{|m_j(\mathbf{k})|^2}{|\mathbf{k}|^{n-2}} < \infty, \quad \forall\ j=1,\cdots,n. \quad \Rightarrow \quad \sum_{\mathbf{k}\in\mathbb{Z}^n\setminus\{0\}} \frac{|\mathbf{m}(\mathbf{k})|^2}{|\mathbf{k}|^{n-2}} < \infty. \tag{5.4}$$

On the other hand, we have $div\,\bar{\mathscr{L}} = I$. So $\mathbf{m}$ also satisfies

$$\sum_{j=1}^{n} k_j m_j(\mathbf{k}) = 1, \quad \forall\ \mathbf{k}\in\mathbb{Z}^n.$$

A Cauchy-Schwarz inequality implies $|\mathbf{m}(\mathbf{k})|^2 \geq |\mathbf{k}|^{-2}$. Plug this into (5.4), we get

$$\sum_{\mathbf{k}\in\mathbb{Z}^n\setminus\{0\}} \frac{1}{|\mathbf{k}|^n} < \infty,$$

which leads to a contradiction.

5.2 Hierarchical of construction bounded solution

In this section, we construct a bounded solution for (5.1). We shall set the dimension $n=2$ for now on for simplicity. Similar results can be derived in higher dimension with no additional difficulty, consulting [91].

As we know from the previous section, the construction can not be linear. In particular, Helmholtz solution (5.2) fails to be uniformly bounded.

However, if F lies in a smaller space $BV_{\#}(\mathbb{T}^2)$, then Helmholtz solution is uniformly bounded.

Proposition 5.2.1. *If $F\in BV_{\#}(\mathbb{T}^2)$, then $U_{Hel}\in L^{\infty}(\mathbb{T}^2)$.*

Proof. F is BV-bounded and hence, [24, 98], $F\in BV\subset L^{2,1}$. On the other hand, $U_{\mathrm{Hel}} = \nabla\Delta^{-1}F = K*F$ where $K\in L^{2,\infty}$. By Young's inequality for Lorentz spaces, [76, 98], U_{Hel} is uniformly bounded. □

It implies that U_{Hel} fails only for $F \in L^2 \backslash BV$, which is the space containing noises in the context of image processing, as discussed in section 4.1.1.

Inspired by the hierarchical decompositions (4.3) in the context of image processing, Tadmor [91] utilized such decompositions as a constructive procedure to solve (5.1): the solution is given in terms of *hierarchical decomposition*, $U_{\text{Bdd}} = \sum \mathbf{u}_j$, where the $\{\mathbf{u}_j\}$'s can be computed recursively as the following minimizers,

$$\mathbf{u}_{j+1} = \underset{\mathbf{u}}{arg\,min} \left\{ \|\mathbf{u}\|_{L^\infty} + \lambda_1 2^j \|F - div\,(\sum_{k=1}^{j} \mathbf{u}_k) - div\,\mathbf{u}\|_{L^2}^2 \right\}, \quad j = 0, 1, \cdots. \tag{5.5}$$

Here, λ_1 is any sufficiently large parameter, $\lambda_1 > \dfrac{1}{2\mathsf{TV}[F]}$, which guarantees that the hierarchical decomposition starts with a non-trivial solution of (5.5), consult (5.21) below.

Our starting point for the construction of a uniformly bounded solution of (5.1), $\mathbf{x} \in L^\infty(\mathbb{T}^2, \mathbb{R}^2)$, is a decomposition of F,

$$F = div\,\mathbf{u}_1 + r_1, \qquad F \in L^2_\#(\mathbb{T}^2) := \big\{ g \in L^2(\mathbb{T}^2) \;\big|\; \int_{\mathbb{T}^2} g(x)dx = 0 \big\}, \tag{5.6a}$$

where $[\mathbf{u}_1, r_1]$ is a minimizing pair of the functional,

$$[\mathbf{u}_1, r_1] = \underset{div\,\mathbf{u} + r = F}{arg\,min} \Big\{ \|\mathbf{u}\|_{L^\infty} + \lambda_1 \|r\|_{L^2}^2 \Big\}. \tag{5.6b}$$

Here, λ_1 is a fixed parameter at our disposal where we distinguish between two cases, consult (5.21) below. If $\lambda_1 \le \dfrac{1}{2\mathsf{TV}[F]}$, then the minimizer of (5.6b) is the trivial one, $\mathbf{u}_1 \equiv 0, r_1 = F$; otherwise, by choosing λ_1 large enough, $\lambda_1 > \dfrac{1}{2\mathsf{TV}[F]}$, then (5.6b) admits a non-trivial minimizer, $[\mathbf{u}_1, r_1]$, which is characterized by a residual satisfying $\mathsf{TV}[r_1] = \dfrac{1}{2\lambda_1}$. By Gagliardo-Nirenberg iso-perimetric inequality, e.g., [109, §2.7], there exists $\beta > 0$ such that

$$\|g\|_{L^2} \le \beta \mathsf{TV}[g], \qquad \int_{\mathbb{T}^2} g(x)dx = 0. \tag{5.7}$$

It follows that r_1 is L^2-bounded:

$$\|r_1\|_{L^2} \leq \beta \mathsf{TV}[r_1] = \frac{\beta}{2\lambda_1}. \tag{5.8}$$

Moreover, since F has a zero mean so does the residual r_1. We conclude that the residual $r_1 \in L^2_{\#}(\mathbb{T}^2)$, and we can therefore implement the same variational decomposition of F in (5.6), and use it to decompose r_1. To this end, we use the same variational statement, $\Big\{\|\mathbf{u}\|_{L^\infty} + \lambda_2 \|r\|^2_{L^2}\Big\}$, with a new parameter, $\lambda = \lambda_2 > \lambda_1$,

$$r_1 = div\,\mathbf{u}_2 + r_2, \qquad [\mathbf{u}_2, r_2] = \underset{div\,\mathbf{u}+r=r_1}{arg\,min} \Big\{\|\mathbf{u}\|_{L^\infty} + \lambda_2 \|r\|^2_{L^2}\Big\}. \tag{5.9}$$

Borrowing the terminology from our earlier work on image processing [92, 93] (discussed in section 4.2), the decomposition (5.9) has the effect of "zooming" on the residual r_1, and it is here that we use the refined scale $\lambda_2 > \lambda_1$. Combining (5.9) with (5.6a) we obtain $F = div\,\mathbf{x}_2 + r_2$ with $\mathbf{x}_2 := \mathbf{u}_1 + \mathbf{u}_2$, which is viewed as an improved *approximate solution* of (5.1). Indeed, the "zooming" effect $\lambda_2 > \lambda_1$ implies that $\mathbf{x}_2$ has a smaller residual $\mathsf{TV}[r_2] = 1/(2\lambda_2)$ compared with $\mathsf{TV}[r_1] = 1/(2\lambda_1)$ in (5.8). In particular,

$$\|r_2\|_{L^2} \leq \beta \mathsf{TV}[r_2] = \frac{\beta}{2\lambda_2}.$$

This process can be repeated: if $r_j \in L^2_{\#}(\mathbb{T}^2)$ is the residual at step j, then we decompose it

$$r_j = div\,\mathbf{u}_{j+1} + r_{j+1}, \tag{5.10a}$$

where $[\mathbf{u}_{j+1}, r_{j+1}]$ is a minimizing pair of

$$[\mathbf{u}_{j+1}, r_{j+1}] = \underset{div\,\mathbf{u}+r=r_j}{arg\,min} \Big\{\|\mathbf{u}\|_{L^\infty} + \lambda_{j+1} \|r\|^2_{L^2}\Big\}, \qquad j = 0, 1, \ldots. \tag{5.10b}$$

For $j = 0$, the decomposition (5.10) is interpreted as (5.6a) by setting $r_0 := F$. Note that the recursive decomposition (5.10a) depends on the invariance that the residuals $r_j \in L^2_\#(\mathbb{T}^2)$: indeed, if r_j has a zero mean then so does r_{j+1}, and since by (5.21) the minimizer r_{j+1} has a bounded variation, $r_{j+1} \in L^2_\#(\mathbb{T}^2)$. The iterative process depends on a sequence of increasing scales, $\lambda_1 < \lambda_2 < \ldots \lambda_{j+1}$, which are yet to be determined.

The telescoping sum of the first k steps in (5.10a) yields an improved approximate solution, $\mathbf{x}_k := \sum_{j=1}^k \mathbf{u}_j$:

$$F = div\,\mathbf{x}_k + r_k, \qquad \|r_k\|_{L^2} \le \beta \mathsf{TV}[r_k] = \frac{\beta}{2\lambda_k} \downarrow 0, \qquad k = 1,2,\ldots. \tag{5.11}$$

The key question is whether the $\mathbf{x}_k$'s remain uniformly bounded, and it is here that we use the freedom in choosing the scaling parameters λ_k: comparing the minimizing pair $[\mathbf{u}_{j+1}, r_{j+1}]$ of (5.10b) with the trivial pair $[\mathbf{u} \equiv 0, r_j]$, we find

$$\begin{aligned} \|\mathbf{u}_{j+1}\|_{L^\infty} + \lambda_{j+1}\|r_{j+1}\|^2_{L^2} &\le \|0\|_{L^\infty} + \lambda_{j+1}\|r_j\|^2_{L^2}, \\ r_j = div\,\mathbf{u}_{j+1} + r_{j+1} &= div\,(0) + r_j. \end{aligned}$$

It remains to upper-bound the energy norm of the r_j's: for $j = 0$ we have $r_0 = F$; for $j > 0$, (5.11) implies that $\|r_j\|_{L^2} \le \beta/(2\lambda_j)$. We end up with

$$\|\mathbf{u}_{j+1}\|_{L^\infty} + \lambda_{j+1}\|r_{j+1}\|^2_{L^2} \le \lambda_{j+1}\|r_j\|^2_{L^2} \le \begin{cases} \lambda_1 \|F\|^2_{L^2}, & j = 0, \\ \\ \dfrac{\beta^2 \lambda_{j+1}}{4\lambda_j^2}, & j = 1,2,\ldots. \end{cases} \tag{5.12}$$

We conclude that by choosing a sufficiently fast increasing λ_j's such that $\sum_j \lambda_{j+1}\lambda_j^{-2} < \infty$, then the approximate solutions $\mathbf{x}_k = \sum_{j=1}^k \mathbf{u}_j$ form a Cauchy sequence in L^∞ whose limit, $\mathbf{x} = \sum_{j=1}^\infty \mathbf{u}_j$, satisfies the following.

Theorem 5.2.2 ([91]). *Fix β such that* (5.7) *holds. Then, for any given $F \in L^2_{\#}(\mathbb{T}^2)$, there exists a uniformly bounded solution of* (5.1)*,*

$$div\,\mathbf{x} = F, \qquad \|\mathbf{x}\|_{L^\infty} \leq 2\beta \|F\|_{L^2}.$$

The solution $\mathbf{x}$ *is given by* $\mathbf{x} = \sum_{j=1}^{\infty} \mathbf{u}_j$*, where the* $\{\mathbf{u}_j\}$*'s are constructed recursively as minimizers of*

$$[\mathbf{u}_{j+1}, r_{j+1}] = \underset{div\,\mathbf{u}+r=r_j}{arg\,min} \left\{ \|\mathbf{u}\|_{L^\infty} + \lambda_1 2^j \|r\|^2_{L^2} \right\}, \qquad r_0 := F, \quad \lambda_1 = \frac{\beta}{\|F\|_{L^2}}. \tag{5.13}$$

Proof. Set $\lambda_j = \lambda_1 2^{j-1}, j = 1,2,\ldots$, then, $\|\mathbf{x}_k - \mathbf{x}_\ell\|_{L^\infty} \lesssim 2^{-k},\ k > \ell \gg 1$. Let $\mathbf{x}$ be the limit of the Cauchy sequence $\{\mathbf{x}_k\}$ then $\|\mathbf{x}_j - \mathbf{x}\|_{L^\infty} + \|div\,\mathbf{x}_j - F\|_{L^2} \lesssim 2^{-j} \to 0$, and since div has a closed graph on its domain $D := \{\mathbf{u} \in L^\infty :\ div\,\mathbf{u} \in L^2(\mathbb{T}^2)\}$, it follows that $div\,\mathbf{x} = F$. By (5.12) we have

$$\|\mathbf{x}\|_{L^\infty} \leq \sum_{j=1}^{\infty} \|\mathbf{u}_j\|_{L^\infty} \leq \lambda_1 \|F\|^2_{L^2} + \frac{\beta^2}{4\lambda_1} \sum_{j=2}^{\infty} \frac{1}{2^{j-3}} = \lambda_1 \|F\|^2_{L^2} + \frac{\beta^2}{\lambda_1}.$$

Here $\lambda_1 > \dfrac{1}{2\mathsf{TV}[F]}$ is a free parameter at our disposal: we choose $\lambda_1 := \beta / \|F\|_{L^2}$ which by (5.7) is admissible, $\lambda_1 = \dfrac{\beta}{\|F\|_{L^2}} > \dfrac{1}{2\mathsf{TV}[F]}$, and the result follows. □

Remark 5.2.1. [Energy decomposition] By squaring the refinement step (5.6a), $r_j = r_{j+1} + div\,\mathbf{u}_{j+1}$, and using the characterization of $[\mathbf{u}_{j+1}, r_{j+1}]$ as an *extremal pair* (consult remark 5.3.2 below), we find

$$\|r_j\|^2_{L^2} - \|r_{j+1}\|^2_{L^2} = 2(r_{j+1}, div\,\mathbf{u}_{j+1}) + \|div\,\mathbf{u}_{j+1}\|^2_{L^2} = \frac{1}{\lambda_{j+1}} \|\mathbf{u}_{j+1}\|_{L^\infty} + \|div\,\mathbf{u}_{j+1}\|^2_{L^2}.$$

A telescoping sum of the last equality yields the "energy decomposition"

$$\sum_{j=1}^{\infty} \frac{1}{\lambda_j} \|\mathbf{u}_j\|_{L^\infty} + \sum_{j=1}^{\infty} \|div\,\mathbf{u}_j\|^2_{L^2(\mathbb{T}^2)} = \|F\|^2_{L^2(\mathbb{T}^2)} \tag{5.14}$$

Remark 5.2.2. We note that the constructive proof of theorem 5.2.2 does not assume the existence of bounded solution for (5.15): it is deduced from the Gagliardo-Nirenberg inequality (5.7). The hierarchical construction of solutions for $\mathscr{L}\mathbf{x} = F$, in the general setup of linear closed operators, $\mathscr{L} : \mathbb{B} \mapsto L^p_{\#},\ 1 < p < \infty$, with bounded invertible duals $\mathscr{L}^*$, was proved in [91].

In [12], Bourgain and Brezis proved that (5.1) admits a bounded solution in the *smaller* space, $\mathbb{B} = L^\infty \cap H^1$. This requires a considerably more delicate argument, which could be justified by the refined dual estimate (compared with (5.7)), $\|g\|_{L^2(\mathbb{T}^2)} \lesssim \|\nabla g\|_{L^1+H^{-1}(\mathbb{T}^2)}$. The proof of [12] is *not constructive*: it is based on an intricate Littlewood-Paley decomposition, which cannot be readily implemented in actual computations.

5.3 Construction of hierarchical minimizers

In this section, we study the minimization step of the hierarchical decompositions (5.5). It is challenge to implement the minimization problem as it involves both L^2 and L^∞ norms. We proceed with a duality argument.

5.3.1 The minimization problem and its dual

We rewrite each minimization step of the hierarchical decompositions (5.5) in the following form,

$$\bar{\mathbf{u}} = \underset{\mathbf{u}:\mathbb{T}^2\to\mathbb{R}^2}{arg\,min} \left\{ \|\mathbf{u}\|_{L^\infty} + \lambda \|f - div\,\mathbf{u}\|^2_{L^2} \right\}, \qquad \|\mathbf{u}\|_{L^\infty} := \operatorname*{ess\,sup}_{x,y} \sqrt{u_1^2 + u_2^2}. \tag{5.15}$$

Here, f is an L^2 function with zero mean which stands for $F - div\,(\sum_{k=1}^{j} \mathbf{u}_k)$ in (5.5), and λ stands for the dyadic scales, $\lambda_1 2^j, j = 0, 1, \cdots$.

To circumvent the difficulty of handling the L^∞ norm in (5.15), we concentrate on the dual problem associated with (5.15). We let $\mathscr{N}(\mathbf{u}) = \|\mathbf{u}\|_{L^\infty} : V \mapsto \bar{\mathbb{R}}$, $\mathscr{E}(p) = \|f-p\|_{L^2}^2 : Y \mapsto \bar{\mathbb{R}}$, and $\Lambda = div : V \mapsto Y$ with $V = L^\infty(\mathbb{T}^2)$ and $Y = L^2(\mathbb{T}^2)$. By duality theorem, [32, §3,Remark 4.2], the variational problem (5.15),

$$(\mathscr{P}): \qquad \inf_{u\in V}[\mathscr{N}(u)+\mathscr{E}(\Lambda u)]$$

is equivalent to its dual problem

$$(\mathscr{P}^*): \qquad \sup_{p^*\in Y^*}[-\mathscr{N}^*(\Lambda^* p^*)-\mathscr{E}^*(-p^*)];$$

moreover, if $\bar{u}$ and $\bar{p}^*$ are solutions of $(\mathscr{P})$ and $(\mathscr{P}^*)$ respectively, then $\Lambda^*\bar{p}^* \in \partial\mathscr{N}(\bar{u})$, and $-\bar{p}^* \in \partial\mathscr{E}(\Lambda\bar{u})$. Here, $\mathscr{N}^*, \mathscr{E}^*$ are conjugate functions of $\mathscr{N}, \mathscr{E}$, expressed in terms of the usual L^2 pairing $\langle \mathbf{w}_1, \mathbf{w}_2\rangle := \displaystyle\int_{\mathbb{T}^2} \mathbf{w}_1\cdot\mathbf{w}_2\, dx$,

$$\begin{aligned}
\mathscr{N}^*(\mathbf{u}^*) &= \sup_{\mathbf{u}}\{\langle \mathbf{u},\mathbf{u}^*\rangle - \|\mathbf{u}\|_{L^\infty}\} \\
&= \sup_{\mathbf{u}}\{\|\mathbf{u}\|_{L^\infty}\|\mathbf{u}^*\|_{\mathscr{M}} - \|\mathbf{u}\|_{L^\infty}\} = \chi_{\{\|\mathbf{u}^*\|_{\mathscr{M}}\le 1\}} = \begin{cases} 0, & \text{if } \|\mathbf{u}^*\|_{\mathscr{M}} \le 1 \\ +\infty, & \text{otherwise} \end{cases}; \\
\mathscr{E}^*(p^*) &= \sup_{p}\{\langle p, p^*\rangle - \lambda\|f-p\|_{L^2}^2\} \\
&= \sup_{p}\{-\lambda\langle p,p\rangle + \langle p^* + 2\lambda f, p\rangle - \lambda\langle f,f\rangle\} = \left\langle f + \frac{1}{4\lambda}p^*, p^*\right\rangle,
\end{aligned}$$

and $\Lambda^* = -\nabla$ is the dual operator of Λ.

We end up with the dual $(\mathscr{P}^*)$ problem

$$\inf_{\{p^*:\|\nabla p^*\|_{\mathscr{M}}\le 1\}} \left\langle \frac{1}{4\lambda}p^* - f, p^*\right\rangle$$

or

$$\inf_{p^*}\sup_{\mu\ge 0}\left[\left\langle \frac{1}{4\lambda}p^* - f, p^*\right\rangle + \mu(\|\nabla p^*\|_{\mathscr{M}} - 1)\right]. \tag{5.16}$$

Moreover, $-\bar{p}^* \in \partial\mathscr{E}(\Lambda\bar{u})$, meaning that $p^* = 2\lambda r$, where r is the residual, $r = f - div\,\mathbf{u}$. So, we can express the dual problem (5.16) in terms of r,

$$\bar{r} = \underset{r}{arg\,min} \sup_{\mu\geq 0} L(r,\mu), \qquad L(r,\mu) := \lambda \langle r - 2f, r\rangle + \mu\left(\mathsf{TV}[r] - \frac{1}{2\lambda}\right), \tag{5.17}$$

where $\bar{r} := f - div\,\bar{\mathbf{u}}$, is the residual corresponding to the optimal minimizer $\bar{u}$. The $\mathsf{TV}[\cdot]$ semi-norm is defined in (4.1), and $\|\nabla r\|_{\mathscr{M}} = \mathsf{TV}[r]$.

Since $L(\cdot,\mu)$ is convex and $L(r,\cdot)$ is concave and, for $r \in BV$ continuous, we can apply the minimax theorem, e.g., [32, §6], which allows us to interchange the infimum and supremum in (5.17), yielding

$$\sup_{\mu\geq 0} \min_{r} \left[\lambda \langle r - 2f, r\rangle + \mu\left(\mathsf{TV}[r] - \frac{1}{2\lambda}\right)\right]. \tag{5.18}$$

The dual problem, (5.18), can be solved in two steps. An inner minimization problem

$$r_\mu = \underset{r}{arg\,min} \left[\lambda \langle r - 2f, r\rangle + \mu\left(\mathsf{TV}[r] - \frac{1}{2\lambda}\right)\right]. \tag{5.19a}$$

Here, for any given $\mu \geq 0$, there exists a unique $r = r_\mu$ such that (μ, r_μ) is a saddle point of L. The optimal $\mu = \mu^*$ is determined by an outer maximization problem,

$$\mu^* = \underset{\mu\geq 0}{arg\,max}\,[P(\mu) + \mu Q(\mu)],$$
$$P(\mu) := \lambda \left\langle r_\mu - 2f, r_\mu \right\rangle, \;\; Q(\mu) := \mathsf{TV}[r_\mu] - \frac{1}{2\lambda}. \tag{5.19b}$$

Once μ^* is found, then $\bar{r} = r_{\mu^*}$ is the optimal residual which is sought as the solution of (5.17) .

5.3.2 The outer maximization problem

We begin by characterizing the maximizer, $\mu = \mu^*$, of the outer problem (5.19b). Fix μ: since r_μ minimizes $L(r,\mu)$ we have

$$P(\mu) + \mu Q(\mu) \leq P(\nu) + \mu Q(\nu).$$

Similarly, $P(\nu) + \nu Q(\nu) \leq P(\mu) + \nu Q(\mu)$. Sum the last two inequalities to get, $(\mu - \nu)[Q(\mu) - Q(\nu)] \leq 0$, which yields that $Q(\cdot)$ is non-increasing.

Let μ^* be a maximizer of (5.19b). Then $\forall \mu \geq 0$,

$$P(\mu) + \mu Q(\mu) \leq P(\mu^*) + \mu^* Q(\mu^*) \leq P(\mu) + \mu^* Q(\mu),$$

which implies $(\mu^* - \mu)Q(\mu) \geq 0$. We distinguish between two cases.

Case #1: $\mu^* > 0$. We have $Q(\mu) \leq 0$ if $\mu > \mu^*$ and $Q(\mu) \geq 0$ if $0 \leq \mu < \mu^*$. We conclude that μ^* is determined as a root of $Q(\cdot)$,

$$Q(\mu^*) = 0, \quad \text{i.e.} \quad \mathsf{TV}[r_{\mu^*}] = \frac{1}{2\lambda}. \tag{5.20}$$

Case #2: $\mu^* = 0$. In this case, r_0 minimizes $\langle r - 2f, r\rangle$, namely, $r_0 = f$. This corresponds to the trivial minimizer of (5.15), $\bar{\mathbf{u}} \equiv 0$, which is the case we want to avoid. Case #2 happens when $Q(0) \leq 0$, i.e.

$$\mu^* \;\leftrightarrow\; \mathsf{TV}[r_0] - \frac{1}{2\lambda} \leq 0 \;\leftrightarrow\; \mathsf{TV}[f] \leq \frac{1}{2\lambda}.$$

So, to make sure that we pick a non-trivial minimizer, $\bar{\mathbf{u}} \not\equiv 0$, we must pick a sufficiently large λ such that

$$\lambda > \frac{1}{2\mathsf{TV}[f]} \;\leftrightarrow\; \bar{\mathbf{u}} \not\equiv 0,\ \mathsf{TV}[\bar{r}] = \frac{1}{2\lambda}. \tag{5.21}$$

It has the same flavor as the Meyer's result for ROF model in image processing, consulting proposition 4.1.2(c). It also coincides with the same lower bound on λ's which yield non-trivial minimizers, asserted in [91, Lemma 5.3].

5.3.3 The inner minimization problem

We return to the inner minimization problem (5.19a). For fixed $\mu = \mu^*$, (5.19a) has the system expression as the ROF model (4.2). Following the procedure in section 4.1.3, we derive the following Euler-Lagrange equations

$$2\lambda (r_{\mu^*} - f) - \mu^* div \left(\frac{\nabla r_{\mu^*}}{|\nabla r_{\mu^*}|} \right) = 0. \tag{5.22}$$

Take the L^2-inner product of (5.22) with r_{μ^*} to get

$$2\lambda \left\langle r_{\mu^*} - f, r_{\mu^*} \right\rangle - \mu^* \left\langle div \left(\frac{\nabla r_{\mu^*}}{|\nabla r_{\mu^*}|} \right), r_{\mu^*} \right\rangle = 0.$$

Using (5.20) (and in the non-periodic case, the Neumann boundary condition $\nabla r_{\mu^*} \cdot \mathbf{n} = 0$), we find

$$\left\langle div \left(\frac{\nabla r_{\mu^*}}{|\nabla r_{\mu^*}|} \right), r_{\mu^*} \right\rangle = - \left\langle \frac{\nabla r_{\mu^*}}{|\nabla r_{\mu^*}|}, \nabla r_{\mu^*} \right\rangle = - \int_{\mathbb{T}^2} |\nabla r_{\mu^*}| dx = -\frac{1}{2\lambda}.$$

This yields, $\mu^* = 4\lambda^2 \left\langle f - r_{\mu^*}, r_{\mu^*} \right\rangle$, and the governing equation (5.22) for the optimal residual, $\bar{r} = r_{\mu^*}$, amounts to

$$(\bar{r} - f) - 2\lambda \left\langle f - \bar{r}, \bar{r} \right\rangle div \left(\frac{\nabla \bar{r}}{|\nabla \bar{r}|} \right) = 0. \tag{5.23}$$

Remark 5.3.1. This system has two solutions: one solution, $\bar{r} = f$, corresponds to the trivial case, $\bar{\mathbf{u}} \equiv 0$. The other is the target solution, i.e., the optimal residual $\bar{r}$ for (5.17). We will discuss numerical algorithms to solve system (5.23) in section 5.4.

5.3.4 From r to $\mathbf{u}$: recovering the uniformly bounded solution

So far, we identified the residual, $\bar{r} = f - div\,\bar{\mathbf{u}}$, corresponding to the uniformly bounded solution $\bar{\mathbf{u}}$ of (5.15). To recover $\bar{\mathbf{u}}$ itself, we substitute $\bar{r} - f = -div\,\bar{\mathbf{u}}$ as the first term of (5.23), and get

$$div\left(\bar{\mathbf{u}} - 2\lambda\,\langle \bar{r} - f, \bar{r}\rangle\,\frac{\nabla \bar{r}}{|\nabla \bar{r}|}\right) = 0. \tag{5.24}$$

Therefore, we can recover a solution $\bar{\mathbf{u}}$ of (5.15),

$$\bar{\mathbf{u}} = 2\lambda\,\langle \bar{r} - f, \bar{r}\rangle\,\frac{\nabla \bar{r}}{|\nabla \bar{r}|}. \tag{5.25}$$

Observe that this $\bar{\mathbf{u}}$ is indeed uniformly bounded:

$$\|\bar{\mathbf{u}}\|_{L^\infty} = 2\lambda\,|\,\langle \bar{r} - f, \bar{r}\rangle\,| < \infty. \tag{5.26}$$

Remark 5.3.2. The explicit expression of $\bar{\mathbf{u}}$ in (5.25) shows that $[\bar{\mathbf{u}}, \bar{r}]$ forms an extremal pair, in the sense of achieving an equality in the duality inequality of pairing $div\,\bar{\mathbf{u}}$ and $\bar{r}$:

$$|\langle div\,\bar{\mathbf{u}}, \bar{r}\rangle| = \|\bar{\mathbf{u}}\|_{L^\infty}\frac{1}{2\lambda} = \|\bar{\mathbf{u}}\|_{L^\infty}\|\nabla \bar{r}\|_{\mathscr{M}}.$$

Similar argument is addressed in [93, Theorem 2.3], [91, Theorem 5.1], and proposition 4.1.2(c) in the previous chapter.

5.4 Numerical algorithms for the hierarchical solution

We solve problem (5.1) using its hierarchical decomposition. In each iteration, we solve the minimization problem (5.15). Each iteration consists of three stages:

Stage 1. Find the non-trivial solution, r_j, of Euler-Lagrange equations (5.23) with $\lambda = \lambda_j$ and $f = f_j$;

Stage 2. Recover $\mathbf{u}_j$ from r_j using equation (5.25);

Stage 3. Update $\lambda_{j+1} \leftarrow 2\lambda_j$, $f_{j+1} \leftarrow r_j$.

Initially, we set λ_1 sufficiently large so that $\lambda_1 > (2\|F\|_{L^2})^{-1}$, and $f_1 := f$. The iterations terminate when $\|f_j\|_{L^2}$ is sufficiently small. The final solution U for (5.1) is given by the sum of all $\mathbf{u}_j$'s.

5.4.1 Numerical discretization for the PDE system

We begin with *regularization*: to avoid the singularity in (5.19a) when $|\nabla r| = 0$, a standard approach is to regularize the problem using a small parameter $\varepsilon > 0$,

$$r_{\mu,\varepsilon} = \underset{r}{arg\,min} \left\{ \lambda \langle r - 2f, r \rangle + \mu \left(\int_{\mathbb{T}^2} \sqrt{\varepsilon^2 + |\nabla r|^2} dxdy - \frac{1}{2\lambda} \right) \right\}. \tag{5.27}$$

At stage 1 of each regularized iteration, we find the minimizer $r = r_{\mu^*,\varepsilon}$. The corresponding Euler-Lagrange equations of the regularized problem read,

$$(r - f) - 2\lambda \langle f - r, r \rangle \cdot div \left(\frac{\nabla r}{\sqrt{\varepsilon^2 + |\nabla r|^2}} \right) = 0. \tag{5.28}$$

In the non-periodic case, these equations are augmented with Neumann boundary condition, $\nabla r \cdot \mathbf{n} = 0$.

To solve (5.28), we cover $\mathbb{T}^2$ with a computational grid with cell size h. Let D_{+x}, D_{-x} and D_{0x} be the usual forward, backward and centered divided difference operator on x, namely, $D_{\pm x} r_{i,j} = \pm(r_{i\pm1,j} - r_{i,j})/h$, $D_{0x} r_{i,j} = (r_{i+1,j} - r_{i-1,j})/2h$. Similarly,

we can define $D_{\pm y}$ and D_{0y}. A straightforward discretization of (5.28) yields,

$$\begin{aligned} r_{i,j} &= f_{i,j} - K(r)\cdot D_{-x}\left[\frac{1}{\sqrt{\varepsilon^2+(D_{+x}r_{i,j})^2+(D_{0y}r_{i,j})^2}}D_{+x}r_{i,j}\right] \\ &\quad - K(r)\cdot D_{-y}\left[\frac{1}{\sqrt{\varepsilon^2+(D_{0x}r_{i,j})^2+(D_{+y}r_{i,j})^2}}D_{+y}r_{i,j}\right] \qquad (5.29) \\ &= f_{i,j} - \frac{K(r)}{h^2}\left[\frac{r_{i+1,j}-r_{i,j}}{\sqrt{\varepsilon^2+(D_{+x}r_{i,j})^2+(D_{0y}r_{i,j})^2}} - \frac{r_{i,j}-r_{i-1,j}}{\sqrt{\varepsilon^2+(D_{+x}r_{i-1,j})^2+(D_{0y}r_{i-1,j})^2}}\right] \\ &\quad - \frac{K(r)}{h^2}\left[\frac{r_{i,j+1}-r_{i,j}}{\sqrt{\varepsilon^2+(D_{0x}r_{i,j})^2+(D_{+y}r_{i,j})^2}} - \frac{r_{i,j}-r_{i,j-1}}{\sqrt{\varepsilon^2+(D_{0x}r_{i,j-1})^2+(D_{+y}r_{i,j-1})^2}}\right]. \end{aligned}$$

Here, $K(r) := 2\lambda\,\langle r-f, r\rangle$, which is approximated using any appropriate numerical quadrature.

5.4.2 Computing the residuals r by implicit iterations

We use implicit iteration method to solve the nonlinear system (5.29),

$$\begin{aligned} r_{i,j}^{(n+1)} &= f_{i,j} \qquad (5.30) \\ &\quad - \frac{K(r^{(n)})}{h^2}\left[\frac{r_{i+1,j}^{(n+1)}-r_{i,j}^{(n+1)}}{\sqrt{\varepsilon^2+(D_{+x}r_{i,j}^{(n)})^2+(D_{0y}r_{i,j}^{(n)})^2}} - \frac{r_{i,j}^{(n+1)}-r_{i-1,j}^{(n+1)}}{\sqrt{\varepsilon^2+(D_{+x}r_{i-1,j}^{(n)})^2+(D_{0y}r_{i-1,j}^{(n)})^2}}\right] \\ &\quad - \frac{K(r^{(n)})}{h^2}\left[\frac{r_{i,j+1}^{(n+1)}-r_{i,j}^{(n+1)}}{\sqrt{\varepsilon^2+(D_{0x}r_{i,j}^{(n)})^2+(D_{+y}r_{i,j}^{(n)})^2}} - \frac{r_{i,j}^{(n+1)}-r_{i,j-1}^{(n+1)}}{\sqrt{\varepsilon^2+(D_{0x}r_{i,j-1}^{(n)})^2+(D_{+y}r_{i,j-1}^{(n)})^2}}\right], \end{aligned}$$

subject to initial condition which we set to be $r^{(0)} = f/2$.

Remark 5.4.1. Recall that $K(r)$ is continuous, and $K(\bar{r}) < 0$ while $K(f) = 0$. To avoid the convergence of $r^{(n)}$ to the trivial solution, $\bar{r} = f$ (mentioned in remark (5.3.1)), we set $r^{(0)}$ small enough, $K(r^{(0)}) < K(\bar{r}) < K(f)$, so that $r^{(n)}$ is expected to reach the non-trivial solution $\bar{r}$, rather than f. As $\underset{r}{arg\,min}\, K(r) = f/2$, a good choice of the initial condition of the iteration is $r^{(0)} = f/2$.

In the non-periodic case, we also need to apply Neumann boundary condition $\nabla r \cdot \mathbf{n} = 0$. To this end, we mirror r at the boundary, meaning $r_{0,j} = r_{2,j}, r_{N+1,j} = r_{N-1,j}$, etc, where the size of the grid is $N \times N$. So we only need to add the weight of the outer points to their corresponding inner points.

In summary, at the nth iteration amounts to an $N \times N$ linear system, $A(r^{(n)})\tilde{r}^{(n+1)} = \tilde{f}$, for the discretized nodes, $\{r^{(n+1)}\}$. Here, A is a sparse matrix with at most 5 non-zero entries every row or column, whose values depend on $r^{(n)}$.

5.4.3 Recovering $\mathbf{u}$ from r and control of errors

After we get a non-trivial solution r at stage 1, we move to stage 2 to recover $\mathbf{u}$ by (5.25). Normally, we apply centered divided difference operator on r to compute the discrete gradient, ∇r. However, this will cause a significant error of the solution $\mathbf{u}$.

For example, consider $u^1_{i,j} = K \cdot \dfrac{r_{i+1,j} - r_{i-1,j}}{2h\sqrt{\varepsilon^2 + |\nabla r_{i,j}|^2}}$. Suppose the error for r in stage 1 is $e(r)$. Then, at points (x,y) such that $|\nabla r(x,y)| \approx 0$, the error for u^1 is of order $Ke(r)/(h\varepsilon)$. Therefore, dividing by $h\varepsilon$ with $\varepsilon \approx 0$, the error bound of u^1 can be significantly amplified at stage 2 of recovering $\mathbf{u}$, even if we obtain a sufficiently small $e(r)$ at stage 1. This amplification will get worse as we refine the mesh and h becomes smaller.

In order to get a reliable solution for $\mathbf{u}$, we cannot carry out stage 2 independent of the discretization stencil of stage 1. To this end, let

$$u^{1,(n+1)}_{i+1/2,j} = \frac{K^{(n)}}{h} \cdot \frac{r^{(n+1)}_{i+1,j} - r^{(n+1)}_{i,j}}{\sqrt{\varepsilon^2 + (D_{+x}r^{(n)}_{i,j})^2 + (D_{0y}r^{(n)}_{i,j})^2}}, \tag{5.31a}$$

$$u^{2,(n+1)}_{i,j+1/2} = \frac{K^{(n)}}{h} \cdot \frac{r^{(n+1)}_{i,j+1} - r^{(n+1)}_{i,j}}{\sqrt{\varepsilon^2 + (D_{0x}r^{(n)}_{i,j})^2 + (D_{+y}r^{(n)}_{i,j})^2}}. \tag{5.31b}$$

We then have

$$r_{i,j} = f_{i,j} - \frac{u^1_{i+1/2,j} - u^1_{i-1/2,j}}{h} - \frac{u^2_{i,j+1/2} - u^2_{i,j-1/2}}{h}.$$

The last two terms represent a numerical discretization of $div\,\mathbf{u}$. Therefore, we use (5.31) to recover $\mathbf{u}$ from the residual $r = f - div\,\mathbf{u}$ calculated at (5.30).

5.5 Numerical implementation

We apply our algorithm for the hierarchically constructed uniformly bounded solution for the example of $F \in L^2_{\#}$ defined at (5.3) with

$$\mathbb{T}^2 = [-1,1]\times[-1,1], \quad \theta = 1/3, \quad \zeta(r)\begin{cases} = e^{-\frac{1}{1-r^2}}, & |r|<1, \\ \equiv 0, & |r|\geq 1. \end{cases} \tag{5.32}$$

5.5.1 Hierarchical solution vesus Helmholtz solution

We concentrate on the first component of the solution U, denoted by U^1. Figure 5.1 shows Helmholtz solution, U^1_{Hel}, which slowly diverges at the origin. Figure 5.2 provides the hierarchical solution U^1_{Bdd} which remains uniformly bounded.

The computed hierarchical solution $\|U^{1,N}_{\text{Bdd}}\|_{L^\infty}/\|F^N\|_{L^2}$ remains uniformly bounded when N increases ($U^{1,N}_{\text{Bdd}}$ stands for the first component of hierarchical solution with grid size $N\times N$.) In contrast, table 5.1 illustrates the (slow) growth of the ratio $\|U^{1,N}_{\text{Hel}}\|_{L^\infty}/\|F^N\|_{L^2}$.

5.5.2 Hierarchical solution meets Helmholtz solution

The hierarchical solution is uniformly bounded. However, as observed in figure 5.2, the hierarchical solution U^1_{Bdd} is oscillatory outside the support of F. As each step of

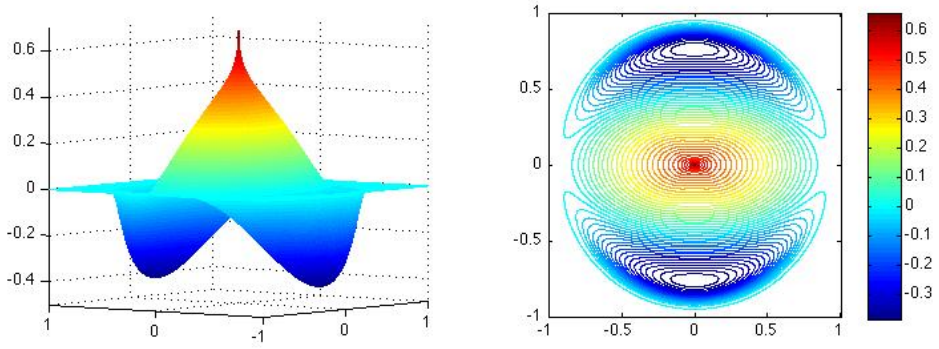

Figure 5.1: Helmholtz solution U^1_{Hel} of example (5.3),(5.32). A blowup is observed at the origin.

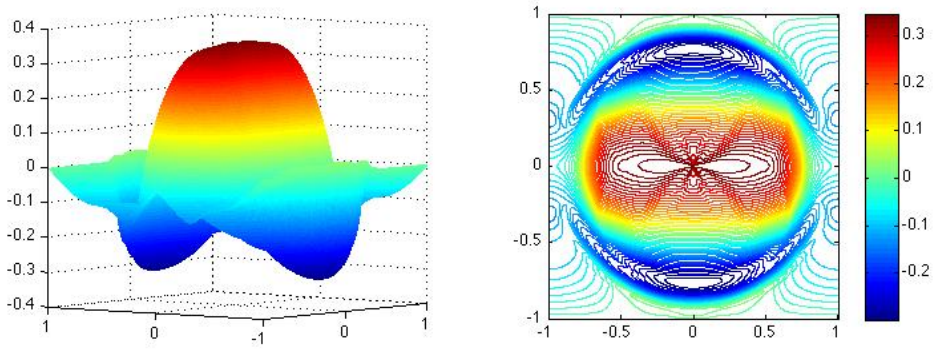

Figure 5.2: Hierarchical solution U^1_{Bdd} of (5.3),(5.32). It is clearly uniformly bounded.

the hierarchical decomposition relies on the previous steps, these oscillations will grow throughout the iterations. To limit their effect, we introduce a new, two-step method to construct bounded solutions of (5.1). It consists of one hierarchical decomposition step, whose residual is treated using Helmholtz decomposition:

Step 1. Solve minimization problem

$$\mathbf{u}_1 := argmin\mathbf{u}\{\|\mathbf{u}\|_{L^\infty} + \lambda_1\|F - div\,\mathbf{u}\|^2_{L^2}\}. \tag{5.33a}$$

The $N \times N$ grid	50×50	100×100	200×200	400×400	800×800
$\frac{\|U^{1,N}_{\text{Hel}}\|_{L^\infty}}{\|F^N\|_{L^2}}$	0.2295	0.2422	0.2540	0.2650	0.2752
$\frac{\|U^{1,N}_{\text{Bdd}}\|_{L^\infty}}{\|F^N\|_{L^2}}$	0.1454	0.1451	0.1455	0.1458	0.1451

Table 5.1: L^∞ norm of numerical solutions for different grids: Helmholtz vs. hierarchical construction. For Helmholtz solution, the value growth as with finer mesh, indicating a blowup, while for hierarchical constructed solution, the value is independent of mesh size.

Step 2. Find the Helmholtz solution for $div\,\mathbf{u}_r = r_1$, i.e.

$$\mathbf{u}_r := \nabla\Delta^{-1} r_1, \quad r_1 = F - div\,\mathbf{u}_1. \tag{5.33b}$$

The two-step solution, $U_{2\text{step}} = \mathbf{u}_1 + \mathbf{u}_r$, satisfies $div\,U = F$. Furthermore, it is uniformly bounded.

Proposition 5.5.1. *The two-step solution, $U_{2step} = \mathbf{u}_1 + \mathbf{u}_r$ given in* (5.33)*, is a uniformly bounded solution of* (5.1)*.*

Proof. Clearly, $\mathbf{u}_1$, as the first iteration of the hierarchical solution, is uniformly bounded. Next, equation (5.8) implies $r_1 \in BV_\#$. Applying proposition 5.2.1, $\mathbf{u}_r$ is uniformly bounded as well. □

From Proposition 5.5.1, we know that $U_{2\text{step}}$ is also a solution of (5.1). As the minimization problem is solved only once, we expect fewer oscillations in $U_{2\text{step}}$ than

U_{Bdd}.

Figure 5.3 shows the two-step solution of the example in Section 5.5.1. From the contour plot, we observe fewer oscillations than the hierarchical solution U_{Bdd}. Yet, the solution is not as smooth as U_{Bdd} at the origin. Table 5.2 reports that the ratio $\|U^{1,N}_{\text{2step}}\|_{L^\infty}/\|F^N\|_{L^2}$ is also stable when N is large. This verifies the uniformly boundedness of the two-step solution.

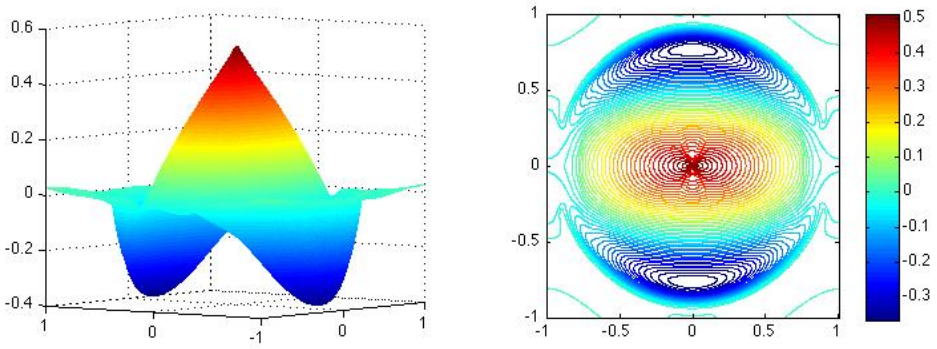

Figure 5.3: Two-step solution U^1_{2step} of (5.3),(5.32). One hierarchical step is enough to generate a bounded solution.

The $N \times N$ grid	50×50	100×100	200×200	400×400	800×800
$\dfrac{\|U^{1,N}_{\text{2step}}\|_{L^\infty}}{\|F^N\|_{L^2}}$	0.2096	0.2128	0.2144	0.2151	0.2154

Table 5.2: L^∞ norm of two-step solution for different grids. It is uniformly bounded.

5.6 Potential applications

The hierarchically constructed solution U_{Bdd} is a new candidate for solving $div\,\mathbf{u} = F$, other than the preferred Helmholtz solution U_{Hel}. Being uniformly bounded, it has the potential to replace Helmholtz solution in the critical spaces. In this section, we briefly discuss possible applications, which might lead to valuable future study.

We start with Euler-Poisson equations (2.8). The Poisson forcing reads

$$\mathsf{F} = -k\nabla\phi = k\nabla\Delta^{-1}\rho,$$

namely, F is the Helmholtz solution of the equation $div\,\mathsf{F} = k\rho$.

As discussed in section 2.8, the global existence theory for 2D Euler-Poisson equations is incomplete. An additional logarithmic growth for the L^∞ estimate on Reisz transform is the main reason for the failure of closing the loop. Such logarithmic growth is also suffered by U_{Hel}: in section 5.1.2, we know U_{Hel} is not bounded, only subject to a fractional logarithmic growth.

We modify Euler-Poisson equation by setting the forcing F as the uniformly bounded solution of $div\,\mathsf{F} = k\rho$. The open problem is whether the new choice of F will enable us to close the loop for the global existence theory for Euler-Poisson equation.

There are other systems with a substructure of $div\,U = F$ type. A usual assumption on U is $curl\,U = 0$ and the Helmholtz solution is used therein. One question is whether the presence of vorticity will regularize the system and prevent blowups. See [65] for some related arguments.

Consider 2D Keller-Segel system for chemotaxis [52, 78] of parabolic-elliptic type:

$$\rho_t = \Delta\rho - div(\rho\nabla c),$$

$$-\Delta c = \rho.$$

This reaction-diffusion system has been well-studied: if the mass is less than 8π, there exists global weak solutions. If the mass is greater than 8π, then there is a finite time concentration. Consult the review [9] and references therein.

The velocity field $\mathbf{u} = \nabla c$ has divergence $-\rho$ and vorticity zero. If we add vorticity to the velocity field, numerical evidence suggests possibility of no finite time concentration for some data with mass larger than 8π, *e.g.*, [68]. It is an indication that vorticity prevents concentration.

The modified system reads

$$\rho_t = \Delta\rho - div(\rho\mathbf{u}),$$

$$div\,\mathbf{u} = -\rho, \quad curl\,\mathbf{u} = \omega.$$

where ω could be modeled in various ways. For instance, if ω satisfies

$$\omega_t + \mathbf{u}\cdot\nabla\omega = \nu\Delta\omega,$$

the coupled system is called Keller-Segel-Navier-Stokes system. We refer [66] for studies on related problems. It is not clear if the coupled system has a larger critical mass.

Another way to introduce vorticity to the system is to define the velocity field $\mathbf{u}$ as the uniformly bounded solution of $div\,\mathbf{u} = -\rho$. Whether the corresponding system consists a larger critical mass is open for future studies.

Bibliography

[1] J. A. ACEBRN, L. L. BONILLA, C. J. P. VICENTE, F. RITORT AND R. SPIGLER, *The Kuramoto model: A simple paradigm for synchronization phenomena*, Reviews of modern physics, 77, no. 1, (2005): 137.

[2] I. AOKI, *A simulation study on the schooling mechanism in fish*, Bull. Jpn. Soc. Sci. Fish., 48, no. 8, (1982): 1081.

[3] C. APPERT-ROLLAND, P. DEGOND AND S. MOTSCH, *Two-way multi-lane traffic model for pedestrians in corridors*, Networks and Heterogeneous Media, 6, no. 3, (2011): 351-381.

[4] D. BALAGUE, J. A. CARRILLO AND Y. YAO, *Confinement for repulsive-attractive kernels*, arXiv:1210.0602

[5] M. BALLERINI, N. CABIBBO, R. CANDELIER, A. CAVAGNA, E. CISBANI, I. GIARDINA, V. LECOMTE, A. ORLANDI, G. PARISI, A. PROCACCINI, M. VIALE AND V. ZDRAVKOVIC, *Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study*, Proceedings of the National Academy of Sciences, 105, no. 4, (2008): 1232–1237.

[6] D. BENEDETTO, E. CAGLIOTI AND M. PULVIRENTI, *A kinetic equation for granular media*, ESAIM: Mathematical Modelling and Numerical Analysis, 33, no. 2, (1999): 439–441.

[7] A. BERTOZZI AND T. LAURENT, *The behavior of solutions of multidimensional aggregation equations with mildly singular interaction kernels*, Chinese Annals of Mathematics, Series B, 30, no. 5, (2009): 463–482.

[8] B. BIRNIR, *An ODE model of the motion of pelagic fish*, J. Stat. Phys., 128, no. 1, (2007): 535-568.

[9] A. BLANCHET, J. DOLBEAULT AND B. PERTHAME, *Two-dimensional Keller-Segel model: Optimal critical mass and qualitative properties of the solutions*, Electronic Journal of Differential Equations, 2006.

[10] E. BONABEAU, G. THERAULAZ, J. L. DENEUBOURG, S. ARON AND S. CAMAZINE, *Self-organization in social insects*, Trends in Ecology and Evolution, 12, no. 5, (1997): 188–193.

[11] J. BOURGAIN AND H. BREZIS, *On the equation* $\operatorname{div} Y = f$ *and application to control of phases*, J. Amer. Math. Soc. 16, no. 2, (2003): 393–426.

[12] J. BOURGAIN AND H. BREZIS, *New estimates for elliptic equations and Hodge type systems*, J. Eur. Math. Soc., 9, no. 2, (2007): 277–315.

[13] A. BRUCKSTEIN, N. COHEN AND A. EFRAT, *Ants, crickets and frogs in cyclic pursuit*, Center Intell. Syst., Technion-Israel Inst. Technol., Haifa, Israel, Tech. Rep 9105 (1991).

[14] J. BUHL, D. J. SUMPTER, D. COUZIN, J. J. HALE, E. DESPLAND, E. R. MILLER, AND S. J. SIMPSON, *From disorder to order in marching locusts*, Science, 312, no. 5778, (2006): 1402–1406.

[15] L. CAFFARELLI AND A. VASSEUR, *Drift diffusion equations with fractional diffusion and the quasi-geostrophic equation*, Annals of Math., 171, no. 3, (2010): 1903–1930.

[16] C. CANUTO, F. FAGNANI AND P. TILLI, *An Eulerian approach to the analysis of Krause's consensus models*, SIAM Journal on Control and Optimization, 50, no. 1, (2012): 243–265.

[17] R. CARLI, F. FAGNANI, A. SPERANZON AND S. ZAMPIERI, *Communication constraints in the average consensus problem*, Automatica, 44, (2008): 671–684.

[18] J. A. CARRILLO, M. DI FRANCESCO, A. FIGALLI, T. LAURENT AND D. SLEPCEV, *Global-in-time weak measure solutions and finite-time aggregation for nonlocal interaction equations*, Duke Math. J., 156, (2011): 229–271.

[19] J. A. CARRILLO, M. FORNASIER, J. ROSADO, AND G. TOSCANI, *Asymptotic flocking dynamics for the kinetic Cucker-Smale model*, SIAM Journal on Mathematical Analysis, 42, no. 1, (2010): 218–236.

[20] A. CHAMBOLLE AND P. L. LIONS, *Image recovery via total variational minimization and related problems*, Numer. Math., 76, (1997): 167–188.

[21] T. F. CHAN AND J. SHEN, *Image processing and analysis: variational, PDE, wavelet, and stochastic methods*, SIAM, 2005.

[22] B. CHENG AND E. TADMOR, *An improved local blow-up condition for Euler-Poisson equations with attractive forcing*, Physica D., 238, no. 20, (2009): 2062–2066.

[23] B. COCKBURN, C.-W. SHU, *TVB Runge-Kutta Local projection discontinuous Galerkin finite element method for conservation law II: General framework*, Mathematics of Computation, 52, (1989): 411–435.

[24] A. COHEN, *private communication.*

[25] P. CONSTANTIN AND VLAD VICOL, *Nonlinear maximum principles for dissipative linear nonlocal operators and applications*, Geometric And Functional Analysis, 22, no. 5, (2012): 1289–1321.

[26] F. CUCKER AND S. SMALE, *Emergent behavior in flocks*, IEEE Trans. Autom. Control, 52, no. 5, (2007): 852.

[27] F. CUCKER AND S. SMALE, *On the mathematics of emergence*, Jpn. J. Math., 2, no. 1, (2007): 197–227.

[28] B. DACOROGNA, N. FUSCO AND L. TARTAR, *On the solvability of the equation* $div\, u = f$ *in* L^1 *and in* C^0, Rend. Mat. Acc. Lincei., 14, no. 3, (2003): 239–245.

[29] P. DEGOND, A. FROUVELLE AND J.-G. LIU, *Macroscopic limits and phase transition in a system of self-propelled particles*, Journal of nonlinear science, 23, no. 3, (2013): 427–456.

[30] P. DEGOND AND S. MOTSCH, *Continuum limit of self-driven particles with orientation interaction*, Mathematical Models and Methods in Applied Sciences, 18, no. 1 (2008):1193–1215.

[31] M. DORSOGNA, Y. L. CHUANG, A. BERTOZZI, AND L. CHAYES, *Self-propelled particles with soft-core interactions: patterns, stability, and collapse*, Physical review letters, 96, no. 10, (2006): 104–302.

[32] I. EKELAND AND R. TÉMAM, *Convex Analysis and Variational Problems*, SIAM, 1999.

[33] S. ENGELBERG, H. LIU AND E. TADMOR, *Critical threshold in Euler-Poisson equations*, Indiana Univ. Math. J., 50, (2001): 109–157.

[34] L. EVANS, *Partial Differential Equations*, 2nd ed., American Mathematical Society, 2010.

[35] A. FROUVELLE AND J.-G. LIU, *Dynamics in a kinetic model of oriented particles with phase transition*, SIAM Journal on Mathematical Analysis, 44, no. 2, (2012): 791–826.

[36] I. GASSER, P.E. JABIN AND B. PERTHAME, *Regularity and propagation of moments in some nonlinear Vlasov systems*, Proc. Roy. Soc. Edinburgh Sect. A, 130, (2000): 1259–1273.

[37] R. GLASSEY AND J. SCHAEFFER, *The relativistic Vlasov-Maxwell system in two space dimensions: Part I*, Arch. Rational Mech. Anal., 141, no. 4 (1998): 331–354.

[38] R. GLASSEY AND J. SCHAEFFER, *The relativistic Vlasov-Maxwell system in two space dimensions: Part II*, Arch. Rational Mech. Anal., 141, no. 4 (1998): 355–374.

[39] R. C. GONZALEZ AND R. E. WOODS, *Digital Image Processing*, Addison-Wesley, New York, 1992.

[40] S. GOTTLIEB, C.-W. SHU AND E. TADMOR, *Strong stability preserving high-order time discretization methods*, SIAM Review, 43, (2001): 89–112.

[41] L. GRAFAKOS, *Classical Fourier Analysis*, Vol 86, Springer, 2008.

[42] V. GRIMM AND S. F. RAILSBACK, *Individual-based modeling and ecology*, Princeton university press, 2013.

[43] Y. GUO, A. TAHVILDAR-ZADEH, *Formation of singularities in relativistic fluid dynamics and in spherically symmetric plasma dynamics*, Nonlinear partial differential equations,(AMS, Providence, 1999) 4 (1999).

[44] S.-Y. HA AND J.-G. LIU, *A simple proof of the Cucker-Smale flocking dynamics and mean-field limit*, Commun. Math. Sci., 7, (2009): 297–325.

[45] S.-Y. HA AND E. TADMOR, *From particle to kinetic and hydrodynamic descriptions of flocking*, Kinetic and Related Models, 1, no. 3, (2008): 415–435.

[46] D. HELBING *Traffic and related self-driven many-particle systems*, Reviews of modern physics, 73, no. 4, (2001): 1067–1141.

[47] C. K. HEMELRIJK AND H. HILDENBRANDT, *SelfOrganized Shape and Frontal Density of Fish Schools*, Ethology, 114, no. 3, (2008): 245–254.

[48] P.-E. JABIN, *Large time concentrations for solutions to kinetic equations with energy dissipation*, Communications in Partial Differential Equations, 25, no. 3-4 (2000): 541–557.

[49] M. JI AND M. B. EGERSTEDT, *Distributed coordination control of multi-agent systems while preserving connectedness*, IEEE Trans. Robot., 23, no. 4, (2007): 693–703.

[50] J. CORTÉS, S. MARTÍNEZ AND F. BULLO, *Robust rendezvous for mobile autonomous agents via proximity graphs in arbitrary dimensions*, IEEE Trans. Automat. Control, 51, no. 8, (2006): 1289–1298.

[51] T. KARPER, A. MELLET AND K. TRIVISA, *Hydrodynamic limit of the kinetic Cucker-Smale flocking model*, arXiv:1205.6831.

[52] E.F. KELLER AND L.A. SEGEL, *Initiation of slime mold aggregation viewed as an instability*, J. Theor. Biol., 26, no. 3, (1970): 399–415.

[53] A. KISELEV, F. NAZAROV AND R. SHTERENBERG, *Blow up and regularity for fractal Burgers equation*, Dyn. Partial Differ. Equ., 5, no. 3, (2008): 211–240.

[54] T. KATO AND G. PONCE, *Commutator estimates and the Euler and Navier-Stokes equations*, Communications on Pure and Applied Mathematics, 41, no. 7, (1988): 891–907.

[55] N. KATZ AND N. PAVLOVIC, *A cheap Caffarelli-Kohn-Nirenberg inequality for the Navier-Stokes equation with hyper-dissipation*, Geom. Funct. Anal., 12, no. 2, (2002): 355–379.

[56] A. L. KOCH, D. WHITE, *The social lifestyle of myxobacteria*, Bioessays, 20, no. 12, (1998): 1030–1038.

[57] U. KRAUSE, *Soziale Dynamiken mit vielen Interakteuren*, Eine Problemskizze. In Modellierung und Simulation von Dynamiken mit vielen interagierenden Akteuren, (1997): 37-51.

[58] Y. KURAMOTO, *Self-entralnment of a population of coupled non-llnear oscillators*, H. Araki, ed., Lecture Notes in Physics, International Symposium on Mathematical Problems in Theoretical Physics 39, Springer-Verlag, New York, (1975): 420.

[59] J. S. LIM, *Two-dimensional signal and image processing*, Prentice-Hall, Upper Saddle River, NJ, 1989.

[60] F. LIN AND P. ZHANG, *On the hydrodynamic limit of Ginzburg-Landau vortices*, Discrete and Continuous Dynamical Systems, 6, no. 1, (2000): 121–142.

[61] P. L. LIONS, B. PERTHAME, *Propagation of moments and regularity of the 3-dimensional Vlasov-Poisson system*, Invent. math., 105, (1991): 415–430.

[62] H. LIU AND E. TADMOR, *Critical thresholds in convolution model for nonlinear conservation laws*, SIAM J. Math. Anal., 33, no. 4, (2001): 930–945.

[63] H. LIU AND E. TADMOR, *Spectral dynamics of velocity gradient field in restricted flows*, Commun. Math. Phys., 228, (2002): 435–466.

[64] H. LIU AND E. TADMOR, *Critical threshold in 2D restricted Euler-Poisson equations*, SIAM J. Math. Anal., 63, no. 6, (2003): 1889–1910.

[65] H. LIU AND E. TADMOR, *Rotation prevents finite-time breakdown*, Physica D, 188, (2004): 262–276.

[66] J.-G. LIU AND A. LORZ, *A coupled chemotaxis-fluid model: Global existence*, Ann. I. H. Poincare, AN, 28, (2011): 643–652.

[67] T.-P. LIU, Z. XIN AND T. YANG, *Vacuum states of compressible flow*, Discrete and Continuous Dynamical Systems, 4, (1998): 1–32.

[68] A. LORZ, *A coupled KellerSegelStokes model: global existence for small initial data and blow-up delay*, Communications in Mathematical Sciences, 10, no. 2, (2012).

[69] Y. MADAY, *L^∞-stable approximation of a solution to $divY = f$ for $f \in L^2$ in two dimensions*, J. Sci. Comput., 28, no. 2–3, (2006): 451–458.

[70] R. MCLACHLAN, G. QUISPEL, *Splitting methods*, Acta Numerica, 11, (2002): 341–434.

[71] Y. MEYER, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations*, University Lecture Series Volume 22, AMS 2002.

[72] S. MOTSCH AND E. TADMOR, *A new model for self-organized dynamics and its flocking behavior*, J. Stat. Phys, 144, no. 5, (2011): 923–947.

[73] S. MOTSCH AND E. TADMOR, *Heterophilious dynamics enhances consensus*, arXiv:1301.4123.

[74] A. MELLET, A. VASSEUR, *Existence and Uniqueness of global strong solutions for one-dimensional compressible Navier-Stokes equations*, SIAM J. Math. Anal., 39, no. 4, (2008): 1344–1365.

[75] C. MOUHOT AND C. VILLANI, *On landau damping*, Acta mathematica, 207, no. 1, (2011):29–201.

[76] R. O'NEIL, *Convolution operators and $L(p,q)$ spaces*, Duke Math, J., 30, (1963): 129–143.

[77] J. K. PARRISH, S. V. VISCIDO AND D. GRÜNBAUM, *Self-organized fish schools: an examination of emergent properties*, The biological bulletin, 202, no. 3, (2002): 296–305.

[78] C. S. PATLAK, *Random walk with persistence and external bias*, Bull. Math. Biophys., 15, (1953): 311-338.

[79] K. PFAFFELMOSER, *Global classical solutions of the Vlasov-Poisson system in three dimensions for general initial data*, J. Diff. Eq., 95, no. 2, (1992): 281–303.

[80] B. PICCOLI AND A. TOSIN, *Time-evolving measures and macroscopic modeling of pedestrian flow*, Archive for Rational Mechanics and Analysis, 199, no. 3 (2011): 707–738.

[81] W.H. REED, T.R. HILL, *Triangular mesh methods for the Neutron transport equation*, Los Alamos Scientific Laboratory Report LA-UR-73-479, Los Alamos, NM, 1973.

[82] CW. REYNOLDS, *Flocks, herds and schools: A distributed behavioral model*, ACM SIGGRAPH Computer Graphics, 21, no. 4, (1987): 25–34.

[83] L. RUDIN AND S. OSHER, *Total variation based image restoration with free local constraints*, Proceedings. ICIP-94., IEEE International Conference, 1, (1994): 31–35.

[84] L. RUDIN, S. OSHER AND E. FATEMI, *Nonlinear total variation based noise removal algorithms*, Physica D, 60, (1992): 259–268.

[85] E. RUSS, *A Survey About the Equation $div\, u = f$ in Bounded Domains of $\mathbb{R}^n$*, Vietnam Journal of Mathematics, 41, no. 4, (2013): 369–381.

[86] J. SCHAEFFER, *Global existence of smooth solutions to the Vlasov-Poisson system in three dimensions*, Comm. P. D. E., 16, (1991): 1313–1335.

[87] J. SCHELLINCK AND T. WHITE, *A review of attraction and repulsion models of aggregation: Methods, findings and a discussion of model validation*, Ecological Modelling, 222, no. 11, (2011): 1897–1911.

[88] C.-W. SHU, *Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws*, Springer Berlin Heidelberg, 1998.

[89] C.-W. SHU AND S. OSHER, *Efficient implementation of essentially non-oscillatory shockcapturing schemes*, Journal of Computational Physics, 77, (1988): 439–471.

[90] G. STRANG, *On the construction and comparison of difference schemes*, SIAM Journal on Numerical Analysis 5, no. 3, (1968): 506–517.

[91] E. TADMOR, *Hierarchical construction of bounded solutions in critical regularity spaces*, arXiv:1003.1525.

[92] E. TADMOR, S. NEZZAR AND L. VESE, *A Multiscale Image Representation Using Hierarchical (BV,L2) Decomposition*, Multiscale Modeling and Simulation: A SIAM Interdisciplinary Journal, 2, no. 4, (2004): 554–579.

[93] E. TADMOR, S. NEZZAR AND L. VESE, *Hierarchical decomposition of images with applications to deblurring, denoising and segmentation*, Communications in Math. Sciences, 6, no. 2, (2008): 281–307.

[94] E. TADMOR AND C. TAN, *Hierarchical Construction of Bounded Solutions of div U=F in Critical Regularity Spaces*, Nonlinear Partial Differential Equations. Springer Berlin Heidelberg, (2012): 255–269.

[95] E. TADMOR AND C. TAN, *Critical thresholds in flocking hydrodynamics with nonlocal alignment*, arXiv:1403.0991.

[96] E. TADMOR AND D. WEI, *On the global regularity of sub-critical Euler-Poisson equations with pressure*, J. European Math. Society, 10, 757–769, 2008.

[97] C. TAN, *A discontinuous Galerkin method on kinetic flocking models*, in preparation.

[98] L. TARTAR, *Lorentz spaces and applications*, Lecture notes, Carnegie Mellon University, 1989.

[99] C. M. TOPAZ, A. J. BERNOFF, S. LOGAN AND W. TOOLSON, *A model for rolling swarms of locusts*, The European Physical Journal Special Topics, 157, no. 1, (2008): 93–109.

[100] C. M. TOPAZ AND A. L. BERTOZZI, *swarming patterns in a two-dimensional kinematic model for biological groups*, SIAM J. Appl. Math., 65, (2004): 152–174.

[101] C. M. TOPAZ, A. L. BERTOZZI AND M. A. LEWIS, *A nonlocal continuum model for biological aggregation*, Bulletin of Mathematical Biology, 68, no. 7, (2006): 1601–1623.

[102] T. VICSEK, A. CZIROK, E. BEN-JACOB, I. COHEN AND O. SHOCHET, *Novel type of phase transition in a system of self-driven particles*, Physical review letters, 75, no. 6, (1995): 1226.

[103] J. WU, *Generalized MHD equations*, J. Differ. Eq. 195, (2003): 284–312.

[104] T. YANG, Z. YAO AND C. ZHU, *Compressible Navier-Stokes equations with density-dependent viscosity and vacuum*, Comm. Partial Differential Equations, 26, (2001): 965–981.

[105] Y. YANG AND C.-W. SHU, *Discontinuous Galerkin method for hyperbolic equations involving δ-singularities: negative-order norm error estimates and applications*, Numerische Mathematik, 124, no. 4, (2013): 753–781.

[106] Y. YANG, D. WEI AND C.-W. SHU, *Discontinuous Galerkin method for Krause's consensus models and pressureless Euler equations*, Journal of Computational Physics, 252, (2013): 109–127.

[107] T. YANG AND C. ZHU, *Compressible Navier-Stokes equations with degenerate viscosity coeffcient and vacuum*, Comm. Math. Phys., 230, no. 2, (2002): 329–363.

[108] X. ZHANG AND C.-W. SHU, *Maximum-principle-satisfying and positivity-preserving high order schemes for conservation laws: Survey and new developments*, Proceedings of the Royal Society A, 467, (2011): 2752–2776.

[109] W. ZIEMER, *Weakly Differentiable Functions*, Graduate Texts in Math., Springer, 1989.