Global existence for the semigeostrophic equations via Sobolev estimates for Monge-Ampère

Alessio Figalli[*][†]

Abstract

These notes record and extend the lectures for the CIME Summer Course held by the author in Cetraro during the week of June 2-7, 2014. Our goal is to show how some recent developments in the theory of the Monge-Ampère equation play a crucial role in proving existence of global weak solutions to the semigeostrophic equations.

Contents

[*]The University of Texas at Austin, Mathematics Dept. RLM 8.100, 2515 Speedway Stop C1200, Austin, Texas 78712-1202 USA. (figalli@math.utexas.edu)

[†]The author is supported by NSF Grant DMS-1262411.

1 The semigeostrophic equations

The semigeostrophic (in short, SG) equations are a simple model used in meteorology to describe large scale atmospheric flows. As explained for instance in [5, Section 2.2] (see also [16] for a more complete exposition), these equations can be derived from the 3-D incompressible Euler equations, with Boussinesq and hydrostatic approximations, subject to a strong Coriolis force. Since for large scale atmospheric flows the Coriolis force dominates the advection term, the flow is mostly bi-dimensional. For this reason, the study of the SG equations in 2-D or 3-D is pretty similar, and in order to simplify our presentation we focus here on the 2-dimentional periodic case.

1.1 The classical SG system

The 2-dimensional SG system can be written as

$$\begin{cases} \partial_t \nabla p_t + (\boldsymbol{u}_t \cdot \nabla)\nabla p_t + \nabla^\perp p_t + \boldsymbol{u}_t = 0 & \text{in } [0,\infty)\times\mathbb{R}^2, \\ \operatorname{div} \boldsymbol{u}_t = 0 & \text{in } [0,\infty)\times\mathbb{R}^2, \\ p_0 = \bar p & \text{on } \mathbb{R}^2, \end{cases} \tag{1.1}$$

where $\boldsymbol{u}_t = (\boldsymbol{u}_t^1, \boldsymbol{u}_t^2) : \mathbb{R}^2 \to \mathbb{R}^2$ and $p_t : \mathbb{R}^2 \to \mathbb{R}$ are time-dependent periodic[1] functions respectively corresponding to the velocity and the pressure.

In the above system the notation $\nabla^\perp p_t$ denotes the rotation of the vector ∇p_t by $\pi/2$, that is $\nabla^\perp p_t = (\partial_2 p_t, -\partial_1 p_t)$. Also, $(\boldsymbol{u}_t \cdot \nabla)$ denotes the operator $\boldsymbol{u}_t^1\partial_1 + \boldsymbol{u}_t^2\partial_2$. Hence, in components the first equation in (1.1) reads as

$$\partial_t\partial_1 p_t + \sum_{j=1,2} \boldsymbol{u}_t^j \partial_j\partial_1 p_t + \partial_2 p_t + \boldsymbol{u}_t^1 = 0,$$

$$\partial_t\partial_2 p_t + \sum_{j=1,2} \boldsymbol{u}_t^j \partial_j\partial_2 p_t - \partial_1 p_t + \boldsymbol{u}_t^2 = 0.$$

[1] By "periodic" we shall always mean $\mathbb{Z}^2$-periodic.

Notice that in (1.1) we have 3 equations (the two above, together with $\operatorname{div} \boldsymbol{u}_t = 0$) for the 3 unknowns $(p_t, \boldsymbol{u}_t^1, \boldsymbol{u}_t^2)$. Also, while in many equations in fluid mechanics one usually prescribes the evolution of the velocity field $\boldsymbol{u}_t$ and p_t acts as a Lagrange multiplier for the incompressibility constraint, here we are prescribing the evolution of the gradient of p_t and $\boldsymbol{u}_t$ acts as a Lagrange multiplier in order to ensure that the vector field ∇p_t remains a gradient along the evolution.

As shown in [16], energetic considerations show that it is natural to assume that p_t is (-1)-convex, i.e., the function $P_t(x) := p_t(x) + |x|^2/2$ is convex on $\mathbb{R}^2$. Hence, noticing that

$$\nabla p_t = \nabla P_t - x, \qquad \partial_t \nabla p_t = \partial_t \nabla P_t, \qquad (\boldsymbol{u}_t \cdot \nabla) x = \boldsymbol{u}_t,$$

we are led to the following extended system for P_t:

$$\begin{cases} \partial_t \nabla P_t + (\boldsymbol{u}_t \cdot \nabla) \nabla P_t + (\nabla P_t - x)^\perp = 0 & \text{in } [0,\infty) \times \mathbb{R}^2, \\ \operatorname{div} \boldsymbol{u}_t = 0 & \text{in } [0,\infty) \times \mathbb{R}^2, \\ P_t \text{ convex} & \text{in } [0,\infty) \times \mathbb{R}^2, \\ P_0 = \bar{p} + |x|^2/2 & \text{on } \mathbb{R}^2, \end{cases} \tag{1.2}$$

with the "boundary conditions" that both $P_t - |x|^2/2$ and $\boldsymbol{u}_t$ are periodic.

The existence theory for this equation is extremely complicated, and so far nobody has been able to attack directly this equation. Instead, there is a way to find a "dual equation" to this system for which existence of solutions is much easier to obtain, and then one can use this "dual solution" to go back and construct a solution to the original system. This is the goal of the next sections.

1.2 An evolution equation for the density associated to P_t: the dual SG system

Notice that ∇P_t can be seen a map from $\mathbb{R}^2$ to $\mathbb{R}^2$. Motivated by the fact that, in optimal transport theory, gradients of convex functions are optimal transport maps (see Theorem 2.1 below), it is natural to look at the image of the Lebesgue measure under this map and try to understand its behavior. Hence, denoting by dx denote Lebesgue measure on $\mathbb{R}^2$, we define the measure ρ_t as $(\nabla P_t)_\sharp dx$, that is, for any test function $\chi : \mathbb{R}^2 \to \mathbb{R}$,

$$\int_{\mathbb{R}^2} \chi(y)\, d\rho_t(y) := \int_{\mathbb{R}^2} \chi\big(\nabla P_t(x)\big)\, dx. \tag{1.3}$$

Before describing the properties of ρ_t, we make a simple observation that will be useful later.

Remark 1.1. Since $\nabla P_t - x$ is periodic, it is easy to check that the measure ρ_t is periodic on $\mathbb{R}^2$ and

$$\int_{[0,1]^2} d\rho_t = \int_{[0,1]^2} dx = 1,$$

so one can also identify it as a probability measure on the 2-dimensional torus $\mathbb{T}^2$.

Our goal now is to find an evolution equation for ρ_t. To this aim we take $\varphi \in C_c^\infty(\mathbb{R}^2)$ and we compute the time derivative of $\int \varphi\, d\rho_t$:

$$
\begin{aligned}
\frac{d}{dt}\int \varphi\, d\rho_t &\overset{(1.3)}{=} \frac{d}{dt}\int \varphi\big(\nabla P_t\big)\, dx \\
&= \int \nabla\varphi\big(\nabla P_t\big)\cdot \partial_t \nabla P_t\, dx \\
&\overset{(1.2)}{=} -\int \nabla\varphi\big(\nabla P_t\big)\cdot (\boldsymbol{u}_t\cdot\nabla)\nabla P_t\, dx - \int \nabla\varphi\big(\nabla P_t\big)\cdot(\nabla P_t - x)^\perp\, dx \\
&= -\int \nabla\varphi\big(\nabla P_t\big)\cdot D^2 P_t\cdot \boldsymbol{u}_t\, dx - \int \nabla\varphi\big(\nabla P_t\big)\cdot(\nabla P_t - x)^\perp\, dx \qquad (1.4)\\
&= -\int \nabla\big[\varphi\circ\nabla P_t\big]\cdot \boldsymbol{u}_t\, dx - \int \nabla\varphi\big(\nabla P_t\big)\cdot(\nabla P_t - x)^\perp\, dx \\
&= \int \big[\varphi\circ\nabla P_t\big]\operatorname{div}\boldsymbol{u}_t\, dx - \int \nabla\varphi\big(\nabla P_t\big)\cdot(\nabla P_t - x)^\perp\, dx \\
&\overset{(1.2)}{=} -\int \nabla\varphi\big(\nabla P_t\big)\cdot(\nabla P_t - x)^\perp\, dx,
\end{aligned}
$$

where at the last step we used that $\operatorname{div}\boldsymbol{u}_t = 0$.

In order to continue in the computations we need to introduce the Legendre transform of P_t:

$$P_t^*(y) := \sup_{x\in\mathbb{R}^2}\{x\cdot y - P_t(x)\} \qquad \forall\, y\in\mathbb{R}^2.$$

Notice that the function P_t^* is also convex, being the supremum of linear functions. Also, at least formally, the gradient of P_t and P_t^* are inverse to each other:[2]

$$\nabla P_t(\nabla P_t^*(y)) = y, \qquad \nabla P_t^*(\nabla P_t(x)) = x. \qquad (1.5)$$

Thanks to this fact we obtain that the last term in (1.4) is equal to

$$-\int \Big[\nabla\varphi\cdot(\nabla P_t^* - y)^\perp\Big]\circ\nabla P_t(x)\, dx,$$

which by (1.3) can also be rewritten as

$$-\int \nabla\varphi\cdot(\nabla P_t^* - y)^\perp\, d\rho_t(y).$$

[2]The relation (1.5) is valid only at point where the gradients of P_t and P_t^* both exist. There is however a weaker way to formulate such a relation that is independent of any regularity assumption: define the sub-differential of a convex function $\phi:\mathbb{R}^n\to\mathbb{R}$ as

$$\partial\phi(x) := \{p\in\mathbb{R}^n \,:\, \phi(z)\geq\phi(x) + p\cdot(z-x) \quad \forall\, z\in\mathbb{R}^n\}.$$

Then, using the definition of P_t^* it is not difficult to check that

$$y\in\partial P_t(x) \quad\Leftrightarrow\quad x\in\partial P_t^*(y).$$

Noticing that $\partial\phi(x) = \{\nabla\phi(x)\}$ whenever ϕ is differentiable at x, the above relation reduces exactly to (1.5) at differentiability points.

Hence, if we set

$$\boldsymbol{U}_t(y) := \big(\nabla P_t^*(y) - y\big)^{\perp},$$

(1.4) and an integration by parts give

$$\frac{d}{dt}\int \varphi \, d\rho_t = -\int \varphi \, \mathrm{div}(\boldsymbol{U}_t \rho_t),$$

and by the arbitrariness of φ we conclude that $\partial_t \rho_t + \mathrm{div}(\boldsymbol{U}_t \rho_t) = 0$. Thus we have shown that ρ_t satisfies the following *dual problem*:

$$\begin{cases} \partial_t \rho_t + \mathrm{div}(\boldsymbol{U}_t \rho_t) = 0 & \text{in } [0,\infty) \times \mathbb{R}^2, \\ \boldsymbol{U}_t(y) = \big(\nabla P_t^*(y) - y\big)^{\perp} & \text{in } [0,\infty) \times \mathbb{R}^2, \\ \rho_t = (\nabla P_t)_\sharp dx & \text{in } [0,\infty) \times \mathbb{R}^2, \\ P_0 = \bar{p} + |x|^2/2 & \text{on } \mathbb{R}^2. \end{cases} \tag{1.6}$$

2 Global existence for the dual SG system

The global existence of weak solutions to the dual problem (1.6) was obtained by Benamou and Brenier in [5]. The aim of this section is to review their result.

2.1 Preliminaries on transport equations

The system (1.6) is given by a transport equation for ρ_t where the vector field $\boldsymbol{U}_t$ is coupled to ρ_t via the relation $\rho_t = (\nabla P_t)_\sharp dx$. Notice that because $\boldsymbol{U}_t = (\boldsymbol{U}_t^1, \boldsymbol{U}_t^2)$ is the rotated gradient of the function $p_t^*(y) := P_t^*(y) - |y|^2/2$, it is divergence free: indeed,

$$\mathrm{div}\, \boldsymbol{U}_t = \partial_1 \boldsymbol{U}_t^1 + \partial_2 \boldsymbol{U}_t^2 = \partial_1 \partial_2 p_t^* - \partial_2 \partial_1 p_t^* = 0.$$

We now review some basic facts on linear transport equations with Lipschitz divergence-free vector fields. Since the dimension does not play any role, we work directly in $\mathbb{R}^n$.

Let $\boldsymbol{v}_t : \mathbb{R}^n \to \mathbb{R}^n$ be a time-dependent Lipschitz divergence-free vector field. Given a measure $\bar{\sigma}$ in $\mathbb{R}^n$, our goal is to solve the equation

$$\begin{cases} \partial_t \sigma_t + \mathrm{div}(\boldsymbol{v}_t \sigma_t) = 0 & \text{in } [0,\infty) \times \mathbb{R}^n, \\ \sigma_0 = \bar{\sigma} & \text{on } \mathbb{R}^n. \end{cases} \tag{2.1}$$

Notice that because $\boldsymbol{v}_t$ is divergence free, the above equation can also be rewritten as a standard transport equation:

$$\partial_t \sigma_t + \boldsymbol{v}_t \cdot \nabla \sigma_t = 0.$$

To find a solution, we first apply the classical Cauchy-Lipschitz theorem for ODEs in order to construct a flow for $\boldsymbol{v}_t$:

$$\begin{cases} \dot{Y}(t,y) = \boldsymbol{v}_t(Y(t,y)) \\ Y(0,y) = y, \end{cases} \tag{2.2}$$

and then we define $\sigma(t) := Y(t)_\#\bar\sigma$. Let us check that this definition provides a solution to (2.1): take $\varphi \in C_c^\infty(\mathbb{R}^n)$ and observe that

$$\begin{aligned}\frac{d}{dt}\int \varphi(y)\,d\sigma_t(y) &\overset{(1.3)}{=} \frac{d}{dt}\int \varphi(Y(t,y))\,d\bar\sigma(y)\\ &= \int \nabla\varphi(Y(t,y))\cdot \dot Y(t,y)\,d\bar\sigma(y)\\ &= \int \nabla\varphi(Y(t,y))\cdot \boldsymbol{v}_t(Y(t,y))\,d\bar\sigma(y)\\ &\overset{(1.3)}{=} \int \nabla\varphi(y)\cdot \boldsymbol{v}_t(y)\,d\sigma_t(y).\end{aligned}$$

By the arbitrariness of φ, this proves the validity of (2.1).

It is interesting to observe that the curve of measures $t \mapsto Y(t)_\#\bar\sigma$ is the unique solution of (2.1). A possible way to prove this is to consider σ_t an arbitrary solution of (2.1) and define $\hat\sigma_t := [Y(t)^{-1}]_\#\sigma_t$. Then a direct computation shows that

$$\frac{d}{dt}\int \varphi(y)\,d\hat\sigma_t(y) = 0 \qquad \forall \in C_c^\infty(\mathbb{R}^2), \tag{2.3}$$

therefore

$$[Y(t)^{-1}]_\#\sigma_t = \hat\sigma_t \overset{(2.3)}{=} \hat\sigma_0 = \bar\sigma \quad \Rightarrow \quad \sigma_t = Y(t)_\#\bar\sigma,$$

as desired.

We also notice that, if we assume that $\bar\sigma$ is not just a measure but a function, then we can give a more explicit formula for σ_t. Indeed the fact that $\operatorname{div}\boldsymbol{v}_t = 0$ implies that $\det\nabla Y(t) = 1$,[3] and the classical change of variable formula gives

$$\int \varphi(y)\,d\sigma_t(y) \overset{(1.3)}{=} \int \varphi(Y(t,y))\,\sigma(y)\,dy \overset{z=Y(t,y)}{=} \int \varphi(z)\,\bar\sigma\big(Y(t)^{-1}(z)\big)\,dz.$$

Since φ is arbitrary we deduce that σ_t is a function (and not just a measure) and it is given by the formula $\sigma_t = \bar\sigma\circ Y(t)^{-1}$, or equivalently

$$\sigma_t(Y(t,y)) = \bar\sigma(y) \qquad \forall\, y. \tag{2.5}$$

This implies in particular that any pointwise bound on $\bar\sigma$ is preserved in time, that is

$$\lambda \le \bar\sigma \le \Lambda \quad \Rightarrow \quad \lambda \le \sigma_t \le \Lambda \qquad \forall\, t. \tag{2.6}$$

[3] To show this, one differentiates (2.2) with respect to y and uses the classical identity

$$\frac{d}{d\varepsilon}_{|\varepsilon=0}\det(A+\varepsilon BA) = \operatorname{tr}(B)\,\det(A),$$

to get

$$\begin{cases}\frac{d}{dt}\big(\det\nabla Y(t,y)\big) = [\operatorname{div}\boldsymbol{v}_t(Y(t,y))]\,\big(\det\nabla Y(t,y)\big) = 0,\\ \det\nabla Y(0,y) = y.\end{cases} \tag{2.4}$$

2.2 Preliminaries on optimal transport

Let μ, ν denote two probability measures on $\mathbb{R}^n$. The optimal transport problem (with quadratic cost) consists in finding the "optimal" way to transport μ onto ν, given that the transportation cost to move a point from x to y is $|x-y|^2$. Hence one is naturally led to minimize

$$\int_{\mathbb{R}^n} |S(x) - x|^2 \, d\mu(x)$$

among all maps S which send μ onto ν, that is $S_\# \mu = \nu$. By a classical theorem of Brenier [9] existence and uniqueness of optimal maps always hold provided μ is absolutely continuous and both μ and ν have finite second moments. In addition, the optimality of the map is equivalent to the fact that T is the gradient of a convex function. This is summarized in the next theorem:

Theorem 2.1. *Let μ, ν be probability measures on $\mathbb{R}^n$ with $\mu = f\,dx$ and*

$$\int |x|^2 \, d\mu(x) + \int |y|^2 \, d\nu(y) < \infty.$$

Then:
(1) There exists a unique optimal transport map T.
(2) There exists a convex function $\phi : \mathbb{R}^n \to \mathbb{R}$ such that $T = \nabla\phi$.
(3) The fact that T is the gradient of a convex function is equivalent to its optimality. More precisely, if $\psi : \mathbb{R}^n \to \mathbb{R}$ is a convex function such that $\nabla\psi_\# \mu = \nu$ then $\nabla\psi$ is optimal and $T = \nabla\psi$. In addition, if $f > 0$ a.e. then $\psi = \phi + c$ for some additive constant $c \in \mathbb{R}$.

We now show the connection between optimal transport and the Monge-Ampère equation. Assume that both μ and ν are absolutely continuous, that is $\mu = f\,dx$ and $\nu = g\,dy$, let $\varphi \in C_c^\infty(\mathbb{R}^n)$, and suppose that $T = \nabla\phi$ is a smooth diffeomorphism. Then, using the definition of push-forward and the standard change of variable formula, we get

$$\int \varphi(T(x))\, f(x)\, dx \overset{(1.3)}{=} \int \varphi(y)\, g(y)\, dy \overset{y=T(x)}{=} \int \varphi(T(x))\, g(T(x)) \, \big|\det \nabla T(x)\big| \, dx.$$

By the arbitrariness of φ, this gives the Jacobian equation

$$f(x) = g(T(x)) \big|\det \nabla T(x))\big|,$$

that combined with the condition $T = \nabla\phi$ implies that ϕ solves the Monge-Ampère equation

$$\det(D^2\phi) = \frac{f}{g \circ \nabla\phi}. \tag{2.7}$$

Notice that the above computations are formal since we needed to assume a priori T to be smooth in order to write the above equation. Still, this fact is the starting point behind the regularity theory of optimal transport maps. We shall not enter into this but we refer to the survey paper [21] for more details.

Notice that in Section 1.2 we started from the Lebesgue measure on $\mathbb{R}^2$ and we sent it onto ρ_t using the gradient of the convex function P_t. If we could apply Theorem 2.1(3) above we would know that ∇P_t is the unique optimal map sending the Lebesgue measure onto ρ_t. However in our case we do not have probability measure but rather periodic measures on $\mathbb{R}^n$, hence Theorem 2.1 does not directly apply. However, since both the Lebesgue measure and ρ_t are probability measures on the torus (see Remark 1.1), we can apply [15] (see also [2, Theorem 2.1]) to obtain the following:

Theorem 2.2. *Let μ, ν be probability measures on $\mathbb{T}^2$, and assume that $\mu = f\,dx$ and that $f > 0$ a.e. Then there exists a unique (up to an additive constant) convex function $P : \mathbb{R}^2 \to \mathbb{R}$ such that $(\nabla P)_\sharp \mu = \nu$ and $P - |x|^2/2$ is periodic.*

2.3 Dual SG vs. 2-D Euler

Before entering into the proof of existence of solutions to (1.6), let us first make a parallel with the 2-dimensional Euler equations. Starting from the Euler system

$$\begin{cases} \partial_t \boldsymbol{u}_t + (\boldsymbol{u}_t \cdot \nabla)\boldsymbol{u}_t + \nabla p_t = 0 \\ \operatorname{div} \boldsymbol{u}_t = 0 \end{cases} \tag{2.8}$$

one can consider the curl of $\boldsymbol{u}_t$ given by $\omega_t := \partial_1 \boldsymbol{u}_t^2 - \partial_2 \boldsymbol{u}_t^1$. Then, by taking the curl of the first equation in (2.8), one finds that ω_t solves the equation

$$\begin{cases} \partial_t \omega_t + \operatorname{div}(\boldsymbol{u}_t\, \omega_t) = 0 \\ \omega_t = \operatorname{curl} \boldsymbol{u}_t \\ \operatorname{div} \boldsymbol{u}_t = 0. \end{cases} \tag{2.9}$$

Since $\boldsymbol{u}_t$ is divergence free, it follows that $\operatorname{curl} \boldsymbol{u}_t^\perp = 0$, hence $\boldsymbol{u}_t^\perp$ is the gradient of a function ψ_t, or equivalently $\boldsymbol{u}_t = -\nabla^\perp \psi_t$. Then, inserting this information inside the relation $\omega_t = \operatorname{curl} \boldsymbol{u}_t$ we deduce that $\omega_t = -\operatorname{curl} \nabla^\perp \psi_t = \Delta \psi_t$, and (2.9) rewrites as

$$\begin{cases} \partial_t \omega_t + \operatorname{div}(\boldsymbol{u}_t\, \omega_t) = 0 \\ \boldsymbol{u}_t = \nabla^\perp \psi_t \\ \Delta \psi_t = \omega_t \end{cases} \tag{2.10}$$

(see [18, Section 1.2.1] for more details).

If we now compare (2.10) and (1.6), we can notice that the two systems are very similar. More precisely, since the linearization of the determinant around the identity matrix gives the trace, we see that (1.6) is a nonlinear version of (2.10). This can be formalized in the following way (see [32, Section 6] for a rigorous result in this direction):

Exercise: Assume that $(\rho_t^\varepsilon, P_t^{*,\varepsilon})_{\varepsilon>0}$ is a family of solutions to (1.6) with

$$\rho_t^\varepsilon = 1 + \varepsilon\, \omega_{\varepsilon t} + o(\varepsilon), \qquad P_t^{*,\varepsilon} = |y|^2/2 + \varepsilon\, \psi_{\varepsilon t} + o(\varepsilon),$$

for some couple of smooth functions (ω_t, ψ_t). Then (ω_t, ψ_t) solve (2.10).

2.4 Global existence of weak solutions to (1.6)

In order to construct solutions to (1.6) one uses the following splitting method:
(1) Given ρ_0, we construct the vector field $\boldsymbol{U}_0$ using the optimal transport map sending ρ_0 to dx, and we use (a regularization of) it to let ρ_0 evolve over a time interval $[0,\varepsilon]$.
(2) Starting from ρ_ε, we construct $\boldsymbol{U}_\varepsilon$ as before and we use it to let ρ_ε evolve over the time interval $[\varepsilon, 2\varepsilon]$.
(3) Iterating this procedure, we obtain an approximate solution on $[0,\infty)$, and letting $\varepsilon \to 0$ produces the desired solutions.

We now describe in detail this construction.

2.4.1 Construction of approximate solutions

Assume that $\rho_0 := (x + \nabla \bar{p})_\# dx$ satisfies

$$\lambda \le \rho_0 \le \Lambda \tag{2.11}$$

for some constants $0 \le \lambda \le \Lambda$.[4] Since ρ_0 is absolutely continuous, we can apply Theorem 2.2 in order to find a convex function P_0^* whose gradient sends ρ_0 to dx, and we define

$$\boldsymbol{U}_0(y) := \big(\nabla P_0^*(y) - y\big)^\perp.$$

The idea is to fix $\varepsilon > 0$ a small time step and to solve the transport equation in (1.6) over the time interval $[0,\varepsilon]$ but with $\boldsymbol{U}_0$ in place of $\boldsymbol{U}_t$, using the theory described in Section 2.1. However, since $\boldsymbol{U}_0$ is not smooth, we shall first regularize it.[5] For this reason with introduce a regularization parameter $\delta > 0$.[6]

Hence, we fix a smooth convolution kernel $\chi \in C_c^\infty(\mathbb{R}^2)$ and, for $t \in [0,\varepsilon]$, we define

$$\boldsymbol{U}_t^{\varepsilon,\delta}(y) := \boldsymbol{U}_0 * \chi_\delta(y) = \int_{\mathbb{R}^2} \boldsymbol{U}_0(z)\, \chi_\delta(y-z)\, dz, \qquad \chi_\delta(y) := \frac{1}{\delta^2}\, \chi\Big(\frac{y}{\delta}\Big).$$

Notice that $\boldsymbol{U}_t^{\varepsilon,\delta} \in C^\infty(\mathbb{R}^2)$ and $\operatorname{div} \boldsymbol{U}_t^{\varepsilon,\delta} = (\operatorname{div} \boldsymbol{U}_0) * \chi_\delta = 0$, hence we can apply the theory discussed in Section 2.1 in the following way: we denote by $Y^{\varepsilon,\delta}(t)$ the flow of $\boldsymbol{U}_t^{\varepsilon,\delta}$ on $[0,\varepsilon]$, that is

$$\begin{cases} \dot{Y}^{\varepsilon,\delta}(t,y) = \boldsymbol{U}_t^{\varepsilon,\delta}(Y^{\varepsilon,\delta}(t,y)) \quad \text{for } t \in [0,\varepsilon], \\ Y^{\varepsilon,\delta}(0,y) = y, \end{cases}$$

and define

$$\rho_t^{\varepsilon,\delta} := Y^{\varepsilon,\delta}(t)_\# \rho_0 \qquad \forall\, t \in [0,\varepsilon].$$

[4] In this proof the lower bound on ρ_0 is not crucial and this is why we are allowing for $\lambda = 0$ as a possible value. However, instead of just writing $\rho_0 \le \Lambda$ we have decided to write (2.11) in order to emphasize that both the upper and the lower bound will be preserved along the flow.

[5] This regularization procedure is not strictly necessary, since in this situation one could also apply the theory of flows for divergence-free BV vector fields [1]. However, to keep the presentation elementary, we shall not use these advanced results.

[6] One could decide to choose $\delta = \varepsilon$ and to have only one small parameter. However, for clarity of the presentation, we prefer to keep these two parameter distinct.

Since $\boldsymbol{U}_t^{\varepsilon,\delta}$ is divergence-free, it follows from (2.11) and (2.6) that

$$\lambda \le \rho_t^{\varepsilon,\delta} \le \Lambda \qquad \forall\, t \in [0,\varepsilon].$$

We now "update" the vector field: we apply Theorem 2.2 to find a convex function $P_\varepsilon^{*,\varepsilon,\delta}$ whose gradient send $\rho_\varepsilon^{\varepsilon,\delta}$ to dx, we set

$$\boldsymbol{U}_t^{\varepsilon,\delta}(y) := \big(\nabla P_\varepsilon^{*,\varepsilon,\delta} - y\big)^\perp * \chi_\delta(y) \qquad \forall\, t \in [\varepsilon, 2\varepsilon],$$

we consider $Y^{\varepsilon,\delta}(t)$ the flow of $\boldsymbol{U}_t^{\varepsilon,\delta}$ on $[\varepsilon, 2\varepsilon]$,

$$\begin{cases} \dot{Y}^{\varepsilon,\delta}(t,y) = \boldsymbol{U}_t^{\varepsilon,\delta}(Y^{\varepsilon,\delta}(t,y)) & \text{for } t \in [\varepsilon, 2\varepsilon], \\ Y^{\varepsilon,\delta}(\varepsilon, y) = y, \end{cases}$$

and we define

$$\rho_t^{\varepsilon,\delta} := Y^{\varepsilon,\delta}(t)_\# \rho_\varepsilon^{\varepsilon,\delta} \qquad \forall\, t \in [\varepsilon, 2\varepsilon].$$

This allows us to update again the vector field on the time interval $[2\varepsilon, 3\varepsilon]$ using the optimal map from $\rho_{2\varepsilon}^{\varepsilon,\delta}$ to dx, and so on. Iterating this procedure infinitely many times and defining

$$P_t^{*,\varepsilon,\delta} := P_{k\varepsilon}^{*,\varepsilon,\delta} \qquad \text{for } t \in [k\varepsilon, (k+1)\varepsilon),\ k \in \mathbb{N},$$

we construct a family of densities $\{\rho_t^{\varepsilon,\delta}\}_{t\ge 0}$ and vector fields $\{\boldsymbol{U}_t^{\varepsilon,\delta}\}_{t\ge 0}$ satisfying

$$\begin{cases} \partial_t \rho_t^{\varepsilon,\delta} + \operatorname{div}(\boldsymbol{U}_t^{\varepsilon,\delta} \rho_t^{\varepsilon,\delta}) = 0 & \text{in } [0,\infty) \times \mathbb{R}^2, \\ \boldsymbol{U}_t^{\varepsilon,\delta} = \big(\nabla P_t^{*,\varepsilon,\delta} - y\big)^\perp * \chi_\delta & \text{in } [0,\infty) \times \mathbb{R}^2, \\ (\nabla P_t^{*,\varepsilon,\delta})_\# \rho_{k\varepsilon} = dx & \text{for } t \in [k\varepsilon, (k+1)\varepsilon), \\ \lambda \le \rho_t^{\varepsilon,\delta} \le \Lambda & \text{in } [0,\infty) \times \mathbb{R}^2, \\ \rho_0 = (x + \nabla \bar{p})_\# dx & \text{on } \mathbb{R}^2. \end{cases} \tag{2.12}$$

2.4.2 Taking the limit in the approximate system

Notice that, because $\nabla P_t^{*,\varepsilon,\delta}$ are optimal transport maps between probability densities on the torus, it is not difficult to show that

$$|\nabla P_t^{*,\varepsilon,\delta}(y) - y| \le \operatorname{diam}(\mathbb{T}^2) = \frac{\sqrt{2}}{2} \qquad \forall\, y \in \mathbb{R}^2 \tag{2.13}$$

(see [2, Theorem 2.1]), which implies that the vector fields $\boldsymbol{U}_t^{\varepsilon,\delta}$ are uniformly bounded. Hence, given an arbitrary sequence $\varepsilon, \delta \to 0$, up to extracting a subsequence we can find densities ρ_t and a bounded vector field $\boldsymbol{U}_t$ such that

$$\rho_t^{\varepsilon,\delta} \rightharpoonup^* \rho_t \qquad \text{in } L^\infty_{\mathrm{loc}}\big([0,\infty) \times \mathbb{R}^2\big),$$

$$\boldsymbol{U}_t^{\varepsilon,\delta} \rightharpoonup^* \boldsymbol{U}_t \qquad \text{in } L^\infty_{\mathrm{loc}}\big([0,\infty) \times \mathbb{R}^2; \mathbb{R}^2\big).$$

In particular, since $\lambda \le \rho_t^{\varepsilon,\delta} \le \Lambda$, it follows immediately that ρ_t satisfies

$$\lambda \le \rho_t \le \Lambda \qquad \text{for a.e. } t \ge 0.$$

• **Step 1: find the limit of $\boldsymbol{U}_t^{\varepsilon,\delta}\rho_t^{\varepsilon,\delta}$.** In order to take the limit into (2.12), the most difficult term to deal with is the product $\boldsymbol{U}_t^{\varepsilon,\delta}\rho_t^{\varepsilon,\delta}$ inside the divergence, since in general it is not true that under weak convergence this would converge to $\boldsymbol{U}_t\rho_t$. However in this case we can exploit extra regularity.

More precisely, since both $\rho_t^{\varepsilon,\delta}$ and $\boldsymbol{U}_t^{\varepsilon,\delta}$ are uniformly bounded, we see that for any smooth function $\psi : \mathbb{R}^2 \to \mathbb{R}$ it holds

$$\int \operatorname{div}(\boldsymbol{U}_t^{\varepsilon,\delta}\rho_t^{\varepsilon,\delta})\psi\, dy = -\int \boldsymbol{U}_t^{\varepsilon,\delta} \cdot \nabla\psi\, \rho_t^{\varepsilon,\delta}\, dy \le C\|\psi\|_{W^{1,1}(\mathbb{R}^2)}.$$

This means that $\operatorname{div}(\boldsymbol{U}_t^{\varepsilon,\delta}\rho_t^{\varepsilon,\delta})$ belongs to $[W^{1,1}(\mathbb{R}^2)]^*$ (the dual space of $W^{1,1}(\mathbb{R}^2)$) uniformly in time, which implies that

$$\partial_t\rho_t^{\varepsilon,\delta} = -\operatorname{div}(\boldsymbol{U}_t^{\varepsilon,\delta}\rho_t^{\varepsilon,\delta}) \in L^\infty\big([0,\infty), [W^{1,1}(\mathbb{R}^2)]^*\big) \subset L^p_{\rm loc}\big([0,\infty), [W^{1,q}_{\rm loc}(\mathbb{R}^2)]^*\big)$$

for any $p \in [1,\infty]$ and $q \ge 1$ (here we used that $W^{1,q}_{\rm loc}(\mathbb{R}^2) \subset W^{1,1}_{\rm loc}(\mathbb{R}^2)$ to get the opposite inclusion between dual spaces). Combining this regularity in time with the bound

$$\rho_t^{\varepsilon,\delta} \in L^\infty\big((0,\infty), L^\infty(\mathbb{R}^2)\big) \subset L^p_{\rm loc}\big([0,\infty), L^p_{\rm loc}(\mathbb{R}^2)\big),$$

by the Aubin-Lions Lemma (see for instance [34]) we deduce that

$$\text{the family } \rho_t^{\varepsilon,\delta} \text{ is precompact in } L^p_{\rm loc}\big([0,\infty), [W^{s,q}_{\rm loc}(\mathbb{R}^2)]^*\big) \text{ for any } p < \infty,\ q > 1,\ s > 0,$$

hence

$$\rho_t^{\varepsilon,\delta} \to \rho_t \qquad \text{in } L^p_{\rm loc}\big([0,\infty), [W^{s,q}_{\rm loc}(\mathbb{R}^2)]^*\big) \text{ for any } p < \infty,\ q > 1,\ s > 0. \tag{2.14}$$

In order to exploit this strong compactness we need to gain some regularity in space on $\boldsymbol{U}_t^{\varepsilon,\delta}$.

To this aim, observe that being $P_t^{*,\varepsilon,\delta}$ smooth convex functions, one can easily get an a-priori bound on their Hessian: since for a non-negative symmetric matrix the norm is controlled by the trace, using the divergence theorem and the uniform local Lipschitzianity of $P_t^{*,\varepsilon,\delta}$ (see (2.13)) we get

$$\int_{B_R} \|D^2 P_t^{*,\varepsilon,\delta}\|\, dy \le \int_{B_R} \Delta P_t^{*,\varepsilon,\delta}\, dy \le \int_{\partial B_R} |\nabla P_t^{*,\varepsilon,\delta}|\, dy \le C_R \qquad \forall\, R > 0. \tag{2.15}$$

By fractional Sobolev embeddings (see [7, Chapter 6]) we deduce that, uniformly with respect to ε and δ,

$$\boldsymbol{U}_t^{\varepsilon,\delta} \in L^\infty\big((0,\infty), W^{1,1}_{\rm loc}(\mathbb{R}^2)\big) \subset L^\infty\big((0,\infty), W^{s,q}_{\rm loc}(\mathbb{R}^2)\big)$$

for all $s \in (0,1)$ and $1 \le q < \frac{2}{1+s}$. In particular, choosing for instance $s = 1/2$ and $q = 5/4$ we deduce that

$$\boldsymbol{U}_t^{\varepsilon,\delta} \rightharpoonup^* \boldsymbol{U}_t \qquad \text{in } L^\infty\big((0,\infty), W^{1/2,5/4}_{\rm loc}(\mathbb{R}^2)\big),$$

that combined with (2.14) with $s = 1/2$ and $q = 5/4$ implies that

$$\boldsymbol{U}_t^{\varepsilon,\delta} \rho_t^{\varepsilon,\delta} \rightharpoonup^* \boldsymbol{U}_t \rho_t \qquad \text{in } L^\infty_{\text{loc}}\big((0,\infty) \times \mathbb{R}^2; \mathbb{R}^2\big).$$

This allows us to pass to the limit in the transport equation in the distributional sense and deduce that

$$\partial_t \rho_t + \operatorname{div}(\boldsymbol{U}_t \rho_t) = 0 \qquad \text{in } (0,\infty) \times \mathbb{R}^2.$$

• **Step 2: show that $\boldsymbol{U}_t = (\nabla P_t^* - y)^\perp$.** To conclude the proof we need so show that if P_t^* is the convex function sending ρ_t onto dx (see Theorem 2.2) then

$$\boldsymbol{U}_t = (\nabla P_t^* - y)^\perp \qquad \text{for a.e. } t \geq 0. \tag{2.16}$$

To this aim notice that (2.14) implies that, up to a subsequence,

$$\rho_t^{\varepsilon,\delta} \to \rho_t \qquad \text{in } [W^{s,q}_{\text{loc}}(\mathbb{R}^2)]^* \text{ for a.e. } t \geq 0,$$

hence, being $\rho_t^{\varepsilon,\delta}$ uniformly bounded in L^∞, we also deduce that

$$\rho_t^{\varepsilon,\delta} \rightharpoonup^* \rho_t \qquad \text{in } L^\infty(\mathbb{R}^2) \text{ for a.e. } t \geq 0.$$

By stability of optimal transport maps (see for instance [36, Corollary 5.23]) it follows that[7]

$$\nabla P_t^{*,\varepsilon,\delta} \to \nabla P_t^* \qquad \text{in } L^1_{\text{loc}}(\mathbb{R}^2) \text{ for a.e. } t \geq 0.$$

Recalling that $\boldsymbol{U}_t^{\varepsilon,\delta} = (\nabla P_t^{*,\varepsilon,\delta} - y)^\perp * \chi_\delta$ we deduce that

$$\boldsymbol{U}_t^{\varepsilon,\delta} \to (\nabla P_t^* - y)^\perp \qquad \text{in } L^1_{\text{loc}}(\mathbb{R}^2) \text{ for a.e. } t \geq 0,$$

which shows the validity of (2.16) and concludes the proof of the existence of weak solutions.

Notice that, as a consequence of (2.13), the uniform bound

$$|\nabla P_t^*(y) - y| \leq \frac{\sqrt{2}}{2} \tag{2.17}$$

holds. This will be useful in the sequel.

3 Back from dual SG to SG

In the previous section we have constructed a weak solution (ρ_t, P_t^*) to the dual system (1.6). Also, we have seen that if $\rho_0 := (x + \nabla\bar{p})_\# dx$ satisfies $\lambda \leq \rho_0 \leq \Lambda$ then these bounds are propagated in time, that is

$$\lambda \leq \rho_t \leq \Lambda \qquad \text{for a.e. } t \geq 0. \tag{3.1}$$

In this section we shall assume that $\lambda > 0$.

[7] Actually, if one assumes that $\lambda > 0$ (that is the densities $\rho_t^{\varepsilon,\delta}$ are uniformly bounded away from zero and infinity) then the convergence of $\nabla P_t^{*,\varepsilon,\delta}$ to ∇P_t^* holds even in $W^{1,1}_{\text{loc}}(\mathbb{R}^2)$, see [20].

3.1 A formula for $(p_t, \boldsymbol{u}_t)$

We want to use the solution (ρ_t, P_t^*) to construct a couple $(p_t, \boldsymbol{u}_t)$ solving the original SG systems (1.2).

3.1.1 Construction of p_t

Recalling the procedure used to go from p_t to P_t^* (adding $|x|^2/2$ to p_t and taking a Legendre transform), it is easy to perform the inverse construction and define p_t from P_t^*: indeed, if we define[8]

$$P_t(x) := \sup_{y\in\mathbb{R}^2} \{y \cdot x - P_t^*(y)\} \tag{3.2}$$

and set

$$p_t(x) := P_t(x) - \frac{|x|^2}{2}, \tag{3.3}$$

thanks to the periodicity of $P_t^* - |y|^2/2$ it is easy to check that p_t is periodic too. Hence, constructing p_t given P_t^* is relatively simple.

3.1.2 Construction of $\boldsymbol{u}_t$

More complicated is the formula for $\boldsymbol{u}_t$. Let us start from (1.2). From the first equation and the fact that D^2P_t is a symmetric matrix, we get

$$D^2P_t \cdot \boldsymbol{u}_t = -\partial_t \nabla P_t - (\nabla P_t - x)^\perp. \tag{3.4}$$

In order to invert D^2P_t, we differentiate (1.5) with respect to x to find that

$$D^2P_t^*(\nabla P_t)\, D^2P_t = \mathrm{Id}, \tag{3.5}$$

while differentiating (1.5) with respect to t gives

$$[\partial_t \nabla P_t^*](\nabla P_t) + D^2P_t^*(\nabla P_t) \cdot \partial_t \nabla P_t = 0 \tag{3.6}$$

Hence, thanks to (3.5), multiplying both sides of (3.4) by $D^2P_t^*(\nabla P_t)$ we get

$$\boldsymbol{u}_t = -D^2P_t^*(\nabla P_t) \cdot \partial_t \nabla P_t - D^2P_t^*(\nabla P_t) \cdot (\nabla P_t - x)^\perp,$$

that combined with (3.6) gives

$$\boldsymbol{u}_t = [\partial_t \nabla P_t^*](\nabla P_t) - D^2P_t^*(\nabla P_t) \cdot (\nabla P_t - x)^\perp. \tag{3.7}$$

Hence we have found an expression of $\boldsymbol{u}_t$ in terms of derivatives of P_t^* and its Legendre transform. However the problem is to give a meaning to such terms.

First of all one may ask what is $D^2P_t^*(\nabla P_t)$. Notice that being P_t^* a convex function, *a priori* $D^2P_t^*$ is only a matrix-valued measure, thus it is not clear what $D^2P_t^*(\nabla P_t)$ means. However, if we

[8] Recall that the Legendre transform is an involution on convex functions, that is, if $\phi : \mathbb{R}^n \to \mathbb{R}$ is convex then $(\phi^*)^* = \phi$.

remember that $(\nabla P_t^*)_\# \rho_t = dx$, it follows by the discussion in Section 2.2 (see in particular (2.7)) that

$$\det(D^2 P_t^*) = \rho_t. \tag{3.8}$$

Hence, recalling (3.1) we deduce that

$$\lambda \le \det(D^2 P_t^*) \le \Lambda.$$

As we shall see in Section 4.1 below, this condition implies that

$$P_t^* \in W^{2,1+\gamma}_{\mathrm{loc}}(\mathbb{R}^2) \quad \text{for some } \gamma = \gamma(n,\lambda,\Lambda) > 0. \tag{3.9}$$

We claim that this estimate allows us to give a meaning to $D^2 P_t^*(\nabla P_t)$ and prove that

$$D^2 P_t^*(\nabla P_t) \in L^\infty\big((0,\infty), L^{1+\gamma}_{\mathrm{loc}}(\mathbb{R}^2)\big).$$

Indeed, since $(\nabla P_t^*)_\# \rho_t = dx$, it follows from (1.5) that $(\nabla P_t)_\# dx = \rho_t$. Also, since p_t is periodic we see that $\nabla P_t = x + \nabla p_t$ is a bounded perturbation of the identity, hence there exists $C > 0$ such that

$$\nabla P_t(B_R) \subset B_{R+C} \qquad \forall\, R > 0.$$

These two facts imply that

$$\int_{B_R} \|D^2 P_t^*(\nabla P_t)\|^{1+\gamma}\, dx \overset{(1.3)}{=} \int_{\nabla P_t(B_R)} \|D^2 P_t^*\|^{1+\gamma} \rho_t(y)\, dy \overset{(3.1)}{\le} \Lambda \int_{B_{R+C}} \|D^2 P_t^*\|^{1+\gamma}\, dy \overset{(3.9)}{<} \infty$$

for all $R > 0$, hence

$$D^2 P_t^*(\nabla P_t) \in L^\infty\big((0,\infty), L^{1+\gamma}_{\mathrm{loc}}(\mathbb{R}^2)\big).$$

Recalling that $(\nabla P_t - x)^\perp$ is globally bounded (see (2.17)), we deduce that

$$D^2 P_t^*(\nabla P_t) \cdot (\nabla P_t - x)^\perp \in L^\infty\big((0,\infty), L^{1+\gamma}_{\mathrm{loc}}(\mathbb{R}^2)\big),$$

so the last term in (3.7) is a well defined function.

Concerning the term $[\partial_t \nabla P_t^*](\nabla P_t)$, as explained in Section 4.2 one can show that

$$\partial_t \nabla P_t^* \in L^{1+\kappa}_{\mathrm{loc}} \qquad \text{for } \kappa = \frac{\gamma}{2+\gamma} > 0, \tag{3.10}$$

and arguing as above one deduces that

$$[\partial_t \nabla P_t^*](\nabla P_t) \in L^\infty\big((0,\infty), L^{1+\kappa}_{\mathrm{loc}}(\mathbb{R}^2)\big).$$

In conclusion we have seen that, using (3.9) and (3.10), the formula (3.7) defines a function $\boldsymbol{u}_t \in L^\infty\big((0,\infty), L^{1+\kappa}_{\mathrm{loc}}(\mathbb{R}^2)\big)$, which is easily seen to be periodic.

Hence, modulo the validity of (3.9) and (3.10), we have constructed a couple of functions $(p_t, \boldsymbol{u}_t)$ that we expect to solve (1.1). In the next section we shall see that the functions $(p_t, \boldsymbol{u}_t)$ defined in (3.3) and (3.7) are indeed solutions of (1.1), and then in Section 4 we will prove both (3.9) and (3.10).

3.2 $(p_t, \boldsymbol{u}_t)$ solves the semigeostrophic system

In order to prove that $(p_t, \boldsymbol{u}_t)$ is a distributional solution of (1.1) we need to find some suitable test functions to use in (1.6).

More precisely, we first write (1.6) in distributional form:

$$\int \int_{\mathbb{T}^2} \Big\{ \partial_t \varphi_t(x) + \nabla \varphi_t(x) \cdot \boldsymbol{U}_t(x) \Big\} \, \rho_t(x) \, dx \, dt \, dx = 0 \tag{3.11}$$

for every $\varphi \in W^{1,1}((0,\infty) \times \mathbb{R}^2)$ periodic in the space variable.

We now take $\phi \in C_c^\infty((0,\infty) \times \mathbb{R}^2)$ a function periodic in space, and we consider the test function $\varphi : (0,\infty) \times \mathbb{R}^2 \to \mathbb{R}^2$ defined as

$$\varphi_t(y) := J\,(y - \nabla P_t^*(y))\, \phi_t\big(\nabla P_t^*(y)\big), \tag{3.12}$$

where J denotes the matrix corresponding to the rotation by $\pi/2$, that is

$$J := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Notice that $Jv = -v^\perp$ for any $v \in \mathbb{R}^2$, hence φ_t can be equivalently written as

$$\varphi_t := (\nabla P_t^* - y)^\perp \, \phi_t(\nabla P_t^*)$$

We compute the derivatives of φ:

$$\begin{cases} \partial_t \varphi_t = [\partial_t \nabla P_t^*]^\perp \, \phi_t(\nabla P_t^*) + (\nabla P_t^* - y)^\perp \, \partial_t \phi_t(\nabla P_t^*) + (y - \nabla P_t^*)^\perp \big[\nabla \phi_t(\nabla P_t^*) \cdot \partial_t \nabla P_t^*\big], \\ \nabla \varphi_t = J\big(Id - D^2 P_t^*\big)\, \phi_t(\nabla P_t^*) + (\nabla P_t^* - y)^\perp \otimes \big(\nabla \phi_t(\nabla P_t^*) \cdot D^2 P_t^*\big). \end{cases} \tag{3.13}$$

Since $\boldsymbol{U}_t = (\nabla P_t^* - y)^\perp$ and $(\nabla P_t^*)_\sharp \rho_t = dx$, recalling (1.5) we can use (3.13) to rewrite (3.11) as

$$\begin{aligned} 0 &= \int_0^\infty \int_{\mathbb{T}^2} \Big\{ \partial_t \varphi_t + \nabla \varphi_t \cdot \boldsymbol{U}_t \Big\} \rho_t(y) \, dy \, dt \\ &= \int_0^\infty \int_{\mathbb{T}^2} \Big\{ [\partial_t \nabla P_t^*]^\perp (\nabla P_t) \, \phi_t + (x - \nabla P_t)^\perp \, \partial_t \phi_t + (x - \nabla P_t)^\perp \big[\nabla \phi_t \cdot [\partial_t \nabla P_t^*](\nabla P_t)\big] \\ &\qquad + \Big[J\big(Id - D^2 P_t^*(\nabla P_t)\big)\, \phi_t + (x - \nabla P_t)^\perp \otimes \big(\nabla \phi_t \cdot D^2 P_t^*(\nabla P_t)\big)\Big] (x - \nabla P_t)^\perp \Big\} \, dx \, dt. \end{aligned}$$

Taking into account the formula (3.7) for $\boldsymbol{u}_t$, after rearranging the terms the above expression yields

$$0 = \int_0^\infty \int_{\mathbb{T}^2} \Big\{ -\nabla^\perp p_t \left(\partial_t \phi_t + \boldsymbol{u}_t \cdot \nabla \phi_t \right) + \left(-\nabla p_t + \boldsymbol{u}_t^\perp \right) \phi_t \Big\} \, dx \, dt,$$

hence $(p_t, \boldsymbol{u}_t)$ solve the first equation in (1.1). The fact that $\boldsymbol{u}_t$ is divergence free is obtained in a similar way, using the test function

$$\varphi_t := \phi(t) \, \psi(\nabla P_t{}^*)$$

where $\phi \in C_c^\infty((0,\infty))$, and $\psi \in C_c^\infty(\mathbb{R}^2)$ is periodic.

This shows that $(p_t, \boldsymbol{u}_t)$ is a distributional solution of (1.1), and we obtain the following result (see [2, Theorem 1.2]):

Theorem 3.1. *Let $\bar{p} : \mathbb{R}^2 \to \mathbb{R}$ be a periodic function such that $\bar{p}(x) + |x|^2/2$ is convex, and assume that the measure $\bar{\rho} := (Id + \nabla \bar{p})_\sharp dx$ is absolutely continuous with respect to the Lebesgue measure and satisfies $0 < \lambda \le \bar{\rho} \le \Lambda$.*

Let (ρ_t, P_t^*) be a solution of (1.6) *starting from $\bar{\rho}$ satisfying $0 < \lambda \le \rho_t \le \Lambda$, and let P_t be as in* (3.2)*. Then the couple $(p_t, \boldsymbol{u}_t)$ defined in* (3.3) *and* (3.7) *is a distributional solution of* (1.1)*.*

Although the vector field $\boldsymbol{u}_t$ provided by the previous theorem is only $L^{1+\kappa}_{\mathrm{loc}}$, in [2] the authors showed how to associate to it a measure-preserving Lagrangian flow. In particular, this allowed them to recover (in the particular case of the 2-dimensional periodic setting) the result of Cullen and Feldman [17] on the existence of Lagrangian solutions to the SG equations in physical space (see also [24, 25] for some recent results on the existence of weak Lagrangian solutions).

As shown in [3], the above result can also be generalized to bounded convex domain $\Omega \subset \mathbb{R}^3$. However this extension presents several additional difficulties. Indeed, first of all in 3-dimensions the SG system becomes much less symmetric compared to its 2-dimensional counterpart, because the action of Coriolis force regards only the first and the second space components. Moreover, even considering regular initial data and velocities, several arguments in the proofs require a finer regularization scheme. Still, under suitable assumptions on the initial data one can prove global existence of distributional solutions (see [3] for more details).

4 Regularity estimates for the Monge-Ampère equation

The aim of this section is to give a proof of the key estimates (3.9) and (3.10) used in the previous section to obtain global existence of distributional solutions to (1.1).

We shall first prove (3.9), and then show how (3.9) combined with (1.6) yields (3.10).

4.1 Sobolev regularity for Monge-Ampère: proof of (3.9)

Our goal is to show that, given $0 < \lambda \le \Lambda$, solutions to

$$\begin{cases} \lambda \le \det(D^2\phi) \le \Lambda \\ \phi \text{ convex} \\ \phi - |x|^2/2 \text{ periodic} \end{cases} \tag{4.1}$$

belong to $W^{2,1+\gamma}_{\mathrm{loc}}$ for some $\gamma > 0$ [19, 22, 33]. This result is valid in any dimension and restricting to dimension 2 would not simplify the proof. Also, since we want to prove an a-priori estimate on solutions to (4.1), one can assume that ϕ is smooth. Hence, we shall assume that $\phi : \mathbb{R}^n \to \mathbb{R}$ is a C^2 solution of (4.1) and we will show that

$$\int_{[0,1]^n} \|D^2\phi\|^{1+\gamma} \le C,$$

for some constant C depending only on n, λ, Λ. (From now on, any constant which depends only on n, λ, Λ will be called *universal*).

We shall mainly follow the arguments in [19], except for Step 2 in Section 4.1.4 which is inspired by [33].

4.1.1 Sections and normalized solutions

An important role in the regularity theory of Monge-Ampère is played by the sections of the function ϕ: given $x \in \mathbb{R}^n$ and $t > 0$, we define the section centered at x with height t as

$$S(x,t) := \{y \in \Omega \,:\, u(y) \le u(x) + \nabla u(x) \cdot (y - x) + t\}. \tag{4.2}$$

Moreover, given $\tau > 0$, we use the notation $\tau S(x, p, t)$ to denote the dilation of $S(x, p, t)$ by a factor τ with respect to x, that is

$$\tau S(x,t) := \left\{ y \in \mathbb{R}^n \,:\, x + \frac{y - x}{\tau} \in S(x,t) \right\}. \tag{4.3}$$

Notice that, because $\phi - |x|^2/2$ is periodic, ϕ has quadratic growth at infinity. In particular its sections $S(x,t)$ are all bounded.

We say that an open bounded convex set $Z \subset \mathbb{R}^n$ is *normalized* if

$$B(0,1) \subset Z \subset B(0,n).$$

By John's Lemma [30], for every open bounded convex set there exists an (invertible) orientation preserving affine transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ such that $T(Z)$ is normalized.

Notice that in the sequel we are not going to notationally distinguish between an affine transformation and its linear part, since it will always be clear to what we are referring to. In particular, we will use the notation

$$\|T\| := \sup_{|v|=1} |Av|, \qquad Tx = Ax + b.$$

One useful property which we will use is the following identity: if we denote by T^* the adjoint of T, then

$$\|T^*T\| = \|T^*\|\|T\|. \tag{4.4}$$

(This can be easily proved using the polar decomposition of matrices.)

Given a section $S(x,t)$, we can consider T an affine transformation which normalizes $S(x,t)$ and define the function

$$v(z) := (\det T)^{2/n} \left[u(T^{-1}z) - u(x) - \nabla u(x) \cdot (T^{-1}z - x) - t \right]. \tag{4.5}$$

Then it is immediate to check that v solves

$$\begin{cases} \lambda \le \det D^2 v \le \Lambda & \text{in } Z, \\ v = 0 & \text{on } \partial Z, \end{cases} \tag{4.6}$$

with $Z := T(S(x,t))$ normalized. We are going to call v a *normalized solution*.

As shown in [12] and [29] (see also [28, Chapter 3]), sections of solution of (4.1) satisfy strong geometric properties. We briefly recall here the ones that we are going to use:[9]

[9] Usually all these properties are stated for small sections (say, when $t \le \rho$ for some universal ρ). However, since in our case ϕ is a global solution which has quadratic growth at infinity, it is immediate to check that all the properties are true when t is large.

Proposition 4.1. *Let ϕ be a solution of* (4.1)*. Then the following properties hold:*

(i) There exists a universal constant $\beta \in (0,1)$ such that

$$\frac{1}{2}S(x,t) \subset S(x,t/2) \subset \beta S(x,t) \qquad \forall\, x \in \mathbb{R}^n,\, t > 0.$$

(ii) There exists a universal constant $\theta > 1$ such that

$$S(x,t) \cap S(y,t) \neq \emptyset \quad \Rightarrow \quad S(y,t) \subset S(x,\theta t) \qquad \forall\, x,y \in \mathbb{R}^n,\, t > 0.$$

(iii) There exists a universal constant $K > 1$ such that such that

$$\frac{t^{n/2}}{K} \leq |S(x,t)| \leq K\, t^{n/2} \qquad \forall\, x \in \mathbb{R}^n,\, t > 0.$$

(iv) $\bigcap_{t>0} S(x,t) = \{x\}$.

4.1.2 A preliminary estimate for normalized solutions

In this section we consider v a solution of (4.6) with $B_r \subset Z \subset B_R$ and we prove the following classical lemma due to Alexandrov:

Lemma 4.2. *Assume that v a solution of* (4.6) *with $B_r \subset Z \subset B_R$ for some universal radii $0 < r \leq R$. There exist two universal constants $c_1, c_2 > 0$ such that*

$$c_1 \leq \left|\inf_Z v\right| \leq c_2, \tag{4.7}$$

Proof. Set

$$g_-(z) := \frac{\lambda^{1/n}}{4}\big(|z|^2 - r^2\big).$$

We claim that $v \leq g_-$. Indeed, if not, let $c > 0$ be the smallest constant such that $v - c \leq g_-$ in Z, so that

$$v - c \leq g_- \quad \text{in } Z, \qquad v(\bar z) - c = g_-(\bar z) \quad \text{for some } \bar z \in \overline{Z}.$$

Notice that because $g_- \leq 0$ on ∂Z (since $B_r \subset Z$), the contact point $\bar z$ must be in the interior of Z. Hence, since the functions $g_- - (v - c)$ attains a local minimum at $\bar z$, its Hessian at $\bar z$ is non-negative definite, thus

$$D^2 g_-(\bar z) \geq D^2(v-c)(\bar z) = D^2 v(\bar z) \geq 0$$

which implies that

$$\frac{\lambda}{2^n} = \det\big(D^2 g_-(\bar z)\big) \geq \det\big(D^2 v(\bar z)\big) \geq \lambda,$$

a contradiction.

Set now

$$g_+(z) := \Lambda^{1/n}\big(|z|^2 - R^2\big).$$

A completely analogous argument based on the fact that $\det(D^2 g_+) = 2^n \Lambda$ and $g_+ \geq 0$ on ∂Z (since $Z \subset B_R$) shows that $g_+ \leq v$ in Z.

This proves that

$$g_+ \leq v \leq g_- \qquad \text{in } Z,$$

and the result follows. □

4.1.3 Two key estimates on the size of the Hessian

The following two lemmas are at the core of the proof of the $W^{2,1+\gamma}_{\mathrm{loc}}$ regularity. The first lemma estimates the L^1-size of $\|D^2\phi\|$ on a section $S(x,t)$, while the second one says that on a large fraction of points in $S(x,t)$ the value of $\|D^2\phi\|$ is comparable to its average.

Lemma 4.3. *Fix $x \in \mathbb{R}^n$, $t > 0$, and let T be the affine map which normalizes $S(x,t)$. Then there exists a positive universal constant C_1 such that*

$$\fint_{S(x,t)} \|D^2\phi\| \le C_1 \frac{\|T\|\|T^*\|}{(\det T)^{2/n}} \tag{4.8}$$

Proof. Consider the function $v : \mathbb{R}^n \to \mathbb{R}$ defined as in (4.5), and notice that

$$D^2 v(z) = (\det T)^{2/n} \left[(T^{-1})^* D^2\phi(T^{-1}z)T^{-1}\right], \tag{4.9}$$

and

$$\begin{cases} \lambda \le \det D^2 v \le \Lambda & \text{in } T(S(x,2t)), \\ v = \text{const.} & \text{on } \partial\big(T(S(x,2t))\big). \end{cases} \tag{4.10}$$

Although the convex set $T(S(x,2t))$ is not normalized in the sense defined before, it is almost so: indeed, since T normalizes $S(x,t)$, we have that

$$B_1 \subset T(S(x,2t)). \tag{4.11}$$

Also, because

$$|S(x,2t)| \le K(2t)^{n/2} \le 2^{n/2}K^2|S(x,t)|$$

(by Propoposition 4.1(iii)) and $T(S(x,t))$ is normalized, it follows that

$$|T(S(x,2t))| \le 2^{n/2}K^2|T(S(x,t))| \le 2^{n/2}K^2|B_n| =: C_0,$$

where C_0 is universal. Since $T(S(x,2t))$ is convex, the above estimate on its volume combined with (4.11) implies that

$$B_1 \subset T(S(x,2t) \subset B_R. \tag{4.12}$$

for some universal radius R. Hence, it follows from (4.10) and Lemma 4.2 that

$$\operatorname{osc}_{T(S(x,2t))} v \le c', \tag{4.13}$$

with c' universal.

Since v is convex, the size of its gradient is controlled by its oscillation in a slightly larger domain (see for instance [28, Lemma 3.2.1]), thus it follows from Proposition 4.1(i) and (4.13) that

$$\sup_{T(S(x,t))} |\nabla v| \le \sup_{\beta T(S(x,2t))} |\nabla v| \le \frac{\operatorname{osc}_{T(S(x,2t))} v}{\operatorname{dist}\Big(\beta T(S(x,2t)),\, \partial\big(T(S(x,2t))\big)\Big)} \le c'' \tag{4.14}$$

for some universal constant c''. Moreover, since $T(S(x,t))$ is a normalized convex set, it holds

$$|T(S(x,t))| \ge c_n \qquad \mathcal{H}^{n-1}\big(\partial T(S(x,t))\big) \le C_n, \tag{4.15}$$

where $c_n, C_n > 0$ are dimensional constants. Finally, since $D^2v(y)$ is non-negative definite (by the convexity of v) its norm is controlled by its trace, that is

$$\|D^2v(z)\| \leq \Delta v(z). \tag{4.16}$$

Thus, combining all these informations together we get

$$\begin{aligned}\fint_{T(S(x,t))} \|D^2v(z)\|\,dz &\overset{(4.16)}{\leq} \fint_{T(S(x,t))} \Delta v(z)\,dz \\ &= \frac{1}{|T(S(x,t))|}\int_{\partial T(S(x,t))} \nabla v(z)\cdot\nu\, d\mathcal{H}^{n-1}(z) \\ &\overset{(4.15)}{\leq} \frac{C_n}{c_n}\Big(\sup_{T(S(x,t))}|\nabla v|\Big) \overset{(4.14)}{\leq} C_1,\end{aligned} \tag{4.17}$$

that together with (4.9) gives

$$\begin{aligned}\fint_{S(x,t)} \|D^2\phi(y)\|\,dy &= \frac{1}{(\det T)^{2/n}}\fint_{S(x,t)} \|T^*D^2v(Ty)T\|\,dy \\ &\leq \frac{\|T^*\|\|T\|}{(\det T)^{2/n}}\fint_{T(S(x,t))}\|D^2v(z)\|\,dz \leq C_1\,\frac{\|T^*\|\|T\|}{(\det T)^{2/n}},\end{aligned}$$

concluding the proof. □

Lemma 4.4. *Fix $x \in \mathbb{R}^n$, $t > 0$, and let T be the affine map which normalizes $S(x,t)$. Then there exists a universal positive constant c_1 and a Borel set $A(x,t) \subset S(x,t)$, such that*

$$\frac{|A(x,t)\cap S(x,t)|}{|S(x,t)|} \geq \frac{1}{2} \tag{4.18}$$

and

$$\|D^2\phi(y)\| \geq c_1\,\frac{\|T\|\|T^*\|}{(\det T)^{2/n}} \qquad \forall\, y \in A(x,t). \tag{4.19}$$

Proof. Let $v : \mathbb{R}^n \to \mathbb{R}$ defined as in (4.5), and recall that

$$\fint_{T(S(x,t))} \|D^2v(z)\|\,dz \leq C_1$$

for some universal constant C_1 (see (4.17)). Set

$$E := \{z \in T(S(x,t))\ :\ \|D^2v(z)\| \geq 2C_1\}.$$

Then

$$2C_1\frac{|E|}{|T(S(x,t))|} \leq \frac{1}{|T(S(x,t))|}\int_E \|D^2v(z)\|\,dz \leq \fint_{T(S(x,t))}\|D^2v(z)\|\,dz \leq C_1,$$

which implies that

$$|E| \leq \frac{1}{2}|T(S(x,t))|.$$

Define $F := T(S(x,t)) \setminus E$ and notice that

$$\frac{|F|}{|T(S(x,t))|} \geq \frac{1}{2} \tag{4.20}$$

and (by (4.6) and the definition of E)

$$\begin{cases} \|D^2 v\| \leq 2C_1 & \text{inside } F, \\ \det(D^2 v) \geq \lambda. \end{cases}$$

If we denote by $\alpha_1 \leq \ldots \leq \alpha_n$ the eigenvalues of $D^2 v$, the first information tells us that $\alpha_n \leq 2C_1$, while the second one that $\prod_i \alpha_i \geq \lambda$, from which it follows that

$$\alpha_1 \geq \frac{\lambda}{\prod_{i=2}^n \alpha_i} \geq \frac{\lambda}{(2C_1)^{n-1}} =: c_1,$$

therefore

$$c_1 \mathrm{Id} \leq D^2 v \leq 2C_1 \mathrm{Id} \qquad \text{inside } F. \tag{4.21}$$

Recalling the definition of v (see (4.5)) this implies that

$$D^2\phi(y) = \frac{T^* D^2 v(Ty) T}{(\det T)^{2/n}} \geq c_1 \frac{TT^*}{(\det T)^{2/n}} \qquad \forall\, y \in A := T^{-1}(F),$$

so in particular

$$\|D^2\phi(y)\| \geq c_1 \frac{\|T\|\|T^*\|}{(\det T)^{2/n}} \qquad \forall\, y \in A.$$

Finally, thanks to (4.20) we get

$$\frac{|A|}{|S(x,t)|} = \frac{|T(A)|}{|T(S(x,t))|} = \frac{|F|}{|T(S(x,t))|} \geq \frac{1}{2},$$

concluding the proof. □

4.1.4 Harmonic analysis related to sections and the $W^{2,1+\gamma}_{\mathrm{loc}}$ regularity

In this section we show how Lemmas 4.3 and 4.4 can be combined to obtain the desired result. Since the covering argument is slightly technical and may hide the ideas behind the proof, we prefer to give a formal argument and refer to the papers [19, 22, 33] for more details (see also [26]).

The basic idea behind the proof is that we can think of a section $S(x,t)$ as a "ball of radius t centered at x", and the properties stated in Proposition 4.1 ensure that sections are suitable objects to do harmonic analysis. Indeed it is possible to show that a Vitali Covering Lemma holds in this context (see for instance [22]), and that many standard quantities in harmonic analysis still enjoy all the properties that we are used to have in $\mathbb{R}^n$.

For instance, to $\|D^2\phi\|$ we can associated a "maximal function" using the sections:

$$\mathcal{M}(x) := \sup_{t>0} \fint_{S(x,t)} \|D^2\phi(y)\|\, dy \qquad \forall\, x \in \mathbb{R}^n.$$

Noticing that $D^2\phi$ is periodic, in order to deal with sets of finite volume we shall see both $D^2\phi$ and $\mathcal{M}$ as functions on the torus $\mathbb{T}^n$. In the same way, also the sections will be seen as subsets of $\mathbb{T}^n$ by considering the canonical projection $\pi : \mathbb{R}^n \to \mathbb{T}^n$.

The fact that sections behave like usual balls allows us to obtain the validity of a classical fact in harmonic analysis, that is that the L^1 norm of $\|D^2\phi\|$ on a super level sets $\{\|D^2\phi\| \geq \sigma\}$ is controlled by the measure where $\mathcal{M}$ is above σ (up to a universal constant). More precisely, by applying [35, Chapter 1, Section 4, Theorem 2] and [35, Chapter 1, Section 8.14], we deduce that the following holds: there exist universal constants $K, \sigma_0 > 0$ such that, for any $\sigma \geq \sigma_0$,

$$\int_{\{\|D^2\phi\|\geq\sigma\}} \|D^2\phi(y)\|\, dy \leq K\sigma \left|\left\{\mathcal{M} \geq \tfrac{\sigma}{K}\right\}\right|. \tag{4.22}$$

Our goal is to combine this estimate with Lemmas 4.3 and 4.4 to show that $D^2\phi \in W^{2,1+\gamma}_{\mathrm{loc}}$.

• **Step 1: replace $\mathcal{M}$ with $\|D^2\phi\|$ in the right hand side of** (4.22)**.** As we shall see, this is the step where we use Lemmas 4.3 and 4.4.

Fix $\sigma \geq \sigma_0$. By the definition of $\mathcal{M}$, for any $x \in \{\mathcal{M} \geq \sigma/K\}$ we can find a section $S(x, t_x)$ such that

$$⨍_{S(x,t_x)} \|D^2\phi(y)\|\, dy \geq \frac{\sigma}{2K}. \tag{4.23}$$

Consider the family of sections $\{S(x, t_x)\}_{x\in\{\mathcal{M}\geq\sigma/K\}}$ constructed in this way, and extract a subfamily $\{S(x_i, t_{x_i})\}_{i\in I}$ such that

$$\left\{\mathcal{M} \geq \tfrac{\sigma}{K}\right\} \subset \textstyle\bigcup_{i\in I} S(x_i, t_{x_i}) \tag{4.24}$$

and the sections $\{S(x_i, t_{x_i})\}_{i\in I}$ have bounded overlapping, that is,

$$\forall\, z \in \mathbb{T}^n \quad \#\{i \in I \,:\, z \in S(x_i, t_{x_i})\} \leq N \tag{4.25}$$

for some $N \in \mathbb{N}$ universal.[10]

Then, Lemmas 4.3 and 4.4 applied to the sections $S(x_i, t_{x_i})$ yield sets $A(x_i, t_{x_i}) \subset S(x_i, t_{x_i})$ such that

$$\frac{|A(x_i, t_{x_i})|}{|S(x_i, t_{x_i})|} \geq \frac{1}{2}, \qquad \#\{i \in I \,:\, z \in A(x_i, t_{x_i})\} \leq N \quad \forall\, z \in \mathbb{T}^n, \tag{4.26}$$

(the finite overlapping property is an immediate consequence of (4.25)), and

$$\frac{\sigma}{2K} \overset{(4.23)}{\leq} ⨍_{S(x_i,t_{x_i})} \|D^2\phi(y)\|\, dy \leq C_1 \frac{\|T_i\|\|T_i^*\|}{(\det T_i)^{2/n}} \leq \frac{C_1}{c_1}\|D^2\phi(y)\| \qquad \forall\, y \in A(x_i, t_{x_i}) \tag{4.27}$$

[10] It is actually unknown whether, given a family of sections, one can extract a subfamily with finite overlapping. Here we are assuming that this can be done just to make the presentation simpler. However, there are at least to ways to circumvent this issue: either one slightly reduces t_{x_i} by a factor $(1 - \varepsilon)$ with $\varepsilon > 0$ so that the finite overlapping property holds (see [14, Lemma 1] and how this is applied in [19]), or one shrink t_{x_i} by a universal factor $\eta < 1$ and then the sections can be made disjoint (see [22, 26]).

(here T_i denotes the affine map which normalizes $S(x_i, t_{x_i})$). Thanks to these facts we deduce that

$$\begin{aligned}\big|\{\mathcal{M} \geq \tfrac{\sigma}{K}\}\big| &\overset{(4.24)}{\leq} \sum_{i\in I} |S(x_i, t_{x_i})| \overset{(4.25)}{\leq} 2\sum_{i\in I} |A(x_i, t_{x_i})| \\ &\overset{(4.27)}{\leq} 2\sum_{i\in I} \big|A(x_i, t_{x_i}) \cap \{\|D^2\phi\| \geq \tfrac{c_1\sigma}{2KC_1}\}\big| \\ &\overset{(4.26)}{\leq} 2N\,\big|\{\|D^2\phi\| \geq \tfrac{c_1\sigma}{2KC_1}\}\big|.\end{aligned}$$

Hence, if we set $K_1 := \max\{2NK, 2KC_1/c_1\}$, this allows us to replace $\mathcal{M}$ with $\|D^2\phi\|$ in the right hand side of (4.22) and get

$$\int_{\{\|D^2\phi\|\geq\sigma\}} \|D^2\phi(y)\|\,dy \leq K_1\sigma\big|\{\|D^2\phi\| \geq \tfrac{\sigma}{K_1}\}\big| \qquad \forall\,\sigma \geq \sigma_0. \tag{4.28}$$

• **Step 2: a Gehring-type lemma.** Equation (4.28) is a sort of reverse Chebyshev's inequality for $\|D^2\phi\|$. We now show how this allows us to obtain higher integrability of $\|D^2\phi\|$.

Set $g(s) := |\{\|D^2\phi\| \geq s\}|$. By the layer-cake formula we have

$$\int_{\{\|D^2\phi\|\geq\sigma\}} \|D^2\phi(y)\|\,dy = \sigma|\{\|D^2\phi\| \geq \sigma\}| + \int_\sigma^\infty |\{\|D^2\phi\| \geq s\}|\,ds = g(\sigma)\sigma + \int_\sigma^\infty g(s)ds, \tag{4.29}$$

hence (4.28) implies that

$$\int_\sigma^\infty g(s)ds \leq K_1\sigma\, g\big(\tfrac{\sigma}{K_1}\big) \qquad \forall\,\sigma \geq \sigma_0. \tag{4.30}$$

Also, noticing that $g(\sigma) \leq |\mathbb{T}^n| = 1$, again by the layer-cake formula we get

$$\int_{\mathbb{T}^n} \|D^2\phi(y)\|^{1+\gamma}\,dy = (1+\gamma)\int_0^\infty \sigma^\gamma g(\sigma)\,d\sigma \leq \sigma_0^{1+\gamma} + (1+\gamma)\int_{\sigma_0}^\infty \sigma^\gamma g(\sigma)\,d\sigma.$$

Hence, to prove that $\|D^2\phi\| \in L^{1+\gamma}(\mathbb{T}^n)$ we have to show that

$$\int_{\sigma_0}^\infty \sigma^\gamma g(\sigma)\,d\sigma < \infty \tag{4.31}$$

for some $\gamma > 0$.

To this aim, performing an integrations by parts and using that $s \mapsto g(s)$ in non-increasing, we see that

$$
\begin{aligned}
\int_{\sigma_0}^{\infty} \sigma^\gamma g(\sigma)\,d\sigma &= -\int_{\sigma_0}^{\infty} \sigma^\gamma \frac{d}{d\sigma}\biggl(\int_{\sigma}^{\infty} g(s)\,ds\biggr)\,d\sigma \\
&= \sigma_0^\gamma \int_{\sigma_0}^{\infty} g(s)\,ds + \gamma \int_{\sigma_0}^{\infty} \sigma^{\gamma-1}\biggl(\int_{\sigma}^{\infty} g(s)\,ds\biggr)\,d\sigma \\
&\overset{(4.30)}{\le} \sigma_0^\gamma \int_{\sigma_0}^{\infty} g(s)\,ds + K_1\gamma \int_{\sigma_0}^{\infty} \sigma^\gamma g\bigl(\tfrac{\sigma}{K_1}\bigr)\,d\sigma \\
&\le \sigma_0^\gamma \int_{\sigma_0}^{\infty} g(s)\,ds + K_1\gamma\, g\bigl(\tfrac{\sigma_0}{K_1}\bigr) \int_{\sigma_0}^{K_1\sigma_0} \sigma^\gamma\,d\sigma + K_1\gamma \int_{K_1\sigma_0}^{\infty} \sigma^\gamma g\bigl(\tfrac{\sigma}{K_1}\bigr)\,d\sigma \\
&\overset{\tau=\sigma/K_1}{=} \sigma_0^\gamma \int_{\sigma_0}^{\infty} g(s)\,ds + K_1\,\frac{\gamma}{\gamma+1}\,(K_1\sigma_0)^{\gamma+1} g\bigl(\tfrac{\sigma_0}{K_1}\bigr) + K_1^{2+\gamma}\,\gamma \int_{\sigma_0}^{\infty} \tau^\gamma g(\tau)\,d\tau.
\end{aligned}
$$

Hence, recalling that $g \le 1$, we can choose $\gamma > 0$ small enough so that $K_1^{2+\gamma}\gamma \le 1/2$ and notice that

$$
\sigma_0^\gamma \int_{\sigma_0}^{\infty} g(s)\,ds \overset{(4.29)}{\le} \int_{\{\|D^2\phi\|\ge\sigma\}} \|D^2\phi(y)\|\,dy \le \int_{\mathbb{T}^n} \|D^2\phi(y)\|\,dy < \infty
$$

(to get the finiteness of $\|D^2\phi\|_{L^1(\mathbb{T}^n)}$ simply apply (4.8) with t large enough so that $S(x,t) \supset [0,1]^n$) to obtain that

$$
\int_{\sigma_0}^{\infty} \sigma^\gamma g(\sigma)\,d\sigma \le 2\int_{\mathbb{T}^n} \|D^2\phi(y)\|\,dy + 2K_1\,\frac{\gamma}{\gamma+1}\,(K_1\sigma_0)^{\gamma+1} < \infty.
$$

This shows the validity of (4.31) and concludes the proof of the $W^{2,1+\gamma}_{\mathrm{loc}}$ regularity of ϕ.

4.2 Regularity for time-dependent solutions of Monge-Ampère: proof of (3.10)

To deal with the term $\partial_t \nabla P_t^*$, we shall use an idea of Loeper [31, Theorem 5.1] to combine (3.9) and (1.6) and prove the following:

Theorem 4.5. *There exists a universal constant C such that, for almost every $t \ge 0$,*

$$
\int_{\mathbb{T}^2} \rho_t |\partial_t \nabla P_t^*|^{1+\kappa} \le C, \qquad \kappa := \frac{\gamma}{2+\gamma}. \tag{4.32}
$$

Notice that, since $\rho_t \ge \lambda > 0$ (see (3.1)), (4.32) implies immediately (3.10).

Proof. In order to justify the following computations one needs to perform a careful regularization argument. Here we show just the formal computations, referring to [2, Section 3] for more details.

We begin by differentiating in time the relation (3.8) to get

$$
\sum_{i,j=1}^{2} M_{ij}(D^2 P_t^*)\,\partial_t \partial_{ij} P_t^* = \partial_t \rho_t,
$$

where $M_{ij}(A) := \frac{\partial \det(A)}{\partial A_{ij}}$ is the cofactor matrix of A. Taking into account (1.6) and the well-known divergence-free property of the cofactor matrix

$$\sum_{i=1}^{2} \partial_i \big[M_{ij}(D^2 P_t{}^*)\big] = 0, \qquad j = 1, 2,$$

(see for instance [23, Chapter 8.1.4.b] for a proof), we can rewrite the above equation as

$$\sum_{i,j=1}^{2} \partial_i \big(M_{ij}(D^2 P_t^*)\, \partial_t \partial_j P_t^*\big) = -\mathrm{div}(\boldsymbol{U}_t \rho_t).$$

Then, recalling that for invertible matrices the cofactor matrix $M(A)$ is equal to $\det(A)\, A^{-1}$, using again the relation (3.8) we get

$$\mathrm{div}\big(\rho_t (D^2 P_t^*)^{-1} \partial_t \nabla P_t^*\big) = -\mathrm{div}(\rho_t \boldsymbol{U}_t). \tag{4.33}$$

We now multiply (4.33) by $\partial_t P_t^*$ and integrate by parts to obtain[11]

$$\begin{aligned} \int_{\mathbb{T}^2} \rho_t |(D^2 P_t^*)^{-1/2} \partial_t \nabla P_t^*|^2 \, dx &= \int_{\mathbb{T}^2} \rho_t \, \partial_t \nabla P_t^* \cdot (D^2 P_t^*)^{-1} \partial_t \nabla P_t^* \, dx \\ &= -\int_{\mathbb{T}^2} \rho_t \, \partial_t \nabla P_t^* \cdot \boldsymbol{U}_t \, dx. \end{aligned} \tag{4.34}$$

From Cauchy-Schwartz inequality, the right-hand side of (4.34) can be estimated as

$$\begin{aligned} &-\int_{\mathbb{T}^2} \rho_t \, \partial_t \nabla P_t^* \cdot (D^2 P_t^*)^{-1/2} (D^2 P_t^*)^{1/2} \boldsymbol{U}_t \, dx \\ &\qquad \leq \left(\int_{\mathbb{T}^2} \rho_t |(D^2 P_t^*)^{-1/2} \partial_t \nabla P_t^*|^2 \, dx \right)^{1/2} \left(\int_{\mathbb{T}^2} \rho_t |(D^2 P_t^*)^{1/2} \boldsymbol{U}_t|^2 \, dx \right)^{1/2}, \end{aligned} \tag{4.35}$$

hence (4.34) and (4.35) give

$$\int_{\mathbb{T}^2} \rho_t |(D^2 P_t^*)^{-1/2} \partial_t \nabla P_t^*|^2 \, dx \leq \int_{\mathbb{T}^2} \rho_t |(D^2 P_t^*)^{1/2} \boldsymbol{U}_t|^2 \, dx. \tag{4.36}$$

We now observe that

$$\int_{\mathbb{T}^2} \rho_t |(D^2 P_t^*)^{1/2} \boldsymbol{U}_t|^2 \, dx = \int_{\mathbb{T}^2} \rho_t \, \boldsymbol{U}_t \cdot D^2 P_t^* \boldsymbol{U}_t \, dx \leq \sup_{\mathbb{T}^2} \big(\rho_t |\boldsymbol{U}_t|^2\big) \int_{\mathbb{T}^2} \|D^2 P_t^*\| \, dx. \tag{4.37}$$

Hence, recalling that $\boldsymbol{U}_t$ and ρ_t are bounded and noticing that $\int_{\mathbb{T}^2} \|D^2 P_t^*\| \, dx < \infty$,[12] it follows from (4.36) and (4.37) that

$$\int_{\mathbb{T}^2} \rho_t |(D^2 P_t^*)^{-1/2} \partial_t \nabla P_t^*|^2 \, dx \leq C. \tag{4.38}$$

[11] Since the matrix $D^2 P_t{}^*$ is positive definite, both its square root and the square root of its inverse are well-defined.
[12] This obviously follows by (3.9), but a direct proof can be given arguing as for (2.15).

Thus, applying Hölder's inequality and noticing that $\frac{1+\kappa}{1-\kappa} = 1 + \gamma$, we get

$$\begin{aligned}
\int_{\mathbb{T}^2} \rho_t |\partial_t \nabla P_t^*|^{1+\kappa}\, dx &\le \int_{\mathbb{T}^2} \Big(\sqrt{\rho_t} |(D^2 P_t^*)^{-1/2} \partial_t \nabla P_t^*| \Big)^{1+\kappa} \left(\frac{\|(D^2 P_t^*)^{1/2}\|}{\sqrt{\rho_t}} \right)^{1+\kappa} dx \\
&\le \left[\int_{\mathbb{T}^2} \rho_t |(D^2 P_t^*)^{-1/2} \partial_t \nabla P_t^*|^2\, dx \right]^{(1+\kappa)/2} \left[\int_{\mathbb{T}^2} \left(\frac{\|D^2 P_t^*\|}{\rho_t} \right)^{\frac{1+\kappa}{1-\kappa}} dx \right]^{(1-\kappa)/2} \\
&\overset{(4.38)+(3.1)}{\le} \left(\frac{C}{\lambda} \right)^{(1+\kappa)/2} \left[\int_{\mathbb{T}^2} \|D^2 P_t^*\|^{1+\gamma}\, dx \right]^{(1-\kappa)/2} \overset{(3.9)}{\le} \bar{C},
\end{aligned}$$

which proves (4.32). □

5 Short-time existence and uniqueness of smooth solutions for dual SG

In this section we discuss the results of Loeper in [32] concerning the short-time existence and uniqueness of smooth solutions for the dual SG system (1.6). As we have seen in the previous sections there is a strict correspondence between solutions of (1.6) and solutions of the original SG system (1.2), hence these results can be easily read back in the original framework.

We shall prove that if ρ_0 is Hölder continuous then there exists a unique Hölder solution (1.6) on some time interval $[0, T]$, where T depends only on the bounds on ρ_0. Using higher regularity theory for elliptic equations, it is not difficult to check that if ρ_0 is more regular (say, $C^{k,\alpha}$ for some $k \ge 0$ and $\alpha \in (0, 1)$), then the solution that we constructed enjoys the same regularity.

The following result is contained in [32, Theorem 3.3, Corollary 3.4, Theorem 4.1]:

Theorem 5.1. *Assume that*

$$0 < \lambda \le \rho_0 \le \Lambda \quad \textit{and} \quad \rho_0 \in C^{0,\alpha}(\mathbb{T}^2) \tag{5.1}$$

for some $\alpha \in (0, 1)$. Then there exists $T > 0$, depending only on $\lambda, \Lambda, \|\rho_0\|_{C^{0,\alpha}(\mathbb{T}^2)}$, such that (1.6) has a unique solution (ρ_t, P_t^*) on $[0, T]$ satisfying

$$0 < \lambda \le \rho_t \le \Lambda, \qquad \rho_t \in L^\infty\big([0, T], C^{0,\alpha}(\mathbb{T}^2)\big), \quad P_t^* \in L^\infty\big([0, T], C^{2,\alpha}(\mathbb{T}^2)\big). \tag{5.2}$$

We first discuss the existence part and then we deal with uniqueness.

5.1 Short-time existence of smooth solutions

The proof of existence given in [32, Theorem 3.3] is based on a fixed point argument. Here we give a different proof, more in the spirit of the argument used in Section 2.4.

Set $K_0 := 2\|\rho_0\|_{C^{0,\alpha}(\mathbb{T}^2)}$, let $T > 0$ small (to be fixed later), let $j \in \mathbb{N}$, and exactly as in Section 2.4.1 construct a family of approximate solutions by "freezing" the vector fields over time intervals of length $\frac{T}{j}$. More precisely, for $t \in \big[0, \frac{T}{j}\big]$ we define $P_t^{*,j}$ as the unique map whose gradient sends ρ_0 to dx (see Theorem 2.2), and we set

$$\boldsymbol{U}_t^j := \big(\nabla P_t^{*,j}(y) - y\big)^\perp \qquad \forall\, t \in \big[0, \tfrac{T}{j}\big]. \tag{5.3}$$

Notice that, by Caffarelli's regularity theory for the Monge-Ampère equation [11] we have

$$\|D^2 P_t^{*,j}\|_{C^{0,\alpha}(\mathbb{T}^2)} \le K_1 = K_1(K_0,\lambda,\Lambda) \qquad \forall\, t\in[0,T/t],$$

hence

$$\|\nabla \boldsymbol{U}_t^j\|_{L^\infty(\mathbb{T}^2)} \le K_2 := 1+K_1. \tag{5.4}$$

We now consider the flow of $\boldsymbol{U}_t^j$ over the time interval $\left[0,\frac{T}{j}\right]$,

$$\begin{cases} \dot Y^j(t,y) = \boldsymbol{U}_t^j(Y^j(t,y)) & \text{for } t\in\left[0,\frac{T}{j}\right],\\ Y^j(0,y)=y, \end{cases} \tag{5.5}$$

and define

$$\rho_t^j := Y^j(t)_\#\rho_0 \qquad \forall\, t\in\left[0,\tfrac{T}{j}\right].$$

Recall that, since $\boldsymbol{U}_t^j$ is divergence free, ρ^j can also be written as

$$\rho_t^j = \rho_0\circ Y^j(t)^{-1} \tag{5.6}$$

(see (2.5)). Recalling (5.1), this implies in particular that $\lambda\le\rho_t^j\le\Lambda$.

We now differentiate (5.5) with respect to y to get

$$\begin{cases} \frac{d}{dt}\nabla Y^j(t,y) = \Big(\nabla\boldsymbol{U}_t^j(Y^j(t,y))\Big)\nabla Y^j(t,y),\\ \nabla Y^j(0,y)=\mathrm{Id}, \end{cases}$$

so (5.4) yields

$$\begin{cases} \frac{d}{dt}\|\nabla Y^j(t,y)\| \le K_2\|\nabla Y^j(t,y)\|,\\ \|\nabla Y^j(0,y)\|=1, \end{cases}$$

and by Gronwall's Lemma we deduce that

$$e^{-K_2 t}\le\|\nabla Y^j(t,y)\|\le e^{K_2 t},$$

that is $Y^j(t)$ is a bi-Lipschitz homeomorphism with bi-Lipschitz norm controlled by $e^{K_2 t}$. Inserting this information into (5.6) we deduce that, provided T is small enough so that

$$e^{K_2 T}\le 2, \tag{5.7}$$

it holds

$$\|\rho_t^j\|_{C^{0,\alpha}(\mathbb{T}^2)}\le e^{K_2 t}\|\rho_0\|_{C^{0,\alpha}(\mathbb{T}^2)}\le K_0 \qquad \forall\, t\in\left[0,\tfrac{T}{j}\right]. \tag{5.8}$$

(recall that, by definition, $K_0 = 2\|\rho_0\|_{C^{0,\alpha}(\mathbb{T}^2)}$).

We now repeat this procedure over the time interval $t\in\left[\frac{T}{j},2\frac{T}{j}\right]$. More precisely, for $t\in\left[\frac{T}{j},2\frac{T}{j}\right]$ we consider $P_t^{*,j}$ the unique map whose gradient sends $\rho_{T/j}^j$ to dx, we define $\boldsymbol{U}_t^j$ for $t\in\left[\frac{T}{j},2\frac{T}{j}\right]$ as in (5.3), we consider its flow $Y^j(t)$, and we use this flow to let $\rho_{T/j}^j$ evolve over the time interval

$\left[\frac{T}{j}, 2\frac{T}{j}\right]$. Notice that, thanks to (5.8), $\|\rho_t^j\|_{C^{0,\alpha}(\mathbb{T}^2)} \leq K_0$ so we still have $\|\nabla \boldsymbol{U}_t^j\|_{L^\infty(\mathbb{T}^2)} \leq K_2$. Hence, by the same argument as above,

$$\|\rho_t^j\|_{C^{0,\alpha}(\mathbb{T}^2)} \leq \|\rho_{T/j}^j\|_{C^{0,\alpha}(\mathbb{T}^2)} e^{K_2\left(t-\frac{T}{j}\right)} \qquad \forall\, t \in \left[\tfrac{T}{j}, 2\tfrac{T}{j}\right].$$

In particular, combining this bound with (5.8) and (5.7), we get

$$\|\rho_t^j\|_{C^{0,\alpha}(\mathbb{T}^2)} \leq e^{K_2\left(t-\frac{T}{j}\right)} e^{K_2\frac{T}{j}} \|\rho_0\|_{C^{0,\alpha}(\mathbb{T}^2)} = e^{K_2 t}\|\rho_0\|_{C^{0,\alpha}(\mathbb{T}^2)} \leq K_0 \qquad \forall\, t \in [0, 2\tfrac{T}{j}].$$

Iterating this procedure j times we construct a family $(\rho_t^j, P_t^{*,j})$, with

$$\lambda \leq \rho_t^j \leq \Lambda, \quad \|\rho_t^j\|_{C^{0,\alpha}(\mathbb{T}^2)} \leq K_0, \quad \|D^2 P_t^{*,j}\|_{C^{0,\alpha}(\mathbb{T}^2)} \leq K_1 \qquad \forall\, t \in [0,T], \tag{5.9}$$

such that

$$\begin{cases} \partial_t \rho_t^j + \operatorname{div}(\boldsymbol{U}_t^j \rho_t^j) = 0 & \text{in } [0,T]\times\mathbb{R}^2, \\ \boldsymbol{U}_t^j(y) = \left(\nabla P_t^{*,j}(y) - y\right)^{\perp} & \text{in } [0,T]\times\mathbb{R}^2, \\ \rho_{iT/j}^j = (\nabla P_t^j)_\sharp dx & \text{for } t \in \left[i\frac{T}{j}, (i+1)\frac{T}{j}\right), \\ \rho_0^j = \rho_0 & \text{on } \mathbb{R}^2. \end{cases} \tag{5.10}$$

Thanks to the bounds (5.9) it is easy to show that, up to subsequences, $(\rho_t^j, P_t^{*,j})$ converge to a solution of (1.6) that will satisfy (5.2) (compare with Section 2.4.2). This concludes the proof of the existence part.

5.2 Uniqueness of smooth solutions

Let $(\rho_t^1, P_t^{*,1})$ and $(\rho_t^2, P_t^{*,2})$ be two solutions of (1.6) satisfying (5.2). Our goal is to show that they coincide. Because the argument is pretty involved, we shall split it into three steps.

5.2.1 A Gronwall argument

Recalling that ρ_t^i are given by $Y^i(t)_\# \rho_0$ where $Y^i(t)$ is the flow of $\boldsymbol{U}_t^i = (\nabla P_t^{*,i} - y)^\perp$, $i = 1,2$ (see Section 2.1), it is enough to show that $Y^1(t) = Y^2(t)$. So we compute

$$\begin{aligned} \frac{d}{dt}\int_{\mathbb{T}^2} \left|Y^1(t) - Y^2(t)\right|^2 dy &= 2\int_{\mathbb{T}^2} \left(Y^1(t) - Y^2(t)\right)\cdot\left(\dot Y^1(t) - \dot Y^2(t)\right) dy \\ &= 2\int_{\mathbb{T}^2} \left(Y^1(t) - Y^2(t)\right)\cdot\left(\boldsymbol{U}_t^1(Y^1(t)) - \boldsymbol{U}_t^2(Y^2(t))\right) dy \\ &= 2\int_{\mathbb{T}^2} \left(Y^1(t) - Y^2(t)\right)\cdot\left(\boldsymbol{U}_t^1(Y^1(t)) - \boldsymbol{U}_t^1(Y^2(t))\right) dy \\ &+ 2\int_{\mathbb{T}^2} \left(Y^1(t) - Y^2(t)\right)\cdot\left(\boldsymbol{U}_t^1(Y^2(t)) - \boldsymbol{U}_t^2(Y^2(t))\right) dy \\ &\leq 2\|\nabla \boldsymbol{U}_t^1\|_{L^\infty(\mathbb{T}^2)} \int_{\mathbb{T}^2} \left|Y^1(t) - Y^2(t)\right|^2 dy \\ &+ \int_{\mathbb{T}^2} \left|Y^1(t) - Y^2(t)\right|^2 dy + \int_{\mathbb{T}^2} \left|\boldsymbol{U}_t^1(Y^2(t)) - \boldsymbol{U}_t^2(Y^2(t))\right|^2 dy, \end{aligned}$$

where at the last step we used that $2a \cdot b \leq |a|^2 + |b|^2$. Notice that (5.2) implies that $\nabla \boldsymbol{U}_t^1$ is bounded, hence the above estimate gives

$$\frac{d}{dt}\int_{\mathbb{T}^2}\left|Y^1(t)-Y^2(t)\right|^2 dy \leq C\int_{\mathbb{T}^2}\left|Y^1(t)-Y^2(t)\right|^2 dy + \int_{\mathbb{T}^2}\left|\boldsymbol{U}_t^1(Y^2(t))-\boldsymbol{U}_t^2(Y^2(t))\right|^2 dy. \quad (5.11)$$

We now want to bound the last term in the right hand side. For this we first notice that

$$|\boldsymbol{U}_t^1-\boldsymbol{U}_t^2| = |(\nabla P_t^{*,1}-y)^\perp - (\nabla P_t^{*,2}-y)^\perp| = |(\nabla P_t^{*,1}-\nabla P_t^{*,2})^\perp| = |\nabla P_t^{*,1}-\nabla P_t^{*,2}|, \quad (5.12)$$

hence, recalling that $\rho_t^2 = Y^2(t)_\# \rho_0$, we get

$$\begin{aligned}\int_{\mathbb{T}^2}\left|\boldsymbol{U}_t^1(Y^2(t))-\boldsymbol{U}_t^2(Y^2(t))\right|^2 dy &\overset{(5.2)}{\leq} \frac{1}{\lambda}\int_{\mathbb{T}^2}\left|\boldsymbol{U}_t^1(Y^2(t))-\boldsymbol{U}_t^2(Y^2(t))\right|^2 \rho_t^2\, dy \\ &\overset{(1.3)}{=} \frac{1}{\lambda}\int_{\mathbb{T}^2}\left|\boldsymbol{U}_t^1-\boldsymbol{U}_t^2\right|^2 \rho_0\, dy \\ &\overset{(5.2)}{\leq} \frac{\Lambda}{\lambda}\int_{\mathbb{T}^2}\left|\boldsymbol{U}_t^1-\boldsymbol{U}_t^2\right|^2 dy \\ &\overset{(5.12)}{\leq} \frac{\Lambda}{\lambda}\int_{\mathbb{T}^2}\left|\nabla P_t^{*,1}-\nabla P_t^{*,2}\right|^2 dy.\end{aligned} \quad (5.13)$$

Thus we are left with estimating the L^2 norm of $\nabla P_t^{*,1}-\nabla P_t^{*,2}$.

5.2.2 An interpolation argument

To estimate $\|\nabla P_t^{*,1}-\nabla P_t^{*,2}\|_{L^2(\mathbb{T}^2)}$, the idea is to find a curve $[1,2] \ni \theta \mapsto \nabla P_t^{*,\theta}$ which interpolates between these two functions, write

$$\nabla P_t^{*,1}-\nabla P_t^{*,2} = \int_1^2 \partial_\theta \nabla P_t^{*,\theta}\, d\theta$$

so that by Holder's inequality

$$\|\nabla P_t^{*,1}-\nabla P_t^{*,2}\|_{L^2(\mathbb{T}^2)}^2 \leq \left(\int_1^2 \|\partial_\theta \nabla P_t^{*,\theta}\|_{L^2(\mathbb{T}^2)}\, d\theta\right)^2 \leq \int_1^2 \|\partial_\theta \nabla P_t^{*,\theta}\|_{L^2(\mathbb{T}^2)}^2\, d\theta, \quad (5.14)$$

and try to control $\|\partial_\theta \nabla P_t^{*,\theta}\|_{L^2(\mathbb{T}^2)}$ with $\|Y^1(t)-Y^2(t)\|_{L^2(\mathbb{T}^2)}$ in order to close the Gronwall argument in (5.11).

To this aim, we consider a curve of measure $[1,2] \ni \theta \mapsto \rho_t^\theta$ (to be chosen) which interpolates between ρ_t^1 and ρ_t^2 and define $\nabla P_t^{*,\theta}$ as the optimal map sending ρ_t^θ onto dx (see Theorem 2.2). Assume that the measures ρ_t^θ satisfy

$$\frac{1}{K_2} \leq \rho_t^\theta \leq K_2, \quad \|\rho_t^\theta\|_{C^{0,\alpha}(\mathbb{T}^2)} \leq K_2, \quad (5.15)$$

for some universal constant $K_2 > 0$, so that[13]

$$\|D^2 P_t^{*,\theta}\|_{L^\infty(\mathbb{T}^2)} \leq K_3, \qquad \|(D^2 P_t^{*,\theta})^{-1}\|_{L^\infty(\mathbb{T}^2)} \leq K_3. \quad (5.16)$$

[13]The bound on $D^2 P_t^{*,\theta}$ follows by the $C^{2,\alpha}$ regularity for Monge-Ampère [11], while the bound for $(D^2 P_t^{*,\theta})^{-1}$ follows exactly as in the proof of (4.21).

Also, assume that there is a vector field $\boldsymbol{V}_t^\theta$ such that

$$\partial_\theta \rho_t^\theta + \operatorname{div}(\boldsymbol{V}_t^\theta \rho_t^\theta) = 0 \qquad \text{on } [1,2]\times\mathbb{R}^2 \tag{5.17}$$

(Notice that here t is just a fixed parameter, while θ is playing the role of the time variable).

Then, by the very same computations as in the proof of Theorem 4.5 we obtain

$$\int_{\mathbb{T}^2} \rho_t^\theta |(D^2 P_t^{*\theta})^{-1/2} \partial_\theta \nabla P_t^{*,\theta}|^2\,dx \le \int_{\mathbb{T}^2} \rho_t^\theta |(D^2 P_t^{*,\theta})^{1/2} \boldsymbol{V}_t^\theta|^2\,dx$$

(compare with (4.36)), and using (5.15) and (5.16) we deduce that

$$\int_{\mathbb{T}^2} |\partial_\theta \nabla P_t^{*,\theta}|^2\,dx \le K_4 \int_{\mathbb{T}^2} |\boldsymbol{V}_t^\theta|^2 \rho_t^\theta\,dx,$$

that combined with (5.13) and (5.14) gives

$$\int_{\mathbb{T}^2} \big|\boldsymbol{U}_t^1(Y^2(t)) - \boldsymbol{U}_t^2(Y^2(t))\big|^2\,dy \le K_4 \frac{\Lambda}{\lambda} \int_1^2 \left(\int_{\mathbb{T}^2} |\boldsymbol{V}_t^\theta|^2 \rho_t^\theta\,dy \right) d\theta. \tag{5.18}$$

Hence, our goal is to choose $(\rho_t^\theta, \boldsymbol{V}_t^\theta)$ in such a way that (5.15)-(5.17) hold, and the right hand side above is controlled by $\|Y^1(t) - Y^2(t)\|_{L^2(\mathbb{T}^2)}$.[14]

5.2.3 Construction of the interpolating curve

The key observation is that, since $Y^1(t)_\# \rho_0 = \rho_t^1$ and $Y^2(t)_\# \rho_0 = \rho_t^2$, the map $S_t := Y^2(t) \circ [Y^1(t)]^{-1}$ satisfies

$$(S_t)_\# \rho_1^t = \rho_2^t.$$

Hence, if $T_t = \nabla \Phi_t : \mathbb{T}^2 \to \mathbb{T}^2$ denotes the optimal transport map from ρ_t^1 to ρ_t^2, by the definition of optimal transport (see Section 2.2) we have

$$\int_{\mathbb{T}^2} |S_t - y|^2\,\rho_t^1(y)\,dy \ge \int_{\mathbb{T}^2} |T_t - y|^2\,\rho_t^1(y)\,dy.$$

Also, since $[Y^1(t)^{-1}]_\# \rho_t^1 = \rho_0$,

$$\begin{aligned}\int_{\mathbb{T}^2} |S_t - y|^2\,\rho_t^1(y)\,dy &= \int_{\mathbb{T}^2} \big|Y^2(t)\circ[Y^1(t)]^{-1} - y\big|^2\,\rho_t^1(y)\,dy \\ &\overset{(1.3)}{=} \int_{\mathbb{T}^2} |Y^2(t) - Y^1(t)|^2\,\rho_0(y)\,dy \overset{(5.1)}{\le} \Lambda \int_{\mathbb{T}^2} |Y^2(t) - Y^1(t)|^2\,dy,\end{aligned}$$

[14]The reader familiar with optimal transport theory may recognize in (5.18) the dynamic formulation of optimal transportation discovered by Benamou and Brenier [6]:

$$\min\left\{ \int_1^2 \left(\int_{\mathbb{T}^2} |\boldsymbol{V}_t^\theta|^2 \rho_t^\theta\,dx \right) d\theta \,:\, (\rho_t^\theta, \boldsymbol{V}_t^\theta) \text{ satisfy (5.17)} \right\} = \min\left\{ \int_{\mathbb{R}^n} |S(x) - x|^2\,d\mu(x) \,:\, S_\# \rho_t^1 = \rho_t^2 \right\}.$$

Although we shall not use this fact, the argument in Section 5.2.3 is strongly inspired by it.

therefore

$$\Lambda \int_{\mathbb{T}^2} |Y^2(t) - Y^1(t)|^2\, dy \geq \int_{\mathbb{T}^2} |T_t - y|^2\, \rho_t^1(y)\, dy. \tag{5.19}$$

Also, since both ρ_t^1 and ρ_t^2 satisfy (5.2), the bounds

$$\|D^2\Phi_t\|_{C^{0,\alpha}(\mathbb{T}^2)} \leq \hat{K}, \qquad \|(D^2\Phi_t)^{-1}\|_{L^\infty(\mathbb{T}^2)} \leq \hat{K} \tag{5.20}$$

hold (compare with (5.16), see also Footnote 13).

We now would like to relate $\boldsymbol{V}_t^\theta$ to $T_t(y) - y$, and this suggests the following definition of ρ_t^θ (as already mentioned in Footnote 14, this is strongly inspired by [6]):

$$\rho_t^\theta := \big[y + (\theta - 1)(T_t(y) - y)\big]_{\#}\rho_1^t \qquad \forall\, \theta \in [1,2],$$

or equivalently, since since $T_t = \nabla\Phi_t$,

$$\rho_t^\theta = [\nabla\Phi_t^\theta]_{\#}\rho_1^t, \qquad \Phi_t^\theta := (2-\theta)\frac{|y|^2}{2} + (\theta - 1)\Phi_t.$$

Let

$$\Phi_t^{\theta,*}(y) := \sup_{x\in\mathbb{R}^2}\{x \cdot y - \Phi_t^\theta(x)\}.$$

Recalling that $\nabla\Phi_t^{\theta,*} = (\nabla\Phi_t^\theta)^{-1}$ (see (1.5)), one can check that with these definitions the following properties hold:[15]

$$\begin{cases} \text{(A)} \quad \text{(5.17) holds with } \boldsymbol{V}_t^\theta := (T_t - y) \circ \nabla\Phi_t^{\theta,*}, \\ \text{(B)} \quad \int_{\mathbb{T}^2} |\boldsymbol{V}_t^\theta|^2 \rho_t^\theta\, dy = \int_{\mathbb{T}^2} |T_t - y|^2\, \rho_t^1\, dy \qquad \forall\, \theta \in [1,2], \\ \text{(C)} \quad \det(D^2\Phi_t^\theta) = \frac{\rho_t^1}{\rho_t^\theta \circ \nabla\Phi_t^\theta} \qquad \forall\, \theta \in [1,2]. \end{cases} \tag{5.21}$$

5.2.4 Bounds on the interpolating curve: proof of (5.15)

We now prove that the measures ρ_t^θ satisfy all properties in (5.15).

First of all we notice that, thanks to (5.20),

$$\frac{1}{\hat{K}}\, \mathrm{Id} \leq D^2\Phi_t \leq \hat{K}\, \mathrm{Id},$$

therefore, since $D^2\Phi_t^\theta = (2-\theta)\mathrm{Id} + (\theta - 1)D^2\Phi_t$, it follows immediately that

$$\frac{1}{\hat{K}}\, \mathrm{Id} \leq D^2\Phi_t^\theta \leq \hat{K}\, \mathrm{Id} \qquad \forall\, \theta \in [1,2]. \tag{5.22}$$

In particular $\det(D^2\Phi_t^\theta) \in \left[\frac{1}{\hat{K}^n}, \hat{K}^n\right]$, that combined with (5.21)-(C) and the fact that ρ_t^1 satisfies (5.2) gives

$$\rho_t^\theta = \frac{\rho_t^1}{\det(D^2\Phi_t^\theta)} \circ \nabla\Phi_t^{\theta,*} \in \left[\frac{\lambda}{\hat{K}^n}, \frac{\hat{K}^n}{\lambda}\right]. \tag{5.23}$$

[15] Property (A) follows by a direct computation very similar to what we already did in Section 2.1 to show that $Y(t)_{\#}\bar{\sigma}$ solves (2.1). Property (B) is a direct consequence of (1.3) and the fact that $[(\nabla\Phi_t^\theta)^{-1}]_{\#}\rho_t^\theta = \rho_1^t$, while (C) follows by (2.7). We leave the details to the interested reader.

Also, the Hölder continuity of $D^2\Phi_t$ (see (5.20)) implies that $D^2\Phi_t^\theta \in C^{0,\alpha}$, from which it follows that

$$\left\|\det(D^2\Phi_t^\theta)\right\|_{C^{0,\alpha}(\mathbb{T}^2)} \leq \hat{K}_0,$$

so by (5.2) and the fact that $\det(D^2\Phi_t^\theta) \geq 1/\hat{K}^n$ we get

$$\left\|\frac{\rho_t^1}{\det(D^2\Phi_t^\theta)}\right\|_{C^{0,\alpha}(\mathbb{T}^2)} \leq \hat{K}_1.$$

Finally, it suffices to observe that $\|D^2\Phi_t^{\theta,*}\| \leq \hat{K}$ (this simply follows from (5.22) and (3.5)) to deduce that $\nabla\Phi_t^{\theta,*}$ is uniformly Lipschitz, thus

$$\left\|\frac{\rho_t^1}{\det(D^2\Phi_t^\theta)} \circ \nabla\Phi_t^{\theta,*}\right\|_{C^{0,\alpha}(\mathbb{T}^2)} \leq \hat{K}_2.$$

Recalling (5.23), this concludes the proof of (5.15).

5.2.5 Conclusion

The fact that the measures ρ_t^θ satisfy the properties in (5.15) allows us to justify all the previous computations. In particular, thanks to (5.19), (5.18), and (5.21)-(B), we get

$$\int_{\mathbb{T}^2}\left|\boldsymbol{U}_t^1(Y^2(t)) - \boldsymbol{U}_t^2(Y^2(t))\right|^2 dy \leq K_4\,\frac{\Lambda^2}{\lambda}\int_{\mathbb{T}^2}|Y^2(t) - Y^1(t)|^2\,dy.$$

Inserting this bound into (5.11), we finally obtain

$$\frac{d}{dt}\int_{\mathbb{T}^2}\left|Y^1(t) - Y^2(t)\right|^2 dy \leq \bar{C}\int_{\mathbb{T}^2}|Y^2(t) - Y^1(t)|^2\,dy,$$

so by Gronwall's inequality

$$\int_{\mathbb{T}^2}\left|Y^1(t) - Y^2(t)\right|^2 dy \leq e^{\bar{C}\,t}\int_{\mathbb{T}^2}\left|Y^1(0) - Y^2(0)\right|^2 dy = 0,$$

as desired.

6 Open problems

In this last section we state some open problems related to the Monge-Ampère and semigeostrophic equations.

1. Our global existence result for weak solutions of SG was based on regularity results for Monge-Ampère that are valid in every dimension. However, the regularity theory for Monge-Ampère provides stronger results in 2-D. For instance, Alexandrov showed in [4] (see also [27, Theorem 2.1]) that a convex function $\phi : \mathbb{R}^2 \to \mathbb{R}$ is continuous differentiable provided the

upper bound $\det(D^2\phi) \leq \Lambda$ holds (this result is false when $n \geq 3$, see [37]). Hence, in relation to the theorem proved in Section 4.1, a natural question becomes the following:

Is it possible to prove $W^{2,1}_{\mathrm{loc}}$ regularity of ϕ in the 2-D case assuming only an upper bound on $\det(D^2\phi)$?

Apart from its own interest, such a result could help in extending Theorem 3.1 outside of the periodic setting.

2. As shown in Section 5, the existence of smooth solutions for the dual SG system is known only for short time. However, for the 2-D incompressible Euler equations it is well-known that smooth solutions exist globally in time (see for instance [8, Corollary 3.3]). By the analogy between the dual SG system and the Euler equations (see Section 2.3) one may hope to say that global smooth solutions exist also for the dual SG system, at least for initial data which are sufficiently close to 1 in some strong norm. Whether this fact holds true is an interesting open problem.

3. As proved in [3], the results described here can be extended to the case when the domain is the whole $\mathbb{R}^2$,[16] provided the initial datum $\rho_0 = (\nabla P_0)_\sharp dx\lfloor_\Omega$ is strictly positive on the whole space. It would nice to remove this assumption in order to deal with the case when ρ_0 is compactly supported (which is the most interesting case from a physical point of view). However, the nontrivial evolution of the support of the solution ρ_t does not permit to apply the regularity regularity results from [19, 22, 33], so completely new ideas need to be introduced in order to prove existence of distributional solutions to (1.1) in this case. As already mentioned above, solving Problem 1 could be extremely helpful in this direction.

References

[1] L. Ambrosio: *Transport equation and Cauchy problem for BV vector fields.* Invent. Math., **158** (2004), 227–260.

[2] L. Ambrosio, M. Colombo, G. De Philippis, A. Figalli: *Existence of Eulerian solutions to the semigeostrophic equations in physical space: the 2-dimensional periodic case*, Comm. Partial Differential Equations, **37** (2012), no. 12, 2209–2227.

[3] L. Ambrosio, M. Colombo, G. De Philippis, A. Figalli: *A global existence result for the semi-geostrophic equations in three dimensional convex domains*, Discrete Contin. Dyn. Syst., **34** (2014), no. 4, 1251–1268.

[4] A. D. Alexandrov: *Smoothness of the convex surface of bounded Gaussian curvature.* C. R. (Doklady) Acad. Sci. URSS (N. S.), **36** (1942). 195–199.

[5] J.-D. Benamou, Y. Brenier: *Weak existence for the semigeostrophic equation formulated as a coupled Monge-Ampère/transport problem.* SIAM J. Appl. Math., **58** (1998), 1450–1461.

[16] To be precise, the results in [3] are three dimensional, but they also hold in 2-D.

[6] J.-D. Benamou, Y. Brenier: *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem.* Numer. Math., **84** (2000), no. 3, 375–393.

[7] J. Bergh, J. Löfström: *Interpolation spaces. An introduction.* Grundlehren der Mathematischen Wissenschaften, No. 223. Springer-Verlag, Berlin-New York, 1976. x+207 pp.

[8] A. L. Bertozzi, A. J. Majda: *Vorticity and incompressible flow.* Vol. 27 of Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge, 2002.

[9] Y. Brenier: *Polar factorization and monotone rearrangement of vector-valued functions.* Comm. Pure Appl. Math., **44** (1991), no. 4, 375–417.

[10] L. Caffarelli: *A localization property of viscosity solutions to the Monge-Ampère equation and their strict convexity.* Ann. of Math. (2), **131** (1990), no. 1, 129–134.

[11] L. Caffarelli: *Interior $W^{2,p}$ estimates for solutions of the Monge-Ampère equation.* Ann. of Math. (2), **131** (1990), no. 1, 135–150.

[12] L. Caffarelli: *Some regularity properties of solutions to Monge-Ampère equations.* Comm. Pure Appl. Math., **44** (1991), 965–969.

[13] L. Caffarelli: *A note on the degeneracy of convex solutions to Monge-Ampère equation.* Comm. Partial Differential Equations, **18** (1993), no. 7-8, 1213–1217.

[14] L. Caffarelli, C. Gutierrez: *Real analysis related to the Monge-Ampère equation.* Trans. Amer. Math. Soc., **348** (1996), no. 3, 1075–1092.

[15] D. Cordero Erausquin: *Sur le transport de mesures périodiques.* C. R. Acad. Sci. Paris Sér. I Math., **329** (1999), 199–202.

[16] M. Cullen: *A mathematical theory of large-scale atmosphere/ocean flow.* Imperial College Press (2006).

[17] M. Cullen, M. Feldman: *Lagrangian solutions of semigeostrophic equations in physical space.* SIAM J. Math. Anal., **37** (2006), 1371–1395.

[18] S. Daneri, A. Figalli: *Variational models for the incompressible Euler equations.* HCDTE Lecture Notes. Part II. Nonlinear HYperboliC PDEs, Dispersive and Transport Equations. AIMS Book Series, Applied Mathematics, to appear.

[19] G. De Philippis, A. Figalli: *$W^{2,1}$ regularity for solutions of the Monge-Ampère equation*, Invent. Math., **192** (2013), no. 1, 55–69.

[20] G. De Philippis, A. Figalli: *Second order stability for the Monge-Ampère equation and strong Sobolev convergence of optimal transport maps.* Anal. PDE, **6** (2013), no. 4, 993–1000.

[21] G. De Philippis, A. Figalli: *The Monge-Ampère equation and its link to optimal transportation.* Bull. Amer. Math. Soc. (N.S.), **51** (2014), no. 4, 527–580.

[22] G. De Philippis, A. Figalli, O. Savin: *A note on interior $W^{2,1+\epsilon}$ estimates for the Monge-Ampère equation*, Math. Ann., **357** (2013), no. 1, 11–22.

[23] L. C. Evans: *Partial differential equations. Second edition.* Graduate Studies in Mathematics, 19. American Mathematical Society, Providence, RI, 2010. xxii+749 pp.

[24] M. Feldman, A. Tudorascu,: *On Lagrangian solutions for the semi-geostrophic system with singular initial data.* SIAM J. Math. Anal., **45** (2013), no. 3, 1616–1640.

[25] M. Feldman, A. Tudorascu,: *On the Semi-Geostrophic system in physical space with general initial data.* Preprint, 2014.

[26] A. Figalli: *Sobolev regularity for the Monge-Ampère equation, with application to the semi-geostrophic equations.* Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI), **411** (2013), Teoriya Predstavlenii, Dinamicheskie Sistemy, Kombinatornye Metody. XXII, 103–118, 242; translation in J. Math. Sci. (N. Y.), **196** (2014), no. 2, 175–183

[27] A. Figalli, G. Loeper: *C^1 regularity of solutions of the Monge-Ampère equation for optimal transport in dimension two.* Calc. Var. Partial Differential Equations, **35** (2009), no. 4, 537–550.

[28] C. Gutierrez: *The Monge-Ampére equation.* Progress in Nonlinear Differential Equations and their Applications, 44. Birkhäuser Boston, Inc., Boston, MA, 2001.

[29] C. Gutierrez, Q. Huang: *Geometric properties of the sections of solutions to the Monge-Ampère equation.* Trans. Amer. Math. Soc., **352** (2000), no. 9, 4381–4396.

[30] F. John: *Extremum problems with inequalities as subsidiary conditions.* In Studies and Essays Presented to R. Courant on his 60th Birthday, January 8, 1948, pages 187–204. Interscience, New York, 1948.

[31] G. Loeper: *On the regularity of the polar factorization for time dependent maps.* Calc. Var. Partial Differential Equations, **22** (2005), 343–374.

[32] G. Loeper: *A fully non-linear version of the incompressible Euler equations: The semi-geostrophic system.* SIAM J. Math. Anal., **38** (2006), 795–823.

[33] T. Schmidt: *$W^{2,1+\epsilon}$ estimates for the Monge-Ampère equation*, preprint (2012).

[34] J. Simon: *Compact sets in the space $L^p(0,T;B)$.* Ann. Mat. Pura Appl. (4), **146** (1987), 65–96.

[35] E. Stein: *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals.* With the assistance of Timothy S. Murphy. Princeton Mathematical Series, 43. Monographs in Harmonic Analysis, III. Princeton University Press, Princeton, NJ, 1993.

[36] C. Villani: *Optimal Transport. Old and new.* Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 338. Springer-Verlag, Berlin, 2009.

[37] X.-J. Wang: *Some counterexamples to the regularity of Monge-Ampère equations*, Proc. Amer. Math. Soc., **123** (1995), no. 3, 841–845.