CIRM

Courbure discrète : théorie et applications

Rencontre organisée par :
Laurent Najman and Pascal Romon

18-22 novembre 2013

Joseph H.G. Fu

Piecewise linear approximation of smooth functions of two variables

Vol. 3, n^{o} 1 (2013), p. 11-16.

<http://acirm.cedram.org/item?id=ACIRM_2013__3_1_11_0>

Centre international de rencontres mathématiques
U.M.S. 822 C.N.R.S./S.M.F.
Luminy (Marseille) France

Actes des rencontres du C.I.R.M.
Vol. **3** n^{o} 1 (2013) 11-16

Piecewise linear approximation of smooth functions of two variables

Joseph H.G. Fu

Abstract

The normal cycle of a singular subset X of a smooth manifold is a basic tool for understanding and computing the curvature of X. If X is replaced by a singular function on $\mathbb{R}^n$ then there is a natural companion notion called the *gradient cycle* of f, which has been introduced by the author and by R. Jerrard. We discuss a few fundamental facts and open problems about functions f that admit gradient cycles, with particular attention to the first nontrivial dimension $n = 2$.

1. Introduction

The Federer-Fleming theory of integral currents (developed in detail in Chapter 4 of [6]) is a mathematical tool designed to extend certain notions of differential geometry to spaces with singularities. Typically it is used to study first order problems in the calculus of variations such as the Plateau problem. However, it also works spectacularly well in the study of curvature for subspaces with singularities, providing the natural setting for Federer's theory of curvature measures and its extensions [5, 16, 9]. The key idea here is that of the *normal cycle* $N(X)$ of a singular subspace X embedded in a smooth manifold M. The normal cycle is an integral current living in the tangent sphere bundle of M that functions as a substitute for the manifold of unit normals of a smooth submanifold. It has been applied effectively in surface modeling, particularly in the problem of approximating a given surface, given either formally as a smooth submanifold or empirically in terms of collections of data points, by a polyhedron [4].

Despite its many advantages, the natural scope of the theory remains murky in the sense that a clear geometric characterization of the class of sets X admitting a normal cycle is unknown. This general problem is essentially analytic. In order to study it without getting distracted by secondary topological questions, it is convenient to consider a closely related problem in which the singular subset X is replaced by a singular *function* $f : \mathbb{R}^n \to \mathbb{R}$. In this case the normal cycle $N(X)$ is replaced by the *gradient cycle* $\mathbb{D}(f)$, an integral current living in the cotangent bundle $T^*\mathbb{R}^n \simeq \mathbb{R}^n \times \mathbb{R}^{n*}$ that serves a substitute for the graph of the gradient of f (or, more correctly, its differential, although we will conflate the two in the present note). If $\mathbb{D}(f)$ exists then f is said to be a *Monge-Ampère function.* This class has been studied by the present author and his collaborator Ryan Scott [8, 10, 11] as well as by R. Jerrard [12, 13]. We describe here some basic issues and progress in the subject, with particular attention to the case $n = 2$.

2. The normal cycle and the gradient cycle

For simplicity let us take the ambient smooth manifold M to be $\mathbb{R}^n$, and assume that $X \subset \mathbb{R}^n$ is compact. The normal cycle $N(X)$ is an integral current of dimension $n-1$ living in $\mathbb{R}^n \times S^{n-1}$

Text presented during the meeting "Discrete curvature: Theory and applications" organized by Laurent Najman and Pascal Romon. 18-22 novembre 2013, C.I.R.M. (Luminy).
2000 *Mathematics Subject Classification.* 00X99.
Key words. graph theory, shape recognition, optimal transportation.
Funded by NSF Grant .

satisfying a few inevitable conditions. Let $x_1, \dots, x_n$ be standard coordinates for $\mathbb{R}^n$ and $y_1, \dots, y_n$ the companion coordinates for the $\mathbb{R}^{n*} \simeq \mathbb{R}^n$ that contains the sphere S^{n-1}. Then

- the support of $N(X)$ must be compact.
- $N(X)$ has boundary zero in the sense of Stokes' theorem, i.e. evaluation of $N(X)$ against any exact $(n-1)$-form vanishes.
- $N(X)$ is *Legendrian*, i.e. evaluation of $N(X)$ against any $(n-1)$-form expressible as a wedge product with the canonical 1-form $\alpha = \sum y_i \, dx_i$ vanishes.
- Finally, $N(X)$ yields the expected Morse theory of height functions restricted to X.

The precise form of the last condition is somewhat awkward, so we refrain from stating it here. The upshot is that these four conditions are enough to determine $N(X)$ uniquely. For truly pathological subsets X this current will not exist at all. It only exists for certain "tame" subsets X, but when it does exist it is defined unambiguously.

2.1. **Monge-Ampère functions.** The companion theory for singular functions may be described in analogous terms, with the advantage that the last condition is easier to understand. Note that the geometric and the functional settings are closely related: if f is smooth and $X := \{(x,t) : x \in \mathbb{R}^n,\ t \le f(x)\}$ then $N(X)$ is the image of graph$(\nabla f) \subset \mathbb{R}^n \times \mathbb{R}^n$ under the map

$$(x;\xi) \mapsto \left(x, f(x); \frac{(-\xi;1)}{\sqrt{1+\xi^2}} \right)$$

and conversely. Another major conjecture states that this remains true also for singular f.

In order to state the fundamental uniqueness theorem we recall that $\mathbb{R}^n \times \mathbb{R}^n$ carries a natural *symplectic 2-form*

$$\omega := \sum_{i=1}^n dx_i \wedge dy_i.$$

Theorem 1. [8, 12] *Suppose $f : \mathbb{R}^n \to \mathbb{R}$ with $\nabla f \in L^1_{loc}$. Then there exists at most one closed integral current T of dimension n in $\mathbb{R}^n \times \mathbb{R}^n$ such that*

$$\int_T \omega \wedge \psi = 0 \qquad \textit{for all } \psi \in \Omega^{n-2}(\mathbb{R}^n \times \mathbb{R}^n) \tag{2.1}$$

$$\text{volume}(T \cap \pi^{-1}K) < \infty \qquad \textit{for all compact } K \subset \mathbb{R}^n \tag{2.2}$$

$$T \cap \pi^{-1}(p) = \{(p, \nabla f(p))\} \qquad \textit{for a.e. } p \in \mathbb{R}^n. \tag{2.3}$$

Here $\pi : \mathbb{R}^n \times \mathbb{R}^n \mathbb{R}^n$ is the projection to the first factor. If it exists, the current T of Theorem 1 is the *gradient cycle* of f, denoted $\mathbb{D}(f)$, and f is said to be a **Monge-Ampère (MA) function**. We denote this class by MA $=$ MA$(\mathbb{R}^n)$.

Condition (2.1) says that T is *Lagrangian.* The point is that if $V \subset \mathbb{R}^n \times \mathbb{R}^n$ is a smooth oriented submanifold of dimension n then V is Lagrangian in the usual sense iff the current T defined by integration over V satisfies (2.1).

The models of MA functions are the C^2 functions, with $\mathbb{D}(f) = \text{graph}(\nabla f)$ (here and elsewhere we identify the manifold graph(∇f) with the current obtained by integration over it with respect to the orientation induced by the standard orientation of $\mathbb{R}^n$). In this case

$$\int_{\mathbb{D}(f)} \phi(x,y)\, dy_1 \wedge \cdots \wedge dy_n = \int_{\mathbb{R}^n} \phi(x, \nabla f(x)) \det D^2 f(x)\, dx.$$

for $\phi \in C_c(\mathbb{R}^n \times \mathbb{R}^n)$, thus motivating the name of the class. The Lagrangian condition is equivalent in this case to the calculus rule $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$. The area formula yields the approximate relation

$$\text{volume}(\mathbb{D}(f)) \simeq \sum_{\#I = \#J} \int \left| \det \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)_{i \in I, j \in J} \right|. \tag{2.4}$$

In the case $n = 1$ this class is nothing new: $f \in \text{MA}(\mathbb{R})$ iff the derivative f' has locally bounded variation. Alternatively, this class may be described as the set of all functions that may be expressed as $f = g - h$ where g, h are convex. Returning to the analogy with the geometry of singular subsets of $\mathbb{R}^n$, this corresponds to the fact that a curve in $\mathbb{R}^n$ has finite total curvature iff its unit tangent vector has bounded variation.

On the other hand, for $n \geq 2$ the class MA does not fit neatly into any known analytic category. For one thing, we know that $\mathrm{MA}(\mathbb{R}^n)$ is not closed under addition for $n \geq 2$ (cf. [10]). While any $f \in \mathrm{MA}(\mathbb{R}^2)$ must be continuous, this is not known for $n \geq 3$. While it follows directly from the definitions that any $f \in \mathrm{MA}(\mathbb{R}^n)$ must have gradient $\nabla f \in \mathrm{BV}_{loc}$ (the class of such functions is sometimes denoted BV^2_{loc}), it is easy to construct examples of $f \in \mathrm{BV}^2_{loc}(\mathbb{R}^n), n \geq 2$, that are not MA. While D. Pokorný and J. Rataj [14] have recently shown that any function on $\mathbb{R}^n$ that is expressible as the difference of two convex functions must be MA, examples [10] show that not every MA function of two or more variables is of this type. Other known subclasses of MA include the Sobolev class $W^{2,n}_{loc}(\mathbb{R}^n)$ of functions with two derivatives that are locally nth power summable, and the class of all locally Lipschitz subanalytic functions.

2.2. **Strong C^2 and PL approximations.** As a consequence of Theorem 1 and the Federer-Fleming compactness theorem for integral currents, if $f_1, f_2, \dots \in C^2(\mathbb{R}^n)$ converge in L^1_{loc} to f_0, with $\mathrm{volume}(\mathbb{D}(f_i) \cap \pi^{-1}K) \leq C_K, \quad i = 1, 2, \dots$ for all compact $K \subset \mathbb{R}^n$, then $f_0 \in \mathrm{MA}$ and $\mathbb{D}(f) = \lim \mathbb{D}(f_i)$. Such f_0 is called C^2 *strongly approximable.* **All known examples of MA functions arise in this way.** Thus another fundamental conjecture states:

$$f \in \mathrm{MA} \implies f \text{ is } C^2 \text{ strongly approximable.} \tag{2.5}$$

Since piecewise linear (PL) functions are locally Lipschitz and subanalytic— in fact semialgebraic— these are always MA (in this case the C^2 strong approximability of any $p \in \mathrm{PL}$ is easy to prove using the Tarski-Seidenberg theorem). On the other hand it is also easy to construct $\mathbb{D}(p)$ directly in this case [13]. For $n = 2$ this process goes as follows. Let $\mathcal{T}$ be a triangulation of $\mathbb{R}^2$ with triangles τ_i, edges σ_j and vertices ρ_k, such that p is affine on each of these elements. We construct $\mathbb{D}(p)$ as $D_2 + D_1 + D_0$, where D_i is supported over the i-skeleton of $\mathcal{T}$.

1. Put $D_2 := \sum_i \tau_i \times \{\nabla(p|_{\tau_i})\}$. This current is Lagrangian, and satisfies (2.3), but has nonzero boundary supported above the edges σ_j.
2. For each edge σ_j with adjacent faces τ_0, τ_1, let s_j be the line segment in $\mathbb{R}^2$ joining $\nabla(p|_{\tau_0}), \nabla(p|_{\tau_1})$. Put $D_1 := \sum_j \sigma_j \times s_j$. Since the affine functions $p|_{\tau_0}, p|_{\tau_1}$ agree along σ_j, we see that $\sigma_j \perp s_j$, which implies that D_1 is Lagrangian. Clearly $\partial D_1 = \sum_j \partial\sigma_j \times s_j - \sigma_j \times \partial s_j$; the latter terms cancel ∂D_2.
3. It remains to cancel the former terms. For each vertex ρ_k, let $P_k \subset \mathbb{R}^2$ be the bounded polygonal region with multiplicities whose boundary is equal to the union of the oriented segments s_j corresponding to edges σ_j incident to ρ_k. Put $D_0 := \sum_k \rho_k \times P_k$, whose boundary provides the desired cancellation. Note that the addition of $D_1 + D_0$ leaves (2.3) unchanged.

We may think of the mass of D_0 (resp. D_1) as the integral of the absolute value of the Hessian of p (resp. the integral of the norm of the Hessian of p), which are in turn closely analogous to the total absolute Gauss curvature (resp. the integral of the norm of the second fundamental form).

Thus it would also be natural to take PL, instead of C^2 as the models for MA functions, and to say that f_0 is *PL strongly approximable* if the condition above holds with the C^2 functions f_i replaced by PL functions p_i. Again we conjecture

$$f \in \mathrm{MA} \implies f \text{ is PL strongly approximable.} \tag{2.6}$$

It is difficult (at least for us) to imagine that conjectures (2.5) and (2.6) could possibly fail, but a proof seems far away (aside from the trivial case $n = 1$). Finding ourselves in this position we must ask: are the two conjectures are equivalent? Even this problem seems difficult, although it is true for $n = 2$. This is a consequence of the following two facts.

Theorem 2 (Brehm-Kühnel [1])**.** *There is a universal constant C with the following property. Given $p \in \mathrm{PL}(\mathbb{R}^2)$ there exists a sequence $C^2(\mathbb{R}^2) \ni f_1, f_2, \dots \to p$ locally uniformly, with*

$$\limsup \mathrm{area}(\mathbb{D}(f_i|_U)) \leq C\, \mathrm{area}(\mathbb{D}(p|_U))$$

for any relatively compact open set $U \subset \mathbb{R}^2$.

Brehm and Kühnel state this result in different language, but this is an essentially equivalent formulation. Clearly this theorem yields: PL strongly approximable $\implies$ C^2 strongly approximable.

Theorem 3 (Fu-Scott [11]). *Given $f \in C^2(\mathbb{R}^2)$ there exists a sequence $\mathrm{PL}(\mathbb{R}^2) \ni p_1, p_2, \dots \to f$ locally uniformly, with*

$$\limsup \operatorname{area}(\mathbb{D}(p_i|_U)) \leq \int_U 1 + 2\sqrt{2}\,\|D^2 f\| + |\det D^2 f|$$

for any relatively compact open set $U \subset \mathbb{R}^2$.

Recall that by (2.4)

$$\operatorname{area}(\mathbb{D}(f|_U)) \simeq \int_U 1 + \|D^2 f\| + |\det D^2 f|.$$

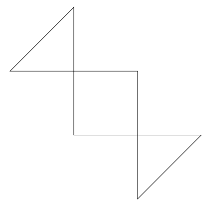
Thus Theorem 3 yields: C^2 strongly approximable $\implies$ PL strongly approximable.

2.3. **Sketch of the proof of Theorem 3.** The basic strategy is to pick an appropriate sequence of fat triangulations $\mathcal{T}_i$ of the domain of f with mesh size $\to 0$. For each i we set the values of the PL function p_i at the vertices of $\mathcal{T}_i$ equal to those of f, then extend to each triangle by linear interpolation.

The trick lies in giving meaning to the word "appropriate". If we simply take a sequence of triangulations $\mathcal{T}_i$ of the plane with mesh size $\to 0$ and uniformly positive fatness, and let p_i be the PL function obtained by linear interpolation from the values of f at the vertices of $\mathcal{T}_i$, then

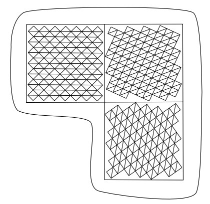
$$\limsup \operatorname{area}(\mathbb{D}(p_i|_U)) \simeq \int_U 1 + \|D^2 f\| + \|D^2 f\|^2, \tag{2.7}$$

the last term replacing the desired term $|\det D^2 f|$. Although the first two terms are acceptable, in general the last term is too big, as may be seen in the following simple example. Let $f(x, y) = (x - y)^2$ and $\mathcal{T}_i$ be a subdivided square grid aligned with the coordinate axes. Construct the PL function p_i as we have just described, and consider the gradient cycle $\mathbb{D}(p_i) = D_2 + D_1 + D_0$ given by the procedure above. The polygons P_k making up the summand D_0 are all congruent copies of the figure



of size comparable to the mesh size of $\mathcal{T}_i$. Here the square has multiplicity $+1$, while the two triangles have multiplicity -1. Thus the *algebraic* area of the figure is zero, in accord with the value $\det D^2 f = 0$, but the contribution to the mass of $\mathbb{D}(p_i)$ is twice the area of the square. Adding these contributions over all the vertices of $\mathcal{T}_i$ yields a term on the order of the integral of the last summand of (2.7). The corresponding term in the estimate we want is zero.

Fortunately, if the subdivided square grid is nearly aligned with the eigenvectors of $D^2 f$ then this bad behavior does not occur. So we construct such grids locally in regions where the eigenvectors don't vary too much, keeping the different grids separated by a distance proportional to the mesh size but together covering most of U:



Then we invoke the guarantee in a mesh interpolation algorithm of Chew [2] to conclude that the interstices can be filled in by a triangulation of the same mesh size and uniform fatness. Since the area covered by the interpolated triangles is small, the estimate (2.7) tells us that the price we pay here is not too great.

The primary obstacle to extending this argument to $n \geq 3$ is the absence of a Chew-type algorithm in higher dimensions. From our (superficial) knowledge of the relevant literature this appears to be a fundamental and poorly understood issue in the theory of mesh generation; cf. e.g. [15].

3. Further remarks and questions

3.1. **What do these questions have to do with geometric modeling?** Three dimensional modeling was one of the primary motives in the origins of surface theory in the 18th and 19th centuries. Physical objects were supposed to look like smooth domains, once irrelevant irregularities were ignored. The curvature (or the second fundamental form) provided an appealing mathematical tool with serious practical applications.

In the modern era, when computers are widely available and we no longer expect nature to behave necessarily in a smooth regular fashion, the assumption that messy natural formations can be thought of as C^2 smooth seems quaint. In this setting it is desirable to possess a more robust but still natural mathematical model that would nonetheless retain some of the main measurements such as curvature. The normal/gradient cycle of X or f provides such a tool. To put it another way, objects and functions that are regular enough to be associated to such cycles provide a model for what a natural geometric object should look like: the total volume of the cycle gives a gross numerical measure of "total curvedness" of the object, which may be distributed either smoothly or else in some irregular fashion. This tool seems uncannily applicable to physical configurations over a wide range of scales. It is tempting to take the existence of this cycle as a certificate of citizenship in the country of "geometrically valid" objects.

Conjectures (2.5) and (2.6) may be rephrased colloquially as: can we use classical mathematical analysis (C^2) or quasi-discrete computer models (PL) to survey this country to any arbitrarily given degree of accuracy? This would be roughly analogous to some basic facts from integration theory: a given signed Radon measure may be approximated weakly either by discrete or by absolutely continuous measures of the same mass.

3.2. **Towards a proof of** (2.5) **for** $n = 2$**.** Can this method be adapted to construct a strong PL approximation of a general $f \in \mathrm{MA}(\mathbb{R}^2)$?

3.3. **Is there a more natural approach to Theorem 2?** The proof of Theorem 2 in [1] seems somewhat ad hoc. A more natural proof might be possible, based on a certain well known and alluring but almost completely unexplored smoothing strategy.

The basic idea seems to have been mentioned first in [7]: if $X \subset \mathbb{R}^n$ is a compact set, and $r > 0$ is a regular value in the sense of Clarke [3] of the distance function $\delta_X := \mathrm{dist}(\cdot, X)$, then the superlevel set $\overline{X_r} := \{\delta_X \geq r\}$ has positive reach. If such r is small then for $1 \gg r \gg s > 0$ the set $X_{r,s} := \{\delta_{\overline{X_r}} \geq s\}$ is a $C^{1,1}$ domain that is close (with respect to the Hausdorff metric) to X. Furthermore the mass of the normal cycle of $X_{r,s}$ is close to that of $\overline{X_r}$. It is then easy to find a C^2 (or even C^∞) domain close to $X_{r,s}$ whose normal cycle has almost the same mass.

Supposing X to admit a normal cycle in its own right, it is tempting to carry out this procedure to try to construct a smooth domain close to X whose normal cycle has mass close to that of $N(X)$. The missing ingredient is a good estimate for the mass of $N(\overline{X_r})$ in terms of the mass of $N(X)$. In certain tightly circumscribed settings, a weaker kind of estimate is available: if X is subanalytic (e.g. a polyhedron) then the masses of the $N(\overline{X_r})$ are uniformly bounded for small $r > 0$ (this is the basis for the discussion of this subject in [9]). However, the known bound is not geometric in nature, depending instead on the complexity of the description of X as a subanalytic set. Thus it may behave badly with respect to the mass of $N(X)$.

Therefore (passing from the geometric to the functional realm) at present this approach does not yield a proof of Theorem 2. Although for each particular PL function p it yields a sequence of smooth f_i with a uniform bound on the $\mathbb{D}(f_i)$, this bound depends on the complexity of the description of p. For example, if $\mathbb{D}(p)$ is very close to the gradient cycle of a constant function, but p consists of a great many small affine pieces, the known bound on the masses of the approximating $\mathbb{D}(f_i)$ will be very large. No such general bounds in terms of the mass of $\mathbb{D}(p)$ or $N(X)$ in any nontrivial instance have been given in the literature, whether proved or conjectured.

The simplest case is that of a PL function of two variables. For this it would be enough to prove such a bound in the neighborhood of a vertex, or in other words for PL functions that are homogeneous. Since the question is now being phrased in terms of functions, it seems convenient to replace the tube construction $X \mapsto \overline{X_r}$ above by the functional analogue $p \mapsto p_r$, where for each $r > 0$ we put

$$p_r(x) := \sup_y \ p(y) - \frac{1}{r}|y - x|^2$$

Note that each $p_r, r > 0$, is semiconvex. Semiconvexity is the functional analogue of the positive reach condition.

Let $p \in PL(\mathbb{R}^2)$, with $p(tx) = tp(x)$ for $t \geq 0$ and $x \in \mathbb{R}^2$. Is there a universal local bound on the area of $\mathbb{D}(p_r)$ in terms of that of $\mathbb{D}(p)$, valid for r small? By homogeneity it is enough to understand the case $r = 1$.

REFERENCES

[1] U. Brehm and W. Kühnel. Smooth approximation of polyhedral surfaces regarding curvatures. *Geom. Dedicata*, 12:435–461, 1982.

[2] L.P. Chew. Guaranteed quality triangular meshes., 1989. `http://www.cs.sandia.gov/~samitch/unm_math_579/course_papers/chew_guaranteed_triangular_89-983.pdf`.

[3] F.H. Clarke. *Optimization and nonsmooth analysis.* Wiley, New York, 1990.

[4] D. Cohen-Steiner and J.-M. Morvan. Second fundamental measure of geometric sets and local approximation of curvatures. *J. Differential Geom.*, 74:363–394, 2006.

[5] H. Federer. Curvature measures. *Trans. Amer. Math. Soc.*, 93:418–491, 1959.

[6] H. Federer. *Geometric measure theory.* Springer, New York, 1969.

[7] J.H.G. Fu. Tubular neighborhoods in Euclidean spaces. *Duke Math. J.*, 52:1025–1046, 1985.

[8] J.H.G. Fu. Monge-Ampère functions I, II. *Indiana Univ. Math. J.*, 38:745–771, 773–789, 1989.

[9] J.H.G. Fu. Curvature measures of subanalytic sets. *Amer. J. Math.*, 116:819–880, 1994.

[10] J.H.G. Fu. An extension of Alexandrov's theorem on second derivatives of convex functions. *Adv. Math.*, 228:2258–2267, 2011.

[11] J.H.G. Fu and R.C. Scott. Piecewise linear approximation of smooth functions of two variables. *Adv. Math.*, 248:435–461, 2013.

[12] R. L. Jerrard. Some remarks on Monge-Ampère functions. In *Singularities in PDE and the calculus of variations*, volume 44 of *CRM Proc. Lecture Notes*, pages 89–112. Amer. Math. Soc., Providence, RI, 2008.

[13] R. L. Jerrard. Some rigidity results related to Monge-Ampère functions. *Canad. J. Math.*, 12:320–354, 2010.

[14] D. Pokorný and J. Rataj. Normal cycles and curvature measures of sets with d.c. boundary. *Adv. Math.*, 248:963–985, 2013.

[15] J. Shewchuk. Tetrahedral mesh generation by Delaunay refinement. In *Proceedings of the Fourteenth Annual Symposium on Computational Geometry (Minneapolis, Minnesota)*, pages 86–95, 1998.

[16] M. Zähle. Curvatures and currents for unions of sets with positive reach. *Geom. Dedicata*, 23:155–171, 1987.

Department of Mathematics, University of Georgia, Athens GA 30602, USA • fu@math.uga.edu