L^1TV computes the flat norm for boundaries

Simon P. Morgan* and Kevin R. Vixie**

*University of Minnesota and **Los Alamos National Laboratory

Abstract

We show that the recently introduced L^1TV functional can be used to explicitly compute the flat norm for co-dimension one boundaries. While this observation alone is very useful, other important implications for image analysis and shape statistics include a method for denoising sets which are not boundaries or which have higher co-dimension and the fact that using the flat norm to compute distances not only gives a distance, but also an informative decomposition of the distance. This decomposition is made to depend on scale using the *flat norm with scale* which we define in direct analogy to the L^1TV functional. We illustrate the results and implications with examples and figures.

1 Introduction

In this research announcement, we point out that the L^1TV functional, introduced and studied in the continuous setting in [3, 12, 1, 13] and earlier in the discrete setting in [2, 9], provides a beautiful way of computing both the value of, and the optimal decomposition required by, the *flat norm* from geometric measure theory.

The L^1TV functional was introduced as an improvement to the now classic Rudin-Osher-Fatemi total variation based denoising. Among its many nice properties are the way it handles binary images, making it useful for shape processing. Theoretically speaking, the clean geometric structure of the functional and its minimizers is very attractive. See [12, 1] for details.

Joan Glaunès was, as far as we know, the first to suggest and use the *flat norm* from geometric measure theory as a distance in shape space. In his dissertation [6] and a couple of application papers [11, 5] with collaborators, the dual formulation of the flat norm is used to compute distances between shapes.

For maximum usefulness to those without much background in geometric measure theory, we will include expository examples and a bit more explanation than we would for those with that background. In several longer papers with co-authors Allard, Esedoglu and Yin, we will explore the full implications and technical details. We end this paper with an outline of some of those directions.

2 L^1TV gives the Flat Norm

The L^1TV functional introduced and studied in [3] is given by:

$$F(u) = \int |\nabla u| dx + \lambda \int |u - u_0| dx \tag{1}$$

where u and u_0 are functions from $\mathbb{R}^n$ (often $\mathbb{R}^2$) to $\mathbb{R}$ with u_0 being the input or measured function (e.g. image intensity). The optimal u minimizing (1) can be thought of as a "denoised" version of u_0. Typically, one chooses the parameter λ based on noise levels since this choice effectively chooses the scale below which features or oscillations are ignored. In the case that the input function is binary, Chan and Esedoglu observe that the functional reduces to:

$$F^{\lambda}_{CE}(\Sigma) = \text{Per}(\Sigma) + \lambda |\Sigma \triangle \Omega|. \tag{2}$$

where $\triangle$ denotes the symmetric difference. Now, let $\Sigma(\Omega, \lambda)$ be a binary minimizer of (2). I.e.

$$\Sigma(\Omega, \lambda) \equiv \text{argmin}\, F^{\lambda}_{CE}(\Sigma) = \text{Per}(\Sigma) + \lambda |\Sigma \triangle \Omega|. \tag{3}$$

For our convenience, we record the optimal decomposition of Ω into $\{\Sigma(\Omega, \lambda)$ and $(\Sigma(\Omega, \lambda) \triangle \Omega)\}$ as the pair $\{\partial\Omega, \Sigma(\Omega, \lambda) \triangle \Omega\}$.

The *flat norm* of an n-current T, denoted by $\mathbb{F}(T)$, is given by

$$\mathbb{F}(T) \equiv \min_S \{\mathbf{M}(S) + \mathbf{M}(T - \partial S)\}) \tag{4}$$

where S varies over a suitable set of $n+1$-currents and $\mathbf{M}$ is the *mass* of the indicated currents. We refer to $\{T, S\}$ as the flat norm induced, optimal decomposition. We will denote the current carried by the manifold or rectifiable set E, by T_E or S_E. (One can informally think of *currents* as oriented submanifolds and *mass* as n-volume, but for those without experience with currents, we suggest focusing the examples in the next section.) Now for results.

Theorem 1. *For the current $T_{\partial\Omega}$,*

$$\mathbb{F}(T_{\partial\Omega}) = F^1_{CE}(\Sigma(\Omega, 1)) \tag{5}$$

and

$$\{T_{\partial\Omega}, S_{\Sigma(\Omega,1)\triangle\Omega}\} \tag{6}$$

is the optimal decomposition that the flat norm requires.

This immediately suggests a very useful generalization of the flat norm which we call the *flat norm with scale*.

Definition 1.

$$\mathbb{F}_\lambda(T) \equiv \min_S\{\lambda\mathbf{M}(S) + \mathbf{M}(T - \partial S)\}) \tag{7}$$

It is known that the λ controls very precisely the maximum curvature of $T - \partial S$. We get the obvious scaled version of the first result:

Theorem 2. *For the current $T_{\partial\Omega}$,*

$$\mathbb{F}_\lambda(T_{\partial\Omega}) = F^\lambda_{CE}(\Sigma(\Omega, \lambda)) \tag{8}$$

and

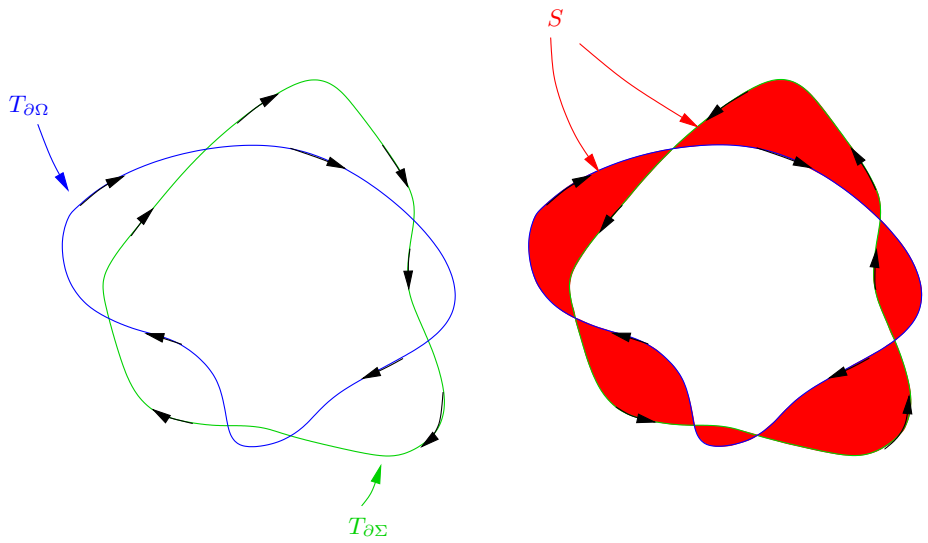
$$\{T_{\partial\Omega}, S_{\Sigma(\Omega,\lambda)\triangle\Omega}\} \tag{9}$$

is the optimal decomposition that the flat norm with scale requires.

Proof. The proof is really simply checking that the picture one can draw holds after the definitions of *mass* ($\mathbf{M}$) and *perimeter* (Per) are used to translate the picture into analytic terms. Very briefly, we have:

$$\begin{aligned} F^\lambda_{CE}(\Sigma) &= \lambda|\Sigma \triangle \Omega| + \mathrm{Per}(\Sigma) \\ &= \lambda M(S_{\Sigma\triangle\Omega}) + M(T_{\partial\Sigma}) \\ &= \lambda M(S_{\Sigma\triangle\Omega}) + M(T_{\partial\Omega} - \partial S_{\Sigma\triangle\Omega}) \end{aligned}$$

Figure 1 illustrates this pictorially. □



$$T_{\partial\Sigma} = T_{\partial\Omega} \text{ - } \partial S_{\Omega\triangle\Sigma}$$

$$\begin{aligned} F^{\lambda}_{CE}(\Sigma) &= \lambda|\Sigma \triangle \Omega| + \mathrm{Per}(\Sigma) \\ &= \lambda M(S_{\Sigma\triangle\Omega}) + M(T_{\partial\Sigma}) \\ &= \lambda M(S_{\Sigma\triangle\Omega}) + M(T_{\partial\Omega} - \partial S_{\Sigma\triangle\Omega}) \end{aligned}$$

Figure 1: In this figure we illustrate the translation of the L^1TV view to the flat norm view. The perimeter of Σ becomes the mass of $T_{\partial\Omega} - \partial S_{\Sigma\triangle\Omega}$ and the volume of $\Sigma \triangle \Omega$ becomes the mass of $S_{\Sigma\triangle\Omega}$. Note: this figure does not depict a minimizer. Rather, we depict Ω and any candidate Σ.

Our final observation is that the *flat norm with scale* gives us the same decomposition as we would get if we first scaled T, computed the decomposition and then reversed the scaling. More precisely

Lemma 1. *Denote the optimal $\mathbb{F}_\lambda$ decomposition by $\{T, S\}_\lambda$. Then*

$$\{T, S\}_\lambda = d_{\frac{1}{\lambda}\#}\{d_{\lambda\#}(T), d_{\lambda\#}(S)\} \tag{10}$$

where d_λ denotes the λ-dilation of $\mathbb{R}^m$, and $d_{\lambda\#}(T_M) = T_{d_\lambda(M)}$.

Proof. The minimizing decomposition $\{T, S\}_\lambda$ which minimizes $\lambda\mathbf{M}(S) + \mathbf{M}(T - \partial S)$ also minimizes $\lambda^n\mathbf{M}(S) + \lambda^{n-1}\mathbf{M}(T - \partial S) = \mathbf{M}(d_{\lambda\#}S) + \mathbf{M}(d_{\lambda\#}(T - \partial S))$. Therefore, to get the minimizer for $\mathbb{F}_\lambda$ run the optimization required by $\mathbb{F}_1$ using $d_{\lambda\#}(T)$ as input, and then contract with $d_{\frac{1}{\lambda}\#}$. □

We now discuss applications and examples with pictures which should clarify things for those with less exposure to geometric measure theory.

3 Micro-tutorial on Currents and the Flat Norm

If you know something about currents and have a clear picture of the flat norm, this section can be skipped. The reference for this section is Frank Morgan's nice introduction [8]. The definitive, though formidable, treatise for a fair bit of geometric measure theory is still Federer's 1969 tome [4]. References between Morgan and Federer include [7, 10].

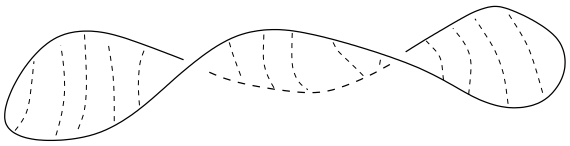
Rectifiable sets An n-rectifiable subset N of $\mathbb{R}^m$ is the union of 1) an H^n negligible set and 2) a countable collection of subsets of C^1 n-submanifolds of $\mathbb{R}^m$ such that $H^n(N) < \infty$. (H^n is the n-dimensional Hausdorff measure.) Intuitively, an n-rectifiable set looks a great deal like an n-manifold at most of its points.

Currents n-Currents in $\mathbb{R}^m$, denoted $\mathcal{D}_n$, are the duals to $\mathcal{D}^n$, the smooth, compactly supported n-forms on $\mathbb{R}^m$. We restrict ourselves to integer multiplicity rectifiable currents T, which have the following representation: $T(\phi) = \int_N \Theta(x)\,\langle\phi(x), \xi(x)\rangle\, d\mathcal{H}^n, \forall\phi \in \mathcal{D}^n$ where N is an n-rectifiable set in $\mathbb{R}^m$, $\Theta(x)$ is an integer multiplicity density function, always ± 1 in this paper, ϕ is the

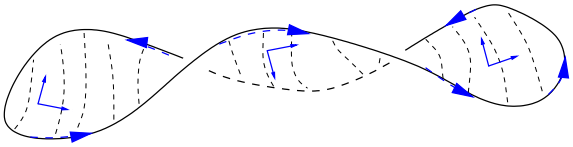
form T is operating on, and $\xi(x)$ is the n-vector defining the orientation on N. Changing the sign of the density function has the effect of reversing the orientation on N which can also be achieved by replacing ξ with $-\xi$.

Currents are naturally equipped with a boundary operator, $\partial T(\phi) \equiv T(d\phi)$. ∂T is therefore an $(n-1)$-current which operates on $(n-1)$-forms. (Note the consistency of this definition with Stokes' theorem.)

Intuitively, a current is an oriented manifold or union of oriented manifolds, each with a oriented boundary whose orientation is inherited from the orientation of the manifold. (There are of course wild beasts in the current menagerie, but one can go a long ways with unions of C^1 manifolds with boundary.) See Figure 2.



Manifold M, with boundary ∂M



The Current T_M showing orientation of T_M and ∂T_M

Figure 2: Example 2-current T_M. Notice that $\partial T_M = T_{\partial M}$. The orientation on the boundary ∂M is simply that inherited from the orientation on M.

Mass and the Flat Norm The *mass* of a current is defined as

$$M(T) = \sup_{|\phi| \leq 1, \phi \in \mathcal{D}^n} T(\phi). \tag{11}$$

Informally, the *mass* is simply the n-dimensional volume of the rectifiable set carrying the n-current. The *flat norm* is defined in two equivalent ways:

$$\mathbb{F}(T) = \min_{S \in \mathcal{D}_{n+1}} (M(S) + M(T - \partial S)) \text{ (given above)} \tag{12}$$

and

$$\mathbb{F}(T) = \sup_{|\phi| \leq 1, |d\phi| \leq 1, \phi \in \mathcal{D}^n} (T(\phi)) \text{ (mentioned above)} \tag{13}$$

(This equivalence is proved in Federer (4.1.12 in [4]) where it is called the flat seminorm.) The corresponding dual definition of the flat norm with scale is given by

$$\mathbb{F}_\lambda(T) = \sup_{|\phi| \leq 1, |d\phi| \leq \lambda, \phi \in \mathcal{D}^n} (T(\phi)) \tag{14}$$

Examples of the flat norm decomposition The flat norm involves the optimal decomposition of the n-current T into an n-current $(T - \partial S)$ and an $(n+1)$-current S. (We use the term *decomposition* in reference to the fact that $T - \partial S$ and S are the components explicitly measured by the flat norm, even though $T = (T - \partial S) + \partial S$ rather than $T = (T - \partial S) + S$.) See Figure 3.

Examples of maximizing forms The dual formulation of the flat norm involves finding the supremum over appropriately constrained forms. Figure 4 shows a maximizing form for the 2-dimensional disk of radius r. We will discuss the computation of optimizing forms and the extraction of S from those optimizing forms in section 4.3

Relation of the flat norm to L^1 The flat norm is very much like the L^1 norm for small differences between currents. The value of the flat norm of a difference between two currents ends up (roughly) being the L^1 difference between the close parts plus the sum of the n-volumes of what is left. See Figure 5.

4 Applications and Illustrating Examples

The value of the above results is fully realized by exploring their use in applications to images and shapes. As observed by Glaunès and others, the flat norm is a very natural candidate for distances between shapes. We now very briefly explore and illustrate applications of the above observations.

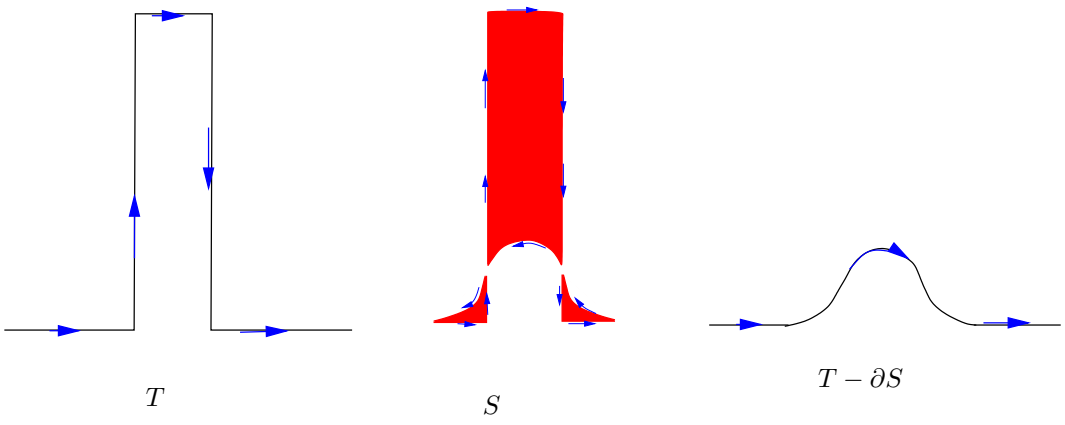


Figure 3: Example flat norm decomposition. T is the 1-current we are computing the flat norm of, and S gives the optimal decomposition into S and $T - \partial S$. Finally, the flat norm is simply $M(S) + M(T - \partial S) =$ length of the right-most curve and area of the red region.

4.1 Generalized Flat Norm: flat norm with scale

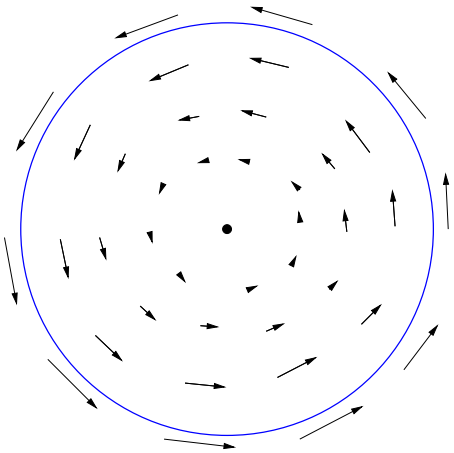
As mentioned in section 2, by letting the $\lambda \neq 1$, one has a natural way to vary the intrinsic scale in the flat norm.

$$\begin{aligned} \mathbb{F}_\lambda(T) &\equiv \min_S \{\lambda \mathbf{M}(S) + \mathbf{M}(T - \partial S)\} \\ &= F^\lambda_{CE}(E) \end{aligned}$$

when $T = \partial E$. The main point here is that this is easy to compute, given the connection to the L^1TV functional. Varying λ gives us the ability to choose what scale is big and worth keeping. See Figure 6.

4.2 Flat norm via L^1TV

We can use the L^1TV functional to very easily calculate *both* the flat norm of differences between surfaces which are boundaries *and* the optimal decomposition of that difference into surface and area parts. (See for example T_1 and T_2 and the optimal



Maximizing form ϕ for disk of radius r

$$\phi = \frac{x}{r}dy - \frac{y}{r}dx$$

Figure 4: A maximizing form for the disk in 2-D. This form satisfies the constraints as long as $\frac{2}{r} \leq \lambda$. The λ is, of course, the scale in the *flat norm with scale* introduced above.

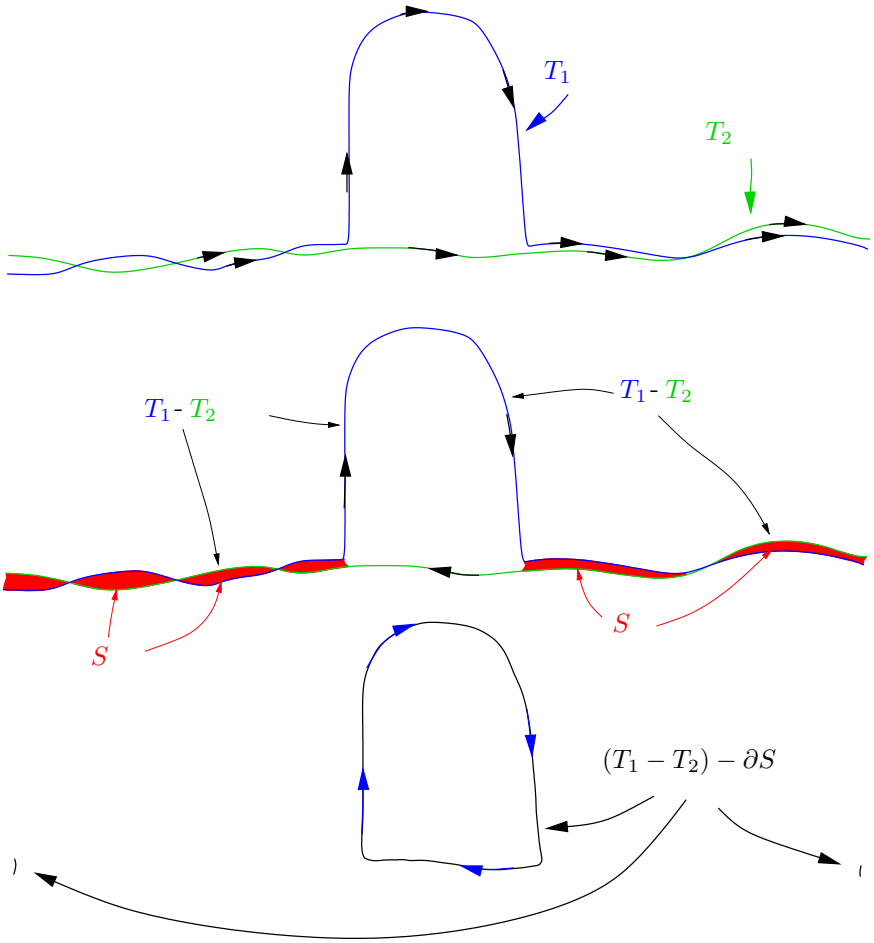


Figure 5: The flat norm of the difference between these two currents is the sum of the area of red region (L^1 like) and the length of the loop that is left.

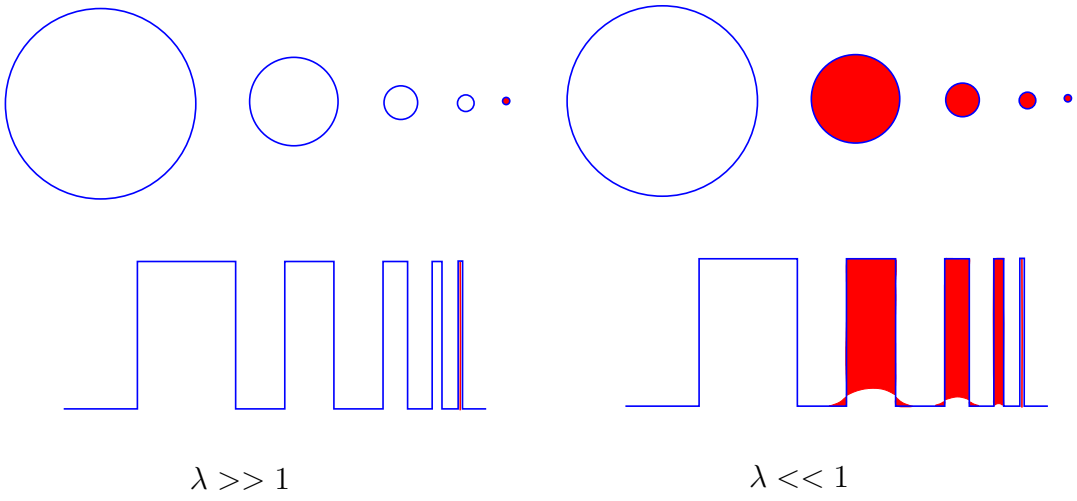


Figure 6: Flat norm decomposition with scale λ:

S in Figure 12). In Figure 7 the decomposition of a boundary current $T_{\partial\Omega}$ into the diminished boundary $T_{\partial\Sigma}$ and the symmetric difference current $S_{\Sigma\Delta\Omega}$ is illustrated.

4.3 L^1TV by the dual form of the flat norm

The dual formulation of the flat norm can be used to compute L^1TV minimizers. The following results establish that maximizing forms (or maximizing sequences of forms) contain the decomposition into S and $T - \partial S$.

Proposition 1. *Suppose that* $T(\phi) = F(T) = \min_S(M(S) + M(T - \partial S))$, *where* ϕ *is a smooth, compactly supported* n*-form satisfying* $|\phi| \leq 1, |d\phi| \leq 1$, *and* T *is an* n*-rectifiable current with density* $\Theta = 1$ $\mathcal{H}^n$ *almost everywhere, then on the support of* S, $|d\phi| = 1$, *and on the support of* $T - \partial S$, $|\phi| = 1$.

Proof. For a minimizing choice of S, we have that

$$\begin{aligned} T(\phi) &= F(T) = M(S) + M(T - \partial S) \\ T(\phi) &= \partial S(\phi) + (T - \partial S)(\phi) \end{aligned}$$

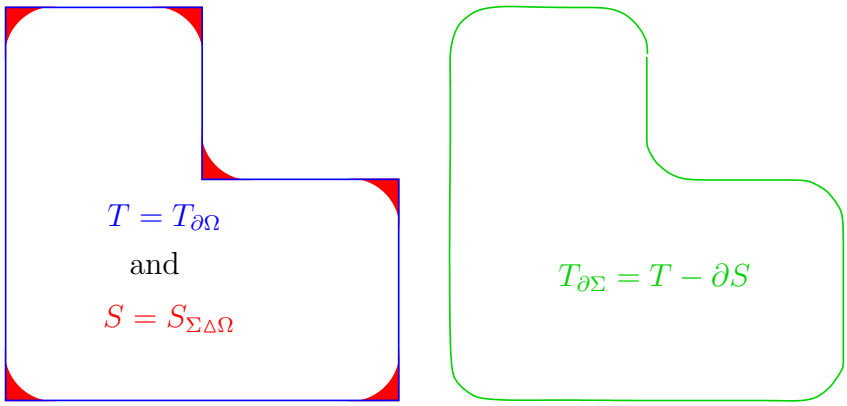


Figure 7: A simple reinterpretation of the L^1TV input and minimizer gives us the flat norm of $T_{\partial\Omega}$ and the decomposition into the diminished boundary $T_{\partial\Sigma}$ and the symmetric difference current $S_{\Sigma\Delta\Omega}$.

and

$$\partial S(\phi) = S(d\phi)$$

so that

$$S(d\phi) + (T - \partial S)(\phi) = M(S) + M(T - \partial S) \tag{15}$$

Now since $|\phi| \leq 1$ and $|d\phi| \leq 1$, we immediately get that in fact $|\phi| = 1$ H^n almost everywhere on $T - \partial S$ and $|d\phi| = 1$ H^{n+1} almost everywhere on S. □

See Figure 8 for an example maximizing form. Note that when S is not unique, this proposition implies that $|d\phi| = 1$ on the union of all possible minimizing S's. If we do not have a maximizing form, we can use the following modified proposition together with the fact that there will be sequence of forms ϕ_i such that $T(\phi_i) + \epsilon_i = F(T)$, $\epsilon_i > 0$, and $\epsilon_i \to 0$:

Proposition 2. *suppose that* $T(\phi) + \varepsilon = F(T) = \min_S(M(S) + M(T - \partial S))$, *where* ϕ *is a smooth, compactly supported* n*-form satisfying* $|\phi| \leq 1, |d\phi| \leq 1$, *and* T *is an* n*-rectifiable current with density* $\Theta = 1$ $\mathcal{H}^n$ *almost everywhere, then* $M(S) - \epsilon \leq S(d\phi) \leq M(S)$ *and* $M(T - \partial S) - \epsilon \leq (T - \partial S)(\phi) \leq M(T - \partial S)$.

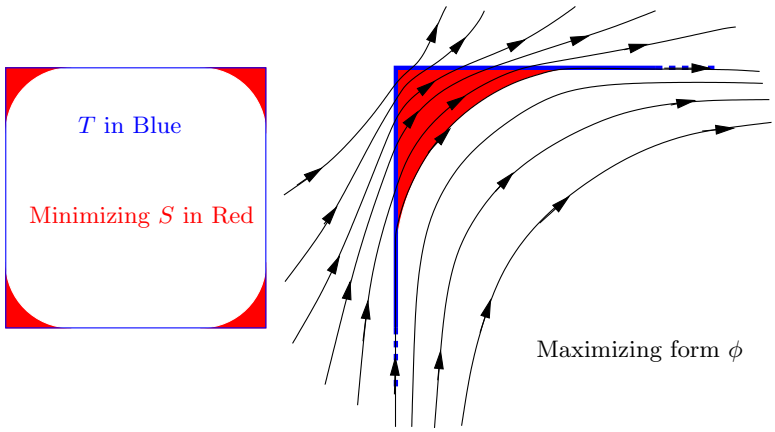


Figure 8: Example of a maximizing form for a square T as input. On the right hand side of the figure, details of the form (which we visualize as a vector field) are shown. On $T - \partial S$, $|\phi| = 1$ everywhere and on S (red), $|d\phi| = 1$ everywhere, while off of these sets, both $|\phi| < 1$ and $|d\phi| < 1$.

This modified proposition is necessary since there are easily constructed examples having no smooth maximizing form. In fact, the example shown in Figure 8 is not actually smooth. The optimizing $d\phi$ we show is actually Lipschitz, so it can be arbitrarily well approximated by smooth forms even though it is not itself smooth. The non-smoothness originates at the points of $T - \partial S$ where the circular arcs (tangentially) join the sides of the square. At these points, the boundary is merely $C^{1,1}$.

We now state a conjecture, the confirmation of which will be very useful for computations.

Conjecture 1. *Suppose Φ is the collection of all Lipschitz forms ϕ maximizing $T(\phi)$ and satisfying $|\phi| \leq 1$ and $|d\phi| \leq 1$. Define X to be the closure of the set on which $|d\phi| = 1$ for every $\phi \in \Phi$. Finally, define $\mathbb{S}$ to be the collection of all optimizing currents S such that $\mathbb{F}(T) = M(S) + M(T - \partial S)$. Then*

I $\Phi \neq \emptyset$

II $X = \cup_{S \in \mathbb{S}} S$

III *When a point $t \in T$ is in the interior of X, then for each $\phi \in \Phi$ the orientation of $d\phi$ will switch sign on the set underlying T at t.*

Part I says that there are always maximizing forms if we permit them to be merely Lipschitz instead of smooth. Part II enables us to see where the possible locations for an S might be, and part III helps us decompose these possible places into individual candidates for S.

A very simple example where Lipschitz forms are necessary and sufficient for optimality is the case in which the current is three equally spaced circles on a torus. See Figure 9. Note that the metric on the torus is chosen such that the circles are parallel. Figure 10 shows the f of a Lipschitz maximizing form $f(x)dy$. In the case that $b = 2a$ and $a < 2$, a little bit of calculation and thought shows that we can't maximize with a smooth form and that a maximizing sequence must approach the form plotted in Figure 10 with $b = 2a$ and $\alpha = 1$. Finally, the above propositions, example and conjecture have obvious analogs for the flat norm with scale.

4.4 L^1TV for co-dimension > 1

We can use the dual formulation of the flat norm to extend the L^1TV denoising to sets which are not boundaries or have co-dimension greater than 1. This depends on

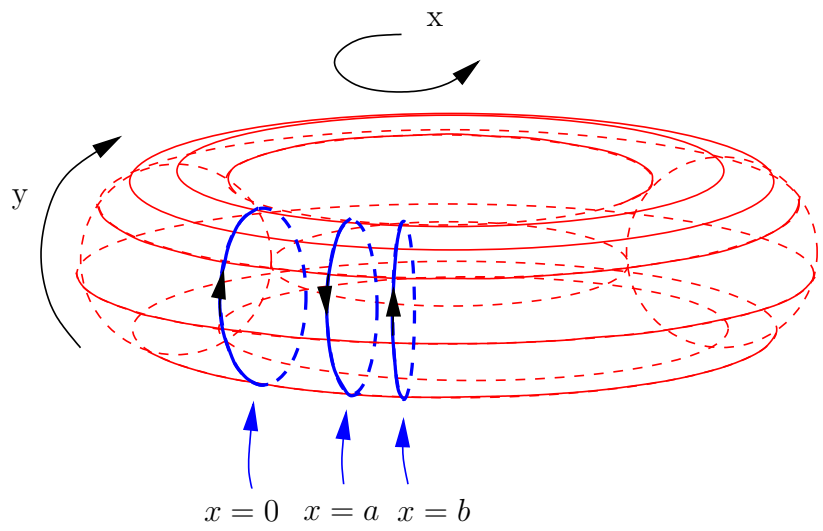


Figure 9: The 3 circles example. The metric is chosen so that the circles are parallel to each other.

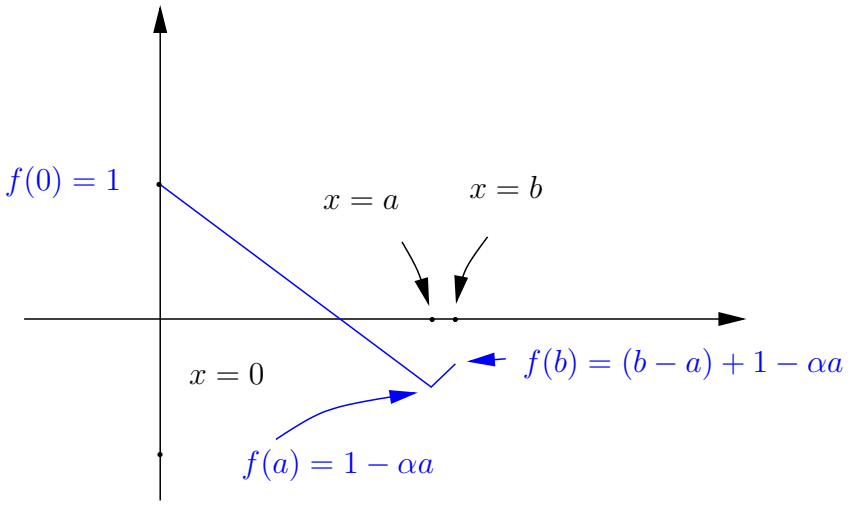


Figure 10: A maximizing form for the 3 circles example. We plot f for the form given by $f(x)dy$. In the case that $b = 2a$ and $a < 2$, we are forced to use $\alpha = 1$ and the Lipschitz form plotted here.

extracting the optimal decomposition from the optimizing form, as discussed in the previous subsection. Figure 11 shows the decomposition of a 1-current in 3D that results when the flat norm is computed. This example actually illustrates both of the useful generalizations possessed by the flat norm decomposition: regularization or denoising of higher co-dimension and non-boundary subsets. Notice that the use of the flat norm with scale permits the choice of what scale is small (and hence greatly diminished) and what scales are large and therefore preserved. In the case of sets which are co-dimension 1 boundaries, we know that in a very precise sense, the regularized surface $(T - \partial S)$ is the best λ-curvature approximation to T. See [12, 1] for details.

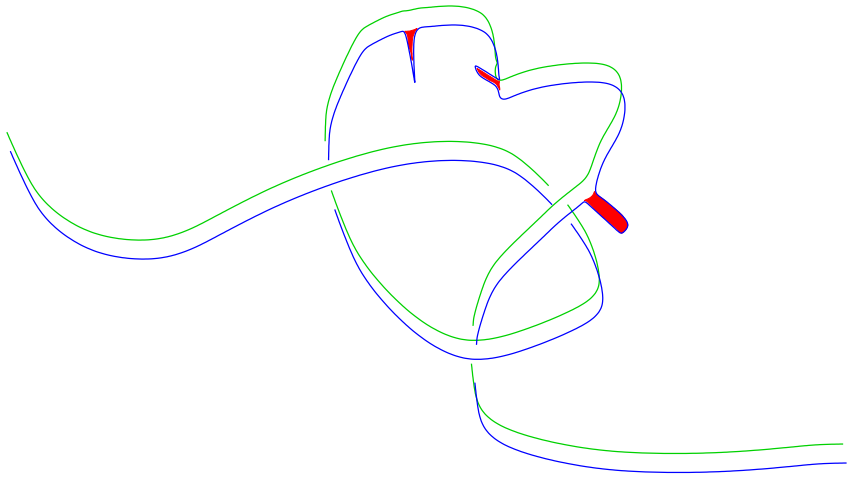


Figure 11: The green curve is the denoised version of the blue, where we have translated the green to make visualization easier.

4.5 Shape Statistics

We can use the L^1TV algorithms to compute both the flat norm distance in shape space and the flat norm decomposition that gives this distance. (As noted above, the flat norm was previously suggested for shape comparisons in [6, 11] and then used in [5] for the purpose of computing shape statistics.) Our observation permits the

use of the various algorithmic approaches to TV minimization for the calculation of both the flat norm and the minimizers when those are of interest.

In particular, we can immediately and easily compute the flat norm difference $T_1 - T_2$ between any two shapes $T_1 = \partial\Omega_1$ and $T_2 = \partial\Omega_2$. The decomposition that we get for free, so to speak, shows us where the difference is big with respect to λ and where it is small. See Figure 12.

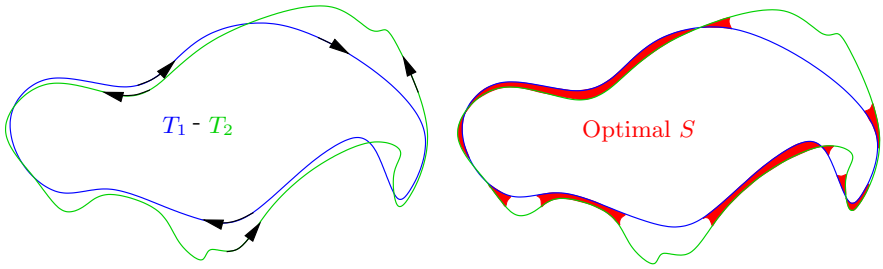


Figure 12: The flat norm via the L^1TV functional provides us with both a distance and an informative optimal decomposition into S and $T_1 - T_2 - \partial S$.

5 Summary

The program suggested by the relatively simple observation of the relation between L^1TV and the flat norm promises many new benefits. In order to reap the full benefits, there are challenges to be met that include efficient computation of the optimizing n-forms (the dual formulation of the flat norm), extraction of the optimal decomposition from the optimizing differential form (required for the generalization of L^1TV), and calculation of the gradient of $F^{\lambda}_{CE}(\Sigma(\Omega \triangle \Omega^*, \lambda))$ with respect to Ω (certain shape statistics approaches require this). These, as well as the general expansion of the above announcement (there are technical details!) is the subject of several papers that are in preparation or in planning with collaborators Allard, Esedoḡlu, and Yin.

References

[1] William K. Allard. On the regularity and curvature properties of level sets of minimizers for denoising models using total variation regularization; I. Theory. Preprint, 2006.

[2] S. Alliney. A property of the minimum vectors of a regularizing functional defined by means of the absolute norm. *IEEE Trans. Signal Process.*, 45:913–917, 1997.

[3] Tony F. Chan and Selim Esedoḡlu. Aspects of total variation regularized L^1 function approximation. *SIAM J. Appl. Math.*, 65(5):1817–1837, 2005.

[4] Herbert Federer. *Geometric Measure Theory*. Classics in Mathematics. Springer-Verlag, 1969.

[5] J. Glaunès and S. Joshi. Template estimation form unlabeled point set data and surfaces for computational anatomy. *JMIV*, 2007. Proceedings of Mathematical Foundations of Computational Anatomy, MICCAI 2006.

[6] Joan Glaunès. *Transport par difféomorphismes de points, de mesures et de courants pour la comparaison de formes et l' anatomie numérique.* PhD thesis, l' Université Paris 13 en Mathématiques, 2005.

[7] Fanghua Lin and Xiaoping Yang. *Geometric Measure Theory – an Introduction*. Science Press and International Press, 2002.

[8] Frank Morgan. *Geometric Measure Theory: A Beginner's Guide*. Academic Press, 3rd edition, 2000.

[9] Mila Nikolova. Minimizers of cost-functions involving nonsmooth data-fidelity terms. *SIAM J. Numer. Anal.*, 40:965–994, 2003.

[10] Leon Simon. *Lectures on Geometric Theory*. Centre for Mathematical Analysis, Australian National University, 1984. ISBN 0-86784-429-9.

[11] Marc Vaillant and Joan Glaunès. Surface matching via currents. In *Proceedings of Information Processing in Medical Imaging (IPMI 2005)*, volume 3565 of *Lecture Notes in Computer Science*. Springer, 2005.

[12] Kevin R. Vixie and Selim Esedoḡlu. Some properties of minimizers for the L^1TV functional. *Inverse Problems*, 2006. submitted.

[13] Wotao Yin, Donald Goldfarb, and Stanley Osher. Image cartoon-texture decomposition and feature selection using the total variation regularized L^1 functional. *Submitted to SIAM MMS*, 2006. UCLA CAM Tech Report 05-47.