

Chapter 9

Deformable Objects and Matching Functionals



9.1 General Principles

In the previous two chapters, we introduced and studied basic tools related to deformations and their mathematical representation using diffeomorphisms. In this chapter, we start investigating relations between deformations and the objects they affect, which we will call deformable objects, and discuss the variations of matching functionals, which are cost functions that measure the quality of the registration between two deformable objects.

Let Ω be an open subset of \mathbb{R}^d and G a group of diffeomorphisms on Ω . Consider a set \mathcal{I} of structures of interest, on which G has an action: for every I in \mathcal{I} and every $\varphi \in G$, the result of the action of φ on I is denoted $\varphi \cdot I$ and is a new element of \mathcal{I} . This requires (see Sect. B.5) that $\text{id} \cdot I = I$ and $\varphi \cdot (\psi \cdot I) = (\varphi \circ \psi) \cdot I$. Elements of \mathcal{I} will be referred to as *deformable objects*.

A matching functional is based on a function $D : \mathcal{I} \times \mathcal{I} \rightarrow [0, +\infty)$ such that $D(I, I')$ measures the discrepancy between the two objects I and I' , and is defined over G by

$$E_{I,I'}(\varphi) = D(\varphi \cdot I, I'). \quad (9.1)$$

So $E_{I,I'}(\varphi)$ measures the difference between the *target* object I' and the deformed one $\varphi \cdot I$. Because it is mapped onto the target by the deformation, the object I will often be referred to as the *template* (and $\varphi \cdot I$ as the deformed template).

Even if our discussion of matching principles and algorithms is rather extensive, and occupies a large portion of this book, the size of the literature, and our choice of privileging methods that implement diffeomorphic matching prevents us from providing an exhaustive account of the registration methods that have been proposed over the last few decades. The interested reader can refer to a few starting points in order to complement the presentation that is made here, including [12, 13, 22, 27, 28, 41, 42, 111, 125, 240, 244, 275], and textbooks such as [132, 139, 208, 214].

9.2 Differentiation with Respect to Diffeomorphisms

We will review, starting with the next section, a series of matching functionals that are adapted to different types of deformable objects (landmarks, images, curves, etc.). We will also compute the derivative of each of them with respect to the diffeomorphism φ .

We also introduce a special form of differential which is adapted to variational problems over diffeomorphisms. This shape, or Eulerian differential, as we will call it, is a standard tool in shape optimization [80], and we will interpret it later on as a gradient for a specific Riemannian metric over diffeomorphisms.

Recall that we have defined $\text{Diff}^{p,\infty} = \text{Diff}^{p,\infty}(\Omega)$ to be the set of diffeomorphisms ψ such that

$$\max(\|\psi - \text{id}\|_{p,\infty}, \|\psi^{-1} - \text{id}\|_{p,\infty}) < \infty.$$

We have also defined $\text{Diff}_0^{p,\infty}$ as the subgroup of $\text{Diff}^{p,\infty}$ whose elements converge to the identity at infinity.

Definition 9.1 A function $\varphi \mapsto U(\varphi)$ is (p, ∞) -compliant if it is defined for all φ in $\text{Diff}_0^{p,\infty}$.

A (p, ∞) -compliant U is locally (p, ∞) -Lipschitz if, for all $\varphi \in \text{Diff}_0^{p,\infty}$, there exist positive numbers $\varepsilon(\varphi)$ and $C(\varphi)$ such that

$$|U(\psi) - U(\tilde{\psi})| \leq C(\varphi)\|\psi - \tilde{\psi}\|_{p,\infty}$$

whenever ψ and $\tilde{\psi}$ are diffeomorphisms such that

$$\max(\|\psi - \varphi\|_{p,\infty}, \|\tilde{\psi} - \varphi\|_{p,\infty}) < \varepsilon(\varphi).$$

Note that a (p, ∞) -compliant (resp. locally Lipschitz) U is (q, ∞) -compliant (resp. locally Lipschitz) for any q larger than p .

Because $\text{Diff}_0^{p,\infty}$ is an open subset of $\text{id} + C_0^p(\Omega, \mathbb{R}^d)$, both Gâteaux and Fréchet derivatives are well defined for functions defined on this set (see Sect. C.1). In the following, whenever we speak of a derivative (without a qualifier), this will always mean in the strong (Fréchet) sense. A function U is C^1 on $\text{Diff}_0^{p,\infty}$ if and only if U is Fréchet differentiable and $dU(\psi)$ is continuous in ψ , which is equivalent (by Proposition C.5) to U being Gâteaux differentiable and $dU(\psi)$ continuous in ψ . Note also that U being C^1 implies that U is (p, ∞) -Lipschitz.

Using the group structure of $\text{Diff}_0^{p,\infty}$, we can define another type of differential using the infinitesimal action of vector fields. If V is an admissible vector space and $v \in V$, we will denote by $\varphi_{0_t}^v$ the flow associated to the equation

$$\partial_t y = v(y).$$

Note that this is the same notation as the flow associated to a differential equation $\partial_t y = v(t, y)$, where v is now a time-dependent vector field. This is not a conflict of notation if one agrees to identify vector fields, v , in V and the associated constant time-dependent vector field defined by $\tilde{v}(t, \cdot) = v$ for all t .

Definition 9.2 Let V be an admissible Hilbert space continuously embedded in $C_0^p(\Omega, \mathbb{R}^d)$ (so that $\text{Diff}_V \subset \text{Diff}_0^{p,\infty}$). We say that a (p, ∞) -compliant function U over diffeomorphisms has an Eulerian differential in V at ψ if there exists a linear form $\bar{\partial}U(\psi) \in V^*$ such that, for all $v \in V$,

$$(\bar{\partial}U(\psi) \mid v) = \partial_\varepsilon U(\varphi_{0\varepsilon}^v \circ \psi)|_{\varepsilon=0}. \tag{9.2}$$

If the Eulerian differential exists, the V -Eulerian gradient of U at ψ , denoted $\bar{\nabla}^V U(\varphi) \in V$, is defined by

$$\langle \bar{\nabla}^V U(\varphi), v \rangle_V = (\bar{\partial}U(\varphi) \mid v). \tag{9.3}$$

In this case, $\bar{\nabla}^V U(\varphi) = \mathbb{K}\bar{\partial}U(\varphi)$, where \mathbb{K} is the kernel operator of V .

The following proposition indicates when Eq.(9.2) remains valid with time-dependent vector fields v .

Proposition 9.3 *Let V be an admissible Hilbert space continuously embedded in $C_0^{p+1}(\Omega, \mathbb{R}^d)$. Let V and U satisfy the hypotheses of Definition 9.2. If U is (p, ∞) -locally Lipschitz and has a V -Eulerian differential at ψ and if $v(t, \cdot)$ is a time-dependent vector field such that*

$$\lim_{\varepsilon \rightarrow \infty} \frac{1}{\varepsilon} \int_0^\varepsilon \|v(t, \cdot) - v(0, \cdot)\|_V dt = 0, \tag{9.4}$$

then

$$(\bar{\partial}U(\psi) \mid v(0, \cdot)) = \partial_\varepsilon U(\varphi_{0\varepsilon}^v \circ \psi)|_{\varepsilon=0}. \tag{9.5}$$

Proof Letting $v_0 = v(0, \cdot)$, we need to prove that

$$\frac{1}{\varepsilon} (U(\varphi_{0\varepsilon}^v \circ \psi) - U(\varphi_{0\varepsilon}^{v_0} \circ \psi)) \rightarrow 0$$

as $\varepsilon \rightarrow 0$. From Proposition 7.4, we know that if $\psi, \varphi, \tilde{\varphi}$ are in $\text{Diff}_0^{p,\infty}$, there exists a constant $C_p(\psi)$ such that

$$\|\varphi \circ \psi - \tilde{\varphi} \circ \psi\|_{p,\infty} \leq C_p(\psi) \|\varphi - \tilde{\varphi}\|_{p,\infty}.$$

Now, since U is Lipschitz, we have, for small enough ε ,

$$\begin{aligned}
|U(\varphi_{0\varepsilon}^v \circ \psi) - U(\varphi_{0\varepsilon}^{v_0} \circ \psi)| &\leq C(\psi) \|\varphi_{0\varepsilon}^v \circ \psi - \tilde{\varphi}_{0\varepsilon}^v \circ \psi\|_{p,\infty} \\
&\leq C(\psi) C_p(\psi) \|\varphi_{0\varepsilon}^v - \varphi_{0\varepsilon}^{v_0}\|_{p,\infty} \\
&\leq C(\psi) C_p(\psi) \tilde{C}(v_0) \int_0^\varepsilon \|v(t, \cdot) - v_0\|_V dt,
\end{aligned}$$

where $\tilde{C}(v_0)$ depends on $\|v_0\|_{p+1,\infty}$ and can be derived from Eq. (7.16), Noting that $\|v\|_{\mathcal{X}^{p+1,1,\varepsilon}} \leq \varepsilon(C'' + \|v_0\|_{p+1,\infty})$ for small enough ε , this proves the proposition. \square

Note also that, if U is C^1 , then the chain rule implies that

$$(\bar{\partial}U(\varphi) \mid v) = (dU(\varphi) \mid v \circ \varphi). \quad (9.6)$$

To the Eulerian gradient of U , we associate a “gradient descent” process (that we will formally interpret as a Riemannian gradient descent for a suitable metric in Sect. 11.4.3) which generates a time-dependent element of G by setting

$$\partial_t \varphi(t, x) = -\bar{\nabla}^V U(\varphi(t))(\varphi(t, x)). \quad (9.7)$$

As long as $\int_0^t \left\| \bar{\nabla}^V U(\varphi(s)) \right\|_V ds$ is finite, this generates a time-dependent element of Diff_V . This therefore provides an evolution within the group of diffeomorphisms, an important property. Assuming that Proposition 9.3 applies at time t (e.g., if U is C^1), we can write

$$\partial_t U(\varphi(t)) = \langle \bar{\nabla}^V U(\varphi(t)), \partial_t \varphi \rangle_V = - \left\| \bar{\nabla}^V U(\varphi(t)) \right\|_V^2,$$

so that $U(\varphi(t))$ decreases with time.

9.3 Relation with Matching Functionals

As pointed out in the introduction, matching functionals take the form

$$U(\varphi) = U_I(\varphi) = Z(\varphi \cdot I), \quad (9.8)$$

where I is a fixed deformable object for some function Z (e.g., $Z(I) = D(I, I_1)$ for a fixed I_1). Using the group action property, we have

$$U_I(\psi) = U_{\varphi, I}(\psi \circ \varphi^{-1}).$$

Using this property and the fact that the mapping $\psi \circ \varphi^{-1}$ is smooth (infinitely differentiable) from $\text{Diff}_0^{p,\infty}$ onto itself, we find that if U_I is Gâteaux or Fréchet

differentiable at $\psi = \text{id}$ for any $I \in \mathcal{I}$, then it is differentiable at all $\psi \in \text{Diff}_0^{p,\infty}$, and

$$(dU_I(\varphi) \mid h) = (dU_{\varphi \cdot I}(\text{id}) \mid h \circ \varphi^{-1}).$$

A similar statement holds for the Eulerian differential, with

$$(\bar{\partial}U_I(\psi) \mid v) = (\bar{\partial}U_{\psi \cdot I}(\text{id}) \mid v).$$

Notice that when U is differentiable at $\psi = \text{id}$, then $\bar{\partial}U(\text{id}) = dU(\text{id})$. Finally, if we assume that \mathcal{I} is itself a Banach space (or an open subset of a Banach space), that Z is differentiable and that the action $R_I : \varphi \mapsto \varphi \cdot I$ is also differentiable, we have, using the chain rule

$$(dU_I(\varphi) \mid v) = (dZ(\varphi \cdot I) \mid dR_I(\varphi)v)$$

or $dU_I(\varphi) = dR_I(\varphi)^* dZ(\varphi \cdot I)$. At $\varphi = \text{id}$, $dR_I(\text{id})v$ is the *infinitesimal action* of v on I , which we denote by $v \cdot I$, so that

$$(dU_I(\text{id}) \mid v) = (dZ(I) \mid v \cdot I)$$

and

$$(\bar{\partial}U_I(\psi) \mid v) = (dZ(\psi \cdot I) \mid v \cdot (\psi \cdot I)).$$

We now present a series of matching problems, involving different types of deformable objects. In each case, we will introduce adapted matching functionals and compute their differentials. As just remarked, derivatives with respect to the diffeomorphisms can all be derived from that of the function Z , on which we will, whenever possible, focus the computations.

9.4 Labeled Point Matching

The simplest way to represent a visual structure is with configurations of labeled points, or *landmarks* attached to the structure. Anatomical shapes or images are typical examples of structures on which landmarks can be easily defined; this includes specific locations in faces (corners of the eyes, tip of the nose, etc.), fingertips for hands, apex of the heart, etc. Many man-made objects, like cars or other vehicles, can be landmarked too. Finally, landmarks can represent the centers of simple objects, like cells in biological images.

In the labeled point-matching problem, objects are ordered collections of N points $x_1, \dots, x_N \in \Omega$, where N is fixed. Diffeomorphisms act on such objects by:

$$\varphi \cdot (x_1, \dots, x_N) = (\varphi(x_1), \dots, \varphi(x_N)). \quad (9.9)$$

The landmark-matching problem is not to find correspondences between two objects, say $I = (x_1, \dots, x_N)$ and $I' = (x'_1, \dots, x'_N)$, since we know that x_i and x'_i are homologous, but to extrapolate these correspondences to the rest of the space.

Here we can take $\mathcal{I} = (\mathbb{R}^d)^N$, or, if one restricts to distinct landmarks, the open subset

$$\mathcal{I} = \{(x_1, \dots, x_N) \in (\mathbb{R}^d)^N, x_i \neq x_j \text{ if } i \neq j\}.$$

For $I = (x_1, \dots, x_N)$, the action $\varphi \mapsto \varphi \cdot I$ is C^1 on $\text{Diff}_0^{p,\infty}$ for any $p \geq 0$ (it is the restriction of a linear map, with $|\varphi \cdot I|^2 \leq \sqrt{N} \|\varphi\|_\infty$). The simplest matching functional that we can consider for this purpose is associated with

$$Z(I) = |I - I'|^2 = \sum_{k=1}^N |x_k - x'_k|^2$$

with $(dZ(I) | h) = 2(I - I')^T h$ (considering I, I' and h as dN -dimensional column vectors). We have

$$U_I(\varphi) = E_{I,I'}(\varphi) = |I' - \varphi \cdot I|^2 = \sum_{i=1}^N |x'_i - \varphi(x_i)|^2, \quad (9.10)$$

$$(dU_I(\varphi) | v) = 2 \sum_{i=1}^N (\varphi(x_i) - x'_i)^T v(x_i). \quad (9.11)$$

This can be written as

$$dU_I(\varphi) = 2 \sum_{i=1}^N (\varphi(x_i) - x'_i) \delta_{x_i}.$$

From (9.6), we have

$$\left(\bar{\partial} U_I(\varphi) \middle| h \right) = 2 \sum_{i=1}^N (\varphi(x_i) - x'_i)^T h \circ \varphi(x_i) \quad (9.12)$$

or

$$\bar{\partial} U_I(\varphi) = 2 \sum_{i=1}^N (\varphi(x_i) - x'_i) \delta_{\varphi(x_i)},$$

and (9.3) gives

$$\bar{\nabla}^V U_I(\varphi) = 2 \sum_{i=1}^N K(\cdot, \varphi(x_i)) (\varphi(x_i) - x'_i). \quad (9.13)$$

The gradient descent algorithm (9.7) takes a very simple form:

$$\partial_t \varphi(t, x) = -2 \sum_{i=1}^N K(\varphi(t, x), \varphi(t, x_i))(\varphi(t, x_i) - x'_i). \quad (9.14)$$

This system can be solved in two steps: let $y_i(t) = \varphi(t, x_i)$. Applying (9.14) at $x = x_j$ yields

$$\partial_t y_j = -2 \sum_{i=1}^N K(y_j, y_i)(y_i - x'_i).$$

This is a differential system in y_1, \dots, y_N . The first step is to solve it with initial conditions $y_j(0) = x_j$. Once this is done, the extrapolated value of $\varphi(t, x)$ for a general x is the solution of the differential equation

$$\partial_t y = -2 \sum_{i=1}^N K(y, y_i)(y_i - x'_i)$$

initialized at $y(0) = x$. Figure 9.1 gives an example obtained by running this procedure, providing an illustration of the impact of the choice of the kernel for the solution. The last panel in Fig. 9.1 also shows the limitations of this algorithm, in the sense that it is trying to move the points in the direction of their targets at each step, while a more indirect path can sometimes be found generating less distortion (these results should be compared to Fig. 10.1 in Chapter 10).

9.5 Image Matching

Images, or more generally multivariate functions, are also important and widely used instances of deformable objects. They correspond to functions I defined on Ω with values in \mathbb{R} . Diffeomorphisms act on them by:

$$(\varphi \cdot I)(x) = I(\varphi^{-1}(x))$$

for $x \in \Omega$. Fixing two such functions I and I' , the simplest matching functional which can be considered is the squared L^2 norm of the difference $Z(I) = \|I - I'\|_2^2$, yielding

$$U_I(\varphi) = E_{I,I'}(\varphi) = \int_{\Omega} |I \circ \varphi^{-1}(x) - I'(x)|^2 dx. \quad (9.15)$$

We will need the derivative of the mapping $\text{Inv} : \varphi \mapsto \varphi^{-1}$. Considering Inv as a mapping from $\text{Diff}_0^{p+1, \infty}$ to $\text{Diff}_0^{p, \infty}$, it is given by (see Proposition 7.8)

$$d\text{Inv}(\varphi)h = - (d\varphi \circ \varphi^{-1})^{-1} h \circ \varphi^{-1} = -d(\varphi^{-1})h \circ \varphi^{-1}. \quad (9.16)$$

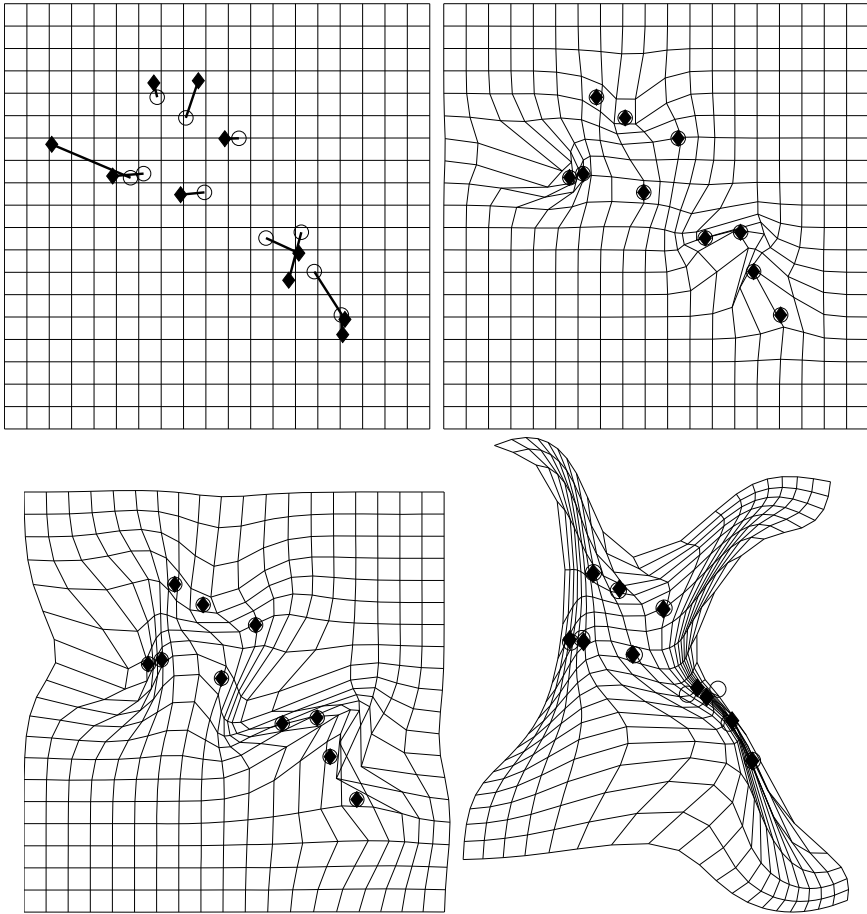


Fig. 9.1 Greedy landmark matching. Implementation of the gradient descent algorithm in (9.14), starting with $\varphi = \text{id}$, for the correspondences depicted in the upper-left image (diamonds moving to circles). The following three images provide the result after numerical convergence for Gaussian kernels $K(x, y) = \exp(-|x - y|^2/2\sigma^2)\text{Id}$ with $\sigma = 1, 2, 4$ in grid units. Larger σ induce increasing smoothness in the final solution, and deformations affecting a larger part of the space. As seen in the figure for $\sigma = 4$, the evolution can result in huge deformations

Similarly Inv is C^k from $\text{Diff}_0^{p+k, \infty}$ to Diff_0^p .

We now compute the derivative of U_I under the assumption that I' is square integrable, I is compactly supported (on some set Q_I) and continuously differentiable. One can relax the differentiability assumption on I (considering, for example, piecewise smooth images), but the analysis is much more difficult, and we refer the reader to results in [293–295] for more details. Define

$$\tilde{U}_I(\varphi) = \int_{\Omega} |I \circ \varphi(x) - I'(x)|^2 dx$$

so that $U_I = \tilde{U}_I \circ \text{Inv}$. Fixing $\varphi \in \text{Diff}_0^{p+1}$, $h \in C_0^{p+1}(\Omega, \mathbb{R}^d)$ and letting $\varphi_\varepsilon = \varphi + \varepsilon h$ for $|\varepsilon| \leq 1$, we have

$$\partial_\varepsilon \tilde{U}_I(\varphi_\varepsilon) = 2 \int_\Omega (I \circ \varphi_\varepsilon(x) - I'(x)) \nabla I \circ \varphi_\varepsilon(x)^T h(x) dx.$$

The integrand in the right-hand side is dominated by the integrable upper bound $(\|I\|_\infty + |I'(x)|) \|\nabla I\|_\infty \|h\|_\infty \mathbf{1}_{\tilde{Q}}(x)$, where \tilde{Q} is any compact set that contains $\varphi_\varepsilon^{-1}(Q_I)$ for $|\varepsilon| \leq 1$, which justifies the derivation. Taking $\varepsilon = 0$, we obtain the directional derivative of \tilde{U}_I , which is

$$\left(d\tilde{U}_I(\varphi) \mid h \right) = 2 \int_\Omega (I \circ \varphi(x) - I'(x)) \nabla I \circ \varphi(x)^T h(x) dx.$$

Our hypotheses on I imply that $d\tilde{U}_I(\varphi)$ is continuous in $\varphi \in \text{Diff}_0^{p,\infty}$ for any $p \geq 0$. Fixing φ and assuming $\|\varphi' - \varphi\|_\infty \leq 1$, we have

$$\begin{aligned} & \left| \left(d\tilde{U}_I(\varphi) - d\tilde{U}_I(\varphi') \mid h \right) \right| \\ & \leq 2 \int_\Omega (I \circ \varphi(x) - I \circ \varphi'(x)) \nabla I \circ \varphi'(x)^T h(x) dx \\ & \quad + 2 \int_\Omega (I \circ \varphi(x) - I'(x)) (\nabla I \circ \varphi(x) - \nabla I \circ \varphi'(x))^T h(x) dx \\ & \leq 2|\tilde{Q}| \|I \circ \varphi - I \circ \varphi'\|_\infty \|\nabla I\|_\infty \|h\|_\infty \\ & \quad + 2\sqrt{|\tilde{Q}|} \|I \circ \varphi - I'\|_2 \|\nabla I \circ \varphi - \nabla I \circ \varphi'\|_\infty \|h\|_\infty, \end{aligned}$$

where \tilde{Q} is a compact set containing all $(\varphi')^{-1}(Q_I)$ for $\|\varphi' - \varphi\|_\infty \leq 1$. The facts that I and ∇I are uniformly continuous on Q_I imply that $\|I \circ \varphi - I \circ \varphi'\|_\infty$ and $\|\nabla I \circ \varphi - \nabla I \circ \varphi'\|_\infty$ tend to 0 as $\|\varphi - \varphi'\|_\infty$ tends to 0. (We have denoted by $|\tilde{Q}|$ the Lebesgue measure of the set \tilde{Q} .)

As a composition of C^1 functions, we find that U_I is C^1 on $\text{Diff}_0^{p+1,\infty}$ for any $p \geq 0$, with (applying the chain rule)

$$\begin{aligned} (dU_I(\varphi) \mid h) &= \\ & - 2 \int_\Omega (I \circ \varphi^{-1}(x) - I'(x)) \nabla I \circ \varphi^{-1}(x)^T d(\varphi^{-1})h \circ \varphi^{-1}(x) dx. \end{aligned} \quad (9.17)$$

Notice that $\nabla(I \circ \varphi^{-1})^T = (\nabla I^T \circ \varphi^{-1})d(\varphi^{-1})$, so that

$$(dU_I(\varphi) \mid h) = -2 \int_\Omega (I \circ \varphi^{-1}(x) - I'(x)) \nabla(I \circ \varphi^{-1})(x)^T h \circ \varphi^{-1}(x) dx$$

and we retrieve the formula

$$(dU_I(\varphi) \mid h) = (dU_{\varphi \cdot I}(\text{id}) \mid h \circ \varphi^{-1}).$$

The Eulerian derivative is given by

$$\begin{aligned} (\bar{\partial}U_I(\varphi) \mid v) &= (dU_I(\text{id}) \mid v \circ \varphi) \\ &= -2 \int_{\Omega} (I \circ \varphi^{-1}(x) - I'(x)) \nabla(I \circ \varphi^{-1})(x)^T v(x) dx. \end{aligned}$$

We introduce a notation that will be used throughout this chapter and that generalizes the one given for point measures in Eq. (8.4). If μ is a measure on Ω and $z : \Omega \rightarrow \mathbb{R}^d$ a μ -measurable function, the vector measure $(z\mu)$ is the linear form over vector fields on Ω defined by

$$(z\mu \mid h) = \int_{\Omega} z^T h d\mu. \quad (9.18)$$

With this notation we can write

$$\bar{\partial}U_I(\varphi) = -2((I \circ \varphi^{-1} - I') \nabla(I \circ \varphi^{-1})) dx. \quad (9.19)$$

Notice also that, making a change of variable in (9.17), we have

$$dU_I(\varphi) = -2(\det(d\varphi) (I - I' \circ \varphi) d\varphi^{-T} \nabla I) dx. \quad (9.20)$$

To compute the Eulerian gradient of U_I , we need to apply the kernel operator, \mathbb{K} , to $\bar{\partial}U_I(\varphi)$, which requires the following lemma.

Lemma 9.4 *If V is a reproducing kernel Hilbert space (RKHS) of vector fields on Ω with kernel operator \mathbb{K} and kernel K , μ is a measure on Ω and z a μ -measurable function from Ω to \mathbb{R}^d , then, for all $x \in \Omega$,*

$$\mathbb{K}(z\mu)(x) = \int_{\Omega} K(x, y) z(y) d\mu(y).$$

Proof From the definition of the kernel, we have, for any $a \in \mathbb{R}^d$:

$$\begin{aligned} a^T \mathbb{K}(z\mu)(x) &= (a\delta_x \mid \mathbb{K}(z\mu)) \\ &= (z\mu \mid \mathbb{K}(a\delta_x)) \\ &= (z\mu \mid K(\cdot, x)a) \\ &= \int_{\Omega} z^T(y) K(y, x) a d\mu(y) \\ &= a^T \int_{\Omega} K(x, y) z(y) d\mu(y), \end{aligned}$$

which proves Lemma 9.4. □

The expression of the Eulerian gradient of U_I is now given by Lemma 9.4:

$$\bar{\nabla}^V U_I(\varphi) = -2 \int_{\Omega} (I \circ \varphi^{-1}(y) - I'(y))K(\cdot, y)\nabla(I \circ \varphi^{-1})(y)dy. \tag{9.21}$$

This provides the following “greedy” image-matching algorithm [67, 278].

Algorithm 9.5 Greedy image matching Start with $\varphi(0) = \text{id}$ and solve the evolution equation

$$\partial_t \varphi(t, y) = 2 \int_{\Omega} (J(t, x) - I'(x))K(\varphi(t, y), x)\nabla J(t, x)dx \tag{9.22}$$

with $I(t, \cdot) = I \circ (\varphi(t))^{-1}$.

This algorithm can also be written uniquely in terms of the evolving image, J , using $\partial_t J \circ \varphi + (J \circ \varphi)^T \partial_t \varphi = 0$. This yields

$$\partial_t J(t, y) = -2 \int_{\Omega} K(y, x)(J(t, x) - I'(x))\nabla J(t, x)^T \nabla I(t, y)dx.$$

In contrast to what we did in the landmark case, this algorithm should not be run indefinitely (or until numerical convergence). The fundamental difference is that, in the landmark case, there is an infinity of solutions to the diffeomorphic interpolation problem, and the greedy algorithm would generally run until it finds one of them and then stabilize. In the case of images, it is perfectly possible (and even typical) that there is no solution to the matching problem, i.e., no diffeomorphism φ such that $I \circ \varphi^{-1} = I'$. In that case, Algorithm 9.5 will run indefinitely, creating huge deformations while trying to solve an impossible problem.

To decide when the evolution should be stopped, an interesting suggestion has been made in [278]. Define

$$v(t, x) = 2 \int_{\Omega} (J(t, x) - I'(x))K(y, x)\nabla J(t, x)dx$$

so that (9.22) reduces to $\partial_t \varphi = v(t) \circ \varphi$. As we know from Chap. 7, the smoothness of φ at time t can be controlled by

$$\int_0^t \|v(s)\|_V^2 ds,$$

the norm being explicitly given by

$$\begin{aligned} & \|v(s)\|_V^2 \\ &= 2 \int_{\Omega \times \Omega} K(y, x)(J(s, x) - I'(x))(J(s, y) - I'(y))\nabla J(s, x)^T \nabla J(s, y)dxdy. \end{aligned}$$

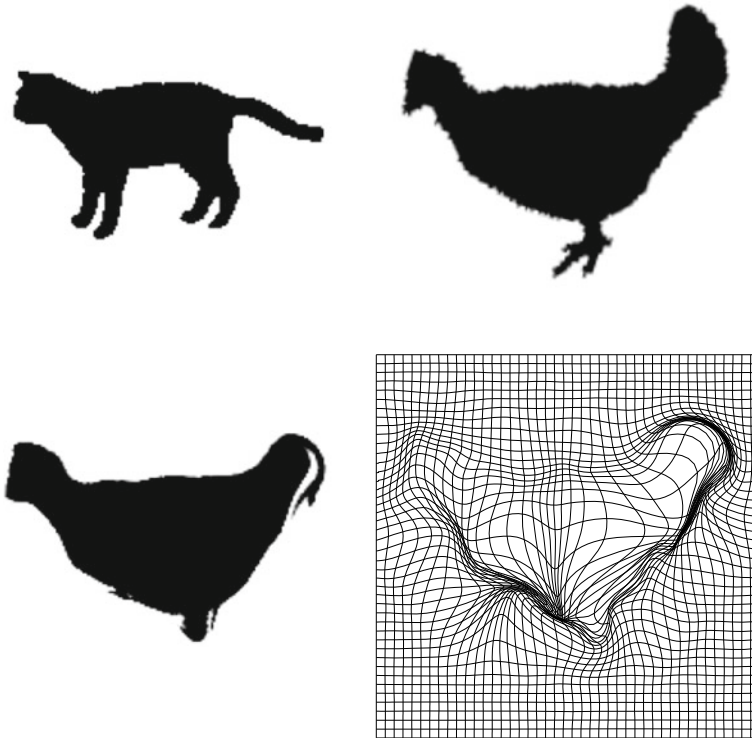


Fig. 9.2 Greedy image matching. Output of Algorithm 9.5 when estimating a deformation of the first image to match the second one. The third image is the obtained deformation of the first one and the last provides the deformation applied to a grid

Define, for some parameter λ ,

$$E(t) = \frac{1}{t} \int_0^t \|v(s)\|_V^2 ds + \lambda \int_{\Omega} (J(t, y) - I'(y))^2 dy.$$

Then, the stopping time proposed in [278] for Algorithm 9.5 is the first t at which $E(t)$ stops decreasing. Some experimental results using this algorithm and stopping rule are provided in Fig. 9.2.

There are many other possible choices for a matching criterion, least squares being, as we wrote, the simplest one. Among other possibilities, comparison criteria involving histograms provide an interesting option, because they allow for contrast-invariant comparisons.

Given a pair of images, I, I' , associate to each $x \in \Omega$ and image values λ and λ' the local histogram $H_x(\lambda, \lambda')$, which counts the frequency of simultaneous occurrence of values λ in I and λ' in I' at the same location in a small window around x . One computationally feasible way to define it is to use the kernel estimator

$$H_{I,I'}(x, \lambda, \lambda') = \int_{\Omega} f(|I(y) - \lambda|)f(|I'(y) - \lambda'|)g(x, y)dy$$

in which f is a positive function such that $\int_{\mathbb{R}} f(t)dt = 1$ and f vanishes when t is far from 0, and $g \geq 0$ is such that for all x , $\int_{\Omega} g(x, y)dy = 1$ and $g(x, y)$ vanishes when y is far from x .

For each x , $H_{I,I'}(x, \cdot, \cdot)$ is a bi-dimensional probability function, and there exist several ways of measuring the degree of dependence between its components. The simplest one, which is probably sufficient for most applications, is the correlation ratio, given by

$$C_{I,I'}(x) = 1 - \frac{\int_{\mathbb{R}^2} \lambda\lambda' H_{I,I'}(x, \lambda, \lambda')d\lambda d\lambda'}{\sqrt{\int_{\mathbb{R}^2} \lambda^2 H_{I,I'}(x, \lambda, \lambda')d\lambda d\lambda' \int_{\mathbb{R}^2} (\lambda')^2 H_{I,I'}(x, \lambda, \lambda')d\lambda d\lambda'}}$$

It is then possible to define the matching function by

$$U_I(\varphi) = \int_{\Omega} C_{I \circ \varphi^{-1}, I'}(x)dx.$$

The differential of U_I with respect to φ can be obtained after a lengthy (but elementary) computation. Some details can be found in [145]. A slightly simpler option is to use criteria based on the global histogram, which is defined by

$$H_{I,I'}(\lambda, \lambda') = \int_{\Omega} f(|I(y) - \lambda|)f(|I'(y) - \lambda'|)dy,$$

and the matching criterion is simply $U_I(\varphi) = C_{I \circ \varphi^{-1}, I'}$ or, as introduced in [185, 298], the mutual information computed from the joint histogram.

9.6 Measure Matching

The running assumption in Sect. 9.4 was that the point sets (x_1, \dots, x_N) were labeled, so that, when comparing two of them, the correspondences were known and the problem was to extrapolate them to the whole space.

In some applications, correspondences are not given and need to be inferred as part of the matching problem. One way to handle this is to include them as new unknowns (in addition to the unknown diffeomorphism), add extra terms to the energy that measures the quality of correspondences, and minimize the whole thing. Such an approach is taken, for example, in [240, 241].

Another point of view is to start with a representation of the point set that does not depend on how the points are ordered. A natural mathematical representation of a subset of \mathbb{R}^d is by the uniform measure on this set, at least when this is well-defined.

For a very general class of sets, this corresponds to the Hausdorff measure for the appropriate dimension [107], which, for finite sets, simply provides the sum of Dirac measures at each point, i.e., $x = (x_1, \dots, x_N)$ is represented by

$$\mu_x = \sum_{i=1}^N \delta_{x_i}.$$

For us, this raises the issue of comparing measures using diffeomorphisms, which will be referred to as the measure-matching problem.

In line with all other matching problems we are considering in this chapter, specifying the measure-matching problem requires, first, defining the action of diffeomorphisms on the considered objects, and second, using a good comparison criterion between two objects.

Let us start with the action of diffeomorphisms. The only fact we need here concerning measures is that they are linear forms acting on functions on \mathbb{R}^d via

$$(\mu \mid f) = \int_{\mathbb{R}^d} f d\mu.$$

In particular, if μ_x is as above, then

$$(\mu_x \mid f) = \sum_{i=1}^N f(x_i).$$

If φ is a diffeomorphism of Ω and μ a measure, we define a new measure $\varphi \cdot \mu$ by

$$(\varphi \cdot \mu \mid f) = (\mu \mid f \circ \varphi).$$

It is straightforward to check that this provides a group action. If $\mu = \mu_x$, we have

$$(\varphi \cdot \mu_x \mid f) = \sum_{i=1}^N f \circ \varphi(x_i) = (\mu_{\varphi(x)} \mid f),$$

so that the transformation of the measure associated to a point set x is the measure associated to the transformed point set, which is reasonable.

When μ has a density with respect to Lebesgue measure, say $\mu = z dx$, this action can be translated to a resulting transformation over densities as follows.

Proposition 9.6 *If $\mu = z dx$, where z is a positive, Lebesgue integrable function on $\Omega \subset \mathbb{R}^d$, and φ is a diffeomorphism of Ω , then*

$$\varphi \cdot \mu = \det(d(\varphi^{-1})) z \circ \varphi^{-1} dx. \quad (9.23)$$

The proposition is an immediate consequence of the definition of $\varphi \cdot \mu$ and of the change of variable formula (details are left to the reader). Note that the action of diffeomorphisms does not change the total mass of a positive measure, that is $(\varphi \cdot \mu)(\Omega) = \mu(\Omega)$ if φ is a diffeomorphism of Ω .

Now that we have defined the action, we need to choose a function $D(\mu, \mu')$ to compare two measures μ and μ' . Many such functions exist, especially when measures are normalized to have a unit mass, since this allows for the use of many comparison criteria defined in probability or information theory (such as the Kullback–Leibler divergence [75]). A very general example is the Wasserstein distance [238, 301], which is associated to a positive, symmetric, cost function $\rho : \Omega \times \Omega \rightarrow [0, +\infty)$ and defined by

$$d_\rho(\mu, \mu') = \inf_\nu \int_{\Omega^2} \rho(x, y) \nu(dx, dy), \tag{9.24}$$

where the minimization is over all ν with the first marginal given by μ , and the second one by μ' . If μ and μ' are uniform measures on discrete point sets, i.e.,

$$\mu = \frac{1}{N} \sum_{k=1}^N \delta_{x_k}, \quad \mu' = \frac{1}{M} \sum_{k=1}^M \delta_{x'_k},$$

then computing the Wasserstein distance reduces to minimizing

$$\sum_{k=1}^N \sum_{l=1}^M \rho(x_k, x'_l) \nu(x_k, x'_l)$$

subject to the constraints

$$\sum_{l=1}^M \nu(x_k, x'_l) = 1/N \text{ and } \sum_{k=1}^N \nu(x_k, x'_l) = 1/M.$$

This *linear assignment* problem is solved by finite-dimensional linear programming. If this is combined with diffeomorphic interpolation, i.e., if one tries to compute a diffeomorphism φ minimizing $d_\rho(\varphi \cdot x, x')$, this results in a formulation that mixes discrete and continuous optimization problems, similar to the methods introduced in [240]. The Wasserstein distance is also closely related to the mass transport problem, which can also be used to estimate diffeomorphisms, and will be discussed in the next chapter. For the moment, we focus on matching functionals associated with measures, and start with the case in which the compared measures are differentiable with respect to Lebesgue measure, i.e., with the problem of matching densities.

9.6.1 Matching Densities

Since densities are scalar-valued functions, we can use standard norms to design matching functionals for them. As an example, we can take the simplest case of the L^2 norm, as we did with images. The difference with the image case is that the action is different, and has the interesting feature of involving the derivative of the diffeomorphism, via the Jacobian determinant.

So, let us consider the action $\varphi \star \zeta$ given by

$$\varphi \star \zeta = \det(d(\varphi^{-1})) \zeta \circ \varphi^{-1}$$

and use the matching functional

$$U_\zeta(\varphi) = E_{\zeta, \zeta'}(\varphi) = \int_{\Omega} (\varphi \star \zeta - \zeta')^2 dx.$$

Since we will need it for the differentiation of the Jacobian, we recall the following standard result on the derivative of the determinant.

Proposition 9.7 *Let $F(A) = \det(A)$ be defined over $\mathcal{M}_n(\mathbb{R})$, the space of all n by n matrices. Then, for any $A, H \in \mathcal{M}_n(\mathbb{R})$,*

$$dF(A)H = \text{trace}(\text{Adj}(A)H), \quad (9.25)$$

where $\text{Adj}(A)$ is the adjugate matrix of A , i.e., the matrix with (i, j) entry given by the determinant of A with the j th row and i th column removed, multiplied by $(-1)^{i+j}$. (When A is invertible $\text{Adj}(A) = \det(A) A^{-1}$.)

For $A = \text{Id}$, we have

$$dF(\text{Id})H = \text{trace}(H). \quad (9.26)$$

Proof To prove this proposition, start with $A = \text{Id}$ and use the facts that, if δ_{ij} is the matrix with 1 as the (i, j) entry and 0 everywhere else, then $\det(\text{Id} + \varepsilon\delta_{ij}) = 1 + \varepsilon$ if $i = j$ and 1 otherwise, which directly gives (9.26). Then, prove the result for an invertible A using

$$\det(A + \varepsilon H) = \det(A) \det(\text{Id} + \varepsilon A^{-1}H)$$

and the fact that, when A is invertible, $\det(A)A^{-1} = \text{Adj}(A)$. This also implies the result for a general (not necessarily invertible) A because the determinant is a polynomial in the entries of a matrix, and so are its partial derivatives, and the coefficients of these polynomials are fully determined by the values taken on the dense set of invertible matrices. \square

We have

$$U_\zeta(\varphi) = \int_{\Omega} (\det(d(\varphi^{-1})) \zeta \circ \varphi^{-1} - \zeta')^2 dx.$$

Under the assumptions that ζ is C^1 and compactly supported and that ζ' is square integrable, one can prove that $E_{\zeta, \zeta'}$ is C^1 when defined over $\text{Diff}_0^{p+2, \infty}$ with $p \geq 0$ (the details are left to the reader). To compute the derivative at any given φ , it will be convenient to use the trick described at the end of Sect. 9.2, starting with the computation of the differential at the identity and deducing from it the differential at any φ by replacing ζ by $\varphi \cdot \zeta$.

If $\varphi(\varepsilon, \cdot)$ is a diffeomorphism that depends on a parameter ε , such that $\varphi(0, \cdot) = \text{id}$ and $\partial_\varepsilon \varphi(0, \cdot) = h \in C_0^{p+2}(\Omega, \mathbb{R}^d)$, then, at $\varepsilon = 0$, $\partial_\varepsilon \zeta \circ \varphi(\varepsilon, \cdot)^{-1} = -\nabla \zeta^T h$ and $\partial_\varepsilon \det(d(\varphi(\varepsilon, \cdot)^{-1})) = -\text{trace}(dh) = -\text{div } h$. This implies that

$$\partial_\varepsilon (\zeta \circ \varphi(\varepsilon, \cdot)^{-1} \det(d(\varphi(\varepsilon, \cdot)^{-1})) = -\nabla \zeta^T h - \zeta \text{div } h = -\text{div}(\zeta h)$$

at $\varepsilon = 0$ and

$$\partial_\varepsilon U_\zeta(\varphi_\varepsilon)|_{\varepsilon=0} = -2 \int_\Omega (\zeta - \zeta') \text{div}(\zeta h) \, dx.$$

So this gives

$$(dU_\zeta(\text{id}) \mid h) = -2 \int_\Omega (\zeta - \zeta') \text{div}(\zeta h) \, dx$$

and

$$(dU_\zeta(\varphi) \mid h) = -2 \int_\Omega (\varphi \star \zeta - \zeta') (\text{div}((\varphi \star \zeta)h)) \, dx. \tag{9.27}$$

We can use the divergence theorem to obtain an alternative expression (using the fact that h vanishes on $\partial\Omega$ or at infinity), yielding

$$(dU_\zeta(\varphi) \mid h) = 2 \int_\Omega \nabla(\varphi \star \zeta - \zeta') (\varphi \star \zeta) h \, dx \tag{9.28}$$

or

$$dU_\zeta(\varphi) = 2(\varphi \star \zeta) \nabla(\varphi \star \zeta - \zeta') \, dx. \tag{9.29}$$

One can appreciate the symmetry of this expression compared with the one obtained with images in (9.19).

9.6.2 Dual RKHS Norms on Measures

One of the limitations of functional norms, such as the L^2 norm, is that they do not apply to singular objects such as the Dirac measures that motivated our study of the measure-matching problem. It is certainly possible to smooth out singular objects and transform them into densities that can be compared using the previous matching functional. For example, given a density function ρ (a Gaussian, for example) and a point set (x_1, \dots, x_N) , one can compute a density

$$\zeta_x(y) = \sum_{k=1}^N \rho\left(\frac{y - x_k}{\sigma}\right), \tag{9.30}$$

where σ is a positive scale parameter (this is a standard kernel density estimator). One can then compare two point sets, say x and x' , by comparing the associated ζ_x and $\zeta_{x'}$ using the previous method.

The representation in (9.30) is somewhat imperfect, in the sense that, for the natural actions we have defined, we have in general $\varphi \star \zeta_x \neq \zeta_{\varphi \cdot x}$: the density associated to a deformed point set is not the deformed density. If the goal is to compare two point sets, it makes more sense to use $\zeta_{\varphi \cdot x}$ instead of $\varphi \cdot \zeta_x$ as a density resulting from the deformation, and to rather use the cost function

$$U_x(\varphi) = E_{x,x'}(\varphi) = \int_{\mathbb{R}^d} (\zeta_{\varphi \cdot x} - \zeta_{x'})^2 dy, \tag{9.31}$$

which can be written, if $x = (x_1, \dots, x_N)$ and $x' = (x'_1, \dots, x'_M)$, and introducing the function

$$\xi(z, z') = \int_{\mathbb{R}^d} \rho\left(\frac{y - z}{\sigma}\right) \rho\left(\frac{y - z'}{\sigma}\right) dy, \tag{9.32}$$

as

$$\begin{aligned} U_x(\varphi) &= \sum_{k,l=1}^N \xi(\varphi(x_k), \varphi(x_l)) \\ &\quad - 2 \sum_{k=1}^N \sum_{l=1}^M \xi(\varphi(x_k), x'_l) + \sum_{k,l=1}^M \xi(x'_k, x'_l). \end{aligned} \tag{9.33}$$

Before computing the variations of this energy, we make the preliminary remark that the obtained expression is a particular case of what comes from a representation of measures as linear forms over RKHSs of scalar functions. Indeed, since measures are linear forms on functions, we can evaluate their dual norm, given by

$$\|\mu\| = \sup \{(\mu \mid f) : \|f\| = 1\}. \tag{9.34}$$

Following [128], assume that the function norm in (9.34) is that of an RKHS. More precisely, let W be an RKHS of real-valued functions, so that we have an operator $\mathbb{K}_W : W^* \rightarrow W$ with $\mathbb{K}_W \delta_x := \xi(\cdot, x)$ and with the identity $(\mu \mid f) = \langle \mathbb{K}_W \mu, f \rangle_W$ for $\mu \in W^*, f \in W$. With this choice, (9.34) becomes

$$\begin{aligned} \|\mu\|_{W^*} &= \sup \{(\mu \mid f) : \|f\|_W = 1\} \\ &= \sup \{ \langle \mathbb{K}_W \mu, f \rangle_W : \|f\|_W = 1 \} \\ &= \|\mathbb{K}_W \mu\|_W. \end{aligned}$$

This implies that

$$\|\mu\|_{W^*}^2 = \langle \mathbb{K}_W \mu, \mathbb{K}_W \mu \rangle_W = (\mu | \mathbb{K}_W \mu).$$

If μ is a measure, this expression is very simple and is given by

$$\|\mu\|_{W^*}^2 = \int \xi(x, y) d\mu(x) d\mu(y).$$

This is because $\mathbb{K}_W \mu(x) = (\delta_x | \mathbb{K}_W \mu) = (\mu | \mathbb{K}_W \delta_x) = \int \xi(y, x) d\mu(y)$. So we can take

$$U_\mu(\varphi) = E_{\mu, \mu'}(\varphi) = \|\varphi \cdot \mu - \mu'\|_{W^*}^2. \tag{9.35}$$

Expanding the norm, we get

$$\begin{aligned} U_\mu(\varphi) &= \langle \varphi \cdot \mu, \varphi \cdot \mu \rangle_{W^*} - 2\langle \varphi \cdot \mu, \mu' \rangle_{W^*} + \langle \mu', \mu' \rangle_{W^*} \\ &= (\varphi \cdot \mu | \xi(\varphi \cdot \mu)) - 2(\varphi \cdot \mu | \xi \mu') + (\mu' | \xi \mu') \\ &= \int \xi(\varphi(x), \varphi(y)) d\mu(x) d\mu(y) - 2 \int \xi(\varphi(x), y) d\mu(x) d\mu'(y) \\ &\quad + \int \xi(x, y) d\mu'(x) d\mu'(y). \end{aligned}$$

We retrieve (9.33) when μ and μ' are sums of Dirac measures and ξ is chosen as in (9.32), but the RKHS formulation is more general.

Assume that μ is bounded and that ξ is continuously differentiable and bounded, with bounded derivatives. Then (leaving the proof to the reader) U_I is C^1 on $\text{Diff}_0^{p, \infty}$ for any $p \geq 0$ with derivative

$$\begin{aligned} (\partial U_\mu(\varphi) | h) &= 2 \int \nabla_1 \xi(\varphi(x), \varphi(y))^T h(x) d\mu(x) d\mu(y) \\ &\quad - 2 \int \nabla_1 \xi(\varphi(x), z)^T h(x) d\mu(x) d\mu'(z). \end{aligned}$$

In particular,

$$dU_\mu(\text{id}) = \bar{\partial} U_\mu(\text{id}) = 2 \left(\int \nabla_1 \xi(\cdot, y) d\mu(y) - \int \nabla_1 \xi(\cdot, z) d\mu'(z) \right) \mu. \tag{9.36}$$

To obtain the Eulerian differential at a generic φ , it suffices to replace μ by $\varphi \cdot \mu$, which yields:

Proposition 9.8 *The Eulerian derivative and gradient of (9.35) are*

$$\bar{\partial} U_\mu(\varphi) = 2 \left(\int \nabla_1 \xi(\cdot, \varphi(y)) d\mu(y) - \int \nabla_1 \xi(\cdot, z) d\mu'(z) \right) (\varphi \cdot \mu) \tag{9.37}$$

and

$$\bar{\nabla}^V U_\mu(\varphi)(\cdot) = 2 \int K(\cdot, \varphi(x)) \left(\int \nabla_1 \xi(\varphi(x), \varphi(y)) d\mu(y) - \nabla_1 \xi(\varphi(x), z) d\mu'(z) \right) d\mu(x). \quad (9.38)$$

The derivative of the expression in (9.33) can be directly deduced from this expression. This leads to the following unlabeled point-matching evolution for point sets $x = (x_1, \dots, x_N)$ and $x' = (x'_1, \dots, x'_M)$:

$$\partial_t \varphi(z) = -2 \sum_{i=1}^N K(\varphi(z), \varphi(x_i)) \left(\sum_{j=1}^N \nabla_1 \xi(\varphi(x_i), \varphi(x_j)) - \sum_{h=1}^M \nabla_1 \xi(\varphi(x_i), x'_h) \right). \quad (9.39)$$

As discussed in the case of labeled point sets, this equation may be solved in two stages: letting $z_i(t) = \varphi(t, x_i)$, first solve the system

$$\partial_t z_q = -2 \sum_{i=1}^N K(z_q, z_i) \left(\sum_{j=1}^N \nabla_1 \xi(z_i, z_j) - \sum_{h=1}^M \nabla_1 \xi(z_i, x'_h) \right).$$

Once this is done, the trajectory of an arbitrary point $z(t) = \varphi_t(z_0)$ is

$$\partial_t z = -2 \sum_{i=1}^N K(z, z_i) \left(\sum_{j=1}^N \nabla_1 \xi(z_i, z_j) - \sum_{h=1}^M \nabla_1 \xi(z_i, x'_h) \right).$$

9.7 Matching Curves and Surfaces

Curves in two dimensions and surfaces in three dimensions are probably the most natural representations of shapes, and their comparison using matching functionals is a fundamental issue. In this section, we discuss a series of representations that can be seen as extensions of measure-matching methods. (This is not the unique way to compare such objects, and we will see a few more methods in the following chapters, especially for curves.)

Note that we are looking here for correspondences between points in the curves and surfaces that derive from global diffeomorphisms of the ambient space. The curve- (or surface-) matching problems are often studied in the literature as attempts to find diffeomorphic correspondences between points along the curve (or surface)

only. Even if such restricted diffeomorphisms can generally be extended to diffeomorphisms of the whole space, the two approaches generally lead to very different algorithms. The search for correspondences within the structures is often implemented as a search for correspondences between parametrizations. This is easier for curves (looking, for example, for correspondences of the arc-length parametrizations), than for surfaces, which may not be topologically equivalent in the first place (a sphere cannot be matched to a torus); when matching topologically equivalent surfaces, special parametrizations, like conformal maps [72, 269] can be used. In this framework, once parametrizations are fixed, one can look for diffeomorphisms in parameter space that optimally align some well-chosen, preferably intrinsic, representation. In the case of curves, one can choose the representation $s \mapsto \kappa_\gamma(s)$, where κ_γ is the curvature of a curve γ , with the curve rescaled to have length 1 to fix the interval over which this representation is defined. One can then use image-matching functionals to compare them, i.e., find φ (a diffeomorphism of the unit interval) such that $\varphi \cdot \kappa_\gamma \simeq \kappa_{\gamma'}$.

But, as we wrote, the main focus in this chapter is the definition of matching functionals for deformable objects in \mathbb{R}^d , and we now address this problem for curves and surfaces.

9.7.1 Curve Matching with Measures

We can arguably make a parallel between point sets and curves in that labeled point sets correspond to parametrized curves and unlabeled point sets to curves modulo parametrization. In this regard we have a direct generalization of the labeled point-matching functional to parametrized curves (assumed to be defined over the same interval, say $[0, 1]$), simply given by

$$E_{\gamma, \gamma'}(\varphi) = \int_0^1 |\varphi(\gamma(u)) - \gamma'(u)|^2 du.$$

But being given two consistent parametrizations of the curves (to allow for direct comparisons as done above) almost never happens in practice. Interesting formulations of the curve matching problem should therefore consider curves modulo parametrization, so that the natural analogy is with unlabeled point sets. The counterpart of a uniform measure over a finite set of points is the uniform measure on the curve, defined by, if γ , parametrized over an interval $[a, b]$, is C^1 and regular

$$(\mu_\gamma \mid f) = \int_\gamma f d\sigma_\gamma = \int_a^b f(\gamma(u)) |\dot{\gamma}(u)| du.$$

This is clearly a parametrization-independent representation. Now, if φ is a diffeomorphism, we have, by definition of the action of diffeomorphisms on measures

$$(\varphi \cdot \mu_\gamma \mid f) = \int_\gamma f \circ \varphi d\sigma_\gamma = \int_a^b f(\varphi(\gamma(u))) |\dot{\gamma}(u)| du.$$

However, we have

$$(\mu_{\varphi \cdot \gamma} \mid f) = \int_{\varphi(\gamma)} f d\sigma_{\varphi(\gamma)} = \int_a^b f(\varphi(\gamma(u))) |d\varphi(\gamma(u))\dot{\gamma}(u)| du.$$

So, in contrast to point sets, for which we had $\varphi \cdot \mu_x = \mu_{\varphi(x)}$, the image of the measure associated to a curve is not the measure associated to the image of a curve. When the initial goal is to compare curves, and not measures, it is more natural to use the second definition, $\mu_{\varphi \cdot \gamma}$, rather than the first one. Using the notation of the previous section, and introducing a target curve γ' defined on $[a', b']$, we can set

$$\begin{aligned} E_{\gamma, \gamma'}(\varphi) &= \|\mu_{\varphi \cdot \gamma} - \mu_{\gamma'}\|_{W^*}^2 & (9.40) \\ &= \langle \mu_{\varphi \cdot \gamma}, \mu_{\varphi \cdot \gamma} \rangle_{W^*} - 2\langle \mu_{\varphi \cdot \gamma}, \mu_{\gamma'} \rangle_{W^*} + \langle \mu_{\gamma'}, \mu_{\gamma'} \rangle_{W^*} \\ &= (\mu_{\varphi \cdot \gamma} \mid \xi(\mu_{\varphi \cdot \gamma})) - 2(\mu_{\varphi \cdot \gamma} \mid \xi\mu_{\gamma'}) + (\mu_{\gamma'} \mid \xi\mu_{\gamma'}) \\ &= \int_a^b \int_a^b \xi(\varphi(\gamma(u)), \varphi(\gamma(v))) |d\varphi(\gamma(u))\dot{\gamma}(u)| |d\varphi(\gamma(v))\dot{\gamma}(v)| dudv \\ &\quad - 2 \int_a^b \int_{a'}^{b'} \xi(\varphi(\gamma(u)), \gamma'(v)) |d\varphi(\gamma(u))\dot{\gamma}(u)| |\dot{\gamma}'(v)| dudv \\ &\quad + \int_{a'}^{b'} \int_{a'}^{b'} \xi(\gamma'(u), \gamma'(v)) |\dot{\gamma}'(u)| |\dot{\gamma}'(v)| dudv. \end{aligned}$$

If ξ is C^1 , then $E_{\gamma, \gamma'}$ is C^1 on $\text{Diff}_0^{p+1, \infty}$ for any $p \geq 0$. To explicitly compute the derivative, take $\varphi(\varepsilon, \cdot)$ such that $\varphi(0, \cdot) = \text{id}$ and $\partial_\varepsilon \varphi(0, \cdot) = h$, so that

$$\partial_\varepsilon E(\varphi(\varepsilon, \cdot)) = 2\partial_\varepsilon \langle \mu_{\varphi(\varepsilon, \cdot) \cdot \gamma} - \mu_{\gamma'}, \mu_\gamma - \mu_{\gamma'} \rangle_{W^*} = 2\partial_\varepsilon \langle \mu_{\varphi(\varepsilon, \cdot) \cdot \gamma}, \mu_\gamma - \mu_{\gamma'} \rangle_{W^*},$$

the derivatives being computed at $\varepsilon = 0$. Introduce

$$\tilde{E}(\varphi) = \langle \mu_{\varphi \cdot \gamma}, \mu_\gamma - \mu_{\gamma'} \rangle_{W^*}$$

and let, for a given curve $\tilde{\gamma}$,

$$Z^{\tilde{\gamma}}(\cdot) = \int_{\tilde{\gamma}} \xi(\cdot, p) d\sigma_{\tilde{\gamma}}(p). \quad (9.41)$$

Let also $\zeta = Z^\gamma - Z^{\gamma'}$ and, for further use, $\zeta^\varphi = Z^{\varphi \cdot \gamma} - Z^{\gamma'}$. With this notation, we have

$$\tilde{E}(\varphi) = \int_{\varphi(\gamma)} \zeta(p) d\sigma_{\varphi(\gamma)}(p)$$

and we can use Theorem 5.2 to derive, letting p_0 and p_1 be the extremities of γ ,

$$\begin{aligned} \partial_\varepsilon \tilde{E}(\varphi(\varepsilon, \cdot))|_{\varepsilon=0} &= \zeta(p_1)h(p_1)^T T^\gamma(p_1) - \zeta(p_0)h(p_0)^T T^\gamma(p_0) \\ &+ \int_\gamma (\nabla \zeta^T N^\gamma - \zeta \kappa^\gamma) h^T N^\gamma dl. \end{aligned}$$

Replacing γ by $\varphi \cdot \gamma$, this provides the expression of the Eulerian derivative of E at φ , namely

$$\begin{aligned} \frac{1}{2} \bar{\partial} E_{\gamma, \gamma'}(\varphi) &= \zeta^\varphi T^\gamma (\delta_{p_1} - \delta_{p_0}) \\ &+ ((\nabla \zeta^\varphi)^T N^{\varphi \cdot \gamma} - \zeta^\varphi \kappa^{\varphi \cdot \gamma}) N^{\varphi \cdot \gamma} \mu_{\varphi \cdot \gamma}. \end{aligned} \quad (9.42)$$

The Eulerian gradient on V therefore is

$$\begin{aligned} \frac{1}{2} \bar{\nabla} E_{\gamma, \gamma'}(\varphi) &= K(\cdot, p_1) \zeta^\varphi(p_1) T^\gamma(p_1) - K(\cdot, p_0) \zeta^\varphi(p_0) T^\gamma(p_0) \\ &+ \int_{\varphi \cdot \gamma} \left(\nabla \zeta^\varphi(p)^T N^{\varphi \cdot \gamma}(p) - \zeta^\varphi(p) \kappa^{\varphi \cdot \gamma}(p) \right) K(\cdot, p) N^{\varphi \cdot \gamma}(p) d\sigma_{\varphi \cdot \gamma}(p). \end{aligned} \quad (9.43)$$

To write this expression, we have implicitly assumed that γ is C^2 . In fact, we can give an alternative expression for the Eulerian gradient that does not require this assumption, by directly computing the variation of $\tilde{E}(\varphi(\varepsilon, \cdot))$ without applying Theorem 5.2. This yields, using the fact that, if z is a function of a parameter ε , then $\partial_\varepsilon |z| = (\dot{z}^T z) / |z|$,

$$\begin{aligned} \partial_\varepsilon |d\varphi(\varepsilon, \cdot)(\gamma) \dot{\gamma}|_{\varepsilon=0} &= (T^\gamma)^T dh(\gamma) \dot{\gamma} = (T^\gamma)^T dh(\gamma) T^\gamma |\dot{\gamma}| \\ \text{and } \partial_\varepsilon \tilde{E}(\varphi(\varepsilon, \cdot)) &= \int_\gamma (\nabla \zeta^T h + \zeta (T^\gamma)^T dh T^\gamma) d\sigma_\gamma. \end{aligned}$$

The term involving dh can be written in terms of V -dot products of h with derivatives of the kernel, K , since (we use the notation introduced in Sect. 8.1.3, Eq. (8.9))

$$a^T dh(x)b = \langle h, \partial_2 K(\cdot, x)(a, b) \rangle_V. \quad (9.44)$$

This gives

$$\begin{aligned} \partial_\varepsilon \tilde{E}(\varphi(\varepsilon, \cdot)) &= \int_\gamma \left(\langle K(\cdot, p) \nabla \zeta(p), h \rangle_V \right. \\ &\quad \left. + \langle \zeta(p) \partial_2 K(\cdot, p)(T^\gamma(p), T^\gamma(p)), h \rangle_V \right) d\sigma_\gamma(p) \end{aligned}$$

and a new expression of the Eulerian gradient

$$\frac{1}{2} \bar{\nabla}^V E_{\gamma, \gamma'}(\varphi) = \int_{\varphi \cdot \gamma} \left(K(\cdot, p) \nabla \zeta^\varphi(p) + \zeta^\varphi(p) \partial_2 K(\cdot, p)(T^{\varphi \cdot \gamma}(p), T^{\varphi \cdot \gamma}(p)) \right) d\sigma_\gamma(p). \quad (9.45)$$

To be complete, let us consider the variation of a discrete form of $E_{\gamma, \gamma'}(\varphi)$. If a curve γ is discretized with points x_0, \dots, x_N (with $x_N = x_0$ if the curve is closed), one can define the discrete measure, still denoted μ_γ

$$(\mu_\gamma \mid f) = \sum_{i=1}^N f(c_i) |\tau_i|$$

with $c_i = (x_i + x_{i-1})/2$ and $\tau_i = x_i - x_{i-1}$. Use a similar expression for the measure associated to a discretization of $\varphi \cdot \gamma$, with $c_i^\varphi = (\varphi(x_i) + \varphi(x_{i-1}))/2$ and $\tau_i^\varphi = \varphi(x_i) - \varphi(x_{i-1})$. Finally, let γ' be discretized in x'_1, \dots, x'_M , and define

$$\begin{aligned} E_{\gamma, \gamma'}(\varphi) &= \sum_{i,j=1}^N \xi(c_i^\varphi, c_j^\varphi) |\tau_i^\varphi| |\tau_j^\varphi| \\ &\quad - 2 \sum_{i=1}^N \sum_{j=1}^N \xi(c_i^\varphi, c'_j) |\tau_i^\varphi| |\tau'_j| + \sum_{i,j=1}^M \xi(c'_i, c'_j) |\tau'_i| |\tau'_j| \end{aligned} \quad (9.46)$$

in which we identify indices 1 and $N + 1$ or $M + 1$ (assuming closed curves). Note that this functional depends on $\varphi \cdot x$ and x' . The computation of the differential proceeds as above. Define, for a point set $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_Q)$

$$Z^{\tilde{x}}(\cdot) = \sum_{j=1}^Q \xi(\cdot, \tilde{c}_j) |\tilde{\tau}_j|,$$

and $\zeta = Z^x - Z^{x'}$, $\zeta^\varphi = Z^{\varphi \cdot x} - Z^{x'}$. We then obtain

$$\begin{aligned} \frac{1}{2} dE_{\gamma, \gamma'}(\text{id}) &= \sum_{i=1}^N (\nabla \zeta(c_i) |\tau_i| + \nabla \zeta(c_{i+1}) |\tau_{i+1}|) \delta_{x_i} \\ &\quad - 2 \sum_{i=1}^N \left(\zeta(c_{i+1}) \frac{\tau_{i+1}}{|\tau_{i+1}|} - \zeta(c_i) \frac{\tau_i}{|\tau_i|} \right) \delta_{x_i}. \end{aligned}$$

The Eulerian differential at $\varphi \neq \text{id}$ is obtained by replacing ζ, c_i, τ_i by $\zeta^\varphi, c_i^\varphi, \tau_i^\varphi$ and the Eulerian gradient by applying the V -kernel to it.

9.7.2 Curve Matching with Vector Measures

Instead of describing a curve with a measure, which is a linear form on functions, it is possible to represent it by a vector measure, which is a linear form on vector fields. Given a parametrized curve $\gamma : [a, b] \rightarrow \mathbb{R}^d$, we define a vector measure ν_γ , which associates to each vector field f on \mathbb{R}^d a number $(\nu_\gamma | f)$ given by

$$(\nu_\gamma | f) = \int_a^b \dot{\gamma}(u)^T f \circ \gamma(u) du,$$

i.e., $\nu_\gamma = T^\gamma \mu_\gamma$ where T^γ is the unit tangent to γ and μ_γ is the line measure along γ , as defined in the previous section. This definition is invariant under a change of parametrization, but depends on the orientation of γ . If φ is a diffeomorphism, we then have

$$\nu_{\varphi \circ \gamma}(f) = \int_a^b (d\varphi(\gamma(u))\dot{\gamma}(u))^T f \circ \varphi(\gamma(u)) du.$$

As done with scalar measures, we can use a dual norm for the comparison of two vector measures. Such a norm is defined by

$$\|\nu\|_{W^*} = \sup \{(\nu | f) : \|f\|_W = 1\},$$

where W is now an RKHS of vector fields, and we still have $\|\nu\|_{W^*}^2 = (\nu | \mathbb{K}_W \nu)$. Still letting ξ denote the kernel of W (which is now matrix-valued), we have

$$\|\nu_\gamma\|_{W^*}^2 = \int_a^b \int_a^b \dot{\gamma}(u)^T \xi(\gamma(u), \gamma(v)) \dot{\gamma}(v) du dv$$

and

$$\|\nu_{\varphi \circ \gamma}\|_{W^*}^2 = \int_a^b \int_a^b \dot{\gamma}(u)^T d\varphi(\gamma(u))^T \xi(\varphi(\gamma(u)), \varphi(\gamma(v))) d\varphi(\gamma(v)) \dot{\gamma}(v) du dv.$$

Define $E_{\gamma, \gamma'}(\varphi) = \|\nu_{\varphi \circ \gamma} - \nu_{\gamma'}\|_{W^*}^2$. We follow the same pattern as in the previous section and define

$$\tilde{E}(\varphi) = \langle \nu_{\varphi \circ \gamma}, \nu_\gamma - \nu_{\gamma'} \rangle_{W^*},$$

which (introducing $\varphi(\varepsilon, \cdot)$ with $\varphi(0, \cdot) = \text{id}$ and $\partial_\varepsilon \varphi(0, \cdot) = h$) is such that $\partial_\varepsilon E(\varphi(\varepsilon, \cdot)) = 2\partial_\varepsilon \tilde{E}(\varphi(\varepsilon, \cdot))$ at $\varepsilon = 0$. Define

$$Z^{\tilde{\gamma}}(\cdot) = \int_{\tilde{\gamma}} \xi(\cdot, p) N^{\tilde{\gamma}}(p) dp,$$

and $\zeta = Z^\gamma - Z^{\gamma'}$, $\zeta^\varphi = Z^{\varphi \circ \gamma} - Z^{\gamma'}$, so that (using $(T^{\varphi \circ \gamma})^T T^{\gamma'} = (N^{\varphi \circ \gamma})^T N^{\gamma'}$)

$$\tilde{E}(\varphi) = \int_{\varphi \cdot \gamma} \zeta^T N^{\varphi \cdot \gamma} d\sigma_{\varphi \cdot \gamma}.$$

We can use Theorem 5.2, Eq. (5.4), to find

$$\partial_\varepsilon E(\varphi(\varepsilon, \cdot)) = -[\det(\zeta, h)]_0^A + \int_\gamma \operatorname{div}(\zeta)(N^\gamma)^T h dl.$$

This yields in turn (replacing γ by $\varphi \cdot \gamma$, and letting p_0 and p_1 be the extremities of γ)

$$\frac{1}{2} \bar{\partial} E_{\gamma, \gamma'}(\varphi) = -(R_{\pi/2} \zeta^\varphi)(\delta_{\varphi(p_1)} - \delta_{\varphi(p_0)}) + \operatorname{div}(\zeta^\varphi) \nu_{\varphi \cdot \gamma}, \tag{9.47}$$

where $R_{\pi/2}$ is a 90° rotation. This final expression is remarkably simple, especially for closed curves, for which the first term cancels. A discrete version of the matching functional can also be defined, namely, using the notation of the previous section:

$$\begin{aligned} E_{\gamma, \gamma'}(\varphi) &= \sum_{i, j=1}^N \xi(c_i^\varphi, c_j^\varphi)(\tau_i^\varphi)^T \tau_j^\varphi \\ &\quad - 2 \sum_{i=1}^N \sum_{j=1}^N \xi(c_i^\varphi, c_j^\varphi)(\tau_i^\varphi)^T \tau_j^\varphi + \sum_{i, j=1}^M \xi(c'_i, c'_j)(\tau'_i)^T \tau'_j. \end{aligned}$$

We leave the computation of the associated Eulerian differential (which is a slight variation of the one we made with measures) to the reader.

9.7.3 Surface Matching

We now extend to surfaces the matching functionals that we just studied for curves. The construction is formally very similar. If S is a surface in \mathbb{R}^3 , one can compute a measure μ_S and a vector measure ν_S defined by

$$(\mu_S \mid f) = \int_S f(x) d\sigma_S(x) \text{ for a scalar } f \tag{9.48}$$

and

$$(\nu_S \mid f) = \int_S f(x)^T N(x) d\sigma_S(x) \text{ for a vector field } f, \tag{9.49}$$

where $d\sigma_S$ is the volume measure on S and N is the unit normal (S being assumed to be oriented in the definition of ν_S).

We state without proof the following result:

Proposition 9.9 *If S is a surface and φ a diffeomorphism of \mathbb{R}^3 that preserves the orientation (i.e., with positive Jacobian), we have*

$$(\mu_{\varphi(S)} \mid f) = \int_S f \circ \varphi(x) |d\varphi(x)^{-T} N| \det(d\varphi(x)) d\sigma_S(x)$$

for a scalar f and for a vector-valued f ,

$$(\nu_{\varphi(S)} \mid f) = \int_S f \circ \varphi(x)^T d\varphi(x)^{-T} N \det(d\varphi(x)) d\sigma_S(x).$$

If $e_1(x), e_2(x)$ is a basis of the tangent plane to S at x , we have

$$d\varphi(x)^{-T} N \det(d\varphi(x)) = (d\varphi(x)e_1 \times d\varphi(x)e_2) / |e_1 \times e_2|. \tag{9.50}$$

The last formula implies in particular that if S is parametrized by $(u, v) \mapsto m(u, v)$, then (since $N = (\partial_1 m \times \partial_2 m) / |\partial_1 m \times \partial_2 m|$ and $d\sigma_S = |\partial_1 m \times \partial_2 m| dudv$)

$$\begin{aligned} (\nu_S \mid f) &= \int f(x)^T (\partial_1 m \times \partial_2 m) dudv \\ &= \int \det(\partial_1 m, \partial_2 m, f) dudv \end{aligned}$$

and

$$(\nu_{\varphi(S)} \mid f) = \int \det(d\varphi \partial_1 m, d\varphi \partial_2 m, f \circ \varphi) dudv.$$

If W is an RKHS of scalar functions or vector fields, we can compare two surfaces by using the norm of the difference of their associated measures on W^* . So define (in the scalar measure case)

$$E_{S,S'}(\varphi) = \|\mu_{\varphi \cdot S} - \mu_{S'}\|_{W^*}^2 \tag{9.51}$$

and the associated

$$\tilde{E}(\varphi) = \langle \mu_{\varphi \cdot S}, \mu_S - \mu_{S'} \rangle_{W^*}$$

so that, for $\varphi(\varepsilon, \cdot)$ such that $\varphi(0, \cdot) = \text{id}$ and $\partial_\varepsilon \varphi(0, \cdot) = h$

$$\partial_\varepsilon E_{\gamma, \gamma'}(\varphi(\varepsilon, \cdot)) = 2\partial_\varepsilon \tilde{E}(\varphi(\varepsilon, \cdot)).$$

To a given surface \tilde{S} , associate the function

$$Z^{\tilde{S}}(\cdot) = \int_{\tilde{S}} \xi(\cdot, p) d\sigma_{\tilde{S}}(p)$$

and $\zeta = Z^S - Z^{S'}$, $\zeta^\varphi = Z^{\varphi \cdot S} - Z^{S'}$. Since

$$\tilde{E}(\varphi) = \int_{\varphi \cdot S} \zeta(p) d\sigma_{\varphi \cdot S}(p),$$

Theorem 5.4 yields

$$\begin{aligned} \partial_\varepsilon \tilde{E}(\varphi(\varepsilon, \cdot)) &= - \int_{\partial S} \zeta(n^S)^T h d\sigma_{\partial S} \\ &\quad + \int_S (-2\zeta H^S + \nabla \zeta^T N^S)(N^S)^T h d\sigma_S \end{aligned}$$

where H^S is the mean curvature on S . This implies

$$\frac{1}{2} \bar{\partial} E_{S,S'}(\varphi) = -\zeta^\varphi n^{\varphi \cdot S} \mu_{\varphi \cdot \partial S} + (-2\zeta^\varphi H^{\varphi \cdot S} + (\nabla \zeta^\varphi)^T N^{\varphi \cdot S}) \nu_{\varphi \cdot S}. \quad (9.52)$$

If we now use vector measures, so that

$$E_{S,S'}(\varphi) = \|\nu_{\varphi \cdot S} - \nu_{S'}\|_{W^*}^2 \quad (9.53)$$

and

$$\tilde{E}(\varphi) = \langle \nu_{\varphi \cdot S}, \nu_S - \nu_{S'} \rangle_{W^*},$$

we need to define

$$Z^{\tilde{S}}(\cdot) = \int_{\tilde{S}} \xi(\cdot, p) N^{\tilde{S}} d\sigma_{\tilde{S}}(p)$$

and $\zeta = Z^S - Z^{S'}$, $\zeta^\varphi = Z^{\varphi \cdot S} - Z^{S'}$, so that

$$\tilde{E}(\varphi) = \int_{\varphi \cdot S} \zeta^T N^{\varphi \cdot S} d\sigma_{\varphi \cdot S}.$$

Variations derive again from Theorem 5.4, yielding

$$\begin{aligned} \partial_\varepsilon \tilde{E} &= - \int_{\partial S} ((\zeta^T N^S)(h^T n^S) - (\zeta^T n^S)(h^T N^S)) d\sigma_{\partial S} \\ &\quad + \int_S \operatorname{div}(\zeta)(N^S)^T h d\sigma_S. \end{aligned}$$

We therefore have

$$\frac{1}{2} \bar{\partial} E_{S,S'}(\varphi) = -((\zeta^\varphi)^T N^{\varphi \cdot S} n^{\varphi \cdot S} - (\zeta^\varphi)^T n^{\varphi \cdot S} N^{\varphi \cdot S}) \mu_{\varphi \cdot \partial S} + \operatorname{div}(\zeta^\varphi) \nu_{\varphi \cdot S}. \quad (9.54)$$

Again the expression is remarkably simple for surfaces without boundary.

Consider now the discrete case and let S be a triangulated surface [289]. Let x_1, \dots, x_N be the vertices of S and f_1, \dots, f_Q be the faces (triangles) which are ordered triples of vertices $f_i = (x_{i1}, x_{i2}, x_{i3})$. Let c_i be the center of f_i , N_i its oriented unit normal and a_i its area. Define the discrete versions of the previous measures by

$$(\mu_S | h) = \sum_{i=1}^Q h(c_i) a_i, \quad \text{for a scalar } h \quad (9.55)$$

and

$$(\nu_S | h) = \sum_{i=1}^Q (h(c_i)^T N_i) a_i, \quad \text{for a vector field } h. \quad (9.56)$$

The previous formulae can be written as

$$(\mu_S | h) = \sum_{i=1}^K h\left(\frac{x_{i1} + x_{i2} + x_{i3}}{3}\right) |(x_{i2} - x_{i1}) \times (x_{i3} - x_{i1})|$$

and

$$(\nu_S | h) = \sum_{i=1}^K h\left(\frac{x_{i1} + x_{i2} + x_{i3}}{3}\right)^T (x_{i2} - x_{i1}) \times (x_{i3} - x_{i1}),$$

where the last formula requires that the vertices or the triangles are ordered consistently with the orientation (see Sect. 4.2). The transformed surfaces are now represented by the same expressions with x_{ik} replaced by $\varphi(x_{ik})$. If, given two triangulated surfaces, one defines $E_{S,S'}(\varphi) = \|\mu_{\varphi \cdot S} - \mu_{S'}\|_{W^*}^2$, then (leaving the computation to the reader)

$$\frac{1}{2} \bar{\partial} E_{S,S'}(\operatorname{id}) = \sum_{k=1}^N \left(\sum_{i: x_k \in f_i} (\nabla \zeta(c_i) \frac{a_i}{3} - \zeta(c_i) e_{ik} \times N_i) \right) \delta_{x_k},$$

where e_{ik} is the edge opposite x_k in f_i (oriented so that (x_k, e_{ik}) is positively ordered), and $\zeta = Z^S - Z^{S'}$, with

$$Z^{\bar{S}}(\cdot) = \sum_{i=1}^{\bar{K}} \xi(\cdot, \tilde{c}_i) \tilde{a}_i$$

for a triangulated surface \tilde{S} . The Eulerian differential at φ is obtained by replacing all x_k 's by $\varphi(x_k)$.

For the vector-measure form, $E_{S,S'}(\varphi) = \|\nu_{\varphi \cdot S} - \nu_{S'}\|_{W^*}^2$, we get

$$\frac{1}{2} \bar{\partial} E_{S,S'}(\text{id}) = \sum_{k=1}^N \left(\sum_{i: x_k \in f_i} (d\zeta(c_i) N_i) \frac{a_i}{3} - e_{ik} \times \zeta(c_i) \right) \delta_{x_k}$$

still with $\zeta = Z^S - Z^{S'}$, but with

$$Z^{\tilde{S}}(\cdot) = \sum_{i=1}^{\tilde{K}} \xi(\cdot, \tilde{c}_i) N_i a_i.$$

9.7.4 Induced Actions and Currents

We have designed the action of diffeomorphisms on measures by $(\varphi \cdot \mu | h) = (\mu | h \circ \varphi)$. Recall that we have the usual action of diffeomorphisms on functions defined by $\varphi \cdot h = h \circ \varphi^{-1}$, so that we can write $(\varphi \cdot \mu | h) = (\mu | \varphi^{-1} \cdot h)$. In the case of curves, we have seen that this action on the induced measure did not correspond to the image of the curve by a diffeomorphism, in the sense that $\mu_{\varphi \cdot \gamma} \neq \varphi \cdot \mu_\gamma$. Here, we discuss whether the transformations $\mu_\gamma \rightarrow \mu_{\varphi \cdot \gamma}$ or $\nu_\gamma \rightarrow \nu_{\varphi \cdot \gamma}$ (and the equivalent transformations for surfaces) can be described by a similar operation, e.g., whether one can write $(\varphi \cdot \mu | h) = (\mu | \varphi^{-1} \star h)$ where \star would represent another action of diffeomorphisms on functions (or on vector fields for vector measures).

For μ_γ , the answer is negative. We have, letting $T(\gamma(u))$ be the unit tangent to γ ,

$$\begin{aligned} (\mu_{\varphi \cdot \gamma} | h) &= \int_a^b h(\varphi(\gamma(u))) |d\varphi(\gamma(u)) \dot{\gamma}(u)| du \\ &= \int_a^b h(\varphi(\gamma(u))) |d\varphi(\gamma(u)) T(u)| |\dot{\gamma}(u)| du, \end{aligned}$$

so that $(\mu_{\varphi \cdot \gamma} | h) = (\mu_\gamma | h \circ \varphi | d\varphi T |)$, with some abuse of notation in the last formula, since T is only defined along γ . The important fact here is that the function h is transformed according to a rule which depends not only on the diffeomorphism φ , but also on the curve γ , and therefore the result cannot be put in the form $\varphi^{-1} \star h$.

The situation is different for vector measures. Indeed, we have

$$\begin{aligned} \nu_{\varphi \cdot \gamma}(h) &= \int_a^b (d\varphi(\gamma(u)) \dot{\gamma}(u))^T h \circ \varphi(\gamma(u)) du \\ &= (\nu_\gamma | d\varphi^T h \circ \varphi). \end{aligned}$$

So, if we define $\varphi \star h = d(\varphi^{-1})^T h \circ \varphi^{-1}$, we have $(\nu_{\varphi \cdot \gamma} \mid h) = (\nu_\gamma \mid \varphi^{-1} \star h)$. The transformation $(\varphi, h) \mapsto \varphi \star h$ is a valid action of diffeomorphisms on vector fields, since $\text{id} \star h = h$ and $\varphi \star (\psi \star h) = (\varphi \circ \psi) \star h$, as can easily be checked.

The same analysis can be made for surfaces; scalar measures do not transform in accordance to an action, but vector measures do. Let us check this last point by considering the formula in a local chart, where

$$\begin{aligned} (\nu_{\varphi(S)} \mid h) &= \int \det(d\varphi \partial_1 m, d\varphi \partial_2 m, h \circ \varphi) dudv \\ &= \int \det(d\varphi) \det(\partial_1 m, \partial_2 m, (d\varphi)^{-1} h \circ \varphi) dudv \\ &= (\nu_S \mid \det(d\varphi)(d\varphi)^{-1} h \circ \varphi). \end{aligned}$$

So, we need here to define

$$\varphi \star h = \det(d(\varphi^{-1}))(d\varphi^{-1})^{-1} h \circ \varphi^{-1} = (d\varphi h / \det(d\varphi)) \circ \varphi^{-1}.$$

Here again, a direct computation shows that this is an action.

We have just proved that vector measures are transformed by a diffeomorphism φ according to a rule $(\varphi \cdot \mu \mid h) = (\mu \mid \varphi^{-1} \star h)$, the action \star being apparently different for curves and surfaces. In fact, all these actions (including the scalar one) can be placed within a single framework if one replaces vector fields by differential forms and measures by currents [126, 127, 289].

The reader may refer to Sects. B.7.1 and B.7.2 for basic definitions of linear and differential forms, in which the space of differential k -forms on \mathbb{R}^d is denoted Ω_k , or Ω_k^d . We can consider spaces of smooth differential k -forms, and in particular, reproducing kernel Hilbert spaces of such forms: a space $W \subset \Omega_k$ is an RKHS if, for every $x \in \mathbb{R}^d$ and $e_1, \dots, e_k \in \mathbb{R}^d$, the evaluation function

$$(e_1, \dots, e_k)\delta_x : q \mapsto (q(x) \mid e_1, \dots, e_k)$$

belongs to W^* . Introduce the duality operator, so that $\mathbb{K}_W((e_1, \dots, e_k)\delta_x) \in W$. Introduce, for $x, y \in \mathbb{R}^d$, the $2k$ -linear form $\xi(x, y)$ defined by

$$(\xi(x, y) \mid e_1, \dots, e_k; f_1, \dots, f_k) = (\mathbb{K}_W((e_1, \dots, e_k)\delta_x)(y) \mid f_1, \dots, f_k).$$

Notice that this form is skew-symmetric with respect to its first k and its last k variables and that

$$\langle \xi_x(e_1, \dots, e_k), \xi_y(f_1, \dots, f_k) \rangle_W = (\xi(x, y) \mid e_1, \dots, e_k; f_1, \dots, f_k),$$

so that ξ may be called the reproducing kernel of W . Similar to vector fields, kernels for differential k -forms can be derived from scalar kernels by letting

$$\begin{aligned}
 (\xi(x, y) \mid e_1, \dots, e_k; f_1, \dots, f_k) = \\
 \xi(x, y) \langle e_1 \times \dots \times e_k, f_1 \times \dots \times f_k \rangle_{\Lambda_{d-k}},
 \end{aligned}
 \tag{9.57}$$

where the dot product on the space of k -linear forms, Λ_k , is the product of coefficients of the forms over a basis formed by all cross products of subsets of k elements of an orthonormal basis of \mathbb{R}^d , as described in Sect. B.7.1.

Elements of the dual space, W^* , to W are therefore linear forms over differential k -forms, and are special instances of k -currents [107, 210] (k -currents are bounded differential forms over C^∞ differential k -forms with compact support, which is less restrictive than being bounded on W). Important examples of currents are those associated to submanifolds of \mathbb{R}^d , and are defined as follows. Let M be an oriented k -dimensional submanifold of \mathbb{R}^d . To a differential k -form q , associate the quantity

$$(\eta_M \mid q) = \int_M (q(x) \mid e_1(x), \dots, e_k(x)), d\sigma_M(x)$$

where e_1, \dots, e_k is, for all x , a positively oriented orthonormal basis of the tangent space to M at x (by Eq. (B.16), the result does not depend on the chosen basis).

If W is an RKHS of differential k -forms, η_M belongs to W^* and we can compute the dual norm of η_M , which is

$$\|\eta_M\|_{W^*}^2 = \int_M \int_M (\xi(x, y) \mid e_1(x), \dots, e_k(x); e_1(y), \dots, e_k(y)) d\sigma_M(x) d\sigma_M(y)$$

or, for a scalar kernel defined by (9.57),

$$\begin{aligned}
 \|\eta_M\|_{W^*}^2 = \int_M \int_M \xi(x, y) \langle e_1(x) \times \dots \times e_k(x), e_1(y) \times \dots \times e_k(y) \rangle_{\Lambda_{d-k}} \\
 d\sigma_M(x) d\sigma_M(y).
 \end{aligned}$$

The expressions of η_M and its norm in a local chart of M are quite simple. Indeed, if (u_1, \dots, u_k) is the parametrization in the chart and $(\partial_1 m, \dots, \partial_k m)$ the associated tangent vectors (assumed to be positively oriented), we have, for a k -form q (using (B.16))

$$(q \mid \partial_1 m, \dots, \partial_k m) = (q \mid e_1, \dots, e_k) \det(\partial_1 m, \dots, \partial_k m)$$

which immediately yields

$$(q \mid \partial_1 m, \dots, \partial_k m) du_1 \dots du_k = (q \mid e_1, \dots, e_k) d\sigma_M.$$

We therefore have, in the chart,

$$(\eta_M \mid q) = \int (q \mid \partial_1 m, \dots, \partial_k m) du_1 \dots du_k$$

and similar formulas for the norm.

Now consider the action of diffeomorphisms. If M becomes $\varphi(M)$, the formula in the chart yields

$$(\eta_{\varphi(M)} | q) = \int (q \circ \varphi | d\varphi \partial_1 m, \dots, d\varphi \partial_k m) du_1 \dots du_k$$

so that $(\eta_{\varphi(M)} | q) = (\eta_M | \tilde{q})$ with

$$(\tilde{q}(x) | f_1, \dots, f_k) = (q(\varphi(x)) | d\varphi f_1, \dots, d\varphi f_k).$$

As we did with vector measures, we can introduce the left action on k -forms (also called the push-forward of the k -form):

$$(\varphi \star q | f_1, \dots, f_k) = (q \circ \varphi^{-1} | d(\varphi^{-1})f_1, \dots, d(\varphi^{-1})f_k)$$

and the resulting action on p -currents

$$(\varphi \cdot \eta | q) = (\eta | \varphi^{-1} \star q), \quad (9.58)$$

so that we can write $\eta_{\varphi(M)} = \varphi \cdot \eta_M$.

This is reminiscent of what we have obtained for measures, and for vector measures with curves and surfaces. We now check that these examples are particular cases of the previous discussion.

Measures are linear forms on functions, which are also differential 0-forms. The definition $(\varphi \cdot \mu | f) = (\mu | f \circ \varphi)$ is exactly the same as in (9.58).

Consider now the case of curves, which are 1D submanifolds, so that $k = 1$. If γ is a curve, and T is its unit tangent, we have

$$(\eta_\gamma | q) = \int_\gamma (q(\gamma) | T) d\sigma_\gamma = \int_a^b (q(\gamma(u)) | \dot{\gamma}(u)) du.$$

To a vector field h on \mathbb{R}^d , we can associate the differential 1-form q_h defined by $(q_h(x) | v) = h(x)^T v$. In fact all differential 1-forms can be expressed as q_h for some vector field h . Using this identification and noting that $(\nu_\gamma | h) = (\eta_\gamma | q_h)$, we can see that the vector measure for curve matching is a special case of the currents that we have considered here.

For surfaces in three dimensions, we need to take $k = 2$, and if S is a surface, we have

$$(\eta_S | q) = \int_S (q(x) | e_1(x), e_2(x)) d\sigma_S(x).$$

Again, a vector field f on \mathbb{R}^3 induces a 2-form q_f , defined by $(q_f | v_1, v_2) = \det(f, v_1, v_2) = f^T(v_1 \times v_2)$, and every 2-form can be obtained this way. Using the

fact that, if (e_1, e_2) is a positively oriented basis of the tangent space to the surface, then $e_1 \times e_2 = N$, we retrieve $(\nu_S | f) = (\eta_S | q_f)$.

9.7.5 Varifolds

A differential k -form ω on \mathbb{R}^d uniquely defines a function on the product space $\mathbb{R}^d \times \widetilde{\text{Gr}}(d, k)$, the product space of \mathbb{R}^d with the set of all oriented k -dimensional subspaces of \mathbb{R}^d (called the oriented Grassmannian, on which a manifold structure similar to the one discussed in Sect. B.6.7 for the Grassmann manifold can be defined). One can indeed assign to any pair (x, α) in that set the scalar $F_q(x, \alpha) = (q(x) | e_1, \dots, e_k)$ where e_1, \dots, e_k is any positively oriented orthonormal basis of α , and the value does not depend on the chosen basis. Given an oriented k -dimensional submanifold of \mathbb{R}^d , one can define the linear form on continuous functions F defined on $\mathbb{R}^d \times \widetilde{\text{Gr}}(d, k)$, given by

$$(\tilde{\rho}_M | F) = \int_M F(p, T_p M) d\sigma_M,$$

where $T_p M$ is considered with its orientation. The current η_M defined in the previous section is such that $(\eta_M | q) = (\tilde{\rho}_M | F_q)$.

When one wants to disregard orientation, which may be convenient, and sometimes necessary in practice, it is natural to replace $\widetilde{\text{Gr}}(d, k)$ by $\text{Gr}(d, k)$ (the Grassmannian) and define the same linear form (that we now call ρ_M) on functions defined on $\mathbb{R}^d \times \text{Gr}(d, k)$. The linear form ρ_M is a special case of a *varifold*, where varifolds are defined as (Radon) measures on $\mathbb{R}^d \times \text{Gr}(d, k)$.

From this point, and following [61], one can make a construction analogous to the one just described for measures on \mathbb{R}^d . Given a reproducing kernel Hilbert space W of functions defined on $\mathbb{R}^d \times \text{Gr}(d, k)$, define the square distance between two k -dimensional submanifolds of \mathbb{R}^d by

$$D(M, M') = \|\rho_M - \rho_{M'}\|_{W^*}^2.$$

For the approach to be practical, one needs to have explicit kernels on $\mathbb{R}^d \times \text{Gr}(d, k)$. Referring to [61] for a complete discussion, we note here that a class of such kernels can be designed based on the following observations.

- (i) The function defined for $\alpha, \beta \in \widetilde{\text{Gr}}(d, k)$ by

$$\tilde{\xi}(\alpha, \beta) = \langle e_1(x) \times \dots \times e_k(x), f_1(y) \times \dots \times f_k(y) \rangle_{\Lambda_{d-k}},$$

where (e_1, \dots, e_k) and (f_1, \dots, f_k) are positively oriented orthonormal bases of $\tilde{\alpha}$ and $\tilde{\beta}$, is a positive definite kernel. Hence, the function defined for $\alpha, \beta \in \text{Gr}(d, k)$

$$\xi(\alpha, \beta) = \left\langle e_1(x) \times \cdots \times e_k(x), f_1(y) \times \cdots \times f_k(y) \right\rangle_{\Lambda_{d-k}}^2,$$

where (e_1, \dots, e_k) and (f_1, \dots, f_k) are orthonormal bases of α and β , is also definite positive. More generally, if $\tilde{\xi}$ is positive definite on $\widetilde{\text{Gr}}(d, k)$, then $\xi = f(\tilde{\xi})$ is definite positive on $\text{Gr}(d, k)$ for any even analytic function f whose derivatives at 0 are all non-negative, and at least one of them positive. These statements simply use the fact that products of positive kernels remain positive.

(ii) If η is a positive kernel on differential p -forms, then $\tilde{\xi}$ defined by

$$\tilde{\xi}(x, \alpha; y, \beta) = (\eta(x, y) \mid e_1, \dots, e_k; f_1, \dots, f_k)$$

is a positive kernel on $\mathbb{R}^d \times \widetilde{\text{Gr}}(d, k)$.

(iii) If $\xi^{(1)}$ is a reproducing kernel on \mathbb{R}^d and $\xi^{(2)}$ a reproducing kernel on $\text{Gr}(d, k)$, then ξ defined by

$$\xi(x, \alpha; y, \beta) = \xi^{(1)}(x, y)\xi^{(2)}(\alpha, \beta)$$

is a reproducing kernel on $\mathbb{R}^d \times \text{Gr}(d, k)$.

Applying this to surfaces, for example, and using the discussion at the end of the previous section, we find that taking

$$\langle \rho_S, \rho_{\tilde{S}} \rangle_{W^*} = \int_S \int_{\tilde{S}} \xi(x, \tilde{x}) \left(1 + a \left(N(x)^T \tilde{N}(\tilde{x}) \right)^2 \right) d\sigma_{\tilde{S}}(\tilde{x}) d\sigma_S(x),$$

where ξ is a reproducing kernel on \mathbb{R}^d , provides an RKHS dual inner-product on varifolds. The discretization of such a norm is similar to those detailed for scalar and vector measures and is left to the reader.

9.8 Matching Vector Fields

We now study vector fields as deformable objects. They correspond, for example, to velocity fields (that can be observed for weather data), or to gradient fields that can be computed for images. Orientation fields (that can be represented by unit vector fields) are also interesting. They can correspond, for example, to fiber orientations in tissues observed in medical images.

We want to compare two vector fields f and f' , i.e., two functions from \mathbb{R}^d to \mathbb{R}^d . To simplify, we restrict ourselves to $E_{f, f'}(\varphi)$ being the L^2 norm between $\varphi \cdot f$ and $\varphi \cdot f'$, and focus our discussion on the definition of the action of diffeomorphisms on vector fields.

The simplest choice is to use the same action as in image matching and take $\varphi \cdot f = f \circ \varphi^{-1}$, where f is a vector field on \mathbb{R}^d . It is, however, natural (and more consistent with applications) to combine the displacement of the points at which f

is evaluated with a reorientation of f , also induced by the transformation. Several choices can be made for such an action and all may be of interest depending on the context.

For example, we can interpret a vector field as a velocity field, assuming that each point in Ω moves on to a trajectory $x(t)$ and that $f(x) = \dot{x}(t)$, say at time $t = 0$. If we make the transformation $x \mapsto x' = \varphi(x)$, and let f' be the transformed vector field, such that $\dot{x}'(0) = f'(x')$, we get: $\dot{x}'(0) = d\varphi(x)\dot{x}(0) = f' \circ \varphi(x)$ so that $f' = (d\varphi f) \circ \varphi^{-1}$. The transformation $f \mapsto (d\varphi f) \circ \varphi^{-1}$ is an important Lie group operation, called the adjoint representation ($\text{Ad}_\varphi f$). This is anecdotal here, but we will use it again later as a fundamental tool. So, our first action is

$$\varphi * f = (d\varphi f) \circ \varphi^{-1}.$$

To define a second action, we now consider vector fields that are obtained as gradients of a function $I: f = \nabla I$. If I becomes $\varphi \cdot I = I \circ \varphi^{-1}$, then f becomes $d(\varphi^{-1})^T \nabla I \circ \varphi^{-1}$. This defines a new action

$$\varphi \star f = d(\varphi^{-1})^T f \circ \varphi^{-1} = (d\varphi^{-T} f) \circ \varphi^{-1}.$$

This action can be applied to any vector field, not only gradients, but one can check that the set of vector fields f such that $\text{curl } f = 0$ is left invariant by this action.

Sometimes, it is important that the norms of the vector fields at each point remain invariant under the transformation, when dealing, for example, with orientation fields. This can be achieved in both cases by normalizing the result, and we define the following normalized actions:

$$\begin{aligned} \varphi \bar{*} f &= \left(|f| \frac{d\varphi f}{|d\varphi f|} \right) \circ \varphi^{-1} \\ \varphi \bar{\star} f &= \left(|f| \frac{d\varphi^{-T} f}{|d\varphi^{-T} f|} \right) \circ \varphi^{-1} \end{aligned}$$

(taking, in both cases, the right-hand side equal to 0 if $|f| = 0$).

We now evaluate the differential of $E_{f,f'}(\varphi) = \|\varphi \cdot f - f'\|_2^2$, where $\varphi \cdot f$ is one of the actions above. We will make the computation below under the assumption that f is C^1 and compactly supported. For the $*$ action, we can observe that, for $\varphi = \text{id} + h$,

$$\varphi * f - f = dh f \circ (\text{id} + h)^{-1} + f \circ (\text{id} + h)^{-1} - f,$$

so that

$$\begin{aligned} \varphi * f - f - dh f + df h &= dh(f \circ (\text{id} + h)^{-1} - f) + f \circ (\text{id} + h)^{-1} \\ &\quad - f \circ (\text{id} - h) + f \circ (\text{id} - h) - f + df h. \end{aligned}$$

Using the fact that $\|(\text{id} + h)^{-1} - \text{id}\|_\infty = \|h\|_\infty$,

$$\|(\text{id} + h)^{-1} - (\text{id} - h)\|_\infty = \|h \circ (\text{id} + h) - h\|_\infty \leq \|h\|_{1,\infty}^2$$

and letting

$$\omega_f^{(1)}(\varepsilon) = \sup_{x \in \mathbb{R}^d} \sup_{|\delta| < \varepsilon} |f(x + \delta) - f(x) - df(x)\delta|$$

we find that

$$\|\varphi * f - f - dh f + df h\|_\infty \leq \|f\|_\infty \|h\|_{1,\infty}^2 + \|f\|_{1,\infty} \|h\|_{1,\infty}^2 + \omega_f^{(1)}(\|h\|_\infty).$$

Noting that $\omega_f^{(1)}(\varepsilon) = o(\varepsilon)$, we find that

$$\|\varphi * f - f - dh f + df h\|_\infty = o(\|h\|_{1,\infty}). \tag{9.59}$$

Using this estimate, it is now easy to show that $E_{f,f'} : \text{Diff}_0^{1,\infty} \rightarrow \mathbb{R}$ is differentiable at $\varphi = \text{id}$ with derivative

$$\begin{aligned} (dE_{f,f'}(\text{id}) \mid h) &= 2\langle dh f - df h, f - f' \rangle_2 \\ &= 2 \int_\Omega (dh f - df h)^T (f - f') dx. \end{aligned}$$

For $f \mapsto \varphi * f$ to map compactly supported C^1 vector fields into vector fields with the same property, we need to take twice-differentiable diffeomorphisms, i.e., $\varphi \in \text{Diff}_0^{2,\infty}$. Over this group, we find that $E_{f,f'}$ is differentiable everywhere, with $(dE_{f,f'}(\psi) \mid h) = (dE_{\psi * f, f'}(\text{id}) \mid h \circ \psi^{-1})$.

The Eulerian derivative is then given by

$$\begin{aligned} (\bar{\partial} E_{f,f'}(\psi) \mid v) &= (dE_{\psi * f, f'}(\text{id}) \mid v) \\ &= 2\langle dv(\psi * f) - d(\psi * f)v, \psi * f - f' \rangle_2. \end{aligned}$$

This expression can be combined with (9.44) to obtain the Eulerian gradient of U , namely

$$\begin{aligned} \bar{\nabla}^V E_{f,f'}(\psi) &= \\ &= 2 \int_\Omega (\partial_2 K(\cdot, x)(\psi * f - f', \psi * f) - K(\cdot, x)d(\psi * f)^T(\psi * f - f')) dx. \end{aligned}$$

The Eulerian differential can be rewritten in another form to avoid the intervention of the differential of h . The following lemma is a consequence of the divergence theorem.

Lemma 9.10 *If Ω is a bounded open domain of \mathbb{R}^d and v, w, h are smooth vector fields on \mathbb{R}^d , then*

$$\int_{\Omega} v^T dh w \, dx = \int_{\partial\Omega} (v^T h)(w^T N) d\sigma_{\partial\Omega} - \int_{\Omega} (w^T dv^T h + (\operatorname{div} w)(v^T h)) dx. \quad (9.60)$$

Equation (9.60) can be rewritten as

$$\langle dh w, v \rangle_2 = ((w^T N)v\sigma_{\partial\Omega} \mid h) - \langle dv w + (\operatorname{div} w)v, h \rangle_2. \quad (9.61)$$

Proof To prove this, introduce the coordinates h^1, \dots, h^d for h and v^1, \dots, v^d for v so that

$$v^T dh w = \sum_{i=1}^d v^i (\nabla h^i)^T w.$$

Now, use the fact that

$$\begin{aligned} \operatorname{div}(v^T h w) &= \operatorname{div}\left(\sum_{i=1}^d v^i h^i w\right) \\ &= \sum_{i=1}^d (v^i (\nabla h^i)^T w + h^i (\nabla v^i)^T w + h_i v^i \operatorname{div} w) \\ &= v^T dh w + h^T dv w + (h^T v) \operatorname{div} w \end{aligned}$$

and the divergence theorem to obtain the result. \square

Using this lemma with Ω large enough so that f vanishes on $\partial\Omega$, we find

$$(dE_{f,f'}(\operatorname{id}) \mid h) = -2\langle (df - df')f + \operatorname{div} f(f - f') + df^T(f - f'), h \rangle_2,$$

which directly provides a new version of the Eulerian derivative at an arbitrary φ , with the corresponding new expression of the Eulerian gradient:

$$\begin{aligned} \bar{\nabla}^V E_{f,f'}(\varphi) &= -2 \int_{\Omega} K(\cdot, x) \left(d(\varphi * f - f')(\varphi * f) \right. \\ &\quad \left. + \operatorname{div}(\varphi * f)(\varphi * f - f') + d(\varphi * f)^T(\varphi * f - f') \right) dx. \end{aligned}$$

Let us now consider the normalized version of this action. We will make the computation under a few additional assumptions on f , namely, that f is compactly supported and $f/|f|$ can be replaced by a smooth unit vector field that can be extended to an open set that contains the support of f . More precisely, we will assume that there exists a scalar function ρ , continuously differentiable and supported by a compact set Q , and a vector field u such that $|u(x)| = 1$ for all x in an open set Ω containing Q , u is continuously differentiable on Ω and $f(x) = \rho(x)u(x)$. With this notation,

we have

$$\varphi \bar{*} f = \rho \circ \varphi^{-1} \frac{\varphi * u}{|\varphi * u|}.$$

Note that, for $z \neq 0$, the derivative of $z/|z|$ is

$$h \mapsto h/|z| - zz^T h/|z|^3 = \frac{1}{|z|} \pi_{z^\perp}(h),$$

where π_{z^\perp} is the orthogonal projection on the space of vectors perpendicular to z .

Consider $\varphi = \text{id} + h$ for some $h \in \text{Diff}_0^{1,\infty}$. Let $\delta = \text{dist}(Q, \Omega^c)$ and assume that

$$\|h\|_\infty = \|\varphi^{-1} - \text{id}\|_\infty < \delta/2.$$

From (9.59), we have, letting Ω' be the set of $x \in \mathbb{R}^d$ such that $\text{dist}(x, Q) < \delta$,

$$\sup_{x \in \Omega'} |(\varphi * u)(x) - u(x) - dh(x)u(x) + du(x)h(x)| = o(\|h\|_{1,\infty}),$$

from which we deduce

$$\sup_{x \in \Omega'} \left| \frac{(\varphi * u)(x)}{|(\varphi * u)(x)|} - u(x) - \pi_{u^\perp}(dh(x)u(x) - du(x)h(x)) \right| = o(\|h\|_{1,\infty}).$$

Since $\|\rho \circ \varphi^{-1} - \rho - \nabla \rho^T h\|_\infty = o(\|h\|_\infty)$, we obtain the fact that $E_{f,f'}$ is differentiable at $\varphi = \text{id}$ with

$$\begin{aligned} (dE_{f,f'}(\text{id}) \mid h) &= 2(-\nabla \rho^T h u + \rho \pi_{u^\perp}(dh u - du h) \mid f - f') \\ &= -2(\nabla \rho^T h u \mid f - f') - 2(\rho(dh u - du h) \mid \pi_{u^\perp}(f')). \end{aligned}$$

Assuming now that $\psi \in \text{Diff}_0^{2,\infty}$, we obtain the fact that $E_{f,f'}$ is differentiable at ψ , with

$$(dE_{f,f'}(\psi) \mid h) = (dE_{\psi \bar{*} f, f'}(\text{id}) \mid h \circ \psi^{-1}).$$

The Eulerian derivative is

$$(\bar{\partial} E_{f,f'}(\psi) \mid v) = (dE_{\psi \bar{*} f, f'}(\text{id}) \mid v).$$

Finally, we note that after integration by parts, we can write

$$\begin{aligned} dE_{f,f'}(\text{id}) &= 2(-u^T(f - f')\nabla \rho + \rho du^T(\pi_{u^\perp}(f')) \\ &\quad + d(\pi_{u^\perp}(f'))f + \text{div}(f)\pi_{u^\perp}(f'))dx. \end{aligned}$$

The computations for $\varphi \star f = (d\varphi^{-T}f) \circ \varphi^{-1}$ and its normalized version are very similar. One only needs to note that (9.59) is now replaced by

$$\|\varphi \star f - f + dh^T f + df h\|_\infty = o(\|h\|_{1,\infty}). \quad (9.62)$$

As a consequence, the formulas for the differentials of the \star and $\bar{\star}$ can be deduced from the $*$ and $\bar{*}$ actions by replacing $(dh f - df f)$ by $(-dh^T f - df h)$.

For the unnormalized action, this yields

$$\begin{aligned} (dE_{f,f'}(\text{id}) \mid h) &= -2\langle dh(f - f'), f \rangle_2 - 2\langle df^T(f - f'), h \rangle_2 \\ &= 2\langle (df - df^T)(f - f') + \text{div}(f - f')f, h \rangle_2 \end{aligned}$$

and $\bar{d}E_{f,f'}(\varphi)$ is obtained by replacing f by $\varphi \star f$. To obtain the differential of $E_{f,f'}$ for the normalized \star action, we get

$$(dE_{f,f'}(\text{id}) \mid h) = -2(\nabla \rho^T h u \mid f - f') + 2(\rho(dh^T u + du h) \mid \pi_{u^\perp}(f')),$$

where $f = \rho u$ as above, which can also be written as

$$dE_{f,f'}(\text{id}) = 2(-u^T(f - f')\nabla \rho + (\rho du^T - df)(\pi_{u^\perp}(f')) - \text{div}(\pi_{u^\perp}(f'))f)dx.$$

As an example of application of vector field matching, let us consider contrast-invariant image registration [90]. If $I : \Omega \rightarrow \mathbb{R}$ is an image, a change of contrast is a transformation $I \mapsto q \circ I$, where q is a scalar diffeomorphism of the image intensity range. The level sets $I_\lambda = \{x, I(x) \leq \lambda\}$ are simply relabeled by a change of contrast, and one obtains a contrast-invariant representation of the image by considering the normals to these level sets, i.e., the vector field

$$f = \nabla I / |\nabla I|$$

with the convention that $f = 0$ when $\nabla I = 0$. Two images represented in this way can now be compared using vector field matching. Since we are using normalized gradients, the natural action is $(\varphi, f) \mapsto \varphi \bar{\star} f$. For our results to hold, some regularization needs to be applied, replacing f by $\rho \tilde{f}$ where $\rho = 1$ and $\tilde{f} = f$ when $|f| = 1$, \tilde{f} is a unit vector field that smoothly extends f over a neighborhood of the domain over which $|f| = 1$, ρ is smooth and vanishes outside this neighborhood.

9.9 Matching Fields of Frames

We now extend vector field deformation models to define an action of diffeomorphisms on fields of positively oriented orthogonal matrices, or frames. We will restrict ourselves to dimension 3, so that the deformable objects considered in this section are mappings $x \mapsto R(x)$, with, for all $x \in \Omega$, $R(x) \in \text{SO}_3(\mathbb{R})$ (the group of rotation matrices).

The $*$ and \star actions we have just defined on vector fields have the nice property of conserving the Euclidean dot product when combined, that is

$$(\varphi * f)^T (\varphi \star g) = (f^T g) \circ \varphi^{-1}.$$

Since $\bar{*}$ and $\bar{\star}$ also conserve the norm, we find that $(\varphi \bar{*} f, \varphi \bar{\star} g)$ is orthonormal as soon as (f, g) is.

We now define an action of diffeomorphisms on fields on frames. Writing $R(x) = (f_1(x), f_2(x), f_3(x))$, we let

$$\varphi \cdot R = (\varphi \bar{*} f_1, (\varphi \bar{\star} f_3) \times (\varphi \bar{*} f_1), \varphi \bar{\star} f_3). \tag{9.63}$$

That this defines an action is a straightforward consequence of $\bar{*}$ and $\bar{\star}$ being actions.

The action can be interpreted as follows. Given a local chart in \mathbb{R}^3 , which is a diffeomorphic change of coordinates $x = m(s, t, u)$, one uniquely specifies a positively oriented frame $R_m = (f_1, f_2, f_3)$ by $f_1 = \partial_1 m / |\partial_1 m|$ and $f_3 = (\partial_1 m \times \partial_2 m) / |\partial_1 m \times \partial_2 m|$. Then, the action we have just defined is such that $\varphi \cdot R$ is the frame associated to the change of coordinates $\varphi \circ m$, i.e.,

$$R_{\varphi \circ m} \circ \varphi = \varphi \cdot R_m.$$

The transformation $m \rightarrow R_m$ has in turn the following interpretation, which is relevant for some medical imaging modalities. Let the change of coordinates be adapted to the following stratified description of a tissue. Curves $s \mapsto m(s, t, u)$ correspond to tissue fibers, and surfaces $(s, t) \mapsto m(s, t, u)$ describe a layered organization. The cardiac muscle, for example, exhibits this kind of structure. Then f_1 in R_m represents the fiber orientation, and f_3 the normal to the layers; $\varphi \cdot R_m$ then corresponds to the tissue to which the deformation φ has been applied.

Frame fields are typically observed over some object-dependent subregion of the observation domain. To account for this, we assume that we are dealing with weighted fields of frames, taking the form $A = \rho R$, where the weight ρ vanishes outside a compact set and R is smooth over a neighborhood of this compact set. We will then consider the action

$$\varphi \cdot A = (\rho \circ \varphi^{-1}) \varphi \cdot R.$$

The computations of the previous sections can now be applied to each column of A . In particular, letting $\varphi = \text{id} + h$, we have $\varphi \cdot A = A + (w_1, w_2, w_3) + o(\|h\|_{1,\infty})$ with, writing $R = (f_1, f_2, f_3)$,

$$\begin{cases} w_1 = -(\nabla \rho^T h) f_1 + \rho \pi_{f_1^\perp}(dh f_1 - df_1 h), \\ w_3 = -(\nabla \rho^T h) f_3 - \rho \pi_{f_3^\perp}(dh^T f_3 + df_3 h), \\ w_2 = -(\nabla \rho^T h) f_2 - \rho \left(\pi_{f_3^\perp}(dh^T f_3 + df_3 h) \right) \times f_1 \\ \quad + \rho f_3 \times \left(\pi_{f_1^\perp}(dh f_1 - df_1 h) \right). \end{cases}$$

Noticing that, for any vector $u \in \mathbb{R}^3$,

$$\left(\pi_{f_3^\perp} u\right) \times f_1 = \left((u^T f_1) f_1 + (u^T f_2) f_2\right) \times f_1 = -(u^T f_2) f_3$$

and similarly $f_3 \times \left(\pi_{f_1^\perp} u\right) = -(u^T f_2) f_1$, we can simplify the expression of w_2 , yielding

$$\begin{cases} w_1 = -(\nabla \rho^T h) f_1 + \rho \pi_{f_1^\perp} (dh f_1 - df_1 h), \\ w_2 = -(\nabla \rho^T h) f_2 + \rho ((dh^T f_3 + df_3 h)^T f_2) f_3 \\ \quad - \rho ((dh f_1 - df_1 h)^T f_2) f_1, \\ w_3 = -(\nabla \rho^T h) f_3 - \rho \pi_{f_3^\perp} (dh^T f_3 + df_3 h). \end{cases} \quad (9.64)$$

Consider the matching functional

$$E_{A,A'}(\varphi) = \int_{\mathbb{R}^3} |\varphi \cdot A - A'|^2 dx$$

with $|A|^2 = \text{trace}(A^T A)$. If $A = \rho R$ and $A' = \rho' R'$, then

$$\begin{aligned} |A - A'|^2 &= 3\rho^2 - 2\rho\rho' \text{trace}(R^T R') + 3\rho'^2 \\ &= 3(\rho - \rho')^2 + 2\rho\rho' \text{trace}(\text{Id} - R^T R'). \end{aligned}$$

Introducing the rotation angle, θ , from R to R' , defined by

$$\text{trace}(R^T R') = 1 + 2 \cos \theta, \quad (9.65)$$

we get

$$|A - A'|^2 = 3(\rho - \rho')^2 + 4\rho\rho'(1 - \cos \theta).$$

Obviously, if $A = (u_1, u_2, u_3)$ and $A' = \rho(u'_1, u'_2, u'_3)$, we also have

$$|A - A'|^2 = |u_1 - u'_1|^2 + |u_2 - u'_2|^2 + |u_3 - u'_3|^2.$$

Using this, one gets the expression of the differential of $E_{A,A'}$ at $\varphi = \text{id}$,

$$(dE_{A,A'}(\text{id}) | h) = 2 \int_{\mathbb{R}^2} \text{trace}(W^T (A - A')) dx,$$

where $W = (w_1, w_2, w_3)$ is given by (9.64). In particular,

$$\begin{aligned} \text{trace}(W^T (A - A')) &= -\nabla \rho^T h (3\rho - \rho' \text{trace}(R^T R')) \\ &\quad + \rho\rho' (dh f_1 - df_1 h)^T \left(-\pi_{f_1^\perp} (f'_1) + (f_1^T f'_2) f_2 \right) \\ &\quad - \rho\rho' (dh^T f_3 + df_3 h)^T \left(-\pi_{f_3^\perp} (f'_3) + (f_3^T f'_2) f_2 \right). \end{aligned}$$

Letting

$$\begin{aligned} u^1_{A,A'} &= \rho\rho' \left(-\pi_{f_1^\perp}(f'_1) + (f_1^T f'_2) f_2 \right) \\ u^3_{A,A'} &= \rho\rho' \left(-\pi_{f_3^\perp}(f'_3) + (f_3^T f'_2) f_2 \right) \end{aligned}$$

and using Lemma 9.10 to eliminate dh , we find

$$\begin{aligned} dE_{A,A'}(\text{id}) &= 2 \left(- (3\rho - \rho' \text{trace}(R^T R')) \nabla \rho - du^1_{A,A'} f_1 - \text{div}(f_1) u^1_{A,A'} \right. \\ &\quad \left. - df_1^T u^1_{A,A'} + \text{div}(u^3_{A,A'}) f_3 \right) dx. \end{aligned} \tag{9.66}$$

9.10 Matching Tensors

The last class of deformable objects we will consider in this chapter are fields of matrices (or tensor fields). For general matrices, we can use the actions we have defined on vector fields, and apply them to each column of M , where M is a field of matrices. The differential of matching functionals is then computed as done in the previous two sections.

One sometimes needs to consider subclasses of tensors, and therefore define an action that leaves this subclass invariant. Here we consider symmetric matrices, which have especially been studied in diffusion tensor imaging (DTI) [6]. The previous actions applied to each column do not work, because they would break the symmetry. A simple choice to address this is to make the diffeomorphism also act on the right, in transpose form, defining, for a field $x \mapsto S(x)$ of symmetric matrices

$$\begin{aligned} \varphi * S &= (d\varphi S d\varphi^T) \circ \varphi^{-1} \\ \varphi \star S &= (d\varphi^{-T} S d\varphi^{-1}) \circ \varphi^{-1}. \end{aligned}$$

We leave to the reader the computation of the differentials of objective functions derived from these actions.

These actions are not necessarily well adapted to DTI data, though, for which alternative options may be considered. DTI produces, at each point x in space, a symmetric positive definite matrix $S(x)$ that measures the diffusion of water molecules in the imaged tissue. Roughly speaking, the tensor $S(x)$ is such that if a water molecule is at s at time t , the probability of being at $x + dx$ at time $t + dt$ is centered Gaussian with variance $dt^2 dx^T S(x) dx$.

If we return to the structured tissue model discussed in the last section (represented by the parametrization $x = m(s, t, u)$), we can assume that molecules travel more easily along fibers, and with most difficulty across layers. So the direction of $\partial_1 m$ is the direction of largest variance, and $\partial_1 m \times \partial_2 m$ of smallest variance, so

that the frame $R_m = (f_1, f_2, f_3)$ associated to the parametrization is such that f_1 is an eigenvector of S for the largest eigenvalue, and f_3 for the smallest eigenvalue, which implies that f_2 is an eigenvector for the intermediate eigenvalue. According to our discussion in the last section, a diffeomorphism φ should transform S so that the frame R_S formed by the eigenbasis of S transforms according to the action of diffeomorphisms on frames, namely, $R_{\varphi \cdot S} = \varphi \cdot R_S$ defined in (9.63).

So, if we express the decomposition of S in the form

$$S = \lambda_1 f_1 f_1^T + \lambda_2 f_2 f_2^T + \lambda_3 f_3 f_3^T$$

with $\lambda_1 \geq \lambda_2 \geq \lambda_3$, we should take

$$\varphi \cdot S = \tilde{\lambda}_1 \tilde{f}_1 \tilde{f}_1^T + \tilde{\lambda}_2 \tilde{f}_2 \tilde{f}_2^T + \tilde{\lambda}_3 \tilde{f}_3 \tilde{f}_3^T \quad (9.67)$$

with $(\tilde{f}_1, \tilde{f}_2, \tilde{f}_3) = \varphi \cdot (f_1, f_2, f_3)$ and $\tilde{\lambda}_i = \lambda_i \circ \varphi^{-1}$, $i = 1, 2, 3$. The action on eigenvalues expresses that intrinsic tissue properties have not been affected by the deformation. If there are reasons to believe that variations in volume should affect the intensity of water diffusion, using the action of diffeomorphisms on densities may be a better option, namely $\tilde{\lambda}_i = \det d(\varphi^{-1}) \lambda_i \circ \varphi^{-1}$.

The action with $\tilde{\lambda}_i = \lambda_i \circ \varphi^{-1}$ is identical to the eigenvector-based tensor reorientation discussed in [6]. One of the important (and required) features of the construction is that, although the eigen-decomposition of S is not unique (when two or three eigenvalues coincide) the transformation $S \mapsto \varphi \cdot S$ is defined without ambiguity. This will be justified below.

The following computations require that $\lambda_1, \lambda_2, \lambda_3$ are C^1 and vanish outside a compact set, and that R_S is smooth in a small neighborhood of this compact set. They are sketchily justified here, as they strongly resemble the computations that have been done before. It will be convenient to introduce the three-dimensional rotation $U_S(\varphi) = ((\varphi \cdot R_S) \circ \varphi) R_S^T$, so that

$$\varphi \cdot S = (U_S(\varphi) S U_S(\varphi)^T) \circ \varphi^{-1}.$$

Taking $\varphi = \text{id} + h$, we have $U_S(\varphi) - S = \omega_S(h) + o(\|h\|_{1,\infty})$, where

$$\begin{aligned} \omega_S(h) = & \pi_{f_1^\perp} dh f_1 f_1^T - ((\pi_{f_3^\perp} dh^T f_3) \times f_1) f_2^T \\ & + (f_3 \times (\pi_{f_1^\perp} dh f_1)) f_2^T - (\pi_{f_3^\perp} dh^T f_3) f_3^T \end{aligned}$$

is a skew-symmetric matrix. With this notation, we can write

$$\varphi \cdot S - S = \omega_S(h) S - S \omega_S(h) - dS h + o(\|h\|_{1,\infty}).$$

(Here dSh is the matrix with coefficients $(\nabla S^{ij})^T h$.) Letting

$$E_{S,S'}(\varphi) = \int_{\Omega} \text{trace}((\varphi \cdot S - S')^2) dx,$$

we then get

$$(dE_{S,S'}(\text{id}) | h) = 2 \int_{\Omega} \text{trace}((S - S')(\omega_S(h)S - S\omega_S(h) - dSh)) dx,$$

with, as usual, for $\psi \in \text{Diff}_0^{2,\infty}$, $(dE_{S,S'}(\psi) | h) = (dE_{\psi \cdot S, S'}(\text{id}) | h \circ \psi^{-1})$ and $(\bar{\partial}E_{S,S'}(\psi) | v) = (dE_{\psi \cdot S, S'}(\text{id}) | v)$.

Here again, the derivatives of h that are involved in $\omega_S(h)$ can be integrated by parts using the divergence theorem. Let us sketch this computation at $\psi = \text{id}$, which leads to a vector measure form for the differential. We focus on the term

$$(\eta | h) := \int_{\Omega} \text{trace}((S - S')(\omega_S(h)S - S\omega_S(h))) dx = \int_{\Omega} \text{trace}(A\omega_S(h)) dx,$$

where $A = S(S - S') - (S - S')S = SS' - S'S$, and want to express η as a vector measure. We have (using the fact that A is skew symmetric and that (f_1, f_2, f_3) is orthonormal)

$$\begin{aligned} -\text{trace}(A\omega_S(h)) &= (\omega_S(h)f_1)^T Af_1 + (\omega_S(h)f_2)^T Af_2 + (\omega_S(h)f_3)^T Af_3 \\ &= (\pi_{f_1^\perp} dh f_1)^T Af_1 - ((\pi_{f_3^\perp} dh^T f_3) \times f_1)^T Af_2 \\ &\quad + (f_3 \times (\pi_{f_1^\perp} dh f_1))^T Af_2 - (\pi_{f_3^\perp} dh^T f_3)^T Af_3 \\ &= (dh f_1)^T u_{S,S'}^1 - (dh^T f_3)^T u_{S,S'}^3, \end{aligned}$$

with

$$\begin{aligned} u_{S,S'}^1 &= \pi_{f_1^\perp}(Af_1 + (Af_2 \times f_3)) \\ \text{and } u_{S,S'}^3 &= \pi_{f_3^\perp}(Af_3 + (f_1 \times Af_2)). \end{aligned}$$

It now remains to use Lemma 9.10 to identify η as

$$\eta = (du_{S,S'}^1 f_1 + \text{div}(f_1)u_{S,S'}^1 - df_3 u_{S,S'}^3 - \text{div}(u_{S,S'}^3) f_3) dx.$$

To write the final expression of $dE_{S,S'}(\text{id})$, define $(S - S') \odot dS$ to be the vector

$$(S - S') \odot dS = \sum_{i,j=1}^3 (S^{ij} - (S')^{ij}) \nabla S^{ij},$$

so that we have

$$dE_{S,S'}(\text{id}) = 2(du_{S,S'}^1 f_1 + \text{div}(f_1)u_{S,S'}^1 - df_3 u_{S,S'}^3 - \text{div}(u_{S,S'}^3) f_3 - (S - S') \odot dS) dx. \quad (9.68)$$

We now generalize this action to arbitrary dimensions, in a way that will provide a new interpretation of the three-dimensional case. Decompose a field of d by d symmetric matrices S in \mathbb{R}^d in the form

$$S(x) = \sum_{k=1}^d \lambda_k(x) f_k(x) f_k(x)^T$$

with $\lambda_1 \geq \dots \geq \lambda_d$ and (f_1, \dots, f_d) orthonormal. The matrices $f_k f_k^T$ represent the orthogonal projections on the one-dimensional space $\mathbb{R} f_k$ and, letting

$$W_k = \text{span}(f_1, \dots, f_k),$$

and noting that the projection on W_k , π_{W_k} is equal to $f_1 f_1^T + \dots + f_k f_k^T$, we can obviously write

$$S(x) = \sum_{k=1}^d \lambda_k(x) (\pi_{W_k(x)} - \pi_{W_{k-1}(x)}),$$

where we have set $W_0 = \{0\}$.

Define the action $S \mapsto \varphi \cdot S$ by

$$\varphi \cdot S = \left(\sum_{k=1}^d \lambda_k \left(\pi_{d\varphi(W_k)} - \pi_{d\varphi(W_{k-1})} \right) \right) \circ \varphi^{-1}.$$

In three dimensions, because

$$(d\varphi f_2)^T \tilde{f}_3 \circ \varphi = (f_2^T f_3) / |d\varphi^{-T} f_3| = 0,$$

we see that $d\varphi f_2 \in \text{span}(\tilde{f}_1 \circ \varphi, \tilde{f}_2 \circ \varphi)$. Since $\tilde{f}_1 \circ \varphi$ is proportional to $d\varphi f_1$, we can conclude that

$$d\varphi \text{span}(f_1, f_2) = \text{span}(\tilde{f}_1 \circ \varphi, \tilde{f}_2 \circ \varphi).$$

This proves that the action we have just defined coincides with the one we have considered for the case $d = 3$.

Returning to the general d -dimensional case, the definition we just gave does not depend on the choice made for the basis f_1, \dots, f_d . Indeed, if we let $\mu_1 > \dots > \mu_q$ denote the distinct eigenvalues of S , and $\Lambda_1, \dots, \Lambda_q$ the corresponding eigenspaces, then, regrouping together the terms with identical eigenvalues in the decomposition

of S and $\varphi \cdot S$, and letting

$$\Gamma_k = \Lambda_1 + \cdots + \Lambda_k, \quad \Gamma_0 = \{0\},$$

we clearly have

$$S(x) = \sum_{k=1}^q \mu_k(x) (\pi_{\Gamma_k(x)} - \pi_{\Gamma_{k-1}(x)})$$

and

$$\varphi \cdot S = \left(\sum_{k=1}^q \mu_k \left(\pi_{d\varphi(\Gamma_k)} - \pi_{d\varphi(\Gamma_{k-1})} \right) \right) \circ \varphi^{-1}.$$

Since the decomposition of S in terms of its eigenspaces is uniquely defined, we obtain the fact that the definition of $\varphi \cdot S$ is non-ambiguous.

9.11 Pros and Cons of Greedy Algorithms

We have studied in this chapter a series of deformable objects, by defining the relevant action(s) that diffeomorphisms have on them and computing the variations of associated matching functionals.

This computation can be used, as we did with landmarks and images, to design “greedy” registration algorithms, which implement gradient descent to progressively minimize the functionals within the group of diffeomorphisms. These algorithms have the advantage of providing relatively simple implementations, and of requiring a relatively limited computation time.

Most of the time, however, this minimization is an ill-posed problem. Minimizers may fail to exist, for example. This has required, for image matching, the implementation of a suitable stopping rule that prevents the algorithm from running indefinitely. Even when a minimizer exists, it is generally not unique (see the example we gave with landmarks). Greedy algorithms provide the minimizer corresponding to the path of steepest descent from where they have been initialized (usually the identity). This solution does not have to be the “best one”, and we will see that other methods can find much smoother solutions when large deformations are involved.

To design potentially well-posed problems, the matching functionals need to be combined with regularization terms that measure the smoothness of the registration. This will be discussed in detail in the next chapter.