Eurographics Symposium on Geometry Processing 2013
Yaron Lipman and Richard Hao Zhang
(Guest Editors)

Volume 32 (2013), Number 5

Dirichlet Energy for Analysis and Synthesis of Soft Maps

Justin Solomon Leonidas Guibas Adrian Butscher

Geometric Computing Group, Stanford University, Stanford, CA

Abstract

Soft maps taking points on one surface to probability distributions on another are attractive for representing surface mappings in the presence of symmetry, ambiguity, and combinatorial complexity. Few techniques, however, are available to measure their continuity and other properties. To this end, we introduce a novel Dirichlet energy for soft maps generalizing the classical map Dirichlet energy, which measures distortion by computing how soft maps transport probabilistic mass from one distribution to another. We formulate the computation of the Dirichlet energy in terms of a differential equation and provide a finite elements discretization that enables all of the quantities introduced to be computed. We demonstrate the effectiveness of our framework for understanding soft maps arising from various sources. Furthermore, we suggest how these energies can be applied to generate continuous soft or point-to-point maps.

1. Introduction

Algorithms for tasks as varied as symmetry detection, motion capture, and medical imaging require reliable methods for mapping between surfaces. However, there are difficulties with even the best mapping algorithms in the form of ambiguities and combinatorial complexity. Symmetry, slippage, discretization issues, a lack of user guidance, and inappropriate regularization assumptions like isometry or conformality contribute to the near-impossibility of finding smooth maps.

One way to overcome these issues is to develop a new *representation* of maps that is better suited for continuity and acknowledges ambiguity. Of particular interest is the probabilistic representation called *soft mapping*. Rather than committing to exact correspondences, a soft map assigns to each point on a source surface a continuous, non-parametric *probability distribution* representing likely matches on a target surface. This contrasts with graphical approaches that perform inference on a predefined class of distributions, e.g. [ASP*04].

Soft maps have several advantages over conventional maps: (1) probability densities can encode superpositions of mapping targets and express uncertainty when the exact target is unknown; (2) probability densities are positive scalar functions that can be discretized easily; (3) there are convex strategies for measuring soft map continuity, descriptor matching, bijectivity, and other desiderata. Moreover, we expect this representation to enjoy a sparsity property: a good soft map has strong peaks near a few targets and is zero elsewhere.

Soft maps can arise in many ways. Most commonly, they can be derived from shape similarity measures such as descriptor differences. On the computationally expensive end, recent methods like [KLM*12, SNB*12] compute soft maps using optimizations balancing different desirable properties. Even so, few theoretical or computational tools exist for analyzing soft maps. Most prominently, the notion of continuity must be redefined appropriately. A key theoretical contribution of this paper is a convex *continuity measure* based on the theory of optimal transportation, along with a principled discretization of this measure using a finite elements solution of a differential equation.

Measuring map continuity is a well-explored topic in differential geometry through the use of the classical *Dirichlet energy* expressing “intrinsic stretching." Although maps minimizing this energy are closely related to conformal maps and have appeared in the geometry processing literature for mapping to simple domains (e.g. planar parametrizations), the formulation of a straightforward and convergent discretization of the Dirichlet energy for maps between non-flat surfaces has received little attention.

In this paper, we introduce a *generalized Dirichlet energy* for measuring the continuity of soft maps. We use this energy together with a *bijectivity energy* to analyze the structure of soft maps arising from different sources. We also suggest how these quantities can be used to regularize the generation of soft maps from an optimization perspective.

© 2013 The Author(s)
Computer Graphics Forum © 2013 The Eurographics Association and Blackwell Publishing Ltd. Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.

1.1. Contributions

The introduction of a soft map Dirichlet energy advances theory and practice in the field of surface mapping. Some of the ideas and techniques we introduce are listed below.

Theory: We define the Dirichlet energy for soft maps using ideas from optimal transportation and discuss its properties. Writing the energy in differential form also reveals finer tools for analyzing soft maps. Secondly, we define a bijectivity energy for soft maps.

Discretization: The soft map Dirichlet energy has a finite elements discretization for maps on triangle meshes. To make the computations tractable, we introduce a partition of unity basis for expressing probabilistic maps.

Map evaluation: We show how our energies and their associated densities can be used to analyze and visualize properties of soft maps. Such methodology is demonstrated on soft maps arising from a variety of sources.

Map generation: We show how soft maps can be used to extract point-to-point images of paths given a single match. Also, we suggest how optimizing our energies can produce continuous soft maps given a few ground truth matches.

1.2. Overview

After discussing related work (§2), we introduce the continuous soft map Dirichlet energy and show how it can be computed using a PDE (§3); we also introduce the continuous bijectivity energy (§3.8). We discretize these objects for computation on a triangle mesh (§4). With the basic theory in place, we consider three applications. First, we use solutions of the equations defining the Dirichlet energy to trace localized correspondences using path integration (§5.1). We then discuss how the Dirichlet and bijectivity energies can be used to analyze soft and point-to-point maps (§5.2) and suggest how they can be optimized to generate soft maps (§5.3).

2. Previous Work

Our main object of study is the *soft map*. Soft maps have existed implicitly in geometry and vision for some time [WL78]. In geometry processing, soft maps have been used to track correspondences before constructing a dense map [BBM05,LF09]; objects like soft maps also appear when relaxing 0/1 integer programs for mapping [WSSC11]. Most recently, works like [TBW*11, KLM*12, SNB*12] consider the problem of constructing soft (or "fuzzy") maps independently of a pipeline for point-to-point correspondence, optimizing for alignment, continuity, or descriptor preservation at least at a coarse scale. [HZG*12] also constructs soft maps to analyze a collection of shapes.

Other than the neighborhood continuity penalties used in [SNB*12, HZG*12], few methods encourage continuity while computing soft maps. By contrast, continuity and stretch are key components for point-to-point mapping methods. For example, isometric and conformal mapping models can lead to successful point-to-point mapping algorithms under these strong assumptions [OMMG10, KLF11, OBCS*12]. Other methods optimize extrinsic stretching directly but suffer with respect to efficiency and convexity due to combinatorial issues [ZSCO*08]. Methods like [CPSS10] also measure stretch for surface deformation, focusing on extrinsic motions of a single surface rather than maps between distinct objects.

The Dirichlet energy is a popular map stretching energy in the mathematical literature [HW08] but is less explored in geometry processing except for special cases where it simplifies. For instance, [BCWG09] computes harmonic maps — critical points of the Dirichlet energy — between volumetric domains for surface deformation. Techniques for planar and spherical parameterization use Dirichlet energies to minimize distortion [EDD*95, DMA03, FH05]. A more general discretization appears in [Bar10], but this approach remains challenging for application in geometry processing. While harmonic maps between surfaces of spherical topology are exactly the conformal maps [GY02], little is understood computationally about the behavior of these maps under additional constraints, in more general topologies, or with additional energy terms.

We discuss work related to optimal transportation in §3.4.

3. The Mathematics of Soft Maps

We assume that we have two smooth, compact surfaces embedded in $\mathbb{R}^3$, namely a "source" M_0 and a "target" M. Denote their gradient operators by ∇_0 and ∇, respectively. Suppose that M_0 and M are normalized to have unit area.

3.1. Background: The Dirichlet Energy of a Map

The *Dirichlet energy* of a conventional map $\phi : M_0 \to M$ is analogous to the Dirichlet energy of a scalar function; i.e.

$$\mathcal{E}_D(\phi) := \int_{M_0} \|\nabla_0 \phi\|^2 dx. \tag{1}$$

The extrema of (1) are called *harmonic maps* and can be characterized as the maps from M_0 into M causing the "least intrinsic stretching." The field of harmonic maps is a long-standing area of research in mathematics. We refer the interested reader to [HW08] for a recent survey.

REMARK: An important distinction between the map and scalar cases is that the derivatives in (1) are *covariant* and so $\|\nabla_0 \phi\|^2$ involves the metrics of M_0 and M. Thus apart from some special cases (e.g. either M_0 or M is Euclidean) it is not known how to discretize (1) in a principled yet simple manner when M_0 and M are replaced by triangle meshes.

3.2. Soft Maps

A *soft map* from M_0 to M is a function $\mu : x \mapsto \mu_x$ assigning a probability measure $\mu_x \in \text{Prob}(M)$ to each $x \in M_0$. Thus if $\mathcal{U} \subseteq M$ we interpret the number $\mu_x(\mathcal{U})$ as the probability that a randomly sampled $y \in \mathcal{U}$ corresponds to x. In this way, correspondences can be expressed with a degree of uncertainty. For all $x \in M_0$ and $\mathcal{U} \subseteq M$, we require that $\mu_x(\mathcal{U}) \geq 0$ and $\mu_x(M) = 1$. A soft map as defined here is the continuous analog of the soft map defined in [SNB*12].

Soft maps generalize conventional point-to-point maps between surfaces. In particular, every map $\phi : M_0 \to M$ yields soft map μ by requiring that $\mu_x(\mathcal{U}) = 1$ if and only if $\phi(x) \in \mathcal{U}$. That is, μ_x is a unit Dirac mass centered at $\phi(x)$.

Soft maps can encode a wider variety of mapping behavior than conventional maps. For example, the precise location of the point corresponding to x might not be known. Then μ_x would have a peak at $\phi(x)$ with nonzero width representing uncertainty in the location of $\phi(x)$. If M admits self-symmetries, then x might correspond equally well to multiple points on M. Then μ_x would be the convex combination of two or more peaked probability measures, as in Fig. 4(a).

A second advantage of soft maps is that they can be represented by positive scalar functions, i.e. their *densities*. The density of the soft map μ is the function $\rho : M_0 \times M \to \mathbb{R}_+$ satisfying $\mu_x(\mathcal{U}) = \int_{\mathcal{U}} \rho(x,y)\,dy$ for all $x \in M_0$ and $\mathcal{U} \subseteq M$. Here dy is the area measure of M. Note $\int_M \rho(x,y)\,dy = 1$ must hold for all $x \in M_0$. Henceforth, we will use the abbreviation $d\mu_x(y) := \rho(x,y)\,dy$. This scalar function representation makes it straightforward to discretize soft maps.

REMARK: Recall that not all probability distributions admit densities. But we consider in this paper only distributions that are at least *weak limits* of smooth densities. Our analysis below remains valid with this assumption.

3.3. Dirichlet Energy for Soft Maps

Let μ be a soft map from M_0 to M. We would like to quantify the degree of smoothness of the function $x \mapsto \mu_x$ in the x-variable. To do so, we construct a *Dirichlet energy* for soft maps in line with the general framework suggested in [Jos94] for Dirichlet energies in metric spaces. We choose a distance metric for $\text{Prob}(M)$, namely the 2-Wasserstein distance $\mathcal{W}_2 : \text{Prob}(M) \times \text{Prob}(M) \to \mathbb{R}_+$ (discussed below).

Definition 1. The *Dirichlet energy density* of a soft map μ from M_0 to M at $x \in M_0$ is

$$e_\mu(x) := \lim_{r \to 0} \frac{1}{\text{Area}(B_r(x))} \int_{B_r(x)} \left[\frac{\mathcal{W}_2(\mu_x, \mu_{x'})}{\text{dist}_0(x,x')}\right]^2 dx' \qquad (2)$$

where dist_0 is the geodesic distance of M_0. The *Dirichlet energy* of μ is then

$$\mathcal{E}_D(\mu) := \frac{1}{\text{Area}(M_0)} \int_{M_0} e_\mu(x)\,dx\,. \qquad (3)$$

3.4. Wasserstein Distances

Wasserstein distances between probability measures arise in the theory of *optimal transportation* and are suitable for our purposes because they take the geometry of the underlying space M into account. We summarize briefly the theory behind these distances.

The optimal transportation problem, first posed by Monge [Mon81], defines the distance between probability distributions as the minimal amount of "work" it takes to carry mass from one distribution to the other. Work is measured by a cost function such as a power of the distance traveled during the transport. To be precise, let $\mu, \nu \in \text{Prob}(M)$ and let $p \in [0,\infty)$. The p-Wasserstein distance is defined as

$$\mathcal{W}_p(\mu,\nu) := \inf_{\pi \in \Pi(\mu,\nu)} \left(\iint_{M \times M} [\text{dist}(y,\bar{y})]^p \, d\pi(y,\bar{y}) \right)^{\frac{1}{p}}, \quad (4)$$

where $\Pi(\mu,\nu)$ is the set of *measure couplings* of μ and ν, i.e. the distributions $\pi \in \text{Prob}(M \times M)$ marginalizing to μ and ν:

$$\int_{x \in M} d\pi(x,y) = d\nu(y) \text{ and } \int_{y \in M} d\pi(x,y) = d\mu(x)\,.$$

We interpret π as a "transport plan" for moving an infinitesimal mass $d\pi(x,y)$ from x to y, with cost $[\text{dist}(x,y)]^p$. The p-Wasserstein distance is the total cost of the optimal plan.

In computer science, the 1-Wasserstein distance is the "Earth Mover's Distance" introduced to the vision community in [RTG98] and applied to geometry problems in [LD11, LPD11, Mém11, dGCSAD11]. We choose the 2-Wasserstein distance because it generalizes the traditional Dirichlet energy for point-to-point maps (see Prop. 2), and because the solution of the optimal transportation problem for this distance is theoretically well-understood and extends to compact, non-flat surfaces (see §3.6).

REMARK: Distances in $\text{Prob}(M)$ can be measured in other ways, e.g. L^p distances or diffusion distances like [LO06, YWL*07]. However, these distances are insufficiently aware of the underlying geometry of M for our purposes.

3.5. Properties of the Dirichlet Energy

We will appeal to three important properties of the Dirichlet energy for soft maps in this paper. We state the first two here and give a more thorough discussion of the key third property in the next section. Please see the Appendix for proofs.

Proposition 2. *The behavior of the Dirichlet energy in two special limiting cases is as follows.*

1. *Let μ be a soft map from M_0 to M with constant density, i.e. $\rho(x,y) = \rho(y)$ for all $x \in M_0$. Then $\mathcal{E}_D(\mu) = 0$.*
2. *Let $\phi : M_0 \to M$ be a map. The Dirichlet energy of the associated soft map equals the Dirichlet energy of ϕ.*

Proposition 3. *The Dirichlet energy is convex under linear combination. That is, if μ_1 and μ_2 are soft maps from M_0 to M and $\alpha \in [0,1]$, then*

$$\mathcal{E}_D((1-\alpha)\mu_1 + \alpha\mu_2) \leq (1-\alpha)\mathcal{E}_D(\mu_1) + \alpha\mathcal{E}_D(\mu_2)\,.$$

3.6. Simplification of the Dirichlet Energy

The expression in (2) for the Dirichlet energy density is unwieldy and does not adapt well to a discrete setting. In particular, the infimum introduces many auxiliary variables for representing measure couplings. Therefore, the discretization of the Dirichlet energy in this form scales poorly with problem size — a key limitation of methods like [SNB*12].

By exploiting the properties of the 2-Wasserstein distance, however, we can simplify the Dirichlet energy density (2) into a form that enables a tractable discretization. To do so, we must introduce a new mathematical object, which we call the *transportation potential* and denote by Q. This object takes as input a point $x \in M_0$ and a tangent vector $V \in T_xM_0$ and outputs a function on M, with linear dependence on V. Given (x,V) we write $y \mapsto Q(x,y) \cdot V$ for the output function.

Proposition 4. *Let* $d\mu_x(y) := \rho(x,y)dy$ *be a soft map from* M_0 *to* M *satisfying a suitable regularity condition. Then its Dirichlet energy satisfies*

$$\mathcal{E}_D(\mu) = \int_{M_0}\int_M \rho(x,y)\|\nabla Q(x,y)\|^2\,dy\,dx, \qquad (5)$$

where Q *is the transportation potential of* μ. *It is found by solving the partial differential equation*

$$\begin{aligned}\nabla\cdot\big(\rho(x,y)\nabla Q(x,y)\cdot V\big) &= -\langle\nabla_0\rho(x,y),V\rangle\\ \int_M \rho(x,y)Q(x,y)\cdot V\,dy &= 0\end{aligned} \qquad (6)$$

for every $x \in M_0$ *and* $V \in T_xM_0$.

The derivation of (5) uses properties of the solution of the optimal transportation problem for the 2-Wasserstein distance [Vil03, Ch. 2] and will be sketched in the Appendix.

REMARK: The differential operator in (6) is linear and has the constant functions in its kernel. The second equation in (6), however, ensures that solutions are transverse to the constant functions. Thus the solutions of (6) are unique.

3.7. Discussion of the Transportation Potential

Intuition. Despite its involved mathematical definition, the intuition for Q is straightforward. For each $x \in M_0$, we visualize the distribution $d\mu_x(y) = \rho(x,y)\,dy$ as a collection of particles on M whose density near y is proportional to $\rho(x,y)$. If we choose a small vector $V \in T_xM_0$ and displace x to $x' := \exp_x(V)$, we can track the motion of these particles on M as they rearrange themselves in a Wasserstein-optimal manner from μ_x to $\mu_{x'}$. The vector field $\rho(x,y)\nabla Q(x,y)\cdot V$ is the momentum of these particles as they move and the cost (5) is twice their kinetic energy. This interpretation aligns with the Benamou-Brenier formulation of optimal transportation in terms of fluid flow [BB00], see also [Vil03, Ch. 8].

The nature of the defining equation. The partial differential equation (6) satisfied by Q is an anisotropic version of Poisson's equation. We can invoke standard theory when $\rho > 0$ to establish existence and uniqueness of Q. If this inequality does not hold, these properties can fail. Then, solvability of (6) can be restored by considering the $\varepsilon \to 0$ limit of the equation for $(1-\varepsilon)\rho + \varepsilon\rho_0$ where $\rho_0(x,y)$ is uniform on M for all $x \in M_0$, see [AGS05, Ch. 8]. We return to this point in Sec. 4.4 when we discuss the discretization of (6).

3.8. Bijectivity Energy for Soft Maps

A bijectivity energy for soft maps should promote the equal distribution of probabilistic mass pushed forward under the soft map. To this end, let $d\mu_x(y) := \rho(x,y)\,dy$ be a soft map between M_0 and M, and consider the quantity $b(y) := \int_{M_0}\rho(x,y)\,dx$. Note that $b(y) \geq 0$ and $\int_M b(y)\,dy = 1$ so $b(y)\,dy$ is a probability measure on M. Indeed $\int_{\mathcal{U}} b(y)\,dy$ gives the probability that a randomly sampled $y \in \mathcal{U} \subseteq M$ receives mass from *somewhere* in M_0. We view b as a *bijectivity energy density* whose square integral yields the bijectivity energy, a convex quadratic function of μ.

Definition 5. The bijectivity energy of a soft map μ from M_0 to M, with $d\mu_x(y) := \rho(x,y)dy$ is

$$\mathcal{E}_b(\mu) := \int_M\left[\int_{M_0}\rho(x,y)\,dx\right]^2 dy. \qquad (7)$$

We can see that μ has small bijectivity energy when $b(y)$ is nearly constant and $b(y)\,dy$ is as spread out as possible. Thus for such μ, most $y \in M$ receive mass from M_0 and no y receives a large amount of mass. A final desirable property of the bijectivity energy is that it has the “correct" limit for soft maps arising from conventional maps.

Proposition 6. *Suppose* $\phi : M_0 \to M$ *is a diffeomorphism and let* μ *be the associated soft map. Then*

$$\mathcal{E}_b(\mu) = \int_M \left[\det(\nabla_0\phi)\right]^{-2}\circ\phi^{-1}(y)dy.$$

See the Appendix for proof. Therefore a conventional map for which $\det(\nabla_0\phi)$ is too small will have large bijectivity energy, and so $\mathcal{E}_b$ penalizes the failure of local injectivity. Moreover, the formula

$$\det(\nabla_0\phi(x)) = \lim_{r\to 0}\frac{\text{Area}(\phi(B_r(x)))}{\text{Area}(B_r(x))},$$

where $B_r(x)$ is the ball of radius r centered at x, tells us that energy-optimal maps will be such that $\det(\nabla_0\phi(x))$ is as spread out as possible, which promotes non-zero relative values of $\text{Area}(\phi(B_r(x)))$ for each x, i.e. local surjectivity.

4. Discretization

With the theory of soft map Dirichlet energies in place, we derive in this section a suitable discretization using finite elements. Henceforth we will assume that the “source” surface M_0 and the “target” surface M are represented as triangle meshes with unit area. We first construct larger-scale

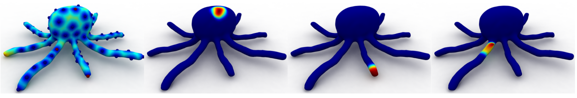


Figure 1: *(left) One hundred geodesic farthest point samples on a mesh colored by distance to the closest sample; (right) partition of unity basis functions corresponding to the three red samples.*

piecewise-linear bases on M_0 and M to decrease the degrees of freedom during our calculations (§4.2). We expand ρ and Q on these bases and pose a family of linear systems for the discretized Q allowing computation of the Dirichlet energy (§4.5). Finally, we discretize the bijectivity energy (§4.6).

4.1. Preliminaries

At the finest level we represent functions on a triangle mesh as linear combinations of piecewise linear "hat" functions whose coefficients equal function values at the vertices. Gradients of these functions are per-face vector fields; the inner product of two such vector fields is taken with respect to the Euclidean metric of the face. We represent a density on a triangle mesh as a set of positive weights per triangle integrating to one. These choices are the simplest possible given the orders of differentiability required to compute Q from ρ.

4.2. Basis Elements

Since soft maps are represented by scalar functions on $M_0 \times M$, the space of soft maps will have dimensionality on the order of the product of the number of triangles of the two surfaces. This is often a prohibitively large number from the point of view of storage and optimization.

We therefore employ smaller bases for our computations. We obtain these from *partitions of unity* (POUs) — collections of smooth, positive functions with overlapping supports whose sum is equal to one at each point. To construct these, we partition the surface into geodesic farthest-point Voronoi cells and solve Poisson's equation $\Delta f = -1$ on each cell's one-ring with Dirichlet boundary conditions. The result is a nonnegative f, strictly positive on the original cell by the maximum principle. Denote the minimum value by m. To remove values of f smaller than εm, we take $\varepsilon \in (0,1)$ and replace f by $\gamma \circ f$ where $\gamma : \mathbb{R} \to \mathbb{R}$ is smooth and monotone, with $\gamma(x) = x$ for $x \geq m$ and $\gamma(x) = 0$ for $x < \varepsilon m$. After performing these steps for all cells, we have a collection of smooth, positive functions supported in their one-rings; and at least one function is non-zero at each vertex. We normalize by the sum of the values of the functions at each vertex to obtain the desired POU. Examples are shown in Fig. 1.

Applying this procedure to M_0 and M yields two POU bases $\{\beta_{0i} : i = 1, \ldots, n_0\}$ and $\{\beta_j : j = 1, \ldots, n\}$ for M_0 and M, respectively. We can assume that both sets of functions are linearly independent (but not orthogonal) in the L^2 sense. Additionally, we introduce a per-face version of the POU on M by averaging the values at the three vertices of each triangle, which we subsequently re-normalize to have integral one. Denote this new set of densities by $\{\mu_j : j = 1, \ldots, n\}$.

REMARK: Previous approaches make use of a variety of strategies for dimensionality reduction. [OBCS*12] uses low-frequency Laplace-Beltrami eigenfunctions for representing maps. This basis is suboptimal for nonnegative and sparse probability distributions because of negative lobes, non-locality, and ringing effects. Alternatively, [SNB*12] uses a basis of indicator functions relative to a coarse surface partition. These functions are nonnegative and local; but their non-smoothness makes them undesirable in our setting.

4.3. Discrete Soft Maps

We express a discrete soft map from M_0 to M in terms of its density function as $d\mu_x(y) = \rho(x,y)\,dy$ and expand ρ on the basis elements as

$$\begin{aligned} \rho(x,y) &:= \sum_{ij} C_{ij} \beta_{0i}(x) \mu_j(y) \\ \text{where} \quad C_{ij} &\geq 0 \; \forall i,j \quad \text{and} \quad \sum_j C_{ij} = 1 \; \forall i \end{aligned} \tag{8}$$

Thus $\rho(x,y)$ is non-negative and integrates to one for each x.

4.4. Discrete Transportation Potential

Representation. The discrete transportation potential Q takes as input a vector tangent to a face of M_0 and outputs a function on M. To discretize this functionality, we choose an orthonormal basis $\{V^s_f : s = 1, 2\}$ for each $f \in F_0$ (the set of faces of M_0) and introduce the one-form ω^s_f defined by $\omega^s_f(X) := \langle V^s_f, X \rangle$ for any vector X tangent to f. Now we can expand Q with coefficients $q^s_{fj} \in \mathbb{R}$ as

$$Q(f,y) := \sum_{s=1,2} \sum_j q^s_{fj} \omega^s_f \beta_j(y) \,. \tag{9}$$

Defining equation. The discretization of the differential equation (6) involves straightforward algebra. We substitute the expansions for ρ and Q and integrate (by parts on the left hand side) against the test function $\beta_k(y)$. We then project onto each face f and direction V^s_f. This yields

$$\sum_{ijj'} C_{ij} W_{jj'k} B^+_{0if} q^s_{fj'} = \sum_{ij} C_{ij} U_{jk} B^-_{0ifs} \tag{10}$$

for all $f \in F_0$ and $s = 1, 2$, where

$$\begin{aligned} B^+_{0if} &:= \int_f \beta_{0i}(x)\,dx \\ B^-_{0ifs} &:= \int_f \langle \nabla_0 \beta_{0i}(x), V^s_f \rangle\,dx \\ U_{jk} &:= \int_M \mu_j(y) \beta_k(y)\,dy \\ W_{jj'k} &:= \int_M \mu_j(y) \left\langle \nabla \beta_{j'}(y), \nabla \beta_k(y) \right\rangle dy \,. \end{aligned} \tag{11}$$

Note that the quantities in (11) depend only on the geometry of M_0 and M and can be pre-computed easily using triangle areas and inner products of hat functions.

The uniqueness constraint. For the constraint appearing in (6), we proceed as above to find

$$\sum_{ijj'} q^s_{fj'} C_{ij} U_{jj'} B^+_{0if} = 0 \tag{12}$$

for all $f \in F_0$ and $s = 1,2$.

The linear solve. The majority of the computation time in our pipeline is spent solving the system (10) subject to (12). As we have seen in §3.7, if the strict inequality $\rho > 0$ holds then the equation for Q is uniquely solvable. This condition corresponds to a positive definite solve for Q in the discrete system. When ρ is near-zero, the system becomes poorly conditioned and eventually admits a null space. We avoid this by computing Q for $(1-\varepsilon)\rho + \varepsilon\rho_0$ as discussed in §3.7, with $\varepsilon = 10^{-5}$ in our implementation. Also note that the system for Q decouples over faces. Thus, we solve a linear system per face sized proportionally to the number of basis functions.

4.5. Discrete Dirichlet Energy

For the discrete Dirichlet energy we again substitute the expansions for ρ and Q and integrate over both surfaces, giving

$$\mathcal{E}_D(C) = \sum_{ijkk'} \sum_{s=1,2} \sum_{f \in F_0} C_{ij} q^s_{fk} q^s_{fk'} B^+_{0if} W_{jkk'} \,. \tag{13}$$

For optimization, we also need the gradient of the discrete Dirichlet energy (13). This is complicated since q^s_{fj} depends on C_{ij} via the defining equation (10). Thus we state the result here and defer the proof for the Appendix. Note that the gradient is easy to evaluate given Q.

Proposition 7. *The gradient of the energy* (13) *is*

$$D_{ij} := \sum_{k'} \sum_{s=1,2} \sum_{f \in F_0} \left(2U_{jk'} B^-_{0ifs} - \sum_k q^s_{fk} W_{jkk'} B^+_{0if} \right) q^s_{fk'} \,.$$

4.6. Discrete Bijectivity Energy

We discretize the bijectivity energy by substituting the expansion (8) into (7). The resulting expression is

$$\mathcal{E}_b(C) := \sum_{ii'jj'} C_{ij} C_{i'j'} B_{0i} B_{0i'} V_{jj'}$$

where $B_{0i} = \int_{M_0} \beta_{0i}(x)\,dx$ and $V_{jj'} := \int_M \mu_j(y) \mu_{j'}(y)\,dy$. These quantities can also be pre-computed.

5. Applications for Soft Map Analysis and Synthesis

In this section, we show how the transportation potential Q and the Dirichlet energy can be used for analyzing and synthesizing soft maps.

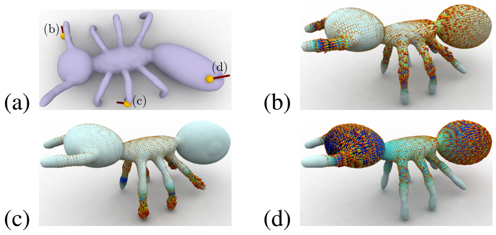


Figure 2: *(a) An ant model with three source points and directions; (b,c,d) corresponding momentum fields for a soft map from the heat kernel signature. The blue shading of the targets shows the magnitude of the soft map in each case.*

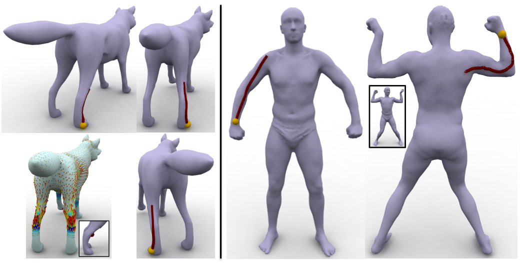


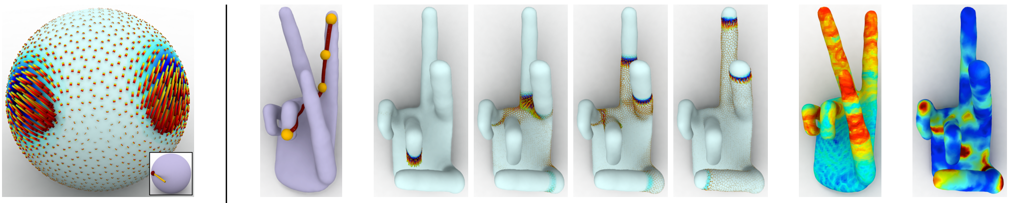
Figure 3: *(left) The soft map (lower left) in blue does not distinguish between the two legs of the wolf, nor between radial points at the same height on a given leg. The velocity vector field ∇Q, however, deduces second-order information from the soft map. Given a path (upper left) on the source, corresponding paths (middle left) can be integrated along ∇Q from a single match; the soft map encodes forward and orientation-reversing maps, which can be traced depending on the initial match shown in yellow. (right) Following ∇Q transfers the path on the left model to the path on the right model. Path integration disambiguates the soft map, which does not differentiate points on rings around the arms.*

5.1. Extracting Correspondences

The soft map density ρ is a function on the four-dimensional product space $M_0 \times M$ and is therefore hard to visualize. Picking a few salient source points $x \in M_0$ and showing the corresponding distributions on M gives some sense of the correspondence behavior suggested by ρ. But it remains difficult to see *dynamic* behavior in these correspondences as a source point is moved in a given direction.

As discussed in §3.7 however, the momentum field $\rho\nabla Q$ of ρ captures exactly this type of behavior. In Figures 2, and 4, we illustrate how the momentum might be used to illustrate the dynamics of a soft map.

In particular, the fluid flow interpretation of optimal transportation views the probability density ρ as a collection of particles whose aggregate motion is by advection under the momentum field. We leverage this idea to extract pointwise correspondences from a soft map as follows. Suppose we are given a soft map density ρ from M_0 to M and a single point-to-point correspondence $x \in M_0 \mapsto y \in M$. We can assume that



(a) (b) (c) (d) (e)

Figure 4: *(a) Soft maps are capable of representing superpositions of uncertain point-to-point maps. We simulate such a soft map (the average two soft maps representing the identity and an orientation-reversing map of the sphere); the momentum field $\rho\nabla Q$ in red tracks the motion of both peaks of the map simultaneously (source point and direction boxed). (b) Four points marked along a geodesic on a source surface; (c) the corresponding images of a soft map from (b) constructed using the wave kernel signature [ASC11] on the target shaded from light to dark and with the momentum vector field in red; (d) the Dirichlet energy density of the map in log scale; (e) the reciprocal of the bijectivity energy density. Note in (c) that $\rho\nabla Q$ shows where mass of the soft map moves: up and down the target fingers and along the hand; the WKS cannot distinguish between the index and middle fingers. The Dirichlet energy density (c) is highest exactly where the map induces distortion at the joints.*

y is near a peak of μ_x. Now given a path $\gamma_0(t) : [0,T) \to M_0$ starting at x, we can trace a path $\gamma : [0,T) \to M$ of corresponding points on M using the soft map velocity as a guide. In particular, we obtain an ordinary differential equation (ODE) for γ by substituting γ_0 for x and its derivative $\dot\gamma_0$ for V into the expression for the velocity:

$$\dot\gamma = \nabla Q(\gamma_0,\gamma)\cdot\dot\gamma_0 \quad \text{with} \quad \gamma(0) = y \qquad (14)$$

This ODE can be integrated in t using Euler's method. Q in the right hand side of (14) can be found rapidly since the only q^s_{fj} needed are those for the face f containing $\gamma_0(t)$. Fig. 3 show two examples of this process. The left subfigure also shows that different choices of $y \in M$ yield different paths; this provides a strategy for isolating symmetries in ρ.

5.2. Analyzing Maps

The transportation potential, Dirichlet energy density, and bijectivity energy density all can be used to visualize and analyze characteristics of soft maps. We demonstrate our visualization techniques on maps from two sources.

First, when we are given a point-to-point map $\phi : M_0 \to M$. Prop. 2 shows that the traditional Dirichlet energy of ϕ is approximated by the Dirichlet energy of any sufficiently "close" soft map. This suggests that a strategy for analyzing the traditional Dirichlet energy of ϕ is to construct a soft map approximating ϕ and computing its Dirichlet energy. To do so, we represent ϕ at the finest level using a density ρ^{fine} that is sparse in the basis of hat functions on M_0 and per-triangle densities on M. For efficiency, we then project into the bases $\{\beta_{0i}\}$ and $\{\mu_j\}$ using a discretization of the formula

$$\rho^{\text{coarse}}(x,y) := \sum_{ij} \beta_{0i}(x)\mu_j(y) \int_M \rho^{\text{fine}}(x,\bar y)\beta_j(\bar y)\,d\bar y$$

This formula is justified by a simple sampling strategy that assigns weight in proportion to the height of the basis function and avoids negative values and ringing that would arise from least-squares projection. Fig. 5 shows how the behavior of the traditional Dirichlet energy for maps aligns with that of the Dirichlet energy of the projected soft maps.

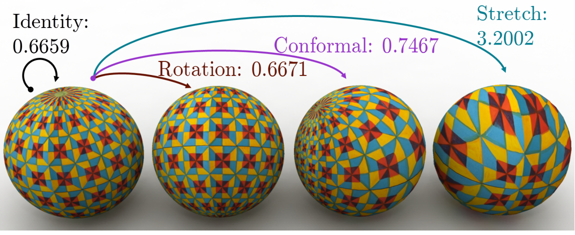


Figure 5: *Four sphere maps illustrated by transferring textures. The Dirichlet energies of the approximating soft maps is indicated. The identity, rotation, and conformal maps yield nearly the same energy while the stretched map is larger. We expect low values for the first three maps because they are critical points of the point-to-point Dirichlet energy; the only map with exactly zero Dirichlet energy in the point-to-point case sends the entire source to a single point on the target.*

Second, soft maps can be obtained by constructing the maximum-entropy probability distribution derived from descriptor differences. That is, suppose $f_0 : M_0 \to \mathbb{R}^d$ and $f : M \to \mathbb{R}^d$ are descriptors for M_0 and M. Then the distribution we have in mind is $\rho(x,y) = Z(x)^{-1}\exp(-\|f_0(x) - f(y)\|^2/\sigma^2)$, where $Z(x)$ is a normalization factor. We use this as our soft map and project it onto the bases as above. In our examples, we use the heat kernel and wave kernel signatures [SOG09, ASC11], similar to [SNB*12].

In Figs. 5 & 6, we use Dirichlet and bijectivity energy densities to visualize and quantify point-to-point map continuity; the remaining figures are constructed using descriptors. Fig. 4 shows the Dirichlet density as a function on a mesh. Fig. 7 shows how it can be used to find poorly-generated soft maps and points where distortion is undesirably high.

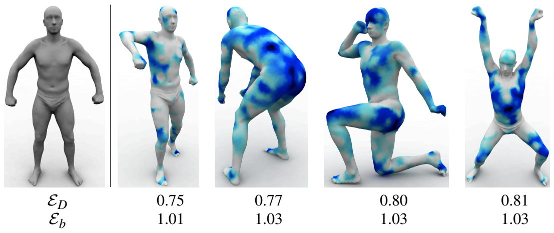


Figure 6: *Four deformations of the mesh on the left with ground-truth correspondences colored by the reciprocal of the bijectivity energy density. Below the meshes are the Dirichlet energy $\mathcal{E}_D$ and bijectivity energy $\mathcal{E}_b$.*

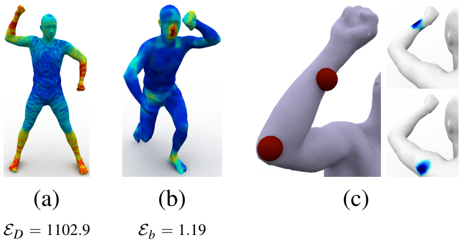


Figure 7: *Constructing a soft map from (a) to (b) using the wave kernel signature unexpectedly yields a large Dirichlet energy $\mathcal{E}_D$. Examining the energy density (a; in log scale) shows large distortion on the model's forearms; examining the map more closely in (c) reveals that the map inadvertently stretches the source's forearm to the target's entire arm.*

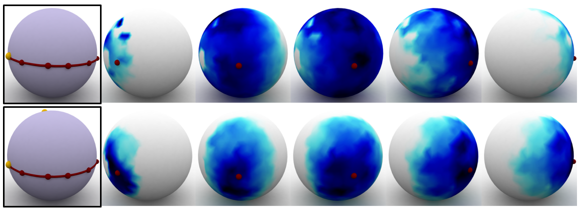


Figure 8: *(top) A map constructed by optimizing $\mathcal{E}(\mu)$ with one constraint fixing the yellow point on the left; images of the points on the red curve are shown to the right. (bottom) The same after fixing three points; one yellow fixed point is occluded. As might be expected, fixing one point yields rings around the constraint since it is impossible to disambiguate the rotational symmetry of the sphere. Fixing three points, however, yields distributions with single peaks.*

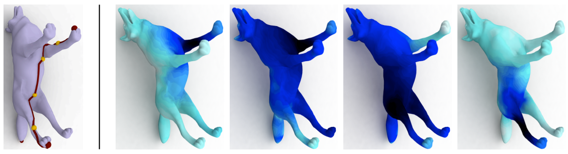


Figure 9: *A soft map (right) between a straight and a bent wolf model constructed by optimizing $\mathcal{E}(\mu)$ after fixing three points marked in red (left). The resulting distributions, constructed solely using continuity and constraints, are wide but have peaks aligned with their counterparts on the source.*

5.3. Generating Optimal Soft Maps

We have constructed and analyzed two convex energies that encode desirable properties of a soft map. We now propose a convex optimization strategy for generating soft maps by minimizing these energies with geometric constraints.

To be precise, we consider soft maps minimizing $\mathcal{E}(\mu) := \mathcal{E}_D(\mu) + \lambda\mathcal{E}_b(\mu)$ where λ trades off between the Dirichlet energy and the bijectivity energy. The behavior we expect for minima of this energy can be characterized as follows. The minimum of $\mathcal{E}_D$ is zero and highly degenerate, achieved by soft maps with constant density independent of x. When $\lambda > 0$, this degeneracy is removed since $\mathcal{E}_B$ favors a unique minimum, the uniform distribution. Adding geometric constraints as correspondences between points or regions pulls the optimum away from the uniform distribution, yielding soft maps that are as spread out as possible while interpolating continuously in the x-variable. $\mathcal{E}_D$ and $\mathcal{E}_b$ are similarly scaled, and we find that $\lambda \approx 1$ achieves a reasonable compromise.

We minimize $\mathcal{E}$ using the L-BFGS algorithm [NW06]. Figs. 8 and 9 show examples of output from the optimization; computation of Q limits the speed of each iteration, so convergence criteria were purposefully loose. Although the maps are broad, clear peaks encode the likely correspondences from which maps can be generated. Additional work, however, will be required to evaluate the properties of this optimization systematically. In particular, while the correspondence constraints prevent the map from reaching the minimum energy achieved by constant or uniform mapping targets, sparsity terms or sharper basis functions may be needed to yield optimal maps with clear peaks. Testing such hypotheses will require more tuned optimization iterations or larger steps without recomputing the full Q matrix.

6. Discussion

We have implemented the techniques discussed above in MATLAB using C++ for matrix-tensor products. Finding the full matrix of the transportation potential requires two sparse positive definite linear solves per triangle of M_0 as explained in §4.5. On a 2.4 GHz i5 notebook processor with 4 GB of memory, solving for Q takes approximately five minutes for a mesh with 1300 vertices equipped with 300 basis functions.

The largest limitation in our setup is the time needed to compute Q. While this paper focuses on the theory and applications of soft Dirichlet energies, there are two potential engineering solutions to this problem that likely would alleviate timing issues. Most obviously, our implementation uses threading to parallelize the calculations on a dual-core processor, but in reality these solves are all independent and could be carried out in parallel given a specialized implementation. A second alternative is to introduce broader bases for one-forms on M_0, bringing about savings similar to those achieved using partitions of unity. This basis need not be localized, so eigenbases of the one-form Laplacian [Hir03] or

Killing operator [BCBSG10] would be acceptable. But now the solves for the transportation potential no longer decouple, so additional considerations will be needed.

Beyond timing, we find that our finite elements discretization is relatively stable for many different types of soft maps. Poorly-shaped triangles can create localized issues, but we have not found them to affect the energies in question significantly for the meshes we have used for testing. Noise in the mapping matrix C also has little effect on computation.

Finally, while our main focus is on analyzing preexisting soft maps, our outlined mapping method in §5.3 shows considerable promise but requires refinement. Our energies are convex and nonlinear, and thus evaluating the Dirichlet energy must be accomplished as fast as possible to maximize the feasible number of iterations of L-BFGS or any other optimization. An additional sparsity term such as [PEC12] would also encourage more localized mapping behavior.

7. Conclusion

Our techniques formulate a notion soft map continuity and propose a Dirichlet energy for its measurement. This energy is well-defined for smooth surfaces and has a principled discretization based on finite elements. We show how our energies and their associated densities can be used to analyze and visualize soft maps arising from a variety of sources. They also can be used to extract point-to-point correspondences and suggest how an optimization framework could be used to compute soft maps given a few geometric constraints.

Together, these methods provide a toolkit for exploring and generating soft maps and indicate promising avenues for future research in surface mapping and analysis. As suggested earlier, many geometry processing algorithms implicitly make use of soft maps through descriptor differencing or by accumulating potential matches, and our proposed techniques can be used to understand the quality and structure of these constructions, including their discontinuities and locations where additional mapping evidence might increase bijectivity or sharpness. They also provide methods for displaying local variations of soft maps using momenta rather than small differences between probability distributions. With more specialized optimizations, it also may be possible to refine the technique in §5.3 to compute dense, continuous soft maps in analogy to the coarse maps computed in [SNB*12]. In the end, this work represents a considerable step toward the design of a pipeline for generating and understanding soft maps backed by a convergent theory characterizing discrete and continuous behavior.

Acknowledgments. The authors gratefully acknowledge the support of NSF grants IIS 914833, CCF 1161480, DMS 1228304, NSF/NIH grant 900700, AFOSR grant 1156110-1-TACAD, a Google Research Award, the National Defense Science and Engineering Graduate Fellowship, the Hertz Foundation Fellowship, the NSF GRF program, and the Max Planck Center for Visual Computing and Communications.

References

[AGS05] AMBROSIO L., GIGLI N., SAVARÉ G.: *Gradient flows: in metric spaces and in the space of probability measures*. Springer, 2005. 4

[ASC11] AUBRY M., SCHLICKEWEI U., CREMERS D.: The wave kernel signature: A quantum mechanical approach to shape analysis. In *Computer Vision Workshops (ICCV)* (2011), pp. 1626–1633. 7

[ASP*04] ANGUELOV D., SRINIVASAN P., PANG H.-C., KOLLER D., THRUN S., DAVIS J.: The correlated correspondence algorithm for unsupervised registration of nonrigid surfaces. In *NIPS* (2004). 1

[Bar10] BARTELS S.: Numerical analysis of a finite element scheme for the approximation of harmonic maps into surfaces. *Math. Computation 79*, 271 (2010), 1263–1301. 2

[BB00] BENAMOU J.-D., BRENIER Y.: A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numerische Mathematik 84*, 3 (2000), 375–393. 4

[BBM05] BERG A., BERG T., MALIK J.: Shape matching and object recognition using low distortion correspondences. In *CVPR* (2005), vol. 1, pp. 26–33. 2

[BCBSG10] BEN-CHEN M., BUTSCHER A., SOLOMON J., GUIBAS L.: On discrete Killing vector fields and patterns on surfaces. *CGF 29*, 5 (2010), 1701–1711. 9

[BCWG09] BEN-CHEN M., WEBER O., GOTSMAN C.: Variational harmonic maps for space deformation. *TOG 28*, 3 (2009), 34:1–34:11. 2

[CPSS10] CHAO I., PINKALL U., SANAN P., SCHRÖDER P.: A simple geometric model for elastic deformations. *TOG 29*, 4 (July 2010), 38:1–38:6. 2

[dGCSAD11] DE GOES F., COHEN-STEINER D., ALLIEZ P., DESBRUN M.: An optimal transport approach to robust reconstruction and simplification of 2d shapes. In *CGF* (2011), vol. 30, pp. 1593–1602. 3

[DMA03] DESBRUN M., MEYER M., ALLIEZ P.: Intrinsic parameterizations of surface meshes. In *CGF* (2003), vol. 21, pp. 209–218. 2

[EDD*95] ECK M., DEROSE T., DUCHAMP T., HOPPE H., LOUNSBERY M., STUETZLE W.: Multiresolution analysis of arbitrary meshes. In *SIGGRAPH* (1995), pp. 173–182. 2

[FH05] FLOATER M., HORMANN K.: Surface parameterization: a tutorial and survey. In *Advances in Multiresolution for Geometric Modelling*, Dodgson N., Floater M., Sabin M., (Eds.), Mathematics and Visualization. Springer, 2005, pp. 157–186. 2

[GY02] GU X., YAU S.-T.: Computing conformal structures of surfaces. 121–146. 2

[Hir03] HIRANI A.: *Discrete Exterior Calculus*. PhD thesis, Caltech, 2003. 8

[HW08] HÉLEIN F., WOOD J.: Harmonic maps. In *Handbook of global analysis*. Elsevier Science, 2008, pp. 417–491. 2

[HZG*12] HUANG Q.-X., ZHANG G.-X., GAO L., HU S.-M., BUTSCHER A., GUIBAS L.: An optimization approach for extracting and encoding consistent maps in a shape collection. *TOG 31*, 6 (2012), 167:1–167:11. 2

[Jos94] JOST J.: Equilibrium maps between metric spaces. *Calc. Var. and PDE 2* (1994), 173–204. 3

[KLF11] KIM V. G., LIPMAN Y., FUNKHOUSER T.: Blended intrinsic maps. *TOG 30*, 4 (2011), 79:1–79:12. 2

[KLM*12] KIM V. G., LI W., MITRA N. J., DIVERDI S., FUNKHOUSER T.: Exploring collections of 3d models using fuzzy correspondences. *TOG 31*, 4 (2012), 54:1–54:11. 1, 2

[LD11] LIPMAN Y., DAUBECHIES I.: Conformal Wasserstein distances: Comparing surfaces in polynomial time. *Advances in Mathematics 227*, 3 (2011), 1047–1077. 3

[LF09] LIPMAN Y., FUNKHOUSER T.: Möbius voting for surface correspondence. *TOG 28*, 3 (July 2009), 72:1–72:12. 2

[LO06] LING H., OKADA K.: Diffusion distance for histogram comparison. In *CVPR* (2006), vol. 1, pp. 246–253. 3

[LPD11] LIPMAN Y., PUENTE J., DAUBECHIES I.: Conformal Wasserstein distance: Ii. computational aspects and extensions. *arXiv preprint:1103.4681* (2011). 3

[Mém11] MÉMOLI F.: Gromov–Wasserstein distances and the metric approach to object matching. *Foundations of Computational Mathematics 11*, 4 (2011), 417–487. 3

[Mon81] MONGE G.: Mémoire sur la théorie des déblais et remblais, mémoires acad. *Royale Sci. Paris* (1781). 3

[NW06] NOCEDAL J., WRIGHT S.: *Numerical Optimization*, 2 ed. Springer, 2006. 8

[OBCS*12] OVSJANIKOV M., BEN-CHEN M., SOLOMON J., BUTSCHER A., GUIBAS L.: Functional maps: A flexible representation of maps between shapes. *TOG 31*, 4 (2012). 2, 5

[OMMG10] OVSJANIKOV M., MÉRIGOT Q., MÉMOLI F., GUIBAS L.: One point isometric matching with the heat kernel. In *CGF* (2010), vol. 29, pp. 1555–1564. 2

[PEC12] PILANCI M., EL GHAOUI L., CHANDRASEKARAN V.: Recovery of sparse probability measures via convex programming. In *NIPS* (Dec. 2012). 9

[RTG98] RUBNER Y., TOMASI C., GUIBAS L. J.: A metric for distributions with applications to image databases. In *ICCV* (1998), pp. 59–66. 3

[SNB*12] SOLOMON J., NGUYEN A., BUTSCHER A., BEN-CHEN M., GUIBAS L.: Soft maps between surfaces. In *CGF* (2012), vol. 31, pp. 1617–1626. 1, 2, 3, 4, 5, 7, 9

[SOG09] SUN J., OVSJANIKOV M., GUIBAS L.: A concise and provably informative multi-scale signature based on heat diffusion. In *SGP* (2009), pp. 1383–1392. 7

[TBW*11] TEVS A., BERNER A., WAND M., IHRKE I., SEIDEL H.-P.: Intrinsic shape matching by planned landmark sampling. *CGF 30*, 2 (2011), 543–552. 2

[Vil03] VILLANI C.: *Topics in optimal transportation*, vol. 58. American Mathematical Soc., 2003. 4, 10

[WL78] WITTE F. P., LUCAS D.: Probabilistic tracking in a multitarget environment. In *IEEE Conf. on Decision and Control* (1978), vol. 17, pp. 1212–1216. 2

[WSSC11] WINDHEUSER T., SCHLICKWEI U., SCHIMDT F. R., CREMERS D.: Large-scale integer linear programming for orientation preserving 3d shape matching. *CGF 30*, 5 (2011), 1471–1480. 2

[YWL*07] YAN W., WANG Q., LIU Q., LU H., MA S.: Topology-preserved diffusion distance for histogram comparison. In *British Machine Vision Conf.* (2007). 3

[ZSCO*08] ZHANG H., SHEFFER A., COHEN-OR D., ZHOU Q., VAN KAICK O., TAGLIASACCHI A.: Deformation-driven shape correspondence. In *SGP* (2008), pp. 1431–1439. 2

Appendix A: Proofs

Proof of Prop. 2. If μ is a soft map from M_0 to M with constant density then $\mu_x = \mu_{x'}$ for all $x, x' \in M_0$. Thus $\mathcal{W}_2(\mu_x, \mu_{x'}) = 0$ because the product distribution $\pi := \mu_x \otimes \mu_{x'}$ is a measure coupling with zero cost. Next, if $\phi : M_0 \to M$ is a map, then the associated soft map is $\delta_{\phi(x)}(y)\,dy$ where δ_p is the Dirac δ-density centered at p. There is only on measure coupling of $\delta_{\phi(x)}(y)\,dy$ and $\delta_{\phi(x')}(y)\,dy$, namely the product distribution for which the cost is $\mathrm{dist}(\phi(x), \phi(x'))$. The Dirichlet energy density now reduces to the conventional Dirichlet energy density in the limit as $r \to 0$.

Proof of Prop. 3. Let μ_1, μ_2 be soft maps, π_k for $k = 1,2$ be optimal measure couplings of $[\mu_k]_x$, $[\mu_k]_{x'}$, and $\alpha \in [0,1]$. Then $\pi_\alpha := (1-\alpha)\pi_1 + \alpha\pi_2$ is a measure coupling of $[(1-\alpha)\mu_1 + \alpha\mu_2]_x$ and $[(1-\alpha)\mu_1 + \alpha\mu_2]_{x'}$. Since $\iint_{M \times M} [\mathrm{dist}(y, \bar{y})]^2 d\pi_\alpha(y, \bar{y}) = (1-\alpha) W_2^2([\mu_1]_x, [\mu_1]_{x'}) + \alpha W_2^2([\mu_2]_x, [\mu_2]_{x'})$, the proposition follows.

Proof of Prop. 4. Assume first that $\rho > 0$. We recall from the theory of optimal transportation that the solution of the optimal transportation problem for the 2-Wasserstein distance on a compact surface can be characterized as follows. The transport between two measures with positive density μ_1 and μ_2 is achieved by a map $\psi : M \to M$ of the form $\psi(y) := \exp_y(\nabla q(y))$ where $q : M \to \mathbb{R}$ is a convex function and $\exp_y : T_y M \to M$ is the geodesic exponential map. Moreover, the transportation cost can be expressed as $[\mathcal{W}_2(\mu_1, \mu_2)]^2 = \int_M \|\nabla q\|^2 d\mu_1$. See [Vil03, Ch. 2] for details. We can use these ideas to simplify the Dirichlet density (2) by setting $\mu_1 := \mu_x$ and $\mu_2 := \mu_{x'}$, yielding $q := q_{x,x'}$ that achieves the transport from μ_x to $\mu_{x'}$. When x and x' are sufficiently close, we can write $x' = \exp_x(\varepsilon V)$ where $\varepsilon = \mathrm{dist}_0(x, x')$ and $V \in T_x M_0$. To first order, $q_{x,x'}(y) \approx \varepsilon Q(x, V, y)$ where $y \mapsto Q(x, V, y)$ is a function on M. Also Q is linear V. Substituting q into the expression for the cost given above and differentiating in ε leads to the PDE (6) for each $x \in M_0$ and $V \in T_x M_0$. Finally, taking the limit as $\varepsilon \to 0$ in the expression for the Dirichlet energy density (2) yields the desired simplification. Limiting arguments can be made to handle the cases of densities which fail to be everywhere non-zero and weak limits of smooth densities.

Proof of Prop. 6. Let $d\mu_x(y) := \delta_{\phi(x)}(y)\,dy$ and perform a simple change of variables in the inner integral of (7).

Proof of Prop. 7. Write $C_\varepsilon := C + \varepsilon X$, where X has coefficients X_{ij} satisfying $\sum_j X_{ij} = 0$ for all i. Also let $q^s_{f,\varepsilon} := \varepsilon Y^s_f + \mathcal{O}(\varepsilon^2)$. Differentiating (10) with respect to ε yields an expression for Y in terms of C, q_f and X, namely

$$\sum_{ijk} Y^s_{fk} C_{ij} W_{jkk'} B^+_{0if} = \sum_{ij} X_{ij} \left(U_{jk'} B^-_{0ifs} - \sum_k q^s_{fk} W_{jkk'} B^+_{0if} \right)$$

for all $f \in F_0$ and $s = 1,2$. If we differentiate (13), we find that $\frac{d}{d\varepsilon} \mathcal{E}_D(C_\varepsilon)\big|_{\varepsilon=0}$ is a sum of two terms, one containing X and the other containing Y. Eliminating Y using this expression results the expression for the gradient of $\mathcal{E}_D$ at C.