UNIVERSITY OF IOANNINA
DEPARTMENT OF MATHEMATICS

Vasileios Kalivopoulos

The semigeostrophic equation in twodimensional periodic space and its relation to the Euler equation

MASTER THESIS

Ioannina, 2023

The present dissertation thesis was carried out under the postgraduate program of the Department of Mathematics of the University of Ioannina in order to obtain the master degree.

Ioannis Giannoulis Associate Professor (Supervisor), Department of Mathematics, University of Ioannina

Stamatakis Marios Assistant Professor, Department of Mathematics, University of Ioannina

Saroglou Christos Associate Professor, Department of Mathematics, University of Ioannina

Statutory Declaration
I lawfully declare here with statutorily that the present dissertation thesis was carried out under the international ethical and academical rules and under the protection of intellectual property. According to these rules, I avoided plagiarism of any kind and I made reference to any source which I used in this thesis.

Vasileios Kalivopoulos

Περίληψη

Το εισαγωγικό πρώτο κεφάλαιο παρουσιάζει τα συστήματα της ημιγεωστροφικής και της δυϊκής ημιγεωστροφικής εξίσωσης. Γίνεται η εξαγωγή των δύο συστημάτων και αναφέρεται η προϋπόθεση της κυρτότητας των λύσεων στον χώρο για τη μετάβαση από την πρώτη στη δεύτερη. Στην παρούσα εργασία η δυϊκή ημιγεωστροφική εξίσωση αποτελεί το βασικό αντικείμενο ενασχόλησης.

Για να εξαγάγουμε τη δυϊκή ημιγεωστροφική εξίσωση από αυτήν σε φυσικές συντεταγμένες θα κινηθούμε μέσω ενός μέτρου εικόνα και θα καταλήξουμε σε μία εξίσωση συνέχειας για μέτρα με πυκνότητα. Ακόμη, θα αποδείξουμε ότι το σύστημα παραμένει ασυμπίεστο, δηλαδή ότι η δυϊκή ταχύτητα είναι πεδίο μηδενικής απόκλισης.

Εν συνεχεία θα ορίσουμε τι είναι μία ασθενής λύση για τη δυϊκή ημιγεωστροφική εξίσωση. Στην πορεία αναζήτησης της σχέσης που πρέπει να ικανοποιεί μία λύση της δυϊκής ημιγεωστροφικής εξίσωσης θα ορίσουμε τη λύση της συνήθους διαφορικής εξίσωσης για την αντίστοιχη ροή, η οποία θα παίξει σημαντικό ρόλο. Μάλιστα, θα δείξουμε και ενδιαφέρουσες ιδιότητες της λύσης αυτής. Έπειτα, θα προβούμε στην επίλυση του προβλήματος ύπαρξης λύσεων. Ξεκινάμε με την εύρεση ασθενών λύσεων ολικά στον χρόνο. Για να το πετύχουμε αυτό θα κατασκευάσουμε μια οικογένεια προσεγγιστικών λύσεων και θα βρούμε μια συγκλίνουσα υπακολουθία, το όριο της οποίας θα είναι η ζητούμενη λύση.

Όσον αφορά ισχυρότερες λύσεις θα δείξουμε ότι μπορούμε να έχουμε λείες λύσεις (όχι με την κλασική έννοια, αλλά με την ασθενή έννοια στον χώρο και στον χρόνο, όπου τώρα σε κάθε χρονική στιγμή η λύση είναι λεία στον χώρο) τοπικά όμως στον χρόνο. Η συγκεκριμένη λύση, η ύπαρξη της οποίας προκύπτει μέσω των ίδιων βημάτων και επιχειρημάτων όπως προηγουμένως, αποδεικνύεται ότι είναι μοναδική. Αυτό επιτυγχάνεται δείχνοντας ότι για δύο λύσεις, οι αντίστοιχες λύσεις (ροές) της προαναφερθείσας συνήθους διαφορικής εξίσωσης είναι ίσες μέσω της χρήσης ενός επιχειρήματος Gronwall και με την βοήθεια καμπυλών παρεμβολής (interpolating curves).

Στο τέλος της διατριβής παρατίθεται ένα παράρτημα όπου έχουν καταγραφεί όσο το δυνατόν περισσότερες μαθηματικές έννοιες και προτάσεις, οι οποίες χρησιμοποιήθηκαν στην παρούσα διατριβή.

Abstract

We begin with the introduction of the equations that we are going to study. We start by mentioning the Semi-Geostrophic equation (which we abbreviate as SG) in physical variables, for which we explain thoroughly the notations we are going to use throughout the thesis. After that, we make a formal derivation of the aforementioned SG system and we insert the convexity-in-space requirement for their solutions.

Then, we move on to deriving the dual SG system, which will be the main object of study in this thesis. The reason one moves past the SG system is that, at a first glance at least, it provides no evolution equation for the velocity. In order to obtain the dual SG equations, we first try to understand the continuity equation for a measure with density. Lastly, we show that the dual velocity (velocity of the dual SG system) is divergence free as well.

In the second chapter we formulate the equation of a weak solution to the dual SG system, taking the Lagrangian point of view (for the coordinates describing the image of the physical flow). We then proceed to solve the dual SG system, in the weak sense (sometimes referred to as distributional) we have just discussed. We show that we can have global in time weak solutions, but we do not show any uniqueness result. To obtain these solutions we construct a family of approximate ones and we prove that their limit leads to a solution for the dual SG system. We do so with subsequences, which do not yield uniqueness, unless they are shown to yield the same limit. The approximate solutions are obtained by solving the measure continuity equation we obtained, with the help of ordinary differential equations. We also show some interesting properties while studying the existence of weak solutions to the dual SG equation.

In the next chapter we prove the existence and uniqueness of a smooth solution, though this time our solution is only local in time. We follow the same steps as in the proof of existence of weak solutions. We build an approximate sequence and then we take its limit. Moving on, this time, we can prove uniqueness. We show that if two solutions exist, then they coincide. We reduce the question of the existence of a unique solution to the uniqueness of the respective flow, that is, the solution of the aforementioned ODE. To achieve our goal we implement a Gronwall type argument and an interpolation argument.

In the final chapter, we try to relate the dual SG system (rewritten as a coupled system of a continuity equation and a Monge-Ampère equation) to the 2d incompressible Euler in vorticity-stream formulation. Before we work on this, we briefly present some facts about the Euler and the Navier-Stokes equations. At last, we show that local smooth solutions of the dual SG system converge, under some norm, to the 2d incompressible Euler equation in vorticity-stream formulation.

Finally, this thesis contains an appendix, where there was made an effort to gather together mathematical notions and results used in this thesis.

Contents

CHAPTER 1

The semigeostrophic equations

The semigeostrophic (hereafter SG) equations are used in meteorology to describe atmospheric flows in large scale. The SG equations can be derived (with Boussinesq and hydrostatic approximations, under a strong Coriolis force) from those of the 3d incompressible Euler system.

To make our first glance simpler we will present the 2-dimensional periodic SG system.

These equations can be found in [23] [20] [29] [7] [16] [14]

1.1 The SG system in physical variables

The 2d periodic SG system is:

$$\begin{cases} \partial_t \nabla p_t + \langle u_t, \nabla \rangle \nabla p_t + \nabla^{\perp} p_t + u_t = \vec{0} & (x,t) \in \mathbb{R}^2 \times [0,+\infty) \\ \mathrm{div} u_t = 0 & (x,t) \in \mathbb{R}^2 \times [0,+\infty) \\ p_0 = \bar{p} & x := (x_1, x_2) \in \mathbb{R}^2 \end{cases} \tag{1.1.1}$$

where we omit the spatial variable (argument) x and we use the subscript t to denote the time variable.

Remark.
From now on, when we write zero 0 with no subscripts or superscripts, we will mean the corresponding zero of the space we work on.

1.1.1 Explaining the notation

Having an insight on the SG system (1.1.1) , it consists of the time dependent functions $u_t : \mathbb{R}^2 \to \mathbb{R}^2$ and $p_t : \mathbb{R}^2 \to \mathbb{R}$ denoting the velocity and pressure respectively.

We choose to notate the time dependence by writing the subscript t. So, we identify a function $f(x,t)$ as $f_t(x)$. Sometimes it is useful to identify the function $f(x,t)$ as $f_x(t)$ (e.g. when differentiating with respect to time).

A convention

We view vectors either as rows or as columns.

With this convention in mind we use the following notations:

$$\forall t \geq 0 \text{ and } \forall x = (x_1, x_2) \in \mathbb{R}^2$$

The velocity vector field $u_t : \mathbb{R}^2 \to \mathbb{R}^2$

$$u(x,t) = u_t(x) = u_t(x_1,x_2) := \big(u_t^1(x_1,x_2), u_t^2(x_1,x_2)\big) = \big(u_t^1(x), u_t^2(x)\big)$$

Remark.
Early on we "quietly" utilize the convention of considering $\mathbb{R}^n$ as the vector space containing the row vectors or the column vectors depending on the usefulness regarding the presentation and correctness in the mathematical context. If we wanted to be consistent with the definition of vector-valued functions, then we should have written u_t as a column vector. But column vectors are rather lenghty and for this case it does not affect us to view u_t as a row vector.

The pressure function $p_t : \mathbb{R}^2 \to \mathbb{R}$

$$p(x,t) = p_t(x) = p_t(x_1,x_2)$$

The use and no use of subscripts

We "split" the derivatives depending on time t and space x as well. We continue to use the subscript t to refer to everything about time. We avoid the use of any special symbol to denote the differentiation with respect to the space variables, instad we only abbreviate when possible (partial derivatives).

Thus we have the following:

The time derivative:

$$\frac{\partial}{\partial t} = \partial_t$$

which is a (one out of three) partial derivative for our time-depending (space-depending as well) functions.

The abbreviated spatial partial derivatives ∂_i

$$\frac{\partial}{\partial x_1} = \partial_{x_1} = \partial_1$$

$$\frac{\partial}{\partial x_2} = \partial_{x_2} = \partial_2$$

which denote the differentiation with respect to the corresponding fisrt and second spatial variables x_1 and x_2

The differential operator gradient ∇ , which equals the first derivative D when the function is differentiable, but can be defined even if the function in discuss is assumed to only be partially differentiable

$$\nabla = (\partial_1, \partial_2) = D \text{ when the function is differentiable}$$

and it is used to denote the differentiation with respect to the space variables, notated with the symbol "nabla".

Remark.
recall that the terms gradient and derivative (since they do not have a subscript) refer to the differentiation with repsect to the space variable.

Thus the term gradient of pressure reads:

$$\nabla p_t = (\partial_1, \partial_2)(p_t) = (\partial_1 p_t, \partial_2 p_t)$$

Remark.
Notice that since the pressure $p_t : \mathbb{R}^2 \to \mathbb{R}$ is a real-valued function, its gradient $\nabla p_t : \mathbb{R}^2 \to \mathbb{R}^{1\times 2}$ is a row vector by definition (we do not have to "change" our view of $\mathbb{R}^2$ to view it as such). We do identify it as a column vector on $\mathbb{R}^{2\times 1}$, when we want to differentiate (since it is a vector-valued function).

We also implement the term perpendicular gradient $\nabla^\perp$ denoting the clockwise (mathematically negative direction) "rotation" of the "vector" ∇ by $\pi/2$

$$\nabla^\perp = (\partial_2, -\partial_1)$$

Thus

$$\nabla^\perp p_t = (\partial_2, -\partial_1)(p_t) = (\partial_2 p_t, -\partial_1 p_t)$$

We move on to the time derivative of pressure's space derivative, that is $\partial_t \nabla p_t$. Here and every time we differentiate we must be careful with the dimensions. We view $\nabla p_t : \mathbb{R}^2 \to \mathbb{R}^2$ with respect to its time variable i.e. as $\nabla p_t = \nabla p_x : \mathbb{R} \to \mathbb{R}^2$. Thus its time derivative $\partial_t \nabla p_t = \partial_t \nabla p_x : \mathbb{R} \to \mathbb{R}^{2\times 1}$ is a column vector, which (like the velocity u_t) we view as a row vector.

So,

$$\partial_t \nabla p_t = (\partial_t \partial_1 p_t, \partial_t \partial_2 p_t)$$

Remark.
This needs to be done in order to avoid the use of the traspose matrix, but still be right in terms of mathematical correctness otherwise we wouldn't be able to sum the vector-valued functions in the first equation of the SG system (1.1.1).

And now we proceed to the last term (also a differential operator) for the fisrt equation

$$\langle u_t, \nabla \rangle = \sum_{i=1}^{2} u_t^i \partial_i = u_t^1 \partial_1 + u_t^2 \partial_2$$

Hence

$$\begin{aligned}
\langle u_t, \nabla \rangle \nabla p_t &= \sum_{i=1}^{2} u_t^i \partial_i \nabla p_t \\
&= \sum_{i=1}^{2} u_t^i \partial_i (\partial_1 p_t, \partial_2 p_t) \\
&= \sum_{i=1}^{2} (u_t^i \partial_i \partial_1 p_t, u_t^i \partial_i \partial_2 p_t) \\
&= \left(\sum_{i=1}^{2} u_t^i \partial_i \partial_1 p_t, \sum_{i=1}^{2} u_t^i \partial_i \partial_2 p_t \right)
\end{aligned}$$

And finally, the last-last term, the divergence differential operator:

$$\text{div} u_t = \langle \nabla, u_t \rangle = \big\langle (\partial_1, \partial_2), (u_t^1, u_t^2) \big\rangle = \partial_1 u_t^1 + \partial_2 u_t^2 = \sum_{i=1}^{2} \partial_i u_t^i$$

The initial value data $\bar{p}$ is a time independent function from $\mathbb{R}^2$ to $\mathbb{R}$

The SG system in component form

Combining all the above we can rewrite the SG system in its component form. To do that, we firstly substitute each of the previous into the first equation of the SG system ((1.1.1)). So,

$$\bar{0} = \partial_t \nabla p_t + \langle u_t, \nabla \rangle \nabla p_t + \nabla^{\perp} p_t + u_t =$$

$$= (\partial_t \partial_1 p_t, \partial_t \partial_2 p_t) + \left(\sum_{i=1}^{2} u_t^i \partial_i \partial_1 p_t, \sum_{i=1}^{2} u_t^i \partial_i \partial_2 p_t \right) + (\partial_2 p_t, -\partial_1 p_t) + (u_t^1, u_t^2)$$

Thus, we obtain the SG system in component form:

$$\begin{cases} \partial_t \partial_1 p_t + \sum_{i=1}^{2} u_t^i \partial_i \partial_1 p_t + \partial_2 p_t + u_t^1 = 0 & x \in \mathbb{R}^2 \ \ t \geq 0 \\ \partial_t \partial_2 p_t + \sum_{i=1}^{2} u_t^i \partial_i \partial_2 p_t - \partial_1 p_t + u_t^2 = 0 & x \in \mathbb{R}^2 \ \ t \geq 0 \\ \partial_1 u_t^1 + \partial_2 u_t^2 = 0 & x \in \mathbb{R}^2 \ \ t \geq 0 \\ p_0 = \bar{p} & x \in \mathbb{R}^2 \end{cases} \tag{1.1.2}$$

1.1.2 Derivation of the SG equation

This part can be found on Cullen's book [16].

We will derive the SG equations from the 2d incompressible Euler equations with Boussinesq and hydrostatic approximations under a constant Coriolis force F_C.

The 2d hydrostatic incompressible Boussinesq Euler equations under a constant Coriolis force $\vec{F_C} = (F_C, F_C)$ read:

$$\begin{cases} D_t u_t + \nabla p_t = F_C u_t^{\perp} \\ \text{div} u_t = 0 \end{cases}$$

Remark.
In reality, there is one more equation "$D_t \theta = 0$" mentioned by Cullen, but since we will not make use of it, we omit it. θ denotes the temperature of the fluid/flow.

When it comes to the study of a flow in atmosphere (at a large scale), we consider that the velocity comes from the geostrophic and ageostrophic wind.

Thus, we have:

$$u_t = u_t^g + u_t^{ag}$$

where the ageostrophic wind is the difference between the actual wind and the geostrophic wind, a result of the (geostrophic) balance between the horizontal pressure and the Coriolis force. In nature, due to friction, the geostrophic wind does not equal the total wind. But we consider this disturbance to be small i.e. $D_t u_t^{ag} = 0$

The SG approximation to the above equations are the following:

$$\begin{cases} D_t u_t^g + \nabla p_t = F_C u_t^\perp \\ \nabla^\perp p_t = -F_C u_t^g \\ \text{div} u_t = 0 \end{cases} \tag{1.1.3}$$

Remark.
The second equation is exactly the geostrophic balance.

Thus, the geostrophic balance reads:

$$F_C u_t^g = -\nabla^\perp p_t$$

Expanding the first equation of the SG approximation (1.1.3), while normalizing by setting $F_C = 1$ and inserting the geostrophic balance we have:

$$\begin{aligned} D_t u_t^g + \nabla p_t &= F_C u_t^\perp \\ \partial_t u_t^g + \langle u_t, \nabla \rangle u_t^g + \nabla p_t &= u_t^\perp \\ \partial_t \big(-\nabla^\perp p_t \big) + \langle u_t, \nabla \rangle \big(-\nabla^\perp p_t \big) + \nabla p_t &= u_t^\perp \\ -\partial_t \nabla^\perp p_t - \langle u_t, \nabla \rangle \nabla^\perp p_t + \nabla p_t &= u_t^\perp \end{aligned}$$

We now prove that $\partial_t \nabla^\perp p_t = (\partial_t \nabla p_t)^\perp$ and $\langle u_t, \nabla \rangle \nabla^\perp p_t = \big(\langle u_t, \nabla \rangle \nabla p_t \big)^\perp$

Indeed, since $(a, b)^\perp = (b, -a)$ we have:

$$\begin{aligned} \partial_t \nabla^\perp p_t &= (\partial_t \partial_2 p_t, -\partial_t \partial_1 p_t) \\ &= (\partial_t \partial_1 p_t, \partial_t \partial_2 p_t)^\perp \\ &= (\partial_t \nabla p_t)^\perp \end{aligned}$$

and

$$\begin{aligned} \langle u_t, \nabla \rangle \nabla^\perp p_t &= \left(\sum_{i=1}^{2} u_t^i \partial_i \partial_2 p_t, \sum_{i=1}^{2} u_t^i \partial_i (-\partial_1 p_t) \right) \\ &= \left(\sum_{i=1}^{2} u_t^i \partial_i \partial_2 p_t, -\sum_{i=1}^{2} u_t^i \partial_i \partial_1 p_t \right) \\ &= \left(\sum_{i=1}^{2} u_t^i \partial_i \partial_1 p_t, \sum_{i=1}^{2} u_t^i \partial_i \partial_2 p_t \right)^\perp \end{aligned}$$

$$= \big(\langle u_t, \nabla\rangle \nabla p_t\big)^{\perp}$$

So, the first equation of the SG approximation (1.1.3) becomes:

$$-(\partial_t \nabla p_t)^{\perp} - \big(\langle u_t, \nabla\rangle \nabla p_t\big)^{\perp} + \nabla p_t = u_t^{\perp}$$

Next, we claim that $\big(f^{\perp}\big)^{\perp} = -f$

Indeed, let $f = (f_1, f_2)$, then

$$\begin{aligned}\Big(f^{\perp}\Big)^{\perp} &= \big((f_1, f_2)^{\perp}\big)^{\perp} \\ &= \big((f_2, -f_1)\big)^{\perp} \\ &= (-f_1, -f_2) \\ &= -(f_1, f_2) \\ &= -f\end{aligned}$$

Also the perpendicular is a linear operator, that is $(f+g)^{\perp} = f^{\perp} + g^{\perp}$ and $(af)^{\perp} = af^{\perp}$

Indeed, let $f = (f_1, f_2)$ and $g = (g_1, g_2)$, then

$$\begin{aligned}(f+g)^{\perp} &= (f_1 + g_1, f_2 + g_2)^{\perp} \\ &= \big(f_2 + g_2, -(f_1 + g_1)\big) \\ &= (f_2 + g_2, -f_1 - g_1) \\ &= (f_2, -f_1) + (g_2, -g_1) \\ &= f^{\perp} + g^{\perp}\end{aligned}$$

and

$$\begin{aligned}(af)^{\perp} &= (af_1, af_2)^{\perp} \\ &= (af_2, -af_1) \\ &= a(f_2, -f_1) \\ &= af^{\perp}\end{aligned}$$

Thus, the first equation of the SG approximation (1.1.3) finally reads:

$$\partial_t \nabla p_t + \langle u_t, \nabla\rangle \nabla p_t + \overset{\perp}{\nabla} p_t = -u_t$$

which is the first equation of the SG system in physical variables (1.1.1). Adding the incompressibility condition div$u_t = 0$, we have derived the SG system.

One can find in bibliograpy that the SG system can be rewritten inserting a convex function, which is reasonable, in terms of physics, to consider. Simple calculations, as they will be shown below, will lead to a reformed SG system that "envelopes" the convexity requirement.

1.1.3 SG system and convexity

Energy considerations, as studied in [16] [14] [15], have shown that it is reasonable to assume that p_t is (-1)convex, meaning that

$$P_t(x) := p_t(x) + \frac{||x||^2}{2}$$

is convex.

From p_t to P_t

With P_t defined like this, we try to change our equations "substituting" p_t

We can prove that these four properties hold true:

i) $\nabla p_t = \nabla P_t \; - \; x$ ii) $\partial_t \nabla p_t = \partial_t \nabla P_t$ iii) $\langle u_t, \nabla \rangle x = u_t$
iv) $\nabla^{\perp} p_t = (\nabla P_t - x)^{\perp}$

Remark.
Usually we omit the argument x when writing functions.

Proof.

Indeed, somewhat simple and apparent computations lead to the desired:

i) $\nabla p_t = \nabla P_t \; - \; x$

$$\begin{aligned}
\nabla P_t = (\partial_1, \partial_2)(P_t) &= \big(\partial_1 P_t, \partial_2 P_t\big) \\
&= \left(\partial_1 \Big(p_t + \tfrac{||x||^2}{2}\Big) \; , \; \partial_2 \Big(p_t + \tfrac{||x||^2}{2}\Big) \right) \\
&= \left(\partial_1 p_t + \partial_1 \big(\tfrac{{x_1}^2 + {x_2}^2}{2}\big) \; , \; \partial_2 p_t + \partial_2 \big(\tfrac{{x_1}^2 + {x_2}^2}{2}\big) \right) \\
&= \big(\partial_1 p_t + x_1 \; , \; \partial_2 p_t + x_2\big) \\
&= (\partial_1 p_t, \partial_2 p_t) + (x_1, x_2) \\
&= (\partial_1, \partial_2)(p_t) + (x_1, x_2) \\
&= \nabla p_t + x \qquad \textbf{q.e.d.}
\end{aligned}$$

ii) $\partial_t \nabla p_t = \partial_t \nabla P_t$

Since p_t,P_t : $\mathbb{R}^2 \to \mathbb{R}$ both ∇p_t,∇P_t are vector fields from $\mathbb{R}^2$ to $\mathbb{R}^2$. We now want to differentiate with respect to the time variable t, thus we view our functions as

$\nabla p_x : \mathbb{R} \to \mathbb{R}^2$ and $\nabla P_x : \mathbb{R} \to \mathbb{R}^2$

So their time derivatives $\partial_t \nabla p_t$,$\partial_t \nabla P_t$ are matrices belonging in the space $\mathbb{R}^{2\times 1}$ i.e. they are column vectors. In order to avoid a lengthy proof we actually consider them as row vectors, since this consideration does not impact the arguments nor alter anything meaningful.

$$\begin{aligned}
\partial_t \nabla p_t &= \partial_t\big((\partial_1, \partial_2)(p_t) \big) \\
&= \partial_t\big((\partial_1 p_t, \partial_2 p_t) \big) \\
&= (\, \partial_t \partial_1 p_t \;,\; \partial_t \partial_2 p_t \,)
\end{aligned}$$

$$\begin{aligned}
\partial_t \nabla P_t &= (\partial_t \partial_1 P_t, \partial_t \partial_2 P_t) \\
&= \left(\; \partial_t \partial_1 \Big(p_t + \tfrac{||x||^2}{2} \Big) \;,\; \partial_t \partial_2 \Big(p_t + \tfrac{||x||^2}{2} \Big) \right) \\
&= \left(\; \partial_t \partial_1 p_t + \partial_t \partial_1 \tfrac{||x||^2}{2} \;,\; \partial_t \partial_2 p_t + \partial_t \partial_2 \tfrac{||x||^2}{2} \; \right) \\
&= (\; \partial_t \partial_1 p_t + \partial_t x_1 \;,\; \partial_t \partial_2 p_t + \partial_t x_2 \;) \\
&= (\; \partial_t \partial_1 p_t + 0 \;,\; \partial_t \partial_2 p_t + 0 \;) \;\; \textbf{q.e.d.}
\end{aligned}$$

iii) $\langle u_t, \nabla \rangle x = u_t$

Similarly we view the functions Id (that is x) and u_t as row vectors instead of column vectors.

$$\begin{aligned}
\langle u_t, \nabla \rangle (x) &= u_t^1 \partial_1(x) + u_t^2 \partial_2(x) \\
&= u_t^1 \partial_1(x_1, x_2) + u_t^2 \partial_2(x_1, x_2) \\
&= u_t^1 (1, 0) + u_t^2 (0, 1) \\
&= (u_t^1, 0) + (0, u_t^2) \\
&= (u_t^1, u_t^2) \\
&= u_t
\end{aligned}$$

iv) $\nabla^{\perp} p_t = (\nabla P_t - x)^{\perp}$

$$\begin{aligned}
\nabla^{\perp} p_t &= \nabla^{\perp}\Big(P_t - \tfrac{\|x\|^2}{2} \Big) \\
&= \left(\; \partial_2 \Big(P_t - \tfrac{\|x\|^2}{2} \Big) \;,\; -\partial_1 \Big(P_t - \tfrac{\|x\|^2}{2} \Big) \; \right) \\
&= \left(\; \partial_2 P_t + \partial_1 \Big(-\tfrac{\|x\|^2}{2} \Big) \;,\; -\partial_1 P_t - \partial_1 \Big(-\tfrac{\|x\|^2}{2} \Big) \; \right) \\
&= \left(\; \partial_2 P_t - \partial_1 \Big(\tfrac{\|x\|^2}{2} \Big) \;,\; -\partial_1 P_t + \partial_1 \Big(\tfrac{\|x\|^2}{2} \Big) \; \right) \\
&= \big(\partial_2 P_t - x_2 \;,\; -\partial_1 P_t + x_1 \big)
\end{aligned}$$

$$
\begin{aligned}
&= \big(\partial_2 P_t, -\partial_1 P_t\big) + (-x_2, x_1)\\
&= \big(\partial_2 P_t, -\partial_1 P_t\big) - (x_2, -x_1)\\
&= \nabla^{\perp} P_t - x^{\perp}\\
&= (\nabla P_t - x)^{\perp}
\end{aligned}
$$

which concludes the proof of the properties □

Substituting ii),iii),iv) on the SG system (1.1.1) and omitting the bar symbol over zero, we have:

$$
\begin{cases}
\partial_t \nabla P_t + \langle u_t, \nabla \rangle \nabla p_t + (\nabla P_t - x)^{\perp} + \langle u_t, \nabla \rangle x = 0 & (x,t) \in \mathbb{R}^2 \times [0, +\infty)\\
\mathrm{div} u_t = 0 & (x,t) \in \mathbb{R}^2 \times [0, +\infty)\\
P_t \text{ convex} & (x,t) \in \mathbb{R}^2 \times [0, +\infty)\\
P_0(x) = p_0(x) + \frac{\|x\|^2}{2} & x \in \mathbb{R}^2
\end{cases}
$$

Summing first equation's second term, which still includes the pressure p_t, with the fourth term we get:

$$
\begin{cases}
\partial_t \nabla P_t + \langle u_t, \nabla \rangle (\nabla p_t + x) + (\nabla P_t - x)^{\perp} = 0 & (x,t) \in \mathbb{R}^2 \times [0, +\infty)\\
\mathrm{div} u_t = 0 & (x,t) \in \mathbb{R}^2 \times [0, +\infty)\\
P_t \text{ convex} & (x,t) \in \mathbb{R}^2 \times [0, +\infty)\\
P_0 = p_0 + \frac{\|x\|^2}{2} & x \in \mathbb{R}^2
\end{cases}
$$

Using now i) and substituing p_0 with the initial data $\bar{p}$ we have the SG system involving convexity:

$$
\begin{cases}
\partial_t \nabla P_t + \langle u_t, \nabla \rangle \nabla P_t + (\nabla P_t - x)^{\perp} = 0 & (x,t) \in \mathbb{R}^2 \times [0, +\infty)\\
\mathrm{div} u_t = 0 & (x,t) \in \mathbb{R}^2 \times [0, +\infty)\\
P_t \text{ convex} & (x,t) \in \mathbb{R}^2 \times [0, +\infty)\\
P_0 = \bar{p} + \frac{\|x\|^2}{2} & x \in \mathbb{R}^2
\end{cases} \tag{1.1.4}
$$

with the boundary conditions that $P_t(x) - \frac{\|x\|^2}{2}$ and $u_t(x)$ are periodic.

1.2 The dual SG system

The two aforementioned SG systems (1.1.1) , (1.1.4) are rather "strange", due to the fact that they do not include anything resembling an evolution equation for the velocity u_t. Moreover tackling them seems quite difficult. For this reason we will proceed implementing the dual SG system.

Searching for an other evolution equation

We define the pushforward measure ρ_t of the Lebesgue measure on $\mathbb{R}^2$ by the vector field $\nabla P_t : \mathbb{R}^2 \to \mathbb{R}^2$.

$$\rho_t := (\nabla P_t)_{\#}(l^2) = \nabla P_{t\#} dx$$

that is $\forall\, t \geq 0$ and $\forall\, B \in \mathcal{B}(\mathbb{R}^2)$

$$\rho_t(B) = l^2\Big(\big(\nabla P_t\big)^{-1}(B)\Big)$$

A simplification of the notation

If no parentheses are used in a pushforward measure notation, then it is always implied that the "push function" is whatever appears before the # symbol and the measure comes after this.

Remark.
l^2 denotes the Lebesgue measure on $\mathbb{R}^2$, which (depeding again on the context, in an effort to make the presentation more well-received by the reader) we also denote as dx (especially when integrating).
We also denote $\mathcal{B}(\mathbb{R}^2)$ the Borel $\sigma-$algebra on $\mathbb{R}^2$, which is the smallest $\sigma-$algebra containing the open sets and a subcollection of the σ-algebra of Lebesgue measurable sets on $\mathbb{R}^2$ denoted as $\mathcal{M} \equiv \mathcal{M}_{l^2} \equiv \mathcal{L}(\mathbb{R}^2)$

The derivation of the dual SG system is formal, which means that enough smoothness (classic derivatives) and possibly several other requirements, allowing the calculations to be performed, are met by the quantities involved.

1.2.1 Continuity equation for measures with densities

The evolution equation we want to "achieve" is a continuity equation for ρ_t and U_t (which will be defined later on) i.e.

$$\partial_t \rho_t + \operatorname{div}(\rho_t U_t) = 0$$

Remark.
ρ_t is considered the dual density and U_t is considered the dual velocity. This means that they are density and velocity, respectively, in the space of dual variables.

In order to satisfy the evolution equation above for ρ_t, we have to make sense of it first. Since ρ_t is a measure, we will be understanding the equation in a weak sense.

We rewrite the continuity equation in a more general context and we formulate the equation that a solution (in the weak sense) has to satisfy.

Definition 1.1 (solution to the measure continuity equation)**.**

For every time $t \geq 0$ let $V_t : \mathbb{R}^n \to \mathbb{R}^n$ be a family of $L^1_{\text{loc}}(\mathbb{R}^n)$ functions and σ_t be a family of finite measures on $\mathbb{R}^n$, absolutely continous with respect to the Lebesgue measure l^n. We say that σ_t is a (weak) solution to the continuity equation

$$\partial_t \sigma_t + \text{div}(\sigma_t V_t) = 0$$

$:\iff \forall \varphi \in C_c^\infty(\mathbb{R}^n)$ the following two properties hold true:

$$\text{The function } h(t) = \int_{\mathbb{R}^n} \varphi \, d\sigma_t \text{ is differentiable}$$

$$\text{and} \quad \partial_t \int_{\mathbb{R}^n} \varphi \, d\sigma_t = \int_{\mathbb{R}^n} \langle \nabla \varphi, V_t \rangle \, d\sigma_t$$

Clarification 1.1.1.
The differentiality of the function h, stated in this definition, is the classic one. Even though the solution has been attributed the characterization weak.

Remark.
Note that the functions V_t are not assumed any differentiable at all (even in the weak sense). This will be explained now that we will derive the equation for the weak solution.

Deriving the equation of a solution to the measure continuity

We will follow the same strategy, one would follow to define the weak derivative of a function. We will calculate the integral of test functions with respect to the measure ρ_t. We use the fact that ρ_t has density (with respect to the Lebesgue measure) to obtain a time dependent function inside the integral. We differentiate over time and pass the time derivative inside the integral. We will then integrate by parts to find the desired.

Integration by parts formula can be found in the appendix of Evan's book [18]

Indeed, we (at least) formally deduce:

Let $\varphi \in C_c^\infty(\mathbb{R}^2)$, hence it is integrable (measurable and the integral is finite).

We know that each measure σ_t is absolutely continous with respect to the Lebesgue measure l^n (symbolically $\sigma_t \ll l^n$ or equivalently we also write $\sigma_t = \sigma_t \, dx$), thus there exists an l^n-a.e. unique function (the density, denoted by the same symbol) σ_t, for which it holds:

$$\int_{\mathbb{R}^n} \varphi \, d\sigma_t = \int_{\mathbb{R}^n} \varphi \sigma_t \, dx$$

$$\Rightarrow \partial_t \int_{\mathbb{R}^n} \varphi \, d\sigma_t = \partial_t \int_{\mathbb{R}^n} \varphi \sigma_t \, dx$$

Passing the differentiation inside the integral we have:

$$\partial_t \int_{\mathbb{R}^n} \varphi \, d\sigma_t = \int_{\mathbb{R}^n} \partial_t(\varphi \sigma_t) \, dx$$

Since $\varphi \in C_c^\infty(\mathbb{R}^2)$ has no time dependence we get:

$$\partial_t \int_{\mathbb{R}^n} \varphi \, d\sigma_t = \int_{\mathbb{R}^n} \varphi \, \partial_t \sigma_t \, dx$$

Assuming that the density σ_t satisfies the continuity equation, that is

$$\partial_t \sigma_t + \operatorname{div}(\sigma_t V_t) = 0$$

we get:

$$\partial_t \sigma_t = -\operatorname{div}(\sigma_t V_t)$$

Thus, we are lead to the following:

$$\partial_t \int_{\mathbb{R}^n} \varphi \, d\sigma_t = -\int_{\mathbb{R}^n} \varphi \operatorname{div}(\sigma_t V_t) \, dx$$

Since φ has compact support, there exists $r_0 > 0$ such that $\operatorname{supp}\varphi \subseteq \bar{B}(0, r_0)$. We then choose $r_1 > r_0$ and we set $U := \bar{B}(0, r_1)$. Due to the fact that $\{ \, x \in \mathbb{R}^2 \mid \varphi(x) = 0 \, \} \subseteq \operatorname{supp}\varphi \subseteq U$ we can rewrite the equation above as:

$$\partial_t \int_U \varphi \, d\sigma_t = -\int_U \varphi \operatorname{div}(\sigma_t V_t) \, dx$$

Performing integration by parts on the right hand side we have that:

$$\int_U \varphi \operatorname{div}(\sigma_t V_t) \, dx = \int_{\partial U} \varphi \langle \bar{n}, \sigma_t V_t \rangle \, dS - \int_U \langle \nabla \varphi, \sigma_t V_t \rangle \, dx$$

where $\bar{n}$ denotes the outward pointing unit normal vector field along the surface defined by the smooth boundary of U.

Due to the fact that $x \in \partial U \Rightarrow x \notin \operatorname{supp}\varphi$, the integral over the boundary equals zero.

Hence,

$$-\int_U \varphi \operatorname{div}(\sigma_t V_t) \, dx = \int_U \langle \nabla \varphi, \sigma_t V_t \rangle \, dx$$

$$= \int_U \sigma_t \langle \nabla\varphi, V_t \rangle \, dx$$

since the density σ_t is a real-valued function

Density's integral property implies:

$$\int_U \sigma_t \langle \nabla\varphi, V_t \rangle \, dx = \int_U \langle \nabla\varphi, V_t \rangle \, d\sigma_t$$

Thus, we have:

$$\partial_t \int_U \varphi \, d\sigma_t = \int_U \langle \nabla\varphi, V_t \rangle \, d\sigma_t$$

Finally, because every integrand (integrated quantity) becomes zero (since it involves the compactly supported φ) outside of U (that is the complement of U in $\mathbb{R}^n$) we get:

$$\partial_t \int_{\mathbb{R}^n} \varphi \, d\sigma_t = \int_{\mathbb{R}^n} \langle \nabla\varphi, V_t \rangle \, d\sigma_t$$

which is the property that needs to be satisfied for a time dependent family of measures σ_t, in order to be a solution for the continuity equation

$$\partial_t \sigma_t + \operatorname{div}(\sigma_t V_t) = 0$$

with known V_t

In fact, the opposite direction is also true. This can be shown using the same method (formally).

Let σ_t, V_t satisfy the following proprety for every $\varphi \in C_c^\infty(\mathbb{R}^2)$:

$$\partial_t \int_{\mathbb{R}^n} \varphi \, d\sigma_t = \int_{\mathbb{R}^n} \langle \nabla\varphi, V_t \rangle \, d\sigma_t$$

Assume that σ_t has a density, denoted by the same symbol.

Then, passing the differentiation inside the integral and integrating by parts (like above) we get respectively:

$$\partial_t \int_{\mathbb{R}^n} \varphi \, d\sigma_t = \partial_t \int_{\mathbb{R}^n} \varphi \sigma_t \, dx$$

$$= \int_{\mathbb{R}^n} \varphi \, \partial_t \sigma_t \, dx$$

and

$$\int_{\mathbb{R}^n} \langle \nabla\varphi, V_t\rangle \, d\sigma_t = \int_{\mathbb{R}^n} \sigma_t \langle \nabla\varphi, V_t\rangle \, dx$$
$$= \int_{\mathbb{R}^n} \langle \nabla\varphi, \sigma_t V_t\rangle \, dx$$
$$= -\int_{\mathbb{R}^n} \varphi \mathrm{div}(\sigma_t V_t) \, dx$$

So, combinig we have:

$$\int_{\mathbb{R}^n} \varphi \, \partial_t \sigma_t \, dx = -\int_{\mathbb{R}^n} \varphi \mathrm{div}(\sigma_t V_t) \, dx$$

Thus, for all $\varphi \in C_c^\infty(\mathbb{R}^2)$:

$$\int_{\mathbb{R}^n} \varphi \big(\partial_t \sigma_t + \mathrm{div}(\sigma_t V_t)\big) \, dx = 0$$

Hence, we are lead to the satisfaction of the measure (with density) continuity equation:

$$\partial_t \sigma_t + \mathrm{div}(\sigma_t V_t) = 0$$

1.2.2 Formal passage from SG to dual SG

Resuming back to our target, that is to find an evolution equation (the continuity equation we have mentioned earlier).

We let $\varphi \in C_c^\infty(\mathbb{R}^2)$ and calculate:

$$\partial_t \int_{\mathbb{R}^n} \varphi \, d\rho_t$$

The pushforward measure $\rho_t = \nabla P_{t\#} dx$ satisfies a property similar to the change of variables [1], that is the following equality for all t:

$$\int_{\mathbb{R}^2} \varphi \, d\rho_t \quad = \int_{(\nabla P_t)^{-1}(\mathbb{R}^2)} \varphi \circ \nabla P_t \; dx$$

[1] See PropositionA.23 for the "pushforward change of variables"

Due to the fact that the pre-image (inverse image) of the whole space is the entire domain of the function i.e. $(\nabla P_t)^{-1}(\mathbb{R}^2) = D_{\nabla P_t} = \mathbb{R}^2$ we get:

$$\int_{\mathbb{R}^2} \varphi \, d\rho_t = \int_{\mathbb{R}^2} \varphi \circ \nabla P_t \, dx$$

Then, we pass the differentiation inside the integral [2] to obtain:

$$\partial_t \int_{\mathbb{R}^2} \varphi \circ \nabla P_t \, dx = \int_{\mathbb{R}^2} \partial_t \big(\varphi \circ \nabla P_t\big) \, dx$$

Thus, we have so far:

$$\partial_t \int_{\mathbb{R}^2} \varphi \, d\rho_t = \int_{\mathbb{R}^2} \partial_t \big(\varphi \circ \nabla P_t\big) \, dx$$

We now proceed to calculate the time derivative of the composition $\varphi \circ \nabla P_t$.

To do that in the right way, we have to view the involved vector-valued functions "like we should" i.e. as column vectors.

Since we want to differentiate with respect to time, we view the function $\nabla P(x,t)$ as the time function $\nabla P_x(t)$. So, we have the following:

$$\begin{aligned} &\nabla P_x : \mathbb{R} \to \mathbb{R}^2 \\ &\Rightarrow \partial_t \nabla P_x : \mathbb{R} \to \mathbb{R}^{2\times 1} \\ &\Rightarrow \partial_t \nabla P_t : \mathbb{R}^2 \to \mathbb{R}^{2\times 1} \end{aligned}$$

recall that, since there is no subscript under the nabla, the derivative of φ stated in the chain rule is its spatial (only) derivative $\nabla\varphi = D\varphi \in \mathbb{R}^{1\times 2}$

More on the convention

For all the computations below (until the end of proof at least), we will clarify (we will do this "over-clarification" of the dimensions only in the introductory first chapter) when a vector-valued function on $\mathbb{R}^2$

1. is considered as a column vector on $\mathbb{R}^{2\times 1}$ (usually when it is identifed as a vector-valued function by definition)

2. and when it is viewed as a row vector on $\mathbb{R}^{1\times 2}$ (usually when it is identifed as the derivative of a real-valued function)

[2] This is the Liebniz integral rule, which holds true under the assumptions of PropositionA.31

We return back to the composition $\varphi \circ \nabla P_t$, where $\varphi : \mathbb{R}^2 \to \mathbb{R}$ and we are wiewing $\nabla P_x : \mathbb{R}_0^+ \to \mathbb{R}^2$ with respect to its time dependence.

Differentiating with respect to time t and applying the chain rule [3] we get:

$$\partial_t\big(\varphi(\nabla P_t)\big) \overset{\text{chain rule}}{=} \nabla\varphi(\nabla P_t) \diamond \partial_t \nabla P_t$$

Then, isolating, in the first equation of the SG system with convexity (1.1.4) the first term, we obtain that:

$$\partial_t \nabla P_t = -\langle u_t, \nabla\rangle \nabla P_t - (\nabla P_t - x)^{\perp}$$

Hence, the equality $\partial_t\big(\varphi(\nabla P_t)\big) = \nabla\varphi(\nabla P_t) \diamond \partial_t \nabla P_t$ becomes:

$$\partial_t\big(\varphi(\nabla P_t)\big) = \nabla\varphi(\nabla P_t) \diamond \Big(-\langle u_t, \nabla\rangle \nabla P_t - (\nabla P_t - x)^{\perp}\Big) \Rightarrow$$

$$\partial_t\big(\varphi(\nabla P_t)\big) = -\nabla\varphi(\nabla P_t) \diamond \langle u_t, \nabla\rangle \nabla P_t \; - \nabla\varphi(\nabla P_t) \diamond (\nabla P_t - x)^{\perp}$$

Remark.
Note that for the matrix multiplication to be well-defined (that means we must have the right dimensions e.g. $k \times l$, $l \times m$) we have to consider $\partial_t \nabla P_t = -\langle u_t, \nabla\rangle \nabla P_t - (\nabla P_t - x)^{\perp}$ as a column vector in $\mathbb{R}^{2\times 1}$ since φ is in the space $\mathbb{R}^{1\times 2}$.
As we have said earlier, in order to avoid the use of transpose and "cut in length" of the presentation, we identify vector-valued functions as either column vectors or row vectors while we can (that is as long as nothing is impacted by that consideration) and when it is no more unavoidable we will "roll back" to the dimension where we should have from the beginning.

Since we already have

$$\partial_t \int_{\mathbb{R}^2} \varphi \, d\rho_t = \int_{\mathbb{R}^2} \partial_t\big(\varphi(\nabla P_t)\big) \, dx$$

From the above equality involving $\partial_t\big(\varphi(\nabla P_t)\big)$ we obtain that:

$$\partial_t \int_{\mathbb{R}^2} \varphi \, d\rho_t = \int_{\mathbb{R}^2} - \nabla\varphi(\nabla P_t) \diamond \langle u_t, \nabla\rangle \nabla P_t - \nabla\varphi(\nabla P_t) \diamond (\nabla P_t - x)^{\perp} \; dx$$

$$= -\int_{\mathbb{R}^2} \nabla\varphi(\nabla P_t) \diamond \langle u_t, \nabla\rangle \nabla P_t \, dx - \int_{\mathbb{R}^2} \nabla\varphi(\nabla P_t) \diamond (\nabla P_t - x)^{\perp} \; dx$$

We continue evaluating each quantity seperately.

Before we do so, we will briefly discuss our plan. This "conversation" is a complement to the computations below and not a stand alone proof.

[3]See PropositionA.36 for the details of the chain rule

One target is to show that the first quantity's integral equals zero. We will start by showing that $\nabla\varphi(\nabla P_t)\diamond\Big(\langle u_t,\nabla\rangle\nabla P_t\Big)$ is equal with the inner product of the functions $\nabla(\varphi\circ\nabla P_t)$ and u_t. Writing the standard inner product as its (by definition) sum of the respective component elements, we will then use integration by parts and the incompressibility ($\text{div}u_t = 0$) of the fluid to obtain the wanted result.

The other target is to show that the second quantity equals the inner product of $\nabla\varphi$ and U_t composed with the function ∇P_t (calculated at that point), where U_t is the new "dual" velocity vector field defined "through the help" of the Legendre transform P_t^* for the convex "pressure" P_t. The aforementioned equality is shown using fact that ∇P_t and ∇P_t^* are reverse functions.

At last, we will prove that the newly defined dual velocity U_t is divergence free, which combined with the change of variables for the pushforward measure $\rho_t = \nabla P_{t\#}\, dx$ (this enables us to return in integration with respect to ρ_t instead of dx, since the inner product of $\nabla\varphi$ and U_t is composed with the vector field ∇P_t) shows that the second quantity's l^2-integral equals the ρ_t-integral of $\text{div}\varphi U_t$. This, in turn, will allow us to reach our final destination i.e.

$$\partial_t\int\limits_{\mathbb{R}^2}\varphi\, d\rho_t = \int\limits_{\mathbb{R}^2}\langle\varphi, U_t\rangle\, d\rho_t$$

This equality leads to an evolution equation for ρ_t, that is the measure ρ_t satisfies the continuity equation Definition1.1 in the weak sense we have already discussed.

We begin our computations with the first quantity, starting with $\langle u_t,\nabla\rangle\nabla P_t$

$$\begin{aligned}
\langle u_t,\nabla\rangle\nabla P_t &= u_t^1\cdot\partial_1\nabla P_t + u_t^2\cdot\partial_2\nabla P_t\\
&= u_t^1\partial_1\left(\partial_1 P_t,\partial_2 P_t\right) + u_t^2\partial_2\left(\partial_1 P_t,\partial_2 P_t\right)\\
&= \left(\ u_t^1\partial_1\partial_1 P_t\ ,\ u_t^1\partial_1\partial_2 P_t\ \right) + \left(\ u_t^2\partial_2\partial_1 P_t\ ,\ u_t^2\partial_2\partial_2 P_t\ \right)\\
&= \left(\ u_t^1\partial_1^2 P_t\ ,\ u_t^1\partial_1\partial_2 P_t\ \right) + \left(\ u_t^2\partial_2\partial_1 P_t\ ,\ u_t^2\partial_2^2 P_t\ \right)\\
&= \left(\ u_t^1\partial_1^2 P_t + u_t^2\partial_2\partial_1 P_t\ ,\ u_t^1\partial_1\partial_2 P_t + u_t^2\partial_2^2 P_t\ \right)\\
&= (u_t^1, u_t^2)\diamond\begin{pmatrix}\partial_1^2 P_t & \partial_1\partial_2 P_t\\ \partial_2\partial_1 P_t & \partial_2^2 P_t\end{pmatrix}
\end{aligned}$$

recalling our previous "conversation" (the convention) about the dimensions, we are actually interested in the transpose matrix of the above product. Moreover due to the convention that when we write a vetor-valued function we mean either the row vector or the column vector notation. Here, we view u_t as a column vector.

So,

$$\begin{aligned}
\langle u_t,\nabla\rangle\nabla P_t &= \begin{pmatrix}\partial_1^2 P_t & \partial_2\partial_1 P_t\\ \partial_1\partial_2 P_t & \partial_2^2 P_t\end{pmatrix}\diamond\begin{pmatrix}u_t^1\\ u_t^2\end{pmatrix}\\
&= D^2 P_t\diamond u_t
\end{aligned}$$

Thus, for the first quantity we have

$$\begin{aligned}\nabla\varphi(\nabla P_t)\diamond\langle u_t,\nabla\rangle\nabla P_t &= \nabla\varphi(\nabla P_t)\diamond D^2P_t\diamond u_t\\ &= D\varphi(\nabla P_t)\diamond D\left(\nabla P_t\right)\diamond u_t\\ &= D\left(\varphi\circ\nabla P_t\right)\diamond u_t\\ &= \nabla\left(\varphi\circ\nabla P_t\right)\diamond u_t\end{aligned}$$

Setting $h := \varphi\circ\nabla P_t$, we obtain:

$$\nabla\varphi(\nabla P_t)\diamond\langle u_t,\nabla\rangle\nabla P_t = \nabla h\diamond u_t$$

Since $\varphi : \mathbb{R}^2\to\mathbb{R}$ and $\nabla P_t : \mathbb{R}^2\to\mathbb{R}^2$

their composition $h = \varphi\circ\nabla P_t : \mathbb{R}^2\to\mathbb{R}$ is a real-valued function.

Hence, its spatial derivative $\nabla h : \mathbb{R}^2\to\mathbb{R}^{1\times 2}$ is a (row) vector-valued function.

The matrix multiplication $\nabla h\diamond u_t$ can be viewed as the inner product of the vector-valued function ∇h with the row vecotr (u_t^1, u_t^2) , which we can also denote u_t as well, due to the convention of identifying the space $\mathbb{R}^2$ as either $\mathbb{R}^{1\times 2}$ or $\mathbb{R}^{2\times 1}$ when it comes to the values of a vector-valued function.

Thus, we have shown for the first quantity that:

$$\nabla\varphi(\nabla P_t)\diamond\langle u_t,\nabla\rangle\nabla P_t = \big\langle\nabla h, u_t\big\rangle$$

We now integrate to obtain:

$$\int\limits_{\mathbb{R}^2}\nabla\varphi(\nabla P_t)\diamond\langle u_t,\nabla\rangle\nabla P_t\,dx = \int\limits_{\mathbb{R}^2}\big\langle\nabla h, u_t\big\rangle\,dx$$

The next (and last one regarding the first quantity) “move” is to show that:

$$\int\limits_{\mathbb{R}^2}\big\langle\nabla h, u_t\big\rangle\,dx = 0$$

Indeed, since φ has compact support so does h, so (similarly with the argument followed in subsubsection “Deriving the equation for a weak solution” when integrating by parts too) there exists $B(0,r) := U \supseteq \text{supp}h$

Integrating by parts we get:

$$\int\limits_{U}\big\langle\nabla h, u_t\big\rangle\,dx = \int\limits_{\partial U}h\langle\bar{n}, u_t\rangle\,dS - \int\limits_{U}h\cdot\text{div}u_t\,dx$$

$h \equiv 0$ on ∂U , because of the inclusion supp$h \subseteq U$ and moreover the velocity u_t satisfies the incompressibility condition div$u_t = 0$

Hence we are lead to the fulfilment of the first target.

Thus, we are now left with the term:

$$\partial_t \int_{\mathbb{R}^2} \varphi \, d\rho_t = -\int_{\mathbb{R}^2} \nabla\varphi(\nabla P_t) \diamond (\nabla P_t - x)^{\perp} \; dx$$
$$= \int_{\mathbb{R}^2} \nabla\varphi(\nabla P_t) \diamond (x - \nabla P_t)^{\perp} \; dx$$

Because $\nabla\varphi \in \mathbb{R}^{1\times 2}$ and $x - \nabla P_t \in \mathbb{R}^{1\times 2}$ we can view again the matrix multiplication as the (standard) inner product i.e.

$$\nabla\varphi(\nabla P_t) \diamond (x - \nabla P_t)^{\perp} = \big\langle \nabla\varphi(\nabla P_t), (x - \nabla P_t)^{\perp} \big\rangle$$

Now, we define the Legendre transform (sometimes also called the convex conjugate) of the function P_t i.e.

$$P_t^*(y) := \sup_{x\in\mathbb{R}^2} \big(\langle y, x\rangle - P_t(x)\big)$$

The property we are going to use, in order to achieve our target is the fact that ∇P_t and ∇P_t^* are inverse functions. This result holds true under some assumptions which are mentioned in the appendix (at the corresponding section) and we assume that are satisfied.

Thus, we can write x as $\nabla P_t^*\big(\nabla P_t(x)\big)$, which we abbreviate (omitting the argument variable x) as $\nabla P_t^*(\nabla P_t)$.

So,

$$\nabla\varphi(\nabla P_t) \diamond (x - \nabla P_t)^{\perp} = \Big\langle \; \nabla\varphi(\nabla P_t) \;, \; \big(\nabla P_t^*(\nabla P_t) - \nabla P_t\big)^{\perp} \; \Big\rangle$$
$$= \Big\langle \; \nabla\varphi \circ \nabla P_t \;, \; \big((\nabla P_t^* - Id) \circ \nabla P_t\big)^{\perp} \; \Big\rangle$$
$$= \Big\langle \; \nabla\varphi \, , \, (\nabla P_t^* - Id)^{\perp} \; \Big\rangle \circ \nabla P_t$$

Defining the velocity vector field in the dual space as:

$$U_t := (\nabla P_t^* - Id)^{\perp}$$

we get:

$$\nabla\varphi(\nabla P_t) \diamond (x - \nabla P_t)^{\perp} = \big\langle \nabla\varphi, U_t \big\rangle \circ \nabla P_t$$

Hence, integrating over $\mathbb{R}^2$ we obtain:

$$\int_{\mathbb{R}^2} \nabla\varphi(\nabla P_t) \diamond (x - \nabla P_t)^{\perp} \; dx = \int_{\mathbb{R}^2} \big\langle \nabla\varphi, U_t \big\rangle \circ \nabla P_t \; dx$$

Due to the definition of ρ_t as the pushforward measure $\nabla P_{t\#}dx$ using the formula[4] for change of variables through the pushforward measure we have that:

$$\int_{\mathbb{R}^2} \langle \nabla\varphi, U_t \rangle \circ \nabla P_t \, dx = \int_{\mathbb{R}^2} \langle \nabla\varphi, U_t \rangle \, d\rho_t$$

Thus, we have reached to this:

$$\partial_t \int_{\mathbb{R}^2} \varphi \, d\rho_t = \int_{\mathbb{R}^2} \langle \nabla\varphi, U_t \rangle \, d\rho_t$$

Since this is shown for every φ we are lead to the continuity equation:

$$\partial_t \rho_t + \mathrm{div}(\rho_t U_t) = 0$$

Gathering all the data we have the system:

$$\begin{cases} \partial_t \rho_t + \mathrm{div}(\rho_t U_t) = 0 & (x,t) \in \mathbb{R}^2 \times [0, +\infty) \\ U_t = (\nabla P_t^* - Id)^{\perp} & (x,t) \in \mathbb{R}^2 \times [0, +\infty) \\ \rho_t = \nabla P_{t\#} dx & t \in [0, +\infty) \\ P_0 = \bar{p} + \frac{\|x\|^2}{2} & x \in \mathbb{R}^2 \end{cases}$$

Note that the last equation is just the relation between the initial data of the dual SGsystem and the classic (in physical variables) SG system, as such we don't have to include it in the dual SG system description. We only need to define $\bar{p}$ or P_0 respectively satisfying this equality in order to pass from one SG system formulation to the other.

Remark.
We do not cover the backwards passage, from the dual SG system to the classic SG system

Thus the dual SG system is the following:

$$\begin{cases} \partial_t \rho_t + \mathrm{div}(\rho_t U_t) = 0 & (x,t) \in \mathbb{R}^2 \times [0, +\infty) \\ U_t = (\nabla P_t^* - Id)^{\perp} & (x,t) \in \mathbb{R}^2 \times [0, +\infty) \\ \rho_t = \nabla P_{t\#} dx & t \in [0, +\infty) \end{cases} \tag{1.2.1}$$

Velocity of dual SG equation is divergence free

The dual velocity U_t is divergence free i.e. $\mathrm{div} U_t = 0$ and it satisfies the property

[4] See PropositionA.23

$\langle\varphi, U_t\rangle = \text{div}(\phi U_t)$

Both relations follow from two more general results, which we will state and prove now.

Proposition 1.1 (the rotated gradient of a function is divergence free)**.**

Let $U : \mathbb{R}^2 \to \mathbb{R}^2$ be a function which is written as the rotated gradient of a real valued C^2 function i.e. $U = (\nabla f)^\perp$ with $C^2(\mathbb{R}^2) \ni f : \mathbb{R}^2 \to \mathbb{R}$, then $\text{div}U = 0$

Proof.

Since $f : \mathbb{R}^2 \to \mathbb{R}$ is C^2 we have

$$\nabla f \equiv Df : \mathbb{R}^2 \to \mathbb{R}^2 \text{ with } \nabla f = (d_1 f, d_2 f)$$

Thus,

$$U = (\nabla f)^\perp = (d_2 f, -d_1 f)$$

$$\begin{aligned}\Rightarrow \text{div}U &= d_1 d_2 f + d_2(-d_1 f)\\ &= d_1 d_2 f - d_2 d_1 f\\ &= d_1 d_2 f - d_1 d_2 f\\ &= 0\end{aligned}$$

□

Remark.
In terms of the dual SG system we have that

$$\begin{aligned}U_t &= (\nabla P_t^* - Id)^\perp\\ &= \left(\nabla\left(P_t - \frac{||Id||^2}{2}\right)\right)^\perp\end{aligned}$$

where setting f equal to

$$P_t - \frac{||\cdot||^2}{2} : \mathbb{R}^2 \to \mathbb{R}$$

which can be assumed C^2 since the passage is formal, implies that for each $t \in [0, +\infty)$ the velocity vector field U_t is divergence free.

The second property of the dual velocity is implied from the following:

Proposition 1.2. Let $w : W$ open $\subseteq \mathbb{R}^n \to \mathbb{R}^n$ and $g : V$ open $\subseteq \mathbb{R}^n \to \mathbb{R}$ be two partially differentiable functions with $w = (w_1, \ldots, w_n)$, then

$$\text{div}(g \cdot w) = g \cdot \text{div}w + \langle \nabla g, w\rangle$$

Proof.

$$\begin{aligned}
\mathrm{div}(g\cdot w) &= \partial_1(g\cdot w_1)+\cdots+\partial_n(g\cdot w_n)\\
&= g\cdot\partial_1 w_1+\partial_1 g\cdot w_1+\cdots+g\cdot\partial_n w_n+\partial_n g\cdot w_n\\
&= g\cdot(\partial_1 w_1+\cdots+\partial_n w_n)+\partial_1 g\cdot w_1+\cdots+\partial_n g\cdot w_n\\
&= g\cdot\sum_{i=1}^{n}\partial_i w_i+\sum_{i=1}^{n}\partial_i g\cdot w_i\\
&= g\cdot \mathrm{div} w+\langle\nabla g, w\rangle
\end{aligned}$$

□

Corollary 1.2.1. In particular, if div$w$=0, then

$$\mathrm{div}(g\cdot w)=\langle\nabla g, w\rangle$$

Setting $g=\phi$ and $w=U_t$ for all $t\geq 0$, which is divergence free (as we have just proved that div$U_t=0$) , we have shown the second one.

CHAPTER 2

Global in time weak solutions for the dual SG system

Now we focus our attention on solving the dual SG system (1.2.1). We can show that there exists indeed, globally in time, at least one weak solution for our problem.

Before we do so, we must first introduce what we call a weak solution for the dual SG system:

$$\begin{cases} \partial_t \rho_t + \operatorname{div}(\rho_t U_t) = 0 & (x,t) \in \mathbb{R}^2 \times [0,+\infty) \\ U_t = (\nabla P_t^* - Id)^{\perp} & (x,t) \in \mathbb{R}^2 \times [0,+\infty) \\ \rho_t = \nabla P_{t\#} dx & t \in [0,+\infty) \\ \text{with initial data } \rho_0 & \end{cases}$$

2.1 Formulation of weak solution for the dual SG

We change our view to Lagrangian coordinates, we consider the particle trajectory for the particles of the fluid (aquatic or atmospheric) which we study.

That is, we view the space variable x as a time dependent function $X(t) \in \mathbb{R}^2$ initially located at $x \in \mathbb{R}^2$.

$t \mapsto X(t)$ is called: space trajectory of the fluid particle being at x initially.

Since the velocity is the time derivative of the displacement (change in position), $X(t)$ must satisfy:

$$\begin{cases} \partial_t X(t) = u_t\big(X(t)\big) = u\big(X(t), t\big) \\ X(0) = x \end{cases}$$

One would expect a particle, starting its movement at a specific point of the space, to follow one unique trajectory.

Let us assume that we can uniquely solve this ordinary differential equation for each y, and let us call the solution $X_y(t)$ (since it is dependent on the particular y which we solved it for)

Then, we deduce that the map sending x to $X_x(t)$ is a function, due to the fact that for every x the solution $X_x(t)$ is unique.

We denote $X_x(t)$ also as $X(y,t)$ and $X_t(x)$

Hence, $X(x,t)$ satisfies:

$$\begin{cases} \partial_t X(x,t) = u\big(X(x,t),t\big) \\ X(x,0) = x \end{cases}$$

which we abbreviate like we usually do (omitting the space variable and putting the subscript t to denote time dependence) writing:

$$\begin{cases} \partial_t X_t = u_t(X_t) \\ X_0 = Id \end{cases}$$

We present one important property of the particle trajectory, which is also true in $\mathbb{R}^n$

For all times t the function X_t is measure preserving

Proposition 2.1 (flow is measure preserving)**.** Let $X : \mathbb{R}^n \times [0,+\infty) \to \mathbb{R}^n$ be a smooth function where ∇X_x is invertible for all x i.e. for every x the map $t \mapsto X_x(t)$ is invertible with $\big(\nabla X_x\big)^{-1}$ being the inverse. Also the following is satisfied:

$$\begin{cases} \partial_t X(x,t) = u\big(X(x,t),t\big) \\ X(x,0) = x \end{cases}$$

then

$$\det(\nabla X_t) = 1 \;\; \forall t \in [0,+\infty)$$

Proof.

The Jacobi formula says that, if we consider a matrix A with coefficients depending on time i.e. we can view it as a matrix-valued function $A(t) \equiv A_t$ then

$$\partial_t\big(\det(A_t)\big) = \mathrm{tr}(\mathrm{adj}A_t \diamond \partial_t A_t)$$

and if A_t is invertible

$$\partial_t\big(\det(A_t)\big) = \det(A_t) \cdot \mathrm{tr}\big(A_t^{-1} \diamond \partial_t A_t\big)$$

With this, knowing that X_t is smooth enough and $\nabla X_x(t)$ is invertible with $\big(\nabla X_x\big)^{-1}$ being the inverse, we get:

$$\partial_t\Big(\det\big(\nabla X_t\big)\Big) = \det\big(\nabla X_t\big)\cdot \operatorname{tr}\Big(\big(\nabla X_x\big)^{-1} \diamond \partial_t \nabla X_t\Big)$$

We recall that the flow X_t satisfies the system:

$$\begin{cases} \partial_t X_t = U_t\big(X_t\big) \\ X_0 = Id \end{cases}$$

Differentiating with respect to x, the chain rule and the identity $\nabla\partial_t = \partial_t\nabla$ give the following:

$$\begin{cases} \nabla\partial_t X_t = \nabla U_t\big(X_t\big) \diamond \nabla X_t \\ \det\big(\nabla X_0\big) = 1 \end{cases}$$

Using the first equation, the Jacobi formula now reads:

$$\partial_t\Big(\det\big(\nabla X_t\big)\Big) = \det\big(\nabla X_t\big)\cdot \operatorname{tr}\Big(\big(\nabla X_x\big)^{-1} \diamond \nabla U_t\big(X_t\big) \diamond \nabla X_t\Big)$$

where $\nabla X_t = \nabla X(x,t) = \nabla X_x$

We also know that for any square matrices A, B with B ivertible, the trace satisfies the equality

$$\operatorname{tr}\big(B^{-1}AB\big) = \operatorname{tr}(A)$$

hence:

$$\partial_t\Big(\det\big(\nabla X_t\big)\Big) = \det\big(\nabla X_t\big)\cdot \operatorname{tr}\Big(\nabla U_t\big(X_t\big)\Big)$$

The trace of a matrix satisfies one more property, which comes in handy:

$$\operatorname{tr}(\nabla f) = \operatorname{div} f$$

Implementing this, the Jacobi formula finally becomes:

$$\partial_t\Big(\det\big(\nabla X_t\big)\Big) = \det\big(\nabla X_t\big)\cdot \operatorname{div}\Big(U_t\big(X_t\big)\Big)$$

because $\operatorname{div} U_t = 0$ and the functions are all defined for t in the closed and connected $[0, +\infty)$ we get that $\det\big(\nabla X_t\big)$ is constant with respect to time.

This implies that it is equal to its value at any specific value of t, in particular for $t = 0$, we get that:

$$\text{For every } t \in [0, +\infty)$$
$$\det\big(\nabla X_t\big) = \det\big(\nabla X_0\big) = 1$$

□

Now what we wanted for X_t, that is, measure preservation, will follow from the (below) corollary (of the proposition above).

Corollary 2.1.1. Assume that X_t is $1-1$ and onto $\mathbb{R}^n$, then

$$\int\limits_{\mathbb{R}^n} f(y)\,dy = \int\limits_{\mathbb{R}^n} f\big(X_t(x)\big)\,dx$$

or equivalenty written

$$\int\limits_{\mathbb{R}^n} f\,dy = \int\limits_{\mathbb{R}^n} f(X_t)\,dx$$

for all Lebesgue measurable functions i.e. $f \in L^1(\mathbb{R}^n)$

Proof.

Since $\det(\nabla X_t) = 1$, the change of variable $y = \nabla X_t(x)$ implies:

$$\int\limits_{\nabla X_t(\mathbb{R}^n)} f(y)\,dy = \int\limits_{\mathbb{R}^n} f\big(X_t(x)\big)\,dx$$

∇X_t being onto $\mathbb{R}^n$ means that $\nabla X_t(\mathbb{R}^n) = \mathbb{R}^n$ □

At last, setting $f = \chi_{X_t(\Omega)}$ leads to measure preservation $l^2\big(X_t(\Omega)\big) = l^2(\Omega)$

where χ is the characteristic function of the set noted on its subscript.

$$\chi_S(x) := \begin{cases} 1 & x \in S \\ 0 & x \notin S \end{cases}$$

Because it holds true that $\chi_{X_t(\Omega)}\big(X_t(x)\big) = \chi_\Omega(x)$ for every set Ω

We resume back on finding an equation that a weak solution of the dual SG equation has to satisfy.

Let $\xi \in C_c^\infty(\mathbb{R}^2 \times [0, +\infty))$

We are interested in the time derivative of the function $\xi\Big(\nabla P\big(X(x,t),t\big),t\Big)$ which like usual we abbreviate as $\xi_t\big(\nabla P_t(X_t)\big)$

The reason we "are led to" do that is because we know that ∇P_t satisfies the 1st equation of SG involving convexity (1.1.4) i.e.

$$\partial_t \nabla P_t + \langle u_t, \nabla\rangle \nabla P_t = (x - \nabla P_t)^\perp$$

Setting x to be $X(x,t)$ implies the following identity:

$$\partial_t \nabla P_t(X_t) + \langle u_t(X_t), \nabla \rangle \nabla P_t(X_t) = \big(X_t - \nabla P_t(X_t)\big)^{\perp}$$

where the quantity of each hand side is calculated at the point x, which we usually omit.

Proposition 2.2. Let $f : \mathbb{R}^2 \times [0,+\infty) \to \mathbb{R}^2$ smooth then

$$\begin{aligned}\partial_t\big(f_t(X_t)\big) &= \partial_t f_t(X_t) + \langle u_t(X_t), \nabla \rangle f_t(X_t) \\ &\underset{x=X_t}{\overset{\text{since}}{=}} \partial_t f_t + \langle u_t, \nabla \rangle f_t\end{aligned}$$

Proof. Let

$$f(x,t) = \big(f_1(x,t), f_2(x,t)\big)$$

and consider the auxiliary function:

$$\begin{aligned}g : \mathbb{R}^2 \times [0,+\infty) &\to \mathbb{R}^2 \times [0,+\infty) \\ g(x,t) :=& (X(x,t), t)\end{aligned}$$

where

$$\begin{aligned}X :=& (X_1, X_2) \\ g_i :=& X_i \text{ for } i = 1,2 \\ g_3(x,t) :=& t\end{aligned}$$

or equivalently written (all of the above) with the subscript t and omitting the space variable x

$$\begin{aligned}f_t =& (f_t^1, f_t^2) \\ g_t :=& (X_t \,, t) \\ X_t :=& (X_t^1, X_t^2) \\ g_t^i :=& X_t^i \text{ for } i = 1,2 \\ g_t^3 :=& t\end{aligned}$$

And the chain rule implies:

Remark.

Here D refers to the differentiation with respect to space and time, while $\nabla = (\partial_1, \partial_2)$

$$D(f \circ g) = Df(g) \diamond Dg$$

with

$$f \circ g : \mathbb{R}^2 \times [0,+\infty) \to \mathbb{R}^2$$

$$f \circ g := \big((f \circ g)_1, (f \circ g)_1\big)$$

where for each function the derivative (with respect to both space and time) is as follows:

$$\begin{aligned} D(f \circ g) &= \begin{pmatrix} \partial_1(f \circ g)_1 & \partial_2(f \circ g)_1 & \partial_t(f \circ g)_1 \\ \partial_1(f \circ g)_2 & \partial_2(f \circ g)_2 & \partial_t(f \circ g)_2 \end{pmatrix} \\ &= \Big(\partial_1(f \circ g), \partial_2(f \circ g), \partial_t(f \circ g)\Big) \end{aligned}$$

and

$$Df = \begin{pmatrix} \partial_1 f_1 & \partial_2 f_1 & \partial_t f_1 \\ \partial_1 f_2 & \partial_2 f_2 & \partial_t f_2 \end{pmatrix}$$

Sine $g = (X_t, t)$ we get:

$$Df(g) = \begin{pmatrix} \partial_1 f_1(X_t, t) & \partial_2 f_1(X_t, t) & \partial_t f_1(X_t, t) \\ \partial_1 f_2(X_t, t) & \partial_2 f_2(X_t, t) & \partial_t f_2(X_t, t) \end{pmatrix}$$

with $\partial_i f_j(X_t, t)$ being abbreviated as $\partial_i f_t^j(X_t)$ (meaning that each partial derivative $\partial_i f^j$ is calculated at the point with its last, third in our case, coordinate being the time variable t) for all indices $i \in \{1, 2, t\}$ and $j \in \{1, 2\}$, we rewrite:

$$Df(g) = \begin{pmatrix} \partial_1 f_t^1(X_t) & \partial_2 f_t^1(X_t) & \partial_t f_t^1(X_t) \\ \partial_1 f_t^2(X_t) & \partial_2 f_t^2(X_t) & \partial_t f_t^2(X_t) \end{pmatrix}$$

and

$$\begin{aligned} Dg &= \begin{pmatrix} \partial_1 g_1 & \partial_2 g_1 & \partial_t g_1 \\ \partial_1 g_2 & \partial_2 g_2 & \partial_t g_2 \\ \partial_1 g_3 & \partial_2 g_3 & \partial_t g_3 \end{pmatrix} \\ &= \begin{pmatrix} \partial_1 g_1 & \partial_2 g_1 & \partial_t X_1 \\ \partial_1 g_2 & \partial_2 g_2 & \partial_t X_2 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \partial_1 g_1 & \partial_2 g_1 & \partial_t X_t^1 \\ \partial_1 g_2 & \partial_2 g_2 & \partial_t X_t^2 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

since $g_3 = t$ and $g_i = X_i$ for $i = 1, 2$

So,

$$Df(g) \diamond Dg = \begin{pmatrix} \cdot & \cdot & \displaystyle\sum_{i=1}^{2} \partial_i f_t^1(X_t) \cdot \partial_t X_t^i + \partial_t f_t^1(X_t) \\ \cdot & \cdot & \displaystyle\sum_{i=1}^{2} \partial_i f_t^2(X_t) \cdot \partial_t X_t^i + \partial_t f_t^2(X_t) \end{pmatrix}$$

Thus, we deduce that:

$$\partial_t(f \circ g) = \left(\sum_{i=1}^{2} \partial_i f_t^1(X_t) \cdot \partial_t X_t^i + \partial_t f_t^1(X_t) \ , \ \sum_{i=1}^{2} \partial_i f_t^2(X_t) \cdot \partial_t X_t^i + \partial_t f_t^2(X_t) \right)$$

Due to the fact that:

$$\partial_t X_t = u_t(X_t)$$

it follows that:

$$\partial_t X_t^i = u_t^i(X_t)$$

With this we have for all i and j in $\{1,2\}$:

$$\begin{aligned}&\partial_i f_t^j(X_t)\cdot\partial_t X_t^i + \partial_t f_t^j(X_t)\\ =&\partial_i f_t^j(X_t)\cdot u_t^i(X_t) + \partial_t f_t^j(X_t)\\ =&u_t^i(X_t)\cdot\partial_i f_t^j(X_t) + \partial_t f_t^j(X_t)\end{aligned}$$

Hence, for $j = 1, 2$

$$\sum_{i=1}^{2}\partial_i f_t^j(X_t)\cdot\partial_t X_t^i + \partial_t f_t^j(X_t) = \langle u_t, \nabla\rangle f_t^j(X_t) + \partial_t f_t^j(X_t)$$

Since, we know that

$$\langle u_t, \nabla\rangle f_t = \big(\langle u_t, \nabla\rangle f_t^1, \langle u_t, \nabla\rangle f_t^2\big)$$
$$\text{and } \partial_t f_t = (\partial_t f_t^1, \partial_t f_t^2)$$

we get:

$$\partial_t(f\circ g) = \langle u_t(X_t), \nabla\rangle f_t(X_t) + \partial_t f_t(X_t)$$

i.e.

$$\partial_t\big(f_t(X_t)\big) = \langle u_t(X_t), \nabla\rangle f_t(X_t) + \partial_t f_t(X_t)$$

and the proof is completed. □

Setting $f(x,t) = \nabla P(x,t) \Leftrightarrow f_t = \nabla P_t$ we get that:

$$\begin{aligned}\partial_t\big(\nabla P_t(X_t)\big) &= \langle u_t(X_t), \nabla\rangle\nabla P_t(X_t) + \partial_t\nabla P_t(X_t)\\ &= \big(X_t - \nabla P_t(X_t)\big)^{\perp}\end{aligned}$$

Proposition 2.3.
Let $\xi \in C^1(\mathbb{R}^2\times[0,+\infty))$ and $h : \mathbb{R}^2\times[0,+\infty) \to \mathbb{R}^2\times[0,+\infty)$ with $h(x,t) = \big(h_1(x,t), h_2(x,t), t\big)$ which is also first order differentiable, then

$$\partial_t(\xi\circ h) = \big\langle\nabla\xi(h)\,,\,\partial_t(h_1,h_2)\big\rangle + \partial_t\xi(h)$$

Proof.
The chain rule implies:

$$D(\xi\circ h) = D\xi(h)\diamond Dh$$

For the derivatives we have:

$$D(\xi\circ h) = \big(\partial_1(\xi\circ h), \partial_2(\xi\circ h), \partial_t(\xi\circ h)\big)$$

and

$$D\xi(h) = \big(\partial_1\xi(h), \partial_2\xi(h), \partial_t\xi(h)\big)$$

and

$$\begin{aligned} Dh &= \begin{pmatrix} \partial_1 h_1 & \partial_2 h_1 & \partial_t h_1 \\ \partial_1 h_2 & \partial_2 h_2 & \partial_t h_2 \\ \partial_1 t & \partial_2 t & \partial_t t \end{pmatrix} \\ &= \begin{pmatrix} \partial_1 h_1 & \partial_2 h_1 & \partial_t h_1 \\ \partial_1 h_2 & \partial_2 h_2 & \partial_t h_2 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

Thus,

$$D\xi(h) \diamond Dh = \left(*, *, \sum_{i=1}^{2} \partial_i \xi(h) \cdot \partial_t h_i + \partial_t \xi(h) \right)$$

Hence,

$$\begin{aligned} \partial_t(\xi \circ h) &= \sum_{i=1}^{2} \partial_i \xi(h) \cdot \partial_t h_i + \partial_t \xi(h) \\ &= \big\langle \nabla\xi(h) \, , \partial_t(h_1, h_2) \big\rangle + \partial_t \xi(h) \end{aligned}$$

so we have proven the desired □

Setting $h(x,t) = \big(\nabla P_t(X_t), t\big) \Leftrightarrow (h_1, h_2) = \nabla P_t(X_t)$ we get:

$$\partial_t\Big(\xi \circ \big(\nabla P_t(X_t), t\big)\Big) = \big\langle \nabla\xi\big(\nabla P_t(X_t), t\big) \, , \partial_t\big(\nabla P_t(X_t)\big) \big\rangle + \partial_t \xi\big(\nabla P_t(X_t), t\big)$$

We abbreviate once more, we write $\xi(\cdot, t)$ as ξ_t. This leads to:

$$\begin{aligned} \partial_t\Big(\xi_t\big(\nabla P_t(X_t)\big)\Big) &= \big\langle \nabla\xi_t\big(\nabla P_t(X_t)\big) \, , \partial_t\big(\nabla P_t(X_t)\big) \big\rangle + \partial_t \xi_t\big(\nabla P_t(X_t)\big) \\ &= \big\langle \nabla\xi_t\big(\nabla P_t(X_t)\big) \, , \big(X_t - \nabla P_t(X_t)\big)^{\perp} \big\rangle + \partial_t \xi_t\big(\nabla P_t(X_t)\big) \end{aligned}$$

We integrate over time t to get:

$$\int_0^{+\infty} \partial_t\Big(\xi_t\big(\nabla P_t(X_t)\big)\Big) \, dt =$$

$$= \int_0^{+\infty} \big\langle \nabla\xi_t\big(\nabla P_t(X_t)\big) \, , \big(X_t - \nabla P_t(X_t)\big)^{\perp} \big\rangle + \partial_t \xi_t\big(\nabla P_t(X_t)\big) \, dt$$

Using the fundamental theorem of calculus

$$\int_0^{+\infty} \partial_t\Big(\xi_t\big(\nabla P_t(X_t)\big)\Big) \, dt = \lim_{s \to +\infty} \xi_s\big(\nabla P_s(X_s)\big) - \xi_0\big(\nabla P_0(X_0)\big)$$

Since $\xi \in C_c^\infty(\mathbb{R}^2 \times [0, +\infty))$ there is a $t_0 > 0$ such that $\xi_t \equiv 0$ for all $t > t_0$ and $X_0 = Id \Leftrightarrow X_0(x) = x$ we get:

$$\int_0^{+\infty} \partial_t \Big(\xi_t \big(\nabla P_t(X_t) \big) \Big) \, dt = -\xi_0(\nabla P_0)$$

We, now integrate over the space variable x:

$$\int_{\mathbb{R}^2} - \xi_0(\nabla P_0) \, dx =$$

$$= \int_{\mathbb{R}^2 \times [0,+\infty)} \big\langle \nabla \xi_t \big(\nabla P_t(X_t) \big) \, , \, \big(X_t - \nabla P_t(X_t) \big)^\perp \big\rangle + \partial_t \xi_t \big(\nabla P_t(X_t) \big) \, dt \, dx$$

For the second integral (right hand side) $\det(\nabla X_t) = 1$ (X_t being measure preserving) we have:

$$= \int_{\mathbb{R}^2 \times [0,+\infty)} \big\langle \nabla \xi_t \big(\nabla P_t(X_t) \big) \, , \, \big(X_t - \nabla P_t(X_t) \big)^\perp \big\rangle + \partial_t \xi_t \big(\nabla P_t(X_t) \big) \, dt \, dx$$

$$= \int_{[0,+\infty)} \int_{\mathbb{R}^2} \big\langle \nabla \xi_t \big(\nabla P_t(X_t) \big) \, , \, \big(X_t - \nabla P_t(X_t) \big)^\perp \big\rangle + \partial_t \xi_t \big(\nabla P_t(X_t) \big) \, dx \, dt$$

$$= \int_{[0,+\infty)} \int_{\mathbb{R}^2} \big\langle \nabla \xi_t(\nabla P_t) \, , \, (x - \nabla P_t)^\perp \big\rangle + \partial_t \xi_t(\nabla P_t) \, dx \, dt$$

$$= \int_{\mathbb{R}^2 \times [0,+\infty)} \big\langle \nabla \xi_t(\nabla P_t) \, , \, (x - \nabla P_t)^\perp \big\rangle + \partial_t \xi_t(\nabla P_t) \, dt \, dx$$

Thus, we get:

$$\int_{\mathbb{R}^2} - \xi_0(\nabla P_0) \, dx = \int_{\mathbb{R}^2 \times [0,+\infty)} \big\langle \nabla \xi_t(\nabla P_t) \, , \, (x - \nabla P_t)^\perp \big\rangle + \partial_t \xi_t(\nabla P_t) \, dt \, dx$$

We simplify this even further.

We follow the same method with $t = 0$ for the left hand side's integral.

We perform the change of variables $y = \nabla P_t(x) \Leftrightarrow x = \nabla P_t^*(y)$

Since ∇P_t and ∇P_t^* are inverse to each other we have $\nabla P_t(\nabla P_t^*(y)) = y$ in particular.

With these we obtain:

$$\int_{\mathbb{R}^2} -\xi_0 \big|\det\big(D^2 P_0^*\big)\big|\, dy = \int_{\mathbb{R}^2\times[0,+\infty)} \big(\langle \nabla\xi_t\,, (\nabla P_t^* - y)^\perp\rangle + \partial_t \xi_t\big)\big|\det\big(D^2 P_t^*\big)\big|\, dt\, dy$$

We will later show Proposition4.4 that the pushforward equation of the dual SG equation

$$\rho_t = \nabla P_{t\#} dx$$

implies the Monge-Ampère equation

$$\rho_t = \big|\det\big(D^2 P_t^*\big)\big|$$

Also, for the dual SG, the velocity is given by $U_t = (\nabla P_t^* - Id)^\perp$

So, utilizing them:

$$\int_{\mathbb{R}^2} -\xi_0 \rho_0\, dy = \int_{\mathbb{R}^2\times[0,+\infty)} \big(\langle \nabla\xi_t\,, U_t\rangle + \partial_t \xi_t\big)\rho_t\, dt\, dy$$

and we have finally arrived at the equation of a weak solution to the dual SG system

$$\int_{\mathbb{R}^2\times[0,+\infty)} \big(\partial_t \xi_t + \langle \nabla\xi_t\,, U_t\rangle\big)\rho_t\, dt\, dy \; + \int_{\mathbb{R}^2} \xi_0 \rho_0\, dy = 0$$

Before we move on to the existence of a weak solution satisfying this specific equation, we clearly state the definition of a weak solution to the dual SG system.

Definition 2.1 (weak solution of the dual SG system)**.**

We call ρ_t, P_t^* a weak solution to the dual SG system

$$\begin{cases} \partial_t \rho_t + \operatorname{div}(\rho_t U_t) = 0 & (x,t) \in \mathbb{R}^2 \times [0,+\infty) \\ U_t = (\nabla P_t^* - Id)^\perp & (x,t) \in \mathbb{R}^2 \times [0,+\infty) \\ \rho_t = \nabla P_{t\#} dx & t \in [0,+\infty) \\ \text{with initial data } \rho_0 & \end{cases}$$

iff

$$\forall \xi \in C_c^\infty\big(\mathbb{R}^2 \times [0,+\infty)\big)$$

$$\int_{\mathbb{R}^2\times[0,+\infty)} \big(\partial_t \xi_t + \langle \nabla\xi_t\,, U_t\rangle\big)\rho_t\, dt\, dy \; + \int_{\mathbb{R}^2} \xi_0 \rho_0\, dy = 0$$

and

$$U_t = (\nabla P_t^* - Id)^\perp$$

We can now pursue our target, that is to prove the existence of such (weak in the definition we just gave) a solution for the dual SG.

There is a particular result that will be useful, and can be found in [23] [20] [7] [29] [13]

Theorem 2.1 (Probability measures on the torus)**.**

Let μ,ν be two probability measures on the torus $\mathbb{T}^2$

If $\mu = f\ dx$ with $f \overset{\text{a.e.}}{>} 0$, then there exists an, up to additive constant, unique convex function $P : \mathbb{R}^2 \to \mathbb{R}$ such that:

- $\nu = \nabla P_{\#}\mu$
- $P(x) - \frac{||x||^2}{2}$ is $\mathbb{Z}^2-$periodic
- ∇P is a.e. $\mathbb{Z}^2$-periodic, that is $\nabla P(x)$ is $\mathbb{Z}^2$-periodic for a.e. $x \in \mathbb{R}^2$
- $\nabla P : \mathbb{T}^2 \to \mathbb{T}^2$ is the $\mu-$a.e. unique optimal transport map sending μ onto ν

and

$$||\nabla P(x) - x|| \leq \text{diam}(\mathbb{T}^2) = \frac{\sqrt{2}}{2} \qquad \text{for a.e. } x \in \mathbb{R}^2$$

Additionally, if

$$\begin{cases} \nu = g\ dx \text{ and} \\ \text{there exist constants } \lambda, \Lambda \text{ such that } 0 < \lambda \leq f, g \leq \Lambda \end{cases}$$

then P is a strictly convex Alexandrov solution of the Monge-Ampère equation

$$\det\left(D^2 P\right) = \frac{f}{g(\nabla P)}$$

This theorem will be used on the construction of approximate solutions (for the section, existence of weak solutions below), in order to obtain a convex function from the inital data ρ_0.

Actually it will be used twice, since we will follow the same (logical) steps to build a sequence of approximate solutions at the existence of smooth solutions as well.

2.2 Existence

We proceed now to prove that there is indeed, at least one, weak solution of the dual SG system existing globally in time.

Theorem 2.2 (Existence of global weak solution for the dual SG)**.**

Assume that ρ_0 is absolutely continuous with respect to Lebesgue measure and a probability measure on the torus.

$$\text{If } \exists m, M \in \mathbb{R} \text{ such that } 0 < m \leq \rho_0 \leq M$$

then $\exists\ \rho_t, P_t$ weak solution to the dual SG system on $\mathbb{R}^2 \times [0, +\infty)$, which satisfies the following:

$$0 < m \leq \rho_t \leq M \text{ for a.e. } t \geq 0$$

$$\text{and } \rho_t \in L^\infty\big([0, +\infty), L^\infty(\mathbb{R}^2)\big)$$

Remark.
The condition that ρ_0 is absolutely continuous with respect to the Lebesgue measure and a probability measure on the torus, is not a "tough" one. If we set $\rho_0 = (x+\nabla\bar{p})_{\#}dx$ (recall this is the initial condition connecting the initial data between the SG system and the dual SG system) we can have this requirement fulfilled.

The proof will be split into three parts.

The first part (Part I) consists of the approximate solution construction. In essence, we mollify the initial data U_0 (defined with the help of ρ_0) and we build a sequence of smooth functions that satisfy seperately (not as coupled equations) the equations the dual SG system consists of.

We wiil achieve that by solving the measure continuity equation with the time-frozen, mollified U_0. In this part we will need the so-called flow function, which is the (unique) solution of a non-autonomous first order ode.

We continue taking their limits (under weak convergence). The last two parts belong to the "bigger category" of the limit passage in the distributional sense (thus proving that they are indeed a weak solution to the dual SG).

First (Part II), we show that the product of the density with the velocity (sequence) converges to the product of their limits (which for the weak convergence is not true in general).

And finally (Part III), we show that the limit of the velocity satisfies the condition which “connects” it with the convex conjugate of pressure.

Proof.

We begin with the first part

2.2.1 Part I: Constructing the approximate solution

Applying Theorem2.1 with ρ_0 and the Lebesgue measure, we obtain a unique (up to additive constant) convex P_0 such that:

$$\rho_0 = \nabla P_{0\#} dx$$

$$P_0 - \frac{\|x\|_2^2}{2} \text{ is periodic}$$

$$\|\nabla P_0 - Id\| \leq \frac{\sqrt{2}}{2}$$

We define:

$$U_0 := (\nabla P_0^* - Id)^{\perp}$$

We will utilize U_0 to define the flow and solve $\partial_t \rho_t + \operatorname{div}(\rho_t U_0) = 0$. But we will need to mollify it first, in order to have the needed reguarity.

The reason we must have the velocity mollified, is because the flow actually help us solve the respective transport equation $\partial_t \rho_t + \langle \nabla \rho_t, U_0 \rangle = 0$. This equation is equivalent to our continuity equation, when our functions are smooth enough and the velocity is divergence free.

Let $\varepsilon > 0$

First iteration

We restrict to $t \in [0, \varepsilon]$

We define the time (freezed) and epsilon independent pressure and velocity

$$P_t^{\varepsilon} := P_0$$

$$U_t^{\varepsilon} := U_0$$

We then mollify the velocity defining:

$$U_t^{\varepsilon,\delta} := U_0^{\delta} = \eta_\delta * U_0$$

where $*$ denotes the convolution of function, that is:

$$U_0^\delta(x) = (\eta_\delta * U_0)(x) = \int_{\mathbb{R}^2} \eta_\delta(x-z)U_0(z)\,dz \ \ \forall x \in \mathbb{R}^2$$

Since $U_0 \in \mathbb{R}^2$ we identify the integral above (and any integral of a vector-valued function) as its component integrals:

$$\int_{\mathbb{R}^2} \eta_\delta(x-z)U_0(z)\,dz = \left(\int_{\mathbb{R}^2} \eta_\delta(x-z)U_0^1(z)\,dz, \int_{\mathbb{R}^2} \eta_\delta(x-z)U_0^2(z)\,dz \right)$$

where $U_0 = (U_0^1, U_0^2)$

Next, we proceed to show that $U_t^{\varepsilon,\delta}$ is Lipschitz and divergence free.

Indeed,

Evans "tells us" that $U_t^{\varepsilon,\delta} \in C^\infty(\mathbb{R}^2 : \mathbb{R}^2)$ and

$$\nabla U_t^{\varepsilon,\delta} = \nabla \eta_\delta * U_0$$

Also, since $\nabla P_t \circ \nabla P_t^* = Id$ and $\|\nabla P_t - Id\| \leq \frac{\sqrt{2}}{2}$, by setting x as $\nabla P_0^*(x)$ and $t = 0$ we get:

$$\|U_0\| \leq \frac{\sqrt{2}}{2} \Rightarrow \|U_0\|_{L^\infty(\mathbb{R}^2)} \leq \frac{\sqrt{2}}{2}$$

Thus,

$$\left\| U_t^{\varepsilon,\delta} \right\|_{L^\infty(\mathbb{R}^2)} \leq \frac{\sqrt{2}}{2}$$

and

$$\left\| \nabla U_t^{\varepsilon,\delta} \right\|_{L^\infty(\mathbb{R}^2)} \leq C$$

Hence, $U_t^{\varepsilon,\delta}(x)$ is Lipschitz in $\mathbb{R}^2$ for all times $t \in [0, \varepsilon]$

Epsilon (ε) and delta (δ) do not play any particular role in the next step, so this part will be presented in a more general context.

Solving $\partial_t \rho_t + \mathrm{div}(\rho_t U_t) = 0$ for $t \geq 0$ with konwn U_t

Since the velocity $U(x,t)$ is continuous and U_t is Lipschitz we are able to uniquely solve the initial value problem for every $y \in \mathbb{R}^2$:

$$\begin{cases} \partial_t X(t) = U\big(X(t), t\big) \\ X(0) = y \end{cases}$$

that is $\forall$ initial data $y \in \mathbb{R}^2$ $\exists!$ (time) function $Y_y : \mathbb{R} \to \mathbb{R}^2$ which solves this differential equation i.e.

$$\begin{cases} \partial_t Y_y(t) = U\big(Y_y(t), t\big) \\ Y_y(0) = y \end{cases}$$

where $t \mapsto Y_y$ has one order higher regurality than $(x,t) \mapsto U(x,t)$, since it is given as the composition $t \mapsto (Y_y(t), t) \mapsto U(Y_y(t), t)$

The uniqueness of the solution for this problem allows us to define a (space) function for all times t

$$Y_t : \mathbb{R}^2 \to \mathbb{R}^2 \ \ , \ \forall t \in [0, \infty)$$

This function is the map sending y to the unique $Y_y(t)$ i.e. $y \mapsto Y_y(t)$ which we also identify as $Y_t(y)$.

This map is indeed a function, since for all $t \geq 0$

$$\begin{aligned} y_1 &= y_2 \\ &\xrightarrow[\text{solution}]{\text{unique}} Y_{y_1}(t) = Y_{y_2}(t) \\ &\Rightarrow Y_t(y_1) = Y_t(y_2) \end{aligned}$$

So, actually, we have obtained a time differentiable and space dependent function $Y(y,t)$ which we also denote $Y_t(y)$ or $Y_y(t)$

Now we can rewrite the flow initial value problem in the usual way we have chosen to denote our time and space dependent functions (that is with the time t as a subscript and omitting the space, “main”, variable x or y).

Hence

$$\begin{cases} \partial_t Y_t = U_t(Y_t) \\ Y_0 = Id \end{cases}$$

Then, taking advantage of the flow Y_t, we can obtain a weak solution for the measure continuity equation:

$$\partial_t \rho_t + \operatorname{div}(\rho_t U_t) = 0$$

We define $\forall t \in [0, +\infty)$

$$\rho_t := Y_{t\#}\rho_0$$

Let us check that this measure is indeed a solution i.e. it satisfies the measure continuity equation in the weak sense Definition1.1 we have already discussed.

Obviously, ρ_t is well-defined, since for $t = 0$

$$Y_{0\#}\rho_0 = Id_{\#}\rho_0 = \rho_0$$

We proceed to show that $\forall \varphi \in C_c^\infty(\mathbb{R}^2)$

$$\partial_t \int \varphi \, d\rho_t = \int \langle \nabla\varphi, U_t \rangle \, d\rho_t$$

Let $\varphi \in C_c^\infty(\mathbb{R}^2)$, we compute:

First of all, since φ is continuous, it is also Lebesque measurable, hence it is Borel measurable as well.

Next, we claim that for every time t the flow Y_t is a bi-Lipschitz homeomorphism.

More precisely it satisfies:

$$e^{-Kt}\|y_1 - y_2\| \leq \|Y_t(y_1) - Y_t(y_2)\| \leq e^{Kt}\|y_1 - y_2\|$$

Indeed,

Let $y_1, y_2 \in \mathbb{R}^2$, as we have seen, the initial value problem has a unique solution for all y

Thus, for $i = 1, 2$

$$\begin{cases} \partial_t Y_t(y_i) = U_t\big(Y_t(y_i)\big) \\ Y_0(y_i) = y_i \end{cases}$$

Subtracking the two equations and taking their norms, we have:

$$\begin{cases} \|\partial_t Y_t(y_1) - \partial_t Y_t(y_2)\| = \big\|U_t\big(Y_t(y_1)\big) - U_t\big(Y_t(y_2)\big)\big\| \\ \\ \|Y_0(y_1) - Y_0(y_2)\| = \|y_1 - y_2\| \end{cases}$$

Using the inequality $\Big|\partial_t\|f(t)\|\Big| \leq \|\partial_t f(t)\|$ along with the fact that the derivative is a linear operator, we get:

$$\Big|\partial_t\|Y_t(y_1) - Y_t(y_2)\|\Big| \leq \|\partial_t Y_t(y_1) - \partial_t Y_t(y_2)\|$$

The velocity U_t is K-Lipschitz i.e.

$$\|U_t(x_1) - U_t(x_2)\| \leq K\|x_1 - x_2\| \quad \forall x_1, x_2 \in \mathbb{R}^2$$

Choosing $x_1 = Y_t(y_1)$ and $x_2 = Y_t(y_2)$ combined with the result above, we have:

$$\begin{cases} \Big|\partial_t\|Y_t(y_1) - Y_t(y_2)\|\Big| \leq K\|Y_t(y_1) - Y_t(y_2)\| \\ \|Y_0(y_1) - Y_0(y_2)\| = \|y_1 - y_2\| \end{cases}$$

Expanding the absolute value and utilizing the two Gronwall lemmas for the respective inequalities (K and $-K$) with $\varphi(t) := \|Y_t(y_1) - Y_t(y_2)\|$ in both cases, we have proved the desired.

Thus, Y_t is $\big(\mathcal{B}(\mathbb{R}^2), \mathcal{B}(\mathbb{R}^2)\big)$-measurable, since it is a continuous function. The "change of variables" for the push forward measure implies:

$$\varphi \circ Y_t \text{ is in } L^1(\rho_0)$$

$$\int_{\mathbb{R}^2} \varphi \, d\rho_t = \int_{(Y_t)^{-1}(\mathbb{R}^2)} \varphi \circ Y_t \, d\rho_0$$

The inverse image $(Y_t)^{-1}(\mathbb{R}^2)$ is equal to the function's domain $D_{Y_t} = \mathbb{R}^2$.

Next we show the three needed conditions to apply the Liebniz integral rule PropositionA.31.

i) We have already shown (at the push forward change of variables earlier) that for every t the function $y \mapsto \varphi \circ Y_t$ is in $L^1(\rho_0)$.

ii) For every y the function $t \mapsto \varphi \circ Y_y$ is differentiable (Picard-Lindelöf's theorem for the ordinary differential equation guarantees a classic, in terms of differentiability, solution to the initial value problem).

iii) Moreover, the chain rule PropositionA.36 leads to:

$$\begin{aligned}\partial_t \left(\varphi \circ Y_t\right) &= \nabla\varphi \circ Y_t \diamond \partial_t Y_t \\ &= \left\langle \nabla\varphi \circ Y_t \, , \, \partial_t Y_t \right\rangle\end{aligned}$$

By Cauchy-Schwarz's inequality we get:

$$|\partial_t \left(\varphi \circ Y_t\right)| \leq \|\nabla\varphi \circ Y_t\| \cdot \|\partial_t Y_t\|$$

We also have that:

$$\partial_t Y_t = U_t\big(Y_t\big) \text{ , where } \|U_t\| \leq \frac{\sqrt{2}}{2}$$

and for all t

$$\begin{cases} \|\nabla\varphi \circ Y_t\| \leq M & \text{when } Y_t(y) \in \text{supp}\nabla\varphi \\ \|\nabla\varphi \circ Y_t\| = 0 & \text{when } Y_t(y) \notin \text{supp}\nabla\varphi \end{cases}$$

because $\nabla\varphi$ is continuous with compact support.

Hence, by setting

$$\begin{cases} h(y) = M\frac{\sqrt{2}}{2} & \text{when } y \in (Y_t)^{-1}(\text{supp}\nabla\varphi) \\ h(y) = 0 & \text{when } y \notin (Y_t)^{-1}(\text{supp}\nabla\varphi) \end{cases}$$

which belongs in $L^1(\rho_0)$ we have completed the proof of the three criteria.

Hence,

$$\begin{aligned}\partial_t \int_{\mathbb{R}^2} \varphi \, d\rho_t &= \int_{\mathbb{R}^2} \partial_t \left(\varphi \circ Y_t\right) \, d\rho_0 \\ &= \int_{\mathbb{R}^2} \left\langle \nabla\varphi \circ Y_t \, , \, \partial_t Y_t \right\rangle d\rho_0\end{aligned}$$

$$= \int_{\mathbb{R}^2} \big\langle \nabla\varphi \circ Y_t \,,\, U_t \circ Y_t \big\rangle \, d\rho_0$$

$$= \int_{\mathbb{R}^2} \big\langle \nabla\varphi \,,\, U_t \big\rangle \circ Y_t \, d\rho_0$$

With $\langle \varphi, U_t \rangle$, Y_t being continuous and $(Y_t)^{-1}(\mathbb{R}^2) = \mathbb{R}^2$, we use again the push forward change of variables to obtain:

$$\partial_t \int_{\mathbb{R}^2} \varphi \, d\rho_t = \int_{\mathbb{R}^2} \big\langle \nabla\varphi \,,\, U_t \big\rangle \, d\rho_t$$

Proposition 2.4 ($\sigma_t = Y_{t\#} d\rho_0$ is the unique solution of the continuity equation with initial data ρ_0)**.**

Let σ_t be a solution of

$$\partial_t \rho_t + \operatorname{div}(\rho_t U_t) = 0$$

with $\sigma_0 = \rho_0$, then

$$\sigma_t = Y_{t\#} d\rho_0$$

Proof.
For a proof look at Figalli's [23] section2.1 □

Proposition 2.5 (equation for the density of the measure solution $Y_{t\#} d\rho_0$)**.**

If ρ_0 has a density, then so does ρ_t

and

$$\rho_t(y) = \rho_0\big(Y_t^{-1}(y)\big)$$

Proof.

Let $\varphi \in C_c^\infty(\mathbb{R}^2)$

The pushforward change of variables ($\rho_t = Y_{t\#} d\rho_0$) and the fact that ρ_0 has a desnity implies:

$$\int_{\mathbb{R}^2} \varphi \, d\rho_t = \int_{\mathbb{R}^2} \varphi \circ Y_t \, d\rho_0$$

$$= \int_{\mathbb{R}^2} \rho_0 \cdot (\varphi \circ Y_t)\, dy$$

Since Y_t is measure preserving i.e. $\det(\nabla Y_t) = 1$ for all times t, as shown in Proposition2.1, the classical change of variables $x = Y_t(y) \Leftrightarrow Y_t^{-1}(x) = y$ leads to:

$$\int_{\mathbb{R}^2} \rho_0 \cdot (\varphi \circ Y_t)\, dy = \int_{\mathbb{R}^2} \varphi \cdot (\rho_0 \circ Y_t^{-1})\, dx$$

Thus, for all $\varphi \in C_c^\infty(\mathbb{R}^2)$

$$\int_{\mathbb{R}^2} \varphi\, d\rho_t = \int_{\mathbb{R}^2} \varphi \cdot (\rho_0 \circ Y_t^{-1})\, dx$$

The arbitrariness of φ implies the desired. □

Adding everywhere epsilon,delta (ε, δ) as a superscript, we can return to our case and view.

We define for consistency

$$P_t^{\varepsilon,\delta} := P_t^\varepsilon \underset{def}{\overset{by}{=}} P_0$$

since the next step will have δ playing its role in the definition of pressure.

Remark.
P_0 needed no mollification, due to the fact that a convex function is two times a.e. differentiable.

So, we have built a triplet $\rho_t^{\varepsilon,\delta}, P_t^{\varepsilon,\delta}, U_t^{\varepsilon,\delta}$ which is an approximate solution to the dual SG system i.e. satisfies each equation individually.

$$\begin{cases} \partial_t \rho_t^{\varepsilon,\delta} + \operatorname{div}(\rho_t^{\varepsilon,\delta} U_t^{\varepsilon,\delta}) = 0 & (x,t) \in \mathbb{R}^2 \times [0,\varepsilon] \\ U_t^{\varepsilon,\delta} = \eta_\delta * \left(\nabla P_t^{\varepsilon,\delta,*} - Id\right)^\perp & (x,t) \in \mathbb{R}^2 \times [0,\varepsilon] \\ \rho_0^{\varepsilon,\delta} = \nabla P_t^{\varepsilon,\delta}{}_\# dx & t \in [0,\varepsilon] \\ \text{with initial data } \rho_0 & \end{cases}$$

and it also satisfies:

$$\begin{cases} \rho_t^{\varepsilon,\delta} = Y_t^{\varepsilon,\delta}{}_\# \rho_0 & t \in [0,\varepsilon] \\ m \le \rho_t^{\varepsilon,\delta} \le M & (x,t) \in \mathbb{R}^2 \times [0,\varepsilon] \\ \left\| U_t^{\varepsilon,\delta} \right\| \le \dfrac{\sqrt{2}}{2} & \end{cases}$$

Remark.
Notice that $\rho_t^{\varepsilon,\delta}$ is well-deifned because $Y_0^{\varepsilon,\delta} = Id$

Indeed,

$$\rho_0^{\varepsilon,\delta} = {Y_0^{\varepsilon,\delta}}_{\#}\rho_0 = Id_{\#}\rho_0 = \rho_0$$

We have to "restrict" our triplet, in terms of time, even further, leaving ε out.

The reason we have to do this, is because we want to avoid conflict with the next interval $[\varepsilon, 2\varepsilon)$.

Otherwise our functions would have to coincide in the value ε of time ,which is not guaranteed.

In fact it would have meant that $P_0 \equiv P_\varepsilon$ and $U_0 \equiv U_\varepsilon$ where $P_\varepsilon, U_\varepsilon$ are defined utilizing $\rho_\varepsilon^{\varepsilon,\delta}$ and implementing Theorem2.1 like we did with ρ_0.

Remark.
This is exactly the second step being followed in the procedure to construct the approximate solution. which comes next (see below) in the proof.

We rewrite and "sum up" what we have built so far:

$$\begin{cases} \partial_t \rho_t^{\varepsilon,\delta} + \operatorname{div}(\rho_t^{\varepsilon,\delta} U_t^{\varepsilon,\delta}) = 0 & (x,t) \in \mathbb{R}^2 \times [0,\varepsilon) \\ U_t^{\varepsilon,\delta} = \eta_\delta * \left(\nabla P_t^{\varepsilon,\delta,*} - Id\right)^\perp & (x,t) \in \mathbb{R}^2 \times [0,\varepsilon) \\ \rho_0^{\varepsilon,\delta} = {\nabla P_t^{\varepsilon,\delta}}_{\#} dx & t \in [0,\varepsilon) \\ \rho_t^{\varepsilon,\delta} = {Y_t^{\varepsilon,\delta}}_{\#} \rho_0 & t \in [0,\varepsilon] \\ 0 < m \leq \rho_t^{\varepsilon,\delta} \leq M & (x,t) \in \mathbb{R}^2 \times [0,\varepsilon] \\ \left\| U_t^{\varepsilon,\delta} \right\| \leq \dfrac{\sqrt{2}}{2} \end{cases}$$

Even though pressure and velocity do not necessarily coincide when $t = \varepsilon$, ρ_t will, due to the fact the flow satisfies $Y_\varepsilon = Id$

So, we can repeat those steps above on the next time interval.

Repeating the process, second iteration
$t \in [\varepsilon, 2\varepsilon)$

We apply Theorem2.1 with $\rho_\varepsilon^{\varepsilon,\delta}$ (and the Lebesgue measure) to obtain an up to additive constant unique convex $P^{\varepsilon,\delta}$ such that:

$$\rho_\varepsilon^{\varepsilon,\delta} = {\nabla P^{\varepsilon,\delta}}_{\#} dx$$

$$P^{\varepsilon,\delta} - \frac{\|x\|_2^2}{2} \text{ is periodic}$$

$$\left\|\nabla P^{\varepsilon,\delta} - Id\right\| \leq \frac{\sqrt{2}}{2}$$

We define:

$$U^{\varepsilon,\delta} := \left(\nabla P^{\varepsilon,\delta,*} - Id\right)^{\perp}$$

$$P_t^{\varepsilon,\delta} := P^{\varepsilon,\delta}$$

and we mollify

$$U_t^{\varepsilon,\delta} := \eta_\delta * U^{\varepsilon,\delta}$$

We also define the flow, the unique solution of:

$$\begin{cases} \partial_t Y_t^{\varepsilon,\delta} = U_t^{\varepsilon,\delta}(Y_t^{\varepsilon,\delta}) \\ Y_\varepsilon^{\varepsilon,\delta} = Id \end{cases}$$

We define, once more, the measure

$$\rho_t^{\varepsilon,\delta} := Y_t^{\varepsilon,\delta}{}_{\#} d\rho_\varepsilon^{\varepsilon,\delta}$$

which is well defined since

$$\begin{aligned} Y_\varepsilon^{\varepsilon,\delta} &= Id \Rightarrow \\ \rho_\varepsilon^{\varepsilon,\delta} &= Id_{\#} d\rho_\varepsilon^{\varepsilon,\delta} = \rho_\varepsilon^{\varepsilon,\delta} \end{aligned}$$

and a waek solution to the measure continuity equation

$$\partial_t \rho_t^{\varepsilon,\delta} + \operatorname{div}(\rho_t^{\varepsilon,\delta} U_t^{\varepsilon,\delta}) = 0$$

Thus, for the second iteration we have:

$$\begin{cases} \partial_t \rho_t^{\varepsilon,\delta} + \operatorname{div}(\rho_t^{\varepsilon,\delta} U_t^{\varepsilon,\delta}) = 0 & (x,t) \in \mathbb{R}^2 \times [\varepsilon, 2\varepsilon) \\ U_t^{\varepsilon,\delta} = \eta_\delta * \left(\nabla P_t^{\varepsilon,\delta,*} - Id\right)^{\perp} & (x,t) \in \mathbb{R}^2 \times [\varepsilon, 2\varepsilon) \\ \rho_\varepsilon^{\varepsilon,\delta} = \nabla P_t^{\varepsilon,\delta}{}_{\#} dx & t \in [\varepsilon, 2\varepsilon) \\ \rho_t^{\varepsilon,\delta} = Y_t^{\varepsilon,\delta}{}_{\#} \rho_0 & t \in [\varepsilon, 2\varepsilon] \\ 0 < m \leq \rho_t^{\varepsilon,\delta} \leq M & (x,t) \in \mathbb{R}^2 \times [\varepsilon, 2\varepsilon] \\ \left\|U_t^{\varepsilon,\delta}\right\| \leq \frac{\sqrt{2}}{2} & (x,t) \in \mathbb{R}^2 \times [\varepsilon, 2\varepsilon) \end{cases}$$

Repeating the process we obtain the following approximate solution:

$$\begin{cases} \partial_t \rho_t^{\varepsilon,\delta} + \operatorname{div}(\rho_t^{\varepsilon,\delta} U_t^{\varepsilon,\delta}) = 0 & (x,t) \in \mathbb{R}^2 \times [0, +\infty) \\ U_t^{\varepsilon,\delta} = \eta_\delta * \left(\nabla P_t^{\varepsilon,\delta,*} - Id\right)^{\perp} & (x,t) \in \mathbb{R}^2 \times [0, +\infty) \\ \rho_{k\varepsilon}^{\varepsilon,\delta} = \nabla P_t^{\varepsilon,\delta}{}_{\#} dx & t \in [k\varepsilon, (k+1)\varepsilon) \\ \rho_t^{\varepsilon,\delta} = Y_t^{\varepsilon,\delta}{}_{\#} \rho_0 & t \in [0, +\infty) \\ 0 < m \leq \rho_t^{\varepsilon,\delta} \leq M & (x,t) \in \mathbb{R}^2 \times [0, +\infty) \\ \left\|U_t^{\varepsilon,\delta}\right\| \leq \frac{\sqrt{2}}{2} & (x,t) \in \mathbb{R}^2 \times [0, +\infty) \end{cases}$$

Having constructed the approximate solution, we move on to the next part.

2.2.2 Part II: Taking the limit

This section is based on [23].

For this part, we set $\varepsilon = \delta = \frac{1}{n}$ to obtain a triplet of sequences ρ_t^n, P_t^n, U_t^n that satisfy:

$$\begin{cases} \partial_t \rho_t^n + \operatorname{div}(\rho_t^n U_t^n) = 0 & (x,t) \in \mathbb{R}^2 \times [0,+\infty) \\ U_t^n = \eta_{\frac{1}{n}} * \left(\nabla P_t^{n,*} - Id\right)^{\perp} & (x,t) \in \mathbb{R}^2 \times [0,+\infty) \\ \rho_{k\frac{1}{n}}^n = \nabla P_t^n{}_{\#} dx & t \in [k\frac{1}{n}, (k+1)\frac{1}{n}) \\ \rho_t^n = Y_t^n{}_{\#} d\rho_0 & t \in [0,+\infty) \\ 0 < m \leq \rho_t^n \leq M & (x,t) \in \mathbb{R}^2 \times [0,+\infty) \\ \|U_t^n\| \leq \dfrac{\sqrt{2}}{2} \end{cases}$$

Hence, the sequences ρ_t^n, U_t^n are uniformly bounded in time and space.

Brezis end of page 116 C.(ii) implies that there are functions ρ_t, U_t such that

$$\rho_t^n \rightharpoonup^* \rho_t \text{ in } L^\infty_{loc}\left(\mathbb{R}^2 \times [0,+\infty)\right)$$

$$U_t^n \rightharpoonup^* U_t \text{ in } L^\infty_{loc}\left(\mathbb{R}^2 \times [0,+\infty) : \mathbb{R}^2\right)$$

$\rho_t^n U_t^n \rightharpoonup^* \rho_t U_t$ **in** $L^\infty_{loc}\left(\mathbb{R}^2 \times [0,+\infty) : \mathbb{R}^2\right)$

Proposition 2.6. Let $f : \mathbb{R}^n \to \mathbb{R}$ then the following are true

$\forall q \in [1,+\infty), p \in [1,+\infty]$ and $V \subset\subset \mathbb{R}^2$

1. $\|f\|_{L^1(V)} \leq l^{\frac{q-1}{q}}(V)\|f\|_{L^q(V)}$
2. $L^q(V) \subseteq L^1(V)$
3. $L^q_{loc}(\mathbb{R}^2) \subseteq L^1_{loc}(\mathbb{R}^2)$
4. $\|f\|_{W^{1,q}(V)} \leq c\|f\|_{W^{1,1}(V)}$
5. $W^{1,q}(V) \subseteq W^{1,1}(V)$
6. $W^{1,q}_{loc}(\mathbb{R}^2) \subseteq W^{1,1}_{loc}(\mathbb{R}^2)$
7. $\left(W^{1,1}_{loc}(\mathbb{R}^2)\right)^* \subseteq \left(W^{1,q}_{loc}(\mathbb{R}^2)\right)^*$
8. $\|f\|_Y \leq c\|f\|_X \Rightarrow L^\infty(A,X) \subseteq L^\infty(A,Y)$
9. $L^\infty(B) \subseteq L^p_{loc}(B)\ \ \forall B \subseteq \mathbb{R}^2$

where

$$\|f\|_{A,X} := \sup_{t\in A}\|f(t)\|_X < +\infty$$

Combining the above, we can show that:

$$L^\infty\left([0,+\infty),\left(W^{1,1}_{loc}(\mathbb{R}^2)\right)^*\right)\subseteq L^p_{loc}\left([0,+\infty),\left(W^{1,q}_{loc}(\mathbb{R}^2)\right)^*\right)$$

and

$$L^\infty\left([0,+\infty),L^\infty(\mathbb{R}^2)\right)\subseteq L^p_{loc}\left([0,+\infty),L^p_{loc}(\mathbb{R}^2)\right)$$

We will prove that

$$\rho_t^n\to\rho_t \text{ in } L^p_{loc}\left([0,+\infty),\left(W^{s,q}_{loc}(\mathbb{R}^2)\right)^*\right)\quad\forall q\geq 1, p>1 \text{ and } s>0$$

Indeed,

Let $\psi\in C_c^\infty(\mathbb{R}^2)$, then integration by parts gives:

$$\int\limits_{B(0,r)}\psi\mathrm{div}(\rho_t^nU_t^n)\,dy=-\int\limits_{B(0,r)}\left\langle\nabla\psi,\rho_t^nU_t^n\right\rangle dy+\int\limits_{\partial B(0,r)}\left\langle\psi\rho_t^nU_t^n,\bar{n}\right\rangle dS$$

since ψ has compact support, sending r to infinity yields:

$$\int\limits_{\mathbb{R}^2}\psi\mathrm{div}(\rho_t^nU_t^n)\,dy=-\int\limits_{\mathbb{R}^2}\left\langle\nabla\psi,\rho_t^nU_t^n\right\rangle dy$$

So,

$$\int\limits_{\mathbb{R}^2}\psi\big(-\mathrm{div}(\rho_t^nU_t^n)\big)\,dy=\int\limits_{\mathbb{R}^2}\left\langle\nabla\psi,\rho_t^nU_t^n\right\rangle dy$$

The Cauchy-Schwarz inequality combined with the uniform bounds of ρ_t^n,U_t^n $0<\rho_t^n\leq M$ and $\|U_t^n\|\leq\frac{\sqrt{2}}{2}$ imply that:

$$\begin{aligned}\left\langle\nabla\psi,\rho_t^nU_t^n\right\rangle&\leq\|\nabla\psi\|\cdot\|\rho_t^nU_t^n\|\\&=\|\nabla\psi\|\cdot|\rho_t^n|\cdot\|U_t^n\|\\&=\|\nabla\psi\|\cdot\rho_t^n\cdot\|U_t^n\|\\&\leq\|\nabla\psi\|\cdot M\frac{\sqrt{2}}{2}\end{aligned}$$

Hence,

$$\int\limits_{\mathbb{R}^2}\psi\big(-\mathrm{div}(\rho_t^nU_t^n)\big)\,dy\leq C\|\nabla\psi\|_{W^{1,1}(\mathbb{R}^2)}$$

Thus, $-\text{div}(\rho_t^n U_t^n) \in \left(W^{1,1}_{loc}(\mathbb{R}^2)\right)^*$ uniformly in time i.e.

$$-\text{div}(\rho_t^n U_t^n) \in L^\infty\left([0,+\infty), \left(W^{1,1}_{loc}(\mathbb{R}^2)\right)^*\right)$$

where using $\partial_t \rho_t^n = -\text{div}(\rho_t^n U_t^n)$ and the inclusion we have shown. we get:

$$\partial_t \rho_t^n \in L^p_{loc}\left([0,+\infty), \left(W^{1,q}_{loc}(\mathbb{R}^2)\right)^*\right)$$

Also,

$$\rho_t^n \in L^p_{loc}\left([0,+\infty), L^p_{loc}(\mathbb{R}^2)\right)$$

due to the fact that:

$$\rho_t^n \in L^\infty\left([0,+\infty), L^\infty(\mathbb{R}^2)\right)$$

since $0 < \rho_t^n \le M \Rightarrow \sup\limits_{t\in[0,+\infty)} |\rho_t^n| \le M$ and $\forall t \;\; y \mapsto \rho_t^n(y) \in L^\infty(\mathbb{R}^2)$

Having proved that:

$$\begin{cases} \partial_t \rho_t^n \in L^p_{loc}\left([0,+\infty), \left(W^{1,q}_{loc}(\mathbb{R}^2)\right)^*\right) \\ \rho_t^n \in L^p_{loc}\big([0,+\infty), L^p_{loc}(\mathbb{R}^2)\big) \end{cases}$$

the Aubin-Lions Lemma [23] implies that ρ_t^n is precomapct in the space $L^p_{loc}\left([0,+\infty), \left(W^{s,q}_{loc}(\mathbb{R}^2)\right)^*\right)$ $\forall q \ge 1, p > 1$ and $s > 0$

We will prove that

$$U_t^n \rightharpoonup^* U_t \text{ in } L^\infty\left([0,+\infty), W^{r,k}_{loc}(\mathbb{R}^2)\right) \;\; \forall r \in (0,1) \text{ and } 1 \le k < \tfrac{2}{1+r}$$

Indeed,

Proposition 2.7. Let $f : \mathbb{R}^d \to \mathbb{R}$ smooth, convex and Lipschitz function on every neighbourhood then $\forall r > 0 \; \exists C_r > 0$ such that

$$\int\limits_{B(x_0,r)} \left\|D^2 f\right\| \le C_r$$

Proof.

Let $r > 0$ and $x_0 \in \mathbb{R}^d$

We know that f is smooth, thus there exists its hessian $Hf = D^2 f(x)$ at each point x, which is a symmetric matrix.

Since f is convex, its hessian is positive semi-definite

Combining the two above we get that for every x the matrix $D^2 f(x)$ has non-negative eigenvalues $\lambda_i(x)$ with $i \in T(d)$.

Choosing the matrix norm to be the 2, 2-norm, see DefinitionA.29 discussed in the subsection of matrix norms, then PropositionA.8 implies:

$$\left\|D^2 f\right\|_{2,2} = \max_{i \in T(d)} \lambda_i$$

So, omitting the 2, 2 matrix norm, we get:

$$\left\|D^2 f\right\|_{L^1 B(x_0,r)} = \left\|\max_{i \in T(d)} \{\lambda_i\}\right\|_{L^1 B(x_0,r)} \leq \left\|\sum_{i=1}^{d} \lambda_i\right\|_{L^1 B(x_0,r)}$$

Moreover, we know that:

$$\Delta f(x) = \operatorname{tr}\big(D^2 f(x)\big) = \sum_{i=1}^{d} \lambda_i(x)$$

$$\Rightarrow \|\Delta f\|_{L^1 B(x_0,r)} = \left\|\sum_{i=1}^{d} \lambda_i\right\|_{L^1 B(x_0,r)}$$

Hence,

$$\left\|D^2 f\right\|_{L^1 B(x_0,r)} \leq \|\Delta f\|_{L^1 B(x_0,r)}$$

Green's formula (also known as divergence theorem) implies that:

$$\left\|D^2 f\right\|_{L^1 B(x_0,r)} \leq \|\nabla f\|_{L^1 \partial B(x_0,r)} := \int_{\partial B(x_0,r)} \langle \nabla f, \bar{n} \rangle \, dS$$

where $\bar{n}$ is the outward pointing, unit, normal vector field along the surface of the boundary

Due to the fact that f is Lipschitz on every neighbourhood we have that

$$\exists\, K_r > 0 \quad \forall\, x \in B(x_0, r) \quad \|\nabla f\| \leq K_r$$

Hence, the result follows from the inequality

$$\begin{aligned} \int_{\partial B(x_0,r)} \langle \nabla f, \bar{n} \rangle \, dS &\leq \int_{\partial B(x_0,r)} \|\nabla f\| \cdot \|\bar{n}\| \, dS \\ &\leq K_r \int_{\partial B(x_0,r)} 1 \, dS \end{aligned}$$

□

Setting f equal to $P_t^{n,*}$, which is Lipschitz on every neighbourhood because on the ball $B(x_0, r)$

$$\|\nabla P_t^n(x)\| \leq \frac{\sqrt{2}}{2} + r$$

Indeed,

$$\begin{aligned} \|\nabla P_t^n(x)\| &= \|\nabla P_t^n(x) - x + x\| \\ &\leq \|\nabla P_t^n(x) - x\| + \|x\| \\ &\leq \frac{\sqrt{2}}{2} + r \end{aligned}$$

when $x \in B(x_0, r)$

Thus, we get the inequality:

$$\int_{B(x_0,r)} \left\|D^2 P_t^n\right\| \leq C_r$$

which leads to

$$U_t^n \in L^\infty\left([0, +\infty), W_{loc}^{1,1}(\mathbb{R}^2)\right)$$

By fractional Sobolev emdeddings we have that $\forall r \in (0, 1)$ and $1 \leq k < \frac{2}{1+r}$

$$L^\infty\left([0, +\infty), W_{loc}^{1,1}(\mathbb{R}^2)\right) \subseteq L^\infty\left([0, +\infty), W_{loc}^{r,k}(\mathbb{R}^2)\right)$$

Choosing $s = r = \frac{1}{2}$ and $q = k = \frac{5}{4}(< \frac{4}{3})$ yields the desired.

We proceed with the third and final part, which is to show that the limit U_t and the convex function P_t whose gradient sends ρ_t to dx satisfy the relation that connects them in dual SG.

2.2.3 Part III: $U_t = (\nabla P_t^* - Id)^\perp$

For every time t, we apply Theorem2.1 with ρ_t and the Lebesgue measure to obtain a unique (up to additive constant) convex function such that:

$$\rho_t = \nabla P_{t\#} dx$$

$$P_t - \frac{\|x\|_2^2}{2} \text{ is periodic}$$

$$\|\nabla P_t - Id\| \leq \frac{\sqrt{2}}{2}$$

Since $\rho_t^n \to \rho_t$ in $L^p_{loc}\left([0,+\infty),\left(W^{s,q}_{loc}(\mathbb{R}^2)\right)^*\right)$ $\forall q \geq 1, p > 1$ and $s > 0$ we deduce that:

$$\rho_t^{k_n} \to \rho_t \text{ in } \left(W^{s,q}_{loc}(\mathbb{R}^2)\right)^* \text{ for a.e. } t \geq 0$$

Since $\rho_t^n \in L^\infty\left([0,+\infty), L^\infty(\mathbb{R}^2)\right)$ we deduce that:

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

By stability of optimal transport maps we deduce that:

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

Since $U_t^n = \eta_{\frac{1}{n}} * \left(\nabla P_t^{n,*} - Id\right)^\perp$ it follows that:

$$U_t^n \to \left(\nabla P_t^* - Id\right)^\perp \text{ in } L^1_{loc}(\mathbb{R}^2) \text{ for a.e. } t \geq 0$$

Due to the fact that $U_t^n \rightharpoonup^* U_t$ in $L^\infty\left(\mathbb{R}^2 \times [0,+\infty) : \mathbb{R}^2\right)$

$$U_t = \left(\nabla P_t^* - Id\right)^\perp$$

□

CHAPTER 3

Local in time smooth solutions for the dual SG system

Now, we move past the weak solutions to discover strong/smooth solutions of the SG system (1.2.1). Although we will have to "step down" in terms of the time of existence of our solutions. We have to sacrifice the global, in order to achieve classic solutions for the dual SG system.

Hence, we state below our main theorem for smooth local solutions, which we will prove in two parts. First, we will prove the existence of local smooth solutions following the logic and mimicking the arguments in the proof of weak global solutions. Secondly, we will prove the uniqueness of our existing local smooth solutions, splitting the proof in three parts.

Theorem 3.1. If

$$\exists\, \alpha \in (0,1) \text{ and } \lambda, \Lambda > 0 \text{ such that } 0 < \lambda \leq \rho_0 \leq \Lambda \text{ and } \rho_0 \in C^{0,\alpha}(\mathbb{T}^2)$$

then

$\exists T_{\lambda,\Lambda,\|\rho_0\|_{C^{0,\alpha}(\mathbb{T}^2)}} > 0$,unique ρ_t, P_t^* on $[0,T]$ solving the dual SG system (1.2.1)
and satisfying
$0 < \lambda \leq \rho_t \leq \Lambda$, $\rho_t \in L^\infty\left([0,T], C^{0,\alpha}(\mathbb{T}^2)\right)$, $P_t^* \in L^\infty\left([0,T], C^{2,\alpha}(\mathbb{T}^2)\right)$

Before we begin proving Theorem3.1, let us present the basic ideas and notions used, in order of appearance. We will do so using a sketch of proof paragraph first, in which we will describe our reasoning, followed then by a short diagram "exposing" the very key elements and "expanding" the important steps even further.

3.1 Existence

Sketch of proof

Main steps to build an approximate solution

The idea is to obtain a convex function using Theorem2.1 and freeze in time the velocity vector field U_t. Then, we define the flow Y_t of the velocity and we utilize the pushforward measure of ρ_0 with the flow function to define ρ_t

Why one iteration isn't enough

Of course, by doing so, we will have not solved the dual SG's continuity equation $\partial_t \rho_t + \text{div}(U_t \rho_t) = 0$ that we wanted to. Because the solution which we have found on the previous step needs a given velocity vector field in order to ba obtained (so we only know that it satisfies the equality of Definition1.1). Thus, we do not know if it satisfies the equality of Definition2.1 that is needed in order to be a weak solution of the dual SG system.

The sequence of approximate solutions

That is why we repeat the process. We introduce the (random, with no particular choosing) natural number n, which we fix and then we repeat the steps as described above. Leading us to a family (sequence) of approximate solutions in the same interval, for our problem. After that we will send n to infinity (weak convergence under some norm) giving us (not immediately) the time-evolving solution.

Local in time

The main steps above can be followed in the entire time line $[0, +\infty)$, but the choice of a specific $T > 0$ and the restriction of our study to $[0, T]$, hence the local character of the solution presented, is necessary for the estimate of ρ_t with respect to $C^{0,\alpha}$ norm to hold.

Inequalities

When it is time for us to estimate the approximate solution we created, two things stand out the most. Firstly, Cafarelli's regularity theory. Secondly, the estimates we get through Gronwall's lemma using the flow.

The proof diagram

Main steps, first iteration $(n = 1)$

$C_0 := 2||\rho_0||_{C^{0,\alpha}(\mathbb{T}^2)}$

|

We start off using the initial data ρ_0
by Theorem2.1
$\exists!$ (up to additive constant) convex P_0
such as $\rho_0 = \nabla P_{0\#} dx$
and $P_0 - \frac{\|x\|^2}{2}$ is periodic

|

We define the "time-freezed" velocity vector field
$U_0 := \left(\nabla P_0^* - Id\right)^{\perp}$

|

For all $t \in [0, +\infty)$
we solve $\partial_t \rho_t + \text{div}(\rho_t U_0)$
with the help of Y_t the flow of U_0

|

We define the solution $\rho_t := Y_{t\#}\rho_0$
to the aforementioned measure continuity equation

|

In order to construct the desired sequence
we define its first term:
$P_t^1 := P_0$ and $U_t^1 := U_0$
$\rho_t^1 := \rho_t$ and $Y_t^1 := Y_t$
as well

|

Thus for $n = 1$
we have:

$$\begin{cases} \partial_t \rho_t^n + \text{div}(\rho_t^n U_t^n) = 0 \\ U_t^n = \left(\nabla P_t^{n,*} - Id\right)^{\perp} \\ \rho_0^n = \nabla {P_t^n}_{\#} dx \\ \rho_t^n = {Y_t^n}_{\#} \rho_0 \end{cases}$$

|

We choose the suitable $T > 0$, for which ρ_t^1 remains Holder continuous

|

We then restrict everything on $t \in [0, T] = [0, \frac{T}{n}]$ for $n = 1$

Repeating the procedure $(n = 2)$

We split $[0, T]$ in half, that is $[0, \frac{T}{2}) \cup [\frac{T}{2}, T)$

For $t \in [0, \frac{T}{2}]$ we have exact same process
|
P_0 by Theorem2.1
|
$U_0 := \left(\nabla P_0^* - Id\right)^{\perp}$
|
Y_t flow of U_0
|
$\rho_t := Y_{t\#}\rho_0$
|
We define on $t \in [0, \frac{T}{2})$
$P_t^2 := P_0$
$U_t^2 := U_0$
$\rho_t^2 := \rho_t$
$Y_t^2 := Y_t$

For $t \in [\frac{T}{2}, T)$ utilizing $\rho_{T/2}$
|
we have by Theorem2.1 that
$\exists!$ (up to additive constant) convex $P_{T/2}$
such as $\rho_{T/2} = \nabla P_{T/2 \#} dx$
and $P_{T/2} - \frac{\|x\|^2}{2}$ is periodic
|
We define the "time-freezed" velocity
$U_{T/2} := \left(\nabla P_{T/2}^* - Id\right)^{\perp}$
|
We define Y_t the flow of $U_{T/2}$
|
We define the solution of coninuity equation
$\rho_t := Y_{t\#}\rho_{T/2}$
|
We define on $t \in [\frac{T}{2}, T)$
$P_t^2 := P_{T/2}$
$U_t^2 := U_{T/2}$
$\rho_t^2 := \rho_t$
$Y_t^2 := Y_t$

Thus for $n = 2$ we have:

$$\begin{cases} \partial_t \rho_t^n + \operatorname{div}(\rho_t^n U_t^n) = 0 \text{ for } t \in [0, T] \\ U_t^n = \left(\nabla P_t^{n,*} - Id\right)^{\perp} \text{ for } t \in [0, T] \\ \rho_0^n = \nabla P_t^n{}_{\#} dx \text{ for } t \in [0, \frac{T}{2}) \\ \rho_{T/2}^n = \nabla P_t^n{}_{\#} dx \text{ for } t \in [\frac{T}{2}, T) \\ \rho_t^n = Y_t^n{}_{\#} \rho_0 \text{ for } t \in [0, \frac{T}{2}) \\ \rho_t^n = Y_t^n{}_{\#} \rho_{T/2} \text{ for } t \in [\frac{T}{2}, T) \end{cases} \Leftrightarrow \begin{cases} \partial_t \rho_t^n + \operatorname{div}(\rho_t^n U_t^n) = 0 \\ U_t^n = \left(\nabla P_t^{n,*} - Id\right)^{\perp} \\ \rho_{iT/n}^n = \nabla P_t^n{}_{\#} dx \text{ for } t \in [i\frac{T}{n}, (i+1)\frac{T}{n}) \\ \rho_t^n = Y_t^n{}_{\#} \rho_{iT/n} \text{ for } t \in [i\frac{T}{n}, (i+1)\frac{T}{n}) \end{cases}$$

for $i \in T_0(n-1)$

Estimates

Caffarelli's regularity theory can be found in [23] section 5.1 and [11]

By Caffarelli's regularity theory

$$||D^2 P_t^{n,*}||_{C^{0,\alpha}(\mathbb{T}^2)} \leq C_1(\lambda, \Lambda, C)$$

|

$$||\nabla U_t^n||_{L^\infty(\mathbb{T}^2)} \leq C_1 + 1 := C_2$$

|

$$\begin{cases} \partial_t Y_t^n(y) = U_t^n\left(Y_t^n(y)\right) \\ Y_0^n(y) = y \end{cases}$$

|

differentiating with respect to y

$$\begin{cases} \partial_t \nabla Y_t^n(y) = \Big(\nabla U_t^n\left(Y_t^n(y)\right)\Big) \diamond \nabla Y_t^n(y) \\ \nabla Y_0^n(y) = I_{2\times 2} \end{cases}$$

|

$$\begin{cases} \partial_t ||\nabla Y_t^n(y)|| \leq C_2 ||\nabla Y_t^n(y)|| \\ ||\nabla Y_0^n|| = 1 \end{cases}$$

|

$$e^{-C_2 t} \leq \|\nabla Y_t^n(y)\| \leq e^{C_2 t}$$

|

Also, since

$$\rho_t^n = \rho_0 \circ (Y_t^n)^{-1}$$

$$\lambda \leq \rho_t^n \leq \Lambda \qquad ||\rho_t^n||_{C^{0,\alpha}(\mathbb{T}^2)} \leq C_0$$

So aggregated/collectively/combined we have the following

$$\begin{cases} ||D^2 P_t^{n,*}||_{C^{0,\alpha}(\mathbb{T}^2)} \leq C_1 \\ ||\nabla U_t^n||_{L^\infty(\mathbb{T}^2)} \leq C_2 \\ e^{-C_2 t} \leq ||\nabla Y_t^n(y)|| \leq e^{C_2 t} \\ ||\rho_t^n||_{C^{0,\alpha}(\mathbb{T}^2)} \leq C_0 \\ \lambda \leq \rho_t^n \leq \Lambda \end{cases}$$

Proof.

3.1.1 Constructing the approximate solution

We set $C_0 := 2||\rho_0||_{C^{0,\alpha}(\mathbb{T}^2)}$ and let $n \in \mathbb{N}$. Since ρ_0 and the Lebesgue measure are probability measures on the torus, we apply Theorem2.1 to obtain a unique convex function $P_0 : \mathbb{R}^2 \to \mathbb{R}$ whose gradient sends ρ_0 to dx i.e.

$$\rho_0 = \nabla P_{0\#} dx$$

and such as $P_0 - \frac{\|x\|^2}{2}$ is periodic.

Let $b > 0$ for $t \in [0, b]$ we proceed to "freeze" the velocity vector field on this interval. We define

$$P_t^n := P_0$$

$$U_t^n := \left(\nabla P_t^{n,*} - Id\right)^{\perp}$$

Remark. Notice that by definition both the pressure P_t^n and the velocity U_t^n are constant in terms of t and n. That is why we say that we have "frozen" the velocity, meaning that it is time independent.

By Caffarelli's reguralty theory for the Monge-Ampère equation we also have that

$$\exists\, C_1(\lambda, \Lambda, C_0) > 0 \quad ||D^2 P_t^n||_{C^{0,\alpha}(\mathbb{T}^2)} \le C_1$$

Thus, we can obtain a bound (time t and space y independent) for the gradient of the velocity vector field, indeed:

First, we notice that since $P_t^{n,*}$ is $C^{2,\alpha}$ and due to the definition of the velocity as $U_t^n := \left(\nabla P_t^{n,*} - Id\right)^{\perp}$ we have that U_t^n is $C^{1,\alpha}$. Indeed, we can take the classical gradient.

Then we calculate the gradient using the fact that the gradient of a perpendicular vector equals the perpendicular of the gradient of the vector.

$$U_t^n = \left(\nabla P_t^{n,*} - Id\right)^{\perp} \Rightarrow$$
$$\nabla U_t^n = \nabla\Big(\left(\nabla P_t^{n,*} - Id\right)^{\perp}\Big) = \Big(\nabla\left(\nabla P_t^{n,*} - Id\right)\Big)^{\perp} = \left(D^2 P_t^{n,*} - I_{2\times 2}\right)^{\perp}$$

And now we calculate the L^∞-norm using the fact that the norm of a perpendicular vector is the same as the norm of the vector itself.

$$||\nabla U_t^n||_{L^\infty(\mathbb{T}^2)} = ||(\nabla U_t^n)^\perp||_{L^\infty(\mathbb{T}^2)} = ||D^2 P_t^{n,*} - I_{2\times 2}||_{L^\infty(\mathbb{T}^2)} \overset{\text{triangle inequality}}{\leq}$$
$$\leq ||D^2 P_t^{n,*}||_{L^\infty(\mathbb{T}^2)} + ||I_{2\times 2}||_{L^\infty(\mathbb{T}^2)}$$

It holds that $||D^2 P_t^{n,*}||_{L^\infty(\mathbb{T}^2)} \leq ||D^2 P_t^{n,*}||_{C^{0,\alpha}(\mathbb{T}^2)} \leq C_1$ and

$||I_{2\times 2}||_{L^\infty(\mathbb{T}^2)} = \max\{ \|(1,0)\|_{L^\infty(\mathbb{T}^2)}, \|(0,1)\|_{L^\infty(\mathbb{T}^2)} \} = 1$

So, by combining the two above, we have that $||\nabla U_t^n||_{L^\infty(\mathbb{T}^2)} \leq C_1 + 1$

Remark. As we have already discussed in Subsection2.2.1, since the gradient of $U^n(y,t)$ is bounded in time as well as in the spatial variable y, the flow is indeed well defined (the initial value problem has a unique solution) in the whole interval $[0,b]$. Existence is of course needed, but the uniqueness is also crucial because we want to define a function $y \mapsto Y_t^n(y)$ and to do so we need the initial value problem to have a unique solution for the initial data $y = Y^n(0)$

We then define for every $y \in \mathbb{R}^2$ and the fixed $n \in \mathbb{N}$ the flow Y_t^n of U_t^n, which is the unique solution of the initial value problem

$$\begin{cases} \partial_t Y^n(t) = U^n\left(Y^n(t), t\right) \\ Y^n(0) = y \end{cases}$$

The n isn't a variable (for now), flow is n-invariable in every step. The flow is the first function to be time-dependent so far, and it will help us to make things actually "flow" in time.

To this end, we define the density ρ_t^n using the flow $Y_t^n(y)$ to send it to ρ_0

$$\rho_t^n := {Y_t^n}_{\#}\rho_0$$

Thus, $\rho_t^n := \rho_0 \circ (Y_t^n)^{-1} \xRightarrow[\lambda \leq \rho_0 \leq \Lambda]{} \lambda \leq \rho_t^n \leq \Lambda.$

To obtain a bound for the $C^{0,\alpha}$-norm of ρ_t^n we will "pass through" a bound for the Euclidean norm of the flow Y_t^n. We can rewrite flow's initial value problem to read as:

$$\begin{cases} \partial_t Y_t^n(y) = U_t^n\left(Y_t^n(y)\right) \\ Y_0^n(y) = y \end{cases}$$

Notice that due to the fact that the velocity field U_t^n is C^1, the time derivative of the flow is also C^1. We then differentiate with respect to y and by the chain rule we get:

$$\begin{cases} \nabla\partial_t Y_t^n(y) = \Big(\nabla U_t^n\left(Y_t^n(y)\right)\Big) \diamond \nabla Y_t^n(y) \\ \nabla Y_0^n(y) = I_{2\times 2} \end{cases}$$

Using the symmetry of second derivatives by Schwarz's theorem for mixed partials we have that $\nabla\partial_t Y_t^n(y) = \partial_t \nabla Y_t^n(y)$. Hence,

$$\begin{cases} \partial_t \nabla Y_t^n(y) = \Big(\nabla U_t^n\left(Y_t^n(y)\right)\Big) \diamond \nabla Y_t^n(y) \\ \nabla Y_0^n(y) = I_{2\times 2} \end{cases} \Rightarrow$$

$$\begin{cases} ||\partial_t \nabla Y_t^n(y)|| = \Big\|\Big(\nabla U_t^n\left(Y_t^n(y)\right)\Big) \diamond \nabla Y_t^n(y)\Big\| \\ ||\nabla Y_0^n(y)|| = ||I_{2\times 2}|| \end{cases}$$

Our goal now is to apply Gronwall's lemma on the function $||\nabla Y_t^n(y)||$, to do so we need to find an estimate for its time derivative involving the function itself. We set out to prove that $\Big|\partial_t ||\nabla Y_t^n(y)||\Big| \leq C_2 ||\nabla Y_t^n(y)||$

Proposition 3.1. Let $f : A \subseteq \mathbb{R} \to (\mathbb{R}^n, \langle \cdot, \cdot \rangle)$ with $f \in C^1(A)$, that is f is continuously differentiable then

$$\Big|\partial_t ||f(t)||\Big| \leq ||\partial_t f(t)|| \tag{3.1.1}$$

Proof.

$$\Big|\partial_t ||f(t)||\Big| = \Big|\partial_t \sqrt{\langle f(t), f(t)\rangle}\Big| \overset{\text{chain}}{\underset{\text{rule}}{=}} \left|\frac{1}{2\sqrt{\langle f(t), f(t)\rangle}} \cdot 2\langle \partial_t f(t), f(t)\rangle\right|$$

$$= \left|\frac{1}{\sqrt{\langle f(t), f(t)\rangle}}\right| \cdot \Big|\langle \partial_t f(t), f(t)\rangle\Big| \overset{\sqrt{x}\geq 0}{=} \frac{1}{\sqrt{\langle f(t), f(t)\rangle}} \cdot \Big|\langle \partial_t f(t), f(t)\rangle\Big| \leq$$

$$\overset{\text{Cauchy-Schwarz}}{\underset{\text{inequality}}{\leq}} \frac{||\partial_t f(t)|| \cdot ||f(t)||}{\sqrt{\langle f(t), f(t)\rangle}} = \frac{||\partial_t f(t)|| \cdot ||f(t)||}{||f(t)||} = ||\partial_t f(t)||$$

□

Using the proposition above we obtain: $\Big|\partial_t ||\nabla Y_t^n(y)||\Big| \leq ||\partial_t \nabla Y_t^n(y)||$

Since we have shown that $||\partial_t \nabla Y_t^n(y)|| = \Big\|\Big(\nabla U_t^n\left(Y_t^n(y)\right)\Big) \diamond \nabla Y_t^n(y)\Big\|$

Using the Frobenius norm submultiplicativity PropositionA.6 we have that:

$$\Big\|\Big(\nabla U_t^n\left(Y_t^n(y)\right)\Big) \diamond \nabla Y_t^n(y)\Big\| \leq \Big\|\Big(\nabla U_t^n\left(Y_t^n(y)\right)\Big)\Big\| \cdot ||\nabla Y_t^n(y)||$$

And because of the time and space boundness of the velocity vector field's gradient: $||\nabla U_t^n||_{L^\infty(\mathbb{T}^2)} \leq C_2$ we have aggregately proved that indeed

$$\Big|\partial_t ||\nabla Y_t^n(y)||\Big| \leq C_2 ||\nabla Y_t^n(y)||$$

If we expand the absolute value in the inequality above we have that

$$-C_2||\nabla Y_t^n(y)|| \leq \partial_t ||\nabla Y_t^n(y)|| \leq C_2 ||\nabla Y_t^n(y)||$$

So, right now, we are able to put in use both Gronwall lemmas to obtain the inequalities:
$e^{-C_2 t}||\nabla Y_0^n|| \leq ||\nabla Y_t^n|| \leq e^{C_2 t}||\nabla Y_0^n|| \ \ \forall t \in [0,b] \xRightarrow[||\nabla Y_0^n||=1]{}$
$e^{-C_2 t} \leq ||\nabla Y_t^n|| \leq e^{C_2 t}$

The next step is to show that the defined density (the solution of continuity equation is a function as it has been discussed in the Subsection2.2.1) ρ_t^n is $C^{0,\alpha}$. To do so, we will show that $||\rho_t^n||_{C^{0,\alpha}(\mathbb{T}^2)} \leq e^{C_2 t}||\rho_0||_{C^{0,\alpha}(\mathbb{T}^2)}$ and we will finally choose a particular $T > 0$, which will preserve the Holder continuity for the approximate solutions and make them (and the actual solutions) local in time.

To prove the asserted inequality we will show that the composition of a Lipschitz continuous function with a Holder continuous function is Holder continuous as well.

Before we move on to state and prove the proposition we are going to need, let's check out that this is indeed our case. Thanks to the bound $e^{-C_2 t} \leq ||\nabla Y_t^n|| \leq e^{C_2 t}$ of the spatial derivative of the function $Y_t^n : \mathbb{R}^2 \to \mathbb{R}^2$ we have that it is actually a bi-Lipschitz homeomorphism i.e. we have that $e^{-C_2 t}||x-y|| \leq ||Y_t^n(x) - Y_t^n(y)|| \leq e^{C_2 t}||x-y||$ and Y_t^n is an injective and surjective function from $\mathbb{R}^2$ to $\mathbb{R}^2$. So Y_t^n is invertible, hence $\forall t \in [0,b] \ \ \exists!\, (Y_t^n)^{-1} : \mathbb{R}^2 \underset{\text{onto}}{\overset{1-1}{\to}} \mathbb{R}^2$

Also

Definition 3.1 (reverse Lipschitz)**.** We call a function $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ reverse Lipschitz with constant $K :\iff \exists K > 0 \ \ ||f(x) - f(y)|| \geq K||x-y|| \ \ \forall x, y \in U$

Clarification 3.1.1. The norm symbol appearing in the above inequality refers to the Euclidean (or any equivalent) norm on the respective spaces i.e. $\mathbb{R}^n$ and $\mathbb{R}^m$.

Proposition 3.2. Let f, g be two functions where g is real-valued, f is invertible with $f(D_f) \subseteq D_f = D_g$. If f^{-1} is reverse Lipschitz with constant K and g is $C^{0,\alpha}$, then $g \circ f$ is also $C^{0,\alpha}$.

Moreover if $K \leq 1$ we have that:

$$||g \circ f||_{C^{0,\alpha}(D_f)} \leq \frac{1}{K^\alpha}\,||g||_{C^{0,\alpha}(D_f)} \tag{3.1.2}$$

Proof. For simplicity we will denote the domain D_f of f as U. By definition we know that:

$$||g \circ f||_{C^{0,\alpha}(U)} = \sup_U |g \circ f| \ + \sup_{\substack{x \neq y \\ U}} \frac{|g(f(x)) - g(f(y))|}{||x-y||^\alpha}$$

We proceed estimating each quantity: $\sup\limits_U |g \circ f| = \sup\limits_{f(U)} |g| \overset{f(U) \subseteq U}{\leq} \sup\limits_U |g|$

Since f is invertible, $\exists f^{-1} : f(U) \underset{1-1}{\overset{\text{onto}}{\to}} U$ which means that for all $x, y \in U$ there exist unique $z, w \in f(U)$ such that $f^{-1}(z) = x \Leftrightarrow z = f(x)$ and $f^{-1}(w) = y \Leftrightarrow w = f(y)$.

Also $x = y \Leftrightarrow f^{-1}(x) = f^{-1}(y) \Leftrightarrow z = w$, because f^{-1} is an injective (which implies the straightforward direction of the equivalence) function (justifies the reverse direction). Thus we can rewrite the seminorm as:

$$\sup_{\substack{z \neq w \\ f(U)}} \frac{|g(z) - g(w)|}{||f^{-1}(z) - f^{-1}(w)||^{\alpha}}$$

Now we will make use of the fact that f^{-1} is a reverse K-Lipschitz function, which implies that $\exists K > 0$ such $||f^{-1}(\tilde{x}) - f^{-1}(\tilde{y})|| \geq K||\tilde{x} - \tilde{y}||$ $\forall x, y \in f(U)$ (we use the tilde symbol to avoid confusion and conflict with the previously used x and y). Choosing $\tilde{x} = z$ and $\tilde{y} = w$, since $z \neq w$, we have that:

$$||f^{-1}(z) - f^{-1}(w)|| \geq K||z - w|| \overset{\frac{1}{x}\nearrow}{\underset{x>0}{\Longrightarrow}} \frac{1}{||f^{-1}(z) - f^{-1}(w)||} \leq \frac{1}{K||z - w||} \Rightarrow$$

$$\xrightarrow[\text{positive bases}]{\text{exponent } \alpha>0} \frac{1}{||f^{-1}(z) - f^{-1}(w)||^{\alpha}} \leq \frac{1}{K^{\alpha}\,||z - w||^{\alpha}}$$

Putting together all the above we have shown that:

$$\begin{aligned} ||g \circ f||_{C^{0,\alpha}(U)} = & \sup_{U} |g \circ f| + \sup_{\substack{x \neq y \\ U}} \frac{|g(f(x)) - g(f(y))|}{||x - y||^{\alpha}} \\ \leq & \sup_{U} |g| + \sup_{\substack{z \neq w \\ f(U)}} \frac{|g(z) - g(w)|}{||f^{-1}(z) - f^{-1}(w)||^{\alpha}} \\ \leq & \sup_{U} |g| + \frac{1}{K^{\alpha}} \sup_{\substack{z \neq w \\ f(U)}} \frac{|g(z) - g(w)|}{||z - w||^{\alpha}} \\ \leq & \sup_{U} |g| + \frac{1}{K^{\alpha}} \sup_{\substack{z \neq w \\ U}} \frac{|g(z) - g(w)|}{||z - w||^{\alpha}} \end{aligned}$$

To finalise the proof we discern the three possible cases of $K > 0$

i) If $K = 1$ then immediately we obtain: $||g \circ f||_{C^{0,\alpha}(U)} \leq ||g||_{C^{0,\alpha}(U)}$

ii) If $K < 1$ then $\frac{1}{K^{\alpha}} > 1 \Rightarrow 1 < \frac{1}{K^{\alpha}}$ and since $\sup_{U} |g| > 0$ we obtain:

$$||g \circ f||_{C^{0,\alpha}(U)} \leq \frac{1}{K^{\alpha}} ||g||_{C^{0,\alpha}(U)}$$

iii) If $K > 1$ then $\frac{1}{K^{\alpha}} < 1$ and since the seminorm is positive we obtain:

$$||g \circ f||_{C^{0,\alpha}(U)} \leq ||g||_{C^{0,\alpha}(U)}$$

Remark.
The special case where $\sup_{U} |g| = 0$ or $\sup_{\substack{z \neq w \\ U}} \frac{|g(z) - g(w)|}{||z - w||^{\alpha}} = 0$ does somewhat easily imply

the same result. Indeed if the supremum of a non-negative quantity is zero, then the quantity itself is constant and equals zero. Thus $g \equiv 0$ or $g(z) = g(w)\ \ \forall z, w \in f(U) = D_g$, which both imply that the function g is constant so $g \circ f$ is constant as well, hence the proposition is proven.

□

Returning to our particular case. In order to implement Proposition3.2, we firstly recall that $\rho_t^n = \rho_0 \circ (Y_t^n)^{-1}$. So we readily choose $g = \rho_0 \in C^{0,\alpha}(\mathbb{T}^2)$ and $f = (Y_t^n)^{-1}$ for each time t. Since the flow is a bi-Lipschitz homeomorphism of the whole space to itself, it follows that $(Y_t^n)^{-1}(\mathbb{R}^2) = \mathbb{R}^2$. Now all it remains to be shown is that f^{-1} is reverse Lipschitz. Indeed, $f^{-1} = Y_t^n$ for which it holds that $e^{-C_2 t}||x - y|| \leq ||Y_t^n(x) - Y_t^n(y)||$. Hence f^{-1} is Lipschitz with positive constant $e^{-C_2 t} \leq 1$, since $t \geq 0$ and $C_2 > 0$. Because the Lipschitz constant is also space-independent (i.e. it doesn't depend on the space variable, although it is time-dependent) we can apply the recently proven proposition to obtain the following bound for the $C^{0,\alpha}$-norm of the measure ρ_t^n

$$||\rho_t^n||_{C^{0,\alpha}(\mathbb{T}^2)} \leq \frac{1}{(e^{-C_2 t})^{\alpha}}\,||\rho_0||_{C^{0,\alpha}(\mathbb{T}^2)} = e^{\alpha C_2 t}||\rho_0||_{C^{0,\alpha}(\mathbb{T}^2)}$$

Since $\alpha \in (0,1)$, $C_2 > 0$ and $t \geq 0$ we have that $\alpha C_2 t \leq C_2 t$. The monotonicity of the exponential implies that $e^{\alpha C_2 t} \leq e^{C_2 t}$. Hence

$$||\rho_t^n||_{C^{0,\alpha}(\mathbb{T}^2)} \leq e^{C_2 t}||\rho_0||_{C^{0,\alpha}(\mathbb{T}^2)}$$

At last, it is time to choose the T that will work for us. Our purpose is to show that the measure ρ_t^n is $C^{0,\alpha}$. So, we specifically choose a (there are plenty numbers satisfying this property) positive real number T such $T < \frac{ln2}{C_2}$. The chosen $T > 0$ satisfies the inequality $e^{C_2 T} < 2$

Thus $||\rho_t^n||_{C^{0,\alpha}(\mathbb{T}^2)} \leq 2||\rho_0||_{C^{0,\alpha}(\mathbb{T}^2)} = C_0$

And we restrict the entire previous study in the time interval $[0, \frac{T}{n}]$

Before we repeat the procedure and "initiate" the second iteration, let us gather here what we have so far to help us understand better what we have achieved.

So, collectively, for $t \in [0, \frac{T}{n})$ with $t = \frac{T}{n}$ included for ρ_t , for every $y \in \mathbb{R}^2$ and all $n \in \mathbb{N}$ we have constructed a triplet of sequences $P_t^n(y)$ $U_t^n(y)$ ρ_t^n , for which the followings are true:

$$\begin{cases} \rho_t^n := {Y_t^n}_{\#}\rho_0 \Rightarrow \partial_t \rho_t^n + \operatorname{div}(U_t^n \rho_t^n) = 0 \\ U_t^n = \left(\nabla P_t^{n,*} - Id\right)^{\perp} \\ \rho_0^n = {Y_0^n}_{\#}\rho_0 = Id_{\#}\rho_0 = \rho_0 = \nabla P_{0\#}dx = {\nabla P_t^n}_{\#}dx \\ \rho_t^n = {Y_t^n}_{\#}\rho_0 \end{cases}$$

$$\begin{cases} \partial_t \rho_t^n + \operatorname{div}(U_t^n \rho_t^n) = 0 \\ U_t^n = \left(\nabla P_t^{n,*} - Id\right)^{\perp} \\ \rho_0^n = \nabla P_t^n{}_{\#} dx \\ \rho_t^n = Y_t^n{}_{\#} \rho_0 \end{cases}$$

Second iteration $t \in [\frac{T}{n}, 2\frac{T}{n})$ with $2 \leq n$

We repeat the process with $\rho_{T/n}^n$ in the place of ρ_0

We note that, by restricting in $t \in [0, \frac{T}{n})$ previously with $t = \frac{T}{n}$ included for ρ_t^n, we have shown that:

$$\forall t \in [0, \tfrac{T}{n}]$$

$$\begin{aligned} &\|\rho_t^n\|_{C^{0,\alpha}(\mathbb{T}^2)} \leq e^{C_2 t} \|\rho_0\|_{C^{0,\alpha}(\mathbb{T}^2)} \\ \Rightarrow &\left\|\rho_{T/n}^n\right\|_{C^{0,\alpha}(\mathbb{T}^2)} \leq e^{C_2 \frac{T}{n}} \|\rho_0\|_{C^{0,\alpha}(\mathbb{T}^2)} \end{aligned}$$

and

$$\begin{aligned} &\lambda \leq \rho_t^n \leq \Lambda \\ \Rightarrow &\lambda \leq \rho_{T/n}^n \leq \Lambda \end{aligned}$$

and

$$\begin{aligned} &\|\rho_t^n\|_{C^{0,\alpha}(\mathbb{T}^2)} \leq C_0 \\ \Rightarrow &\left\|\rho_{T/n}^n\right\|_{C^{0,\alpha}(\mathbb{T}^2)} \leq C_0 \end{aligned}$$

We define for

$$t \in [\tfrac{T}{n}, 2\tfrac{T}{n}) \text{ with } 2 \leq n$$

the quantities

$$P_t^n := P_{T/n}$$

$$U_t^n := \left(\nabla P_t^{n,*} - Id\right)^{\perp}$$

and

$$\begin{aligned} &\rho_t^n := Y_t^n{}_{\#} \rho_{T/n}^n \\ \Rightarrow &\rho_t^n = \rho_{T/n}^n \circ (Y_t^n)^{-1} \end{aligned}$$

where

$$\begin{cases} \partial_t Y_t^n = U_t(Y_t^n) \\ Y_{T/n} = Id \end{cases}$$

Due to

$$\lambda \leq \rho^n_{T/n} \leq \Lambda \text{ and } \left\| \rho^n_{T/n} \right\|_{C^{0,\alpha}(\mathbb{T}^2)} \leq C_0$$

Caffarelli's reguarity theory [23] section 5.1, holds true with the same constant C_1, because it only depends on λ, Λ, C_0 which have remained the same.

$$\left\| D^2 P^n_t \right\|_{C^{0,\alpha}(\mathbb{T}^2)} \leq C_1$$

Hence,

$$\|U^n_t\|_{L^\infty(\mathbb{T}^2)} \leq C_2 := C_1 + 1$$

Replacing in the calculations above, $t \geq 0$ with $t - \frac{T}{n} \geq 0$ since $t \in [\frac{T}{n}, 2\frac{T}{n})$ now and $\rho^n_t = \rho_{T/n} \circ (Y^n_t)^{-1}$, we have that:

$$\begin{aligned}
\|\rho^n_t\|_{C^{0,\alpha}(\mathbb{T}^2)} &\leq e^{C_2(t-\frac{T}{n})} \left\|\rho_{T/n}\right\|_{C^{0,\alpha}(\mathbb{T}^2)} \\
&\leq e^{C_2(t-\frac{T}{n})} e^{C_2\frac{T}{n}} \|\rho_0\|_{C^{0,\alpha}(\mathbb{T}^2)} \\
&= e^{C_2 t} \|\rho_0\|_{C^{0,\alpha}(\mathbb{T}^2)} \\
&\leq e^{C_2 T} \|\rho_0\|_{C^{0,\alpha}(\mathbb{T}^2)} \\
&\leq 2 \|\rho_0\|_{C^{0,\alpha}(\mathbb{T}^2)} \\
&= C_0
\end{aligned}$$

We follow the same process to obtain sequences on the whole time interval $[0, T]$ with the same estimates remaining true.

Inductively, let us assume that at the i-th iteration $t \in [i\frac{T}{n}, (i+1)\frac{T}{n})$ with $t = (i+1)\frac{T}{n}$ included for ρ^n_t and $i \leq n$ we have:

$$\|\rho^n_t\|_{C^{0,\alpha}(\mathbb{T}^2)} \leq e^{C_2 t} \|\rho_0\|_{C^{0,\alpha}(\mathbb{T}^2)}$$

and

$$\lambda \leq \rho^n_t \leq \Lambda$$

$$\|\rho^n_t\|_{C^{0,\alpha}(\mathbb{T}^2)} \leq C_0$$

then at the next iteration $t \in [(i+1)\frac{T}{n}, (i+2)\frac{T}{n})$ with $t = (i+2)\frac{T}{n}$ included for ρ^n_t and $i + 1 \leq n$ we will have as well that:

$$\|\rho^n_t\|_{C^{0,\alpha}(\mathbb{T}^2)} \leq e^{C_2 t} \|\rho_0\|_{C^{0,\alpha}(\mathbb{T}^2)}$$

and

$$\lambda \leq \rho^n_t \leq \Lambda$$

$$\|\rho^n_t\|_{C^{0,\alpha}(\mathbb{T}^2)} \leq C_0$$

Indeed, it is true that

$$\left\|\rho_{(i+1)T/n}\right\|_{C^{0,\alpha}(\mathbb{T}^2)} \le e^{C_2 \frac{(i+1)T}{n}} \|\rho_0\|_{C^{0,\alpha}(\mathbb{T}^2)}$$

and

$$\lambda \le \rho_{(i+1)T/n} \le \Lambda$$

$$\left\|\rho^n_{(i+1)T/n}\right\|_{C^{0,\alpha}(\mathbb{T}^2)} \le C_0$$

We define for

$$t \in [(i+1)\tfrac{T}{n}, (i+2)\tfrac{T}{n}) \text{ with } i+1 \le n$$

the quantities

$$P^n_t := P_{(i+1)T/n}$$

$$U^n_t := \left(\nabla P^{n,*}_t - Id\right)^{\perp}$$

$$\rho^n_t := {Y^n_t}_{\#}\rho_{(i+1)T/n}$$

where

$$\begin{cases} \partial_t Y^n_t = U_t(Y^n_t) \\ Y_{(i+1)T/n} = Id \end{cases}$$

Caffarelli's regurality theory holds true with the same constant C_1

$$\left\|D^2 P^n_t\right\|_{C^{0,\alpha}(\mathbb{T}^2)} \le C_1$$

Hence,

$$\|U^n_t\|_{L^\infty(\mathbb{T}^2)} \le C_2 := C_1 + 1$$

Setting t as $t - \frac{(i+1)T}{n}$ since $\rho^n_t = \rho_{(i+1)T/n} \circ (Y^n_t)^{-1}$ we have that:

$$\begin{aligned}
\|\rho^n_t\|_{C^{0,\alpha}(\mathbb{T}^2)} &\le e^{C_2(t - \frac{(i+1)T}{n})} \left\|\rho_{(i+1)T/n}\right\|_{C^{0,\alpha}(\mathbb{T}^2)} \\
&\le e^{C_2(t - \frac{(i+1)T}{n})} e^{C_2 \frac{(i+1)T}{n}} \|\rho_0\|_{C^{0,\alpha}(\mathbb{T}^2)} \\
&= e^{C_2 t} \|\rho_0\|_{C^{0,\alpha}(\mathbb{T}^2)} \\
&\le e^{C_2 T} \|\rho_0\|_{C^{0,\alpha}(\mathbb{T}^2)} \\
&\le 2\|\rho_0\|_{C^{0,\alpha}(\mathbb{T}^2)} \\
&= C_0
\end{aligned}$$

A similar argument to that of Part II & III of the existence of weak solutions shows the existence of smooth solutions as well. □

3.2 Uniqueness

To show that the local smooth solution of Theorem3.1, whose existence we established in the previous section, is unique, we will utilize a combination of many facts and arguments. So, we are again splitting the proof in several parts.

We again start off with a sketch of proof like we did for the existence part.

Sketch of proof

Equality of flows implies equal solutions

Let us assume we have two solutions as in Theorem3.1. It is enough to show that their respective flows are equal (it is hinted in Subsection2.2.1 by the uniqueness of measure solution σ_t). Indeed notice that for a solution with the properties of Theorem3.1 the velocity vector field is C^1 and we can apply the theory discussed in Subsection2.2.1 without mollifying.

Equality of flows will be proven with Gronwall's lemma

In our effort to prove that the respective flows are equal, we want to prove that the integral of the norm, of their difference, squared, over the torus is zero. To achieve that we will show that th time-dependent integral aforementioned satisfies the condition in Gronwall's lemma.

Construction of the interpolating curve and proving its bounds

To show that the Gronwall lemma is satisfied we will have to estimate several integrals. We start with the flows, we "pass through" the velocities leading to the convex conjugates of pressures. In order to "return" to the flows we create an interpolating curve and utilize the minimality of the optimal transport map from the one density to the other. The bounds will be proved using arguments from the Monge-Ampère equation and will help us "get rid"(bound by a constant) of everything else except the integral over the torus of the squared norm, of the flows difference.

We now proceed to prove that the existing solution of Theorem3.1 is indeed unique.

Proof.

Let ρ_t^1,$P_t^{*,1}$ and ρ_t^2,$P_t^{*,2}$ be two (weak) solutions of the dual SG system (1.2.1) both satisfying the properties stated in Theorem3.1 i.e.

$$\begin{cases} \partial_t \rho_t^i + \operatorname{div}(\rho_t^i U_t^i) = 0 & (x,t) \in \mathbb{R}^2 \times [0,+\infty) \\ U_t^i = (\nabla P_t^{i,*} - Id)^{\perp} & (x,t) \in \mathbb{R}^2 \times [0,+\infty) \\ \rho_t^i := \nabla P_t^i{}_{\#} dx & t \in [0,+\infty) \\ P_0^i = \bar{p} + \frac{\|x\|^2}{2} & x \in \mathbb{R}^2 \end{cases} \tag{3.2.1}$$

Satisfying for $i = 1, 2$ the following:

$$0 < \lambda \leq \rho_t^i \leq \Lambda \ , \ \rho_t^i \in L^\infty\big([0,T], C^{0,\alpha}(\mathbb{T}^2)\big) \ , \ P_t^{i,*} \in L^\infty\big([0,T], C^{2,\alpha}(\mathbb{T}^2)\big)$$

Representing the measures ρ_t^i by their flows

Since $P_t^{*,i} \in L^\infty\left([0,T], C^{2,\alpha}(\mathbb{T}^2)\right)$ we have that $\sup_{t\in[0,T]} ||P_t^{*,i}||_{C^{2,\alpha}(\mathbb{T}^2)} < +\infty$

Thus there exist two constants, time and space independent, that act as an upper bound for the $C^{2,\alpha}-$norm of the respective pressures' convex conjugates $P_t^{*,i}$, that is:

$$\exists C_i > 0 \text{ such that } ||P_t^{*,i}||_{C^{2,\alpha}(\mathbb{T}^2)} \leq C_i \ \ \forall t \in [0,T]$$

and now we repeat unaltered the exact same arguments to prove that the respective velocities U_t^i are C^1 and Lipschitz.

First, we notice that since $P_t^{i,*}$ is $C^{2,\alpha}$ and due to the fact that the velocities satisfy the SG equations (3.2.1) i.e. $U_t^i = \Big(\nabla P_t^{i,*} - Id\Big)^\perp$, we have that both U_t^i are $C^{1,\alpha}$. Hence, U_t^i are C^1.

Then we calculate the gradient using the fact that the gradient of a perpendicular vector equals the perpendicular of the gradient of the vector.

$$U_t^i = \Big(\nabla P_t^{i,*} - Id\Big)^\perp \Rightarrow$$
$$\nabla U_t^i = \nabla\Big(\big(\nabla P_t^{i,*} - Id\big)^\perp\Big) = \Big(\nabla\big(\nabla P_t^{i,*} - Id\big)\Big)^\perp = \Big(D^2 P_t^{i,*} - I_{2\times 2}\Big)^\perp$$

And now we calculate the L^∞-norm using the fact that the norm of a perpendicular vector is the same as the norm of the vector itself.

$$||\nabla U_t^i||_{L^\infty(\mathbb{T}^2)} = ||\big(\nabla U_t^i\big)^\perp||_{L^\infty(\mathbb{T}^2)} = ||D^2 P_t^{i,*} - I_{2\times 2}||_{L^\infty(\mathbb{T}^2)} \overset{\text{triangle inequality}}{\leq}$$
$$\leq ||D^2 P_t^{i,*}||_{L^\infty(\mathbb{T}^2)} + ||I_{2\times 2}||_{L^\infty(\mathbb{T}^2)}$$

It holds that $||D^2 P_t^{i,*}||_{L^\infty(\mathbb{T}^2)} \leq ||D^2 P_t^{i,*}||_{C^{0,\alpha}(\mathbb{T}^2)} \leq C_1$ and
$||I_{2\times 2}||_{L^\infty(\mathbb{T}^2)} = \max\{\, \|(1,0)\|_{L^\infty(\mathbb{T}^2)} , \, \|(0,1)\|_{L^\infty(\mathbb{T}^2)} \,\} = 1$
So, by combining the two above, we have that $||\nabla U_t^i||_{L^\infty(\mathbb{T}^2)} \leq C_i + 1$

Thus each velocity U_t^i is Lipschitz. So, as it has already been mentioned earlier in the existence proof using SubsectionA.8.1, we can define the respective flows Y_t^i of the velocities U_t^i. Moreover we can rewrite the satisfied differential equation of the flows as:

$$\begin{cases} \partial_t Y_t^i = U_t^i\left(Y_t^i\right) \\ Y_0^i = Id \end{cases}$$

So, as it has already been shown in Subsection2.2.1, the unique solution for the continuity equation of the dual SG system with initial data ρ_0 is ${Y_t^i}_{\#}\rho_0$.

Hence each measure ρ_t^i in (3.2.1) equals ${Y_t^i}_{\#}\rho_0$

3.2.1 Flows' equality is enough to provide uniqueness

Before we proceed to actually prove that $Y_t^1 = Y_t^2$, let us verify that this equality provides indeed the wanted result.

If we assume that $Y_t^1 = Y_t^2$ then $\rho_t^1 = {Y_t^1}_{\#}\rho_0 = {Y_t^2}_{\#}\rho_0 = \rho_t^2 \Rightarrow \rho_t^1 = \rho_t^2$.

Since ρ_t^i satisfy the equations of the dual SG system (3.2.1), we also have that $\rho_t^1 = \nabla {P_t^1}_{\#}dx$ and $\rho_t^2 = \nabla {P_t^2}_{\#}dx$

Thus, we can write the measure ρ_t^1 as both $\nabla {P_t^1}_{\#}dx$ and $\nabla {P_t^2}_{\#}dx$. Due to the (up to an additive constant) uniqueness of the convex function P that Theorem2.1 states, we obtain that $\exists c \in \mathbb{R}$ such as $P_t^1 = P_t^2 + c$

$$\Rightarrow P_t^{1,*} = P_t^{2,*} + c \Rightarrow \nabla P_t^{1,*} = \nabla\left(P_t^{2,*} + c\right) = \nabla P_t^{2,*} + \nabla c \overset{\nabla c=0}{=} \nabla P_t^{2,*} \Rightarrow$$

$$\Rightarrow \nabla P_t^{1,*} - Id = \nabla P_t^{2,*} - Id \Rightarrow \left(\nabla P_t^{1,*} - Id\right)^{\perp} = \left(\nabla P_t^{2,*} - Id\right)^{\perp} \Rightarrow$$

$$U_t^1 = U_t^2$$

3.2.2 The Gronwall argument

And now we resume to the main purpose, to apply the Gronwall lemma on the function $\phi_t = \int\limits_{\mathbb{T}^2} ||Y_t^1 - Y_t^2||_2^2\, dy$. Thus we calculate its time derivative.

$$\partial_t \int_{\mathbb{T}^2} ||Y_t^1 - Y_t^2||_2^2\, dy =$$

$$\begin{aligned}
&\underset{\text{integral rule}}{\overset{\text{Leibniz}}{=}} \int_{\mathbb{T}^2} \partial_t ||Y_t^1 - Y_t^2||_2^2\, dy \\
&= \int_{\mathbb{T}^2} \partial_t \left\langle\, Y_t^1 - Y_t^2 \;,\; Y_t^1 - Y_t^2 \,\right\rangle dy \\
&= \int_{\mathbb{T}^2} 2 \left\langle\, Y_t^1 - Y_t^2 \;,\; \partial_t\left(Y_t^1 - Y_t^2\right) \,\right\rangle dy \\
&= \int_{\mathbb{T}^2} 2 \left\langle\, Y_t^1 - Y_t^2 \;,\; \partial_t Y_t^1 - \partial_t Y_t^2 \,\right\rangle dy \\
&\underset{\partial_t Y_t^i = U_t^i\left(Y_t^i\right)}{\overset{Y_t^i \text{ flows}}{=}} \int_{\mathbb{T}^2} 2 \left\langle\, Y_t^1 - Y_t^2 \;,\; U_t^1\left(Y_t^1\right) - U_t^2\left(Y_t^2\right) \,\right\rangle dy
\end{aligned}$$

$$= \int_{\mathbb{T}^2} 2 \left\langle \; Y_t^1 - Y_t^2 \;\; , \;\; U_t^1 \left(Y_t^1 \right) - U_t^1 \left(Y_t^2 \right) + U_t^1 \left(Y_t^2 \right) - U_t^2 \left(Y_t^2 \right) \; \right\rangle dy$$

by the linearity of the inner product followed by that of the integral, we have that:

$$\partial_t \int_{\mathbb{T}^2} ||Y_t^1 - Y_t^2||_2^2 \, dy =$$

$$= \int_{\mathbb{T}^2} 2 \left\langle \; Y_t^1 - Y_t^2 \; , \; U_t^1 \left(Y_t^1 \right) - U_t^1 \left(Y_t^2 \right) \; \right\rangle dy \; +$$
$$+ \int_{\mathbb{T}^2} 2 \left\langle \; Y_t^1 - Y_t^2 \; , \; U_t^1 \left(Y_t^2 \right) - U_t^2 \left(Y_t^2 \right) \; \right\rangle dy$$

Using the PropositionA.30 we respectively obtain the inequalities

$$\int_{\mathbb{T}^2} 2 \left\langle \; Y_t^1 - Y_t^2 \; , \; U_t^1 \left(Y_t^1 \right) - U_t^1 \left(Y_t^2 \right) \; \right\rangle dy \leq$$
$$\leq \int_{\mathbb{T}^2} ||Y_t^1 - Y_t^2||_2^2 \, dy + \int_{\mathbb{T}^2} \left\| U_t^1 \left(Y_t^1 \right) - U_t^1 \left(Y_t^2 \right) \right\|_2^2 dy$$

and

$$\int_{\mathbb{T}^2} 2 \left\langle \; Y_t^1 - Y_t^2 \; , \; U_t^1 \left(Y_t^2 \right) - U_t^2 \left(Y_t^2 \right) \; \right\rangle dy \leq$$
$$\leq \int_{\mathbb{T}^2} ||Y_t^1 - Y_t^2||_2^2 \, dy + \int_{\mathbb{T}^2} \left\| U_t^1 \left(Y_t^2 \right) - U_t^2 \left(Y_t^2 \right) \right\|_2^2 dy$$

Since U_t^1 is $||\nabla U_t^1||_{L^\infty(\mathbb{T}^2)}$-Lipschitz , because $||\nabla U_t^1||_{L^\infty(\mathbb{T}^2)} \leq C_1 + 1 =: C_3$, it holds that: $\forall x, z \in \mathbb{R}^2$

$$||U_t^1(x) - U_t^1(z)||_2 \leq ||\nabla U_t^1||_{L^\infty(\mathbb{T}^2)} \cdot ||x - z||_2$$

By choosing $x = Y_t^1(y)$ and $x = Y_t^2(y)$ $\forall \, t \in [0, T]$, we have that:

$$||U_t^1(Y_t^1) - U_t^1(Y_t^2)||_2 \leq ||\nabla U_t^1||_{L^\infty(\mathbb{T}^2)} \cdot ||Y_t^1 - Y_t^2||_2 \leq C_3 \cdot ||Y_t^1 - Y_t^2||_2$$

Since every norm is non-negative and the constant C_3 is positive we square the inequality to obtain:

$$||U_t^1(Y_t^1) - U_t^1(Y_t^2)||_2^2 \leq C_3^2 \cdot ||Y_t^1 - Y_t^2||_2^2$$

Integrating over the torus and combining with the respective inquality above, we have shown that:

$$\int_{\mathbb{T}^2} 2 \left\langle \; Y_t^1 - Y_t^2 \; , \; U_t^1 \left(Y_t^1 \right) - U_t^1 \left(Y_t^2 \right) \; \right\rangle dy \leq (1 + C_3^2) \cdot \int_{\mathbb{T}^2} ||Y_t^1 - Y_t^2||_2^2 \, dy$$

In order to apply the Gronwall lemma, we are left with estimating the $\int_{\mathbb{T}^2} \left\| U_t^1 \left(Y_t^2 \right) - U_t^2 \left(Y_t^2 \right) \right\|_2^2 dy$. This is the demanding part, where it will be needed to make several estimates through constructing interpolating curves.

Before we arrive there we can make an estimate without constructing anything yet.

Utilizing the measures ρ_0 , ρ_t^2 and its bounds to "get rid of" the flow Y_t^2 and then "replace" the velocities U_t^i to remain with $\int_{\mathbb{T}^2} \left\| \nabla P_t^{1,*} - \nabla P_t^{2,*} \right\|_2^2 dy$ to estimate. Indeed

Because $0 \leq \lambda \leq \rho_0 \Rightarrow 1 \leq \dfrac{\rho_0}{\lambda}$ and it also holds true that $0 < \rho_t^2 \leq \Lambda$. So

$$\int_{\mathbb{T}^2} \left\| U_t^1 \left(Y_t^2 \right) - U_t^2 \left(Y_t^2 \right) \right\|_2^2 dy \leq \int_{\mathbb{T}^2} \frac{\rho_0}{\lambda} \cdot \left\| U_t^1 \left(Y_t^2 \right) - U_t^2 \left(Y_t^2 \right) \right\|_2^2 dy =$$

$$= \frac{1}{\lambda} \int_{\mathbb{T}^2} \left\| U_t^1 \left(Y_t^2 \right) - U_t^2 \left(Y_t^2 \right) \right\|_2^2 d\rho_0 = \frac{1}{\lambda} \int_{\mathbb{T}^2} ||(U_t^1 - U_t^2) \circ Y_t^2||_2^2 \, d\rho_0 =$$

$$= \frac{1}{\lambda} \int_{\mathbb{T}^2} ||U_t^1 - U_t^2||_2^2 \circ Y_t^2 \, d\rho_0 \overset{\rho_t^2 = Y_t^2 \# \rho_0}{=} \frac{1}{\lambda} \int_{\mathbb{T}^2} ||U_t^1 - U_t^2||_2^2 \, d\rho_t^2 =$$

$$= \frac{1}{\lambda} \int_{\mathbb{T}^2} \rho_t^2 \cdot ||U_t^1 - U_t^2||_2^2 \, dy \underset{\Lambda > 0}{\overset{\rho_t^2 \leq \Lambda}{\leq}} \frac{\Lambda}{\lambda} \int_{\mathbb{T}^2} ||U_t^1 - U_t^2||_2^2 \, dy$$

And now we evaluate the quantity $||U_t^1 - U_t^2||_2^2$

$$||U_t^1 - U_t^2||_2 = \left\| \left(\nabla P_t^{1,*} - Id \right)^{\perp} - \left(\nabla P_t^{2,*} - Id \right)^{\perp} \right\|_2 =$$

$$= \left\| \left(\nabla P_t^{1,*} - Id - \left(\nabla P_t^{2,*} - Id \right) \right)^{\perp} \right\|_2 = \left\| \left(\nabla P_t^{1,*} - \nabla P_t^{2,*} \right)^{\perp} \right\|_2 =$$

$$= \left\| \nabla P_t^{1,*} - \nabla P_t^{2,*} \right\|_2 \Longrightarrow ||U_t^1 - U_t^2||_2^2 = \left\| \nabla P_t^{1,*} - \nabla P_t^{2,*} \right\|_2^2$$

Thus we have shown that:

$$\int_{\mathbb{T}^2} \left\| U_t^1 \left(Y_t^2 \right) - U_t^2 \left(Y_t^2 \right) \right\|_2^2 dy \leq \frac{\Lambda}{\lambda} \int_{\mathbb{T}^2} \left\| \nabla P_t^{1,*} - \nabla P_t^{2,*} \right\|_2^2 dy$$

To "finish" the Gronwall argument we would like to estimate above the integral $\int_{\mathbb{T}^2} \left\| \nabla P_t^{1,*} - \nabla P_t^{2,*} \right\|_2^2 dy$ by the integral $\int_{\mathbb{T}^2} ||Y_t^1 - Y_t^2||_2^2 \, dy$

To achieve this final goal, we now implement the aforementioned interpolation argument.

3.2.3 The interpolation argument

Before we actually construct the interpolating curves, let us present their properties which we would like to have. After listing the requirements that we want to cover, we "explain" the reasoning behind our thinking process. With those "assumptions" we argue to "show" that they indeed provide the result.

Remark. We don't actually prove that they give us the result, this will be done in the next paragraphs "Constructing the interpolating curves" and "Proving the properties".

The main idea is to construct interpolating curves ρ_t^θ,U_t^θ,P_t^θ for $\theta \in [1,2]$ in such a way that they will satisfy all the following:

$$\begin{cases} \partial_\theta \rho_t^\theta + \operatorname{div}(\rho_t^\theta U_t^\theta) = 0 & \text{(A)} \\ \\ \dfrac{1}{C_4} \leq \rho_t^\theta \leq C_4 \quad \text{and} \quad \|\rho_t^\theta\|_{C^{0,\alpha}(\mathbb{T}^2)} \leq C_4 & \text{(B)} \\ \\ \left|\left|D^2 P_t^{\theta,*}\right|\right|_{L^\infty(\mathbb{T}^2)} \quad , \quad \left|\left|\left(D^2 P_t^{\theta,*}\right)^{-1}\right|\right|_{L^\infty(\mathbb{T}^2)} \leq C_5 & \text{(C)} \\ \\ \int\limits_{\mathbb{T}^2} \rho_t^\theta \cdot \|U_t^\theta\|_2^2 \, dy = \int\limits_{\mathbb{T}^2} \rho_t^1 \cdot ||R_t - Id||_2^2 \, dy & \text{(D)} \\ \text{where } R_t \text{ is the optimal transport map sending } \rho_t^1 \text{ to } \rho_t^2 & \\ \\ \nabla P_t^{1,*} - \nabla P_t^{2,*} = \int\limits_1^2 \partial_\theta \nabla P_t^{\theta,*} \, d\theta & \text{(E)} \end{cases}$$

Because if we have all of the above in our hands , then (E) together with Holder's inequality imply that:

$$\int\limits_{\mathbb{T}^2} \left\|\nabla P_t^{1,*} - \nabla P_t^{2,*}\right\|_2^2 dy \leq \int\limits_1^2 \left\|\partial_\theta \nabla P_t^{\theta,*}\right\|_{L^2(\mathbb{T}^2)}^2 d\theta$$

In addition (A),(B),(C) imply that:

$$\left\|\partial_\theta \nabla P_t^{\theta,*}\right\|_{L^2(\mathbb{T}^2)}^2 d\theta \leq c \int\limits_{\mathbb{T}^2} \rho_t^\theta \cdot \left\|U_t^{\theta,*}\right\|_2^2 dy$$

At last equality (D) gives:

$$\int\limits_{\mathbb{T}^2} \rho_t^\theta \cdot \left\|U_t^{\theta,*}\right\|_2^2 dy = \int\limits_{\mathbb{T}^2} \rho_t^1 \cdot \|R_t - Id\|_2^2 \, dy$$

And since R_t is the optimal transport map sending ρ_t^1 to ρ_t^2, that is it minimizes the integral $\int_{\mathbb{T}^2} \|S(x) - x\|_2^2 \, d\rho_t^1(x)$ over all functions S such as $\rho_t^2 = S_{\#}\rho_t^1$, we can show that

$$\int_{\mathbb{T}^2} \rho_t^1 \cdot \|R_t - Id\|_2^2 \, dy \leq \Lambda \int_{\mathbb{T}^2} ||Y_t^1 - Y_t^2||_2^2 \, dy$$

The above quantities have no dependence on θ, hence integrating with respect to θ over the interval $[1, 2]$ we obtain the desired result i.e.

$$\int_{\mathbb{T}^2} \left\| \nabla P_t^{1,*} - \nabla P_t^{2,*} \right\|_2^2 dy \leq \Lambda C_6 \int_{\mathbb{T}^2} ||Y_t^1 - Y_t^2||_2^2 \, dy$$

Remark.
Benamou and Brenier in Chapter 3 at equation (32) of [6] note that the optimal choice of a flow $X(t, x)$ is given by $X(t, x) = x + \frac{t}{T}(\nabla\Psi(x) - x)$.

Inspired by this, in our effort to relate U_t^θ to $R_t - Id$ and since a solution to the measure continuity equation is obtained utilizing the flow, the definition of ρ_t^θ is quite logical to be the pushforward measure of ρ_t^1 with a similar (to Benamou and Brenier's aforementioned flow X) function.

Constructing the interpolating curves

We move on to construct the interpolating curves ρ_t^θ,U_t^θ,P_t^θ for $\theta \in [1, 2]$

Here, in this subsection we will define each curve and we will restrict ourselves to only "noticing" simple remarks about them. These remarks will be useful in the next subsection where we will prove the previously declared properties (A) (B) (C) (D) (E).

Since ρ_t^1 is the pushforward of the Lebesgue measure ($\rho_t^1 = \nabla P_t^1{}_{\#}dx$), it follows that ρ_t^1 is dominated by (absolutely continuous with respect to) $dx \equiv l^2$. Thus, ρ_t^1 has a non-negative density denoted also ρ_t^1 i.e. $\rho_t^1 = \rho_t^1 dx$

The density is positive almost everywhere, due to the fact that the measure ρ_t^1 satisfies the bound $0 < \lambda \leq \rho_t^1$ (by contradiction)

Also ρ_t^1 and ρ_t^2 are probability measures on the torus, hence they are both finite.

So, by Theorem2.1 we find a ρ_t^1−a.e. unique optimal transport map R_t sending ρ_t^1 onto ρ_t^2, which can be written as the gradient of an, up to additive constant, unique convex function P_t and satisfies the relations:

$$\rho_t^2 = R_{t\#}\rho_t^1 \text{ and } R_t = \nabla P_t$$

Remark.
Be cautious of t which for all calculations and quantities considered below is nothing more than a fixed "parameter". θ is considered the time variable thoughout the construction.

We now define for each t the curve of measures ρ_t^θ in $\mathbb{R}^2$ as the measure ρ_t^1 pushed by the function $y + (\theta - 1)(R_t(y) - y)$ i.e.

$$\rho_t^\theta := Id + (\theta - 1)(R_t - Id)_{\#}\rho_t^1$$

Trying to rewrite the "push"-function as the gradient of some other function we define for each t the curve of functions P_t^θ

$$P_t^\theta(y) := (2-\theta)\frac{\|y\|^2}{2} + (\theta - 1)P_t(y)$$

Now, it is easy to see that $\nabla P_t^\theta = Id + (\theta - 1)(R_t - Id)$, hence

$$\rho_t^\theta = \nabla P_t^\theta{}_{\#}\rho_t^1$$

Indeed, we will prove that $\nabla P_t^\theta(y) = y + (\theta - 1)(R_t(y) - y)$

$$P_t^\theta(y) = (2-\theta)\frac{\|y\|^2}{2} + (\theta - 1)P_t(y) \xrightarrow[=(\partial_1,\partial_2)]{\nabla=} \nabla P_t^\theta(y) =$$

$$= \left(\partial_1\Big((2-\theta)\frac{\|y\|^2}{2} + (\theta - 1)P_t(y)\Big) \;,\; \partial_2\Big((2-\theta)\frac{\|y\|^2}{2} + (\theta - 1)P_t(y)\Big)\right)$$

θ and t play the role of constants for the partial derivatives ∂_1,∂_2 with respect to the spatial variable. So, $\nabla P_t^\theta(y)$ equals

$$\left(\frac{(2-\theta)}{2}\partial_1\|y\|^2 + (\theta-1)\partial_1 P_t(y) \;,\; \frac{(2-\theta)}{2}\partial_2\|y\|^2 + (\theta-1)\partial_2 P_t(y)\right)$$

$y = (y_1, y_2) \Rightarrow \|y\|^2 = y_1^2 + y_2^2 \Rightarrow \partial_i\|y\|^2 = 2y_i$ for $i = 1, 2$. So, $\nabla P_t^\theta(y)$ equals

$$\Big((2-\theta)y_1 + (\theta-1)\partial_1 P_t(y) \;,\; (2-\theta)y_2 + (\theta-1)\partial_2 P_t(y)\Big)$$

$2 - \theta = 1 - (\theta - 1) \Rightarrow (2-\theta)y_i = y_i - (\theta - 1)y_i$ for $i = 1, 2$. So, $\nabla P_t^\theta(y) =$

$$= \Big(y_1 - (\theta-1)y_1 + (\theta-1)\partial_1 P_t(y) \;,\; y_2 - (\theta-1)y_2 + (\theta-1)\partial_2 P_t(y)\Big)$$

$$= \Big(y_1 + (\theta-1)\partial_1 P_t(y) - (\theta-1)y_1 \;,\; y_2 + (\theta-1)\partial_2 P_t(y) - (\theta-1)y_2\Big)$$

$$= (y_1\,, y_2) \;+\; (\theta-1)\Big(\partial_1 P_t(y), \partial_2 P_t(y)\Big) \;-\; (\theta-1)\big(y_1\,, y_2\big)$$

$$= (y_1\,, y_2) \;+\; (\theta-1)\nabla P_t \;-\; (\theta-1)\big(y_1\,, y_2\big)$$

$$= (y_1\,, y_2) \;+\; (\theta-1)R_t \;-\; (\theta-1)\big(y_1\,, y_2\big)$$

$$= (y_1\,, y_2) \;+\; (\theta-1)\Big(R_t - \big(y_1\,, y_2\big)\Big)$$
$$= y + (\theta-1)(R_t - y)$$

Having written ρ_t^θ as $\nabla P_t^\theta{}_{\#}\rho_t^1$ and since we want to obtain property (D) which relates U_t^θ to $R_t - Id$, it seems rational (due to the change of variables property of the pushforward measure) to finally define for each t the curve of velocities U_t^θ as:

$$U_t^\theta := (R_t - Id)\circ \nabla P_t^{\theta,*}$$

where $P_t^{\theta,*}$ is the Legendre transform of P_t^θ satisfying $\nabla P_t^{\theta,*}\big(\nabla P_t^\theta(y)\big) = y$

It remains now is to check that the constructed curves ρ_t^θ,P_t^θ,U_t^θ provide us indeed with the wanted properties (A)$-$(E). After this, we will finalize the proof concluding the Gronwall argument we have started earlier.

Before we do so, we summarize what we have defined/constructed so far in term of θ-curves:

$$\rho_t^\theta = \nabla P_t^\theta{}_{\#}\rho_t^1$$

$$P_t^\theta(y) = (2-\theta)\tfrac{\|y\|^2}{2} + (\theta-1)P_t(y)$$

$$\nabla P_t^\theta(y) = y + (\theta-1)(R_t(y) - y)$$

$$U_t^\theta := (R_t - Id)\circ \nabla P_t^{\theta,*}$$

Proving the properties

We begin by showing that property (A) holds.

To show that $\partial_\theta \rho_t^\theta + \mathrm{div}(\rho_t^\theta U_t^\theta) = 0$, it suffices to prove that for every $\varphi \in C_c^\infty(\mathbb{R}^2)$

$$\partial_\theta \int \varphi \, d\rho_t^\theta = \int \big\langle \nabla\varphi, U_t^\theta \big\rangle \, d\rho_t^\theta$$

Proof. Let $\varphi \in C_c^\infty(\mathbb{R}^2)$, then since $\rho_t^\theta = \nabla P_t^\theta{}_{\#} d\rho_t^1$ formula (change of variables through pushforward measure) of PropositionA.23 implies that:

$$\int \varphi \, d\rho_t^\theta = \int \varphi \circ \nabla P_t^\theta \, d\rho_t^1$$

Setting for each t the function $f(y,\theta) := \varphi\left(\nabla P_t^\theta(y)\right)$ for $(y,\theta) \in \mathbb{R}^2 \times [1,2]$ we ought to prove that it satisfies the conditions of PropositionA.31.

Indeed

$$\partial_\theta \int \varphi \, d\rho_t^\theta = \partial_\theta \int \varphi \circ \nabla P_t^\theta \, d\rho_t^1 = \int \partial_\theta \left(\varphi \circ \nabla P_t^\theta\right) \, d\rho_t^1$$

Using the chain rule we have that $\partial_\theta\left(\varphi\circ\nabla P_t^\theta\right)=\left(\nabla\varphi\circ\nabla P_t^\theta\right)\diamond\left(\partial_\theta\nabla P_t^\theta\right)$. So, $\partial_\theta\int\varphi d\rho_t^\theta$ equals

$$\int\left(\nabla\varphi\circ\nabla P_t^\theta\right)\diamond\left(\partial_\theta\nabla P_t^\theta\right)d\rho_t^1$$

Utilizing the fact that the gradient of pressure's interpolating curve P_t^θ and the gradient of its Legendre transformation are inverse functions i.e. $\nabla P_t^{\theta,*}\left(\nabla P_t^\theta(y)\right)=y$. We obtain that: $\partial_\theta\int\varphi\,d\rho_t^\theta=$

$$=\int\left(\nabla\varphi\circ\nabla P_t^\theta\right)\diamond\left(\partial_\theta\nabla P_t^\theta\circ\nabla P_t^{\theta,*}\circ\nabla P_t^\theta\right)d\rho_t^1$$

$$=\int\left(\nabla\varphi\diamond\left(\partial_\theta\nabla P_t^\theta\circ\nabla P_t^{\theta,*}\right)\right)\circ\nabla P_t^\theta\;d\rho_t^1$$

Since $\rho_t^\theta=\nabla P_t^\theta{}_{\#}\rho_t^1$, PropositionA.23 for the change of variables through the pushforward measure implies that: $\partial_\theta\int\varphi\,d\rho_t^\theta=$

$$=\int\nabla\varphi\diamond\left(\partial_\theta\nabla P_t^\theta\circ\nabla P_t^{\theta,*}\right)d\rho_t^\theta$$

Recalling that $\nabla P_t^\theta=Id+(\theta-1)(R_t-Id)$ differentiating with respect to theta (θ) we get: $\partial_\theta\nabla P_t^\theta=R_t-Id$.

Furthermore we have defined U_t^θ as $(R_t-Id)\circ\nabla P_t^{\theta,*}$, hence

$$\partial_\theta\int\varphi\,d\rho_t^\theta=\int\nabla\varphi\diamond U_t^\theta\;d\rho_t^\theta$$

$\varphi:\mathbb{R}^2\to\mathbb{R}\Rightarrow\nabla\varphi\in\mathbb{R}^{1\times2}$ and $U_t^\theta:\mathbb{R}^2\to\mathbb{R}^2$ where for a vector-valued function we identify $\mathbb{R}^2$ with $\mathbb{R}^{2\times1}$. Due to their dimensions we can rewrite the matrix product as the inner product, that is

$$\partial_\theta\int\varphi\,d\rho_t^\theta=\int\left\langle\nabla\varphi,U_t^\theta\right\rangle d\rho_t^\theta$$

Proving property (A) □

The proof of properties (B) and (C) can be found in paragraph 5.2.4 of [23]. Here, we will show the auxiliary property, which is needed for the proof

$$\det\left(D^2P_t^\theta\right)=\frac{\rho_t^1}{\rho_t^\theta\circ\nabla P_t^\theta}$$

Indeed

We denote the densities of the corresponding measures with the same notation but also putting a tilde above. Then, by PropositionA.25 we have:

$$\int \left(\varphi \circ \nabla P_t^\theta\right) \cdot \tilde{\rho}_t^1 \, dx = \int \varphi \circ \nabla P_t^\theta \, d\rho_t^1$$

Making use again of the equality $\rho_t^\theta = \nabla P_t^\theta{}_{\#}\rho_t^1$ and the change of variables for the pushforward measure PropositionA.23 we obtain:

$$\int \left(\varphi \circ \nabla P_t^\theta\right) \cdot \tilde{\rho}_t^1 \, dx = \int \varphi \, d\rho_t^\theta$$

Using PropositionA.25 one more time we are led to:

$$\int \left(\varphi \circ \nabla P_t^\theta\right) \cdot \tilde{\rho}_t^1 \, dx = \int \varphi \cdot \tilde{\rho}_t^\theta \, dy$$

Setting $y = \nabla P_t^\theta(x)$ a change of variables for the a.e. one-to-one (1-1) and continuously differentiable ∇P_t^θ gives:

$$\int \left(\varphi \circ \nabla P_t^\theta\right) \cdot \tilde{\rho}_t^1 \, dx = \int \left(\varphi \circ \nabla P_t^\theta\right) \cdot \left(\tilde{\rho}_t^\theta \circ \nabla P_t^\theta\right) \cdot \left|\det\left(D^2 P_t^\theta\right)\right| \, dx$$

$$\Rightarrow \int \left(\varphi \circ \nabla P_t^\theta\right) \cdot \left(\tilde{\rho}_t^1 - \left(\tilde{\rho}_t^\theta \circ \nabla P_t^\theta\right) \left|\det\left(D^2 P_t^\theta\right)\right|\right) \, dx = 0$$

PropositionA.32 implies that:

$$\tilde{\rho}_t^1 - \left(\tilde{\rho}_t^\theta \circ \nabla P_t^\theta\right) \left|\det\left(D^2 P_t^\theta\right)\right| = 0 \quad l^2 - \text{a.e.}$$

$$\Rightarrow \left(\tilde{\rho}_t^\theta \circ \nabla P_t^\theta\right) \left|\det\left(D^2 P_t^\theta\right)\right| = \tilde{\rho}_t^1$$

recalling that the measure ρ_t^θ is positive, hence so is its density, thus we have that:

$$\left|\det\left(D^2 P_t^\theta\right)\right| = \frac{\tilde{\rho}_t^1}{\tilde{\rho}_t^\theta \circ \nabla P_t^\theta}$$

recalling the definition of P_t^θ as $(2-\theta)\frac{\|y\|^2}{2} + (\theta - 1)P_t(y)$, since P_t and the squared norm $\|\cdot\|^2$ [due to PropositionA.35] are convex and $\dfrac{2-\theta}{2}, \theta - 1$ are non-negative PropositionA.33 implies that their linear combination i.e. P_t^θ is convex as well.

P_t^θ being convex it follows that its hesian is positive semi-definite, thus the determinant of its hesian is non-negative, that is $\det\left(D^2 P_t^\theta\right) \geq 0$. So,

$$\det\left(D^2 P_t^\theta\right) = \frac{\tilde{\rho}_t^1}{\tilde{\rho}_t^\theta \circ \nabla P_t^\theta}$$

Proving the auxiliary property. □

We then prove that property (D) is satisfied.

Proof. Utilizing once more the pushforward measure $\rho_t^\theta = \nabla P_t^\theta{}_{\#}\rho_t^1$ and the change of variables via the push forward function PropositionA.23 we have that:

$$\int \left\|U_t^\theta\right\|_2^2 d\rho_t^\theta = \int \left\|U_t^\theta\right\|_2^2 \circ \nabla P_t^\theta \, d\rho_t^1 = \int \left\|U_t^\theta \circ \nabla P_t^\theta\right\|_2^2 d\rho_t^1$$

Since $U_t^\theta = (R_t - Id) \circ \nabla P_t^{\theta,*}$ and $\nabla P_t^{\theta,*}, \nabla P_t^\theta$ are inverse functions we get:

$$\int \left\|U_t^\theta\right\|_2^2 d\rho_t^\theta = \int \|R_t - Id\|_2^2 \, d\rho_t^1$$

ρ_t^1 is absolutely continuous with respect to the Lebesgue measure (dy), thus using PropositionA.25 we can insert the densities into the integrals.

$$\int \tilde{\rho}_t^\theta \cdot \left\|U_t^\theta\right\|_2^2 dy = \int \tilde{\rho}_t^1 \cdot \|R_t - Id\|_2^2 \, dy$$

concluding this way the proof. □

Lastly, property (E) is an immediate application of the Fundamental Theorem of calculus.

Concluding the Gronwall and thus the proof

As we have discussed using the bounds of the iinterpolating curves, we deduce:

$$\int_{\mathbb{T}^2} \left\|\nabla P_t^{1,*} - \nabla P_t^{2,*}\right\|_2^2 dy \leq C_6 \Lambda \int_{\mathbb{T}^2} ||Y_t^1 - Y_t^2||_2^2 \, dy$$

Since,

$$\int_{\mathbb{T}^2} \left\|U_t^1\left(Y_t^2\right) - U_t^2\left(Y_t^2\right)\right\|_2^2 dy \leq \frac{\Lambda}{\lambda} \int_{\mathbb{T}^2} \left\|\nabla P_t^{1,*} - \nabla P_t^{2,*}\right\|_2^2 dy$$

we obtain that:

$$\int_{\mathbb{T}^2} \left\|U_t^1\left(Y_t^2\right) - U_t^2\left(Y_t^2\right)\right\|_2^2 dy \leq C_6 \Lambda \frac{\Lambda}{\lambda} \int_{\mathbb{T}^2} ||Y_t^1 - Y_t^2||_2^2 \, dy$$

which in turn leads to:

$$\partial_t \int_{\mathbb{T}^2} ||Y_t^1 - Y_t^2||_2^2 \, dy \leq \tilde{C} \int_{\mathbb{T}^2} ||Y_t^1 - Y_t^2||_2^2 \, dy$$

So, the Gronall PropositionA.40 implies:

$$\int_{\mathbb{T}^2} ||Y_t^1 - Y_t^2||_2^2 \, dy \leq e^{\tilde{C}t} \int_{\mathbb{T}^2} ||Y_0^1 - Y_0^2||_2^2 \, dy = 0$$

CHAPTER 4

Convergence of smooth solutions to the Euler equation

4.1 Preliminaries on the 2d Euler equation

Whatever is mentioned here, is taken (and can be found there in more detail) from Majda's and Bertozzi's book [31].

Here, we will briefly "discuss" some things that will help us have a better understanding of the Euler equation which we are going to use.

Before we proceed to the "depths" of the final chapter, that is, the convergence of smooth solutions to the Euler equation, we will make a short interlude to present a few things about the two-dimensional Euler equation.

Navier-Stokes and Euler

We start off noting the Navier-Stokes for an incompressible, homogenous fluid with constant viscosity ν and external force F_t.

We do so in both two and three dimensions.

$$\begin{cases} \partial_t u_t + \langle u_t, \nabla \rangle u_t = -\nabla p_t + \nu \Delta u_t + F_t & (x,t) \in \mathbb{R}^n \times [0, +\infty) \\ \text{div} u_t = 0 & (x,t) \in \mathbb{R}^n \times [0, +\infty) \end{cases} \tag{4.1.1}$$

with Δ being the Laplace operator

$$\Delta := \sum_{i=1}^{d} \partial_i^2$$

where we have abbreviated (like usual) the second partial derivatives, that is

$$\partial_{ij}^2 := \frac{\partial^2}{\partial x_i \partial x_j}$$

and when $i = j$ we simply write:

$$\partial_i^2 := \frac{\partial^2}{\partial x_i \partial x_i} = \frac{\partial^2}{\partial x_i^2}$$

We assume that there is no external force acting on our fluid i.e $F_t = 0$.

Thus, the Navier-Stokes now reads:

$$\begin{cases} \partial_t u_t + \langle u_t, \nabla \rangle u_t + \nabla p_t = \nu \Delta u_t & (x,t) \in \mathbb{R}^n \times [0, +\infty) \\ \text{div} u_t = 0 & (x,t) \in \mathbb{R}^n \times [0, +\infty) \end{cases} \tag{4.1.2}$$

Setting $\nu = 0$ (no viscosity) the Navier-Stokes reduces to the incompressible Euler equation:

$$\begin{cases} \partial_t u_t + \langle u_t, \nabla \rangle u_t + \nabla p_t = 0 & (x,t) \in \mathbb{R}^n \times [0, +\infty) \\ \text{div} u_t = 0 & (x,t) \in \mathbb{R}^n \times [0, +\infty) \end{cases} \tag{4.1.3}$$

In the following sections we will derive two equivalent formulations of the Navier-Stokes equation.

The first one will provide us with an equation involving only the velocity u_t (Leray's formulation).

The other will consist of an equation involving two quantities, the vorticity ω_t and a stream function ψ_t (vorticity-stream formulation).

4.1.1 Leray's formulation

Taking the divergence on both sides of the equation i.e. letting the operator to act on the function of each hand side, while also using the facts that

$$\text{div} u_t = 0$$

$$\text{div} \left(\langle u_t, \nabla \rangle u_t \right) = \text{tr}\Big((\nabla u_t)^2 \Big)$$

and that when $\text{div} u_t = 0$ we have

$$\text{div} \Delta u_t = 0$$

We can extract a Poisson equation for pressure p_t involving the velocity u_t

$$\Delta p_t = -\,\text{tr}\Big((\nabla u_t)^2 \Big)$$

Assuming that u_t is known we can solve this equation, leading us to the equivalent system (to that of Navier-Stokes) for $(x,t) \in \mathbb{R}^n \times [0,+\infty)$

$$\begin{cases} \partial_t u_t + \langle u_t, \nabla \rangle u_t + \int\limits_{\mathbb{R}^n} g(x-y)\,\mathrm{tr}\Big((\nabla u_t(y))^2 \Big)\,dy = \nu \Delta u_t \\ \mathrm{div} u_t = 0 \end{cases}$$

Since we have reformulated the problem in a form containing only the velocity field (the pressure can then be obtained by solving the above Poisson equation).

Although, we are technically done, there is another way to formulate the Navier-Stokes using the Leray projection $\mathbb{P}$

Proposition 4.1 (Helmholtz decomposition)**.**

Let $F \in L^2(\mathbb{R}^n : \mathbb{R}^n)$ then there exist a divergence free vector field w and a scalar potential h such that F can be written as the sum of w plus ∇h the gradient of the scalar potential i.e.

$$\exists\, w, h \;:\; F = w + \nabla h$$

with

$$\mathrm{div} w = 0$$

Definition 4.1 (Leray projection)**.**

We define the above w to be the Leray projection of F, this means that:

$$\mathbb{P}F := w$$

where w is given by the Helmholtz decomposition of F.

After some formal computations we can derive an equivalent Leray formulation of the Navier-Stokes:

$$\begin{cases} \partial_t u_t + \mathbb{P}\left(\langle u_t, \nabla \rangle u_t \right) = \nu \Delta u_t \\ \mathrm{div} u_t = 0 \end{cases}$$

Remark.
Both Leray formulations are equivalent to the Navier-Stokes equation.

Local in time reguralized solution to the Navier-Stokes

We firstly mollify (in a certain way) our equation, in order to show existence and uniqueness of local in time solution to the Navier-Stokes.
We define the mollification operator J_ε

$$J_\varepsilon(f) := \eta_\varepsilon * f$$

where η is a standard mollifier and the scaling of it $\eta_\varepsilon(x) = \frac{1}{\varepsilon^n}(\frac{x}{\varepsilon})$

Defining the rescaled velocity and pressure as well:

$$u_t^\varepsilon(x) := u_t(\frac{x}{\varepsilon})$$

$$p_t^\varepsilon(x) := p_t(\frac{x}{\varepsilon})$$

We are now ready to consider the mollified Navier-Stokes:

$$[NS_\varepsilon] \begin{cases} \partial_t u_t^\varepsilon + J_\varepsilon\left(\langle J_\varepsilon u_t^\varepsilon, \nabla\rangle J_\varepsilon u_t^\varepsilon\right) + \nabla p_t^\varepsilon = \nu J_\varepsilon(\Delta J_\varepsilon u_t^\varepsilon) \\ \text{div} u_t^\varepsilon = 0 \end{cases} \tag{4.1.4}$$

Projecting on the space of divergence free functions, using the Leray projection, we get (omitting the incompressibility condition):

$$[L - NS_\varepsilon] \quad \partial_t u_t^\varepsilon + \mathbb{P}\big(J_\varepsilon\left(\langle J_\varepsilon u_t^\varepsilon, \nabla\rangle J_\varepsilon u_t^\varepsilon\right)\big) = \nu J_\varepsilon(\Delta J_\varepsilon u_t^\varepsilon) \tag{4.1.5}$$

By defining the operator:

$$F_\varepsilon(x) := \nu J_\varepsilon(\Delta J_\varepsilon f) - \mathbb{P}\big(J_\varepsilon\left(\langle J_\varepsilon f, \nabla\rangle J_\varepsilon f\right)\big)$$

The $L - NS_\varepsilon$ (4.1.5) becomes:

$$\partial_t u_t^\varepsilon = F_\varepsilon(u_t^\varepsilon)$$

Proposition 4.2 (autonomous ODE system in Banach space)**.**

Let $\mathbb{B}$ be a Banach space. Let $F : \mathbb{B} \to \mathbb{B}$ be a locally Lipschitz map. Let also $H : \mathbb{B} \times [0, +\infty) \to \mathbb{B}$ be a locally Lipschitz map, then for the autonomous system (initial value problem)

$$\begin{cases} \partial_t H_t = F(H_t) \\ H_0 = G \end{cases}$$

there exists a time $T > 0$ and a unique map $H \in C^1\big([0, T], \mathbb{B}\big)$ satisfying the above (i.e. it is a solution of the aforementinted autonomous equation)

We denote V^m the space consisting of the functions belonging in the Sobolev space $W^{m,2}$ with (weak) divergence being equal to zero.

recalling that we call the Hilbert space $W^{m,2}$ as H^m, we have that V^m is the space having the divergence free functions of H^m.

Note that V^m as a closed subset of a Sobolev space (Sobolev spaces are Banach spaces) is also a Banach space itself.

Due to its definition F_ε has no dependence on time.

It can be shown that $F_\varepsilon : V^m \to V^m$ and also that F_ε is locally Lipschitz.

Hence, forall $\varepsilon > 0$ there exists a unique, local in time, smooth solution u_t^ε to the mollified Navier-Stokes.

We call such a solution, a reguralized solution.

Local in time solution to the Navier-Stokes

Taking the limit as $\varepsilon \to 0^+$, it has been proved [31] that we can obtain a solution to the Navier-Stokes (not the mollified one) equation such that it belongs in $C\big([0,T], V^m\big) \cap C^1\big([0,T], V^{m-2}\big)$

Global in time smooth solution for the 2d incompressible Euler

Setting $\nu = 0$ and restricting ourselves to the two (spatial) dimensions $n = 2$, we have the following result (see [31]):

Using the Beale-Kato-Majda criterion we can expand the previous local in time, smooth solution into a global in time, smooth solution for the 2d incompressible Euler.

4.1.2 Vorticity-stream formulation

One more useful formulation of the Navier-Stokes equation is the vorticity-stream formulation. We manage to “get rid of” the velocity u_t.

Here, we will mention results for the 2d incompressible Euler only.

We define the vorticity:

$$\omega_t := \text{curl} u_t$$

which in two dimensions is a scalar field (a real-valued, multivariable though function).

Taking the curl on the 2d incompressible Euler equation i.e. letting the operator to act on both sides, we get:

$$\partial_t \omega_t + \langle u_t, \nabla \rangle \omega_t = 0$$

But, since this equation still has the velocity, we have not finished yet.

Due to the fact that the vorticity in two dimensions is a scalar field, we compute to make our equation simpler:

$$\begin{aligned}\langle u_t, \nabla \rangle \omega_t &= \sum_{i=1}^{2} u_t^i \partial_i \omega_t \\ &= \langle u_t, \nabla \omega_t \rangle\end{aligned}$$

Thus, we have:

$$\partial_t \omega_t + \langle u_t, \nabla \omega_t \rangle = 0$$

We will make use of the following fact:

Proposition 4.3.
A conservative vector field can be written as the gradient of a scalar field.

We assert that $-u_t^{\perp} = (-u_t^2, u_t^1) = -u_t^2 e_1 + u_t^1 e_2$ is conservative.

Indeed,

Utilizing the Gauss-Green theorem we have:

$$\begin{aligned}\oint_c - u_t^2 e_1 + u_t^1 e_2 \, dl &= \iint_D \partial_1 u_t^1 - \partial_2(-u_t^2) \, dx \, dy \\ &= \iint_D \partial_1 u_t^1 + \partial_2 u_t^2 \, dx \, dy \\ &= \iint_D \text{div} u_t \, dx \, dy \\ &= 0\end{aligned}$$

Thus,

$$\oint_c u_t^{\perp} \, dl = 0$$

Hence, there exists a scalar field ψ_t, which we will call stream, such that

$$-u_t^{\perp} = \nabla\psi_t$$

Since $\left(f^{\perp}\right)^{\perp} = -f$ we get:

$$u_t = \nabla^{\perp}\psi_t$$

Substituting this into $\partial_t\omega_t + \langle u_t, \nabla\omega_t\rangle = 0$ we get:

$$\partial_t\omega_t + \langle\nabla^{\perp}\psi_t, \nabla\rangle\omega_t = 0$$

Since $u_t = \nabla^{\perp}\psi_t$ is divergence free, it is true that:

$$\langle\nabla^{\perp}\psi_t, \nabla\omega_t\rangle = \mathrm{div}(\omega_t\nabla^{\perp}\psi_t)$$

Also, since $u_t = \nabla^{\perp}\psi_t$ and $\omega_t = \mathrm{curl}u_t$ we get:

$$\omega_t = \Delta\psi_t$$

So, we have obtained the 2d incompressible Euler equation in vorticity-stream formulation:

$$\begin{cases} \partial_t\omega_t + \mathrm{div}(\omega_t\nabla^{\perp}\psi_t) = 0 \\ \omega_t = \Delta\psi_t \end{cases}$$

Remark.
The incompressibility condition $\mathrm{div}(\nabla^{\perp}\psi_t) = 0$ holds true, because we have shown that Proposition1.1 the rotated gradient of a scalar field is divergence free.

4.2 The dual SG equations as a coupled system of continuity and Monge-Ampère equation

In order to "see" that the dual SG equation "looks like" the Euler equation, we have to reformulate it.

We begin by rewriting the equation $\rho_t = \nabla P_{t\#}dx$ of the dual SG system, in its more standard counterpart using the Monge-Ampère equation.

Proposition 4.4 (Formal passage from the pushforward equation to the Monge-Ampère equation)**.**

Let the measure ρ_t satisfy the pushforward equation

$$\rho_t = \nabla P_{t\#} dx$$

where P_t is C^2 then its density ρ_t (denoted by the same symbol) satisfies the Monge-Ampère equation

$$\rho_t = \det\big(D^2 P_t^*\big)$$

Proof.

Let $\varphi \in C_c^\infty(\mathbb{R}^2)$ (thus measurable). Since ∇P_t is continuous, it is also $\big(\mathcal{B}(\mathbb{R}^2), \mathcal{B}(\mathbb{R}^2)\big)$-measurable.

The push forward change of variables thus implies:

$$\int_{\mathbb{R}^2} \varphi \, d\rho_t = \int_{\mathbb{R}^2} \varphi \circ \nabla P_t \, dy$$

We perform one more change of variables setting

$$y = \nabla P_t(x)$$

Due to the fact that for all times t the functions $\nabla P_t, \nabla P_t^*$ are inverse to each other, we get that

$$x = \nabla P_t^*(y)$$

So, we have:

$$\begin{aligned}\int_{\mathbb{R}^2} \varphi \circ \nabla P_t \, dy &= \int_{\mathbb{R}^2} \varphi \big|\det\big(\nabla(\nabla P_t^*)\big)\big| \, dx \\ &= \int_{\mathbb{R}^2} \varphi \big|\det\big(D^2 P_t^*\big)\big| \, dx\end{aligned}$$

Since the measure ρ_t has density it holds true that:

$$\int_{\mathbb{R}^2} \varphi \, d\rho_t = \int_{\mathbb{R}^2} \varphi \rho_t \, dx$$

Thus, we aggregately get:

$$\int_{\mathbb{R}^2} \varphi \rho_t \, dx = \int_{\mathbb{R}^2} \varphi \big|\det\big(D^2 P_t^*\big)\big| \, dx$$

$$\Rightarrow \int_{\mathbb{R}^2} \varphi\big(\rho_t - \big|\det\big(D^2 P_t^*\big)\big|\big) = 0$$

Using PropositionA.32 and/or the arbitrariness of φ we deduce that:

$$\rho_t - \big|\det\big(D^2 P_t^*\big)\big| = 0 \text{ for a.e. } x \in \mathbb{R}^2$$

$$\rho_t = \big|\det\big(D^2 P_t^*\big)\big|$$

At any point, the hessian of any convex real-valued, multivariable function (i.e. scalar field) is positive semi-definite

The result now follows immediately. □

We can prove that the reverse direction is also true, that is, one can pass from Monge-Ampère equation to the pushforward equation.

Hence, both formulations (pushforward and Monge-Ampère) are considered equivalent.

Proposition 4.5 (Formal passage from the Monge-Ampère equation to the pushforward equation)**.**

Let the density ρ_t of the measure (denoted by the same symbol) ρ_t satisfy the Monge-Ampère equation

$$\rho_t = \det\big(D^2 P_t^*\big)$$

then the measure ρ_t satisfies the pushforward equation

$$\rho_t = \nabla P_{t\#} dx$$

Proof.

We define the auxiliary measure:

$$\sigma_t := \nabla P_{t\#} dx$$

Let $\varphi \in C_c^\infty(\mathbb{R}^2)$ (thus measurable). Since ∇P_t is continuous, it is also $\big(\mathcal{B}(\mathbb{R}^2), \mathcal{B}(\mathbb{R}^2)\big)$-measurable.

The push forward change of variables then implies:

$$\int_{\mathbb{R}^2} \varphi \, d\sigma_t = \int_{\mathbb{R}^2} \varphi \circ \nabla P_t \, dy$$

Since σ_t is absolutely continuous with respect to the Lebesgue measure TheoremA.1 provides us with a density, which we denote with the same notation σ_t and satisfies:

$$\int_{\mathbb{R}^2} \varphi \sigma_t \, dx = \int_{\mathbb{R}^2} \varphi \, d\sigma_t$$

$$= \int_{\mathbb{R}^2} \varphi \circ \nabla P_t \, dy$$

Setting x as $\nabla P_t(y)$, due to the fact that for all times t the functions $\nabla P_t, \nabla P_t^*$ are inverses, we get that $x = \nabla P_t^*(y)$

So, we have that:

$$\begin{aligned} \int_{\mathbb{R}^2} \varphi \circ \nabla P_t \, dy &= \int_{\mathbb{R}^2} \varphi \big| \det\big(\nabla(\nabla P_t^*)\big) \big| \, dx \\ &= \int_{\mathbb{R}^2} \varphi \big| \det\big(D^2 P_t^*\big) \big| \, dx \\ &= \int_{\mathbb{R}^2} \varphi \rho_t \, dx \end{aligned}$$

Hence, we deduce that for all $\varphi \in C_c^\infty(\mathbb{R}^2)$

$$\begin{aligned} \int_{\mathbb{R}^2} \varphi \sigma_t \, dx &= \int_{\mathbb{R}^2} \varphi \rho_t \, dx \\ \Rightarrow \int_{\mathbb{R}^2} \varphi(\sigma_t - \rho_t) \, dx &= 0 \end{aligned}$$

The arbitrariness of φ implies that

$$\sigma_t = \rho_t$$

q.e.d. □

So, we have shown that:

$$\begin{cases} \partial_t \rho_t + \operatorname{div}(\rho_t U_t) = 0 \\ U_t = (\nabla P_t^* - Id)^\perp \\ \rho_t = \nabla P_{t\#} dx \end{cases} \Leftrightarrow \begin{cases} \partial_t \rho_t + \operatorname{div}(\rho_t U_t) = 0 \\ U_t = (\nabla P_t^* - Id)^\perp \\ \rho_t = \det\big(D^2 P_t^*\big) \end{cases}$$

With this, we can rewrite the dual SG system, bringing it down to two "tightly-packed" equations.

Inserting the Monge-Ampère equation and substituing the velocity U_t in the continuity equation, the dual SG system

$$\begin{cases} \partial_t \rho_t + \operatorname{div}(\rho_t U_t) = 0 \\ U_t = (\nabla P_t^* - Id)^\perp \\ \rho_t = \nabla P_{t\#} dx \end{cases}$$

becomes

$$\begin{cases} \partial_t \rho_t + \operatorname{div}\big(\rho_t \, (\nabla P_t^* - Id)^{\perp}\big) = 0 \\ \rho_t = \det\big(D^2 P_t^*\big) \end{cases}$$

We define the scalar field:

$$q_t := P_t^* - \frac{\|x\|^2}{2}$$

Thus we can rewrite:

$$\nabla P_t^* - Id \text{ as } \nabla q_t$$

and

$$D^2 P_t^* \text{ as } D^2 q_t + I_2$$

Substituting these as well, we obtain:

$$\begin{cases} \partial_t \rho_t + \operatorname{div}(\rho_t \nabla^{\perp} q_t) = 0 \\ \rho_t = \det\big(D^2 q_t + I_2\big) \end{cases}$$

Remark.
Note that this system above is equivalent to the dual SG system.

The reason for this is that ρ_t remained unchanged. Also, if we have a solution P_t^* of the dual SG equation then we can obtain a solution q_t of the above system, and vice versa, if we have a solution q_t of the above system then we can obtain a solution P_t^* of the dual SG system, just by setting $q_t := P_t^* - \frac{\|x\|^2}{2}$ and $P_t^* = q_t + \frac{\|x\|^2}{2}$ respectively

i.e. we have

$$\begin{cases} \partial_t \rho_t + \operatorname{div}(\rho_t U_t) = 0 \\ U_t = (\nabla P_t^* - Id)^{\perp} \\ \rho_t = \nabla P_{t\#} dx \end{cases} \Leftrightarrow \begin{cases} \partial_t \rho_t + \operatorname{div}(\rho_t \nabla^{\perp} q_t) = 0 \\ \rho_t = \det\big(D^2 q_t + I_2\big) \end{cases}$$

The dual SG system now looks pretty similar to the two dimensional Euler equation in vorticity-stream formulation:

$$\begin{cases} \partial_t \omega_t + \operatorname{div}(\omega_t \nabla^{\perp} \psi_t) = 0 \\ \omega_t = \Delta \psi_t \end{cases}$$

with ρ_t to be the analogous of ω_t, even though the first one is density in a dual space.

The obvious difference between them is that instead of a Poisson (for ψ_t) coupled with the continuity equation, we have to deal with a Monge-Ampère equation (for its analogous $q_t + \frac{\|x\|^2}{2}$) coupled with the continuity equation.

However, we can linearize the Monge-Ampère equation (for $q_t + \frac{\|x\|^2}{2}$) and make her "look like" a Poisson equation (for q_t).

The reason for this is that near identity the determinant behaves like the trace.

Furthermore, the trace of the hessian equals the Laplace operator.

Proposition 4.6 (trace of hessian equals Laplace operator)**.**

$$\operatorname{tr}\big(D^2\big) = \Delta$$

Proof.

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a C^2 function

Since f be a real-valued function we get:

$$Df = (\partial_1 f, \partial_2 f, \dots, \partial_n f)$$

and differentiating one more time, we have:

$$D^2 f = \begin{pmatrix} \partial_1^2 f & \partial_{12} f & \cdots & \partial_{1n} f \\ \partial_{21} f & \partial_2^2 f & \cdots & \partial_{2n} f \\ \vdots & \vdots & \ddots & \vdots \\ \partial_{n1} f & \partial_{n2} f & \cdots & \partial_n^2 f \end{pmatrix}$$

Hence,

$$\operatorname{tr}\big(D^2\big) = \sum_{i=1}^{n} \partial_i^2 = \Delta$$

the arbitrariness of the function proves the desired □

Now it is left to show that $\det\big(D^2 q_t + I_2\big)$ is close to $1 + \Delta q_t$

Motivation

Proposition 4.7 (near identity the determinant behaves like the trace)**.**

Let $A \in \mathbb{R}^{2\times 2}$ be a symmetric matrix then

$$\det(I + \varepsilon A) = 1 + \varepsilon \operatorname{tr}(A) + \varepsilon^2 \det(A)$$

and

$$\det(I + \varepsilon A) = 1 + \varepsilon \operatorname{tr}(A) + O(\varepsilon^2) \text{ for } \varepsilon \to 0^+$$

Proof.

Because A is a real symmetric matrix, there exists an orthonormal basis consisting of eigenvectors v_i for $i = 1, 2$.

Thus, there are 2 distinct eigenvalues λ_i for $i = 1, 2$ corresponding to the respective eigenvectors v_i.

And there exists and orthogonal matrix P ($P^T P = PP^T = I_2$) such as:

$$P^T AP = D$$

where D is the diagonal matrix consisting of the eigenvalues.

It holds true that:

$$tr(A) = \mathrm{tr}\big(P^{-1}AP\big) = \mathrm{tr}\big(P^T AP\big) = \sum_{i=1}^{2} \lambda_i = \lambda_1 + \lambda_2$$

since the inverse of an orthonormal matrix is its transpose.

and

$$\det(A) = \det\big(P^T AP\big) = \prod_{i=1}^{2} \lambda_i = \lambda_1 \lambda_2$$

due to the multiplicativity of determinant.

It is also true that the eigenvalues of εA are $\varepsilon\lambda_i$ for $i = 1, 2$

The characteristic polynomial of εA then reads:

$$\begin{aligned}\det(\varepsilon A - sI_2) &= (s - \varepsilon\lambda_1)(s - \varepsilon\lambda_2)\\ &= s^2 - \varepsilon(\lambda_1 + \lambda_2)s + \varepsilon^2\lambda_1\lambda_2\\ &= s^2 - \varepsilon\,\mathrm{tr}(A)s + \varepsilon^2 \det(A)\end{aligned}$$

Setting $s = -1$ we get:

$$\det(I_2 + \varepsilon A) = 1 + \varepsilon\,\mathrm{tr}(A) + \varepsilon^2 \det(A)$$

Let $\varepsilon_0 \in \mathbb{R}$

Since

$$\frac{\big|\varepsilon^2 \det(A)\big|}{\varepsilon^2} = |\det(A)|$$

Assuming that $\det(A) \neq 0$ we deduce that:

$$\exists\, M := |\det(A)| > 0 \quad \forall\varepsilon \in \mathbb{R} \ \ : \ \ \frac{\big|\varepsilon^2 \det(A)\big|}{\varepsilon^2} \leq M$$

Let $\zeta > 0$, the above implies the following three:

$$\exists\, M > 0 \quad \exists\,\zeta > 0 \quad \forall\,\varepsilon \in (\varepsilon_0 - \zeta, \varepsilon_0 + \zeta) \quad \frac{\big|\varepsilon^2 \det(A)\big|}{\varepsilon^2} \leq M$$

i.e.

$$\varepsilon^2 \det(A) = O(\varepsilon^2) \text{ for } \varepsilon \to 0^+$$

If $\det(A) = 0$, then all the previous work still holds true for every $M > 0$. □

For $\varepsilon \to 0^+$, $\varepsilon^2 \to 0$. Hence, we can say that $O(\varepsilon^2) \simeq 0$

So, Proposition4.7 for the symmetric $D^2 q_t$ gives:

$$\det\big(I_2 + \varepsilon D^2 q_t\big) = 1 + \varepsilon \operatorname{tr}\big(D^2 q_t\big) + O(\varepsilon^2) \text{ for } \varepsilon \to 0^+$$

that is

$$\det\big(I_2 + \varepsilon D^2 q_t\big) \simeq 1 + \varepsilon \Delta q_t$$

leading to

$$\det\big(I_2 + \varepsilon D^2 q_t\big) \text{ being close to } 1 + \varepsilon \Delta q_t \text{ for small enough } \varepsilon$$

We bring back to our minds that $\rho_t = \det\big(D^2 q_t + I_2\big)$.

So, if ρ_t is close to 1, one would expect q_t to be small. In turn, $\det\big(D^2 q_t\big)$ would be small.

From Schwarz's theorem for mixed partial derivatives, $D^2 q_t$ is symmetric and using again Proposition4.7 we get:

$$\det\big(I + D^2 q_t\big) = 1 + \operatorname{tr}\big(D^2 q_t\big) + \det\big(D^2 q_t\big)$$

Since we expect $\det\big(D^2 q_t\big)$ to be small, the above equality can be considered as:

$$\det\big(I + D^2 q_t\big) = 1 + \operatorname{tr}\big(D^2 q_t\big) + O\big(\det\big(D^2 q_t\big)\big)$$
$$\text{with } O\big(\det\big(D^2 q_t\big)\big) \simeq 0$$

Therefore, we have:

$$\rho_t = \det\big(D^2 q_t + I_2\big) = \det\big(I_2 + D^2 q_t\big) \simeq 1 + \Delta q_t$$

That is

$$\rho_t - 1 \simeq \Delta q_t$$

where $\rho_t - 1$ also satsfies:

$$\begin{aligned}\partial_t(\rho_t - 1) + \operatorname{div}\big((\rho_t - 1)\nabla^{\perp} q_t\big) &= \partial_t \rho_t + \operatorname{div}(\rho_t \nabla^{\perp} q_t - \nabla^{\perp} q_t) \\ &= \partial_t \rho_t + \operatorname{div}(\rho_t \nabla^{\perp} q_t) - \operatorname{div}(\nabla^{\perp} q_t)\end{aligned}$$

$$
\begin{aligned}
&= \partial_t \rho_t + \operatorname{div}(\rho_t \nabla^{\perp} q_t) \\
&= 0
\end{aligned}
$$

because the differential operator div is linear, the gradient of a rotated vector field is divergence free Proposition1.1 and ρ_t satisfies the dual SG system.

In other words, if we assume to have initial data ρ_0 which is close to 1 (meaning that $\rho_0 - 1$ is small)

Then someone expects that for a solution ρ_t, q_t of the dual SG system, the quantities $\rho_t - 1, q_t$ would stay close to a solution ω_t, ψ_t (respectively) of the incompressible Euler equation in vorticity-stream formulation:

$$
\begin{cases}
\partial_t \omega_t + \operatorname{div}(\omega_t \nabla^{\perp} \psi_t) = 0 \\
\omega_t = \Delta \psi_t
\end{cases}
$$

To take advantage of the aforementioned information, in order to truly show that a solution $\rho_t - 1, q_t$ of the dual SG system converges (under some norm) to a solution ω_t, ψ_t (respectively) of the incompressible Euler equation in vorticity-stream formulation, we rescale.

For $\varepsilon > 0$ we multiply with $\dfrac{1}{\varepsilon}$ and we rescale in time setting t as $\dfrac{t}{\varepsilon}$

4.3 $\mathrm{SG}_{\varepsilon}$ rescaling the dual SG system

Let ρ_t, q_t be a solution of the dual SG system. Let $\varepsilon > 0$ as well.

We define:

$$
\begin{aligned}
\rho_t^{\varepsilon} &:= \frac{1}{\varepsilon}(\rho_{t/\varepsilon} - 1) \\
q_t^{\varepsilon} &:= \frac{1}{\varepsilon} q_{t/\varepsilon}
\end{aligned}
$$

We compute to describe the equations above in terms of ρ_t and q_t respectively

$$
\begin{aligned}
& \rho_t^{\varepsilon} = \frac{1}{\varepsilon}(\rho_{t/\varepsilon} - 1) \\
\Rightarrow \quad & \varepsilon \rho_t^{\varepsilon} + 1 = \rho_{t/\varepsilon} \\
\Rightarrow \quad & \varepsilon \rho_{\varepsilon t}^{\varepsilon} + 1 = \rho_t
\end{aligned}
$$

where in the first step we multiplied the equation by ε and in the second step we set t as εt

Similarly,

$$
q_t^{\varepsilon} = \frac{1}{\varepsilon} q_{t/\varepsilon}
$$

$$\begin{aligned}\Rightarrow \quad & \varepsilon q^{\varepsilon}_{t} = q_{t/\varepsilon} \\ \Rightarrow \quad & \varepsilon q^{\varepsilon}_{\varepsilon t} = q_{t}\end{aligned}$$

Since ρ_t, q_t satisfy the dual SG system, we have that:

$$\begin{cases}\partial_t \rho_t + \operatorname{div}(\rho_t \nabla^{\perp} q_t) = 0 \\ \rho_t = \det\big(D^2 q_t + I_2\big)\end{cases}$$

Inserting $\rho_t = \varepsilon \rho^{\varepsilon}_{\varepsilon t} + 1$ and $q_t = \varepsilon q^{\varepsilon}_{\varepsilon t}$ we get:

$$\begin{cases}\partial_t(\varepsilon \rho^{\varepsilon}_{\varepsilon t} + 1) + \operatorname{div}\Big((\varepsilon \rho^{\varepsilon}_{\varepsilon t} + 1)\, \nabla^{\perp}(\varepsilon q^{\varepsilon}_{\varepsilon t})\Big) = 0 \\ \varepsilon \rho^{\varepsilon}_{\varepsilon t} + 1 = \det\big(D^2(\varepsilon q^{\varepsilon}_{\varepsilon t}) + I_2\big)\end{cases}$$

We calculate each quantity seperately:

$$\begin{aligned}\partial_t(\varepsilon \rho^{\varepsilon}_{\varepsilon t} + 1) &= \varepsilon \partial_t(\rho^{\varepsilon}_{\varepsilon t}) \\ &= \varepsilon \partial_t \rho^{\varepsilon}_{\varepsilon t} \cdot \partial_t(\varepsilon t) \\ &= \varepsilon \partial_t \rho^{\varepsilon}_{\varepsilon t} \cdot \varepsilon \\ &= \varepsilon^2 \partial_t \rho^{\varepsilon}_{\varepsilon t}\end{aligned}$$

Due to the convention that $\nabla^{\perp}$ refers to differentiation with respect to the space variable only, the function $q^{\varepsilon}_{\varepsilon t}$ is not a composition and its derivative can be computed directly (at time εt)

Also, we know that the differential operator div is linear.

Hence,

$$\begin{aligned}\operatorname{div}\Big((\varepsilon \rho^{\varepsilon}_{\varepsilon t} + 1)\, \nabla^{\perp}(\varepsilon q^{\varepsilon}_{\varepsilon t})\Big) &= \operatorname{div}\Big((\varepsilon \rho^{\varepsilon}_{\varepsilon t} + 1)\, \varepsilon \nabla^{\perp} q^{\varepsilon}_{\varepsilon t}\Big) \\ &= \operatorname{div}\big(\varepsilon^2 \rho^{\varepsilon}_{\varepsilon t}\, \nabla^{\perp} q^{\varepsilon}_{\varepsilon t} + \varepsilon \nabla^{\perp} q^{\varepsilon}_{\varepsilon t}\big) \\ &= \varepsilon^2 \operatorname{div}(\rho^{\varepsilon}_{\varepsilon t}\, \nabla^{\perp} q^{\varepsilon}_{\varepsilon t}) + \varepsilon \operatorname{div}(\nabla^{\perp} q^{\varepsilon}_{\varepsilon t}) \\ &= \varepsilon^2 \operatorname{div}(\rho^{\varepsilon}_{\varepsilon t}\, \nabla^{\perp} q^{\varepsilon}_{\varepsilon t})\end{aligned}$$

because the rotated gradient of a real-valued function (namely $q^{\varepsilon}_{\varepsilon t}$) is divergence free Proposition1.1

Combining them with the fact that $D^2(\varepsilon q^{\varepsilon}_{\varepsilon t}) = \varepsilon D^2 q^{\varepsilon}_{\varepsilon t}$ we have that:

$$\begin{cases}\varepsilon^2 \partial_t \rho^{\varepsilon}_{\varepsilon t} + \varepsilon^2 \operatorname{div}(\rho^{\varepsilon}_{\varepsilon t}\, \nabla^{\perp} q^{\varepsilon}_{\varepsilon t}) = 0 \\ \varepsilon \rho^{\varepsilon}_{\varepsilon t} + 1 = \det\big(\varepsilon D^2 q^{\varepsilon}_{\varepsilon t} + I_2\big)\end{cases}$$

With ε being positive, we divide with $\varepsilon^2 \neq 0$ to get:

$$\begin{cases}\partial_t \rho^{\varepsilon}_{\varepsilon t} + \operatorname{div}(\rho^{\varepsilon}_{\varepsilon t}\, \nabla^{\perp} q^{\varepsilon}_{\varepsilon t}) = 0 \\ \varepsilon \rho^{\varepsilon}_{\varepsilon t} + 1 = \det\big(\varepsilon D^2 q^{\varepsilon}_{\varepsilon t} + I_2\big)\end{cases}$$

"Scaling back" we set εt as t to obtain the following system for $\rho_t^\varepsilon, q_t^\varepsilon$

$$[\text{SG}_\varepsilon \text{ system}] \begin{cases} \partial_t \rho_t^\varepsilon + \operatorname{div}(\rho_t^\varepsilon \, \nabla^\perp q_t^\varepsilon) = 0 \\ \varepsilon \rho_t^\varepsilon + 1 = \det\big(\varepsilon D^2 q_t^\varepsilon + I_2\big) \end{cases} \tag{4.3.1}$$

which is precisely what we will call the SG_ε system from now on.

Remark.
The dual SG system is equivalent to the SG_ε system

The direction dual SG system $\Rightarrow$ SG_ε system has just been shown. We now prove the reverse:

Indeed, if we have a solution $\rho_t^\varepsilon, q_t^\varepsilon$ of the SG_ε system

$$\begin{cases} \partial_t \rho_t^\varepsilon + \operatorname{div}(\rho_t^\varepsilon \, \nabla^\perp q_t^\varepsilon) = 0 \\ \varepsilon \rho_t^\varepsilon + 1 = \det\big(\varepsilon D^2 q_t^\varepsilon + I_2\big) \end{cases}$$

then we can follow the process above (to derive SG_ε system) backwards (exactly as done earlier) to obtain:

$$\begin{cases} \partial_t (\varepsilon \rho_{\varepsilon t}^\varepsilon + 1) + \operatorname{div}\Big((\varepsilon \rho_{\varepsilon t}^\varepsilon + 1) \, \nabla^\perp (\varepsilon q_{\varepsilon t}^\varepsilon)\Big) = 0 \\ \varepsilon \rho_{\varepsilon t}^\varepsilon + 1 = \det\big(D^2 (\varepsilon q_{\varepsilon t}^\varepsilon) + I_2\big) \end{cases}$$

Setting

$$\rho_t := \varepsilon \rho_{\varepsilon t}^\varepsilon + 1$$

$$q_t := \varepsilon q_{\varepsilon t}^\varepsilon$$

we are lead to a solution of the system:

$$\begin{cases} \partial_t \rho_t + \operatorname{div}(\rho_t \nabla^\perp q_t) = 0 \\ \rho_t = \det\big(D^2 q_t + I_2\big) \end{cases}$$

which is equivalent to the dual SG system as it has been previously shown.

Thus,

$$\text{dual SG system} \Leftrightarrow \text{SG}_\varepsilon \text{ system}$$

Having derived the rescaled dual SG system, SG_ε, we proceed to state and prove the main theorem of this chapter.

4.4 Convergence of smooth solutions

We are now ready to state the theorem:

Theorem 4.1.

Let ω_t, ψ_t be a solution to the incompressible Euler equation in vorticity-stream formulation

$$\begin{cases} \partial_t \omega_t + \operatorname{div}(\omega_t \nabla^\perp \psi_t) = 0 \\ \omega_t = \Delta \psi_t \end{cases}$$

such that $\omega \in C^2_{loc}\big(\mathbb{T}^2 \times [0, +\infty)\big)$

Let also $\varepsilon > 0$ with ρ_0^ε be a family of probability measures on the torus, initial data to SG_ε system such that:

$$\exists\, \alpha \in (0,1) \;\; \exists\, \lambda, \Lambda \in \mathbb{R} \; : \; 0 < \lambda \le \rho_0^\varepsilon \le \Lambda \text{ and } \rho_0^\varepsilon \in C^{0,\alpha}(\mathbb{T}^2)$$

and

$$\frac{\rho_0^\varepsilon - \omega_0}{\varepsilon} \text{ is bounded in } W^{1,\infty}(\mathbb{T}^2)$$

then

$$\exists \text{ a family } \rho_t^\varepsilon, q_t^\varepsilon \text{ of solutions to the } \mathrm{SG}_\varepsilon \text{ system such that}$$

$$\forall\, S > 0 \quad \exists\, \varepsilon_S > 0 \quad \forall\, \varepsilon \in (0, \varepsilon_s) \quad : \quad \frac{\rho_t^\varepsilon - \omega_t}{\varepsilon}, \frac{\nabla q_t^\varepsilon - \nabla \psi_t}{\varepsilon}$$

are uniformly bounded (no dependence on t, ε) in $L^\infty\big([0, S], W^{1,\infty}(\mathbb{T}^2)\big)$

Proof.

Before we begin proving anything at all, let us firstly check what is enough to show instead.

We define for all y, t, ε

$$g_t^\varepsilon := \frac{\rho_t^\varepsilon - \omega_t}{\varepsilon} \;\text{ and }\; h_t^\varepsilon := \frac{q_t^\varepsilon - \psi_t}{\varepsilon}$$

So, we actually need to show that $\|g_t^\varepsilon\|_{W^{1,\infty}}$, $\|\nabla h_t^\varepsilon\|_{W^{1,\infty}}$ are uniformly bounded.

Let us assume that the following inequality holds true

$$\|h_t^\varepsilon\|_{C^{2,\alpha}(\mathbb{T}^2)} \le C_{\|\psi_t\|_{C^{2,\alpha}(\mathbb{T}^2)}} \big(1 + \|g_t^\varepsilon\|_{C^{0,\alpha}(\mathbb{T}^2)}\big) \tag{4.4.1}$$

Then, for the quantity $\|\nabla h_t^\varepsilon\|_{W^{1,\infty}}$ which we want to estimate, the following facts hold true.

Due to the inclusions

$$C^{2,\alpha}(\mathbb{T}^2) \subseteq C^2(\mathbb{T}^2) \subseteq C^{1,1}(\mathbb{T}^2)$$

and the fact that the $W^{1,\infty}$ norm is equivalent to the Lipschitz norm $C^{0,1}$ on bounded sets with smooth boundary, that is

$$\|\cdot\|_{W^{1,\infty}(\mathbb{T}^2)} \sim \|\cdot\|_{C^{0,1}(\mathbb{T}^2)}$$

we get:

$$\|\nabla h_t^\varepsilon\|_{W^{1,\infty}(\mathbb{T}^2)} \leq C\|\nabla h_t^\varepsilon\|_{C^{0,1}(\mathbb{T}^2)} \leq C\|h_t^\varepsilon\|_{C^{1,1}(\mathbb{T}^2)} \leq C\|h_t^\varepsilon\|_{C^{2,\alpha}(\mathbb{T}^2)}$$

Also, the inclusion $C^{0,1}(\mathbb{T}^2) \subseteq C^{0,\alpha}(\mathbb{T}^2)$ holds true and using again the equivalency of the aforementioned norms, we have

$$\|g_t^\varepsilon\|_{C^{0,\alpha}(\mathbb{T}^2)} \leq C\|g_t^\varepsilon\|_{C^{0,1}(\mathbb{T}^2)} \leq C\|g_t^\varepsilon\|_{W^{1,\infty}(\mathbb{T}^2)}$$

Thus, utilizing the inequality (4.4.1) we obtain

$$\|\nabla h_t^\varepsilon\|_{W^{1,\infty}(\mathbb{T}^2)} \leq C_{\|\psi_t\|_{C^{2,\alpha}(\mathbb{T}^2)}}\big(1 + \|g_t^\varepsilon\|_{W^{1,\infty}(\mathbb{T}^2)}\big)$$

Poincaré inequaltiy implies that $\|g_t^\varepsilon\|_{L^\infty(\mathbb{T}^2)} \leq C\|\nabla g_t^\varepsilon\|_{L^\infty(\mathbb{T}^2)}$

Since $\|g_t^\varepsilon\|_{W^{1,\infty}}$ is by definition equal to the sum $\|g_t^\varepsilon\|_{L^\infty(\mathbb{T}^2)} + \|\nabla g_t^\varepsilon\|_{L^\infty(\mathbb{T}^2)}$ we deduce that

$$\|\nabla h_t^\varepsilon\|_{W^{1,\infty}(\mathbb{T}^2)} \leq C_{\|\psi_t\|_{C^{2,\alpha}(\mathbb{T}^2)}}\big(1 + \|\nabla g_t^\varepsilon\|_{L^\infty(\mathbb{T}^2)}\big) \tag{4.4.2}$$

Hence, it is enough to prove that $\|\nabla g_t^\varepsilon\|_{L^\infty}$ is uniformly bounded and that the inequality (4.4.1) holds true.

We begin proving the former.

$\|\nabla g_t^\varepsilon\|_{L^\infty}$ is uniformly bounded

$\omega \in C^2_{loc}\big(\mathbb{T}^2 \times [0,+\infty)\big)$ implies that:

$$\forall\, S > 0 \quad \psi \in L^\infty\big([0,S], C^3(\mathbb{T}^2)\big)$$

The condition that each ρ_0^ε are bounded and Hölder continuous provides us for each ε with a local smooth solution $\rho_t^\varepsilon, q_t^\varepsilon$ of the dual SG system.

Since dual SG system $\Leftrightarrow$ SG$_\varepsilon$ system we have a solution $\rho_t^\varepsilon, q_t^\varepsilon$ of the SG$_\varepsilon$ i.e.

$$\begin{cases} \partial_t \rho_t^\varepsilon + \operatorname{div}(\rho_t^\varepsilon\, \nabla^\perp q_t^\varepsilon) = 0 \\ \varepsilon\rho_t^\varepsilon + 1 = \det\big(\varepsilon D^2 q_t^\varepsilon + I_2\big) \end{cases}$$

recall that we have defined:

$$g_t^\varepsilon := \frac{\rho_t^\varepsilon - \omega_t}{\varepsilon} \text{ and } h_t^\varepsilon := \frac{q_t^\varepsilon - \psi_t}{\varepsilon}$$

Thus we have:

$$\rho_t^\varepsilon = \varepsilon g_t^\varepsilon + \omega_t \text{ and } q_t^\varepsilon = \varepsilon h_t^\varepsilon + \psi_t$$

and the SG_ε now reads:

$$\begin{cases} \partial_t(\varepsilon g_t^\varepsilon + \omega_t) + \mathrm{div}\big((\varepsilon g_t^\varepsilon + \omega_t)\nabla^\perp(\varepsilon h_t^\varepsilon + \psi_t)\big) = 0 \\ \varepsilon(\varepsilon g_t^\varepsilon + \omega_t) + 1 = \det\big(\varepsilon D^2(\varepsilon h_t^\varepsilon + \psi_t) + I_2\big) \end{cases} \tag{4.4.3}$$

We compute making calculations for each quantity individually.

$$\partial_t(\varepsilon g_t^\varepsilon + \omega_t) = \varepsilon\partial_t g_t^\varepsilon + \partial_t\omega_t$$

We continue with the divergence differential operator div and the quantity

$$\mathrm{div}\big((\varepsilon g_t^\varepsilon + \omega_t)\nabla^\perp(\varepsilon h_t^\varepsilon + \psi_t)\big)$$

Utilizing the fact that the rotated gradient of a real-valued function is divergence free Proposition1.1 we get:

$$\mathrm{div}\nabla^\perp(\varepsilon h_t^\varepsilon + \psi_t) = 0$$

We make use of Corollary1.2.1 to obtain

$$\begin{aligned} \mathrm{div}\big((\varepsilon g_t^\varepsilon + \omega_t)\nabla^\perp(\varepsilon h_t^\varepsilon + \psi_t)\big) &= \langle\, \nabla(\varepsilon g_t^\varepsilon + \omega_t)\,,\, \nabla^\perp(\varepsilon h_t^\varepsilon + \psi_t)\,\rangle \\ &= \langle\, \varepsilon\nabla g_t^\varepsilon + \nabla\omega_t\,,\, \varepsilon\nabla^\perp h_t^\varepsilon + \nabla^\perp\psi_t\,\rangle \end{aligned}$$

By the linearity of inner product we have:

$$\mathrm{div}\big((\varepsilon g_t^\varepsilon + \omega_t)\nabla^\perp(\varepsilon h_t^\varepsilon + \psi_t)\big) = \varepsilon\langle\nabla g_t^\varepsilon\,,\varepsilon\nabla^\perp h_t^\varepsilon + \nabla^\perp\psi_t\rangle + \varepsilon\langle\nabla\omega_t\,,\nabla^\perp h_t^\varepsilon\rangle + \langle\nabla\omega_t\,,\nabla^\perp\psi_t\rangle$$

Again, since $\mathrm{div}\nabla^\perp\psi_t = 0$ (divergence of rotated gradient), we have that:

$$\langle\nabla\omega_t\,,\nabla^\perp\psi_t\rangle = \mathrm{div}(\omega_t\nabla^\perp\psi_t)$$

So,

$$\partial_t\omega_t + \langle\nabla\omega_t\,,\nabla^\perp\psi_t\rangle = \partial_t\omega_t + \mathrm{div}(\omega_t\nabla^\perp\psi_t) = 0$$

because ω_t, ψ_t are a solution to the incompressible Euler equation in vorticity-stream formulation.

Thus, the first equation of SG_ε

$$\partial_t(\varepsilon g_t^\varepsilon + \omega_t) + \mathrm{div}\big((\varepsilon g_t^\varepsilon + \omega_t)\nabla^\perp(\varepsilon h_t^\varepsilon + \psi_t)\big) = 0$$

becomes
$$\varepsilon\partial_t g_t^\varepsilon + \varepsilon\langle\nabla g_t^\varepsilon\,,\varepsilon\nabla^\perp h_t^\varepsilon + \nabla^\perp\psi_t\rangle + \varepsilon\langle\nabla\omega_t\,,\nabla^\perp h_t^\varepsilon\rangle = 0$$

that is
$$\partial_t g_t^\varepsilon + \langle\nabla g_t^\varepsilon\,,\varepsilon\nabla^\perp h_t^\varepsilon + \nabla^\perp\psi_t\rangle + \langle\nabla\omega_t\,,\nabla^\perp h_t^\varepsilon\rangle = 0$$

Differentiating with respect to space, we get:

$$\begin{aligned}&\nabla\partial_t g_t^\varepsilon + (\varepsilon\nabla^\perp h_t^\varepsilon + \nabla^\perp\psi_t)\diamond\nabla(\nabla g_t^\varepsilon)+\\ &+\nabla(\varepsilon\nabla^\perp h_t^\varepsilon + \nabla^\perp\psi_t)\diamond\nabla g_t^\varepsilon + \nabla(\nabla\omega_t)\diamond\nabla^\perp h_t^\varepsilon + \nabla(\nabla^\perp h_t^\varepsilon)\diamond\nabla\omega_t = 0\end{aligned}$$

Since $\nabla(\varepsilon\nabla^\perp h_t^\varepsilon + \nabla^\perp\psi_t) = \nabla^\perp(\varepsilon\nabla h_t^\varepsilon + \nabla\psi_t)$ and also $\nabla^\perp w\diamond\nabla z = \nabla w\diamond\nabla^\perp z$ for every w, z we have:

$$\begin{aligned}&\partial_t\nabla g_t^\varepsilon + \langle\varepsilon\nabla^\perp h_t^\varepsilon + \nabla^\perp\psi_t\,,\nabla\rangle\nabla g_t^\varepsilon+\\ &+(\varepsilon D^2 h_t^\varepsilon + D^2\psi_t)\diamond\nabla^\perp g_t^\varepsilon + D^2\omega_t\diamond\nabla^\perp h_t^\varepsilon + D^2 h_t^\varepsilon\diamond\nabla^\perp\omega_t = 0\end{aligned}$$

Thus, we obtain

$$\begin{aligned}&\partial_t\nabla g_t^\varepsilon + \langle\varepsilon\nabla^\perp h_t^\varepsilon + \nabla^\perp\psi_t\,,\nabla\rangle\nabla g_t^\varepsilon =\\ &= -(\varepsilon D^2 h_t^\varepsilon + D^2\psi_t)\diamond\nabla^\perp g_t^\varepsilon - D^2\omega_t\diamond\nabla^\perp h_t^\varepsilon - D^2 h_t^\varepsilon\diamond\nabla^\perp\omega_t =: f_t^\varepsilon\end{aligned}$$

where we define the right hand side of the equality as a function f_t^ε.

In order to make "easier" the process of estimating $\|\nabla g_t^\varepsilon\|_{L^\infty}$ through $\|f_t^\varepsilon\|_{L^\infty}$, we simplify our notation a little bit.

Let us define the functions

$$u_t^\varepsilon := \nabla g_t^\varepsilon \text{ and } b_t^\varepsilon := \varepsilon\nabla^\perp h_t^\varepsilon + \nabla^\perp\psi_t$$
$$v_t^\varepsilon := \nabla h_t^\varepsilon \text{ and } \alpha_t := \nabla\psi_t\ ,\ \beta_t := \nabla\omega_t$$

Then, we obtain from the equality above that

$$\partial_t u_t^\varepsilon + \langle b_t^\varepsilon,\nabla\rangle u_t^\varepsilon = -(\varepsilon\nabla v_t^\varepsilon + \nabla\alpha_t)\diamond u_t^\varepsilon - \nabla\beta_t\diamond(v_t^\varepsilon)^\perp - \nabla v_t^\varepsilon\diamond\beta_t^\perp =: f_t^\varepsilon \qquad (4.4.4)$$

So, we begin making computations with $\|f_t^\varepsilon\|_{L^\infty}$

$$\|f_t^\varepsilon\|_{L^\infty} = \left\|-(\varepsilon\nabla v_t^\varepsilon + \nabla\alpha_t)\diamond u_t^\varepsilon - \nabla\beta_t\diamond(v_t^\varepsilon)^\perp - \nabla v_t^\varepsilon\diamond\beta_t^\perp\right\|_{L^\infty}$$

The triangle inequality implies

$\|f_t^\varepsilon\|_{L^\infty}$

$$
\begin{aligned}
&\leq \varepsilon\|(\nabla v_t^\varepsilon)\diamond u_t^\varepsilon\|_{L^\infty} + \|(\nabla\alpha_t)\diamond u_t^\varepsilon\|_{L^\infty} + \left\|\nabla\beta_t\diamond(v_t^\varepsilon)^\perp\right\|_{L^\infty} + \left\|\nabla v_t^\varepsilon\diamond\beta_t^\perp\right\|_{L^\infty}\\
&\leq \varepsilon\|\nabla v_t^\varepsilon\|_{L^\infty}\|u_t^\varepsilon\|_{L^\infty} + \|\nabla\alpha_t\|_{L^\infty}\|u_t^\varepsilon\|_{L^\infty} + \|\nabla\beta_t\|_{L^\infty}\left\|(v_t^\varepsilon)^\perp\right\|_{L^\infty} + \|\nabla v_t^\varepsilon\|_{L^\infty}\left\|\beta_t^\perp\right\|_{L^\infty}\\
&\leq \varepsilon\|v_t^\varepsilon\|_{W^{1,\infty}}\|u_t^\varepsilon\|_{L^\infty} + \|\psi_t\|_{C^2}\|u_t^\varepsilon\|_{L^\infty} + \|\omega_t\|_{C^2}\|v_t^\varepsilon\|_{L^\infty} + \|v_t^\varepsilon\|_{L^\infty}\|\omega_t\|_{C^2}\\
&\leq \varepsilon\|v_t^\varepsilon\|_{W^{1,\infty}}\|u_t^\varepsilon\|_{L^\infty} + \|\psi_t\|_{C^2}\|u_t^\varepsilon\|_{L^\infty} + \|\omega_t\|_{C^2}\|v_t^\varepsilon\|_{W^{1,\infty}} + \|\omega_t\|_{C^2}\|v_t^\varepsilon\|_{W^{1,\infty}}\\
&= \varepsilon\|v_t^\varepsilon\|_{W^{1,\infty}}\|u_t^\varepsilon\|_{L^\infty} + \|\psi_t\|_{C^2}\|u_t^\varepsilon\|_{L^\infty} + 2\|\omega_t\|_{C^2}\|v_t^\varepsilon\|_{W^{1,\infty}}\\
&= \varepsilon\|\nabla h_t^\varepsilon\|_{W^{1,\infty}}\|\nabla g_t^\varepsilon\|_{L^\infty} + \|\psi_t\|_{C^2}\|\nabla g_t^\varepsilon\|_{L^\infty} + 2\|\omega_t\|_{C^2}\|\nabla h_t^\varepsilon\|_{W^{1,\infty}}
\end{aligned}
$$

Hence, using inequality (4.4.2) we get

$$
\begin{aligned}
\|f_t^\varepsilon\|_{L^\infty} \leq &\varepsilon C_{\|\psi_t\|_{C^{2,\alpha}}}\big(1+\|\nabla g_t^\varepsilon\|_{L^\infty}\big)\|\nabla g_t^\varepsilon\|_{L^\infty}+\\
&+\|\psi_t\|_{C^2}\|\nabla g_t^\varepsilon\|_{L^\infty} + 2\|\omega_t\|_{C^2}C_{\|\psi_t\|_{C^{2,\alpha}}}\big(1+\|\nabla g_t^\varepsilon\|_{L^\infty}\big)
\end{aligned}
$$

Utilizing the inclusion $C^{2,\alpha}(\mathbb{T}^2)\subseteq C^2(\mathbb{T}^2)$, we have

$$
\begin{aligned}
\|f_t^\varepsilon\|_{L^\infty} \leq &\varepsilon C_{\|\psi_t\|_{C^{2,\alpha}},\|\omega_t\|_{C^{2,\alpha}}}\big(1+\|g_t^\varepsilon\|_{L^\infty}\big)\|\nabla g_t^\varepsilon\|_{L^\infty}+\\
&+C_{\|\psi_t\|_{C^{2,\alpha}},\|\omega_t\|_{C^{2,\alpha}}}\|\nabla g_t^\varepsilon\|_{L^\infty} + C_{\|\psi_t\|_{C^{2,\alpha}},\|\omega_t\|_{C^{2,\alpha}}}\big(1+\|g_t^\varepsilon\|_{L^\infty}\big)
\end{aligned}
$$

we get

$$
\|f\|_{L^\infty} \leq C(t)(1+\|r\|_{L^\infty}+\varepsilon\|r\|^2_{L^\infty}), \quad t\in[0,T]. \tag{4.4.5}
$$

Of course, since ψ_t and ω_t are in $C^2(\mathbb{T}^2)$ for any $t\in[0,T]$ for any $T>0$ (due to the global existence of smooth solutions for the Euler equation in $\mathbb{T}^2$), the time-dependent constant $C(t)$, $t\in[0,T]$, in (4.4.5) can be estimated for any $T>0$ by

$$
C_T := \max_{t\in[0,T]} C(t), \quad T>0,
$$

such that (4.4.5) becomes

$$
\|f(t,\cdot)\|_{L^\infty} \leq C_T(1+\|g_t^\varepsilon(t,\cdot)\|_{L^\infty}+\varepsilon\|g_t^\varepsilon(t,\cdot)\|^2_{L^\infty}), \quad t\in[0,T]. \tag{4.4.6}
$$

Now, assuming $X(t,x)$, $t\geq 0$, $X(0,x)=x$, is the Lagrangian flow corresponding to the transport equation (4.4.4)

$$
\partial_t g_t^\varepsilon(t,x) + b_t^\varepsilon(t,x)\cdot\nabla g_t^\varepsilon(t,x) = f^\varepsilon(t,x), \quad g_t^\varepsilon(0,x)=r^0(x),
$$

that is,

$$
\dot{X}(t,x) = b_t^\varepsilon(t,X(t,x)), \quad X(0,x)=x,
$$

we obtain for $z(t) := g_t^\varepsilon(t,X(t,x))$

$$
\begin{aligned}
\dot{z}(t) &= \partial_t g_t^\varepsilon(t,X(t,x)) + (\nabla g_t^\varepsilon(t,X(t,x)))\dot{X}(t,x)\\
&= \partial_t g_t^\varepsilon(t,X(t,x)) + b_t^\varepsilon(t,X(t,x))\cdot\nabla g_t^\varepsilon(t,X(t,x))\\
&= f^\varepsilon(t,X(t,x))
\end{aligned}
$$

and hence

$$g_t^\varepsilon(t, X(t,x)) = g_t^\varepsilon(0,x) + \int_0^t f^\varepsilon(s, X(s,x))\, ds$$

or, equivalently,

$$g_t^\varepsilon(t,x) = g_t^\varepsilon(0, X^{-1}(t,x)) + \int_0^t f^\varepsilon(s, X(s, X^{-1}(t,x)))\, ds,$$

from which we obtain

$$\begin{aligned} |g_t^\varepsilon(t,x)| &\le |g_t^\varepsilon(0, X^{-1}(t,x))| + \int_0^t |f^\varepsilon(s, X(s, X^{-1}(t,x)))|\, ds \\ &\le \|g_t^\varepsilon(0,\cdot)\|_{L^\infty} + \int_0^t \|f^\varepsilon(s,\cdot)\|_{L^\infty}\, ds \end{aligned}$$

,

and thus

$$\|g_t^\varepsilon(t,\cdot)\|_{L^\infty} \le \|g_t^\varepsilon(0,\cdot)\|_{L^\infty} + \int_0^t \|f(s,\cdot)\|_{L^\infty}\, ds$$

which by (4.4.6) becomes

$$\|g_t^\varepsilon(t,\cdot)\|_{L^\infty} \le \|g_t^\varepsilon(0,\cdot)\|_{L^\infty} + C_T \int_0^t (1 + \|g_t^\varepsilon(s,\cdot)\|_{L^\infty} + \varepsilon \|g_t^\varepsilon(s,\cdot)\|_{L^\infty}^2)\, ds, \quad t \in [0,T]. \tag{4.4.7}$$

Then a generalized Gronwall estimate in integral form, which is attributed to Bihari (in [5]), yields the desired.

Proof of (4.4.1)

Expanding the second equation of (4.4.3), satisfied by h_t^ε, using the fact that $\det(A+\varepsilon B) = \det A + \varepsilon\big(\operatorname{tr} A \operatorname{tr} B - \operatorname{tr}(AB)\big) + \varepsilon^2 \det B$ we get

$$\Delta h_t^\varepsilon = -\varepsilon\left((\Delta\psi_t)\Delta h_t^\varepsilon - \operatorname{tr}\big((D^2\psi_t)D^2 h_t^\varepsilon\big)\right) - \varepsilon^2 \det D^2 h_t^\varepsilon - \det D^2\psi_t + g_t^\varepsilon. \tag{4.4.8}$$

We will show that from this equation we obtain the estimate (4.4.1)

To obtain (4.4.1) we first get by Schauder estimates, see e.g. [25], that the solution of (4.4.8) satisfies

$$\|h_t^\varepsilon\|_{C^{2,\alpha}(\mathbb{T}^2)} \le C(\varepsilon\|\psi_t\|_{C^{2,\alpha}(\mathbb{T}^2)}\|h_t^\varepsilon\|_{C^{2,\alpha}(\mathbb{T}^2)} + \varepsilon^2\|h_t^\varepsilon\|^2_{C^{2,\alpha}(\mathbb{T}^2)} + \|\bar\phi\|^2_{C^{2,\alpha}(\mathbb{T}^2)} + \|g_t^\varepsilon\|_{C^{0,\alpha}(\mathbb{T}^2)})$$

for all $\varepsilon > 0$, where $C > 0$ is a constant independent of $\varepsilon > 0$ and any of the appearing functions.

Since ψ_t is a known function, which moreover is, under sufficient regular initial data imposed on ω_t, as smooth as we like, recalling $\varphi = \|\psi_t\|_{C^{2,\alpha}(\mathbb{T}^2)}$ and setting $a := \|h_t^\varepsilon\|_{C^{2,\alpha}(\mathbb{T}^2)} \geq 0$, we obtain

$$\begin{aligned} a &\leq C(\varepsilon\varphi a + \varepsilon^2 a^2 + \varphi^2 + \|g_t^\varepsilon\|_{C^{0,\alpha}(\mathbb{T}^2)}) \\ &\leq C(\varepsilon\varphi a + \varepsilon^2 a^2 + \varphi^2 + 1 + \|g_t^\varepsilon\|_{C^{0,\alpha}(\mathbb{T}^2)}) \\ &\leq (1 + \|g_t^\varepsilon\|_{C^{0,\alpha}(\mathbb{T}^2)})C(1 + \varepsilon\varphi a + \varepsilon^2 a^2 + \varphi^2) \end{aligned}$$

and thus

$$a \leq C(\phi, \rho)(1 + \varepsilon a + \varepsilon^2 a^2), \quad \varepsilon > 0, \tag{4.4.9}$$

where

$$, \quad C(\phi, \rho) := C_1(\phi)(1 + \|g_t^\varepsilon\|_{C^{0,\alpha}(\mathbb{T}^2)}), \quad C_1(\phi) := (1 + \varphi + \varphi^2)C.$$

We claim now that the estimate (4.4.9) implies that there exists a constant $C_2(\phi) > 0$ depending only on $C_1(\phi)$ and an $\varepsilon_0 > 0$ such that

$$a \leq C_2(\phi)(1 + \|g_t^\varepsilon\|_{C^{0,\alpha}(\mathbb{T}^2)}), \quad \varepsilon \in (0, \varepsilon_0). \tag{4.4.10}$$

(Clearly, if (4.4.10) holds true, then we obtain (4.4.9) for $\varepsilon \in (0, \varepsilon_0)$ just by replacing $C_1(\phi)$ by $C_2(\phi)$ in $C(\phi, \rho)$.)

Assume now that there is no constant $C_2(\phi) > 0$ and no $\varepsilon_0 > 0$ such that (4.4.10) holds true, but that still (4.4.9) is satisfied. Then, for any $n \in \mathbb{N}$ there exists an $0 < \varepsilon_n < \frac{1}{n}$ such that

$$a \geq nC(\phi, \rho) > 0 \quad \text{for} \quad \varepsilon = \varepsilon_n, \quad n \in \mathbb{N},$$

and dividing (4.4.9) for $\varepsilon = \varepsilon_n$ by $aC(\phi, \rho) > 0$ we obtain

$$\frac{1}{C(\phi, \rho)} - \frac{1}{a} - \varepsilon_n \leq \varepsilon_n^2 a, \quad n \in \mathbb{N},$$

and thus, for $n_0 \in \mathbb{N}$ with $n_0 \geq \max\{4, 2C(\phi, \rho)\}$ we have

$$C_3(\phi, \rho) := \frac{1}{4C(\phi, \rho)} \leq \frac{1}{2C(\phi, \rho)} - \frac{1}{a} \leq \frac{1}{C(\phi, \rho)} - \frac{1}{a} - \varepsilon_n \leq \varepsilon_n^2 a, \quad n \in \mathbb{N}, \quad n \geq n_0. \tag{4.4.11}$$

On the other hand, from regularity theory for the Monge-Ampère equation we obtain that there exists a $C > 0$ independent of $\varepsilon > 0$ and any functions, such that

$$\|q_t^\varepsilon\|_{C^{2,\alpha}(\mathbb{T}^2)} \leq C\|\rho_t^\varepsilon\|_{C^{0,\alpha}(\mathbb{T}^2)}, \quad \varepsilon > 0,$$

where we recall $q_t^\varepsilon = \psi_t + \varepsilon h_t^\varepsilon$, $\rho_t^\varepsilon = \omega_t + \varepsilon g_t^\varepsilon$, such that we obtain by the triangle inequality

$$\varepsilon \|h_t^\varepsilon\|_{C^{2,\alpha}(\mathbb{T}^2)} \leq \|\psi_t\|_{C^{2,\alpha}(\mathbb{T}^2)} + \|q_t^\varepsilon\|_{C^{2,\alpha}(\mathbb{T}^2)} \leq C(\phi, \omega_t, g_t^\varepsilon), \quad \varepsilon \in (0,1),$$

where

$$C(\phi, \omega_t, g_t^\varepsilon) := \|\psi_t\|_{C^{2,\alpha}(\mathbb{T}^2)} + C(\|\omega_t\|_{C^{0,\alpha}(\mathbb{T}^2)} + \|g_t^\varepsilon\|_{C^{0,\alpha}(\mathbb{T}^2)}).$$

Recalling now the abbreviation $a = \|h_t^\varepsilon\|_{C^{2,\alpha}(\mathbb{T}^2)}$ and (4.4.11), which holds true for the sequence (ε_n) introduced above with $0 < \varepsilon_n < \frac{1}{n}$, we thus have

$$nC_3(\phi, \rho) < \frac{C_3(\phi, \rho)}{\varepsilon_n} \leq \varepsilon_n a \leq C(\phi, \omega_t, g_t^\varepsilon), \quad n \in \mathbb{N}, \quad n \geq n_0,$$

which yields an obvious contradiction, since $C_3(\phi, \rho) > 0$. Thus, the assumption that (4.4.10) fails is not true, and establishes the latter. With that, the proof is completed.

□

APPENDIX A

Appendix

This appendix has been created using the following books [1] [2] [3] [27] [12] [9] [34] [4] [18] [19] [26] [33] [36]

A.1 Notations

Here we will summarize the symbols we are going to use in order to denote some notions. Of course, many of the notions below have multiple notations that are being used to describe them.

Definition A.1. Let $n \in \mathbb{N}$

$$\mathbb{R}^n \ni e_i := (0, \dots, 0, 1, 0, \dots, 0)$$

where the 1 is placed in the i-th position, where $i \in T(n)$

Clarification A.1.1. $T(m)$ symbolizes the set containing all natural numbers up to m, including it. Thus,

$$T(m) := \{\, i \in \mathbb{N} \mid i \leq m \} = \{1, 2, 3 \dots, m-2, m-1, m\}$$

we consider 1 (and not 0) to be the smallest natural number.

Remark. Unless otherwise stated the symbol n is used to denote a natural number, so when we write n we shall always mean that $n \in \mathbb{N}$.

Definition A.2 (standard inner product)**.** Let $u, v \in \mathbb{R}^n$ be two vectors with representations $u = (u_1, u_2, \dots, u_{n-1}, u_n)$ and $v = (v_1, v_2, \dots, v_{n-1}, v_n)$ on the standard basis of $\mathbb{R}^n$. We denote their standard inner product as:

$$\langle u, v \rangle := \sum_{i=1}^{n} u_i \cdot v_i = u_1 v_1 + \cdots + u_n v_n$$

Clarification A.2.1. We consider vector to be row vectors and we denote the column vectors with the transpose matrix. $u = (u_1, \dots, u_n)$ and $u^T = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$

Remark. So, we actually identify the vector space $\mathbb{R}^n$ with the matrix space $\mathbb{R}^{1\times n}$, when we are referring to vectors and inner products.

We use a special symbol to denote the matrix multiplication instead of the usual dot $\cdot$

Definition A.3 (matrix multiplication symbol)**.** Let the matrices $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{n\times k}$ with $m, n, k \in \mathbb{N}$ we denote their product as:

$$A \diamond B \equiv A \cdot B$$

Proposition A.1. $u, v \in \mathbb{R}^n \Rightarrow \langle u, v\rangle = u \diamond v^T$

Real-valued function spaces

Definition A.4 (compactly contained $\subset\subset$)**.**

$$V \subset\subset U :\Longleftrightarrow V \subseteq \overline{V} \subseteq U \text{ and } \overline{V} \text{ is compact}$$

Definition A.5 (L^p norm)**.** Let $f : U \to \mathbb{R}$ be a Lebesgue measurable function, then we define

$$\|f\|_{L^p(U)} := \begin{cases} \left(\int_U |f(x)|^p \, dx\right)^{\frac{1}{p}} & p \in [1, +\infty) \\ \operatorname{ess\,sup}_U |f(x)| & p = \infty \end{cases}$$

Definition A.6 (L^p space)**.** For $p \in [1, +\infty]$ we define the space of Lebesgue measurable function with finite L^p-norm, which is the following:

$$L^p(U) := \{\ f : U \to \mathbb{R} \ \mid\ f \text{ is Lebesgue measurable and } \|f\|_{L^p(U)} < +\infty\ \}$$

Definition A.7 (local spaces)**.** Whichever function space on U has the subscript $_{loc}$ contains the functions belonging in the respective space for every V compactly contained in U e.g.

$$L^p_{loc}(U) := \{\ f : U \to \mathbb{R} \ \mid\ f \in L^p(V)\ \forall\, V \subset\subset U\ \}$$

Definition A.8 ($C(U)$)**.**

$$C(U) := \{\ f : U \to \mathbb{R} \ \mid\ f \text{ is continuous on } U\ \}$$

Definition A.9 ($C(\bar{U})$)**.**

$$C(\bar{U}) := \{\ f \in C(U) \ \mid\ f \text{ is uniformly continuous on bounded subsets of } U\ \}$$

Definition A.10 (multi-index)**.** The vector $a = (a_1, \ldots, a_n) \in \mathbb{N}_0^n$ is called a multi-index of order $|a| = \sum_{i=1}^n a_i$

Clarification A.10.1. $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ is the set of natural numbers along with zero and $\mathbb{N}_0^n$ refers to the set $\{\ (b_1, \ldots, b_n) \in \mathbb{R}^n | b_1, \ldots, b_n \in \mathbb{N}_0\}$, the n-dimensional product space of natural numbers including zero.

Definition A.11 (multi-index/partial derivatives)**.** For the multi-index a and the function $f : U \text{ open } \subseteq \mathbb{R}^n \to \mathbb{R}$ we define the derivatives of order $|a|$

$$(D^a f)(x) \equiv D^a f(x) := \frac{\partial^{|a|} f}{\partial x_1^{a_1} \cdots \partial x_n^{a_n}}(x)$$

Clarification A.11.1. When we write D^a we refer to one of the (many) derivatives with order $|a|$, which we will specify when needed.

Example. Let $u : \mathbb{R}^2 \to \mathbb{R}$

$$\text{then } Du(x) \equiv D^1 u(x) \text{ denotes either } \tfrac{\partial u}{\partial x_1}(x) \text{ or } \tfrac{\partial u}{\partial x_2}(x)$$

Clarification A.11.2. In the special scenario where the multi-index is the zero vector, we define the following:

$$D^0 f(x) := f(x) \text{ and } \tfrac{\partial^0}{\partial x_i^0} f(x) := f(x)$$

For every multi-index a, we abide by the convention:

$$\tfrac{\partial^{|a|}}{\partial x_i^0} f(x) := f(x)$$

Definition A.12 (partial derivative, gradient operator and nabla symbol)**.** Let $f : U \text{ open } \subseteq \mathbb{R}^n \to \mathbb{R}$ be a partially differentiable function we define its gradient as:

$$\nabla f(x) = \left(\frac{\partial}{\partial x_1} f, \ldots, \frac{\partial}{\partial x_n} f \right) \equiv (\partial_1 f, \ldots, \partial_n f)$$

Definition A.13 (vector-valued functions)**.** Let U be an open subset of $\mathbb{R}^n$. A vector-valued function $f : U \to \mathbb{R}^m$ is denoted as:

$$f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} = (f_1, f_2, \ldots, f_m)^T$$

which is an abbreviation of the term $f(x) = \big(f_1(x), f_2(x), \ldots, f_{m-1}(x), f_m(x)\big)^T$ where every component function of the vector-valued function f is a real-valued function, meaning that $f_i : U \to \mathbb{R} \ \ \forall i \in T(m)$

Remark. So, in the case of vector-valued functions we identify the space $\mathbb{R}^m$ with the matrix space $\mathbb{R}^{m \times 1}$

Definition A.14 (partial derivatives, gradient of vector-valued functions)**.** Let $f : U \text{ open } \subseteq \mathbb{R}^n \to \mathbb{R}^m$ be a partially differentiable function with $f = (f_1, f_2, \ldots, f_m)^T$, we define its gradient as:

$$\nabla f(x) = \begin{pmatrix} \frac{\partial}{\partial x_1} f_1 & \cdots & \frac{\partial}{\partial x_n} f_1 \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_1} f_m & \cdots & \frac{\partial}{\partial x_n} f_m \end{pmatrix} = \begin{pmatrix} \nabla f_1 \\ \vdots \\ \nabla f_m \end{pmatrix}$$

Definition A.15 (derivative)**.** We call the function $f : U$ open $\subseteq \mathbb{R}^n \to \mathbb{R}^m$ differentiable at $x \in U :\Longleftrightarrow$

$$\exists \text{ linear map } D : \mathbb{R}^n \to \mathbb{R}^m \text{ such as } \lim_{v \to \bar{0}} \frac{f(x+v) - f(x) - D(v)}{||v||_2} = \bar{0}$$

For our purposes we consider the vector spaces $\mathbb{R}^k$ ($k = n, m$ and such...) with the standard bases consisting of the vectors e_i for every $i \in T(k)$. Hence, we identify every linear map with its corresponding matrix regarding the standard bases. Thus for us, the derivative is nothing more than a matrix

$$\begin{pmatrix} d_{11} & \dots & d_{1n} \\ \vdots & \ddots & \vdots \\ d_{m1} & \dots & d_{mn} \end{pmatrix}$$

belonging in the matrix space $\mathbb{R}^{m \times n}$. So, we can rewrite the definition as

$$\exists \text{ matrix } D : \mathbb{R}^n \to \mathbb{R}^{m \times n} \text{ such as } \lim_{v \to \bar{0}} \frac{f(x+v) - f(x) - (v \diamond D)^T}{||v||_2} = \bar{0}$$

By identifying the space $\mathbb{R}^{m \times n}$ with the space $\mathbb{R}^{m \cdot n}$ we can define the k-th derivative of a vector-valued function in the exact same manner.

In the case we have a real-valued function, that is $m = 1$, we define the following function spaces

Definition A.16 ($C^k(U)$)**.**

$$C^k(U) := \{ \ f : U \to \mathbb{R} \ \mid \ f \text{ is k-times continuously differentiable on } U \ \}$$

Definition A.17 ($C^\infty(U)$)**.**

$$C^\infty(U) := \bigcap_{k \in \mathbb{N}} C^k(U)$$

Definition A.18 ($C^k(\bar{U})$)**.**

$$C^k(\bar{U}) := \{ \ f \in C^k(U) \ \mid \ D^a f \text{ is uniformly continuous on bounded subsets of } U$$

$$\forall \text{ multi-index } a : |a| \leq k \ \}$$

Definition A.19 ($C^\infty(\bar{U})$)**.**

$$C^\infty(\bar{U}) := \bigcap_{k \in \mathbb{N}} C^k(\bar{U})$$

Definition A.20 (compact support spaces)**.** Whichever function space on U has the subscript $_c$ contains the functions belonging in the respective space and having compact support i.e.

$$C_c^k(U) := \{ \ f \in C^k(U) \ \mid \ f \text{ has compact support } \}$$

Vector-valued function spaces

Definition A.21. Let $f : U$ open $\subseteq \mathbb{R}^n \to \mathbb{R}^m$ be a vector-valued function, then the respective function spaces are denoted with the same symbols adding $: \mathbb{R}^m$ next to the domain U. And they consist of those vector-valued functions whose each component, real-valued function belongs to the respective real-valued function space i.e.

$$C^k(U : \mathbb{R}^m) := \{ \, f : U \to \mathbb{R}^m \ \mid \ f_i \in C^k(U) \ \ \forall i \in T(m) \, \}$$

$$L^p(U : \mathbb{R}^m) := \{ \, f : U \to \mathbb{R}^m \ \mid \ f_i \in L^p(U) \ \ \forall i \in T(m) \, \}$$

etc

Proposition A.2.

Lef $f : \mathbb{R}^n \to \mathbb{R}^n$ partially differentiable function with $f = (f_1, \dots, f_n)^T$, then

$$\mathrm{tr}(\nabla f) = \mathrm{div} f$$

Indeed

$$\nabla(f) = \begin{pmatrix} \partial_1 f_1 & \cdots & \partial_n f_1 \\ \vdots & \ddots & \vdots \\ \partial_1 f_n & \cdots & \partial_n f_n \end{pmatrix}$$

Thus,

$$\mathrm{tr}(\nabla f) = \sum_{i=1}^{n} \partial_i f_i = \langle \nabla, f \rangle = \mathrm{div} f$$

A.2 Norms and inner product

A.2.1 Inner product

Definition A.22 (Inner product, complex)**.**

Let V be a vector space over the field of complex numbers $\mathbb{C}$ then the map $\langle \cdot, \cdot \rangle$: $V \times V \to \mathbb{C}$ is called an inner product if and only if the following conditions hold true:

1. conjugate symmetry

$$\forall \, x, y \in V$$

$$\langle x, y \rangle = \overline{\langle y, x \rangle}$$

2. linearity in the first argument

$$\forall \, x, y \in V \text{ and } \forall \, a, b \in \mathbb{C}$$

$$\langle ax + by, z \rangle = a\langle x, z \rangle + b\langle y, z \rangle$$

3. positive definiteness

$$\forall\, x \in V \setminus \{0_V\}$$

$$\langle x, x \rangle > 0$$

Remark.
Except some definitions where it is explicitly written, zero 0 refers to the zero of the respective space, without the use of any subscript to notate it.

Proposition A.3.

The properties of its definition immediately imply:

$$\langle x, x \rangle = 0 \Leftrightarrow x = 0$$

and conjugate linearity in the second argument

$$\langle x, ay + bz \rangle = \bar{a}\langle x, z \rangle + \bar{b}\langle y, z \rangle$$

In the case where the field of real numbers is chosen, the definition remains the same but conjugate symmetry reduces to symmetry.

Since $\bar{c} = c$ when $c \in \mathbb{R}$

Definition A.23 (Inner product, real)**.**

Let V be a vector space over the field of real numbers $\mathbb{R}$ then the map $\langle \cdot\, , \cdot \rangle : V \times V \to \mathbb{R}$ is called an inner product if and only if the following conditions hold true:

1. symmetry

$$\forall\, x, y \in V$$

$$\langle x, y \rangle = \langle y, x \rangle$$

2. linearity in the first argument

$$\forall\, x, y \in V \text{ and } \forall\, a, b \in \mathbb{R}$$

$$\langle ax + by, z \rangle = a\langle x, z \rangle + b\langle y, z \rangle$$

3. positive definiteness

$$\forall\, x \in V \setminus \{0_V\}$$

$$\langle x, x \rangle > 0$$

The same properties as above hold true, with the only difference being that now we have linearity in the second argument.

Proposition A.4.

The properties of the definition now imply:

$$\langle x, x \rangle = 0 \Leftrightarrow x = 0$$

and linearity in the second argument

$$\langle x, ay + bz \rangle = a\langle x, z \rangle + b\langle y, z \rangle$$

In contrast to our choice depending the norm notation in each case (using subscripts), for the inner product we stick to the same symbol. The reason we do that is because we usually compute with the standard inner product being involved.

A.2.2 Norms

Most of the time the Euclidean (or any equivalent norm in $\mathbb{R}^n$) is written with the absolute value $|\cdot|$ symbol. And the typical norm $||\cdot||$ symbol is reserved to characterise function spaces' norms.

But we will not oblige by this rule.

We will explicitly "declare" which norm is considered in each case by mentioning it or by putting a suitably chosen subscript. Usually, when we are referring to a non-specific norm or the standard/Euclidean one, then we will use the symbol without a subscript.

Remark.
A non-specific norm is a (generic) norm having no special "structure", that is satisfying only the properties of the definition and their consequences.

Definition A.24 (absolute value $|\cdot|$)**.**

$$|\cdot| \text{ denotes the absolute value on } \mathbb{R}$$

Definition A.25 (norm)**.** Let X be a vector space over a field (for our purposes that will usually be the real numbers). We call norm a non-negative function $||\cdot|| : X \to \mathbb{R}_0^+$ with the following three properties

1. positive definiteness
$$||x|| = 0_{\mathbb{R}} \Leftrightarrow x = 0_X$$

2. absolute homogenity
$$||\lambda x|| = |\lambda| \; ||x|| \;\; \forall \lambda \in \mathbb{R} \text{ and } x \in X$$

3. triangle inquality (or subadditivity)

$$||x+y|| \leq ||x|| + ||y|| \ \ \forall x, y \in X$$

Clarification A.25.1.
$\mathbb{R}_0^+ := \{ \, x \in \mathbb{R} \mid x \geq 0 \, \}$ is the set of non-negative real numbers. The non-negativity of the norm function is also a (the "hidden" fourth) requirement.

Remark. The absolute value is a norm.

We now introduce the notation we are going to use for some well-known and commonly used norms, such as:

Definition A.26 (Euclidean norm or 2-norm $||\cdot||_2$)**.**

$$||\cdot||_2 \text{ denotes the Euclidean (standard) norm on } \mathbb{R}^n$$

Let $x \in \mathbb{R}^n$ with $x = (x_1, \ldots, x_n)$ then

$$||x||_2 := \left(\sum_{i=1}^{n} x_i^2 \right)^{\frac{1}{2}}$$

A.2.3 Matrix norms

The next one is a matrix norm (meaning that the aforementioned vector space X is $\mathbb{R}^{n\times m}$) which goes by the names $L_{2,2}$ norm or Frobenius norm.

Definition A.27 (Frobenius norm $||\cdot||_F$ or $L^{2,2}$-norm $||\cdot||_{L^{2,2}}$)**.**

Let $A \in \mathbb{R}^{n\times m}$ with

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}$$

then

$$||A||_F := \left(\sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij}^2 \right)^{\frac{1}{2}}$$

We now prove that the map we have just defined is indeed a norm.

Proposition A.5 ($||\cdot||_F$ is a norm)**.** The Frobenius norm is indeed a norm.

Proof.

The idea behind the proof is to use known inequalities for the (Euclidean) 2-norm $||\cdot||_2$, which looks very similar to this norm. In fact, $L^{2,2}$ or Frobenious norm $||\cdot||_F$ is a summation of standard 2-norms (Euclidean vector norms). Indeed, if we write

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} = \begin{pmatrix} A_1 \\ \vdots \\ A_n \end{pmatrix}$$

where $A_i = (a_{i1}, \dots, a_{im})$ $\forall i \in T(n)$, then we have that $\sum_{j=1}^{m} a_{ij}^2 = \|A_i\|_2^2$

Hence, $||A||_F = \left(\sum_{i=1}^{n} \|A_i\|_2^2 \right)^{\frac{1}{2}}$. Now we start proving the requirements:

Obviously the Forbenius "norm" is a function $||\cdot||_F : \mathbb{R}^{n\times m} \to \mathbb{R}_0^+$

Next, we prove the three requirements:

i) Let $A \in \mathbb{R}^{n\times m}$

$$||A||_F = 0 \iff \left(\sum_{i=1}^{n} \|A_i\|_2^2 \right)^{\frac{1}{2}} = 0 \iff \sum_{i=1}^{n} \|A_i\|_2^2 = 0 \xLeftrightarrow[\forall i \in T(n)]{\|A_i\|_2^2 \geq 0}$$

$$\iff \forall i \in T(n)\ \ \|A_i\|_2^2 = 0 \iff \|A_i\|_2 = 0 \iff A_i = 0_{\mathbb{R}^m} \iff$$

$$\iff A = 0_{\mathbb{R}^{n\times m}}$$

ii) Let $\lambda \in \mathbb{R}$ and $A \in \mathbb{R}^{n\times m}$

$$||\lambda A||_F = \left(\sum_{i=1}^{n} (\lambda \|A_i\|_2)^2 \right)^{\frac{1}{2}} = \left(\sum_{i=1}^{n} \lambda^2 \|A_i\|_2^2 \right)^{\frac{1}{2}} = \left(\lambda^2 \sum_{i=1}^{n} \|A_i\|_2^2 \right)^{\frac{1}{2}} =$$

$$= (\lambda^2)^{\frac{1}{2}} \left(\sum_{i=1}^{n} \|A_i\|_2^2 \right)^{\frac{1}{2}} = |\lambda| \left(\sum_{i=1}^{n} \|A_i\|_2^2 \right)^{\frac{1}{2}} = |\lambda|\ ||A||_F$$

iii) Let $A, B \in \mathbb{R}^{n\times m}$ then:

Since $||\cdot||_2$ is a norm in $\mathbb{R}^m$ we have that:

$$\forall i \in T(n)\ \ ||A_i + B_i||_2 \leq ||A_i||_2 + ||B_i||_2$$

$$\xRightarrow[s^2 \nearrow \text{ on } s \geq 0]{||\cdot||_2 \geq 0} \forall i \in T(n)\ \ ||A_i + B_i||_2^2 \leq \left(||A_i||_2 + ||B_i||_2 \right)^2$$

$$\xrightarrow[\text{over i}]{\text{summing}} \sum_{i=1}^{n} ||A_i + B_i||_2^2 \leq \sum_{i=1}^{n} \big(||A_i||_2 + ||B_i||_2\big)^2$$

$$\xrightarrow[\text{sum}\geq 0]{\sqrt{s}\nearrow} \left(\sum_{i=1}^{n} ||A_i + B_i||_2^2\right)^{\frac{1}{2}} \leq \left(\sum_{i=1}^{n} \big(||A_i||_2 + ||B_i||_2\big)^2\right)^{\frac{1}{2}}$$

Because $|| \cdot ||_2$ is a norm in $\mathbb{R}^n$, we also have that:

$$\forall x, y \in \mathbb{R}^n \ \ ||x + y||_2 \leq ||x||_2 + ||y||_2$$

If we write x, y as $(x_1, \ldots, x_n)$,$(y_1, \ldots, y_n)$ respectively, then

$$\forall i \in T(n) \ \ \forall x_i, y_i \in \mathbb{R} \quad \left(\sum_{i=1}^{n} (x_i + y_i)^2\right)^{\frac{1}{2}} \leq \left(\sum_{i=1}^{n} x_i^2\right)^{\frac{1}{2}} + \left(\sum_{i=1}^{n} y_i^2\right)^{\frac{1}{2}}$$

Choosing $x_i = ||A_i||_2$ and $y_i = ||B_i||_2$ for $i \in T(n)$, we obtain:

$$\left(\sum_{i=1}^{n} \big(||A_i||_2 + ||B_i||_2\big)^2\right)^{\frac{1}{2}} \leq \left(\sum_{i=1}^{n} ||A_i||_2^2\right)^{\frac{1}{2}} + \left(\sum_{i=1}^{n} ||B_i||_2^2\right)^{\frac{1}{2}}$$

Combining the inequalities with the same term, we have shown that:

$$\left(\sum_{i=1}^{n} ||A_i + B_i||_2^2\right)^{\frac{1}{2}} \leq \left(\sum_{i=1}^{n} ||A_i||_2^2\right)^{\frac{1}{2}} + \left(\sum_{i=1}^{n} ||B_i||_2^2\right)^{\frac{1}{2}}$$

$\Leftrightarrow ||A + B||_2 \leq ||A||_2 + ||B||_2$ □

Proposition A.6 ($||\cdot||_F$ is submultiplicative)**.** The Frobenius norm is sub-multiplicative in the space of square matrices, that is $\forall A, B \in \mathbb{R}^{n\times n}$ the following inequality holds

$$||A \diamond B||_F \leq ||A||_F \cdot ||B||_F$$

Proof. To prove this result, all we are going to need is the Cauchy-Schwarz inequality for the Euclidean norm $|| \cdot ||_2$ on $\mathbb{R}^n$.

Let $u = (u_1, \ldots, u_n) \in \mathbb{R}^n$ and $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$ then

$$|\langle u, v\rangle| \leq ||u||_2 \cdot ||v||_2 \Leftrightarrow \langle u, v\rangle^2 \leq ||u||_2^2 \cdot ||v||_2^2$$

$$\Leftrightarrow \left(\sum_{l=1}^{n} u_l \cdot v_l\right)^2 \leq \left(\sum_{l=1}^{n} u_l^2\right) \cdot \left(\sum_{l=1}^{n} v_l^2\right)$$

Let $A, B \in \mathbb{R}^{n\times n}$ with $A = (a_{ij})$ and $B = (b_{ij})$ then $A \diamond B = (c_{ij})$ where $c_{ij} = \sum\limits_{k=1}^{n} a_{ik} b_{kj}$. So

$$||A \diamond B||_F = \left(\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}^2\right)^{\frac{1}{2}} = \left(\sum_{i=1}^{n}\sum_{j=1}^{n} \Big(\sum_{k=1}^{n} a_{ik} b_{kj}\Big)^2\right)^{\frac{1}{2}}$$

Choosing $l = k$ and $u_l = u_k = a_{ik}$, $v_l = v_k = b_{kj}$ we have from the squared Cauchy-Schwarz inequality that:

$$\left(\sum_{k=1}^{n} a_{ik}b_{kj}\right)^2 \leq \left(\sum_{k=1}^{n} a_{ik}^2\right) \cdot \left(\sum_{k=1}^{n} b_{kj}^2\right)$$

Thus

$$\left(\sum_{i=1}^{n}\sum_{j=1}^{n}(\sum_{k=1}^{n} a_{ik}b_{kj})^2\right)^{\frac{1}{2}} \leq \left(\sum_{i=1}^{n}\sum_{j=1}^{n}(\sum_{k=1}^{n} a_{ik}^2)\cdot(\sum_{k=1}^{n} b_{kj}^2)\right)^{\frac{1}{2}} =$$

Since the quantity $(\sum_{k=1}^{n} a_{ik}^2)$ is j-independent, we can treat it as a constant coefficient with respect to the summation over all j and factor it out

$$= \left(\sum_{i=1}^{n}(\sum_{k=1}^{n} a_{ik}^2)\cdot\sum_{j=1}^{n}(\sum_{k=1}^{n} b_{kj}^2)\right)^{\frac{1}{2}}$$

We do the same "trick", as we now factor out the term $\sum_{j=1}^{n}(\sum_{k=1}^{n} b_{kj}^2)$ which is i-independent

$$= \left(\sum_{j=1}^{n}(\sum_{k=1}^{n} b_{kj}^2) \quad \cdot \quad \sum_{i=1}^{n}(\sum_{k=1}^{n} a_{ik}^2)\right)^{\frac{1}{2}}$$

$$= \left(\sum_{j=1}^{n}\sum_{k=1}^{n} b_{kj}^2 \quad \cdot \quad \sum_{i=1}^{n}\sum_{k=1}^{n} a_{ik}^2\right)^{\frac{1}{2}}$$

$$= \left(\sum_{k=1}^{n}\sum_{j=1}^{n} b_{kj}^2 \quad \cdot \quad \sum_{i=1}^{n}\sum_{k=1}^{n} a_{ik}^2\right)^{\frac{1}{2}}$$

Since we have separated each sum to have different quantities involved in its computation, the index of summation does not play any particular role and we can freely change it (even use the same symbols as indices)

$$= \left(\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}^2 \quad \cdot \quad \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^2\right)^{\frac{1}{2}}$$

$$= \left(\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^2 \quad \cdot \quad \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}^2\right)^{\frac{1}{2}}$$

$$= \left(\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^2\right)^{\frac{1}{2}} \cdot \left(\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}^2\right)^{\frac{1}{2}}$$

$\Rightarrow ||A \diamond B||_F \leq ||A||_F \cdot ||B||_F$ □

recall that an equivalent norm to the Euclidean one is the p-norm.

Likewise we can define the more general case $L^{p,q}$-norm.

Definition A.28 ($L^{p,q}$-norm $||\cdot||_{L^{p,q}}$)**.**

Let $A \in \mathbb{R}^{n\times m}$ with

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}$$

then

$$||A||_F := \left(\sum_{i=1}^{n} \left(\sum_{j=1}^{m} |a_{ij}|^p \right)^{\frac{q}{p}} \right)^{\frac{1}{q}}$$

We continue with another matrix norm, which will be useful in estimating the integral over a ball (L^1 norm) of the Hessian of a Lipshitz, convex real-valued function.

Definition A.29 ($2,2$ norm)**.**

Let A be a matrix in $\mathbb{R}^{n\times n}$, then we define its $2,2$-norm as follows:

$$\|A\|_{2,2} := \sup_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_2}$$

where $Ax \in \mathbb{R}^{n\times 1}$ with x viewed as a column vector in $\mathbb{R}^{n\times 1}$ for the matrix multiplication to be well-defined

And Ax, x also viewed as their transpose counterparts, i.e. row vectors in $\mathbb{R}^{1\times n}$, in order to then take their Euclidean norm.

Both $\|\cdot\|_2$ norms are the same standard, Euclidean vector norm.

Proposition A.7 ($\|\cdot\|_{2,2}$ is a norm)**.**

The above map $A \mapsto \|A\|_{2,2}$ is indeed a norm.

Proof.

Let $A \in \mathbb{R}^{n\times n}$

i) Obviously, $\|A\|_{2,2} \geq 0$

ii)

$$\|A\|_{2,2} = 0 \Rightarrow A = 0$$

Indeed, by definition we have

$$\|A\|_{2,2} = \sup_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2}$$

Hence, for all non-zero vector x

$$\frac{\|Ax\|_2}{\|x\|_2} \leq \|A\|_{2,2}$$

Let $\|A\|_{2,2} = 0$, then for every $x \neq 0$

$$\frac{\|Ax\|_2}{\|x\|_2} \leq 0$$

since $\|Ax\|_2, \|x\|_2 > 0$ for every non-zero vector x

$$\frac{\|Ax\|_2}{\|x\|_2} = 0$$

$$\Rightarrow \|Ax\|_2 = 0$$

$$\xrightarrow[\text{a norm}]{\|\cdot\|_2 \text{ is}} Ax = 0$$

because the last equality is true for all $x \neq 0$, we get:

$$A = 0$$

iii) The triangle inequality holds true

Indeed, let $A, B \in \mathbb{R}^{n \times n}$, then:

$$\|A + B\|_{2,2} = \sup_{x \neq 0} \frac{\|(A + B)x\|_2}{\|x\|_2}$$

$$= \sup_{x \neq 0} \frac{\|Ax + Bx\|_2}{\|x\|_2}$$

Since, $\|\cdot\|_2$ is a norm, its subadditivity "tells" us that:

$$\|Ax + Bx\|_2 \leq \|Ax\|_2 + \|Bx\|_2$$

Thus,

$$\|A+B\|_{2,2} \leq \sup_{x\neq 0} \frac{\|Ax\|_2 + \|Bx\|_2}{\|x\|_2}$$

$$= \sup_{x\neq 0} \left(\frac{\|Ax\|_2}{\|x\|_2} + \frac{\|Bx\|_2}{\|x\|_2} \right)$$

Because $\sup_x (f+g) \leq \sup_x f + \sup_x g$, we have that:

$$\|A+B\|_{2,2} \leq \sup_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_2} + \sup_{x\neq 0} \frac{\|Bx\|_2}{\|x\|_2}$$

$$= \|A\|_{2,2} + \|B\|_{2,2}$$

□

The $\| \ \|_{2,2}$ matrix norm enjoys a useful relation.

Proposition A.8 (matrix norm and eigenvalues)**.**

Let $A \in \mathbb{R}^{n\times n}$ be a real and symmetric matrix, then

$$\|A\|_{2,2} = \max_{i\in T(n)} |\lambda_i|$$

where λ_i are its eigenvalues.

Proof.

Since A is symmetric and real, it has an orthonormal basis consisting of eigenvectrors v_i, $i \in T(n)$ with λ_i being their respactive dicrete eigenvalues i.e.

$$Av_i = \lambda_i v_i$$

Thus, every vector x can be written as a unique linear combination of v_i.

$$x = \sum_{i=1}^{n} c_i v_i$$

with $c_i \in \mathbb{R}$

So, due to the linearity of A viewed as a linear mapping

$$Ax = A\left(\sum_{i=1}^{n} c_i v_i\right)$$

$$= \sum_{i=1}^{n} c_i A v_i$$

Let $\langle \cdot\,, \cdot \rangle$ be the standard inner product, then using its linearity and the fact that the basis $\{\, v_i \mid i \in T(n) \,\}$ is orthonormal we get:

$$\begin{aligned}
\|Ax\|_2^2 &= \langle \sum_{i=1}^{n} c_i A v_i \,, \sum_{j=1}^{n} c_j A v_j \rangle \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} c_i c_j \langle A v_i, A v_j \rangle \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} c_i c_j \langle \lambda_i v_i, \lambda_j v_j \rangle \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} c_i c_j \lambda_i \lambda_j \langle v_i, v_j \rangle \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} c_i c_j \lambda_i \lambda_j \delta_{ij} \\
&= \sum_{i=1}^{n} c_i^2 \lambda_i^2
\end{aligned}$$

where δ_{ij} is the delta of Kronecker

$$\delta_{ij} = \begin{cases} 1 & , i = j \\ 0 & , i \neq j \end{cases}$$

And similarly,

$$\|x\|_2^2 = \sum_{i=1}^{n} c_i^2$$

Thus, for every non zero vector x

$$\begin{aligned}
\frac{\|Ax\|_2^2}{\|x\|_2^2} &= \frac{\sum\limits_{i=1}^{n} c_i^2 \lambda_i^2}{\sum\limits_{i=1}^{n} c_i^2} \\
&\leq \frac{\sum\limits_{i=1}^{n} c_i^2 \max\limits_{i \in T(n)}{}^2 \lambda_i}{\sum\limits_{i=1}^{n} c_i^2} \\
&= \frac{\max\limits_{i \in T(n)}{}^2 \lambda_i \sum\limits_{i=1}^{n} c_i^2}{\sum\limits_{i=1}^{n} c_i^2}
\end{aligned}$$

$$= \max_{i \in T(n)}{}^2 \lambda_i$$

Since $\|\cdot\|_2 \geq 0$ (every norm is non-negative) squaring out we get:

$$\frac{\|Ax\|_2}{\|x\|_2} = \max_{i \in T(n)} |\lambda_i|$$

where the right hand side has no dependence on x, so taking the supremum:

$$\|A\|_{2,2} \leq \max_{i \in T(n)} |\lambda_i|$$

For the opposite inequality, we have:

$$\begin{aligned} \|Av_i\|_2^2 &= \langle Av_i, Av_i \rangle \\ &= \langle \lambda_i v_i, \lambda_i v_i \rangle \\ &= \lambda_i^2 \langle v_i, v_i \rangle \\ &= \lambda_i^2 \|v_i\|_2^2 \end{aligned}$$

So, (the eigenvectors are non zero vectors):

$$|\lambda_i| = \frac{\|Av_i\|_2}{\|v_i\|_2}$$

$$\leq \sup_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2}$$

$$= \|A\|_{2,2}$$

Thus,

$$\|A\|_{2,2} \geq \max_{i \in T(n)} |\lambda_i|$$

□

A.3 Convexity

We present some basic facts concerning the notion of convexity.

Definition A.30 (Convex set)**.**

We call a set $C \subseteq \mathbb{R}^n$ convex $:\iff$

$$\forall x, y \in C \quad \forall \lambda \in (0,1) \ \text{ we have that } \ \lambda x + (1-\lambda) y \in C$$

Definition A.31 (Convex combination)**.**

If $\lambda_i \geq 0$ with $\sum_{i=1}^{m} \lambda_i = 1$ we call the sum

$$\sum_{i=1}^{m} \lambda_i x_i$$

a convex combination of the points x_i

Remark. The previously introduced convex combination depends on the given points $x_i \in \mathbb{R}^n$.

Definition A.32 (Convex hull)**.**

We define the convex hull of a set $A \subseteq \mathbb{R}^n$
conv(A):= $\{ \sum_{i=1}^{m} \lambda_i x_i \mid m \in \mathbb{N},\ \lambda_i \geq 0, x_i \in A \ \ \forall i \in T(m) \ \text{ and } \ \sum_{i=1}^{m} \lambda_i = 1 \}$

Remark. In all of the above $\mathbb{R}^n$ can be replaced by a vector space V.

Definition A.33 (Convex function on a convex set)**.**

Let $f : C \to \mathbb{R} \cup \{+\infty\}$,
C convex $\subseteq \mathbb{R}^n$. We call the function f convex $:\Longleftrightarrow \ \forall x, y \in C \quad \forall \lambda \in (0,1)$ we have that $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$

Remark. In the above definition we need the convexity of C, so that $\forall \lambda \in (0,1)$ $\lambda x + (1-\lambda)y$ lies in C thus the expression $f(\lambda x + (1-\lambda)y)$ has meaning.

Definition A.34 (Convex function on an open set)**.**

Let $f : U \to \mathbb{R} \cup \{+\infty\}$, U open $\subseteq \mathbb{R}^n$. We call the function f convex $:\Longleftrightarrow$ There is an expansion $\tilde{f} : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, which is convex.

Remark. The convexity of $\tilde{f}$ that DefinitionA.34 states is that of DefinitionA.33 i.e. $\tilde{f}$ is a convex function on the convex set $\mathbb{R}^n$

Clarification A.34.1.
The calculations involving infinity are subject to the usual laws governing computations with the "quantity" of infinity i.e.

Proposition A.9.

The term convex function is well defined, since for a convex and open[1] $S \subseteq \mathbb{R}^n$ the DefinitionA.33 is equivalent to the DefinitionA.34.

Proof

[1]Such a set exists, for example the open "box" $(0,1)^n$. In fact, there are plenty of them, infinitely many, the sets $(a,b)^n \ \forall a, b \in \mathbb{R}$ with $a \neq b$.

Let S convex and open $\subseteq \mathbb{R}^n$ and $f : S \to \mathbb{R} \cup \{+\infty\}$

($\Rightarrow$) Let us assume that f is convex by the standards of DefinitionA.33 then

$\forall x, y \in S \quad \forall \lambda \in (0,1)$ we have that $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$

We define $\tilde{f} : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ with $\tilde{f}(x) := \begin{cases} f(x) & x \in S \\ +\infty & x \notin S \end{cases}$ thus $\tilde{f}$ is an expansion of f.

We will now show that $\tilde{f}$ is convex on $\mathbb{R}^n$, in the sense of DefinitionA.33.
Let $x, y \in \mathbb{R}^n$ and $\lambda \in (0,1)$, we discern the following four cases:

i) if $x, y \in S$ then $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$ and $\lambda x + (1-\lambda)y \in S$ because S is convex, thus we have $\tilde{f}(\lambda x + (1-\lambda)y) \leq \lambda \tilde{f}(x) + (1-\lambda)\tilde{f}(y)$

ii) if $x \notin S$, $y \in S$[2] and $\lambda x + (1-\lambda)y \in S$ then $\tilde{f}(\lambda x + (1-\lambda)y) =$ $= f(\lambda x + (1-\lambda)y) \leq +\infty = \lambda \tilde{f}(x) + (1-\lambda)\tilde{f}(y)$ because $\lambda > 0$ and $\tilde{f}(x) = +\infty$ and $(1-\lambda)\tilde{f}(y) \in (-\infty, +\infty]$

iii) if $x \notin S$, $y \in S$ and $\lambda x + (1-\lambda)y \notin S$ then by the same reasoning $\tilde{f}(\lambda x + (1-\lambda)y) = +\infty = \lambda \tilde{f}(x) + (1-\lambda)\tilde{f}(y)$

iv) if $x \notin S$, $y \notin S$ and $\lambda x + (1-\lambda)y \notin S$ then we have accordingly $\tilde{f}(\lambda x + (1-\lambda)y) = +\infty = \lambda \tilde{f}(x) + (1-\lambda)\tilde{f}(y)$

Thus we have shown that f is convex by the standards of DefinitionA.34

($\Leftarrow$) Converesely, let us assume that f is convex by the standards of DefinitionA.34 then there is an expansion $\tilde{f} : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, which is convex in the sense of DefinitionA.33, that is $\forall x, y \in \mathbb{R}^n$ and $\forall \lambda \in (0,1)$ we have that: $\tilde{f}(\lambda x + (1-\lambda)y) \leq \lambda \tilde{f}(x) + (1-\lambda)\tilde{f}(y)$ (1)

Let $x, y \in S \xRightarrow{\text{S convex}} \lambda x + (1-\lambda)y \in S \xRightarrow[(1)]{\tilde{f} = f \text{on S}} f(\lambda x + (1-\lambda)y) \leq \leq \lambda f(x) + (1-\lambda)f(y)$

□

Proposition A.10.

For every $A \subseteq \mathbb{R}^n$ the equality below holds
$conv(A) = \cap \{C \subseteq \mathbb{R}^n \mid C \supseteq A \text{ and } C \text{ convex}\}$

Corollary A.10.1. The convex hull of a set A is the smallest convex set containing A

[2]Similarly, the same result holds true if $x \in S$ and $y \notin S$.

A.3.1 Legendre transform and the subdifferential

Definition A.35 (Legendre transform)**.**

Let a function $f : C \to \mathbb{R}$ where
C convex $\subseteq \mathbb{R}^n$, we then define the Legendre transform f^* as follows

$$f^*(p) := \sup_{x \in C} (\langle p, x\rangle - f(x)) \quad , \quad p \in C^*$$

$$C^* := \{\, p \in \mathbb{R}^n \mid \sup_{x \in C} (\langle p, x\rangle - f(x)) < +\infty \,\}$$

Clarification A.35.1.

Often, the independent variable p is also denoted x^*. But we will stick with this notation (at least for the definition) for historical reasons rooted in analytic mechanics.

Proposition A.11.

$f = f^{**}$ Theorem1.11 (Fenchel-Moreau) Brezis Functional Analysis [9]

Definition A.36 (subdifferential at a point)**.**

Let a function $f : U \to \mathbb{R}$ with U open and convex $\subseteq \mathbb{R}^n$. We define the sub-differential of f at the point $x \in U$ as the set:

$$df(x) = \{z \in \mathbb{R}^n \mid \forall y \in U \;\; f(y) - f(x) \geq \langle z, y - x\rangle\,\}$$

Definition A.37 (subdifferential at a set)**.**

Let f as above and $S \subseteq U$ then

$$df(S) := \bigcup_{x \in S} df(x)$$

Proposition A.12. If f is convex, then the subdifferential is non empty at every point in its domain.

Remark.
If f is not convex, then $df(x)$ can be the empty set. Even $df(S)$ for every S can be the empty set.
Example. $f(x) = -\|x\|^2$

Proposition A.13.

Let f be a convex function, then the following holds
f is differentiable at x. $\iff$ f has a unique subdifferential at x.

Corollary A.13.1.

Whenever f is differentiable we have that:

$$df(x) = \{\nabla f(x)\}$$

For a proof look at [33] page 242 theorem 25.1.

Proposition A.14.

Let U open $\subseteq \mathbb{R}^n$ and $f : U \to \mathbb{R}$ a convex function, then

$$z \in df(x) \iff x \in df^*(z)$$

Proof.

$(\Rightarrow)$ If $z \in df(x) \Rightarrow f(y) - f(x) \geq \langle z, y - x\rangle \quad \forall y \in U$

$$\begin{aligned} &\Rightarrow \langle z, x\rangle - f(x) \geq \langle z, y\rangle - f(y) \quad \forall y \in U \\ &\Rightarrow \sup_{y\in U} (\langle z, y\rangle - f(y)) = \langle z, x\rangle - f(x) \\ &\Rightarrow f^*(z) = \langle z, x\rangle - f(x) \\ &\Rightarrow f(x) = \langle z, x\rangle - f^*(z) \;\; (1) \end{aligned}$$

$$\begin{aligned} \text{Thus } \forall y^* \in U^* \;\; f^*(y^*) &= \sup_{x*\in U} (\langle y^*, x^*\rangle - f(x^*)) \\ &\overset{x\in U}{\geq} \langle y^*, x\rangle - f(x) \\ &\overset{(1)}{\geq} \langle y^*, x\rangle - \langle z, x\rangle + f^*(z) \end{aligned}$$

$$\begin{aligned} \Rightarrow f^*(y^*) - f^*(z) &= \langle y^* - z, x\rangle \\ &= \langle x, y^* - z\rangle \end{aligned}$$

$\Rightarrow x \in df^*(z)$

$(\Leftarrow)$ Conversely, since $f = f^{**}$, all we have to do is follow the same steps □

A.4 Measure Theory

Definition A.38 (σ-algebra)**.** Let X be a set and $\mathcal{A} \subseteq \mathcal{P}(X)$ a collection of subsets of X. We call $\mathcal{A}$ a σ-algebra if the next three conditions are met:

$$\mathcal{A} \neq \varnothing \tag{A.4.1}$$

$$A^c \in \mathcal{A} \quad \forall A \in \mathcal{A} \tag{A.4.2}$$

$$\bigcup_{n=1}^{+\infty} A_n \in \mathcal{A} \quad \forall \big(A_n\big)_{n\in\mathbb{N}} \subseteq \mathcal{A} \tag{A.4.3}$$

Clarification A.38.1. $\mathcal{P}(X)$ denotes the power set of X i.e. $\mathcal{P}(X) := \{S \mid S \subseteq X\}$ In general, a collection (sometimes also called a family) is a set containing sets. In our case the collection $\mathcal{A}$ contains subsets of X. From now on, we will usually denote a collection/family of sets using calligraphic letters.

Proposition A.15. Equivalent definitions result if we replace the according requirements by whichever of the following:
$\varnothing \in \mathcal{A}$ (interchangeable with first condition)
$X \in \mathcal{A}$ (interchangeable with first condition)
$\bigcap_{n=1}^{+\infty} A_n \in \mathcal{A} \ \ \forall \big(A_n\big)_{n\in\mathbb{N}} \subseteq \mathcal{A}$ (interchangeable with third condition)

Proposition A.16. From our definition it is obvious that if $\mathcal{A}$ is a σ-algebra then $A \setminus B \in \mathcal{A} \ \ \forall A, B \in \mathcal{A}$ and $\bigcap_{n=1}^{+\infty} A_n \in \mathcal{A} \ \ \forall \big(A_n\big)_{n\in\mathbb{N}} \subseteq \mathcal{A}$

Proposition A.17. If $\big(\mathcal{A}_i\big)_{i\in I}$ are σ-algebras then $\bigcap_{i\in I} \mathcal{A}$ is a σ-algebra.

Proposition A.18. $\forall \mathcal{E} \subseteq \mathcal{P}(X) \ \exists! \ \sigma$-algebra $\mathcal{A} : \mathcal{A}$ is the minimum σ-algebra containing $\mathcal{E}$.

Definition A.39 ($\sigma(\mathcal{E})$)**.** We call the above unique σ-algebra the σ-algebra produced by the collection $\mathcal{E}$ and we denote it $\sigma(\mathcal{E})$.

Definition A.40 (Borel sets)**.** Let $\big(X, \tau\big)$ be a topological space, we define $\mathcal{B}(X) := \sigma(\tau)$. We call this σ-algebra the Borel σ-algebra of X and the sets contained in it the Borel subsets of X.

Remark. $\mathcal{B}(\mathbb{R}^n) = \sigma\big(\{\text{open subsets of } \mathbb{R}^n\}\big)$

Proposition A.19. $\mathcal{B}(\mathbb{R}^n) = \sigma\big(\{\text{closed subsets of } \mathbb{R}^n\}\big) = \sigma(\mathcal{E}_1) = \sigma(\mathcal{E}_2)$ where

$$\mathcal{E}_1 = \left\{ \prod_{i=1}^{n} [b_i, +\infty) \mid b_i \in \mathbb{R} \ \ \forall i \in T(n) \right\}$$

$$\mathcal{E}_2 = \left\{ \prod_{i=1}^{n} (a_i, b_i] \mid a_i < b_i \ , \ a_i, b_i \in \mathbb{R} \ \ \forall i \in T(n) \right\}$$

Definition A.41 (Measure)**.** Let X be a set and $\mathcal{A}$ a σ-algebra on X. We call a function $\mu : \mathcal{A} \to [0, +\infty]$ measure on $(X, \mathcal{A}) :\Longleftrightarrow$

$\mu(\varnothing) = 0$ and

$\forall \big(A_n\big)_{n\in\mathbb{N}} \subseteq \mathcal{A}$ sequence of two by two disjoint sets, (countably additive) we have that

$$\mu\left(\bigcup_{n=1}^{+\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$$

Clarification A.41.1. We call $(X, \mathcal{A})$ measurable space and $(X, \mathcal{A}, \mu)$ measure space.

Proposition A.20. If μ, ν are measures on $(X, \mathcal{A})$ and $a \in \mathbb{R}$ then
$\mu + \nu$ and $|a|\mu$ are also measures.

Clarification A.41.2. We define $(\mu + \nu)(A) := \mu(A) + \nu(A)$ and $(a\mu)(A) := a\mu(A)$ $\forall A \in \mathcal{A}$

Proposition A.21. If $(X, \mathcal{A}, \mu)$ is a measure space, then $\forall A, B \in \mathcal{A}$
$A \subseteq B \Rightarrow \mu(A) \leq \mu(B)$
$\mu(A) < +\infty \Rightarrow \mu(B \setminus A) = \mu(B) - \mu(A)$

Proposition A.22. If $(X, \mathcal{A}, \mu)$ is a measure space, then $\forall \big(A_n\big)_{n \in \mathbb{N}} \subseteq \mathcal{A}$

$$\mu\left(\bigcup_{n=1}^{+\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu(A_n)$$

(countable subadditivity).

Definition A.42 (Types of Measures)**.** Let $(X, \mathcal{A}, \mu)$ be a measure space, then we call the measure μ

i) finite if $\mu(X) < +\infty$

ii) probability measure if $\mu(X) = 1$

A.4.1 The pushforward measure

Definition A.43 ($(\mathcal{A}, \mathcal{B})$-measurable function)**.** Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ two measurable spaces and a function $f : X \to Y$. We call the function f
$(\mathcal{A}, \mathcal{B})$-measurable $:\iff \forall B \in \mathcal{B} \;\; f^{-1}(B) \in \mathcal{A}$

Definition A.44 (Pushforward measure)**.** Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be two measurable spaces and a $(\mathcal{A}, \mathcal{B})$-measurable function $f : X \to Y$. If μ is a measure on $(X, \mathcal{A})$, then we define the pushforward measure ν on $(Y, \mathcal{B})$ as follows:
$\nu : \mathcal{B} \to [0, +\infty] \;\; \nu(B) := \mu\left(f^{-1}(B)\right)$

Remark. We denote the pushforward measure as $f_{\#}\mu = \mu \circ f^{-1}$

Proposition A.23. Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be two measurable spaces and a $(\mathcal{A}, \mathcal{B})$-measurable function $f : X \to Y$. If μ is a measure on $(X, \mathcal{A})$ and $g : Y \to \tilde{\mathbb{R}}$ a measurable function then

$$\int_B g \, df_{\#}\mu = \int_{f^{-1}(B)} g \circ f \, d\mu \tag{A.4.4}$$

A.4.2 Absolute continuity of measures

Definition A.45 (absolute continuity $\ll$ of measures)**.**

Let $(X, \mathcal{A})$ be a measurable space and μ,ν be two measures in it. We say that ν is absolutely continuous with respect to ν and we write $\nu \ll \mu :\iff$

$$\forall A \in \mathcal{A} \quad \mu(A) = 0 \Rightarrow \nu(A) = 0$$

Remark. We also say that ν is dominated by μ.

Proposition A.24.

Let $(X, \mathcal{A})$ be a measurable space and μ,ν be two measures in it, then

$$\nu \ll \mu \iff \forall \epsilon > 0 \ \ \exists \delta > 0 \ \ \forall A \in \mathcal{A} \ \ \mu(A) < \delta \Rightarrow \nu(A) < \epsilon \tag{A.4.5}$$

Theorem A.1 (Radon-Nikodym on finite measures)**.**

Let $(X, \mathcal{A})$ be a measurable space and μ,ν be two finite measures in it such as $\nu \ll \mu$, then

$$\exists! \mu\text{-a.e. measurable function } f : X \to [0, +\infty) \text{ with } \nu(A) = \int_A f \, d\mu \tag{A.4.6}$$

Remark. The above function f we will call density of the measure ν with respect to μ.

Proposition A.25.

Let $(X, \mathcal{A})$ be a measurable space and μ,ν be two finite measures in it such $\nu \ll \mu$ and f the unique function of TheoremA.1, then

$$\int g \, d\nu = \int g \cdot f \, d\mu \ \ \forall \text{ measurable } g : X \to [0, +\infty] \tag{A.4.7}$$

Proposition A.26.

Assume that $\mu = f_{\#} dx$ where f is $(\mathcal{A}, \mathcal{A})$-measurable and a non-singular (non-degenerate) map i.e. its pre-image (inverse image) preserves null (negligible) sets

$$f^{-1}(A) = 0 \ \ \forall A \in \mathcal{A} : l^n(A) = 0$$

then

$$\mu \ll dx$$

As found in Benamou-Brenier [7] equation (21)

Proposition A.27 (continuous functions are pair measurable)**.**

Let $f : (X, \mathcal{B}(X)) \to (X, \mathcal{B}(X))$ be a continuous function where $\mathcal{B}(X)$ denotes the Borel σ-algebra defined by a topology of X

then f is $(\mathcal{B}(X), \mathcal{B}(X))$-measurable

Proof. We define □

A.5 Weak derivative and Sobolev spaces

Here we will mention some definitions and results, mostly from [18] about weak derivatives and Sobolev spaces.

Definition A.46 (Weak derivative)**.**

Let $u : \mathbb{R}^n \to \mathbb{R}$ be an L^1_{loc} real-valued function and a be a multi-index, then we call $D^a_w u : \mathbb{R}^n \to \mathbb{R}$ the a-th order weak derivative of u iff:

$$\int_{\mathbb{R}^n} u D^a \varphi \, dx = (-1)^{|a|} \int_{\mathbb{R}^n} \varphi D^a_w u \, dx$$

Proposition A.28 (Weak derivative is a.e. unique)**.**

If the weak derivative (of any order) of u exists, then it is uniquely defined up to a set of zero Lebesgue measure (this means that it differs from the other function only in a set with Lebesgue measure zero)

Definition A.47 (Sobolev space $W^{k,p}$)**.**

Let $p \in [1, +\infty]$ and $k \in N_0$, then we define:

$$W^{k,p}(\mathbb{R}^n) := \{ \ f \in L^1_{loc}(\mathbb{R}^n) \ \mid \ \forall \text{ multi-index } a : |a| \leq k \ \ \exists D^a_w f \in L^p(\mathbb{R}^n) \ \}$$

The k, p Sobolev space consists of all locally summable scalar (real-valued) functions $f : \mathbb{R}^n \to \mathbb{R}$ such that for every multiindex a of order less or equal to k the weak derivatives $D^a_w f$ exist and belong in L^p.

Definition A.48 (Sobolev space norm)**.**

Let f be a function belonging in $W^{k,p}(\mathbb{R}^n)$, then we define its $W^{k,p}$ (Sobolev) norm as:

$$\|f\|_{W^{k,p}(\mathbb{R}^n)} = \begin{cases} \left(\displaystyle\sum_{\substack{a \text{ multi} \\ \text{index}}}^{|a| \leq k} \int_{\mathbb{R}^n} |D^a_w f|^p \, dx \right)^{\frac{1}{p}} & p \in [1, \infty) \\ \displaystyle\sum_{\substack{a \text{ multi} \\ \text{index}}}^{|a| \leq k} \text{ess sup}_{\mathbb{R}^n} |D^a_w f| & p = \infty \end{cases}$$

Proposition A.29 (Sobolev space is Banach)**.**

The Sobolev space $W^{k,p}$ with its respective norm $\|\cdot\|_{W^{k,p}}$ is a Banach space

A.6 About the torus $\mathbb{T}^2$

Now we are going to give the definition of the two-dimensional torus and introduce the norm we will use on it.

Definition A.49 ($\mathbb{T}^2$ equivalence relation)**.** Let $X = \mathbb{R}^2$ we then define the equivalence relation $\sim$ as follows:

$$x \sim y \iff x - y \in \mathbb{Z}^2$$

We use this equivalence relation to define the torus as the quotient set of X by $\sim$

Definition A.50.

$$\mathbb{T}^2 \equiv \mathbb{R}^2/\mathbb{Z}^2 := X/_\sim$$

Clarification A.50.1. It is useful to recall that the quotient set is defined with the help of the notion of the equivalence class $X/_\sim := \{ \, [x] \ | \ x \in X \, \}$, where $[x] := \{ \, s \in X \ | \ s \sim x \, \}$

Remark. Notice that the two-dimensional torus is nothing more than the plane $\mathbb{R}^2$ "split" in squares with vertices two consecutive points on the grid defined by the lattice of the integers $\mathbb{Z}^2$.

On torus we define a new distance, in this way we will be able to "count" using "only" the points lying in the set $[0,1]^d$. This along with the previous remark is the reason we consider the integrals calculated on torus to be over the set $[0,1]^d$.

Definition A.51 (distance on torus)**.** Let $[x], [y] \in \mathbb{T}^d$ we define their distance as:

$$d([x],[y]) := \sup_{p \in \mathbb{Z}^d} ||x - y + p||_2$$

A.7 Useful propositions

Proposition A.30. Let $x, y \in \big(\mathbb{R}^n, \langle \cdot\,, \cdot \rangle\big)$, then the following inequality holds true:

$$2\langle x, y\rangle \leq ||x||^2 + ||y||^2$$

where $|| \cdot ||$ is the norm induced by the inner product

Proof. $\langle x - y, x - y\rangle = ||x - y||^2 \geq 0$ and using the properties of inner product we have

$$\langle x - y, x - y\rangle =$$

$$
\begin{aligned}
&= \langle x, x-y \rangle - \langle y, x-y \rangle \\
&= \langle x, x \rangle - \langle x, y \rangle - \langle y, x \rangle + \langle y, y \rangle \\
&= ||x||^2 - \langle x, y \rangle - \langle x, y \rangle + ||y||^2 \\
&= ||x||^2 - 2\langle x, y \rangle + ||y||^2
\end{aligned}
$$

So, the desired inequality is proven. □

Remark. The space $\mathbb{R}^n$ can be replaced with any vector space and the same result still holds.

Proposition A.31 (Liebniz Integral rule on measure spaces)**.**
Let $(X, \mathcal{A}, \mu)$ be a measure space, I an interval of the real numbers and $f : X \times I \to \mathbb{R}$ be a function with the following properties:

i) The map $x \mapsto f(x,t)$ belongs to $L^1(\mu)$ $\forall\, t \in I$

ii) The map $t \mapsto f(x,t)$ is differentiable for almost all $x \in X$
We denote its time derivative as $\partial_t f_t(x) \equiv \frac{\partial}{\partial t} f(x,t)$

iii) $\exists$ an $L^1(\mu)$ function $h : X \to \mathbb{R}_0^+$ such that $|\partial_t f_t(x)| \leq h(x)$
for μ−a.e. x and $\forall\, t \in I$

Then $\partial_t f_t \in L^1(\mu)$ $\forall\, t \in I$ and the function $t \mapsto \displaystyle\int f(x,t)\, d\mu$ is differentiable with derivative

$$\partial_t \int f(x,t)\, d\mu = \int \partial_t f_t(x)\, d\mu$$

[27] page 142 differentiation lemma

Proposition A.32 (integral zero implies f zero a.e.)**.**
Let $(X, \mathcal{A}, \mu)$ be a measure space and $f, g : X \to \tilde{\mathbb{R}}_0^+ := [0, +\infty]$ then:

i) $f \overset{\text{a.e.}}{=} g \Rightarrow \displaystyle\int f\, d\mu = \int g\, d\mu$

ii) $f \overset{\text{a.e.}}{=} 0 \Leftrightarrow \displaystyle\int f\, d\mu = 0$

Proposition A.33 (non-negative linear combination of convex is convex)**.** Let f, g : convex $C\mathbb{R}^n \to \mathbb{R}$ be two convex functions and $a, b \geq 0$ two non-negative constants, then $h := af + bg$ is convex too.

Proof.

If $a = 0$ or $b = 0$, then the result holds true by simply multiplying the inequality of convexity for the respective function with the other constant.

Remark.
In the case where the other constant is zero as well the linear combination equals the constant function zero, which is trivially convex (the inequality is satisfied as an equality. The same result holds true if we multiply convexity's inequality with zero).

If $a \neq 0$ and $b \neq 0$ then since f, g are convex we have that:

$$\forall \lambda \in (0,1) \text{ and } \forall x, y \in C \begin{cases} f\big(\lambda x + (1-\lambda)y\big) \leq \lambda f(x) + (1-\lambda)f(y) \\ g\big(\lambda x + (1-\lambda)y\big) \leq \lambda g(x) + (1-\lambda)g(y) \end{cases}$$

Multiplying with the positive numbers a, b we get:

$$\overset{a,b>0}{\Rightarrow} \begin{cases} af\big(\lambda x + (1-\lambda)y\big) \leq \lambda af(x) + (1-\lambda)af(y) \\ bg\big(\lambda x + (1-\lambda)y\big) \leq \lambda bg(x) + (1-\lambda)bg(y) \end{cases}$$

Adding each hand-side of the two inequalities above we have:

$$af\big(\lambda x + (1-\lambda)y\big) + bg\big(\lambda x + (1-\lambda)y\big)$$

$$\leq$$

$$\lambda af(x) + (1-\lambda)af(y) + \lambda bg(x) + (1-\lambda)bg(y)$$

$\Rightarrow h\big(\lambda x + (1-\lambda)y\big) \leq \lambda h(x) + (1-\lambda)h(y)$ □

Proposition A.34 (every norm is convex)**.**
Let $||\cdot||$ be a norm on a vector space X, then it is convex.

Remark.
Note that it makes sense to examine convexity on a vector space, since for every $x, y \in X$ and $k, l \in \mathbb{R}$ by definition $kx + ly$ belongs to X. Thus a vector space can be viewed as a convex set.

Proof. This result is immediate by the triangle inequality of a norm. Indeed $\forall \lambda \in (0,1)$

$$\|\lambda x + (1-\lambda)y\| \leq \|\lambda x\| + \|(1-\lambda)y\| \underset{1-\lambda>0}{\overset{\lambda\geq 0}{=}} \lambda\|x\| + (1-\lambda)\|y\|$$

Defining the function $f(x) := ||x||$, for all $x \in X$ we have showed that $f\big(\lambda x + (1-\lambda)y\big) \leq \lambda f(x) + (1-\lambda)f(y)$ □

Proposition A.35 (every natural power of the norm is convex)**.**
Let $||\cdot||$ be a norm on a vector space X, then the function $||\cdot||^m$ is convex $\forall m \in \mathbb{N}$

Proof. Let $m \in \mathbb{N}$ we define the two following functions:

$$f : \mathbb{R}_0^+ \to \mathbb{R}_0^+ \text{ with } f(s) = s^m$$

$$g : X \to \mathbb{R}_0^+ \text{ with } g(x) = \|x\|$$

Then f is convex and increasing, and g is convex as well.

Remark.
$\mathbb{R}_0^+$ is convex, hence it makes sense to talk about convexity.

So, for $x, y \in X$ and $\lambda \in (0,1)$ because g is convex we have that:

$$g\big(\lambda x + (1-\lambda)y\big) \leq \lambda g(x) + (1-\lambda)g(y)$$

$$\xrightarrow[g \geq 0]{f \nearrow} f\Big(g\big(\lambda x + (1-\lambda)y\big)\Big) \leq f\big(\lambda g(x) + (1-\lambda)g(y)\big)$$

Since f is also convex and $g(x), g(y) \in \mathbb{R}_0^+$ we have that:

$$f\big(\lambda g(x) + (1-\lambda)g(y)\big) \leq \lambda f\left(g(x)\right) + (1-\lambda) f\left(g(x)\right)$$

Thus we have showed that:

$$f\Big(g\big(\lambda x + (1-\lambda)y\big)\Big) \leq \lambda f\left(g(x)\right) + (1-\lambda) f\left(g(x)\right)$$

That is $f \circ g$ is convex.

Noticing that $\left(f \circ g\right)(x) = f\left(g(x)\right) = f(\|x\|) = \|x\|^m$, this concludes the proof. □

Proposition A.36 (chain rule)**.**
Let f : open $U \subseteq \mathbb{R}^n \to \mathbb{R}^k$ and g : open $V \subseteq \mathbb{R}^k \to \mathbb{R}^m$ be two functions such as $f(U) \subseteq V$ (meaning that their composition can be defined on all U).
If f is differentiable at x and g is differentiable at $f(x)$ then for their composition we have:

$$g \circ f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m \text{ is differentiable at x}$$

$$D\big(g \circ f\big)(x) = Dg\big(f(x)\big) \diamond Df(x)$$

Clarification A.51.1.
By omitting the argument x (like we usually do) the above chain rule can be rewritten as:

$$D(g \circ f) = Dg(f) \diamond Df$$

In our case, we have inserted in one more variable (time t), which we seperate from the spatial variable x. The next result is an immediate application of the chain rule.

Corollary A.36.1.
Let $f : \mathbb{R}^{n+1} \to \mathbb{R}^n$ and $g : \mathbb{R}^n \to \mathbb{R}$ be two functions.

We denote a point of $\mathbb{R}^{n+1}$ as (x,t).

We also denote D the derivative with respect to (x,t)

$$D_x \text{ the derivative with respect to } x$$

$$\text{and } \partial_t \text{ the derivative with respect to } t$$

If f, g are differentiable (in their whole domains), then we have:

$$\partial_t(g \circ f) = D_x g(f) \diamond \partial_t f$$

Proof.

In the vector $(x_1, x_2, \ldots, x_n, x_{n+1})$ of $\mathbb{R}^{n+1}$ we have chosen to seperate the last variable x_{n+1} and denote it $t \in \mathbb{R}$, from the other variables $x_1, x_2, \ldots, x_n$ consisting the vector $x \in \mathbb{R}^n$.

Thus, instead or writting D as $\left(\dfrac{\partial}{\partial x_1}, \dfrac{\partial}{\partial x_2}, \ldots, \dfrac{\partial}{\partial x_n}, \dfrac{\partial}{\partial x_{n+1}}\right)$

$$\text{we write } D = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}, \frac{\partial}{\partial t}\right)$$

which we abbreviate as:

$$D = \big(\partial_1, \partial_2, \ldots, \partial_n, \partial_t\big)$$

So, with the notation D_x in mind we also have:

$$D_x = \big(\partial_1, \ldots, \partial_n\big) \text{ that implies } D = \big(D_x\,, \partial_t\big)$$

The composition $g \circ f : \mathbb{R}^{n+1} \to \mathbb{R}$ is well-defined and chain rule implies that:

$$D\big(g \circ f\big)(x,t) = Dg\big(f(x,t)\big) \diamond Df(x,t)$$

Since $g \circ f : \mathbb{R}^{n+1} \to \mathbb{R}$ we have $D\big(g \circ f\big) \in \mathbb{R}^{1 \times (n+1)}$ with

$$D\big(g \circ f\big)(x,t) = \Big(\partial_1\big(g \circ f\big)(x,t), \ldots, \partial_n\big(g \circ f\big)(x,t), \partial_t\big(g \circ f\big)(x,t)\Big)$$

Since $g : \mathbb{R}^n \to \mathbb{R}$ we have $Dg\big(f(x,t)\big) \in \mathbb{R}^{1 \times n}$ with

$$Dg\big(f(x,t)\big) = \Big(\partial_1 g\big(f(x,t)\big), \ldots, \partial_n g\big(f(x,t)\big)\Big)$$

Since $f : \mathbb{R}^{n+1} \to \mathbb{R}^n$, let $f = (f_1, \ldots, f_n)^T$ we have $Df(x,t) \in \mathbb{R}^{n \times (n+1)}$ with

$$Df(x,t) = \begin{pmatrix} \partial_1 f_1(x,t) & \partial_2 f_1(x,t) & \cdots & \partial_n f_1(x,t) & \partial_t f_1(x,t) \\ \partial_1 f_2(x,t) & \partial_2 f_2(x,t) & \cdots & \partial_n f_2(x,t) & \partial_t f_2(x,t) \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \partial_1 f_n(x,t) & \partial_2 f_n(x,t) & \cdots & \partial_n f_n(x,t) & \partial_t f_n(x,t) \end{pmatrix}$$

Thus, while omitting (x,t) the chain rule reads:

$$\big(\partial_1(g\circ f),\dots,\partial_n(g\circ f),\partial_t(g\circ f)\big) =$$

$$= \big(\partial_1 g(f),\dots,\partial_n g(f)\big)\diamond\begin{pmatrix}\partial_1 f_1 & \partial_2 f_1 & \cdots & \partial_n f_1 & \partial_t f_1\\ \partial_1 f_2 & \partial_2 f_2 & \cdots & \partial_n f_2 & \partial_t f_2\\ \vdots & \ddots & \ddots & \ddots & \vdots\\ \vdots & \ddots & \ddots & \ddots & \vdots\\ \partial_1 f_n & \partial_2 f_n & \cdots & \partial_n f_n & \partial_t f_n\end{pmatrix}$$

$$= \left(\sum_{i=1}^n \partial_i g(f)\cdot\partial_1 f_i, \sum_{i=1}^n \partial_i g(f)\cdot\partial_2 f_i,\dots,\sum_{i=1}^n \partial_i g(f)\cdot\partial_n f_i,\sum_{i=1}^n \partial_i g(f)\cdot\partial_t f_i\right)$$

Hence,

$$\begin{aligned}\partial_t(g\circ f) &= \sum_{i=1}^n \partial_i g(f)\cdot\partial_t f_i\\ &= \big(\partial_1 g(f),\dots,\partial_n g(f)\big)\diamond\begin{pmatrix}\partial_t f_1\\ \vdots\\ \partial_t f_n\end{pmatrix}\\ &= D_x g(f)\diamond\partial_t f\end{aligned}$$

and the proof is completed □

Proposition A.37 (identity of material derivative)**.**

For any smooth function $f:\mathbb{R}^2\times[0,+\infty)\to\mathbb{R}^2$ the following holds true:

$$\partial_t\Big(f\big(X(t),t\big)\Big) = \partial_t f(x,t) + \langle u,\nabla\rangle f(x,t)$$

where $x = X(t)$

Proof. Indeed,

Let $f:\mathbb{R}^2\times[0,+\infty)\to\mathbb{R}^2$ with $f(x,t) = \big(f_1(x,t),f_2(x,t)\big) =: (f_1,f_2)(x,t)$

Since $X:[0,+\infty)\to\mathbb{R}^2$ the composition $f\big(X(t),t\big)$ is well defined.

We set $g := (X, Id)$, where $t\mapsto\big(X(t),t\big) = \big(X_1(t),X_2(t),t\big)$

Thus, $f\big(X(t),t\big)$ can be written as $(f\circ g)(t)$

Then, the chain rule implies that:

$$\partial_t(f\circ g)(t) = Df\big(g(t)\big)\diamond Dg(t)$$

where

$$Df\big(g(t)\big) = \begin{pmatrix} \partial_1 f_1\big(X(t),t\big) & \partial_2 f_1\big(X(t),t\big) & \partial_t f_1\big(X(t),t\big) \\ \partial_1 f_2\big(X(t),t\big) & \partial_2 f_2\big(X(t),t\big) & \partial_t f_2\big(X(t),t\big) \end{pmatrix}$$

and

$$Dg(t) = \begin{pmatrix} \partial_t X_1(t) \\ \partial_t X_2(t) \\ \partial_t t \end{pmatrix} = \begin{pmatrix} u_1\big(X(t),t\big) \\ u_2\big(X(t),t\big) \\ 1 \end{pmatrix}$$

because $\partial_t X(t) = u\big(X(t),t\big)$ where $u = (u_1, u_2)$

So, we get that $\partial_t (f \circ g)(t)$ equals

$$\begin{pmatrix} \sum_{i=1}^{2} \Big(\partial_i f_1\big(X(t),t\big) \cdot u_i\big(X(t),t\big) \Big) + \partial_t f_1\big(X(t),t\big) \\ \\ \sum_{i=1}^{2} \Big(\partial_i f_2\big(X(t),t\big) \cdot u_i\big(X(t),t\big) \Big) + \partial_t f_2\big(X(t),t\big) \end{pmatrix}$$

which can be written as:

$$\begin{pmatrix} \sum_{i=1}^{2} \partial_i f_1\big(X(t),t\big) \cdot u_i\big(X(t),t\big) \\ \\ \sum_{i=1}^{2} \partial_i f_2\big(X(t),t\big) \cdot u_i\big(X(t),t\big) \end{pmatrix} + \begin{pmatrix} \partial_t f_1\big(X(t),t\big) \\ \\ \partial_t f_2\big(X(t),t\big) \end{pmatrix}$$

Since $\langle u, \nabla \rangle = \sum_{i=1}^{2} u_i \partial_i$

$$\langle u, \nabla \rangle f = \sum_{i=1}^{2} u_i \partial_i f = \sum_{i=1}^{2} u_i \partial_i (f_1, f_2) \equiv \sum_{i=1}^{2} u_i \partial_i (f_1, f_2)^T$$

Hence, we have:

$$\partial_t (f \circ g)(t) = \langle u, \nabla \rangle f\big(X(t),t\big) + \partial_t f\big(X(t),t\big)$$

that is:

$$\partial_t \Big(f\big(X(t),t\big) \Big) = \partial_t f\big(X(t),t\big) + \langle u, \nabla \rangle f\big(X(t),t\big)$$

substituing $X(t)$ with x on the right hand side we have proved the desired. □

Proposition A.38 (Taylor theorem)**.**
Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a function in C^k (k-times continuously differentiable) and a point

$x_0 \in \mathbb{R}^n$ then
there exists a function $h_{x_0} : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$f(x) = \sum_{\substack{|a| \le k \\ a \text{ multi index}}} \frac{D^a f(x_0)}{|a|!}(x - x_0)^{|a|} + \sum_{\substack{|a| = k \\ a \text{ multi index}}} h_{x_0}(x - x_0)^{|a|}$$

where

$$\lim_{x \to x_0} h_{x_0}(x) = 0$$

A.8 Ordinary differential equations

Let $F : \mathbb{R}^n \times [0, +\infty) \to \mathbb{R}^n$. We consider the initial value problem

$$\begin{cases} \frac{\partial}{\partial t} X(t) = F\left(X(t), t\right) \\ X(0) = x_0 \end{cases} \tag{A.8.1}$$

The reason we chose t to be non-negative in the definition of the function F is solely because the semigeostrophic equations that we study involve time. This specific initial value problem (actually, the most general first order differential equation form) has been studied on many sets and has a rich theory. Here we are going to present only the results that we will need and use for our purposes.

A.8.1 Initial value problem and Lipschitz continuity

It has been proven that (among many other conditions) the Lipschitzianity of the function F (alone) is enough to provide a unique solution existing in an entire interval $[0, b]$

Definition A.52 (K-Lipschitz on product space)**.**

Let $F : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ be a function.

We say that the function $F_t : \mathbb{R}^n \to \mathbb{R}^n$ is K-Lipschitz on $S \subseteq \mathbb{R}^n$ with $S \times \mathbb{R} \subseteq D_F$

$$:\iff \exists K > 0 \ \forall \, (x, t) \text{ and } (y, t) \in S$$

$$||F(x, t) - F(y, t)|| \le K||x - y||$$

Proposition A.39 (Existence of a unique solution to the ivp)**.**

Consider the flow (A.8.1) where $F : \mathbb{R}^n \times [0, +\infty) \to \mathbb{R}^n$ is a continuous function and $F_t : \mathbb{R}^n \to \mathbb{R}^n$ is a K-Lipschitz function, then the initial value problem has a unique solution $X : [0, b] \to \mathbb{R}^n$

A.8.2 The Gronwall lemma

The Gronwall lemma comes in various "shapes and sizes". The one that we are going to use here is a rather elementary version of the inequality, as it assumes (strong/classic) differentiality for every t in the time interval instead of other weaker assumptions.

Proposition A.40 (Gronwall lemma for C)**.**
Let the function $\phi : [0, +\infty) \to \mathbb{R}$ be differentiable for which we have that
$\exists C > 0$ such that $\phi'(t) \leq C\phi(t)$ (where $' = \partial_t$ denotes the derivative) then:

$$\phi(t) \leq e^{Ct}\phi(0)$$

In this special case of Gronwall lemma we can show that a similar result holds true if we replace C with $-C$, just by simply following the exact same proof. Hence, we will state the other lemma as well and we will prove only this one.

Proposition A.41 (Gronwall lemma for $-C$)**.**
Let the function $\phi : [0, +\infty) \to \mathbb{R}$ be differentiable for which we have that
$\exists C > 0$ such that $\phi'(t) \geq -C\phi(t)$ (where $' = \partial_t$ denotes the derivative) then:

$$\phi(t) \geq e^{-Ct}\phi(0)$$

Proof. We define the function

$$f(t) := e^{\int_0^t -C\,dt} = e^{-Ct} > 0 \ \ , \ \ t \in [0, +\infty)$$
$$\Rightarrow f'(t) = -Ce^{-Ct} = -Cf(t)$$

Since f is positive, we can define for all $t \in [0, +\infty)$ the function $\dfrac{\phi}{f}$ for which by the quotient derivative rule we have that:

$$\partial_t \frac{\phi(t)}{f(t)} = \frac{\phi'(t)f(t) - \phi(t)f'(t)}{f^2(t)} \overset{f,f^2>0}{\geq} \frac{-C\phi(t)f(t) - \phi(t)f'(t)}{f^2(t)} =$$

$$= \frac{-C\phi(t)f(t) - \phi(t)(-Cf(t))}{f^2(t)} = \frac{-C\phi(t)f(t) + C\phi(t)f(t)}{f^2(t)} = 0$$

Thus, the function $\dfrac{\phi}{f}$ is non-decreasing (increasing and/or constant). Thus,

$$t \geq 0 \underset{\nearrow}{\overset{\phi/f}{\Longrightarrow}} \frac{\phi(t)}{f(t)} \geq \frac{\phi(0)}{f(0)} \underset{\text{of f}}{\overset{\text{def}}{\Longrightarrow}} \frac{\phi(t)}{e^{-Ct}} \geq \frac{\phi(0)}{e^0} \underset{e^0=1}{\overset{e^{-Ct}>0}{\Longrightarrow}} \phi(t) \geq e^{-Ct}\phi(0)$$

□

Corollary A.41.1 (Gronwall in other intervals)**.**
With the same "technique" we can have the same result in intervals of the form $[0, b]$ for whatever $b > 0$ (since the right endpoint didn't matter in the proof). And we also obtain a similar property (exactly the same inequality) in the interval $[a, +\infty)$ or $[a, b]$ if we replace $\phi(0)$ with $\phi(a)$

Remark.
Notice that in our case ϕ does not need to be non-negative (which is a usual assumption in Gronwall inequalities)

One can obtain similar results if instead of a constant a function of t "makes its appearence" in the inequality.

Bibliography

[1] Ι. Γιαννούλης: Διανυσματική ανάλυση, Κάλλιπος Ανοιχτές Ακαδημαϊκές Εκδόσεις, 2015

[2] Γ. Κουμουλλής, Σ. Νεγρεπόντης: Θεωρία Μέτρου, Εκδόσεις Συμμετρία Αθήνα, 2005

[3] Χ. Γ. Φίλος: Μια εισαγωγή στις διαφορικές εξισώσεις, Πανεπιστήμιο Ιωαννίνων , 2005

[4] Hans. W. Alt: *Linear functional analysis*, Springer London, 2016

[5] Viorel Barbu: *Differential Equations*, Springer Undergraduate Mathematics Series, 2016

[6] J. D. Benamou, Y. Brenier: *A computational Fluid Mechnics solution to the Monge-Kantorovich mass transfer problem*, Numerische Mathematik, vol 84, 375-393, 2000

[7] J. D. Benamou, Y. Brenier: *Weak existence for the Semigeostriphic equations formulated as a coupled Monge-Ampere/Transport problem*, SIAM Journal of Applied Mathematics, vol. 58, 1450-1461, 1998

[8] J. Bergh, J. Lofstrom: *Interpolation spaces-An introduction*, World Publishing Corporation, 1976

[9] H. Brezis: *Functional analysis Sobolev spaces and partial differential equations*, Springer, 2011

[10] L. A. Caffarelli: *The regularity of mappings with a convex potential*, Journal of the American Mathematical Society, vol.5, 99-104, 1992

[11] L. A. Caffarelli: $W^{2,p}$ *Estimates for solutions of the Monge-Ampere equation*, Annals of Mathematics, vol. 131, 135-150, 1990

[12] D. L. Cohn: *Measure theory second edition*, Birkhauser, 2010

[13] D. Cordero-Erausquin: *Sur le transport de measures peridiques*, Academie de Sciences, vol. 329, 199-202, 1999

[14] M. Cullen, R. J .Purser: *A duality pinciple in Semigeostrophic theory*, Journal of Atmospheric Sciences, vol. 44, 3449-3468, 1987

[15] M. Cullen, R.J.Purser: *Properties of the Lagrangian Semigeostrophic equations*, Journal of Atmospheric Siences, vol. 46, 2684–2697, 1989

[16] M. Cullen: *The mathematics of large-scale Atmosphereand and Ocean*, World Scientific, 2021

[17] C. M. Dafermos: *Hyperbolic conservation laws in Continuum Physics fourth edition*, Springer, 2016

[18] L. C. Evans: *Partial differential equations Second edition*, AMS, 2010

[19] L. C. Evans, R. F. Gariepy, *Measure theory and fine properties of functions*, CRC Press, 2015

[20] A. Figalli, G. de Philipis, L. Ambrosio, M. Colombo: *Existance of Eulerian solutions to the Semigeostrophic equations in Physical space; The 2-dimensional periodic case*, Communications in Partial Differential Equations, 2011

[21] A. Figalli, G. de Philipis: *The Monge-Ampere equation and its link to optimal transportation*, AMS, vol. 51, 2013

[22] A. Figalli: *The Monge-Ampere equation and its applications*, European Mathematical Society, 2017

[23] A. Figalli: *Global existance for the Semigeostrophic equations via Sobolev estimates for Monge-Ampere*, CIME Lecture Notes, 2015

[24] J. F. Fournier, R. A. Adams: *Sobolev spaces, Second edition*, Academic Press, 2003

[25] David Gilbarg, Neil S. Trudinger: *Elliptic Partial Differential Equations of Second Order*, Classics in Mathematics, 2001

[26] P. Hartman: *Ordinary differential equations Second edition*, SIAM, 2002

[27] A. Klenke: *Probability theory Second edition*, Springer, 2013

[28] J.L.Lions: *Quelques Methods de Resoluton des Problemes aux Limites Non-Lineaires*, Dunod, 1969

[29] G. Loeper: *A fully nonlinear version of the incompressible Euler equations; The Semigeostrophic system*, SIAM Journal on Mathematical Analysis, vol. 38, 795-823, 2006

[30] G. Loeper: *Uniqueness of the solution to the Vlasov-Poisson system with bounded density*, Journal de Mathematiques Pures et Appliquees, vol.86, 68-79, 2006

[31] Andrew J. Majda, Andrea L. Bertozzi: *Vorticity and Incompressible Flow*, Cambridge Texts in Applied Mathematics, 2001

[32] J. Norbury, M. Cullen: *The use of an energy principle to define Atmospheric balance*, Journal of Meteorological Society of Japan, vol. 64A, 75-82, 1986

[33] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, 1970

[34] H. L. Royden, P. M. Fitzpatrick: *Real analysis 4th edition*, Pearson, 2017

[35] J. Simon: *Compact sets in the space* $L^p(0,T;B)$, Annali di Mathematica pura ed applicata, vol. 146, 65-69, 1986

[36] W. Walter: *Ordinary differential equations*, Springer, 1998