A SEMI DERIVATION LEMMA ON BV FUNCTIONS

FILIPPO SANTAMBROGIO, FRANÇOIS-XAVIER VIALARD

ABSTRACT. This paper presents a proof of a derivation lemma on the space of BV functions. The origin of this question can be found in the context of the image matching in the framework of large deformation diffeomorphisms. To compute the geodesic equations on the space of diffeomorphisms, one needs this result, which also gives the structure of the initial momentum, i.e. the central tool in the Hamiltonian formulation of geodesic equations.

CONTENTS

1. INTRODUCTION

This work arises from the large deformation diffeomorphisms framework. With broad applications in computational anatomy, the starting point of this growing field is the minimization of an energy on a chosen space of diffeomorphisms. The functional which is usually chosen is

$$E(\phi) = D(Id, \phi)^2 + \frac{1}{\sigma^2}\|I_0 \circ \phi^{-1} - I_{targ}\|^2_{L^2}, \tag{1}$$

where ϕ is the diffeomorphism, I_0 the initial function (or image) and I_{targ} the target function , Id is the identity map, σ is a calibration parameter. The distance D is a Riemannian metric on the diffeomorphisms group, coming from the minimization of a certain geodesic distance. For more details about the group and its metric, one can refer to [TY05]. To derive the optimality equations, we have to compute the variation of the functional with respect to small perturbations of the diffeomorphism ϕ. Although the existence result of a minimizer is easily obtained even if the two images are not smooth, it is more difficult to compute derivatives and first-order variations when dealing with discontinuous functions. An attempt to answer this question is developed in [Via08].

To present the derivation result obtained in such a paper, we need to introduce a functional space, based on Lipschitz continuous functions, which is included in SBV (and it will turn out to be a dense subset of SBV). If U is a Lipschitz open domain, we define $Lip^p(U)$ as the set of piecewise Lipschitz functions: we say $f \in Lip^p(U)$ if there exists a finite partition of U in Lipschitz domains $(V_i)_{i=1\dots n}$ such that the restriction $f_{|V_i}$ of f on each V_i is Lipschitz . Obviously, we have $Lip^p(U) \subset SBV(U)$. The following result is the starting point of this paper and is contained in [Via08].

Lemma 1.1. *Let $(f, g) \in Lip^p(U)^2$, X a Lipschitz vector field on $\mathbb{R}^n$ and ϕ_t the associated flow.*

$$J_t = \int_U f \circ \phi_t^{-1}(x)\, g(x) dx,$$

Date: July 20, 2008.
1991 *Mathematics Subject Classification.* Primary: 58B10; Secondary: 26B30, 49J45, 68T10.
Key words and phrases. SBV, Lipschitz domain, derivation lemma.

then the derivation of J_t gives:

$$\partial_{t|t=0+} J_t = \int_U < \nabla f, -X > g dx + \int (f_+ - f_-) \tilde{g} < \nu_f, -X > dH^{n-1}. \quad (2)$$

with $\tilde{g}(x) := \lim_{t \mapsto 0^+} g(\phi_t(x))$ if the limit exists and $\tilde{g}(x) = 0$ if not.

Notice that this theorem only gets a one-sided derivative (the derivative at $t = 0^-$ could be computed exchanging X with $-X$), which is sufficient for variational purposes. This is the reason for the title of this paper, where "semi derivation" stands for the derivative of the functional in one direction only.

We aim to generalize this lemma in two directions. First the goal is to extend the formula to SBV or BV functions. Second we want to give a version of this lemma in order to be able to derive other terms than the square of the L^2 norm for the penalty term in the functional (1). Sometimes, we will work with a C^1 vector field instead of a Lipschitz vector field. This is not restrictive for current applications. On the other hand, Lipschitz assumptions are the least we must ask since without them we could not define the flow ϕ_t.

2. Presenting the main results

2.1. **Statement of the results and notations.** We need some basic properties of BV functions. If g is a BV function, the precise representative of g is defined $\mathcal{H}^{n-1}$ a.e. We denote by (g^+, g^-, ν) the precise representative of g (g^+ and g^- being the upper and lower value at each point and ν the normal vector to the jump set, denoted by J_g, pointing in the direction of the upper value). To make the notations shorter we introduce the algebra B composed of bounded BV functions on $\mathbb{R}^n$ with compact support. If we denote $BV_c(\mathbb{R}^n)$ the subset of functions of compact support in $BV(\mathbb{R}^n)$, then $\mathrm{B} = BV_c(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$. In the definition below, a vector field is an application from $\mathbb{R}^n$ to $\mathbb{R}^n$ without any further assumption.

Definition 1. If X is a vector field on $\mathbb{R}^n$ and g a BV function, we define g_X by $g_X(x) = g(x)$ if $x \notin J_g$. On J_g, we define $\mathcal{H}^{n-1}$ a.e.

- $g_X(x) = g^+(x)$ if $\langle \nu(x), X(x) \rangle > 0$,
- $g_X(x) = g^-(x)$ if $\langle \nu(x), X(x) \rangle < 0$,
- else $\langle \nu(x), X(x) \rangle = 0$ and $g_X(x) = \frac{g^-(x) + g^+(x)}{2}$.

Hence, g_X lies in $BV(\mathbb{R}^n) \times L^1(J_g; \mathcal{H}^{n-1})$.

Remark 1. In order to make use of change of variables formulas, the action by a diffeomorphism ψ is given by

$$(g \circ \psi)_X \circ \psi^{-1} = g_{d\psi(X \circ \psi^{-1})}.$$

Remark that if X is $\mathcal{H}^{n-1}$ measurable (for example for continous vector fields), then g_X is also measurable. In this article this will always be the case, thanks to the Lipschitz or even C^1 assumptions we will use.

The main result of the paper is the following, which we will refer to as "derivation result". Here $\mathcal{S}_1$ and $\mathcal{S}_2$ represent functional spaces that will be precised in the different extensions of the statement (they could be Lip^p, BV, SBV, B...).

To stress its generality, we give its statement in the time-dependent case. Yet, in the whole paper we will only deal with the autonomous case, but a remark will show how to extend the results to time-dependent vector fields. We will consider time dependent vector fields $X(t,x) : \mathbb{R} \times \mathbb{R}^n \mapsto \mathbb{R}^n$ which are continuous in both variables and Lipschitz in x (with, for simplicity a Lipschitz constant which does not depend on t). Mostly we will use the notation $X_t(x) = X(t,x)$ and X_t for the vector field at time t. For such a time dependent Lipschitz vector field, the flow is defined for all time. We will use the notation $(t,x) \in \mathbb{R} \times \mathbb{R}^n \mapsto \phi_t(x) \in \mathbb{R}^n$ for the flow generated by X. In the derivation result, if X is continuous in time, as we can expect, we will get X_0 instead of X. Obviously, we cannot expect such a result to be true if X is not continuous.

Theorem 2.1. *Let X a Lipschitz time dependent vector field and ϕ_t its associated flow. Take $f \in \mathcal{S}_1$ and $g \in \mathcal{S}_2$. We define the functional*

$$J_t(f,g) = \int_{\mathbb{R}^n} f \circ \phi_t^{-1}(x) g(x) dx\,, \tag{3}$$

then we have

$$\partial_{t=0^+} J_t = \int g_{X_0}(x) \langle -X_0(x), \partial f(dx) \rangle, \tag{4}$$

where ∂f stands for the distributional derivative of f, which is a finite vector measure.

The same result will also be extended to more general situation than simply the product fg. A precise statement for a general function $H(f,g)$ is the following.

Theorem 2.2. *Let H be a locally Lipschitz function $H : \mathbb{R}^2 \mapsto \mathbb{R}$ and C^1 in the first variable such that $H(0,0) = 0$, $(f,g) \in \mathrm{B} \times \mathrm{B}$, X a Lipschitz time dependent vector field C^1 in space and ϕ_t its associated flow. We define the functional*

$$J_t(f,g) = \int_{\mathbb{R}^n} H(f \circ \phi_t^{-1}(x), g(x)) dx\,, \tag{5}$$

then, under the additional assumption that H is C^1 in the first variable (the derivative w.r.t. to such a variable being denoted by $\nabla_1 H$), we have

$$\partial_{t=0^+} J_t = \int \langle \partial_1 H(f(x), g_{X_0}(x)), -X_0 \rangle dx\,, \tag{6}$$

where $\partial_1 H(f, g_{X_0})$ is a part of the BV derivative of $H(f,g)$, defined by

$$\partial_1 H(f(x), l) = \nabla_1 H(f(x), l)(\nabla f(x) + D^c f(x)) + j_H(x) \mathcal{H}^{n-1} \llcorner J_f,$$

with $j_H(x) = (H(f^+(x), l) - H(f^-(x), l))\, \nu_f(x)$.

Remark first that since f and g are bounded, we can replace H with $\tilde{H}$ such that $\tilde{H}$ is Lipschitz on $\mathbb{R}^n$ and $H = \tilde{H}$ on $\mathrm{Im}(f) \times \mathrm{Im}(g)$. This allows to deal only with the case where H is globally Lipschitz on $\mathbb{R}^n$.

Back to the functional (1), the attachement term is the square of the L^2 norm. It is of importance for the applications to be able to differentiate other terms and this is the goal of Theorem 2.2. Yet, it could be interesting to cover a larger set of penalty terms. For example, our theorem allows to deal with all the L^p norms for $p > 1$, but it does not include $p = 1$ (even if the result seems true in this case as well).

On the contrary, the function $G^\alpha : (x,y) \mapsto |x-y|^\alpha$ is not locally Lipschitz on $\mathbb{R}^2$ for $\alpha \in]0,1[$. This penalty term cannot be dealt with in a BV framework because the composition of a non Lipschitz function with a BV function may not be a BV function any more; in the case $\alpha < 1$, here there is an example of non differentiability. Take

$$H(x,y) = x^{\frac{1}{2}}, \quad f(t) = \begin{cases} \frac{1}{n^2} & \text{if } t \in [\frac{1}{2n+1}, \frac{1}{2n}], \\ 0 & \text{otherwise,} \end{cases}$$

then we have $J_0(f,f) = 0$ and for any fixed n_0, for t small enough we have

$$J_t(f,f) = \int_0^1 \sqrt{|f(x-t) - f(x)|} dx \geq 2t \sum_{n \leq n_0} \frac{1}{n},$$

which implies $\partial_{t=0^+} J_t = \sum_{n=1}^\infty \frac{1}{n} = +\infty$.

2.2. **Structure of the paper.** The starting point for the paper is the recent work by the second author in [Via08] where Theorem 2.1 is proven for Lipschitz functions on Lipschitz domains. As it is done in [Via08], this may be easily extended, by additivity, to the case of $\mathcal{S}_1 = \mathcal{S}_2 = Lip^p(U)$, the space of piecewise Lipschitz functions, i.e. the functions which are Lipschitz continuous on a finite partition of U, the sets of the partition being Lipschitz domains as well. This is what was presented in Lemma 1.1.

In Section 2 we will prove that this class of functions is dense in the strong BV topology in the space $SBV(U)$. This will be useful for extending the result by approximation. Actually, extending derivative results by approximation is always very delicate and there is in general no hope to succeed if no uniform estimate is shown. This is why weak approximation by regular functions will not be sufficient to prove the result for more general f and g.

Section 3 will present as a first result a suitable uniform estimate of $J_t(f,g) - J_0(f,g)$ (a very similar estimate will be used to prove that we can generalize to the case of continuous time-dependent vector fields). This estimate involves the BV and L^∞ norms of f and g, and justifies the need for strong approximations. All the section will be devoted to extensions of the derivation result thanks to density and approximation. Since we use strong convergence and SBV is a closed subspace of BV we could not hope, by means of this strategy, for more general results than SBV. Yet, there is sort of a duality between BV and L^∞ in this framework and it turns out that uniform approximations may work as well. This allows to present other extended results with continuous functions. In the end, we get the derivative result for functions which are sum of a continuous one and an SBV one. By the way, this raises an interesting question: can we hope for a decomposition result for arbitrary BV functions into the sum of an SBV and a continuous one? Obviously this is true in the one-dimensional case and it is what we exploit in Section 4, where we summarize the main results in dimension one as a consequence of what previously proven in any dimension. Thanks to the uniform approximation technique we may also extend the result to one of most natural frameworks in dimension one: the space of functions admitting right and left limit at any point (which is required to define g_X).

The result is now proven for a wide class of BV or BV+continuous functions in any dimension, and up to now the vector field X has been supposed to be Lipschitz continuous. Yet, there is some possibilities of extending it to arbitrary BV functions through a different technique.

The integral curves of the vector field X actually determine a partition of the space $\mathbb{R}^n$ into one-dimensional slices (this only works for autonomous vector fields) and it is worthwhile trying to prove the general result by slicing, through a suitable trivializing diffeomorphism that can transform these curves into straight lines. This is what is done in Section 5. Yet, we need a slightly stronger assumption on X, which has to be C^1 in space. This is required because a Jacobian factor will appear and we would like to apply the one-dimensional derivation result to a product which will involve this factor: should it only be L^∞ (as it is the case for Lipschitz vector fields X) we could not, while if it is continuous we can, since we know the result in one dimension for all functions admitting right and left limits.

Thus, if we could recover the general BV result in any dimension through the one-dimensional one, what is the reason of Sections 2 and 3 ?

- First: we must notice that the proof in one dimension, up to some notational simplification, cannot be performed (to our knowledge) through techniques really simpler than the ones we presented in those sections;
- Second: the assumptions in Section 3 are stronger as far as f and g are concerned (SBV instead of BV) but weaker on X, since we require less regularity;
- Third: the generalization to time-dependent vector fields is more natural if we do not pass through the trivialization of the flow (since this is possible for autonomous systems only), and it could be performed without Lemma 4.2 in the framework of Section 3.
- Last, but not least: we feel that the results in these sections and especially the density results in Section 2 deserve their own attention and are interesting in themselves. Moreover, the whole framework let some questions on BV functions (such as the decomposition SBV + continuous) arise.

Finally, the last section presents an extension to the framework of Theorem 2.2.

3. Density lemmas

In this section, we prove the density of the set of Lipschitz piecewise functions in the SBV space. We suppose that U is a Lipschitz domain and all the functions we deal with are of compact support in U.

We want to prove that any SBV function g may be approximated strongly in the BV norm by some functions g_n in $Lip^p(U)$ and that the same sequence of function also gives $\mathcal{H}^{d-1}-$a.e. pointwise convergence of the functions $(g_n)_X$ to g_X (for a fixed vector field X). Actually this last point will be stated with almost-everywhere convergence with respect to an arbitrary finite measure, absolutely continuous w.r.t. $\mathcal{H}^{d-1}$, due to metrizability conditions. Anyway, this is sufficient for letting the approximation procedure of the next section work.

If G is a Lipschitz graph in U such that $d(G,\partial U)>0$, we denote by $SBV_G(U)=\{f\in SBV(U)|J(f)\subset G\}$ the set of functions of $SBV(U)$ whose jump set is included in G. We also denote by $SBV_{G,\delta}(U)=\{f\in SBV(U)||D^sf|(U\setminus G)<\delta\}$ the functions in $SBV(U)$ whose jumps occur on G up to small measure jumps.

There are basically two steps in the proof:

(1) $Lip^p(U)$ is dense in $SBV_G(U)$.
(2) $\mathrm{Vect}(SBV_G(U),G)$ is dense in $SBV(U)$.

Lemma 3.1. *Let f be a function in $SBV_G(U)$: there exists a sequence $(u_n)_{n\in\mathbb{N}}\subset Lip^p(U)$ such that*

$$u_n\to f \text{ in } BV \text{ and } (u_n)_X\to f_X \text{ pointwisely } \mathcal{H}^{d-1}-a.e.$$

Moreover, if instead $f\in SBV_{G,\delta}(U)$, then there exists a sequence $(u_n)_{n\in\mathbb{N}}\subset Lip^p(U)$ such that

$$\limsup_n ||f-u_n||_{BV}\le 2\delta \text{ and } (u_n)_X\to f_X \text{ pointwisely } \mathcal{H}^{d-1}-a.e. \text{ outside } Jf\setminus G.$$

Proof. Let us start from the case $f\in SBV_G(U)$.
First, we can find a partition in Lipschitz domains (V_1,V_2) of U such that $G\subset\partial V_1\cap\partial V_2:=\Gamma$. This is an application of the Lipschitz extension theorem: indeed, there exists an extension of the Lipschitz graph on $\mathbb{R}^n$, which gives a partition on $\mathbb{R}^n$ and by restriction to U, a candidate partition. Be careful that this partition may not be a Lipschitz partition (but this can only be the case if ∇G and $\nabla\partial U$ are colinear). The hypothesis $d(G,\partial U)>0$ ensures that a small perturbation in a neighborhood of ∂U of such an extension gives the desired partition.

Consider now $f\in SBV_G(U)$, it gives by restriction $f_i=f_{|V_i}\in W^{1,1}(V_i)$ for $i=1,2$.
From Meyers-Serrin theorem we know that any $W^{1,1}$ function may be approximated in $W^{1,1}$ by means of its convolutions (after performing an extension of the function itself beyond the boundary, which requires Lipschitz assumption on the domain, which is exactly the situation we face here). Moreover, by the continuity of the trace map from $W^{1,1}(\Omega)$ to $L^1(\partial\Omega,\mathcal{H}^{d-1})$, we may infer from the strong convergence in $W^{1,1}$ the strong L^1 convergence on the boundary. Let us perform this convolution separately on the two domains V_1 ad V_2, thus obtaining two sequences of regular functions which do not coincide on the common boundary Γ. Then, we glue the two functions and, with [AFP00], we get functions u_n in $Lip^p(U)\subset SBV(U)$. The BV distance of this function to f may be estimated by the sum of the BV distances in the domains V_1 and V_2 (which are actually $W^{1,1}$ distances, since no singular part of the derivative is involved) and the L^1 distance between the jumps of f and of u_n on Γ, i.e.

$$||(u_n^+-u_n^-)-(f^+-f^-)||_1=||\,|u_n^{(1)}-u_n^{(2)}|-|f^{(1)}-f^{(2)}|\,||_1\le||u_n^{(1)}-f^{(1)}||_1+||u_n^{(2)}-f^{(2)}||_1$$

(the superscripts (1) and (2) at u and f stand for the values on the two different sides of Γ, while the superscripts $+$ and $-$ stand for the upper and the lower of these two values, respectively). Since these last L^1 norms converge to zero as the $W^{1,1}$ norm goes to zero, we get the that u_n converges to f in the BV distance. Obviously, u_n belongs to $Lip^p(U)$ since it is composed of two regular functions glued together on a Lipschitz boundary. Moreover, from the general theory on BV functions (see Evans and Gariepy, [EG92]), we know that the convolution regularizations also converge pointwisely $\mathcal{H}^{d-1}-$a.e. outside the jump set of the limit function g (and on this set they converge to the average between the upper and lower value of g). This gives $\mathcal{H}^{d-1}-$a.e. convergence of $(u_n)_X$ to g_X outside Γ (since outside Γ the functions u_n and g agree with $(u_n)_X$ and g_X, respectively). The convergence on Γ is easily deduced from the fact that the values $(u_n)_X$ and g_X agree with the values of u_n and g on one side of Γ (up to the points where Γ and X are parallel, where these quantities are not well defined): the boundary values of u_n converge to those of g as a consequence of the L^1 convergence on Γ (with respect to the $(d-1)-$dimensional measure).

Hence the thesis is easily obtained in the first case $f \in SBV_G(U)$.

The general case is considered in the same way: divide U into two domains and consider again convolutions. The only difference lies in the fact that we do not have any more $W^{1,1}$ but BV functions. This means that the convolutions do not converge strongly neither pointwisely. Yet, the pointwise convergence stays true outside Jf, and Lemma 3.2 gives the estimate of the BV distance between a suitable convolution u_n and f. □

Remark that we don't have obtained $Ju_n \subset G$. Yet, we could have slightly modified the proof to get $J(u_n) \subset G$. Anyway, we do not need it in the following.

The following Lemma has been applied in Lemma 3.1 to the case $\Omega = V_1$ and $\Omega = V_2$.

Lemma 3.2. *Let $f \in SBV(\Omega)$ with jump set Jf: then we may obtain by convolution a sequence of smooth function u_n on Ω, such that $\limsup_n |f - u_n|_{BV} \le 2|D^s f|$ and $u_n \to f$ pointwisely $\mathcal{H}^{d-1}-$a.e. outside Jf.*

Proof. The convolution with a sequence of mollifiers ρ_n gives the result. First, remark that we have $\lim_n ||f * \rho_n - f||_{L^1} = 0$, and then:

$$|\partial(f * \rho_n) - \partial f|(U) = |(D^a f) * \rho_n + (D^s f) * \rho_n - \partial f|(U)$$
$$|\partial(f * \rho_n) - \partial f|(U) \le ||(D^a f) * \rho_n - D^a f||_{L^1} + ||(D^s f) * \rho_n||_{L^1} + |D^s f|(U)$$

With the theorem of approximation of BV functions by smooth functions ([EG92]), we know that $\lim_n |\partial(f * \rho_n)|(U) = |\partial f|(U)$. Moreover, $\lim ||(D^a f) * \rho_n - D^a f||_{L^1} = 0$ (thanks to the behaviour of convolutions on L^1 functions), and, as a result of the two preceding assertions, we get $\lim ||(D^s f)*\rho_n||_{L^1} = |D^s f|(U)$. Finally we get

$$\limsup_n |\partial(f * \rho_n) - \partial f|(U) \le \lim_n ||(D^a f) * \rho_n - D^a f||_{L^1} + \lim_n ||(D^s f) * \rho_n||_{L^1} + |D^s f|(U) = 2|D^s f|(U).$$

As in the previous Lemma, we know that the mollified functions $u_n = f * \rho_n$ converge $\mathcal{H}^{d-1}-$a.e. to f outside Jf. □

Now, we can prove the second lemma:

Lemma 3.3. *For any function f in $SBV(U)$ there exists a sequence $(u_n)_n \subset Lip^p(U)$ such that $u_n \to f$ in BV. Moreover, for any finite measure $\mu << \mathcal{H}^{d-1}$ this sequence may be chosen so that $(u_n)_X \to f_X$ pointwisely $\mu-$a.e.*

Proof. We know ([EG92]), that the jump set of f is a countable union of compact Lipschitz graph. Hence, we enumerate the Lipschitz graphs involved: $(G_i)_{i \in \mathbb{N}}$. Take $\delta > 0$: there exists an integer N such that $\sum \int_{G_i:\, i>N} |D^s f| \le \delta$.

We can suppose that $G_i \cap G_j = \emptyset$ for all couples $(i,j) \in [0,N]^2$ with $i \ne j$ (actually, two Lipschitz graphs G and H may be always replaced with a new pair G and $\tilde{H}$, so that $\tilde{H}$ is a finite union of Lipschitz graphs, $G \cap \tilde{H} = \emptyset$ and $\mathcal{H}^{d-1}(\{H \setminus \tilde{H}\})$ is as small as we want).

Take a smooth partition of unity $(\psi_0, \psi_1, \dots, \psi_N)$ such that, for $i \ge 1$, we have $\psi_i = 1$ on G_i and $\psi_i = 0$ on G_k for $k \ne i$, $1 \le k \le N$. Writing $f = \sum_{i=0}^N f\psi_i$, we can restrict our interest to $f_i = f\psi_i$, which is again an SBV function.

Each function f_i has a jump set which is composed by two parts: one part is included in G_i, while the second one has a "small" jump, since $|D^s f_i|(U \setminus G_i) = \int_{U \setminus G_i} \psi_i |D^s f| = \int_{U \setminus \bigcup_{j=1}^N G_j} \psi_i |D^s f|$. Moreover $\sum_{i=0}^N \int_{U \setminus \bigcup_{j=1}^N G_j} \psi_i |D^s f| = \sum_{i=0}^N |D^s f_i|(U \setminus \bigcup_{j=1}^N G_j) \le \delta$.

Thanks to Lemma 3.1, we may approximate each function f_i through functions $u_n^{(i)} \in Lip^p(U)$. By summing together the results (and the functions) we get a sequence of functions u_n in $Lip^p(U)$ such that

$$\limsup_n ||u_n - f||_{BV} \le 2\delta, \qquad \text{and } u_n \to f \text{ pointwisely } \mathcal{H}^{d-1} - a.e. \text{ outside } \bigcup_{j>N} G_j.$$

The result can then be obtained if we take a function u_n from this sequence so that $||u_n - f||_{BV} < 3\delta$ and then we repeat the same construction with smaller values of δ, thus getting a sequence converging to f in BV. Yet, this cannot be performed for the pointwise convergence, because we need a metric to do that. Actually, we know that almost everywhere convergence w.r.t. a measure is a metrizable convergence, provided the measure is finite (or $\sigma-$finite), which is not the case for $\mathcal{H}^{d-1}$. This is why we introduced μ.

If, from the very beginning, we choose N so that $\mathcal{H}^{d-1}\left(\bigcup_{j>N} G_j\right)$ is sufficiently small (and hence $\mu\left(\bigcup_{j>N} G_j\right)$ is sufficiently small as well, say smaller than δ, then we can select a function u_n so that $d(u_n, g) < 2\delta$ (d being for instance the distance in probability $d(f,g) = \inf\{\varepsilon : \mu(\{|f-g| > \varepsilon\}) < \varepsilon\}$) and go on with the same procedure as before. □

4. SBV and continuous functions in any dimension

This Section presents wider and wider generalizations of the derivation result thanks to approximations techniques. The first tool we will use is the following lemma.

Lemma 4.1. *For $t \le t_0$, if $g \in L^\infty(U)$ and $f \in BV(U)$, we have*

$$|J_t(f,g) - J_0(f,g)| \le \int_U |f \circ (\phi_t^{-1}) - f|\,|g|\,dx \le Ct||g||_\infty ||f||_{BV}$$

for a constant C which only depends on the vector field X and on t_0.

On a subset $A \subset U$, the same result is true for the functional $J_t(A; f, g)$:

$$\int_A |f \circ (\phi_t^{-1}) - f|\,|g|\,dx \le Ct||X||_{L^\infty(A_{Ct})}||g||_\infty ||f||_{BV},$$

where A_ε is $\{x \in U : d(x,A) < \varepsilon\}$ and C is again a constant which only depends on the vector field X and on t_0.

Analogously, if on the contrary $f \in L^\infty(U)$ and $g \in BV(U)$, then we have

$$|J_t(f,g) - J_0(f,g)| \le \int_U |f \circ (\phi_t^{-1}) - f|\,|g|\,dx \le Ct||f||_\infty ||g||_{BV}.$$

Proof. Let us start from the case $g \in L^\infty$ and $f \in C^1$. Just consider

$$\int_U |f\circ(\phi_t^{-1}) - f|\,|g|\,dx \le \int_0^t\int_U |\nabla f \circ \phi_s|\,|X \circ \phi_s|\,|g|\,dx\,ds = \int_0^t\int_U |\nabla f|\,|X|\,|g \circ (\phi_s)^{-1}|\,|\mathrm{Jac}(\phi_s)^{-1}|\,dx\,ds$$

and then use the fact that, since for small s the map $(\phi_s)^{-1}$ is close to the identity, the Jacobian $|J(\phi_s)^{-1}|$ is close to one, and hence bounded. Then estimate the second member by

$$||X||_\infty \sup_{s\in[0,t]} ||J(\phi_s)^{-1}||_\infty ||g||_\infty \int |\nabla f|.$$

The proof in the case f $A \subset U$ is similar, since we estimate

$$\int_0^t\int_A |\nabla f \circ \phi_s||X \circ \phi_s|\,|g|\,dx\,ds \le \int_0^t\int_{A_{Ct}} |\nabla f||X|\,|g \circ (\phi_s)^{-1}||\mathrm{Jac}(\phi_s)^{-1}|\,dx\,ds$$

where, in the change of variable, we do not exit the set A_{Ct}, provided $||X||_\infty \le C$. Then we go on with the same estimates.

To pass to the general case $f \in BV(U)$ it is always sufficient to choose a sequence of regular functions $(f_k)_k$ which converges to f in L^1 with the additional property $||f_k||_{BV} \to ||f||_{BV}$ (see [EG92]) and let the previous estimate pass to the limit.

The estimate in the opposite case is only a little bit trickier. Let us perform a change of variables so that

$$J_t(f,g) = \int_U f \circ \phi_t^{-1} g dx = \int_U fg \circ \phi_t \mathrm{Jac}(\phi_t) dx = \int_U fg \circ \phi_t dx + \int_U fg \circ \phi_t (\mathrm{Jac}(\phi_t) - 1) dx.$$

The difference between first term in the last sum and $J_0(f,g)$ may estimated as before (replacing the vector field X with $-X$) by $Ct||f||_\infty||g||_{BV}$. The second may be estimated, after a new change of variable with bounded Jacobian, by $||f||_\infty||g||_1||\mathrm{Jac}(\phi_t)-1||_\infty$. The thesis is obtained as far as one notices $|\mathrm{Jac}(\phi_t)-1|\leq Ct$ and $||g||_1\leq||g||_{BV}$. □

In a very analogous way we can prove the following reduction lemma, that we referred to in the introductory sections. Its aim is proving that the result is true for time-dependent vector fields (continous in t), if it is true for the autonomous case.

Lemma 4.2. *If ψ_t denotes the usual flow associated to a time dependent vector field $X=X(t,x)$ (that we suppose continuous in time and C^1 in space) and ϕ_t the flow associated to the (constant in time) vector field $X_0=X(0,\cdot)$, then we have*

$$\int_U \left(f\circ(\phi_t^{-1})-f\circ(\psi_t^{-1})\right)\,g\,dx=o(t).$$

Proof. Set $\chi_t:=\phi_t^{-1}\circ\psi_t$ and let Y be the vector field hidden behind the flow χ_t, i.e.

$$\dot\chi_t=Y(t,\chi_t)$$

(such a field Y exists since χ_t is a diffeomorphism). As usual, we estimate the difference by

$$\begin{aligned}\int_U \left(f\circ(\phi_t^{-1})-f\circ(\psi_t^{-1})\right)\,g\,dx &= \int_U \left(f\circ(\phi_t^{-1}\circ\psi_t)-f\right)\,g\circ\psi_t\,|\mathrm{Jac}(\psi_t)|\,dx\\ &\leq \int_U\int_0^t |\nabla f\circ\chi_s||Y\circ\chi_s||g\circ\psi_t|\,\mathrm{Jac}(\psi_t)\,dx\\ &\leq t||g||_{L^\infty}||\mathrm{Jac}\psi||_{L^\infty}||\mathrm{Jac}\chi^{-1}||_{L^\infty}||f||_{BV}||Y||_{L^\infty(U\times[0,t])}.\end{aligned}$$

The only thing that we need to conclude is to prove that $||Y||_{L^\infty(U\times[0,t])}\to 0$. Let us look for a while at the regularity of ψ_t and ψ_t: as a consequence of the assumptions on X they are both C^1 functions. Hence $\chi\in C^1$ and $Y\in C^0$. This implies that the condition $Y(0,\cdot)=0$ is sufficient to imply $\lim_{t\to0}||Y||_{L^\infty(U\times[0,t])}=0$. We can compute

$$\dot\chi_t=\frac{\partial(\phi_t^{-1})}{\partial t}+\nabla_x(\phi_t^{-1})\cdot X(t,\psi_t).$$

If we take $t=0$, remembering $\phi_t^{-1}=\phi_{-t}$ (since ϕ is the flow of an autonomous vector field) and $\phi_0=id$, we get

$$Y(0,\cdot)=\dot\chi_{t|t=0}=-X_0+Id\cdot X(0,\cdot)=0.$$

and this concludes the proof. □

Lemma 4.3. *If $g\in Lip^p(U)$ and $f\in SBV(U)$, then the derivation result is true.*

Proof. Take a sequence $f_k\in Lip^p(U)$ converging to f in $BV(U)$. Write $f=f_k+r_k$ with $||r_k||_{BV}\to 0$. By linearity, we have $J_t(f,g)=J_t(f_k,g)+J_t(r_k,g)$. Since we know the derivative of the first term and we can estimate the second we have

$$\limsup_{t\to0}\frac{J_t(f,g)-J_0(f,g)}{t}\leq\int_U\langle\partial f_k,-X\rangle\,g_X\,dx+C||g||_\infty||r_k||_{BV}$$

and analogously

$$\liminf_{t\to0}\frac{J_t(f,g)-J_0(f,g)}{t}\geq\int_U\langle\partial f_k,X\rangle\,g_X\,dx-C||g||_\infty||r_k||_{BV}.$$

When we let k go to infinity, the last term of both inequalities vanishes, while for the first we have the convergence

$$\int_U\langle\partial f_k,X\rangle\,g_X\,dx\to\int_U\langle\partial f,X\rangle\,g_X\,dx.$$

This convergence is a consequence of the fact that the derivatives of f_k strongly converge as measures to the derivative of f, and this is enough to integrate it against any measurable bounded function (such as Xg_X).

These estimates finally imply the existence of the derivative of $J_t(f,g)$ and its equality with the desired formula. □

Lemma 4.4. *If $g \in SBV(U) \cap L^\infty(U)$ and $f \in SBV(U) \cap L^\infty(U)$, then the derivation result is true.*

Proof. The proof of this last step is very similar to the previous one. But now f will be fixed and we will strongly approximate g in BV by a sequence g_k of functions of $Lip^p(U)$. As before, we need to check two facts; first that the remainders $(J_t(f, g - g_k) - J_0(f, g - g_k))/t$ may be made as small as we want; second that the quantities $\int_U \partial f \cdot X(g_k)_X$ actually converge to $\int_U \partial f \cdot X g_X$. For the first point, we will use the second estimate in Lemma 4.1. For the second, we only need pointwise convergence $|\partial f|-$ a.e. of $(g_k)_X$ to g_X. The fact that we can satify this condition by properly choosing the sequence is ensured by Lemma 3.3. □

Lemma 4.5. *If $g \in SBV(U) \cap L^\infty(U)$ and $f \in SBV(U)$, then the derivation result is true.*

Proof. In this case we fix g and approximate f by a sequence of bounded BV functions. Take $f_k = H_k \circ f$, for some functions H_k satisfying: $H_k(z) = z$ for $|z| \leq k-1$, $0 \leq H_k' \leq 1; |H_k(z)| \leq k \vee |z|$, $H_k \in C^1$.

We will use the chain rule for BV functions (see Ambrosio Fusco Pallara, Theorem 3.96) so that we can easily get $f_k \to f$ in $BV(U)$. This is enough to use the same arguments as in Lemma 4.3 and extend the result to such a framework. □

Up to now we have provided results only in the case where both functions f and g belong to $SBV(U)$ and the main reason lies in the fact that we used strong approximation in BV by means of functions in $Lip^p(U)$ and those functions all belong to $SBV(U)$, which is a closed subset of $BV(U)$. Anyway, by means of different and much simpler method it is possible to handle the case where f is a generic BV function and g is continuous.

Lemma 4.6. *Suppose that $f \in BV(U)$ and $g \in C^\infty(U)$, then the derivation result is true.*

Proof. We have, by change of variables

$$\begin{aligned}
J_t(g) &= \int_U f(x)\, g \circ \phi_t(x)\, \mathrm{Jac}(\phi_t) d\mu(x), \\
\partial_{|t=0} J_t(g) &= \int_U f \langle \nabla g, X \rangle + g\, (\nabla \cdot X)\, dx, \\
\partial_{|t=0} J_t(g) &= \int_U [f \nabla \cdot (gX)]\, dx, \\
\partial_{|t=0} J_t(g) &= -\int \langle \partial f,\, gX \rangle. \quad \square
\end{aligned}$$

This proves, as we may have expected, that a sufficiently strong regularity in one of the two functions can compensate weaker assumptions on the other. We go on extending the result to functions g which are only continuous.

Lemma 4.7. *Suppose that $f \in BV(U)$ and $g \in C^0(U)$, then the derivation result is true.*

Proof. Approximate uniformly g by a sequence g_k of C^∞ functions. As usual, one only needs to manage the remainder (and this is done thanks to Lemma 4.1 since we have $||g_k - g||_\infty \to 0$), and to have convergence of the derivative terms. This last convergence is true since ∂f is a fixed finite measure and hence uniform convergence is sufficient (actually pointwise dominated convergence would have been enough). □

A possible interest of the extension to a BV-continuous setting lies in the following question: is it true that all BV functions may be decomposed as the sum of a continuous and an SBV function? this is true in dimension one and it has a priori no hope to be true in higher dimension where even $W^{1,1}$ do not need to be continuous. Yet, possible discontinuities due to this kind of behaviour could be inserted in the SBV part. This question is obviously interesting in itself and does not seem being treated in the literature. We thank Giovanni Alberti for a brief discussion on the subject.

Here in this context a decomposition $f = f_c + f_s$ such as this one, even if with no additional property mimicking what happens in dimension one (i.e. without requiring neither estimates on the dependence of f_c and f_s on f nor any properties on the derivatives of the two addends), would allow to generalize the derivation result to any pair (f,g) of BV functions, thanks to the following statement.

Theorem 4.8. *Suppose that $f = f_c + f_s$ and $g = g_c + g_s$ with $f_c, g_c \in BV(U) \cap C^0(U)$ and $f_s, g_s \in SBV(U)$, then the derivation result is true.*

Proof. We have to manage and derive four terms, and precisely $J_t(f_c, g_c)$, $J_t(f_c, g_s)$, $J_t(f_s, g_c)$ and $J_t(f_s, g_s)$. The thesis is proven if we prove, for all of them, that the derivative is given by the $\int_U \langle \partial f, X\rangle\, g_X$, being f and g replaced by their continuous or SBV parts.

The term $J_t(f_s, g_s)$ does not give any problem since its derivative has been the object of the proof of Lemma 4.4. The terms $J_t(f_s, g_c)$ and $J_t(f_c, g_c)$ can be dealt with thanks to Lemma 4.7. We need to look at the term $J_t(f_c, g_s)$ which does not fit into the the frameworks we considered so far. The idea is to switch the roles of f_c and g_s.

Notice

$$\begin{aligned} J_t(f_c, g_s) &= \int_U f_c \circ \phi_t^{-1} g_s dx = \int_U f_c\, g_s \circ \phi_t \operatorname{Jac}(\phi_t) dx \\ &= \int_U f_c\, g_s \circ \phi_t dx + \int_U f_c g_s (\operatorname{Jac}(\phi_t) - 1) dx + \int_U f_c (g_s \circ \phi_t - g_s)(\operatorname{Jac}(\phi_t) - 1) dx. \end{aligned}$$

The first term in the last sum may be derived as usual, replacing the vector field X with $-X$, thanks to the result in Lemma 4.7. Its derivative gives $\int_U \langle \partial g_s, f_c\rangle$. The second may be derived pointwisely since the only part depending on t is the Jacobian, and the derivative is $\int_U f_c\, g_s\, (\nabla \cdot X)\, dx$. For the third term we have (thanks to the second estimate in Lemma 4.1)

$$\left| \int_U f_c (g_s \circ \phi_t - g_s)(\operatorname{Jac}(\phi_t) - 1) dx \right| \le Ct \int_U |f_c|\, |g_s \circ \phi_t - g_s|\, dx \le Ct^2 ||f_c||_\infty ||g_s||_{BV},$$

and thus its contribution to the derivative at $t = 0$ is zero.

This means that we have

$$\frac{d}{dt} J_t(f_c, g_s) = \int_U \langle \partial g_s, f_c\rangle + \int_U f_c\, g_s\, (\nabla \cdot X)\,.$$

We want this sum to equal $\int \langle \partial f_c, -X\rangle (g_s)_X = \int \langle \partial f_c, -X\rangle g_s$ (replacing $(g_s)_X$ with g_s itself is allowed since ∂f_c does not give mass to $(d-1)-$dimensional sets, as a consequence of f_c being continuous). To get this equality it is sufficient to integrate by part and use the product rules for the derivatives of $g_s\, X$ (a product of a BV function and a Lipschitz vector field). $\square$

Remark 2. The same techniques of the last proofs could be used to prove a statement such as the following: if the derivation result is true for $f \in BV$ and g belonging to a certain functional class $\mathcal{S}$, then the same result stays true if g belongs to the closure of $\mathcal{S}$ for the uniform convergence.

Yet, we will not develop this remark here in this section, since to be precise we should check that the meaning of g_X stays well-defined for the functions that we obtain as uniform limits of BV functions. Actually, g_X was defined for BV functions g and, by uniform convergence, we may go out of this functional space. We will develop the same concept in Section 4 in the one-dimensional case, which is easier to treat and simpler definitions may be given.

5. The one-dimensional case

The results from Section 3 may obviously be applied to the case of dimension one, which is actually much simpler. The main peculiarity is that in dimension one it is true and easy that every BV function is the sum of SBV and continuous functions.

In this section U will be a compact interval of $\mathbb{R}$ and we will always replace a BV function on U with its precise representative. Equivalently, we will use the definition of one-dimensional bounded variation functions through the total variation as a supremum over partitions (and not as a distributional object). This means that one-dimensional BV functions will be defined pointwisely, not only almost everywhere.

For a function $g \in BV(U)$, an equivalent definition for g_X will be the following:

$$g_X(x_0) = \begin{cases} \lim_{x\to x_0^+} g(x) & \text{if } X(x_0) > 0, \\ \lim_{x\to x_0^-} g(x) & \text{if } X(x_0) < 0, \\ g(x_0) & \text{if } X(x_0) = 0. \end{cases} \tag{7}$$

Theorem 5.1. *Suppose that f, $g \in BV(U)$ ($U \subset \mathbb{R}$): then the derivation result is true.*

Proof. It is sufficient to apply Theorem 4.8 since any BV function in one dimension is the sum of an SBV function and a continuous one. This can be easily obtained by taking the cumulative distribution function of the Cantor part of the derivative, which will be a continuous BV function whose derivative is exactly the Cantor part of the original derivative. The remainder is an SBV function by construction. □

In this one-dimensional setting, we will be able to extend furtherly the result thanks to Remark 2. The following closure result is quite known.

Theorem 5.2. *The set of functions on U which are uniform limits of BV functions is the following vector space $RL(U)$:*

$$RL(U) = \{f : U \to \mathbb{R} \,:\, f \text{ admits right and left limits at every point of } U\}\,.$$

Proof. In dimension one the property of admitting limits on the two sides is satisfied by any BV function (since, if x_h converges monotonely to x then $\sum_h |f(x_{h+1}) - f(x_h)|$ is finite and $f(x_h)$ is a Cauchy sequence).

This property is preserved by uniform convergence and hence any function which is a uniform limit of BV functions belongs to $RL(U)$.

On the other hand, if $f \in RL(U)$, for every $\varepsilon > 0$ and every $x \in U$, there exists a neighbourhood $V_{\varepsilon,x} =]a_{\varepsilon,x}, b_{\varepsilon,x}[$ of x such that the oscillations of f on $]a_{\varepsilon,x}, x[$ and on $]x, b_{\varepsilon,x}[$ are smaller than ε. By compactness one can cover U by a finite number of these intervals and hence, by considering the left parts of these intervals, the right ones and the central points, one covers U by a finite number of intervals (or points) where the oscillation of f is smaller than ε.

A function g constant on each one of these intervals may be built so that $||f - g||_\infty < \varepsilon$ and g is BV. This proves the density of BV in $RL(U)$. □

Corollary 5.3. *Suppose that $f \in BV(U)$ and $g \in RL(U)$: then the derivation result is true.*

Proof. It is sufficient to proceed by approximations as in Lemmas 4.3 and 4.4. the function f belongs to $BV(U)$ and is fixed and g will be approximated by a sequence $(g_k)_k$ of BV functions thanks to Theorem 5.2. □

It is interesting to notice that the space $RL(U)$ is exactly the natural space for the function g, since it is the largest space where the definition of g_X given in (7) makes sense.

6. Reduction to the one-dimensional case

We turn now to the proof of the derivation results for BV functions in dimension d, thanks to a one-dimensional reduction technique. The proof of the reduction is based on two arguments:

- The box flow theorem,
- the one dimensional restriction of BV functions.

We begin with a remark which states that the property is local. Thus we only have to focus on the open set of non-equilibrium points of the vector field (i.e. where the vector field is non zero).

Remark 3. If for each $x \in U$ such that $X(x) \neq 0$, there exists a neighborhood V of x such that the result is true for $J_t(V; f, g)$, with

$$J_t(V; f, g) = \int_V f \circ \phi_t^{-1}(x) g(x) dx,$$

then the result is true.

To prove this remark, consider $Z_\varepsilon = \{y \in U; |X(y)| \geq \varepsilon\}$ which is a compact subset (we denote by Z_ε^c its complement in U). By means of a finite covering the result is true on Z_ε. Then, we need to control what happens on Z_ε^c: this is easy by lemma 4.1 for $t \leq t_0$:

$$\left| \frac{J_t(Z_\varepsilon^c; f, g) - J_0(Z_\varepsilon^c; f, g)}{t} \right| \leq Ct(\varepsilon + L||X||_\infty 2t),$$

where L is the Lipschitz constant of X and we used the fact that, if we move from Z_ε^c no more than $t||X||_\infty$, then the value of $|X|$ does not increase more than $tL||X||_\infty$.

This implies than, when we divide by t and let $t \to 0$, we get

$$\limsup_{t\to 0} \left| \frac{J_t(U; f, g) - J_0(U; f, g)}{t} - \int_U \langle \partial f, -X \rangle g_X \right| \leq \left| \int_U \langle \partial f, -X \rangle g_X - \int_{Z_\varepsilon} \langle \partial f, -X \rangle g_X \right| + C\varepsilon.$$

Letting $\varepsilon \to 0$ we get in the end the result we want, since $\int_{Z_\varepsilon} \partial f \cdot X g_X \to \int_{\{X\neq 0\}} \partial f \cdot X g_X = \int_U \partial f \cdot X g_X$.

It is now sufficient to prove the result in the neighborhood of a point $z_0 \in U$ such that $X(z_0) \neq 0$. In that case, we would like to use the standard reduction of the flow of vector field: up to a C^1 diffeomorphism, we just need to treat the case where $\phi_t(x) = x + t\nu$ in a neighborhood of z_0. It has even been proven in the Lipschitz case by Calcaterra and Boldt ([CB03]). (Notice by the way that, with standard arguments, a time dependent vector field can be viewed as a vector field on $\mathbb{R} \times \mathbb{R}^n$ of the same regularity).

Here a slightly different version of the box-flow theorem is presented.

Theorem 6.1. *If X is a C^1 vector field on $\mathbb{R}^n$ and $x_0 \in \mathbb{R}^n$ is a point such that $X(x_0) \neq 0$, then there exist a neighborhood V of x_0, a neighborhood U of 0, a vector $v \in \mathbb{R}^n$, and a C^1 diffeomorphism $\psi : U \mapsto V$ such that for any $x \in U$ we have*

$$\phi_t \circ \psi(x) = \psi(x + tv) \tag{8}$$

for t such that $x + tv \in U$.

Proof. Since this result is well known, we give here a short proof. Trough a diffeomorphism we can assume that $x_0 = 0$, $X(0) = v = e_1$ where $(e_1, \ldots, e_n)$ a basis of $\mathbb{R}^n$. Defining

$$\psi(x_1, \ldots, x_n) = \phi_{x_1}(0, x_2, \ldots, x_n)\,, \tag{9}$$

we have the result. The regularity of ψ is obtained due to the regularity of the flow ϕ_t which is C^1. $\square$

Deriving in time the equality (8) we have

$$(X \circ \phi_t \circ \psi) = \nabla\psi(x + tv) \cdot v\,. \tag{10}$$

Proposition 6.2. *Let $(f, g) \in BV(U)$, X a C^1 vector field on $\mathbb{R}^n$ and ϕ_t the associated flow. Set*

$$J_t = \int_U f \circ \phi_t^{-1}(x) g(x) dx,$$

then the derivation of J_t gives:

$$\partial_{t|t=0+} J_t = \int < \partial f, -X > g_X\,. \tag{11}$$

Proof. Trivializing the flow with the box-flow theorem, we obtain trough a change of variables in the functional J_t,

$$J_t = \int_{\mathbb{R}^n} f \circ \psi(x - tv)\, g \circ \psi\, \mathrm{Jac}(\psi) dx.$$

The product $h = g \circ \psi\, \mathrm{Jac}(\psi)$ is not a BV function because $\mathrm{Jac}(\psi)$ is only continuous. However, we have extended our results to continuous functions. Set $H = \nu^\perp$, the orthogonal hyperplan to ν, and by using coordinates $x = (x', h)$, write

$$\begin{aligned} J_t - J_0 \quad &= \quad \int_{\mathbb{R}^n} (f \circ \psi(x - t\nu) - f \circ \psi(x)) h(x) dx\,, \\ &= \quad \int_{H=\mathbb{R}^{n-1}} \int_{\mathbb{R}} (f \circ \psi(x' + (h - t)\nu) - f \circ \psi(x' + h\nu)) h(x' + h\nu) dx' dh\,. \end{aligned}$$

We want to derive under the integral in dx'. Set

$$\delta j_t(x) = \int_{\mathbb{R}} \frac{f\circ\psi(x+(h-t)\nu) - f\circ\psi(x+h\nu)}{t} h(x+h\nu)dh.$$

We have

$$|\delta j_t(x)| \le \int_{\mathbb{R}} |\partial(f\circ\psi)^\nu| \|h\|_\infty,$$

$$\int_H \int_{\mathbb{R}} |\partial(f\circ\psi)^\nu| \le |\partial(f\circ\psi)|(\mathbb{R}^n) < +\infty$$

This allows to apply dominated convergence on δj_t. Moreover, since we can apply our derivation result in dimension one, we have

$$\lim_{t\mapsto 0+} \delta j_t(x) = \int_{\mathbb{R}} -\partial(f\circ\psi)^\nu (g\circ\psi)_\nu \mathrm{Jac}(\psi)\,,$$

remark that the definition of $(g\circ\psi)_\nu$ is the one dimensional one. A priori, it may not coincide with the definition on $\mathbb{R}^n$. The theorem (3.108 in [AFP00]) of the continuity of the precise representative tells us that the one dimensional restriction of $(g\circ\psi)_\nu$ gives $\mathcal{H}^{d-1}$-a.e. the same function. Thus

$$\begin{aligned}
\partial_{t=0^+} J_t &= \int_H \int_{\mathbb{R}} -\partial(f\circ\psi)^\nu (g\circ\psi)_\nu \mathrm{Jac}(\psi)\,, \\
&= \int \langle \partial(f\circ\psi), -\nu\rangle (g\circ\psi)_\nu \mathrm{Jac}(\psi)\,, \\
&= \int \langle \partial f, -(\nabla\psi)\circ\psi^{-1}\cdot\nu\rangle (g\circ\psi)_\nu \circ\psi^{-1}\,.
\end{aligned}$$

The last equality comes from Lemma 6.3 below, applied to $\phi = v(g\circ\psi)_\nu \mathrm{Jac}(\psi)$. Using the remark 1 and the equality (10) we have the result:

$$\partial_{t=0^+} J_t = \int \langle \partial f, -X_0\rangle g_{X_0}\,. \qquad \square$$

Lemma 6.3. *If $f \in BV(U)$ and ψ is a diffeomorphism of U, then, for any bounded measurable function $\phi : U \to \mathbb{R}^d$, we have*

$$\int_U \langle \partial(f\circ\psi), \phi\rangle = \int_U \langle \partial f, \frac{(D\psi)\cdot\phi}{\mathrm{Jac}\psi} \circ\psi^{-1}\rangle.$$

Proof. Suppose f and ϕ regular. From the chain rule from regular function we get

$$\int_U \langle \partial(f\circ\psi), \phi\rangle = \int_U \langle (\nabla f)\circ\psi, (D\psi)\cdot\phi\rangle dx.$$

Thanks to the change of variable $x = \psi^{-1}(y)$ we get the desired formula. This is valid for $f \in C^1$ but we can recover the same result for $f \in BV$ by weak approximating f with a sequence of regular functions f_n (notice that in this case $\partial f_n \rightharpoonup \partial f$ as measures and the other factor is a continuous function, if $\phi \in C^0(U;\mathbb{R}^d)$).

Once the result is obtained for ϕ continuous, one can approximate pointwisely any measurable function ϕ by a sequence of continuous functions ϕ_n, and the results stays true for ϕ as well. $\square$

7. Extension to other penalty terms

In this last section, we will extend, for the sake of applications, the derivation result to more general functions $H(f,g)$. Here again, we will only prove that what we have shown so far (i.e. the result for $(x,y) = xy$) is sufficient. We present the easy estimation we need in order to follow with a proof of the reduction based on a density argument.

Let us give a simple definition:

Definition 2. If H is a Lipschitz function in two variables, we denote by

$$\mathrm{Lip}_1(H) = \inf\{M \in \mathbb{R}^+ | \forall (x,x',y) \in \mathbb{R}^3\, |H(x,y) - H(x',y)| \le M|x-x'|_1\}.$$

Exchanging the two variables, we define $\mathrm{Lip}_2(H)$ as well.

By definition, we trivially have for each $i = 1, 2$, $\mathrm{Lip}_i(H) \le \mathrm{Lip}(H)$.

To avoid summability problems and to have to possibility to replace locally Lipschitz functions H with globally Lipschitz ones, we will stick to the case f, $g \in \mathrm{B}$, i.e. we suppose them to be bounded..

Lemma 7.1. *If $g \in L^\infty(\mathbb{R}^n)$ with compact support and $f \in \mathrm{B}$, then if $0 \le t \le t_0$*

$$|J_t(f,g) - J_0(f,g)| \le CtLip_1(H)\|f\|_{BV},$$

for a constant C which only depends on the vector field X and on t_0.

Proof. We will follow again the same ideas as in Lemma 4.1.

Let us start from the case $g \in L^\infty$ and $f \in C^1$. Just consider

$$|J_t(f,g) - J_0(f,g)| \le \mathrm{Lip}_1(H) \int_U |f \circ \phi_t^{-1} - f| dx \le \mathrm{Lip}_1(H) ||X||_\infty \int_0^t \int_U |\nabla(f \circ \phi_s^{-1})| dx\, ds.$$

Then, with a change of variables whose Jacobian is bounded, we get

$$|J_t(f,g) - J_0(f,g)| \le c\mathrm{Lip}_1(H) ||X||_\infty \sup_{s \in [0,t]} ||J(\phi_s)^{-1}||_\infty \int |\nabla f|.$$

Generalizing to $f \in BV$ works the same as in Lemma 4.1. □

Lemma 7.2. *Proving Theorem 2.1 (i.e. the case $H(x,y) = xy$) implies Theorem 2.2.*

Proof. The result is true for all the polynomial functions since B is an algebra and we can easily check that the formula coming from Theorem 2.1 in the case $f^k g^h$ is exactly the one we want.

Remark also that we can suppose that $H(0,y) = 0$ for any $y \in \mathbb{R}$ using the fact that the replacing H by $H(x,y) - H(0,y)$ gives the same result for $\partial_{t=0^+} J_t$. Then approximate $\nabla_1 H(x,y)$ with a polynomial function P_ε such that, on $\mathrm{Im}(f) \times \mathrm{Im}(g)$, we have $|\nabla_1 H - P_\varepsilon|_\infty < \varepsilon$. Using the fact that $H(0,y) = 0$ we can integrate P_ε w.r.t. the first variable and get a polynomial function Q_ε such that $\nabla_1 Q_\varepsilon = P_\varepsilon$. This implies $\mathrm{Lip}_1(H - Q_\varepsilon) < \varepsilon$.

Applying the theorem to Q_ε, we obtain the result by the lemma 7.1 letting $\varepsilon \to 0$. □

References

[AFP00] Luigi Ambrosio, Nicola Fusco, and Diego Pallara. *Functions of bounded variation and free discontinuity problems.* Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, 2000.
[Arn78] I Arnold, V. *Mathematical methods of Classical Mechanics.* Springer, 1978. Second Edition: 1989.
[CB03] Craig Calcaterra and Axel Boldt. Lipschitz flow-box theorem, 2003.
[DP98] Thierry De Pauw. On SBV dual. *Indiana Univ. Math. J.*, 47(1):99–121, 1998.
[EG92] Lawrence C. Evans and Ronald F. Gariepy. *Measure theory and fine properties of functions.* Studies in Advanced Mathematics. Boca Raton: CRC Press. viii, 268 p. , 1992.
[TY05] Alain Trouvé and Laurent Younes. Local geometry of deformable templates. *Siam Journal of Mathematical Analysis*, 2005.
[Via08] François-Xavier Vialard. Hamiltonian approach to geodesic image matching. *arXiv:0801.2095v1*, 2008.

(Filippo Santambrogio) Ceremade, Université Paris-Dauphine, Place de Lattre de Tassigny, F-75775 Paris CEDEX
E-mail address: filippo@ceremade.dauphine.fr

(François-Xavier Vialard) CMLA, Ecole Normale Supérieure de Cachan, CNRS, UniverSud, 61, avenue du Président Wilson, F-94 235 Cachan CEDEX
E-mail address: francois.xavier.vialard@normalesup.org