ANTHROPOMORPHIC IMAGE RECONSTRUCTION VIA HYPOELLIPTIC DIFFUSION*

UGO BOSCAIN†, JEAN DUPLAIX‡, JEAN-PAUL GAUTHIER§, AND FRANCESCO ROSSI¶

Abstract. In this paper we study a model of geometry of vision due to Petitot, Citti and Sarti. One of the main features of this model is that the primary visual cortex V1 lifts an image from $\mathbb{R}^2$ to the bundle of directions of the plane. Neurons are grouped into *orientation columns*, each of them corresponding to a point of this bundle.

In this model a corrupted image is reconstructed by minimizing the energy necessary for the activation of the orientation columns corresponding to regions in which the image is corrupted. The minimization process gives rise to an hypoelliptic heat equation on the bundle of directions of the plane.

In the original model directions are considered both with and without orientation giving rise respectively to a problem on the group of rototranslations of the plane $SE(2)$ or on the projective tangent bundle of the plane $PT\mathbb{R}^2$.

We provide a mathematical proof of several important facts for this model. We first prove that the model is mathematically consistent only if direction are considered without orientation. We then prove that generically the convolution of a $L^2(\mathbb{R}^2, \mathbb{R})$ function (e.g. an image) with a 2-D Gaussian is generically a Morse function. This fact is important since the lift of Morse functions to $PT\mathbb{R}^2$ is defined on a smooth manifold. We then provide the explicit expression of the hypoelliptic heat kernel on $PT\mathbb{R}^2$ in terms of Mathieu functions.

Finally we present the main ideas of an algorithm which allows to perform image reconstruction on real non-academic images. A very interesting point is that this algorithm is massively parallelizable and needs no information on where the image is corrupted.

Keywords: sub-Riemannian geometry, image reconstruction, hypoelliptic diffusion

1. Introduction. In this paper we study a model of geometry of vision initially due to Petitot (see [30, 31] and references therein), then refined by Citti and Sarti [13, 14], and by the authors of the present paper in [9]. This model was also studied by Hladky and Pauls [23] and, independently, by Duits et al. in a series of papers mostly for contour completion [16] and contour enhancement [17, 18].

To start with, assume that a grey-level image is represented by a function $\mathcal{I} \in L^2(\mathcal{D}, \mathbb{R})$, where $\mathcal{D}$ is an open bounded domain of $\mathbb{R}^2$. The algorithm that we present here is based on three crucial ideas coming from neurophysiology:

1. It is widely accepted that the retina approximately smoothes the images by making the convolution with a Gaussian function (see for instance [25, 27, 29] and references therein), equivalently solving a certain isotropic heat equation. Moreover, in image processing smoothing by the same technique is a widely used method. Then, it is an interesting question in itself to understand generic properties of these smoothed images. Our first result (proved in in Appendix A) is that, given $G(\sigma_x, \sigma_y)$ the two dimensional Gaussian centered in $(0, 0)$ with standard deviations $\sigma_x, \sigma_y > 0$, then the smoothed image

$$f = \mathcal{I} * G(\sigma_x, \sigma_y) \in L^2(\mathbb{R}^2, \mathbb{R}) \cap C^\infty(\mathbb{R}^2, \mathbb{R}),$$

*This research has been supported by the European Research Council, ERC StG 2009 "GeCoMethods", contract number 239748, by the ANR "GCM", program "Blanc – CSD", and by the DIGITEO project "CONGEO"

†CMAP, École Polytechnique CNRS, Route de Saclay, 91128 Palaiseau Cedex, France, boscain@cmap.polytechnique.fr

‡Laboratoire LSIS, Université de Toulon, France, duplaix@univ-tln.fr

§Laboratoire LSIS, Université de Toulon, France, gauthier@univ-tln.fr

¶Laboratoire LSIS, Université Paul Cézanne, Marseille, France, francesco.rossi@lsis.org

is generically a Morse function (i.e. a smooth function having as critical points only maxima, minima and saddles). The fact that this smoothed image is a Morse function has interesting consequences, as explained in the following.

Remark 1. In several applications, the convolution is made with a Gaussian of small standard deviations. Equivalently, the smoothed image can be obtained as the solution of an isotropic heat equation with small final time.

Remark 2. These results can be generalized to non-Gaussian filters and even to non-linear smoothing processes. What is crucial is the request that this filtering process produces a Morse function. See for instance [15] for some of these generalizations.

2. **The primary visual cortex V1 lifts the image from $\mathbb{R}^2$ to the bundle of directions of the plane $PT\mathbb{R}^2$.**
In a simplified model[1] (see [31, p. 79]), neurons of V1 are grouped into *orientation columns*, each of them being sensitive to visual stimuli at a given point a of the retina and for a given direction p on it. The retina is modeled by the real plane, i.e. each point is represented by $a \in \mathbb{R}^2$, while the directions at a given point are modeled by the projective line, i.e. $p \in P^1$. Hence, the primary visual cortex V1 is modeled by the so called *projective tangent bundle* $PT\mathbb{R}^2 := \mathbb{R}^2 \times P^1$. From a neurological point of view, orientation columns are in turn grouped into *hypercolumns*, each of them being sensitive to stimuli at a given point a with any direction. In the same hypercolumn, relative to a point a of the plane, we also find neurons that are sensitive to other stimuli properties, like colors. In this paper, we focus only on directions and therefore each hypercolumn is represented by a fiber P^1 of the bundle $PT\mathbb{R}^2$. See Figure 1.1.

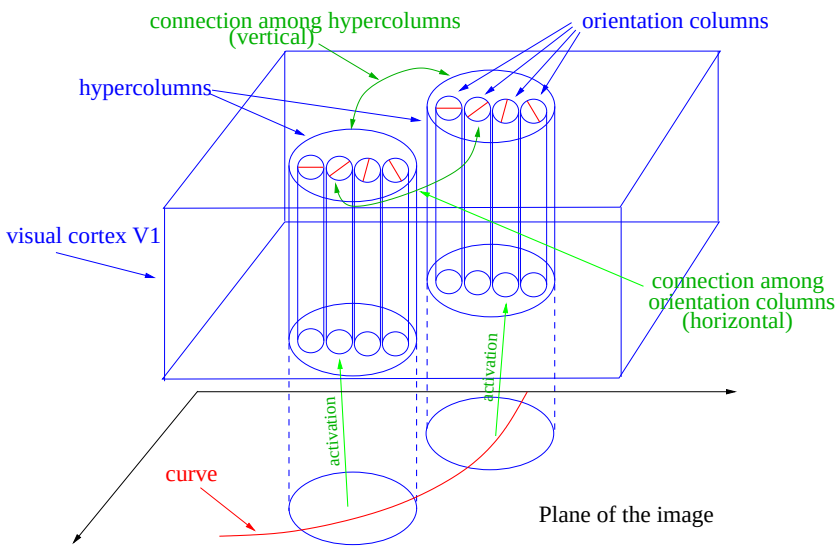

FIGURE 1.1. *A scheme of the primary visual cortex V1.*

This space has the topology of $\mathbb{R}^2 \times P^1$ (it is a trivial bundle) and its points are triples (x, y, θ), where $(x, y) \in \mathbb{R}^2$, $\theta \in \mathbb{R}/(\pi\mathbb{Z})$.

[1] For example, in this model we do not take into account the fact that the continuous space of stimuli is implemented via a discrete set of neurons.

The smoothed image $f : \mathbb{R}^2 \to \mathbb{R}$ is *lifted* to a a function $\bar{f}$ defined as follows:

$$\bar{f}(x, y, \theta) = \begin{cases} f(x, y) & \text{if } \theta \text{ is the direction of the level set of } f, \\ \\ 0 & \text{otherwise.} \end{cases}$$

It follows that $\bar{f}$ has support on a set $S_f \subset PT\mathbb{R}^2$. The following fact constitutes our second result. If f is a Morse function (which happens generically due to the smoothing of the retina as explained above), then S_f is an embedded surface in $PT\mathbb{R}^2$, see Proposition 19.

3. If the image is corrupted or missing on a set $\Omega \subset \mathcal{D}$ (i.e. if $\mathcal{I}$ is defined on $\mathcal{D} \setminus \Omega$), then the reconstruction of $\mathcal{I}$ in Ω is made by minimizing a given cost. This cost represents the energy that the primary visual cortex should spend in order to excite orientation columns which corresponds to points in Ω and hence that are not directly excited by the image. An orientation column is easily excited if it is close to another (already activated) orientation column sensitive to a similar direction in a close position (i.e. if the two are close in $PT\mathbb{R}^2$).
   When the image to be reconstructed is a curve, this gives rise to a sub-Riemannian problem (i.e. an optimal control problem linear in the control and with quadratic cost) on $PT\mathbb{R}^2$, which we briefly discuss in Sections 2.3, 2.4, 2.5.
   When the image is not just a curve, the reconstruction is made by considering the diffusion process naturally associated with the sub-Riemannian problem on $PT\mathbb{R}^2$ (described by an hypoelliptic heat equation). Such a reconstruction makes use of the function $\bar{f}$ as initial condition in a suitable way. The reconstructed image is then obtained by projecting the result of the diffusion from $PT\mathbb{R}^2$ to $\mathbb{R}^2$.

In this paper we study this model providing a mathematical proof of several important facts and adding certain important details with respect to its original version given in [13, 14]. The main improvements are described in the following. Let us first recall that $PT\mathbb{R}^2$ can be seen as the quotient of the group of rototranslations of the plane $SE(2) \simeq \mathbb{R}^2 \times S^1$ by $\mathbb{Z}_2$, where the quotient is the identification of antipodal points in S^1.

- As already mentioned, we start with any function $\mathcal{I} \in L^2(\mathcal{D}, \mathbb{R})$ and we prove that after convolution with a Gaussian of standad deviations[2] $\sigma_x = \sigma_y$, generically, we are left with a Morse function $f \in L^2(\mathbb{R}^2, \mathbb{R}) \cap C^\infty(\mathbb{R}^2, \mathbb{R})$ (see Appendix A). This smoothing process is important to guarantee certain regularity of the domain of definition of the lifted function $\bar{f}$.
- Our definition of the lift is suitable to all smooth functions, since we don't require conditions like nondegenerate gradient (as in [13]) or more complicated condition on the so called non-Legendrian solitary points (as in [23, Thm 1.6]).
- In the first version of this model [13] the image is lifted on $SE(2)$ (i.e. directions are considered with orientation), while in the second one [14] it is lifted on $PT\mathbb{R}^2$ (i.e. directions are considered without orientation). The next contribution of our paper is to show that the problem of reconstruction of images for smooth functions is well posed on $PT\mathbb{R}^2$ while it is not on $SE(2)$. First, on $PT\mathbb{R}^2$ the lift is unique, while on $SE(2)$ it is not, since level sets of

[2] We fix $\sigma_x = \sigma_y$ to guarantee invariance by rotrotranslations of the algorithm.

the image are not oriented curves. Second, the problem on $SE(2)$ cannot be interpreted as a problem of reconstruction of contours (see Remarks 4, 13 and [9]). Third, as proved in Proposition 19, the domain S_f of the lift of a Morse function f is much more natural on $PT\mathbb{R}^2$ than on $SE(2)$. On $PT\mathbb{R}^2$ it is a manifold, while on $SE(2)$ it is a manifold with a boundary (for a continuous choice of the orientation of the level sets of f). The boundary appears on minima, maxima and saddles of f. In the diffusion process, starting with an initial condition which is concentrated on a manifold is much more natural than starting with an initial condition which is concentrated on a manifold with a boundary.

- We show that the sub-Riemannian structure over $PT\mathbb{R}^2$ is not trivializable which means that it cannot be specified by a single global orthonormal frame as in [14]. For a detailed discussion of this issue see Remark 6 and [9].
- We give the expression of the hypoelliptic heat kernel over $PT\mathbb{R}^2$, while, previously, it was known only on $SE(2)$ (see [4] and [16, 17], where it was found independently).
- We provide an effective algorithm for image reconstruction that looks unexpectedly efficient on real non-academic examples as shown in Section 3.3. Moreover our algorithm has the good feature to be massively parallelizable (see Section 3). This is just the materialization of the classical fact that the noncommutative Fourier transform disintegrates the regular representation over $SE(2)$. Moreover the algorithm does not need the information of where the original image is corrupted.

Other numerical methods to compute hypoelliptic diffusion on $SE(2)$ for image processing have been developed. For instance: group convolution methods (see [12, 17, 19]) finite differences [13, 14, 20][3] and finite elements methods. (See [17, 18]. These last works are related to the noncommutative Fourier transform on $SE(2)$ and are extensions of the works by August [7].) Most of these works are about contour enhancement. In our opinion, at the present moment, our results seems to be the most convincing.

Remark 3. Notice that from the very beginning of the algorithm, we deal with the intensity of the image. Other related algorithms [13, Sec. 3.3], [23] are instead composed of two reconstruction steps. After the lift of the image, these algorithms have to deal with a surface in $SE(2)$ or $PT\mathbb{R}^2$ with a hole, corresponding to the corrupted part. The first reconstruction step is thus to fill the hole as a surface, without considering the intensity of the image. The second reconstruction step is then to put the intensity on the reconstructed part. See Remarks 15 and 22.

The results of our algorithm can be compared to the ones coming from psychological experiments. Moreover, they can be useful to reconstruct the geometry of an image, as a preliminary step of exemplar-based methods (see [11]).

Notice that an alternative technique of image processing (in particular for contour completion) based on physiological models of the visual cortex have been developed in [16, 17, 18, 26]. In these models non-isotropic diffusion is associated to an optimal control problem with drift having as solution elastica curves.

The structure of the paper is the following. In Section 2 we present in detail the sub-Riemannian structure defined on $PT\mathbb{R}^2$. We then define the corresponding

[3] Notice that classical finite difference methods "hardly works" to compute hypoelliptic diffusion. This is due to the diffusion at different scales on different directions as a consequence of the non-ellipticity of the diffusion operator.

hypoelliptic diffusion which is one of the main tools used in the algorithm of image reconstruction and we find explicitly the corresponding kernel on $PT\mathbb{R}^2$. At the end, we present in detail the mathematical algorithm.

Section 3 is devoted to the discussion about the numerical integration of the hypoelliptic evolution and to the presentation of samples.

Appendix A is devoted to the detailed proof that, generically, the convolution of a L^2 function over a bounded domain $\mathcal{D} \subset \mathbb{R}^2$ with a Gaussian G is a Morse function. In particular, we prove that the set of functions $\mathcal{I} \in L^2(\mathcal{D})$ whose convolution with a Gaussian is a Morse function in $L^2(\mathcal{D})$ is residual (i.e. it is a countable intersection of open and dense sets). We then prove that the set of functions $\mathcal{I} \in L^2(\mathcal{D})$ such that $\mathcal{I} * G$ restricted to a compact $K \subset \mathbb{R}^2$ is a Morse function is open and dense. Notice again that the proof can be adapted to any reasonable smoothing process, not only Gaussian.

2. The mathematical model and the algorithm.

2.1. Reconstruction of a curve. In this section we briefly describe an algorithm to reconstruct interrupted planar curves. The main interest of this section is the definition of the sub-Riemannian structure over $PT\mathbb{R}^2$, from which we are going to define the sub-elliptic diffusion equation.

Consider a smooth function $\gamma_0 : [a,b] \cup [c,d] \to \mathbb{R}^2$ (with $a < b < c < d$) representing a curve that is partially hidden or deleted in (b,c). We assume that starting and ending points never coincide, i.e. $\gamma_0(b) \neq \gamma_0(c)$, and that initial and final velocities $\dot\gamma(b)$ ad $\dot\gamma(c)$ are well defined and nonvanishing.

We want to find a curve $\gamma : [b,c] \to \mathbb{R}^2$ that completes γ_0 in the deleted part and that minimizes a cost depending both on the length and on the curvature K_γ of γ. Recall that

$$K_\gamma = \frac{\dot x \ddot y - \dot y \ddot x}{(\dot x^2 + \dot y^2)^{3/2}}$$

where (x,y) are the components of γ.

The fact that γ completes γ_0 means that $\gamma(b) = \gamma_0(b)$, $\gamma(c) = \gamma_0(c)$. It is also reasonable to require that the directions of tangent vectors coincide, i.e. $\dot\gamma(b) \approx \dot\gamma_0(b)$, $\dot\gamma(c) \approx \dot\gamma_0(c)$ where

$$v_1 \approx v_2 \text{ if it exists } \alpha \in \mathbb{R} \setminus \{0\} \text{ such that } v_1 = \alpha\, v_2. \tag{2.1}$$

Remark 4. Notice that we have required boundary conditions on initial and final directions without orientation. Alternatively, the problem above can be formulated requiring boundary conditions with orientation, i.e. substituting in (2.1) the condition $\alpha \in \mathbb{R}^+$. However, this choice does not guarantee existence of minimizers for the cost we are interested in, see [9] and Remark 6 below.

In this paper we are interested in the minimization of the following cost, defined for smooth curves γ in $[b,c]$:

$$J[\gamma] = \int_b^c \sqrt{\|\dot\gamma(t)\|^2 + \|\dot\gamma(t)\|^2 K_\gamma^2(t)}\, dt \tag{2.2}$$

This cost is interesting for several reasons:

- It depends on both length and curvature of γ. It is small for curves that are straight and short;

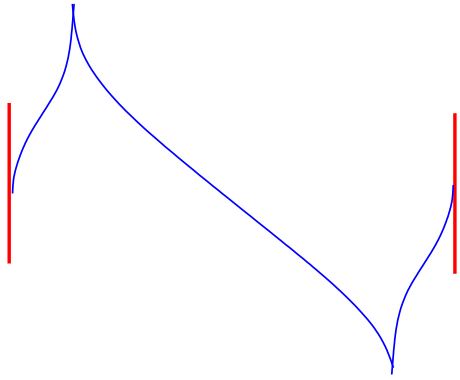

FIGURE 2.1. *A trajectory with two cusps.*

- It is invariant by rototranslation (i.e. under the action of $SE(2)$) and by reparametrization of the curve, as should be any reasonable process of reconstruction of interrupted curves.
- Minimizers for this cost do exist in the natural functional space in which this problem is formulated, without involving sophisticated functional spaces or curvatures that become measures. Indeed, in [9] we have proved:
  Proposition 5. *For every $(x_b, y_b), (x_c, y_c) \in \mathbb{R}^2$ with $(x_b, y_b) \neq (x_c, y_c)$ and $v_b,\ v_c \in \mathbb{R}^2 \setminus \{0\}$, the cost (2.2) has a minimizer over the set*

$$\mathscr{D} := \left\{ \gamma \in C^2([b,c], \mathbb{R}^2) \ s.t. \ \begin{array}{c} \sqrt{\|\dot\gamma(t)\|^2(1+K_\gamma^2(t))} \in L^1([b,c],\mathbb{R}), \\ \gamma(b) = (x_b, y_b),\ \gamma(c) = (x_c, y_c), \\ \dot\gamma(b) \approx v_b,\ \dot\gamma(c) \approx v_c. \end{array} \right\} \tag{2.3}$$

  Remark 6. In [28, 32, 33], it has been proved that minimizers for the cost (2.2) are analytic functions for which $\dot\gamma = 0$ at most for two isolated points. At these points the limit of $\|\dot\gamma(t)\| K_\gamma(t)$ is well defined. They are cusp points, i.e. points at which $\dot\gamma$ becomes opposite. See Figure 2.1.
  Notice that at cusp points the limit direction (regardless of orientation) is well defined. In [9] it is proved that if boundary conditions are required with orientation, then the cost (2.2) has no minimum over the set $\mathscr{D}$.

- Minimizing the cost $J[\gamma]$ is equivalent to minimize the energy-like cost

$$E[\gamma] = \int_b^c \left(\|\dot\gamma(t)\|^2 + \|\dot\gamma(t)\|^2 K_\gamma^2(t) \right) dt.$$

  This is a good model for describing the energy necessary to excite orientation columns which are not directly excited by the image (because they correspond to the corrupted part of the image).

However, the most interesting aspect is that this cost is a Riemannian length for lifts of planar curves over $PT\mathbb{R}^2$ (more precisely $J[\gamma]$ is a sub-Riemannian length, see below). As a consequence, we have a diffusion equation naturally associated with this cost that can be used to reconstruct more complicated images and not just curves.

Remark 7. One could argue that there is no reason to give the same weight to the length term $\|\dot\gamma\|$ and to the curvature term $\|\dot\gamma(t)\|^2 K_\gamma^2(t)$. However, if we define

the cost

$$J_\beta[\gamma] := \int_b^c \sqrt{\|\dot\gamma(t)\|^2 + \beta^2\|\dot\gamma(t)\|^2 K_\gamma^2(t)}\, dt$$

with a fixed $\beta \neq 0$ and if we consider an homothety $(x, y) \mapsto (\beta x, \beta y)$ and the corresponding transformation of a curve $\gamma = (x(t), y(t))$ to $\gamma_\beta = (\beta x(t), \beta y(t))$, then it is easy to prove that $J_\beta[\gamma_\beta] = \beta^2 J_1[\gamma] = \beta^2 J[\gamma]$. Therefore the problem of minimizing J_β is equivalent to the minimization of J with a suitable change of boundary conditions.

Although the mathematical problem is equivalent by changing β, this parameter will play a crucial role in the following, see Remark 21.

Another interesting feature is the uniqueness of this sub-Riemannian distance. Beside the possibility of adding a weight β on the curvature term, that can be removed via an homothety, it is the unique sub-Riemannian distance for lift of planar curves on $PT\mathbb{R}^2$ that is invariant under the action of $SE(2)$. See Proposition 14 below.

2.2. Sub-Riemannian manifolds. In this section we recall some standard definitions of sub-Riemannian geometry, that we use in the following. We start by recalling the definition of sub-Riemannian manifold.

Definition 8. *A (n, m)-sub-Riemannian manifold is a triple $(M, \blacktriangle, \mathbf{g})$, where*

- *M is a connected smooth manifold of dimension n;*
- *$\blacktriangle$ is a smooth distribution of constant rank $m < n$ satisfying the* **Hörmander condition***, i.e. $\blacktriangle$ is a smooth map that associates to $q \in M$ a m-dim subspace $\blacktriangle(q)$ of T_qM and $\forall\ q \in M$ we have*

$$\operatorname{span}\{[X_1, [\ldots [X_{k-1}, X_k]\ldots]](q) \mid X_i \in \mathrm{Vec}_H(M)\} = T_qM$$

  where $\mathrm{Vec}_H(M)$ denotes the set of **horizontal smooth vector fields** *on M, i.e.*

$$\mathrm{Vec}_H(M) = \{X \in \mathrm{Vec}(M) \mid X(q) \in \blacktriangle(q) \quad \forall\ q \in M\}.$$

- *$\mathbf{g}_q$ is a Riemannian metric on $\blacktriangle(q)$, that is smooth as function of q.*

A Lipschitz continuous curve $q(\cdot) : [0, T] \to M$ is said to be **horizontal** if $\dot q(t) \in \blacktriangle(q(t))$ for almost every $t \in [0, T]$. Given an horizontal curve $q(\cdot) : [0, T] \to M$, the *length of* $q(\cdot)$ is

$$l(q(\cdot)) = \int_0^T \sqrt{\mathbf{g}_{q(t)}(\dot q(t), \dot q(t))}\ dt. \tag{2.4}$$

The *distance* induced by the sub-Riemannian structure on M is the function

$$d(q_0, q_1) = \inf\{l(q(\cdot)) \mid q(0) = q_0, q(T) = q_1,\ q(\cdot) \text{ horizontal}\}.$$

The connectedness assumption for M and the Hörmander condition guarantee the finiteness and the continuity of $d(\cdot, \cdot)$ with respect to the topology of M (Chow's Theorem, see for instance [6]). The function $d(\cdot, \cdot)$ is called the *Carnot-Charateodory distance* and gives to M the structure of metric space (see [8, 22]).

It is a standard fact that $l(q(\cdot))$ is invariant under reparametrization of the curve $q(\cdot)$. On one side, if an admissible curve $q(\cdot)$ minimizes the so-called *energy functional*

$$E(q(\cdot)) = \int_0^T \mathbf{g}_{q(t)}(\dot q(t), \dot q(t))\ dt.$$

with fixed T (and initial and final fixed points), then $v = \sqrt{\mathbf{g}_{q(t)}(\dot{q}(t), \dot{q}(t))}$ is constant and $q(\cdot)$ is also a minimizer of $l(\cdot)$. On the other side, a minimizer $q(\cdot)$ of $l(\cdot)$ such that v is constant is a minimizer of $E(\cdot)$ with $T = l(q(\cdot))/v$.

A *geodesic* for the sub-Riemannian manifold is a curve $q(\cdot) : [0, T] \to M$ such that for each sufficiently small interval $[t_1, t_2] \subset [0, T]$, then $q(\cdot)_{|_{[t_1,t_2]}}$ is a minimizer of $E(\cdot)$. A geodesic for which $\mathbf{g}_{q(t)}(\dot{q}(t), \dot{q}(t))$ is identically equal to one is said to be *arclength parameterized.*

Locally, the pair $(\blacktriangle, \mathbf{g})$ can be specified by assigning a set of m smooth vector fields spanning $\blacktriangle$, that are moreover orthonormal for $\mathbf{g}$, i.e.

$$\blacktriangle(q) = \operatorname{span}\{X_1(q), \ldots, X_m(q)\}, \qquad \mathbf{g}_q(X_i(q), X_j(q)) = \delta_{ij}. \tag{2.5}$$

Such a set $\{X_1, \ldots, X_m\}$ is called a local **orthonormal frame** for the sub-Riemannian structure. When $(\blacktriangle, \mathbf{g})$ can be defined by m globally defined vector fields as in (2.5) we say that the sub-Riemannian manifold is *trivializable.*

Given a (n, m)-trivializable sub-Riemannian manifold, the problem of finding a curve minimizing the energy between two fixed points $q_0, q_1 \in M$ is naturally formulated as the following optimal control problem

$$\dot{q}(t) = \sum_{i=1}^{m} u_i(t) X_i(q(t)), \tag{2.6}$$

$$u_i(.) \in L^\infty([0, T], \mathbb{R}), \int_0^T \sum_{i=1}^{m} u_i^2(t)\, dt \to \min,$$

$$q(0) = q_0, \quad q(T) = q_1. \tag{2.7}$$

It is a standard fact that this optimal control problem is equivalent to the minimum time problem with controls $u_1, \ldots, u_m$ satisfying $u_1(t)^2 + \ldots + u_m(t)^2 \leq 1$ in $[0, T]$. When the sub-Riemannian manifold is not trivializable, the equivalence with the optimal control problem (2.6)-(2.7) is just local.

When the manifold is analytic and the orthonormal frame can be assigned by m analytic vector fields, we say that the sub-Riemannian manifold is *analytic.* In this paper we deal with an analytic sub-Riemannian manifold.

A sub-Riemannian manifold is said to be of **3D contact** type if $n = 3$, $m = 2$ and for every $q \in M$ we have $\operatorname{span}\{\blacktriangle(q), [\blacktriangle, \blacktriangle](q)\} = T_q M$. This is the case that we study in this paper. For details, see [5].

Remark 9. As a consequence of the invariance by reparameterization of the cost (2.4), it is equivalent to state the minimization problem in the space of Lipschitz or absolutely continuous curves (i.e. for $u_i(\cdot) \in L^\infty([0, T], \mathbb{R})$ or for $u_i(\cdot) \in L^1([0, T], \mathbb{R})$.) See [9, Lemma 1].

2.2.1. Left-invariant sub-Riemannian manifolds. In this section we present a natural sub-Riemannian structure that can be defined on Lie groups. All along the paper, notations are adapted to group of matrices only. For general Lie groups, by gv with $g \in G$ and $v \in \mathfrak{L}$, we mean $(L_g)_*(v)$ where L_g is the left-translation on the group.

Definition 10. *Let G be a Lie group with Lie algebra $\mathfrak{L}$ and $\mathfrak{P} \subseteq \mathfrak{L}$ a subspace of $\mathfrak{L}$ satisfying the* **Lie bracket generating condition**

$$\operatorname{Lie} \mathfrak{P} := \operatorname{span}\{[\mathfrak{p}_1, [\mathfrak{p}_2, \ldots, [\mathfrak{p}_{n-1}, \mathfrak{p}_n]]] \mid \mathfrak{p}_i \in \mathfrak{P}\} = \mathfrak{L}.$$

Endow $\mathfrak{P}$ with a positive definite quadratic form $\langle ., . \rangle$. Define a sub-Riemannian structure on G as follows:

- *the distribution is the left-invariant distribution*

$$\blacktriangle(g) := g\mathfrak{P};$$

- *the quadratic form* $\mathbf{g}$ *on* $\blacktriangle$ *is given by*

$$\mathbf{g}_g(v_1, v_2) := \langle g^{-1}v_1, g^{-1}v_2\rangle.$$

In this case we say that $(G, \blacktriangle, \mathbf{g})$ *is a left-invariant sub-Riemannian manifold.*

In the following we define a left-invariant sub-Riemannian manifold by choosing a set of m vectors $\{\mathfrak{p}_1, \ldots, \mathfrak{p}_m\}$ which form an orthonormal basis for the subspace $\mathfrak{P} \subseteq \mathfrak{L}$ with respect to the metric from Definition 10, i.e. $\mathfrak{P} = \operatorname{span}\{\mathfrak{p}_1, \ldots, \mathfrak{p}_m\}$ and $\langle \mathfrak{p}_i, \mathfrak{p}_j\rangle = \delta_{ij}$. We thus have

$$\blacktriangle(g) = g\mathfrak{P} = \operatorname{span}\{g\mathfrak{p}_1, \ldots, g\mathfrak{p}_m\}$$

and $\mathbf{g}_g(g\mathfrak{p}_i, g\mathfrak{p}_j) = \delta_{ij}$. Notice that every left-invariant sub-Riemannian manifold is trivializable.

2.3. Lift of a curve on $PT\mathbb{R}^2$ and the sub-Riemannian problem. Consider a smooth planar curve $\gamma : [b, c] \to \mathbb{R}^2$. This curve can be naturally lifted to a curve $\bar{\gamma} : [b, c] \to PT\mathbb{R}^2$ in the following way. Let $(x(t), y(t))$ be the Euclidean coordinates of $\gamma(t)$. Then the coordinates of $\bar{\gamma}(t)$ are $(x(t), y(t), \theta(t))$, where $\theta(t) \in \mathbb{R}/(\pi\mathbb{Z})$ is the direction of the vector $(x(t), y(t))$ measured with respect to the vector $(1, 0)$. In other words,

$$\theta(t) = \arctan\left(\frac{\dot{y}(t)}{\dot{x}(t)}\right) \mod \pi. \tag{2.8}$$

Of course we can extend by continuity the definition to points where $\dot{\gamma}(\bar{t}) = 0$ but $\lim_{t\to\bar{t}} \theta(t)$ is well defined. We assume
[H] $\theta : [b, c] \to \mathbb{R}/(\pi\mathbb{Z})$ is absolutely continuous.
Notice that $\dot{\theta} = \|\dot{\gamma}\|K_\gamma$, hence hypothesis **[H]** is equivalent to require that $\|\dot{\gamma}\|K_\gamma \in L^1([b, c], \mathbb{R})$.

The requirement that a curve $(x(t), y(t), \theta(t))$ satisfies the constraint (2.8) under **[H]** can be slightly generalized by requiring that $(x(t), y(t), \theta(t))$ is an admissible trajectory of the control system on $PT\mathbb{R}^2$:

$$\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \end{pmatrix} = u_1(t) \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \\ 0 \end{pmatrix} + u_2(t) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{2.9}$$

with $u_1, u_2 \in L^1([b, c], \mathbb{R})$. Indeed each smooth trajectory γ satisfying **[H]** is an admissible trajectory of (2.9).

Since $u_1(t)^2 = \|\dot{\gamma}(t)\|^2$, $u_2(t)^2 = \dot{\theta}^2 = \|\dot{\gamma}(t)\|^2 K_\gamma(t)^2$, we have

$$J[\gamma] = \int_b^c \sqrt{u_1(t)^2 + u_2(t)^2}\, dt$$

Hence, the problem of minimizing the cost (2.2) on the set of curves $\mathscr{D}$ is (slightly)

generalized considering the optimal control problem (here $q(\cdot) = (x(\cdot), y(\cdot), \theta(\cdot))$)

$$\dot{q} = u_1(t)X_1(q) + u_2(t)X_2(q), \tag{2.10}$$

$$X_1(q) = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \\ 0 \end{pmatrix}, \quad X_2(q) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \tag{2.11}$$

$$l(q(\cdot)) = \int_b^c \sqrt{u_1(t)^2 + u_2(t)^2}\, dt \to \min, \tag{2.12}$$

$$q(b) = (x_b, y_b, \theta_b), \ q(c) = (x_c, y_c, \theta_c), \tag{2.13}$$

$$(x_b, y_b) \neq (x_c, y_c), \ u_1, u_2 \in L^1([b, c], \mathbb{R}). \tag{2.14}$$

Remark 11. Notice that there are admissible trajectories $q(\cdot) = (x(\cdot), y(\cdot), \theta(\cdot))$ of the control system (2.10) for which the condition $\theta(\bar{t}) = \lim_{t\to\bar{t}} \arctan\left(\frac{\dot{y}(t)}{\dot{x}(t)}\right)$ is not verified (consider for instance the trajectory $x(t) = 0$, $y(t) = 0$, $\theta(t) = t$) or such that $x(\cdot)$ or $y(\cdot)$ fail to be smooth. However, it has been proved in [9] that minimizers of (2.10)-(2.14) are minimizers of (2.2) on the set $\mathscr{D}$ and they are smooth.

Remark 12. (non-trivializability) A certain abuse of notation appears in formulas (2.9), (2.12), and (2.13), as in [14]. Indeed the vector field X_1 is not well defined on $PT\mathbb{R}^2$. For instance, it takes two opposite values in θ and $\theta + \pi$, that are identified. A correct definition of the sub-Riemannian structure requires two charts:

- Chart A: $\theta \in]0 + k\pi, \pi + k\pi[$, $k \in \mathbb{Z}$, $x, y \in \mathbb{R}$.

$$\dot{q} = u_1^A(t)X_1^A(q) + u_2(t)X_2(q), \quad X_1^A = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \\ 0 \end{pmatrix},$$

$$l(q(\cdot)) = \int_b^c \sqrt{u_1^A(t)^2 + u_2(t)^2}\, dt,$$

- Chart B: $\theta \in] - \pi/2 + k\pi, \pi/2 + k\pi[$, $k \in \mathbb{Z}$, $x, y \in \mathbb{R}$.

$$\dot{q} = u_1^B(t)X_1^B(q) + u_2(t)X_2(q), \quad X_1^B = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \\ 0 \end{pmatrix},$$

$$l(q(\cdot)) = \int_b^c \sqrt{u_1^B(t)^2 + u_2(t)^2}\, dt,$$

One can check that the two charts are compatible and that this sub-Riemannian structure is non-trivializable, while $PT\mathbb{R}^2$ is parallelizable.

Since the formal expression of X_1^A and X_1^B are the same, while they are defined on different domains, one can proceed with a single chart (however, one should bear in mind that u_1 changes sign when passing from the chart A to the chart B in $\mathbb{R} \times \mathbb{R} \times]\pi/2, \pi[$). In the following, since we study a "sum of squares" hypoelliptic diffusion on this sub-Riemannian structure, the problem disappears.

This sub-Riemannian manifold is of 3D contact type: the distribution has dimension 2 over a three-dimensional manifold and

$$\mathrm{span}\{X_1(q),\ X_2(q),\ [X_1, X_2](q)\} = T_q PT\mathbb{R}^2.$$

2.4. The sub-Riemannian problem on $SE(2)$. It is convenient to lift the sub-Riemannian problem on $PT\mathbb{R}^2$ (2.10)-(2.14) on the group of rototranslation of the plane $SE(2)$, in order to take advantage of the group structure. It is the group of matrices of the form

$$SE(2) = \left\{ \begin{pmatrix} \cos(\theta) & -\sin(\theta) & x \\ \sin(\theta) & \cos(\theta) & y \\ 0 & 0 & 1 \end{pmatrix} \mid \begin{array}{l} \theta \in \mathbb{R}/(2\pi\mathbb{Z}), \\ x, y \in \mathbb{R} \end{array} \right\}.$$

In the following we often denote an element of $SE(2)$ by $g = (x, y, \theta)$.

A basis of the Lie algebra of $SE(2)$ is $\{p_1, p_2, p_3\}$, with

$$p_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad p_2 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad p_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

We define a trivializable sub-Riemannian structure on $SE(2)$ as presented in Section 2.2.1: consider the two left-invariant vector fields $X_i(g) = gp_i$ with $i = 1, 2$ and set

$$\blacktriangle(g) = \operatorname{span}\{X_1(g), X_2(g)\} \qquad \mathbf{g}_g(X_i(g), X_j(g)) = \delta_{ij}.$$

In coordinates, the optimal control problem

$$\dot{g} \in \blacktriangle(g), \quad l(g(.)) = \int_b^c \sqrt{\mathbf{g}_{g(t)}(\dot{g}, \dot{g})}\, dt \to \min, \tag{2.15}$$

$$g(b) = (x^b, y^b, \theta^b), \quad g(c) = (x^c, y^c, \theta^c), \tag{2.16}$$

$$(x^b, y^b) \neq (x^c, y^c), \tag{2.17}$$

has the form (2.10)-(2.14), but $\theta \in \mathbb{R}/(2\pi\mathbb{Z})$. Notice that the vector field $(\cos(\theta), \sin(\theta), 0)$ is well defined on $SE(2)$.

Remark 13. It is worth mentioning that the problem (2.15)-(2.17) (i.e. the problem (2.10)-(2.14) with $\theta \in \mathbb{R}/(2\pi\mathbb{Z})$) cannot be interpreted as a problem of reconstruction of planar curves where initial and final positions and initial and final direction of velocities (with orientation) are fixed. For instance, consider the curve starting from $(x, y, \theta) = (0, 0, 0)$ and corresponding to controls $u_1(t) = \pi/2 - t$, $u_2(t) = 1$. The corresponding trajectory in the (x, y) plane is $(-\cos(t) + \frac{1}{2}(\pi - 2t)\sin(t) + 1, \pi \sin^2\left(\frac{t}{2}\right) + t\cos(t) - \sin(t))$. Notice that this trajectory has a cusp at time $t = \pi/2$. For $t \in [0, \pi/2[$ we have that θ is the angle with respect to $(1, 0)$ of the vector $(\dot{x}(t), \dot{y}(t))$, while for $t \in]\pi/2, \pi]$, it is not. See Figure 2.2.

The control problem (2.10)-(2.14) defined on $PT\mathbb{R}^2$ is left-equivariant under the action of $SE(2)$. Indeed, topologically, $PT\mathbb{R}^2$ can be seen as the quotient of $SE(2)$ by $\mathbb{Z}_2$ (in other words $SE(2)$ is a double covering of $PT\mathbb{R}^2$). In coordinates, $(x, y, \theta) \in PT\mathbb{R}^2$ corresponds to the two points $(x, y, \theta), (x, y, \theta + \pi) \in SE(2)$. Also, there is a natural transitive action of $SE(2)$ on PTR^2 given by

$$\underbrace{\begin{pmatrix} \cos(\theta) & -\sin(\theta) & x \\ \sin(\theta) & \cos(\theta) & y \\ 0 & 0 & 1 \end{pmatrix}}_{\in SE(2)} \underbrace{\begin{pmatrix} x' \\ y' \\ \theta' \end{pmatrix}}_{\in PT\mathbb{R}^2} = \underbrace{\begin{pmatrix} \cos(\theta)x - \sin(\theta)y + x' \\ \sin(\theta)x + \cos(\theta)y + y' \\ \theta' + \theta \end{pmatrix}}_{\in PT\mathbb{R}^2}$$

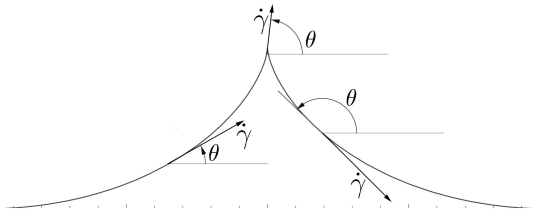

FIGURE 2.2. *A case in which $\theta \in \mathbb{R}/(2\pi\mathbb{Z})$ is not the direction of $\dot\gamma$.*

where $\theta' + \theta$ is intended modulo π. The orthonormal frame for the sub-Riemannian structure on $PT\mathbb{R}^2$ given by X_1 and X_2 in formula (2.10) is indeed left-equivariant under the action of $SE(2)$.

In other words, given $(x, y, \theta) \in PT\mathbb{R}^2$ such that $g \in SE(2)$ satisfies $(x, y, \theta) = g(0, 0, 0)$, then

$$X_1(x, y, \theta) = gp_1, \quad X_2(x, y, \theta) = gp_2. \tag{2.18}$$

The following proposition can be checked directly.

Proposition 14. *Let $(PT\mathbb{R}^2, \blacktriangle, \mathbf{g})$ be a sub-Riemannian manifold and assume that it is left-equivariant under the natural action of $SE(2)$. This means that if $\{F_1, F_2\}$ is an ortnonormal frame for the sub-Riemannian structure then*

$$F_1(x, y, \theta) = gF_1(0, 0, 0), \quad F_2(x, y, \theta) = gF_2(0, 0, 0), \tag{2.19}$$

where $g \in SE(2)$ is such that $(x, y, \theta) = g(0, 0, 0)$. Then, up to a change of coordinates and a rotation of the orthonormal frame, we have that

$$F_1(x, y, \theta) = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \\ 0 \end{pmatrix}. \quad F_2(x, y, \theta) = \begin{pmatrix} 0 \\ 0 \\ 1/\beta \end{pmatrix} \tag{2.20}$$

for some $\beta > 0$. Notice that the problem of finding curves minimizing the length for the sub-Riemannian problem on $PT\mathbb{R}^2$ for which an orthonormal frame is given by (2.20), is equivalent to the optimal control problem (2.10), with the cost (2.4).

2.5. The Sachkov synthesis. The solution of the minimization problem (2.10)-(2.14) on $PT\mathbb{R}^2$, can be obtained from that of the problem on $SE(2)$ (2.15)-(2.17). The latter has been studied by Yuri Sachkov in a series of papers [28, 32, 33] (the first one in collaboration with I. Moiseev).

The authors computed the optimal synthesis for the problem, i.e., they computed the optimal trajectory connecting each pair of points. They also gave a precise description of the cut locus, i.e. the set of points where geodesics lose optimality. The problem is simplified by the group structure: it is enough to study the case where the starting point is the identity.

The complete optimal synthesis and the description of the cut locus for the problem formulated on $PT\mathbb{R}^2$ has not been computed. However, as noticed by Sachkov, if we want to find the optimal trajectory joining (x, y, θ) to $(\bar{x}, \bar{y}, \bar{\theta})$ in $PT\mathbb{R}^2$, it is enough to find the shortest path among the four optimal trajectories joining the following points in $SE(2)$:

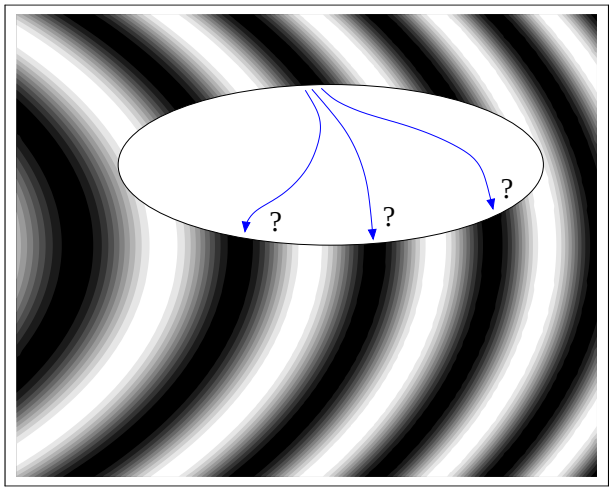

FIGURE 2.3. *The problem of connecting level sets*

- (x, y, θ) to $(\bar{x}, \bar{y}, \bar{\theta})$
- $(x, y, \theta + \pi)$ to $(\bar{x}, \bar{y}, \bar{\theta})$
- (x, y, θ) to $(\bar{x}, \bar{y}, \bar{\theta} + \pi)$
- $(x, y, \theta + \pi)$ to $(\bar{x}, \bar{y}, \bar{\theta} + \pi)$

Moreover, Yuri Sachkov built a numerical algorithm for curve reconstruction on $PT\mathbb{R}^2$. In this paper, we will not go further on the subject of reconstruction of curves. For our purpose of image reconstruction, the sub-Riemannian structure only is important, since it allows to define intrinsically a nonisotropic diffusion process.

2.6. The hypoelliptic heat kernel. When the image is not just a curve, one cannot use the algorithm described above in which curves are reconstructed by solving a sub-Riemannian problem with fixed boundary conditions. Indeed, even if a corrupted image is thought as a set of interrupted curves (the level sets), it is unclear how to connect the different components of the level set among them (see Figure 2.3).

Moreover, if the corrupted part contains the neighborhood of a maximum or minimum, then certain level sets are completely missing and cannot be reconstructed.

Remark 15. The difficulty of reconstructing a portion of an image containing a maximum or a minimum is also the main drawbacks of methods based on sub-Riemannian minimal surfaces. These algorithms (see [13, 14, 23]) consider the boundary of the lift of the corrupted part as a closed curve γ in the space $SE(2)$ or $PT\mathbb{R}^2$. They then "fill the hole" with the surface that has boundary γ and minimizes the surface area computed with respect to the sub-Riemannian metric. As clearly explained in [23], this method can fail for 3 main reasons: the minimal surface does not exist (depending on the regularity of γ), it can be non-unique, or even can exist but its projection on $\mathbb{R}^2$ does not coincide with the corrupted part (either not covering the whole part or covering also a part of the non-corrupted image).

A second problem is that, even if the surface exists and it is computed, one has to choose how diffuse the intensity of the image on the reconstructed surface. See [23, Def 7.3] for the introduction of an "interpolation" function f_t and a "disambiguation" function F.

We then use the original image as the initial condition for the nonisotropic diffusion equation associated with the sub-Riemannian structure.

Roughly speaking, this diffusion equation describes the density of probability of finding the system in the point (x, y, θ) at time t, while replacing the controls in equation (2.10) by independent Wiener processes.

This diffusion equation is

$$\partial_t \phi(x, y, \theta, t) = \Delta_H \phi(x, y, \theta, t) \tag{2.21}$$

where

$$\Delta_H = (X_1)^2 + (X_2)^2 = (\cos(\theta)\partial_x + \sin(\theta)\partial_y)^2 + \partial_\theta^2$$

Since at each point (x, y, θ) we have

$$\text{span}\,\{X_1, X_2, [X_1, X_2]\} = T_{(x_1, x_2, \theta)} PT\mathbb{R}^2,$$

Hörmander theorem [24] implies that the operator Δ_H is hypoelliptic.

The diffusion described by the equation (2.21) is highly non isotropic. Indeed one can estimate the hypoelliptic heat kernel in terms of the sub-Riemannian distance (see for instance [4]), that is highly non isotropic as a consequence of the ball-box theorem (see for instance [8]).

Remark 16. Notice that the sub-elliptic diffusion equation corresponding to the sub-Riemannian structure (2.15)-(2.17) on $SE(2)$, has the same form (2.21). The only difference is that on $SE(2)$, $\theta \in \mathbb{R}/(2\pi\mathbb{Z})$.

Remark 17. In [4] it has been proved that the Laplacian Δ_H is intrinsic on $SE(2)$, meaning that it does not depend on the choice of the orthonormal frame for the sub-Riemannian structure. One can easily prove that this is the case also for Δ_H on $PT\mathbb{R}^2$.

2.7. The hypoelliptic heat kernel on $SE(2)$. The hypoelliptic heat kernel for the equation (2.21) on $SE(2)$ was computed in [4, 16]. More precisely, thanks to the left-invariance of X_1 and X_2, the equation (2.21) admits a a right-convolution kernel $p_t(.)$, i.e. there exists p_t such that

$$e^{t\Delta_H}\phi_0(g) = \phi_0 * p_t(g) = \int_G \phi_0(h) p_t(h^{-1}g)\mu(h) \tag{2.22}$$

is the solution for $t > 0$ of (2.21) with initial condition $\phi(0, g) = \phi_0(g) \in L^1(SE(2), \mathbb{R})$ with respect to the Haar measure μ.

We have computed p_t in [4]:

$$\begin{aligned} p_t(g) = \int_0^{+\infty} \lambda \Bigg(& \sum_{n=0}^{+\infty} e^{a_n^\lambda t} < \text{ce}_n(\theta, \frac{\lambda^2}{4}), \mathfrak{X}^\lambda(g)\text{ce}_n(\theta, \frac{\lambda^2}{4}) > + \\ & + \sum_{n=1}^{+\infty} e^{b_n^\lambda t} < \text{se}_n(\theta, \frac{\lambda^2}{4}), \mathfrak{X}^\lambda(g)\text{se}_n(\theta, \frac{\lambda^2}{4}) > \Bigg)\, d\lambda. \end{aligned} \tag{2.23}$$

Here λ indexes the unitary irreducible representations of the group and

$$\begin{aligned} & \mathfrak{X}^\lambda(g) : L^2(S^1, \mathbb{C}) \to L^2(S^1, \mathbb{C}), \\ & \mathfrak{X}^\lambda(g)\psi(\alpha) = e^{i\lambda(x\cos(\alpha) - y\sin(\alpha))}\psi(\alpha + \theta) \end{aligned}$$

is the representation of the group element $g = (x, y, \theta)$ on $L^2(S^1, \mathbb{C})$.

The functions se_n and ce_n are the 2π-periodic Mathieu cosines and sines, and $< \phi_1, \phi_2 >:= \int_{S^1} \phi_1(\alpha)\overline{\phi_2(\alpha)}\, d\alpha$. The eigenvalues of the hypoelliptic Laplacian are $a_n^\lambda := -\frac{\lambda^2}{4} - a_n\left(\frac{\lambda^2}{4}\right)$ and $b_n^\lambda := -\frac{\lambda^2}{4} - b_n\left(\frac{\lambda^2}{4}\right)$, where a_n and b_n are characteristic values for the Mathieu equation. For details about Mathieu functions see for instance [3, Chapter 20].

Since the operator $\partial_t - \Delta_H$ is hypoelliptic, then the kernel is a C^∞ function of $(t, g) \in \mathbb{R}^+ \times G$. Notice that $p_t(g) = e^{t\Delta_H}\delta_{\mathrm{Id}}(g)$.

The kernel (2.23) has been obtained by using the generalized Fourier transform. Once again, we refer to [4] for a detailed description of the generalized Fourier transform and the method to compute the kernel.

2.8. The hypoelliptic heat kernel on $PT\mathbb{R}^2$. $SE(2)$ is a double covering of $PT\mathbb{R}^2$. To a point $(x, y, \theta) \in PT\mathbb{R}^2$ correspond the two points (x, y, θ) and $(x, y, \theta+\pi)$ in $SE(2)$. From the next proposition it follows that we can interpret the hypoelliptic heat equation on $PT\mathbb{R}^2$ as the hypoelliptic heat equation on $SE(2)$ with a symmetric initial condition. It permits also to compute the heat kernel on $PT\mathbb{R}^2$ starting from the one on $SE(2)$.

Proposition 18. *Let $\phi_0 \in L^1(SE(2), \mathbb{R})$ and assume that $\phi_0(x, y, \theta) = \phi_0(x, y, \theta+\pi)$ a.e. Then the solution at time t of the hypoelliptic heat equation (2.21) on $SE(2)$, having ϕ_0 as initial condition at time zero, satisfies*

$$\phi(t, x, y, \theta) = \phi(t, x, y, \theta + \pi). \tag{2.24}$$

Moreover if $\phi_0 \in L^1(PT\mathbb{R}^2, \mathbb{R})$, then the solution at time t of the hypoelliptic heat equation on $PT\mathbb{R}^2$ (2.21) having ϕ_0 as initial condition at time zero is given by

$$\phi(t, x, y, \theta) = \int_{PT\mathbb{R}^2} \phi_0(\bar{x}, \bar{y}, \bar{\theta}) P_t(x, y, \theta, \bar{x}, \bar{y}, \bar{\theta})\, d\bar{x}\, d\bar{y}\, d\bar{\theta} \tag{2.25}$$

where

$$P_t(x, y, \theta, \bar{x}, \bar{y}, \bar{\theta}) := p_t((\bar{x}, \bar{y}, \bar{\theta})^{-1} \circ (x, y, \theta)) + p_t((\bar{x}, \bar{y}, \bar{\theta})^{-1} \circ (x, y, \theta + \pi)) \tag{2.26}$$

In the right hand side of equation (2.26), the group operations are intended in $SE(2)$.

Proof. Define the element $\Pi = (0, 0, \pi) \in SE(2)$ and observe the following properties:

- Π is idempotent.
- Property (2.24) reads as $\phi_0(g\Pi) = \phi_0(g)$.
- The kernel $p_t(g)$ satisfies $p_t(\Pi g) = p_t(g\Pi)$. Indeed, call $g = (x, y, \theta)$ and observe that, given a real function $\psi(\alpha)$, we have

$$\begin{aligned}\mathfrak{X}^\lambda\,(\Pi \circ g)\,\psi(\alpha) &= \mathfrak{X}^\lambda\,((-x, -y, \theta))\,\psi(\alpha) = \\ &= \overline{\mathfrak{X}^\lambda\,((x, y, \theta + \pi))\,\psi(\alpha)} = \overline{\mathfrak{X}^\lambda\,(g \circ \Pi)\,\psi(\alpha)}.\end{aligned}$$

  Reminding the explicit expression of p_t given in (2.23), we have $p_t(\Pi g) = \overline{p_t(g\Pi)}$. But p_t is real, hence the equality follows.

We compute now $\phi(t, g\Pi)$ in $SE(2)$ with ϕ_0 satisfying (2.24) and we prove that $\phi(t, g\Pi) = \phi(t, g)$. Indeed,

$$\begin{aligned}\phi(t, g\Pi) &= \int_G \phi_0(h) p_t(h^{-1} g\Pi)\, dh = \int_G \phi_0(l\Pi) p_t(\Pi^{-1} l^{-1} g\Pi)\, d(l\Pi) = \\ &= \int_G \phi_0(l) p_t(\Pi l^{-1} g\Pi)\, dl = \int_G \phi_0(l) p_t(l^{-1} g\Pi \circ \Pi)\, dl = \\ &= \int_G \phi_0(l) p_t(l^{-1} g)\, dl = \phi(t, g).\end{aligned}$$

Let us prove now the expression (2.25) for $\phi(t, [g]) \in L^1(PT\mathbb{R}^2, \mathbb{R})$ for initial data $\phi_0([g])$. Consider the function $\psi_0(g) \in L^1(SE(2), \mathbb{R})$ defined by $\psi_0(g) = \phi([g])$, that clearly satisfies (2.24). Consider the unique solution $\psi(t, g)$ of the hypoelliptic equation (2.21), that is given by $\psi(t, g) = \psi_0 * p_t(g)$. Since $\psi(t, g) = \psi(t, g\Pi)$, the function $\phi(t, [g]) := \psi(t, g)$ is well defined.

It remains to show that ϕ defined above is the solution of (2.21) on $PT\mathbb{R}^2$. Indeed $\partial_t \phi = \partial_t \psi = \Delta_H \psi$. Since the vector fields defining Δ_H both on $SE(2)$ and $PT\mathbb{R}^2$ coincide, then the differential operators Δ_H defined on $SE(2)$ and $PT\mathbb{R}^2$ coincide, hence $\Delta_H \psi = \Delta_H \phi$. Thus ϕ satisfies (2.21). Since $\phi(0, [g]) = \phi_0([g])$, then ϕ is the (unique) solution.

The explicit expression (2.25) is a direct consequence of the definition $\phi(t, [g]) := \psi(t, g)$ and of the explicit expression of ψ given in (2.22). Indeed,

$$\begin{aligned}\phi(t, [g]) = \psi(t, g) &= \int_{SE(2)} \psi_0(h) p_t(h^{-1} g) dh = \int_{\mathbb{R}^2} \int_0^{2\pi} \psi_0(h) p_t(h^{-1} g)\, dh = \\ &= \int_{\mathbb{R}^2} \int_0^{\pi} \psi_0(h) p_t(h^{-1} g) + \psi_0(h\Pi) p_t((h\Pi)^{-1} g)\, dh = \\ &= \int_{PT\mathbb{R}^2} \phi_0(h) \left(p_t(h^{-1} g) + p_t(h^{-1} g\Pi) \right)\, dh.\end{aligned}$$

The expression (2.25) is recovered by writing $g = (x, y, \theta)$, $h = (\bar{x}, \bar{y}, \bar{\theta})$ and recalling that $g\Pi = (x, y, \theta + \pi)$. □

2.9. The mathematical algorithm. In this section we describe the main steps of the mathematical algorithm for image reconstruction. In the next section we give some guidelines for numerical implementation.

STEP 1: Smoothing of $\mathcal{I}_c$ Assume that the grey level of a corrupted image is described by a function $\mathcal{I}_c : \mathcal{D}_c := \mathcal{D}^2 \setminus \Omega \to [0, \infty[$. The set Ω represents the region where the image is corrupted. The subscript "$_c$" means "corrupted". After making the convolution with a Gaussian of standard deviations $\sigma_x = \sigma_y > 0$[4], we get a smooth function defined on $\mathbb{R}^2$, which is generically is a Morse function:

$$f_c = \mathcal{I}_c * G(\sigma_x, \sigma_y).$$

We recall that a smooth function $f_c : \mathbb{R}^2 \to \mathbb{R}$ is said to be Morse if it has only isolated critical points with nondegenerate Hessian. Roughly speaking, a Morse function is a function whose level sets are like those of Figure 2.4.

[4] $\mathcal{I}_c$ is considered to be zero outside $\mathcal{D}_c$. Moreover we assume $\sigma_x = \sigma_y$ to guarantee invariance by rototranslations of the algorithm.

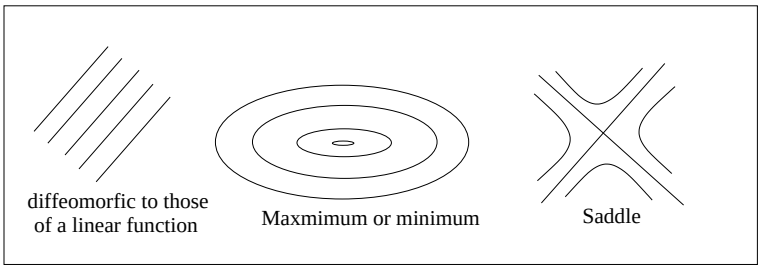

FIGURE 2.4. *Level sets of a Morse function.*

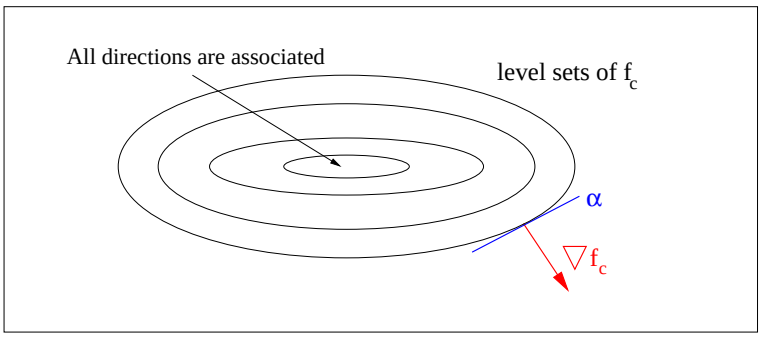

FIGURE 2.5. *Lift of an image with a maximum point.*

STEP 2: The lift of $f_c : \mathbb{R}^2 \to \mathbb{R}$ to a function $\bar{f}_c : PT\mathbb{R}^2 \to \mathbb{R}$

This is made by associating to every point (x, y) of $\mathbb{R}^2$ the direction $\theta \in \mathbb{R}/(\pi\mathbb{Z})$ of the level set of f_c at the point (x, y). This direction is well defined only at points where $\nabla f_c \neq 0$. At points where $\nabla f_c = 0$, we associate all possible directions (see Figure 2.5). More precisely, we define the *lifted support* S_f, associated with the function f as follows,

$$S_f = \{(x, y, \theta) \in \mathbb{R}^2 \times P^1 \text{ s.t. } \nabla f_c(x, y) \cdot (\cos(\theta), \sin(\theta)) = 0\},$$

where the dot means the standard scalar product on $\mathbb{R}^2$. Let $\Pi : S_f \to \mathbb{R}^2$ be the standard projection $(x, y, \theta) \in S_f \to (x, y) \in \mathbb{R}^2$. Notice that if $\nabla f_c(x, y) \neq 0$ then $\Pi^{-1}(x, y)$ is a single point, while if $\nabla f_c(x, y) = 0$ then $\Pi^{-1}(x, y) = \mathbb{R}/(\pi\mathbb{Z})$.

Let us study the set S_f, when f_c is a Morse function. If $(x, y) \in \mathbb{R}^2$ is such that $\nabla f_c(x, y) \neq 0$ and U is a small enough open neighborhood of (x, y), then the lift of S_f is a manifold in $U \times P^1$. See Figure 2.6 A. If (x, y) is an isolated maximum of f_c, and U is a small enough open neighborhood of (x, y) having a level set of f_c as boundary, then S_f is a Möbius strip in $U \times P^1$. See Figure 2.6 B. The same happens when (x, y) is an isolated minimum or saddle point of f_c. Indeed we have:

Proposition 19. *If $f_c : \mathbb{R}^2 \to \mathbb{R}$ is a Morse function, then S_f is an embedded 2-D submanifold of $\mathbb{R}^2 \times P^1$.*

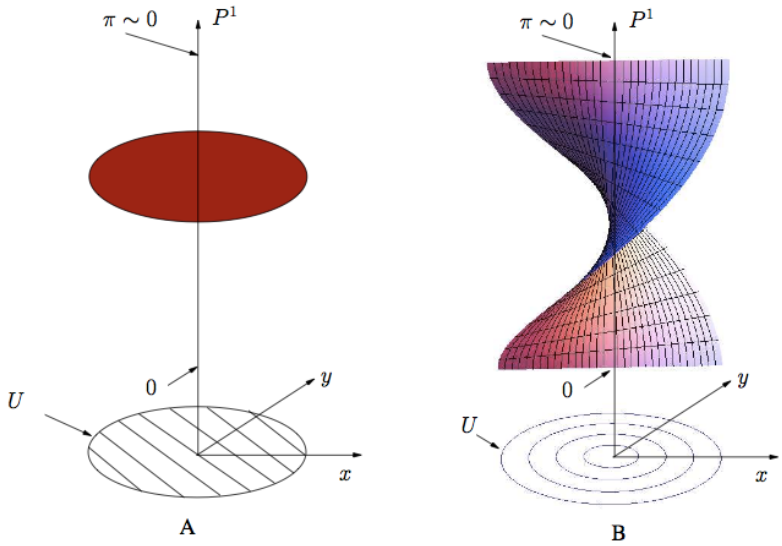

FIGURE 2.6. *In Figure A, we draw the lifted support of a function, the level sets of which are the ones of a linear function with nonvanishing gradient. Figure B presents the lifted support of a function in a neighborhood of a maximum point.*

Proof. Consider the surface $\bar{S} \in SE(2)$ given by the equation

$$g(x, y, \theta) := \cos(\theta)\partial_x f_c + \sin(\theta)\partial_y f_c = 0. \tag{2.27}$$

If $(x, y, \theta) \in \bar{S}$ then $(x, y, \theta + \pi) \in \bar{S}$ as well and $\bar{S}$ is a double covering of S_f. It is enough to show that $\bar{S}$ is a surface.

At points $(x, y, \theta) \in \bar{S}$ where $\nabla f_c \neq 0$ then dg is non-zero. Indeed $\partial_\theta g = 0$ would imply that the vector $\nabla f_c \neq 0$ is orthogonal to two non-zero vectors. At points where $\nabla f_c = 0$ we have

$$\begin{pmatrix} \partial_x g \\ \partial_y g \end{pmatrix} = H_{f_c} \cdot \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}. \tag{2.28}$$

which cannot be zero since the Hessian H_{f_c} of f_c is non-degenerate by the Morse assumption.

STEP 3: Lift of f_c to a distribution in $\mathbb{R}^2 \times P^1$ supported on S_f

Consider the distribution on $\mathbb{R}^2 \times P^1$:

$$\bar{f}_c(x, y, \theta) = f_c(x, y)\delta(g)$$

where $\delta(g)$ is the Dirac-delta distribution associated with $g(x, y, \theta) := \cos(\theta)\partial_x f_c + \sin(\theta)\partial_y f_c$ in the sense of [21, p.222]. This distribution is supported on S_f and it is canonically defined by f_c. Notice that this choice is not crucial and there are other possibilities. For example, in [14] the Dirac delta is replaced by a a power of the cosine of the angle, centered on the angle θ.

Remark 20. This step is formally necessary for the following reason. The surface S_c is 2D in a 3D manifold, hence the real function f_c defined on it is vanishing a.e. as a function defined on $PT\mathbb{R}^2$. Thus the hypoelliptic evolution of f_c (that is, the next STEP 4) produces a vanishing function. Multiplying f_c by a Dirac delta is a natural way to obtain a nontrivial evolution.

STEP 4: Hypoelliptic evolution

Fix $T > 0$. Compute the solution at time T to the Cauchy problem,

$$\begin{cases} \partial_t \phi(g,t) = ((\cos(\theta)\partial_x + \sin(\theta)\partial_y)^2 + \beta^2 \partial_\theta^2)\phi(g,t) \\ \phi(g,0) = \bar{f}_c(g). \end{cases} \tag{2.29}$$

Remark 21. In the formula above, the Laplacian is given by $X_1^2 + \beta^2 X_2^2$, thus it depends on the fixed parameter β. This means that we use evolution depending on the cost J_β rather than J. Tuning this parameter will provide better results of the reconstruction algorithm.

STEP 5: Projecting down

Compute the reconstructed image by choosing the maximum value on the fiber.

$$f_T(x,y) = \max_{\theta \in P^1} \phi(x,y,\theta,T).$$

Again other choices are possible for this projection.

Remark 22. The algorithm depends on two parameters. The first is the time of the evolution T, the second is the relative weight β in formula (2.29). A variant of this algorithm consists of re-iterating the steps above for very short diffusion times. This idea was already presented in [13] to build a minimal surface.

Remark 23. One main features of this algorithm is that it does not need the knowledge of the corrupted part. As a consequence the diffusion acts also in the non-corrupted region. However, due to the highly nonisotropic character of the diffusion, this effect is not too visible as show our results below. Modifications of the algorithm which keep the original image unmodified are possible, by admitting the diffusion only in the corrupted part, like in [13, 14].

3. Numerical implementation and results. First we present the main lines of the algorithm used in our simulations.

3.1. STEPS 2-3: Lift of an image. The formal definition of the lifted function is hard to realize numerically for two reasons: the discretization of the angle variable θ and the presence of a delta function.

Both issues are solved changing the definition of the lifted function:

$$\bar{f}_c(x,y,\theta) = f(x,y)\phi(\nabla f(x,y),\theta),$$

where $\phi(0,\theta) = 1/(2\varepsilon)\,\forall\,\theta \in \mathbb{R}/(\pi\mathbb{Z})$ and $\phi(v,\theta) = \phi^1(\arg(v) - \theta)$ where $\phi^1(\beta)$ is the

π-periodic function assuming the following values over the interval $[0, \pi]$:

$$\phi^1(\beta) := \begin{cases} 1/(2\varepsilon) & \text{if } \beta \in \left[\frac{\pi}{2} - \varepsilon, \frac{\pi}{2} + \varepsilon\right], \\ 0 & \text{otherwise,} \end{cases}$$

for a fixed $\varepsilon > 0$.

Since the space is discretized, the non-zero values of $\bar{f}_c$ are no longer defined over a set of null measure, hence the discretized hypoelliptic diffusion gives non vanishing function for all $T > 0$. Thus, it is not necessary to perform STEP 3.

3.2. STEP 4: Hypoelliptic evolution. In this section we give the crucial ideas to compute efficiently the hypoelliptic evolution (2.29). Here $\langle ., . \rangle$ is the scalar product in $\mathbb{R}^2$ and R_θ is the rotation operator of angle θ.

First of all, the main feature of the noncommutative Fourier transform is to desintegrate the regular representation of $SE(2)$. This was the main ingredient of the computation of the hypoelliptic heat kernel in [4]. Using the Fourier transform again, the hypoelliptic heat equation is transformed into a family of parabolic equations. These are more suitable for standard numerical methods.

Roughly speaking, the non-commutative Fourier transform $\hat{f}(\Lambda)$ of the function $f(x, y, \theta) =: f(X, \theta)$, for $\Lambda \in \mathbb{R}^2$, is an operator meeting:

$$[\hat{f}(\Lambda)\psi](\theta) = \int_X \int_{S^1} f(X, \alpha)\psi(\alpha + \theta) d\alpha e^{2\pi i \langle R_{-\theta}\Lambda, X\rangle} dx = (\widetilde{f *_\theta \psi})(R_{-\theta}\Lambda), \quad (3.1)$$

where $*_\theta$ is the convolution with respect to the angular variable and $\widetilde{}$ is the 2-D Fourier transform with respect to the spatial variables $X = (x, y)$.

Then it is natural to consider the Fourier transform with respect to X. Indeed, apply this transform $u \to \tilde{u}$ to the initial value problem:

$$\begin{cases} \partial_t u = \Delta_H u \\ u(0, X, \theta) = \bar{f}_c(X, \theta), \end{cases} \quad (3.2)$$

that gives

$$\begin{cases} \partial_t \tilde{u} = \beta^2 \partial_\theta^2 \tilde{u} - 4\pi^2 (x \cos(\theta) + y \sin(\theta))^2 \tilde{u} \\ \tilde{u}(0, X, \theta) = \tilde{\bar{f}}_c(X, \theta). \end{cases} \quad (3.3)$$

Hence, for each point in the Fourier space, we have to solve an evolution equation with Mathieu right-hand term.

This is the principle of the algorithm, which is "massively parallelizable" in the sense that we solve simultaneously the equation (3.3) at each point of the Fourier space.

3.3. Results of image reconstruction. In this section we provide results of image reconstruction using the algorithm presented above. For these examples, we have tuned the parameters β, that is the relative weight, and T, the final time of evolution.

Notice again that this algorithm processes the image globally and does not need the information about where the image is corrupted. The counterpart is that it modify also the non-corrupted part.

We present three result.

- Figure 3.1 shows an image which is corrupted in a small piece of it. Then the diffusion can be applied for a rather small time avoiding an important diffusion effect in the noncorrupted part.
- Figure 3.2 shows a strongly corrupted image. In this case a larger diffusion time is necessary to "inpaint" completely the corrupted part. The diffusion effect is clearly much more important. However in our opinion the result is surprisingly good.
- Due to pixelization of the image, one could think that corruption along the diagonal is the worst situation. Figure 3.3 show that this is not the case.

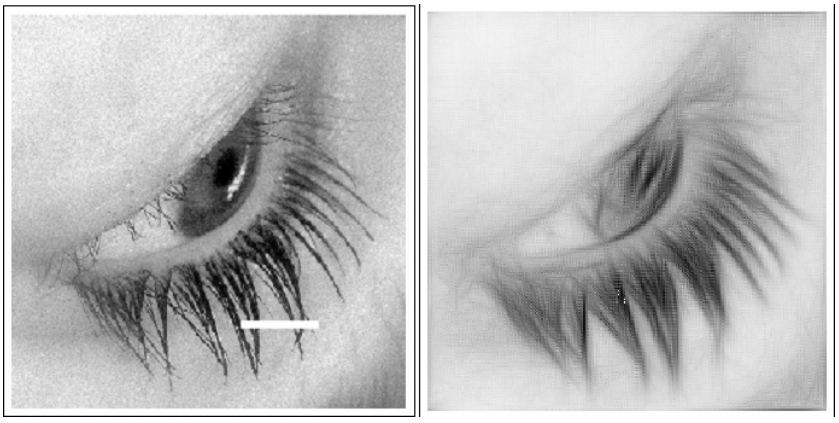

FIGURE 3.1. *Reconstruction of an image corrupted on a small portion. Here the diffusion is applied for a small time*

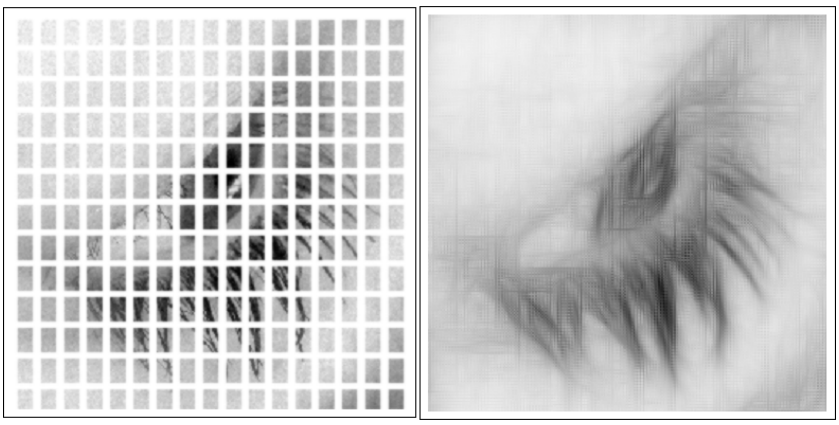

FIGURE 3.2. *Reconstruction of an image deeply corrupted. A larger time of diffusion is necessary*

Acknowledgements The authors are greatly indebted with A. Agrachev for his suggestions in the refinement of the mathematical model. The authors are also grateful to S. Masnou and R. Duits for very helpful discussions.

A Genericity of Morse properties of Gaussian convolution. In this appendix, we prove that, generically, the convolution of a L^2 function over a bounded domain $\mathcal{D} \subset \mathbb{R}^2$ with a Gaussian G is a Morse function. In particular, we first prove

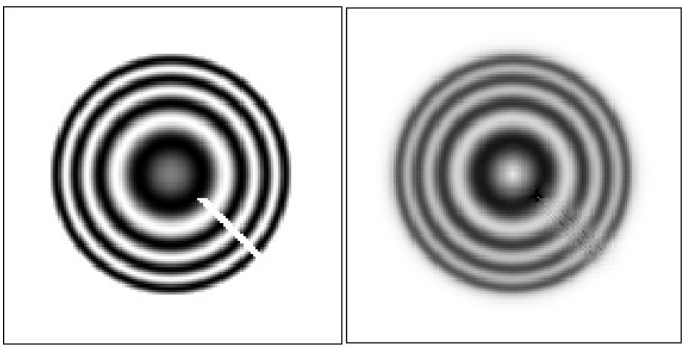

FIGURE 3.3. *Reconstruction of an image corrupted on the diagonal*

in Theorem 26 that the set of functions $\mathcal{I} \in L^2(\mathcal{D}, \mathbb{R})$ the convolution of which with a Gaussian is a Morse function is residual[5] in $L^2(\mathcal{D}, \mathbb{R})$. We then prove in Theorem 28 that the set of functions $\mathcal{I} \in L^2(\mathcal{D}, \mathbb{R})$ such that $\mathcal{I} * G$ **restricted to a compact** $K \subset \mathbb{R}^2$ is a Morse function is open and dense.

Definition 24. *Let Z, Y be C^1 manifolds, $F : Z \to Y$ a C^1 map and $W \subset Y$ a submanifold. We say that F is transversal to W at $z \in Z$, in symbols $F \pitchfork_z W$, if, where $y = F(z)$, either $y \notin W$ or $y \in W$ and*

1. *the inverse image $(T_z F)^{-1}(T_y W)$ splits and*
2. *the image $(T_z F)(T_z Z)$ contains a closed complement to $T_y W$ in $T_y Y$.*

We say that F is transversal to W, in symbols $F \pitchfork W$, if $F \pitchfork_z W$ for every $z \in Z$.

We recall that a closed subspace F of a Banach space E splits when there exists a closed subspace G such that $E = F \oplus G$.

Remark 25. If E is Hilbert, then every closed subspace splits. See [2, Prop. 2.1.15].

Theorem 26. *Let $\mathcal{D}$ be a bounded domain of the plane $\mathbb{R}^2$. Fix $\sigma_x, \sigma_y > 0$. Consider the convolution map*[6]

$$\Gamma : \left\{ \begin{array}{ccc} L^2(\mathcal{D}, \mathbb{R}) & \to & C^\infty(\mathbb{R}^2) \\ \mathcal{I} & \mapsto & \mathcal{I} * G, \end{array} \right.$$

where G is the Gaussian centred at $(0, 0)$

$$G(x, y) := \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2}}.$$

Let $\mathbb{X} := \left\{ \mathcal{I} \in L^2(\mathcal{D}, \mathbb{R}) \text{ s.t. } \rho(\mathcal{I}) \text{ is a Morse function} \right\}$. Then, $\mathbb{X}$ is residual in $L^2(\mathcal{D}, \mathbb{R})$.

Proof. The proof relies on parametric transversality Theorems. The version we use is Abraham's formulation, see [1, Th. 19.1], recalled in the following.

Theorem 27. *Let $\mathbb{A}, X, Y$ be C^r manifolds, $\rho : \mathbb{A} \to C^r(X, Y)$ a C^r representation, $W \subset Y$ a submanifold, and $\mathrm{ev}_\rho : X \times \mathbb{A} \to Y$ the evaluation map. Define $\mathbb{A}_W \subset \mathbb{A}$ by $\mathbb{A}_W = \{a \in \mathbb{A} \mid \rho_a \pitchfork W\}$. Assume that:*

1. *X has a finite dimension n and W has a finite codimension q in Y,*
2. *$\mathbb{A}$ and X are second countable,*

[5] We recall that a subset of a topological space is residual when it is a countable intersection of open and dense sets.

[6] $\mathcal{I}$ is considered to be zero outside $\mathcal{D}$.

3. $r > \max\{0, n-q\}$,

4. $\mathrm{ev}_\rho \pitchfork W$.

Then, $\mathbb{A}_W$ *is residual in* $\mathbb{A}$. We apply Theorem 27 with $\mathbb{A} = L^2(\mathcal{D}, \mathbb{R})$, $X = \mathbb{R}^2$, $r = 2$. We choose $Y = \mathbb{R}^2 \times \mathbb{R} \times \mathbb{R}^2 \times \mathbb{R}^3$ and ρ the 2-jets of $\Gamma(\mathcal{I})$, i.e.,

$$\rho : \left\{ \begin{array}{lcl} \mathbb{A} & \to & C^r(X,Y) \\ \mathcal{I} & \mapsto & (\Pi_1, \Pi_2, \Gamma(\mathcal{I}), \partial_x\Gamma(\mathcal{I}), \partial_y\Gamma(\mathcal{I}), \partial^2_{xx}\Gamma(\mathcal{I}), \partial^2_{xy}\Gamma(\mathcal{I}), \partial^2_{yy}\Gamma(\mathcal{I})) \end{array} \right.$$

where

$$\Pi_1 : \left\{ \begin{array}{lcl} X & \to & \mathbb{R} \\ (x,y) & \mapsto & x \end{array} \right. \quad \text{and} \quad \Pi_2 : \left\{ \begin{array}{lcl} X & \to & \mathbb{R} \\ (x,y) & \mapsto & y \end{array} \right.$$

are the canonical projections.

We fix

$$W = \left\{ (x, y, a, p_1, p_2, q_1, q_2, q_3) \in Y \text{ s.t. } (x,y) \in \mathbb{R}^2,\ p_1 = p_2 = 0,\ q_1 q_3 - q_2^2 = 0 \right\}.$$

A function $\mathcal{I} \in C^2(\mathbb{R}^2)$ is a Morse function if and only if

$$\begin{aligned} \mathrm{ev}_\rho(x, y, \mathcal{I}) = \rho_{\mathcal{I}}(x,y) = (x, y, \Gamma(\mathcal{I})(x,y), \partial_x\Gamma(\mathcal{I})(x,y), \partial_y\Gamma(\mathcal{I})(x,y), \\ \partial^2_{xx}\Gamma(\mathcal{I})(x,y), \partial^2_{xy}\Gamma(\mathcal{I})(x,y), \partial^2_{yy}\Gamma(\mathcal{I})(x,y)) \end{aligned}$$

does not belong to W for all $(x,y) \in \mathbb{R}^2$.

Remark that W is not a manifold. However, it is an algebraic set and hence it is a finite union of manifolds. In the following, we apply Theorem 27 as if W were a manifold, with the understanding that the Theorem is applied to each component.

We now verify each of the conditions 1-4 in Theorem 27. Condition 1 holds with $n = 2$ and $q \geq 3$ for each component of W. Condition 2 holds, since $\mathbb{A}$ and X are separable metric spaces and hence second countable. Condition 3 holds for each component of W.

Now we verify condition 4, that is the transversality condition $\mathrm{ev}_\rho \pitchfork W$. Fix $x, y, \mathcal{I}$ such that $\mathrm{ev}_\rho(\mathcal{I}, (x,y)) \in W$. Condition 1 in Definition 24 holds because of Remark 25. We now verify condition 2 in Definition 24, where $Z = \mathbb{R}^2 \times \mathbb{A}$. In the following, we prove that $(T_{(x,y,\mathcal{I})}\mathrm{ev}_\rho)(T_{(x,y,\mathcal{I})}(\mathbb{R}^2 \times \mathbb{A}))$ is the whole $T_{\mathrm{ev}_\rho(x,y,\mathcal{I})}Y$. The map $T_{(x,y,\mathcal{I})}\mathrm{ev}_\rho$ has the following triangular form

$$T_{(x,y,\mathcal{I})}\mathrm{ev}_\rho = \left(\begin{array}{cc|c} 1 & 0 & * \\ 0 & 1 & * \\ \hline 0 & 0 & T_{\mathcal{I}}\mathrm{ev}_\rho(x,y,\mathcal{I}) \end{array} \right) \tag{3.4}$$

We are left to prove that the tangent mapping $T_{\mathcal{I}}\mathrm{ev}_\rho(x,y,\mathcal{I})$ is surjective in $\mathbb{R} \times \mathbb{R}^2 \times \mathbb{R}^3$, for arbitrary (x,y) fixed. After a suitable change of coordinate, we can assume that $\sigma_x = \sigma_y = 1$ and that $(0,0) \in \mathcal{D}$. Let $\varepsilon > 0$ such that $\mathcal{D} \supset \mathbf{Q} := [-\varepsilon, \varepsilon] \times [-\varepsilon, \varepsilon]$. Define the function in $L^2(\mathcal{D}, \mathbb{R})$

$$\delta\mathcal{I}(\bar{x}, \bar{y}) = \frac{c_0 + c_1\bar{x} + c_2\bar{y} + c_3\bar{x}^2 + c_4\bar{x}\bar{y} + c_5\bar{y}^2}{G(x - \bar{x}, y - \bar{y})}$$

restricted to $\mathbf{Q}$, and zero in $\mathcal{D} \backslash \mathbf{Q}$. The map ρ is linear in $\mathcal{I}$, thus $T_{\mathcal{I}}\mathrm{ev}_\rho(x,y,\mathcal{I})\,[\delta\mathcal{I}] =$

$\mathrm{ev}_\rho(x,y,\delta\mathcal{I})$. Consider the linear operator

$$\mathrm{ev}_\rho(x,y,\delta\mathcal{I}) = \begin{pmatrix} \int_{\mathbf{Q}} \delta\mathcal{I}(\bar{x},\bar{y})G(x-\bar{x},y-\bar{y})\,d\bar{x}d\bar{y} \\ \int_{\mathbf{Q}} \delta\mathcal{I}(\bar{x},\bar{y})\partial_1 G(x-\bar{x},y-\bar{y})\,d\bar{x}d\bar{y} \\ \int_{\mathbf{Q}} \delta\mathcal{I}(\bar{x},\bar{y})\partial_2 G(x-\bar{x},y-\bar{y})\,d\bar{x}d\bar{y} \\ \int_{\mathbf{Q}} \delta\mathcal{I}(\bar{x},\bar{y})\partial^2_{11} G(x-\bar{x},y-\bar{y})\,d\bar{x}d\bar{y} \\ \int_{\mathbf{Q}} \delta\mathcal{I}(\bar{x},\bar{y})\partial^2_{12} G(x-\bar{x},y-\bar{y})\,d\bar{x}d\bar{y} \\ \int_{\mathbf{Q}} \delta\mathcal{I}(\bar{x},\bar{y})\partial^2_{22} G(x-\bar{x},y-\bar{y})\,d\bar{x}d\bar{y} \end{pmatrix}$$

as a function of the 6 variables $(c_0,\dots,c_5)$, and consider the linear system $\mathrm{ev}_\rho(x,y,\delta\mathcal{I}) = (a,p_1,p_2,q_1,q_2,q_3)$, where $(a,p_1,p_2,q_1,q_2,q_3) \in Y$ is fixed. A direct computation shows that the determinant of the system is $\frac{65536\varepsilon^{28}}{164025\sigma_x^8\sigma_y^8} > 0$, thus the system always has a solution, i.e. $T_{\mathcal{I}}\mathrm{ev}_\rho(x,y,\mathcal{I})$ is surjective.

By applying Theorem 27, we get $\mathbb{A}_W$ residual in $\mathbb{A}$. We now prove that $\mathbb{A}_W = \mathbb{X}$. Since $\mathcal{I} \in \mathbb{X}$ implies $\mathrm{ev}_\rho(x,y,\mathcal{I}) \notin W$, then $\rho_{\mathcal{I}} \pitchfork W$, hence $\mathbb{A} \supset \mathbb{X}$.

Now let us prove the inclusion $\mathbb{A} \subset \mathbb{X}$. Let $\mathcal{I} \in \mathbb{A}$ and fix $(x,y) \in \mathbb{R}^2$.

Nonintersection claim : $\rho_{\mathcal{I}}(x,y) \notin W$.

Proof of the claim. By contradiction, let

$$w = \mathrm{ev}_\rho(x,y,\mathcal{I}) \in W.$$

Since $\rho_{\mathcal{I}} \pitchfork_{(x,y)} W$, then $\left(T_{(x,y)}\rho_{\mathcal{I}}\right)\left(T_{(x,y)}\mathbb{R}^2\right)$ contains a closed complement to T_wW in T_wY.

Observe that

$$\dim\left(T_{(x,y)}\rho_{\mathcal{I}}\right)\left(T_{(x,y)}\mathbb{R}^2\right) \le \dim\left(T_{(x,y)}\mathbb{R}^2\right) = 2$$

and $\operatorname{codim} T_wW \ge 3$, thus $\left(T_{(x,y)}\rho_{\mathcal{I}}\right)\left(T_{(x,y)}\mathbb{R}^2\right)$ cannot contain a closed complement to T_wW in T_wY. A contradiction.

By applying the claim for each $(x,y) \in \mathbb{R}^2$, we get that $\rho_{\mathcal{I}}$ is a Morse function. □

Theorem 28. *Let K be a compact subset of $\mathbb{R}^2$ with non-empty interior. Under the hypothesis of Theorem 26, the set $\mathbb{X}_K := \left\{\mathcal{I} \in L^2(\mathcal{D},\mathbb{R})\ s.t.\ \rho_{\mathcal{I}|_K}\ \text{is a Morse function}\right\}$ is open and dense in $L^2(\mathcal{D},\mathbb{R})$.*

Proof. Applying the openness of nonintersection Theorem [1, Th. 18.1] and using the nonintersection claim, we get that $\mathbb{X}_K$ is an open subset of $L^2(\mathcal{D},\mathbb{R})$. Since $\mathbb{X}_K \supset \mathbb{X}$ and $\mathbb{X}$ is dense, then the conclusion holds. □

REFERENCES

[1] R. Abraham, J. Robbin, Transversal mappings and flows, W.A. Bejnamin, Inc. 1967.

[2] R. Abraham, J. E. Marsden, T. S. Ratiu, Manifolds, tensor analysis, and applications, Springer-Verlag, 1988.

[3] M. Abramowitz, I. A. Stegun (Editors), Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables, National Bureau of Standards Applied Mathematics, Washington, 1964.

[4] A. Agrachev, U. Boscain, J.-P. Gauthier, F. Rossi, The intrinsic hypoelliptic Laplacian and its heat kernel on unimodular Lie groups, *J. Funct. Anal.*, 256 (2009), pp. 2621–2655.

[5] A. Agrachev, Exponential mappings for contact sub-Riemannian structures, *J. Dynam. Control Systems*, 2 (1996), pp. 321–358.

[6] A.A. Agrachev, Yu. L. Sachkov, Control Theory from the Geometric Viewpoint, Encyclopedia of Mathematical Sciences, v. 87, Springer, 2004.
[7] J. August The Curve Indicator Random Field, PhD Thesis, http://www.cs.cmu.edu/∼jonas/
[8] A. Bellaiche, The tangent space in sub-Riemannian geometry, in *Sub-Riemannian geometry*, edited by A. Bellaiche and J.-J. Risler, pp. 1–78, Progr. Math., 144, Birkhäuser, Basel, 1996.
[9] U. Boscain, G. Charlot, F. Rossi, Existence of planar curves minimizing length and curvature, *Journal of Mathematical Sciences*, to appear, arXiv:0906.5290.
[10] U. Boscain, F. Rossi, Projective Reeds-Shepp car on S^2 with quadratic cost, *ESAIM: COCV*, 16, no. 2, pp. 275–297, 2010.
[11] F. Cao, Y. Gousseau, S. Masnou, P. Pérez, Geometrically guided exemplar-based inpainting, submitted.
[12] G. S. Chirikjian, A. B. Kyatkin, Engineering applications of noncommutative harmonic analysis, CRC Press, Boca Raton, FL, 2001.
[13] G. Citti, A. Sarti, A cortical based model of perceptual completion in the roto-translation space, *J. Math. Imaging Vision* 24 (2006), no. 3, pp. 307–326.
[14] G. Sanguinetti, G. Citti, A. Sarti Image completion using a diffusion driven mean curvature flow in a sub-riemannian space, in: , Int. Conf. on Computer Vision Theory and Applications (VISAPP'08), FUNCHAL, 2008, pp. 22-25.
[15] J. Damon, Generic structure of two-dimensional images under Gaussian blurring, *SIAM J. Appl. Math.*, Vol. 59, No. 1 (1998), pp. 97-138.
[16] R. Duits, M. van Almsick, The explicit solutions of linear left-invariant second order stochastic evolution equations on the 2D Euclidean motion group. *Quart. Appl. Math.* 66 (2008), 27-67.
[17] R. Duits, E.M.Franken, Left-invariant parabolic evolutions on SE(2) and contour enhancement via invertible orientation scores, Part I: Linear Left-Invariant Diffusion Equations on SE(2) *Quart. Appl. Math.*, 68, (2010), pp. 293-331.
[18] R. Duits, E.M.Franken, Left-invariant parabolic evolutions on SE(2) and contour enhancement via invertible orientation scores, Part II: nonlinear left-invariant diffusions on invertible orientation scores *Quart. Appl. Math.*, 68, (2010), pp. 255-292.
[19] E.M.Franken Enhancement of Crossing Elongated Structures in Images. Ph.D. thesis, Eindhoven University of Technology, 2008, Eindhoven. http://www.bmi2.bmt.tue.nl/Image-Analysis/People/EFranken/PhDThesisErikFranken.pdf .
[20] E.M.Franken, R. Duits Crossing-Preserving Coherence-Enhancing Diffusion on Invertible Orienattion Scores., IJCV, 85(3), 253–278, (2009).
[21] I.M. Gel'fand, G. E. Shilov, Generalized functions Vol. 1, Accademic press, 1964.
[22] M. Gromov, Carnot-Caratheodory spaces seen from within, in *Sub-Riemannian geometry*, Progr. Math., v. 144, pp. 79–323, Birkhäuser, Basel, 1996.
[23] R. K. Hladky, S. D. Pauls, Minimal Surfaces in the Roto-Translation Group with Applications to a Neuro-Biological Image Completion Model, *J Math Imaging Vis* 36, pp. 1–27, 2010.
[24] L. Hörmander, Hypoelliptic Second Order Differential Equations, *Acta Math.*, 119 (1967), pp. 147–171.
[25] M. Langer, Computational Perception, Lecture Notes, 2008, http://www.cim.mcgill.ca/∼langer/646.html
[26] D. Mumford, Elastica and computer vision. Algebraic Geometry and Its Applications. Springer-Verlag, pages 491-506, 1994.
[27] D. Marr; E. Hildreth, Theory of Edge Detection, *Proceedings of the Royal Society of London. Series B, Biological Sciences*, Vol. 207, No. 1167. (Feb. 29, 1980), pp. 187-217.
[28] I. Moiseev, Yu. L. Sachkov, Maxwell strata in sub-Riemannian problem on the group of motions of a plane, *ESAIM: COCV* 16, no. 2, pp. 380–399, 2010.
[29] L. Peichl, H. Wässle, Size, scatter and coverage of ganglion cell receptive field centres in the cat retina, *J Physiol*, Vol. 291, 1979, pp. 117-41.
[30] J. Petitot, Vers une Neuro-géomètrie. Fibrations corticales, structures de contact et contours subjectifs modaux, *Math. Inform. Sci. Humaines*, n. 145 (1999), pp. 5–101.
[31] J. Petitot, Neurogéomètrie de la vision - Modèles mathématiques et physiques des architectures fonctionnelles, Les Éditions de l'École Polythecnique, 2008.
[32] Yu. L. Sachkov, Conjugate and cut time in the sub-Riemannian problem on the group of motions of a plane, *ESAIM: COCV*, Volume 16, Number 4, pp. 1018–1039.
[33] Yu. L. Sachkov, Cut locus and optimal synthesis in the sub-Riemannian problem on the group of motions of a plane, *ESAIM: COCV*, to appear.