On the First Integral Conjecture of René Thom

JACKY CRESSON, ARIS DANIILIDIS & MASAHIRO SHIOTA

Abstract. More than half a century ago R. Thom asserted in an unpublished manuscript that, generically, vector fields on compact connected smooth manifolds without boundary can admit only trivial continuous first integrals. Though somehow unprecise for what concerns the interpretation of the word "generically", this statement is ostensibly true and is nowadays commonly accepted. On the other hand, the (few) known formal proofs of Thom's conjecture are all relying to the classical Sard theorem and are thus requiring the technical assumption that first integrals should be of class C^k with $k \geq d$, where d is the dimension of the manifold. In this work, using a recent nonsmooth extension of Sard theorem we establish the validity of Thom's conjecture for locally Lipschitz first integrals, interpreting genericity in the C^1 sense.

Key words. Structural stability, first integral, o-minimal structure, Sard theorem.

AMS Subject Classification. *Primary* 37C20 ; *Secondary* 34D30, 14P10.

1 Introduction

The purpose of this paper is to discuss the following conjecture attributed to René Thom (see [19], [8], [13], for example) which is part of the *folklore* in the dynamical systems community:

Thom conjecture: *For $1 \leq r \leq \infty$, C^r-generically vector fields on d-dimensional compact, smooth, connected manifolds without boundary do not admit nontrivial continuous first integrals.*

René Thom [19] proposes a scheme for a formal proof relying on the assumption that a C^r closing lemma ($r \geq 1$) is true [8]. The C^1-closing lemma (case $r = 1$) has indeed been proved by Pugh ([14], [15]). Nevertheless, very little is known for a C^r-closing lemma with $r \geq 2$ (see also [9], [16]) so that Thom's strategy should be revised.

In [13], Peixoto proves Thom conjecture for $r = 1$ assuming that first integrals are of class C^k with $k \geq d$, *i.e.* a relation between the regularity class of the first integrals to be considered and the dimension of the underlying compact manifold. As pointed out by Peixoto, this condition is only of technical nature and relates to the use of the classical Sard theorem in a crucial part of the proof. More precisely, denoting by $\mathcal{X}_M$ the set of C^1-vector fields on the d-dimensional compact manifold M, Peixoto's proof is divided in three steps:

Let $X \in \mathcal{X}_M$ and $f \colon M \to \mathbb{R}$ be a first integral of class C^d.

- By Sard's lemma, there exists in $f(M)$ an interval $]a, b[$ made up of regular values. For any $y \in]a, b[$, $f^{-1}(y)$ is an $(d-1)$-dimensional, compact, differentiable manifold, invariant under X.

- By Pugh's general density theorem for each $y \in]a, b[$, $f^{-1}(y)$ does not contain singularities or closed orbits of X since they are generic. As a consequence, singularities and closed orbits are all located at the critical levels of f.

- Any trajectory γ in $f^{-1}(y)$ is such that $\omega(\gamma) \subset f^{-1}(y)$ and cannot be contained in the closure of the set of singularities or closed orbits, in contradiction to Pugh's general density theorem.

T. Bewley [2] extends Peixoto's theorem for $1 \leq r \leq \infty$. The proof has then been simplified by R. Mañe [11]. Mañe's proof seems to have been rediscovered by M. Hurley [10]. However, the technical assumption of Peixoto (C^d-regularity of the first integrals) stays behind all these works, because of the use of Sard's theorem. Nevertheless, Thom conjecture seems to be true in general.

In this paper, we cancel the aforementioned regularity condition by interpreting the word "nontrivial" as "being essentially definable with respect to an o-minimal approximation" (see Definition 2.6 below). In this context we prove the validity of Thom's conjecture for $r = 1$ and for Lipschitz continuous first integrals.

The technic of the proof follows the strategy used by M. Artin and B. Mazur ([1]) to prove that generically the number of isolated periodic points of a diffeomorphism grows at most exponentially. Indeed, using a recently established nonsmooth version of Sard theorem (see [3, Theorem 7] or [4, Corollary 9]) and Peixoto's scheme of proof ([13]) we first show that in an o-minimal manifold (that is, a manifold that is an o-minimal set), generically, o-minimal first integrals are constant. Then by approximating every compact differentiable manifold by a Nash manifold we derive a general statement.

- **Preliminaries in dynamical systems.**

Given a C^1 manifold M we denote by $\mathcal{X}_M$ the space of all C^1-vector fields on M equipped with the C^1 topology. Let $\phi_t\colon M \to M$ be the one-parameter group of diffeomorphisms generated by a vector field X on M. A point $p \in M$ is called *nonwandering*, if given any neighborhood U of p, there are arbitrarily large values of t for which $U \cap \phi_t(U) \neq \emptyset$. Denoting by Ω the set of all nonwandering points we have the following genericity result due to C. Pugh ([14], [15]).

General density theorem (GDT): The set $\mathcal{G}_M$ of vector fields $X \in \mathcal{X}_M$ such that properties (G_1)–(G_4) below hold is residual in $\mathcal{X}_M$.

(G_1) X has only a finite number of singularities, all generic ;

(G_2) Closed orbits of X are generic ;

(G_3) The stable and unstable manifolds associated to the singularities and the closed orbits of X are transversal ;

(G_4) $\Omega = \bar{\Gamma}$, where Γ stands for the union of all singular points and closed orbits of X.

We use the following definition of a *first integral*:

Definition 1.1 (First integral)**.** A first integral of a vector field X on a compact connected manifold M of dimension d is a continuous function $f\colon M \to \mathbb{R}$ which is constant on the orbits of the flow generated by X but it is not constant on any nonempty open set of M.

As mentioned in the introduction, Peixoto [13] only considers C^k-first integrals with $k \geq d$.

- **Preliminaries in o-minimal geometry.**

Let us recall the definition of an o-minimal structure (see [7] for example).

Definition 1.2 (o-minimal structure)**.** An o-minimal structure on the ordered field $\mathbb{R}$ is a sequence of Boolean algebras $\mathcal{O} = \{\mathcal{O}_n\}_{n\geq 1}$ such that for each $n \in \mathbb{N}$

(i) $A \in \mathcal{O}_n \Longrightarrow A \times \mathbb{R} \in \mathcal{O}_{n+1}$ and $\mathbb{R} \times A \in \mathcal{O}_{n+1}$;

(ii) $A \in \mathcal{O}_{n+1} \Longrightarrow \Pi(A) \in \mathcal{O}_n$
($\Pi\colon \mathbb{R}^{n+1} \to \mathbb{R}^n$ denotes the canonical projection onto $\mathbb{R}^n$) ;

(iii) $\mathcal{O}_n$ contains the family of algebraic subsets of $\mathbb{R}^n$, that is, the sets of the form

$$\{x \in \mathbb{R}^n : p(x) = 0\},$$

where $p\colon \mathbb{R}^n \to \mathbb{R}$ is a polynomial function ;

(iv) $\mathcal{O}_1$ consists exactly of the finite unions of intervals and points.

An important example of o-minimal structure is the collection of *semialgebraic sets* (see [6] for example), that is, sets that can be obtained by Boolean combinations of sets of the form

$$\{x \in \mathbb{R}^n : \, p(x) = 0, \, q_1(x) < 0, \ldots, \, q_m(x) < 0\}\,,$$

where $p, q_1, \ldots, q_m$ are polynomial functions in $\mathbb{R}^n$. Indeed, properties (i), (iii) and (iv) of Definition 1.2 are straightforward, while (ii) is a consequence of the Tarski-Seidenberg principle.

A subset A of $\mathbb{R}^n$ is called *definable* (in the o-minimal structure $\mathcal{O}$) if it belongs to $\mathcal{O}_n$. Given any $S \subset \mathbb{R}^n$ a mapping $F\colon S \to \mathbb{R}$ is called *definable* in $\mathcal{O}$ (respectively, semialgebraic) if its graph is a definable (respectively, semialgebraic) subset of $\mathbb{R}^n \times \mathbb{R}$.

- **Preliminaries in variational analysis.**

Let $g\colon U \to \mathbb{R}$ be a Lipschitz continuous function where U is a nonempty open subset of $\mathbb{R}^d$. The *generalized derivative* of g at x_0 in the direction $v \in \mathbb{R}^n$ is defined as follows (see [5, Section 2] for example):

$$g^o(x_0, e) = \limsup_{\substack{x\to x_0 \\ t\searrow 0^+}} \frac{g(x+te) - g(x)}{t} \tag{1}$$

where $t \searrow 0^+$ indicates the fact that $t > 0$ and $t \to 0$. It turns out that the function $v \mapsto g^o(x_0, v)$ is positively homogeneous and convex, giving rise to the *Clarke subdifferential* of g at x_0 defined as follows:

$$\partial g(x_0) \; = \; \{\, x^* \in \mathbb{R}^d : \, g^o(x_0, v) \geq \langle x^*, v\rangle, \, \forall v \in \mathbb{R}^d\}. \tag{2}$$

In case that g is of class C^1 (or more generally, strictly differentiable at x_0) it follows that

$$\partial g(x_0) = \{\nabla g(x_0)\}.$$

A point $x_0 \in U$ is called *Clarke critical*, if $0 \in \partial g(x_0)$. We say that $y_0 \in g(U)$ is a *Clarke critical value* if the level set $g^{-1}(y_0)$ contains at least one Clarke critical point. Given a Lipschitz continuous function $f\colon M \to \mathbb{R}$ defined on a C^1 manifold M we give the following definition of (nonsmooth) critical value.

Definition 1.3 (Clarke critical value)**.** We say that $y_0 \in f(M)$ is a Clarke critical value of the function $f\colon M \to \mathbb{R}$, if there exists $p \in f^{-1}(y_0)$ and a local chart (φ, U) around p such that $0 \in \partial(f \circ \varphi^{-1})(\varphi(p))$. In this case, $p \in M$ is a Clarke critical point for f.

It can be easily shown (see [17, Exercise 10.7], for example) that the above definition does not depend on the choice of the chart.

2 Main results

Throughout this section M will be a C^1 compact connected submanifold of $\mathbb{R}^n$ (without boundary), TM its corresponding tangent bundle and $\mathcal{X}_M$ the space of C^1 vector fields on M equipped with the C^1 topology. Let us recall that submanifolds of $\mathbb{R}^n$ admit ε-tubular neighborhoods for all $\varepsilon > 0$ sufficiently small. We further consider a C^1-submanifold N of $\mathbb{R}^n$, a C^1-diffeomorphism $F\colon M \to N$ and $\varepsilon > 0$.

Definition 2.1 (Approximation of a manifold)**.** (i) We say that (N, F) is a C^1-approximation of M (of precision ϵ), if N belongs to an ϵ-tubular neighborhood U_ϵ of M and F can be extended to a C^1-diffeomorphism $\tilde{F}\colon \mathbb{R}^n \to \mathbb{R}^n$, which is isotopic to the identity id, satisfies $\tilde{F}|_{\mathbb{R}^n \setminus U_r} \equiv \mathrm{id}$ and

$$\max_{x\in\mathbb{R}^n} \left\{ ||\tilde{F}(x) - x|| + ||\, d\tilde{F}(x) - \mathrm{id}\,|| \right\} < \epsilon.$$

(We shall use the notation $\tilde{F} \sim_\epsilon \mathrm{id}$ to indicate that $\tilde{F}$ is ϵ-C^1-closed to the identity mapping.)

(ii) A C^1-approximation (N, F) of M is called semialgebraic (respectively, definable) if the manifold N is a semialgebraic subset of $\mathbb{R}^n$ (respectively, a definable set in an o-minimal structure).

In the sequel, we shall need the following approximation result.

Lemma 2.2 (Semialgebraic approximation)**.** *Let M be a C^1 compact submanifold of $\mathbb{R}^n$. Then for every $\varepsilon > 0$, there exists a semialgebraic ε-approximation of M.*

Proof. Fix $\varepsilon > 0$ and let U be an open ϵ-tubular neighborhood of M in $\mathbb{R}^n$ for some $\epsilon \in (0, \varepsilon)$. Applying [18, Theorem I.3.6] (for $A = \mathbb{R}^n$ and $B = C^1$), we deduce the existence of a C^1-embedding F of M into U which is ϵ-close to the identity map id in the C^1 topology such that $F(M) = N$ is a Nash manifold (that is, N is a C^∞-manifold and a semialgebraic set). Then F can be extended to a C^1 diffeomorphism $\tilde{F}$ of $\mathbb{R}^n$ by a partition of unity of class C^1 such that $\tilde{F} = \mathrm{id}$ on $\mathbb{R}^n \setminus U$. Moreover there exists a C^1 isotopy $\{F_t\}_{t\in[0,1]}$, such that $F_t = \mathrm{id}$ on $\mathbb{R}^n \setminus U$, $F_0 \equiv \mathrm{id}$ and $F_1 = \tilde{F}$ and the map $F_t\colon \mathbb{R}^n \times [0, 1] \to \mathbb{R}^n$ is ϵ-close to the projection to $\mathbb{R}^n$ in the C^1 topology. □

Given a C^1-manifold M and a C^1-vector field $X \in \mathcal{X}_M$, the following result relates generic singularities of hyperbolic type of X with Clarke critical values of Lipschitz continuous first integrals of X.

Lemma 2.3 (Location of singularities)**.** *Assume $f\colon M \to \mathbb{R}$ is a Lipschitz continuous first integral for the vector field $X \in \mathcal{X}_M$. Then all generic singularities and all closed orbits of hyperbolic type are located at the Clarke-critical level sets.*

Proof. Let p_0 be either a singular point of hyperbolic type or any point of a closed orbit of hyperbolic type and let $\pi\colon U \to M$ be the exponential mapping around $p_0 = \pi(0)$, where U is an open neighborhood of $0 \in T_{p_0}M \cong \mathbb{R}^d$ (d denoting the dimension of M). It follows that the function $g = f \circ \pi$ is Lipschitz continuous. Since the stable and unstable manifolds of the flaw of the field X at p_0 are transversal and since f is a first integral, it follows that for some basis $\{e_i\}_{i\in\{1,\dots,d\}}$ of $T_{p_0}M$ one has

$$g^o(0, \pm e_i) \geq g'(0, \pm e_i) := \lim_{t \searrow 0+} \frac{g(\pm t e_i) - g(0)}{t} = 0,$$

where $g^0(0, \cdot)$ is given by (1). In view of (2) we deduce that $0 \in \partial g(0)$, hence $f(p_0)$ is a critical value of f. □

Lemma 2.4 (Density of critical values for GDT fields)**.** *In the situation of Lemma 2.3, let us further assume that $X \in \mathcal{G}_M$. Then the Clarke critical values of f are dense in $f(M)$.*

Proof. Let Γ denote the union of all singular points and closed orbits of the field X. Since $X \in \mathcal{G}_M$, it follows from Lemma 2.3 that the set $f(\Gamma)$ is included to the Clarke critical values. Continuity of f and compactness of M yield $\overline{f(\Gamma)} = f(\overline{\Gamma}) = f(\Omega)$. Since Ω contains all ω-limits of orbits of X, taking any $y \in f(M)$ and any $x \in f^{-1}(y) \diagdown \Gamma$ we denote by γ the orbit passing through x and by γ_∞ the set of ω-limits of γ. Then by continuity $y = f(\gamma) = f(\gamma_\infty) \subset f(\Omega)$. This proves the assertion. □

Corollary 2.5 (Thom conjecture – definable version)**.** *Let $X \in \mathcal{G}_M$. Then X does not admit any Lipschitz continuous definable first integral.*

Proof. Assume f is a Lipschitz continuous first integral of X on M and denote by S the set of its Clarke critical points. If f is o-minimal, then so is M (cf. property (ii) of Definition 1.2), the tangential mappings $\pi\colon U \subset T_{p_0}M \to M$ (around any point $p_0 \in M$) and the composite functions of the form $g = f \circ \pi$ (notation according to the proof of Lemma 2.3). Note that p is a critical point of f if and only if $\pi^{-1}(p)$ is a critical point of $g = f \circ \pi$ where π is any tangential mapping with $p \in \pi(U)$. Applying [4, Corollary 8] we deduce that the set of Clarke critical values of each function g is of measure zero. Using a standard compactness argument we deduce that $f(S)$ is of measure zero, thus in particular $f(M) \setminus f(S)$ contains an interval (y_1, y_2). But this contradicts the density result of Lemma 2.4. □

If the manifold M is not a definable subset of $\mathbb{R}^n$ the above result holds vacuously and gives no information. To deal with this case, the forthcoming notion of essential o-minimality with respect to a given o-minimal approximation turns out to be a useful substitute for our purposes. Let us fix $\epsilon > 0$ and a definable ϵ-approximation (N, F) of M.

Definition 2.6 (Essential o-minimality with respect to a definable approximation)**.** A mapping $f\colon M \to \mathbb{R}$ is called essentially o-minimal with respect to a definable approximation (N, F) of M if the mapping $f \circ F^{-1}\colon N \to \mathbb{R}$ is o-minimal.

Note that every o-minimal function on M is essentially o-minimal with respect to any approximation (N, F) of M for which the diffeomorphism F is o-minimal. Setting

$$M = \{p \in \mathbb{R}^2 : p \in \text{Graph}(h)\}$$

where $h(t) = t^3 \sin(x^{-1})$, if $t \neq 0$ and 0 if $t = 0$, we obtain a nondefinable C^1-submanifold of $\mathbb{R}^2$. Thus the (projection) function $f\colon M \to \mathbb{R}$ with $f((t, h(t)) = t$ is not o-minimal. It can be easily seen that for every $\epsilon > 0$ there exists an ϵ-approximation (N, F) with respect to which f is essentially o-minimal. On the other hand, if χ_K is the characteristic function of the Cantor set K of $(0, 1)$, the function $g\colon \mathbb{R} \to \mathbb{R}$ defined by $g(x) = \int_0^x \chi_K(t)\, dt$ for all $x \in \mathbb{R}$ is not essentially o-minimal with respect to any approximation. Roughly speaking, a function that is not essentially o-minimal contains intrinsic irreparable oscillations.

In view of Lemma 2.2, for every $\epsilon > 0$ there exists a C^1 definable manifold N and a diffeomorphism $F\colon M \to N$ such that $F \sim_\epsilon \mathrm{id}$. Fixing the approximation, we associate to every vector field $X\colon M \to TM$ on M the conjugate C^1-vector field $\tilde{X}\colon N \to TN$ on N defined as follows:

$$\tilde{X}(q) = dF(F^{-1}(q), X(F^{-1}(q)).$$

Note that $\tilde{X}$ is uniquely determined by X. Let us further denote by $\mathcal{G}_N$ the vector fields of N that satisfy the generic GDT assumptions. We are ready to state our main result.

Theorem 2.7 (Genericity of non-existence of first integrals)**.** *Let M be a C^1 compact submanifold of $\mathbb{R}^n$ and $\epsilon > 0$. For the C^1 topology, the set of vector fields in M that do not admit Lipschitz continuous first integrals which are essentially o-minimal with respect to a given definable ϵ-approximation of M is generic.*

Proof. Let us fix any definable ϵ-approximation of M and let us denote by $\mathcal{G}_N$ the vector fields of N that satisfy the generic GDT assumptions. By Pugh's density theorem $\mathcal{G}_N$ is a C^1-residual subset of $\mathcal{X}_N$ and by Corollary 2.5, if $Y \in \mathcal{G}_N$ then Y does not possess any o-minimal Lipschitz continuous first integral. Let $\mathcal{G}$ denote the set of vector fields of X that conjugate inside $\mathcal{G}_N$, that is,

$$\mathcal{G} = \{X \in \mathcal{X}_M :\ \tilde{X} \in \mathcal{G}_N\}.$$

Then $\mathcal{G}$ is residual in $\mathcal{X}_M$. Pick any $X \in \mathcal{G}$ and assume that $f\colon M \to \mathbb{R}$ is a Lipschitz continuous first integral of X. Since the trajectories of X are transported to the trajectories of $\tilde{X} \in \mathcal{G}_N$ through the mapping F, it follows that $\tilde{f} = f \circ F$ is a first integral of $\tilde{X}$ in N. This shows that f cannot be essentially o-minimal with respect to (N, F). □

Acknowledgment. Part of this work has been made during a research visit of the second author to the University of Pau (June 2007). The second author wishes to thank his hosts for hospitality and Olivier Ruatta (University of Limoges) for useful discussions. The first author thanks Jean-Pierre Bourguignon for his help with R. Thom's archive and J-P. Françoise for useful references.

References

[1] Artin M. & Mazur B., On Periodic Points, *The Annals of Mathematics* (Series 2) **81** (1965), 82–99.

[2] Bewley T., *Non genericity of first integrals*, PhD Thesis, University of California (Berkeley, 1971).

[3] Bolte, J., Daniilidis, A., Lewis, A. & Shiota, M., Clarke critical values of subanalytic Lipschitz continuous functions, *Ann. Polon. Math.* **87** (2005), 13–25.

[4] Bolte, J., Daniilidis, A., Lewis, A. & Shiota, M., Clarke subgradients of stratifiable functions, *SIAM J. Optimization* **18** (2007), 556–572.

[5] Clarke, F., *Optimization and Nonsmooth Analysis*, Wiley-Interscience (New York, 1983).

[6] Coste, M., *An introduction to semi-algebraic geometry*, RAAG Notes, 76 pages, Institut de Recherche Mathématiques de Rennes, November 1999.

[7] van den Dries, L. & Miller, C., Geometric categories and o-minimal structures, *Duke Math. J.* **84** (1996), 497–540.

[8] Françoise, J.-P., *Géométrie analytique et systèmes dynamiques*, Cours de troisième cycle, Presses Universitaires de France (Paris, 1995).

[9] Gutierrez C., A counter-example to a C^2-closing lemma, *Erg. Th. and Dyn. Sys.* **7** (1987), 509–530.

[10] Hurley, M., On the generic nonexistence of first integrals, *Proc. Amer. Math. Soc.* **98** (1986), 142–144.

[11] Mañe, R., A note on first integrals of vector fields and endomorphisms, *Bol. Soc. Brasil. Mat.* **3** (1973), 127–132.

[12] Peixoto, M., On structural stability, *Ann. of Math.* **69** (1959), 199–222.

[13] Peixoto M., Qualitative theory of differential equations and structural stability. *Differential Equations and Dynamical Systems (Proc. Internat. Sympos., Mayaguez, P.R., 1965)*, pp. 469–480 (Academic Press, 1967).

[14] Pugh C., The closing lemma, *Am. J. Math.* **89** (1967), 956–1009.

[15] Pugh C., An improved closing lemma and a generic density theorem, *Am. J. Math.* **89** (1967), 1010–1021.

[16] Pugh C., Against the C^2-closing lemma, *J. Differential Equations* **17** (1975), 435–443.

[17] Rockafellar, R.T. & Wets, R., *Variational Analysis*, Grundlehren der Mathematischen, Wissenschaften, Vol. **317**, (Springer, 1998).

[18] Shiota, M., *Nash manifolds*, Lecture Notes in Mathematics **1269** (Springer, 1987).

[19] Thom, R., Sur les intégrales premieres d'un systemes différentiel sur une variété compacte, *Unpublished manuscript*.

Jacky CRESSON

Laboratoire de Mathématiques Appliquées
Université de Pau et des Pays de l'Adour
avenue de l'université, F-64000 Pau, France.

E-mail: jacky.cresson@univ-pau.fr
http://web.univ-pau.fr/~jcresson/

Aris DANIILIDIS

Departament de Matemàtiques, C1/308
Universitat Autònoma de Barcelona
E-08193 Bellaterra, Spain.

E-mail: arisd@mat.uab.es
http://mat.uab.es/~arisd

Research supported by the MEC Grant No. MTM2005-08572-C03-03 (Spain).

Masahiro SHIOTA

Department of Mathematics
Nagoya University (Furocho, Chikusa)
Nagoya 464-8602, Japan.

E-mail: shiota@math.nagoya-u.ac.jp