Euler integration over definable functions

Yuliy Baryshnikov [*], **Robert Ghrist** [†]

[*] Bell Laboratories, Murray Hill, NJ, and [†] University of Pennsylvania, Philadelphia, PA

Submitted to Proceedings of the National Academy of Sciences of the United States of America

We extend the theory of Euler integration from the class of constructible functions to that of "tame" $\mathbb{R}$-valued functions (definable with respect to an o-minimal structure). The corresponding integral operator has some unusual defects (it is not a linear operator); however, it has a compelling Morse-theoretic interpretation. In addition, we show that it is an appropriate setting in which to do numerical analysis of Euler integrals, with applications to incomplete and uncertain data in sensor networks.

definable functions | o-minimal structures | sensor networks | euler characteristic

Integration with respect to Euler characteristic is a homomorphism $\int_X \cdot d\chi : \mathsf{CF}(X) \to \mathbb{Z}$ from the ring of constructible functions $\mathsf{CF}(X)$ ("tame" integer-valued functions on a topological space X) to the integers $\mathbb{Z}$. It is a topological integration theory which uses as a measure the venerable Euler characteristic χ. Euler characteristic, suitably defined, satisfies the fundamental property of a measure:

$$\chi(A \cup B) = \chi(A) + \chi(B) - \chi(A \cap B), \qquad [1]$$

for A and B "tame" subsets of X. We extend the theory to $\mathbb{R}$-valued integrands and demonstrate its utility in managing incomplete data in, e.g., sensor networks.

Constructible integrands

Because the Euler characteristic is only finitely additive, one must continually invoke the word "tame" to ensure that χ is well-defined. One means by which to do so it via an O-MINIMAL STRUCTURE [19], a sequence $\mathcal{O} = (\mathcal{O}_n)$ of Boolean algebras of subsets of $\mathbb{R}^n$ satisfying a small list of axioms: closure under products, closure under projections, and finite decompositions in $\mathcal{O}_1$. Elements of $\mathcal{O}$ are called DEFINABLE sets and these are "tame" for purposes of integration theory. Examples of o-minimal structures include (1) piecewise-linear sets;[1] (2) semi-algebraic sets; and (3) globally subanalytic sets.

Definable functions between spaces are those whose graphs are in $\mathcal{O}$. For X and Y definable spaces, let $\mathsf{Def}(X, Y)$ denote the class of compactly supported definable functions $h : X \to Y$, and fix as a convention $\mathsf{Def}(X) = \mathsf{Def}(X, \mathbb{R})$ and $\mathsf{Def}^n(X) = \mathsf{Def}(X) \cap C^n(X)$. Let $\mathsf{CF}(X) = \mathsf{Def}(X, \mathbb{Z}) \subset \mathsf{Def}(X, \mathbb{R})$ denote the ring of CONSTRUCTIBLE FUNCTIONS: compactly supported $\mathbb{Z}$-valued functions all of whose level sets are definable. Note that in general, definable functions (even definable 'homeomorphisms') are not necessarily continuous.

We briefly recall the theory of Euler integration, established as an integration theory in the constructible setting in [12, 16, 17, 20] and anticipated by a combinatorial version in [2, 10, 11, 15]. Fix an o-minimal structure $\mathcal{O}$ on a space X. The geometric Euler characteristic is the function $\chi : \mathcal{O} \to \mathbb{Z}$ which takes a definable set $A \in \mathcal{O}$ to $\chi(A) = \sum_i (-1)^i \dim H_i^{BM}(A; \mathbb{R})$, where H_*^{BM} is the Borel-Moore homology (equivalently, singular compactly supported cohomology) of A. This also has a combinatorial definition: if A is definably homeomorphic to a finite disjoint union of (open) simplices $\coprod_j \sigma_j$, then $\chi(A) = \sum_j (-1)^{\dim \sigma_j}$. Algebraic topology asserts that χ is independent of the decomposition into simplices. The Mayer-Vietoris principle asserts that χ is a measure (or 'valuation') on $\mathcal{O}$, as expressed in Eqn. [1].

The EULER INTEGRAL is the pushforward of the trivial map $X \mapsto \{pt\}$ to $\int_X d\chi : \mathsf{CF}(X) \to \mathsf{CF}(\{pt\}) \cong \mathbb{Z}$ satisfying $\int_X \mathbb{1}_A \, d\chi = \chi(A)$ for $\mathbb{1}_A$ the characteristic function over a definable set A. From the definitions and a telescoping sum one easily obtains:

$$\int_X h \, d\chi = \sum_{s=-\infty}^{\infty} s\chi\{h = s\} = \sum_{s=0}^{\infty} \chi\{h > s\} - \chi\{h < -s\}. \qquad [2]$$

Because the Euler integral is a pushforward, any definable map $F : X \to Y$ induces $F_* : \mathsf{CF}(X) \to \mathsf{CF}(Y)$ satisfying $\int_X h \, d\chi = \int_Y F_* h \, d\chi$. Explicitly,

$$F_* h(y) = \int_{F^{-1}(y)} h \, d\chi, \qquad [3]$$

as a manifestation of the Fubini Theorem.

The Euler integral has been found to be an elegant and useful tool for explaining properties of algebraic curves [3] and stratified Morse theory [18, 4], for reconstructing objects in integral geometry [17], for target counting in sensor networks [1], and as an intuitive basis for the more general theory of motivic integration [8, 7].

Real-valued integrands

We extend the definition of Euler integration to $\mathbb{R}$-valued integrands in $\mathsf{Def}(X)$ via step-function approximations.

A Riemann-sum definition.

Definition 1. *Given $h \in \mathsf{Def}(X)$, define:*

$$\int_X h \lfloor d\chi \rfloor = \lim_{n\to\infty} \frac{1}{n} \int_X \lfloor nh \rfloor d\chi. \qquad [4]$$

$$\int_X h \lceil d\chi \rceil = \lim_{n\to\infty} \frac{1}{n} \int_X \lceil nh \rceil d\chi. \qquad [5]$$

We establish that these limits exist and are well-defined, though not equal.

Lemma 1. *Given an affine function $h \in \mathsf{Def}(\sigma)$ on an open k-simplex σ,*

$$\int_\sigma h \lfloor d\chi \rfloor = (-1)^k \inf_\sigma h \, ; \int_\sigma h \lceil d\chi \rceil = (-1)^k \sup_\sigma h. \qquad [6]$$

Reserved for Publication Footnotes

[1] Some authors require an o-minimal structure to contain algebraic curves, eliminating this particular example.

Proof: For h affine on σ, $\chi\{\lfloor nh\rfloor > s\} = (-1)^k$ for all $s < n \inf_\sigma h$, and 0 otherwise. One computes

$$\lim_{n\to\infty} \frac{1}{n}\int_\sigma \lfloor nh\rfloor d\chi = \lim_{n\to\infty}\frac{1}{n}\sum_{s=0}^{\infty}\chi\{\lfloor nh\rfloor > s\} = (-1)^k \inf_\sigma h.$$

The analogous computation holds with $\chi\{\lceil nh\rceil > s\} = (-1)^k$ for all $s < n\sup_\sigma h$, and 0 otherwise. ■

This integration theory is robust to changes in coordinates.

Lemma 2. *Integration on* $\mathsf{Def}(X)$ *with respect to* $\lfloor d\chi\rfloor$ *and* $\lceil d\chi\rceil$ *is invariant under the right action of definable bijections of* X.

Proof: This is true for Euler integration on $\mathsf{CF}(X)$; thus, it holds for $\int_X \lfloor nh\rfloor\, d\chi$ and $\int_X \lceil nh\rceil\, d\chi$. ■

Lemma 3. *The limits in Definition 1 are well-defined.*

Proof: The TRIANGULATION THEOREM for $\mathsf{Def}(X)$ [19] states that to any $h \in \mathsf{Def}(X)$, there is a definable triangulation (a definable bijection to a disjoint union of open affine simplices in some Euclidean space) on which h is affine on each open simplex. The result now follows from Lemmas 1 and 2. ■

Integrals with respect to $\lfloor d\chi\rfloor$ and $\lceil d\chi\rceil$ are related to total variation (in the case of compactly supported continuous functions).

Corollary 1. *If* M *is a 1-dimensional manifold and* $h \in \mathsf{Def}^0(M)$, *then*

$$\int_M h\,\lfloor d\chi\rfloor = -\int_M h\,\lceil d\chi\rceil = \frac{1}{2}\mathsf{totvar}(h). \qquad [7]$$

Proof: Apply Lemma 1 to an affine triangulation of h which triangulates M with the maxima $\{p_i\}$ and minima $\{q_j\}$ as 0-simplices and the intervals between them as 1-simplices. To each minimum q_j is associated two open 1-simplicies, since M is a 1-manifold. Thus:

$$\int_M h\,\lfloor d\chi\rfloor = \sum_i h(p_i) + \sum_j h(q_j) - 2\sum_j h(q_j) = \frac{1}{2}\mathsf{totvar}(h).$$

This equals $-\int_M h\,\lceil d\chi\rceil$ via an analogous computation. ■

This result generalizes greatly via Morse theory: see Corollary 5. One notes that $\lfloor d\chi\rfloor$ and $\lceil d\chi\rceil$ give integrals which are conjugate in the following sense.

Lemma 4.

$$\int h\lceil d\chi\rceil = -\int -h\lfloor d\chi\rfloor. \qquad [8]$$

Proof: Apply Lemma 1 to an affine triangulation of h, and note that $\sup_\sigma h = -\inf_\sigma -h$. ■

The temptation to cancel the negatives must be resisted: see Lemma 5 below.

Computation. Definition 1 has a Riemann-sum flavor which extends to certain computational formulae. The following is a definable analogue of Eqn. [2].

Proposition 2. *For* $h \in \mathsf{Def}(X)$,

$$\int_X h\,\lfloor d\chi\rfloor = \int_{s=0}^{\infty} \chi\{h \geq s\} - \chi\{h < -s\}\, ds \qquad [9]$$

$$\int_X h\,\lceil d\chi\rceil = \int_{s=0}^{\infty} \chi\{h > s\} - \chi\{h \leq -s\}\, ds. \qquad [10]$$

Proof: For $h \geq 0$ affine on an open k-simplex σ,

$$\int_\sigma h\,\lfloor d\chi\rfloor = (-1)^k \inf_\sigma h = \int_0^\infty \chi(\sigma\cap\{h\geq s\})ds,$$

and for $h \leq 0$, the equation holds with $-\chi(\sigma\cap\{h < -s\})$. The result for $\int\lceil d\chi\rceil$ follows from Lemma 4. ■

It is not true that $\int_X h\,\lfloor d\chi\rfloor = \int_0^\infty s\chi\{h = s\}ds$: the proper Lebesgue generalization of Eqn. [2] is the following:

Proposition 3. *For* $h \in \mathsf{Def}(X)$,

$$\int_X h\,\lfloor d\chi\rfloor = \lim_{\epsilon\to 0^+}\frac{1}{\epsilon}\int_{\mathbb{R}} s\,\chi\{s \leq h < s+\epsilon\}\, ds \qquad [11]$$

$$\int_X h\,\lceil d\chi\rceil = \lim_{\epsilon\to 0^+}\frac{1}{\epsilon}\int_{\mathbb{R}} s\,\chi\{s < h \leq s+\epsilon\}\, ds. \qquad [12]$$

Proof: For h affine on an open k-simplex σ, and $0 < \epsilon$ sufficiently small, $\int_{\mathbb{R}} s\,\chi\{s \leq h < s+\epsilon\}\, ds = \epsilon\,(-1)^k\left(-\frac{\epsilon}{2} + \inf_\sigma h\right)$ and $\int_{\mathbb{R}} s\,\chi\{s < h \leq s+\epsilon\}\, ds = \epsilon\,(-1)^k\left(-\frac{\epsilon}{2} + \sup_\sigma h\right)$. ■

Morse theory. One important indication that the definition of $\int\lfloor d\chi\rfloor$ is correct for our purposes is the natural relation to Morse theory: the integrals with respect to $\lfloor d\chi\rfloor$ and $\lceil d\chi\rceil$ are Morse index weighted sums of critical values of the integrand. This is a localization result, reducing from an integral over all of X to an integral over an often discrete set of critical points.

Recall that a C^2 function $h : M \to \mathbb{R}$ on a smooth manifold M is MORSE if all critical points of h are nondegenerate, in the sense of having a nondegenerate Hessian matrix of second partial derivatives. Denote by $\mathcal{C}(h)$ the set of critical points of h. For each $p \in \mathcal{C}(h)$, the MORSE INDEX of p, $\mu(p)$, is defined as the number of negative eigenvalues of the Hessian at p, or, equivalently, the dimension of the unstable manifold $W^u(p)$ of the vector field $-\nabla h$ at p.

Stratified Morse theory [9] is a powerful generalization to triangulable spaces, including definable sets with respect to an o-minimal structure [4, 18]. We may interpret $\lfloor d\chi\rfloor$ and $\lceil d\chi\rceil$ in terms of the weighted stratified Morse index of the graph of the integrand.

Definition 2. *For* $X \subset \mathbb{R}^n$ *definable and* $h \in \mathsf{Def}(X)$, *define the co-index of* h, $\mathcal{I}^*h$ *to be the stratified Morse index of the graph of* h, $\Gamma_h \subset X\times\mathbb{R}$, *with respect to the projection* $\pi : X\times\mathbb{R}\to\mathbb{R}$:

$$(\mathcal{I}^*h)(x) = \lim_{\epsilon'\ll\epsilon\to 0^+}\chi\left(\overline{B_\epsilon(x)}\cap\{h < h(x)+\epsilon'\}\right), \qquad [13]$$

where $\overline{B_\epsilon(x)}$ *is the closed ball of radius* ϵ *about* $x \in X$. *The index* $\mathcal{I}_*$ *is the stratified Morse index with respect to height function* $-\pi$: *i.e.,* $\mathcal{I}_*h = \mathcal{I}^*(-h)$ *or,*

$$(\mathcal{I}_*h)(x) = \lim_{\epsilon'\ll\epsilon\to 0^+}\chi\left(\overline{B_\epsilon(x)}\cap\{h > h(x)-\epsilon'\}\right). \qquad [14]$$

Note that $\mathcal{I}_*, \mathcal{I}^* : \mathsf{Def}(X) \to \mathsf{CF}(\overline{X})$, and the restriction of these operators to $\mathsf{CF}(X)$ is the identity (every point of a constructible function is a critical point). The two types of integration on $\mathsf{Def}(X)$ correspond to the Morse indices of the graph with respect to the two orientations of the graph axis — the projections π and $-\pi$.

Theorem 4. *For* $h \in \mathsf{Def}^0(X)$,

$$\int_X h\,\lfloor d\chi\rfloor = \int_{\overline{X}} h\mathcal{I}^*h\, d\chi \quad ; \quad \int_X h\,\lceil d\chi\rceil = \int_{\overline{X}} h\mathcal{I}_*h\, d\chi. \qquad [15]$$

Proof: On an open k-simplex $\sigma \subset X \subset \mathbb{R}^n$ in an affine triangulation of h, the co-index $\mathcal{I}^*h$ equals $(-1)^{\dim(\sigma)}$ times the characteristic function of the closed face of σ determined by $\inf_\sigma h$. Since h is continuous, $\int_{\overline{\sigma}} h\mathcal{I}^*h\, d\chi = (-1)^{\dim(\sigma)}\inf_\sigma h$. Lemma 1 and additivity complete the proof; the analogous proof holds for $\mathcal{I}_*$ and $\lceil d\chi\rceil$. ■

Corollary 5. *If h is a Morse function on a closed n-manifold M, then:*

$$\int_M h\,\lfloor d\chi\rfloor = \sum_{p\in\mathcal{C}(h)} (-1)^{n-\mu(p)} h(p); \qquad [16]$$

$$\int_M h\,\lceil d\chi\rceil = \sum_{p\in\mathcal{C}(h)} (-1)^{\mu(p)} h(p). \qquad [17]$$

Proof: For $p \in \mathcal{C}(h)$ a nondegenerate critical point on an n-manifold, $\mathcal{I}^*h(p) = (-1)^{n-\mu(p)}$ and $\mathcal{I}_*h(p) = (-1)^{\mu(p)}$. ■

From this, one sees clearly that the relationship between $\lfloor d\chi\rfloor$ and $\lceil d\chi\rceil$ is regulated by Poincaré duality. For example, on continuous definable integrands over an n-dimensional manifold M,

$$\int_M h\,\lceil d\chi\rceil = (-1)^n \int_M h\,\lfloor d\chi\rfloor. \qquad [18]$$

The generalization from continuous to general definable integrands is simple, but requires weighting $\mathcal{I}^*h$ by h directly. To compute $\int_X h\lfloor d\chi\rfloor$, one integrates the weighted co-index

$$\lim_{\epsilon'\ll\epsilon\to 0^+} h(x+\epsilon')\chi\left(\overline{B_\epsilon(x)} \cap \{h < h(x)+\epsilon'\}\right) \qquad [19]$$

with respect to $d\chi$.

Corollary 5 can also be proved directly using classical handle-addition techniques or in terms of the Morse complex, using the fact that the restriction of the integrand to each unstable manifold of each critical point is unimodal with a unique maximum at the critical point. It is also possible to express the stratified Morse index — and thus the integral here considered — in terms of integration against a characteristic cycle, *cf.* [9, 18].

One final means of illustrating Corollary 5 is to use a deformation argument. Let h be smooth on X and ϕ_t be the flow of $-\nabla h$. Then the integral is invariant under the action of ϕ_t on h; yet the limiting function $h_\infty = \lim_{t\to\infty} h\circ\phi_t$ is constant on stable manifolds of $-\nabla h$ with values equal to the critical values of h. We have not shown that the limiting function is constructible (this depends on the existence of definable invariant manifolds — we are unaware of relevant results in the literature) and thus do not rely on this method for proof but rather illumination.

The integral operator

We consider properties of the integral operator(s) on $\mathsf{Def}(X)$.

Linearity. One is tempted to apply all the standard constructions of sheaf theory (as in [16, 17]) to $\int_X : \mathsf{Def}(X) \to \mathbb{R}$. However, our formulation of the integral on $\mathsf{Def}(X)$ has a glaring disadvantage.

Lemma 5. $\int_X$: $\mathsf{Def}(X) \to \mathbb{R}$ *(via $\lfloor d\chi\rfloor$ or $\lceil d\chi\rceil$) is not a homomorphism for* $\dim X > 0$.

Proof:

$$1 = \int_{[0,1]} 1\,\lfloor d\chi\rfloor \neq \int_{[0,1]} x\,\lfloor d\chi\rfloor + \int_{[0,1]} (1-x)\,\lfloor d\chi\rfloor = 1+1 = 2.$$

■

This loss of functoriality can be seen as due to the fact that $\lfloor f+g\rfloor$ agrees with $\lfloor f\rfloor + \lfloor g\rfloor$ only up to a set of Lebesgue measure zero, not χ-measure zero. The nonlinear nature of the integral is also clear from Eqn. [15], as Morse data is non-additive.

The Fubini Theorem. In one sense, the change of variables formula trivializes (Lemma 2). The more general change of variables formula encapsulated in the Fubini theorem does not, however, hold for non-constructible integrands.

Corollary 6. *The Fubini theorem fails on* $\mathsf{Def}(X)$ *in general.*

Proof: Let $F : X = Y \coprod Y \to Y$ be the projection map with fibers $\{p\}\coprod\{p\}$. Any $h \in \mathsf{Def}(X)$ is expressible as $f \coprod g$ for $f, g \in \mathsf{Def}(Y)$. The Fubini theorem applied to F is equivalent to the statement

$$\int_Y f + \int_Y g = \int_X h = \int_Y F_*h = \int_Y f+g$$

(where the integration is with respect to $\lfloor d\chi\rfloor$ or $\lceil d\chi\rceil$ as desired). Lemma 5 completes the proof. ■

Fubini holds when the map respects fibers.

Theorem 7. *For $h \in \mathsf{Def}(X)$, let $F : X \to Y$ be definable and h-preserving (h is constant on fibers of F). Then $\int_Y F_*h\lfloor d\chi\rfloor = \int_X h\lfloor d\chi\rfloor$, and $\int_Y F_*h\lceil d\chi\rceil = \int_X h\lceil d\chi\rceil$.*

Proof: An application of the o-minimal Hardt theorem [19] implies that Y has a partition into definable sets Y_α such that $F^{-1}(Y_\alpha)$ is definably homeomorphic to $U_\alpha \times Y_\alpha$ for U_α definable, and that $F : U \times Y_\alpha \to Y_\alpha$ acts via projection. Since h is constant on fibers of F, one computes

$$\int_{Y_\alpha} F_*h\lfloor d\chi\rfloor = \int_{Y_\alpha} h\,\chi(U_\alpha)\lfloor d\chi\rfloor = \int_{U_\alpha\times Y_\alpha} h\lfloor d\chi\rfloor.$$

The same holds for $\int \lceil d\chi\rceil$. ■

Corollary 8. *For $h \in \mathsf{Def}(X)$, $\int_X h = \int_{\mathbb{R}} h_*h$. In other words,*

$$\int_X h\,\lfloor d\chi\rfloor = \int_{\mathbb{R}} s\,\chi\{h=s\}\lfloor d\chi\rfloor, \qquad [20]$$

and likewise for $\lceil d\chi\rceil$.

Continuity. Though the integral operator is not linear on $\mathsf{Def}(X)$, it does retain some nice properties. All properties below stated for $\int \lfloor d\chi\rfloor$ hold for $\int \lceil d\chi\rceil$ via duality.

Lemma 6. *The integral $\int \lfloor d\chi\rfloor : \mathsf{Def}(X) \to \mathbb{R}$ is positively homogeneous.*

Proof: For $f \in \mathsf{Def}(X)$ and $\lambda \in \mathbb{R}^+$, the change of variables variables $s \mapsto \lambda s$ in Eqn. [9] gives $\int \lambda f\,\lfloor d\chi\rfloor = \lambda \int f\,\lfloor d\chi\rfloor$. ■

Integration is not continuous on $\mathsf{Def}(X)$ with respect to the C^0 topology. An arbitrarily large change in $\int h\lfloor d\chi\rfloor$ may be effected by small changes to h on a (large) finite point set. In some situations the "complexity" of the definable functions can be controlled in a way sufficient to ensure continuity.

One example arises in the semialgebraic category. Fix a (finite) semialgebraic stratification $\mathcal{S}$ of a compact definable X, and consider definable semialgebraic functions *algebraic* with respect to this stratification (that is such that the restriction of the function to any stratum $S \in \mathcal{S}$ is a polynomial P_S). The resulting linear space (filtered by the subspaces of polynomials of bounded degree) can be equipped with the structure of a Banach space, by completing the family of seminorms $\|P\|_{S,\mathbf{k}} = \max_{S\in\mathcal{S}} ||P_S||_{C^n}$, where $n = \dim X$. Then $\int_X \cdot\lfloor d\chi\rfloor$ becomes a continuous (non-linear) functional on this Banach space. The proof results, essentially, from the Bézout theorem (mimicking Thom-Milnor theory): the total number of critical points graph of a polynomial of degree D on a fixed semi-algebraic set is bounded by $O(D^n)$. The generalization to increasing (refined) stratifications is straightforward.

Integration itself defines a natural topology for $\mathsf{Def}(X)$ on which integration is continuous. Define the L^1 ϵ-neighborhood of $h \in \mathsf{Def}(X)$ as the intersection of the C^0 ϵ-neighborhood (definable functions with ϵ-close graphs) with those functions

$g \in \mathsf{Def}(X)$ satisfying $\left|\int_X f - g\lfloor d\chi\rfloor\right| < \epsilon$. This provides a basis for an L^1 topology on $\mathsf{Def}(X)$. As a consequence of Lemma 4, the definition is independent of the use of $\lfloor d\chi\rfloor$ or $\lceil d\chi\rceil$.

The interested reader may speculate on other function space topologies on $\mathsf{Def}(X)$.

Duality and links. There is an integral transform on $\mathsf{CF}(X)$ that is the analogue of Poincaré-Verdier duality [18]. It extends seamlessly to integrals on $\mathsf{Def}(X)$ by means of the following definition.

Definition 3. *The* DUALITY OPERATOR *is the integral transform* $\mathcal{D} : \mathsf{CF}(X) \to \mathsf{CF}(X)$ *given by*

$$\mathcal{D}h(x) = \lim_{\epsilon\to 0^+} \int_X h\mathbf{1}_{B_\epsilon(x)} d\chi, \qquad \textbf{[21]}$$

where B_ϵ is an open metric ball of radius ϵ.

We extend the definition to $\mathcal{D} : \mathsf{Def}(X) \to \mathsf{Def}(X)$ by integrating with respect to $\lfloor d\chi\rfloor$ or $\lceil d\chi\rceil$, interchangeably, via:

Lemma 7. *$\mathcal{D}h$ is well-defined on $\mathsf{Def}(X)$ and independent of whether the integration in (21) is with respect to $\lfloor d\chi\rfloor$ or $\lceil d\chi\rceil$.*

Proof: The limit is always well-defined thanks to the Conic Theorem in o-minimal geometry [19]. To show that it is independent of the upper- or lower-semicontinuous approximation, take $\epsilon > 0$ sufficiently small. Note that by triangulation, h can be assumed to be piecewise-affine on open simplices. Pick a point x in the support of h and let $\{\sigma_i\}$ be the set of open simplices whose closures contain x. Then for each i, the limit $h_i(x) := \lim_{y\to x} h(y)$ for $y \in \sigma_i$ exists. Then, applying Eqn. [21], one computes

$$\mathcal{D}h(x) = \sum_i (-1)^{\dim \sigma_i} h_i(x), \qquad \textbf{[22]}$$

independent of the measure $\lfloor d\chi\rfloor$ or $\lceil d\chi\rceil$. ■

For a continuous definable function h on a manifold M, $\mathcal{D}h = (-1)^{\dim M}h$, as one can verify by combining Eqns. [9] and [21]. This is commensurate with the result of Schapira [16] that $\mathcal{D}$ is an involution on $\mathsf{CF}(X)$.

Theorem 9. *Duality is involutive on $\mathsf{Def}(X)$: $\mathcal{D} \circ \mathcal{D}h = h$.*

Proof: Given h, fix a triangulation on which h is affine on open simplices. From Eqn. [22], we see that the dual of h at x is completely determined by the trivialization of h at x. Let $L_x h$ be the constructible function on $B_\epsilon(x)$ which takes on the value $h_i(x)$ on strata $\sigma_i \cap B_\epsilon(x)$. (Though this is not necessarily an integer-valued function, its range is discrete and therefore it is constructible.) As $L_x h$ is close to h in $B_\epsilon(x)$ (this follows from the continuity of h on each of the strata), $\mathcal{D}h$ is close to $\mathcal{D}L_x h$ in $B_\epsilon(x)$: indeed, the total Betti number of intersections of strata with any ball $B_\epsilon(y)$ is bounded, and Euler integral of a function small in absolute value is small as well. Hence the definable function $\mathcal{D}^2h$ is close to the constructible function $\mathcal{D}^L_x h$ with ϵ small. As $\mathcal{D}^2 L_x h(x) = L_x h(x) = h(x)$, the result follows. ■

One can define related integral transforms. For example, the LINK of $h \in \mathsf{CF}(X)$ is defined as

$$\Lambda h(x) = \lim_{\epsilon\to 0^+} \int_X h\mathbf{1}_{\partial B_\epsilon(x)} d\chi. \qquad \textbf{[23]}$$

The link of a continuous function on an n-manifold M is multiplication by $1 + (-1)^n$, as a simple computation shows. In general, $\Lambda = \mathsf{Id} - \mathcal{D}$, where Id is the identity operator.

Integral transforms

Integration with respect to Euler characteristic over $\mathsf{CF}(X)$ has a well-defined and well-studied class of integral transforms, expressed beautifully in Schapira's work on inversion formulae for the generalized Radon transform in $d\chi$ [17]. Integral transforms with respect to $\lfloor d\chi\rfloor$ and $\lceil d\chi\rceil$ are similarly appealing, with applications to signal processing as a primary motivation. Examples of interesting definable kernels for integral transforms over Euclidean $\mathbb{R}^n$ include $\|x - y\|$, $\langle x, y\rangle$, and $g(x - y)$ for some g. These evoke Bessel (Hankel) transforms, Fourier transforms, and convolution with g respectively. The choice between $\lfloor d\chi\rfloor$ and $\lceil d\chi\rceil$ makes a difference, of course, but can be amalgamated. Example: for fixed kernel K, one can consider the mixed integral transform $h \mapsto \int_X hK\lfloor d\chi\rfloor - \int_X hK\lceil d\chi\rceil$. In the case of $K(x,\xi) = \langle x, \xi\rangle$, this transform takes $\mathbf{1}_A$ for A compact and convex to the 'width' of A projected to the ξ-axis.

Convolution. On a vector space V (or Lie group, more generally), a convolution operator with respect to Euler characteristic is straightforward. Given $f, g \in \mathsf{CF}(V)$, one defines

$$(f * g)(x) = \int_V f(t)g(x - t)\, d\chi. \qquad \textbf{[24]}$$

Convolution behaves as expected in $\mathsf{CF}(V)$. By reversing the order of integration, one has immediately that $\int_V f * g\, d\chi = \int_V f\, d\chi \int_V g\, d\chi$. There is a close relationship between convolution and the Minkowski sum, as observed in, *e.g.*, [10]: for A and B convex and closed $\mathbf{1}_A * \mathbf{1}_B = \mathbf{1}_{A+B}$, cf. [20, 16]. Convolution is a commutative, associative operator providing $\mathsf{CF}(V)$ with the structure of an (interesting [3]) algebra.

Convolution is well-defined on $\mathsf{Def}(V)$ by integrating with respect to $\lfloor d\chi\rfloor$ or $\lceil d\chi\rceil$. However, the product formula for $\int f * g$ fails in general, since one relies on the Fubini theorem to prove it in $\mathsf{CF}(V)$.

Linearity. The nonlinearily of the integration operator prevents most straightforward applications of Schapira's inversion formula. Fix a kernel $K \in \mathsf{Def}(X \times Y)$ and consider the integral transform $\mathcal{T}_K : \mathsf{Def}(X) \to \mathsf{Def}(Y)$ of the form $(\mathcal{T}_K h)(y) = \int_X h(x)K(x, y)\lfloor d\chi\rfloor(x)$. In general, this operator is non-linear, via Lemma 5. However, some vestige of (positive) linearity survives within CF^+, the *positive* linear combinations of indicator functions over tame top-dimensional subsets of X.

Lemma 8. *The integral transform $\mathcal{T}_K$ is positive-linear over* $\mathsf{CF}^+(X)$.

Proof: Any $h \in \mathsf{CF}^+(X)$ is of the form $h = \sum_k a_k \mathbf{1}_{U_k}$ for $a_k \in \mathbb{N}$ and $U_k \in \mathsf{Def}(X)$. For $h = \mathbf{1}_A$, $\mathcal{T}_K h = \int_A K\lfloor d\chi\rfloor$. Additivity of the integral in $\lfloor d\chi\rfloor$ (via Eqn. [1]) combined with Lemma 6 completes the proof. ■

This implies in particular that when one convolves a function $h \in \mathsf{CF}^+(\mathbb{R}^n)$ with a smoothing kernel (e.g., a Gaussian) as a means of filtering noise or taking an average of neighboring data points, that convolution may be analyzed one step at a time (decomposing h).

Integral transforms are not linear over all of $\mathsf{CF}(X)$, since $\int -h\lfloor d\chi\rfloor \neq -\int h\lfloor d\chi\rfloor$. However, integral transforms which combine $\lfloor d\chi\rfloor$ and $\lceil d\chi\rceil$ compensate for this behavior. Define the measure $[d\chi]$ to be the average of $\lfloor d\chi\rfloor$ and $\lceil d\chi\rceil$:

$$\int_X h[d\chi] = \frac{1}{2}\left(\int_X h\lfloor d\chi\rfloor + \int_X h\lceil d\chi\rceil\right). \qquad \textbf{[25]}$$

Theorem 10. *Any integral transform of the form*

$$(\mathcal{T}_K h)(y) = \int_X h(x)K(x, y)[d\chi](x) \qquad \textbf{[26]}$$

is a linear operator $\mathsf{CF}(X) \to \mathsf{Def}(Y)$.

Proof: From Lemma 8, $\mathcal{T}$ is positive-linear over $\mathsf{CF}^+(X)$. Full linearity follows from the observation that $\int_X -h[d\chi] = -\int_X h[d\chi]$, which follows from Lemma 1 by triangulating h. ■

As a simple example, consider the transform with kernel $K(x,\xi) = \langle x, \xi\rangle$. The transform of 1_A with respect to $[d\chi]$ for A compact and convex equals a 'centroid' of A along the ξ-axis: the average of the maximal and minimal values of ξ on ∂A. Note how the dependence on critical values of the integrand on ∂A reflects the Morse-theoretic interpretation of the integral in this case.

Integration with respect to $[d\chi]$ seems suitable only for integral transforms over CF. On a continuous integrand, the integral with respect to $[d\chi]$ either returns zero (cf. the integral of Rota [15]) or else the integral with respect to $\lfloor d\chi \rfloor$, depending on the parity of the $\dim X$, via Eqn. [18].

Definable Euler integration and numerical analysis

The Euler calculus on CF has applications to tomography [17] and data aggregation in networks [1]; the extension to Def deepens this utility and opens new potential applications. We highlight applications to numerical approximation of Euler integrals, motivated by work in sensor networks.

Motivation: sensor networks. An application of Euler integration over $\mathsf{CF}(X)$ to sensor network data aggregation problems was initiated in [1]. Consider a space X whose points represent target-counting sensors that scan a workspace W containing a discrete set of targets. Target detection is encoded in a SENSING RELATION $\mathcal{S} \subset W \times X$ where $(w,x) \in \mathcal{S}$ iff a target at $w \in W$ is detected by a sensor at $x \in X$. Assume that sensors count the number of sensed targets, but do not locate or identify the targets. The sensor network therefore induces a constructible TARGET COUNTING FUNCTION $h : X \to \mathbb{N}$ of the form $h = \sum_\alpha 1_{U_\alpha}$, where U_α is the (target's) SUPPORT — the set of sensors which detect target α. Euler integration allows for simple enumeration:

Theorem 11. ([1]) *Assume h as above with $\chi(U_\alpha) = N \neq 0$ for all α. Then the number of targets in W is $\frac{1}{N}\int_X h\, d\chi$.*

This result and its extensions provide motivation to accurately estimate an Euler integral when the integrand h is (1) known not on all of X (a continuum of sensors being unrealistic) but rather on a sufficiently dense set of sample points (physical sensors in a network); or, (2) when h is known only after being modified by noise. We argue that the integral over Def as opposed to CF provides a better theory with which to approximate these integrals.

Instability of constructible Euler integration. Though integration with respect to Euler characteristic has a lengthy history, there appears to be no treatment of numerical integration, even in the simpler setting of $\mathsf{CF}^+(\mathbb{R}^n)$. As in the case of numerical integration for Riemann integrals, one typically assumes something about the features of h and the density and extent of the sampling set. Given our motivation, we restrict to integrands in $\mathsf{CF}^+(\mathbb{R}^n)$ as a model for signal processing over geographically distributed sensor networks.

We begin with the observation that integration with respect to $d\chi$ is numerically unstable, with sensitivity to both discretization and to noise. The canonical example (see Figure 1) is the sum of Heaviside step functions $h(x,y) := H_0(x) + H_0(y) \in \mathsf{CF}^+(\mathbb{R}^2)$. Note: any generic codimension two singularity in $h \in \mathsf{CF}^+(\mathbb{R}^2)$ is locally equivalent to h plus a constant.

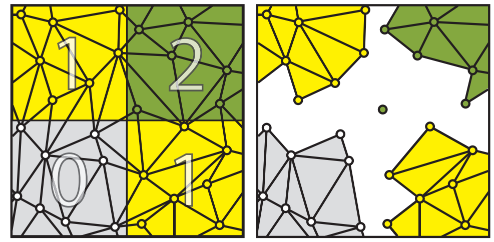

Fig. 1. Sampling over a triangulation [left] leads to an error of $+1$ in the integral with respect to $d\chi$ due to disjoint level sets [right].

A random triangulation will be with positive probability locally equivalent to that in Figure 1. The level sets of h with respect to the triangulation produce a spurious path component, adding 1, wrongly, to the Euler integral. This example is scale-invariant and not affected by increased density of sampling. In a random placement of N discs in a bounded planar domain, the expected number of such intersection points is quadratic in N, leading to potentially large errors in numerical approximations to $\int \cdot d\chi$, regardless of sampling density.

Similar phenomena can happen in the discretizations of smooth definable functions, but the errors are far milder. Assume that sensor nodes $\mathcal{N}$ are the vertex set of a triangulation $\mathcal{T}$. Restricting the smooth function h to $\mathcal{N}$ and extending it affinely on each simplex of $\mathcal{T}$ yields an approximation $h_\mathcal{T}$ of h. If h is smooth enough, the Euler integrals of h and $h_\mathcal{T}$ are close for fine enough triangulations. Given $h \in \mathsf{Def}^2(\mathbb{R}^2)$ and h_{PL} the piecewise-linear function obtained from sampling h on the vertex set of a triangulation $\mathcal{T}$ of $\mathbb{R}^2$. As the sampling and triangulation are refined,

$$\lim_{|\mathcal{T}|\to 0^+} \left| \int_X h_{PL}\lfloor d\chi \rfloor - \int_X h \lfloor d\chi \rfloor \right| = O(|\mathcal{T}|), \qquad \textbf{[27]}$$

as the reader may easily verify. Thus, it is advantageous to work with discrete approximations to smooth integrands.

In like manner, integration over $\mathsf{CF}^+(X)$ is poorly behaved with respect to noise, owing to the fact that points have full measure in $d\chi$. Assume a sampling of $f \in \mathsf{CF}^+(X)$ over a network, with an error of ± 1 on random nodes: specifically, $h = f + e$, where $e : \mathcal{N} \to \{-1, 0, 1\}$ is an error function that is nonzero on a sparse subset $\mathcal{N}' \subset \mathcal{N}$. For typical choices of $\mathcal{N}'$, $\left|\int h\, d\chi - \int f\, d\chi\right|$ will be a normal of variance $O(|\mathcal{N}'|^{\frac{1}{2}})$.

Feature size and convolution. Theorem 11 assumes a sensor modality leading to $\mathbb{N}$-valued integrands; in practice, such are rare. It is much more common to have sensors register real-valued intensities, through both device constraints as well as noise and fluctuation. We propose assuming a convolution $h * K$ of $h \in \mathsf{CF}^+(\mathbb{R}^n)$ with an appropriate kernel K. Ideally, K would be unimodal and of small support relative to the structure of h. To quantify the optimal size of K, we introduce a characteristic length (related to the WEAK FEATURE SIZE of [8]) which encodes the fragility of an integrand $h \in \mathsf{CF}^+(\mathbb{R}^n)$ with respect to $\lfloor d\chi \rfloor$.

Definition 4. *The* FEATURE SIZE *of $h \in \mathsf{CF}^+(\mathbb{R}^n)$ is $FS(h)$, the sup of all R such that, for any closure $\mathcal{C}$ of any connected component of upper or lower excursion sets, and for any point $x \in \mathcal{C}$, the convex hull of bases of outward oriented normals from x to the boundary of $\mathcal{C}$ of length at most R does not contain x.*

It is reasonable to expect in practice that a sensor network returns not a pure $\mathbb{N}$-valued integrand $h \in \mathsf{CF}^+$ but rather a $\mathbb{R}^+$-valued smoothed integrand — a convolution of the form $h*K$ for $h \in \mathsf{CF}^+$ and K some smoothing kernel. The feature size of h is key to showing when convolution does not impact the (definable) integral.

Theorem 12. *For $h \in \mathsf{CF}^+(\mathbb{R}^n)$ and $K \in \mathsf{Def}^1(\mathbb{R}^n)$ a radially-symmetric kernel with support of diameter $R < FS(h)$,*

$$\int_{\mathbb{R}^n} h\, d\chi = \int_{\mathbb{R}^n} h*K\lfloor d\chi\rfloor. \qquad [28]$$

Proof: Note that the convolution $h*K$ is C^1. By Lemma 8, $h*K$ is the sum over α of $\mathbf{1}_{U_\alpha}*K$. Each such convolution is a smoothing of $\mathbf{1}_{U_\alpha}$ on an exterior tubular neighborhood of radius R of U_α. The integral of each $\mathbf{1}_{U_\alpha}*K$ is thus unchanged. Lack of linearity for $\int\lfloor d\chi\rfloor$ forbids concluding that the integral of the full $h*K$ is unchanged. It does suffice, however, to show that no changes in critical points arise in summing these R-neighborhood smoothed supports.

Consider x in the intersection of the exterior R-neighborhoods of (after re-indexing) a collection $\{U_\beta\}$. The gradient $\nabla_x(h*K)$ is a weighted vector sum over β of $\nabla_x(\mathbf{1}_{U_\beta}*K)$. By radial symmetry of K, each such vector is parallel (and oppositely oriented) to the normal from x to U_β. By definition of feature size, this sum cannot be zero, and thus no new critical points arise in the convolution $h*K$. Invoking Theorem 4 and verifying that the indices of previously existing critical sets are unchanged by smoothing, one has that the integrals agree. ■

(This theorem remains valid with convolution $h*K$ in Lebesgue measure as well.) Theorem 12 prompts the question of typical feature size for an arrangement of targets. We claim that feature size is typically small. Assume a collection $\mathcal{B}$ of bodies $B_\ell \subset \mathbb{R}^d$ for $\ell = 1,\ldots,L$ with feature size bounded below[2] and smooth boundaries having curvatures and surface areas bounded above. To model the random placements of elements of $\mathcal{B}$ in a suitably sized domain Λ, we assume that one draws at random elements $g = (g_1,\ldots,g_L)$ from the group of Euclidean motions of $\mathbb{R}^d$ (using product Haar measure) and displaces each B_ℓ by said motions. Denote by $h_g := \sum_\ell \mathbf{1}_{g_\ell\cdot B_\ell}$.

Theorem 13. *For bodies $\mathcal{B} = \{B_\ell\}_1^L$ in a bounded domain $\Lambda \subset \mathbb{R}^d$, the (product) Haar measure of the set $\{g : FS(h_g) < R\}$ is $O(R)$ as $R \to 0^+$.*

Proof: Fix a motion g and say that a point $x \in \mathbb{R}^d$ is *critical* with respect to g if there are $2 \le k \le L$ bodies among $g_1B_1,...,g_LB_L$ (we denote these k bodies as $B_1',\ldots,B_k'$) such that there exist outward normals from x to the boundaries of these bodies of lengths at most R and spanning at most a $(k-1)$-dimensional affine plane. For a point y at distance at most R from x, there exist normals to the boundaries of the B_k''s of length at most $2R$, and such that the angle between the normal to B_1' and the affine plane spanned by the normals to $B_2',\ldots,B_k'$ is at most $O(R)$, with constants depending on the curvature bound for $\mathcal{B}$ and the dimension d. The measure of such collections of k normals (in the k-fold product of $\mathbb{S}^{d-1}$) is $O(R)$, with constant depending on d.

The collections of motions $g_1,\ldots,g_k$ for which y has an R-close critical point can be parameterized by the directions of k normals through y (the measure of such collections of unit vectors in $(\mathbb{S}^{d-1})^k$ is $O(R)$); by the lengths of these normals (each contributing a factor $2R$); by the positions of the bases of the normals on the boundaries of $B_1',\ldots,B_k'$ (each body contributing a constant depending on surface areas of $\mathcal{B}$), and by the auxiliary rotations of the bodies preserving the corresponding normals (contributing factors equal to the volumes of the orthogonal groups in $\mathbb{R}^{d-1}$). Combining and taking into account the $O(L^{d+1})$ choices of the bodies, we arrive an $O((LR)^{d+1})$ estimate for the volume of motions of $\mathcal{B}$ producing a critical point at distance at most R from y.

Let Y be the R-net for Λ, of size $O(|\Lambda|/R^d)$. If $FD(h_g) < R$, then there exists a point $y \in Y$ having a critical point for g at distance at most R. The measure of such displacements is $O((LR)^{d+1})$; the size of Y is $O(|\Lambda|/R^d)$, whence the total volume of motions g with $FS(h_g) < R$ is $O(R)$. ■

Summary.

1. A more realistic model of sensing returns data in Def which is the convolution of a constructible function with a smoothing kernel.
2. Sampling on a triangulation causes errors in the constructible category that do not dissipate under refinement. The passage to smooth data via convolution and integration with respect to $\lfloor d\chi\rfloor$ rectifies this problem.
3. Robustness of the integral $\int\cdot d\chi$ with respect to noise is extremely poor. A Lebesgue smoothing kernel (i.e., local averaging) mitigates this noise in $\lfloor d\chi\rfloor$.
4. Definable integrals are reducible (Theorem 4) to weighted Morse data, allowing a 'focusing' of the integral along (typically small) critical sets.

The development of a complete theory of numerical analysis for Euler integration remains. Additional applications of the definable Euler integral (e.g., to signal processing, valuation theory, and statistics) also await development.

ACKNOWLEDGMENTS. This work supported by DARPA # HR0011-07-1-0002 and ONR N000140810668.

1. Y. Baryshnikov and R. Ghrist, "Target enumeration via Euler characteristic integration," SIAM J. Appl. Math., 70(3), 2009, 825–844.
2. W. Blaschke, Vorlesungen über Integralgeometrie, Vol. 1, 2^{nd} edition. Leipzig and Berlin, Teubner, 1936.
3. L. Bröcker, "Euler integration and Euler multiplication," Adv. Geom. 5(1), 2005, 145–169.
4. L. Bröcker and M. Kuppe, "Integral geometry of tame sets," Geom. Dedicata 82, 2000, 285–323.
5. F. Chazal and A. Lieutier, "Weak feature size and persistent homology: computing homology of solids in $\mathbb{R}^n$ from noisy data samples," in Proceedings of the Twenty-First Annual Symposium on Computational Geometry, 2005, 255-262.
6. B. Chen, "On the Euler measure of finite unions of convex sets," Discrete and Computational Geometry 10, 1993, 79–93.
7. R. Cluckers and M. Edmundo, "Integration of positive constructible functions against Euler characteristic and dimension," J. Pure Appl. Algebra, 208, no. 2, 2006, 691–698.
8. R. Cluckers and F. Loeser, "Constructible motivic functions and motivic integration," Invent. Math., 173(1), 2008, 23–121.
9. M. Goresky and R. MacPherson, Stratified Morse Theory, Springer-Verlag, Berlin, 1988.
10. H. Groemer, "Minkowski addition and mixed volumes," Geom. Dedicata 6, 1977, 141–163.
11. H. Hadwiger, "Integralsätze im Konvexring," Abh. Math. Sem. Hamburg, 20, 1956, 136–154.
12. R. MacPherson, "Chern classes for singular algebraic varieties," Ann. of Math. 100, 1974, 423-432.
13. C. McCrory and A. Parusiński, "Algebraically constructible functions: real algebra and topology," Panoramas & Synthèses 24, Soc. Math. France 2007, 69–85.
14. J. Milnor, Morse Theory, Princeton University Press, 1963.
15. G.-C. Rota, "On the combinatorics of the Euler characteristic," Studies in Pure Mathematics, Academic Press, London, 1971, 221–233.
16. P. Schapira, "Operations on constructible functions," J. Pure Appl. Algebra 72, 1991, 83–93.

[2] For a set $U \subset \mathbb{R}^n$, feature size is defined as $FS(U) = FS(\mathbf{1}_U)$.

17. P. Schapira, "Tomography of constructible functions," in 11th Intl. Symp. on Applied Algebra, Algebraic Algorithms and Error-Correcting Codes, 1995, 427–435.

18. J. Schürmann, Topology of Singular Spaces and Constructible Sheaves, Birkhäuser, 2003.

19. L. Van den Dries, Tame Topology and O-Minimal Structures, Cambridge University Press, 1998.

20. O. Viro, "Some integral calculus based on Euler characteristic," Lecture Notes in Math., vol. 1346, Springer-Verlag, 1988, 127–138.