Research statement

Rodolfo Ríos-Zertuche

Overview

My doctoral studies, under the supervision of Andrei Okounkov, brought me in close contact with the world of dynamics in Teichmüller space, polygonal billiards, and the representation theory at work in the study of the action of $\mathrm{SL}_2(\mathbb{R})$ in this space. This is described in Section 3.

During those studies, I was in constant communication with John N. Mather. I had sought his advice as I applied for doctoral studies, and it was only his sudden illness that ultimately led me to the hard decision of not writing my thesis with him. After my Ph.D. thesis on quadratic differentials, I came back to his line of research, developing the theory of the calculus of variations in the context of closed measures, which is described in Section 2. This has connections with mathematical physics and dynamical systems, especially Lagrangian dynamics, optimal control, and mass transport.

More recently, under the advice of Jérôme Bolte and Edouard Pauwels, I have been able to apply the theory of closed measures also to the analysis of the dynamics and convergence of optimization algorithms, a line of research described in Section 1.1.

After that, I moved on to work on global polynomial optimization methods with Milan Korda, Jean-Bernard Lasserre, and Victor Magron, with projects described in Section 1.2.

1 Algorithm analysis

1.1 First order optimization methods

The advent of huge scale non-convex, non-smooth, and poorly structured optimization problems (e.g. deep learning) has triggered a revival of interest in the explicit subgradient method of Shor [39], whose iterates $\{x_n\}_n$ satisfy

$$x_{n+1} = x_n + \varepsilon_n \theta_n$$

for some θ_n in the Clarke subdifferential [9] of the objective function and small $0 < \varepsilon_n \searrow 0$.

The study of its dynamics turns out to be much more delicate than implicit versions (like the proximal gradient and other proximal decomposition methods). It features indeed highly oscillatory behavior even with rigidity assumptions like semi-algebraicity of the objective function. This makes it hard to directly apply, for the discrete flow, many of the tools that have been developed for the analysis of the continuous flow [14], such as the Kurdyka–Łojasiewicz inequality [7, 24] and the theory of asymptotic pseudotrajectories [3, 4].

In the work [8], we analyze the algorithm under the fairly weak assumption that the objective function is Lipschitz continuous and path-differentiable, and we prove some convergence properties. This shows that a rather weak hypothesis is enough to avoid the very pathological behavior illustrated by recent examples [13] in the Lipschitz continuous case.

To obtain this result we use closed measures, a tool taken from geometric measure theory and further described in Section 2.1, that we expect to be useful also for the analysis of convergence of other algorithms involving sequences in Euclidean spaces.

The paper [8] is accompanied by our other paper [37] that gives very pathological counterexamples that show that the results of [8] are sharp in several senses.

With Pascal Bianchi, we further generalized [6] the arguments in [8] to explain the broader contexts already considered by [3, 4].

1.2 Polynomial optimization

In most applications, the data are semi-algebraic, and in this cases variational and optimal control problems can be attacked [23] using algorithms based in the so-called Lasserre hierarchy of sums of square polynomials. This, however, requires a relaxation of the problem to allow for an occupation measure-valued solution. In our work [38] with Milan Korda, we analyse the possible gap between the classical solutions and those in the (larger) space of occupation measures.

The long-sought solution of the problem depends on the dimensions of the problem: under very lax assumptions, the most important of which being that the integrand be convex in the velocities, there is no gap in the dimension 1 and in the codimension 1 cases, and that there may be a positive gap in all other cases. To prove the no-gap result, we show that every occupation measure in codimension one can be decomposed as a convex superposition of Lipschitz functions. The positive gap result is shown by means of a counterexample in dimension 4, that looks a lot like the Riemann surface of the square-root, and the proof leverages the Poincaré–Wirtinger inequality in a clever way.

2 Calculus of variations done with closed measures

The theory of calculus of variations can be developed in many different contexts, each with its advantages and disadvantages. Following the line of research of Mather [29,40], Fathi–Siconolfi [17,18], Mañé [10,28], and others, I have focused on the so-called *closed measures*. These carry more information than currents and varifolds, the objects that the school of De Giorgi has focused on [19], as they encode "parameterization speeds," and hence constitute the natural home to a rich theory that includes all of classical mechanics.

Closed measures μ are the weak* closure of occupation measures of the space of measures on the tangent bundle TM induced by the jets (γ, γ') of closed curves γ, $\gamma(0) = \gamma(T)$, and the problem of minimizing an integral of the form

$$\frac{1}{T}\int_0^T L(\gamma(t), \gamma'(t))dt$$

can be replaced by the minimization of an integral of the form

$$\int_{TM} L\, d\mu,$$

with many advantages, including compactness properties.

By definition, a Radon probability measure μ on the tangent bundle TM is closed if $\int_{TM} df\, d\mu = 0$ for all $f \in C^\infty(M)$.

In this context, we have strived to understand what the minimizable functionals look like (Section 2.1), what the variations of closed measures look like and the properties of critical closed measures and extremals (Section 2.2), what the link is to the so-called weak KAM theory and how it can be generalized to higher-dimensional contexts (Section 2.3). Leveraging the understanding of critical points derived in those works, we have also endeavoured to bridge the theory with quantum mechanics using Feynman path integrals.

2.1 Minimizable Lagrangian densities

In the case of a Tonelli Lagrangian L, the study of measures minimizing the action functional $\mu \mapsto \int_{TM} L\, d\mu$ has produced substantial fruits, such as an equivalence with the dynamics of twist maps [20], the existence of minimizers

[12], a very developed theory of the Hamilton-Jacobi equation through weak KAM theory [17] and viscosity solutions [11], interesting geometric properties related to the Lyapunov exponents [2], and the Mather–Aubry theory [29, 40], to name a few.

Characterization. We depart from the strictly-Tonelli context, and we find in [36] that the (not necessarily convex) Lagrangian densities L such that the integral $\int_{TM} L\,d\mu$ reaches its minimum in the space of closed measures, can be written as

$$L(x,v) = c + df_x(v) + g(x,v),$$

for a real number c, an exact form df, and a nonnegative function g that vanishes along the support of the minimizer. In the hardest part of the proof, we also prove that df is Lipschitz on the support of the minimizer. Here, f can be understood as a subsolution of the Hamilton-Jacobi equation and $-c$ is Mañé's critical value $-\inf_\gamma \frac{1}{T}\int_0^T L(\gamma(t),\gamma'(t))dt$.

This unifies and generalizes John Mather's Lipschitz lamination result [29,40], the regularity theory of critical subsolutions of the Hamilton-Jacobi equation [5], and the weak KAM theory [17]. The result is sharp in several ways.

This characterization thus gives light to the ubiquitous occurence of the Hamilton-Jacobi equation, while also highlighting the less fundamental nature of the Euler-Lagrange equation and of the Hamiltonian flow, which many minimizers do not follow and which are sometimes not even possible to define. The impossibility arises from the level of generality in which we work — we consider Lagrangian densities that do not, in general, satisfy the hypotheses that are generally used in existence theorems, such as being Tonelli, convex, superlinear, coercive, etc.

Optimal control. A similar characterization [36] for the case of certain optimal control problems also unifies the above with the Pontryagin Maximum Principle and the Hamilton-Jacobi-Bellman equation.

For the Pontryagin Maximum Principle [30], we have a new, concise proof using dual convex cones and, with appropriate conditions, we are able to use the Lipschitz continuity regularity result for df to show that this principle holds for all times t rather than (as is proved almost everywhere else, it seems) for almost all times t.

For the Hamilton-Jacobi-Bellman equation, we deduce that the exact form from the decomposition of L corresponds to a function that is a critical subsolution to this equation, and we prove a $C^{1,1}$ regularity result for the solutions in certain regions.

Finsler metrics. Additionally, we were able to derive some results about the regularity of the distance function associated to a non-strictly-convex Finsler metric. We were able to prove that the distance to a closed set is $C^{1,1}$ away from its obvious singularities, a result that also turned out to be true in the context of non-strictly convex, not necessarily positive, C^2 1-homogeneous functions on TM, as long as they are bounded from below by an exact form. This generalizes results of [26, 27].

2.2 Variations of closed measures

Main result. Given a closed measure μ, which may encode for example a curve, we characterize [31] all the possible families $(\mu_t)_{t\geqslant 0}$ of weakly-differentiable deformations of $\mu = \mu_0$. To do this, we determine the distributions that arise as weak derivatives $d\mu_t/dt|_{t=0}$ of these families, and we use an abstract convex-geometrical result [32] to prove the converse existence statement.

This allows us to identify directions such that when the action is critical as we deform in those directions, that is, if

$$\frac{d}{dt}\int_{TM} L\,d\mu_t\bigg|_{t=0} = 0, \tag{1}$$

this implies, separately:

1. energy conservation,
2. the momentum being given by an exact form,
3. the Lipschitz continuity of the momenta (as in Mather's result [29]), and
4. the Euler-Lagrange equations are satisfied weakly.

In other words, we find subspaces of the tangent bundle to the space of closed measures such that, if (1) holds for all deformations $(\mu_t)_{t\geqslant 0}$ with derivative in that subspace, then the measure must have the corresponding property 1–4. To the best of our knowledge, all preceding texts used a very coarse definition of criticality that implied all of these properties simultaneously; we find that directionally restrictive definitions of criticality may give rise to very specific sets of properties. These are prone to appear, for example, in optimisation with problems with constraints.

We are also able to study in detail the cases in which the Euler-Lagrange equation is applicable, including an almost-variational characterization of measures invariant under the flow it induces on TM. It turns out that there are examples of extremals that do not follow the Euler-Lagrange equation, so a purely-variational characterization is impossible. However, with a natural regularity requirement, we find a relatively small family of directions such that when the action is critical as we deform in those directions, the measure needs to be invariant under the flow induced by the Euler-Lagrange equations on the tangent bundle.

This allows us to recover and generalize the result of R. Mañé that minimizing closed measures are invariant under the Euler-Lagrange flow [10, 28]. Since our cleaner, shorter proof is completely non-dynamical, we are able to extend this result to much more general contexts (we only need the fiberwise Hessian of the Lagrangian to be locally invertible).

Mass transport. In [32] we study the continuity equation in mass transport. The paper characterizes the distributions (in the sense of Schwartz) that are derivatives of weakly-differentiable deformations of probability measures.

While in the case of families absolutely-continuous with respect to the Wasserstein metric the answer is that probabilities can only be flowed along vector fields, slightly less regular families display much richer behavior: they can give rise to distributions of arbitrary degree as their derivative. We thus find that the tangent bundle to the space of probabilities is much larger than the one proposed by Ambrosio–Gigli–Savaré [1] if we allow weakly differentiable curves, rather than only Wasserstein absolutely continuous curves.

Using Colombeau algebras, we also develop a way to understand those distributions as giving a "direction of the movement" in the way a vector field V describes the direction in which the family of densities ρ_t flows when following the continuity equation

$$\frac{d\rho_t}{dt} + \mathrm{div}\rho_t V_t = 0.$$

2.3 Higher-dimensional weak KAM theory

For 1-dimensional objects, the Fathi-Siconolfi weak KAM theory [17] nonperturbatively relates the viscosity solutions of the Hamilton-Jacobi equation, the quasiperiodic properties of minimizers in different homology classes, and the dynamics of orbits close to those minimizers.

We found [33] a generalization of weak KAM to the case of higher-dimensional minimizers. Inspired from the discovery that the criticality with respect to certain variations implied that the momenta were given by an exact form, we tried to look for a general theory that looked more like the one that had been developed for the one-dimensional case. Minimizing candidates of dimension n are encoded as closed measures on the sum

$$T^n M = \underbrace{TM \oplus TM \oplus \cdots \oplus TM}_{n}.$$

The functional they minimize is of the form $\mu \mapsto \int_{T^n M} L\, d\mu$ for a Lagrangian density $L\colon T^n M \to \mathbb{R}$ such as the n-dimensional volume.

The theory we found entails a suitable definition of "slices of cobordisms," and it produces a fixed point weak KAM solution of the action of the Lax-Oleinik semigroup defined there, and it shows that in many circumstances that solution descends to a calibrating differential form of order n on the manifold M, thereby completing the analogy with the 1-dimensional case.

3 Representation theory: quadratic differentials

For my thesis [34], I worked on the problem of the computation of the volumes of the moduli space of quadratic differentials on a Riemann surface. This was a path already well-trodden [15, 16, 22] since the problem was first motivated [21] as a way to obtain information about the Lyapunov exponents of the $\mathrm{SL}_2(\mathbb{R})$ action on Teichmüller space.

The main result of the thesis is a structural formula for the volumes of the strata of the moduli space of quadratic differentials on a Riemann surface, relating them to expectations of a certain point process, and several related results.

Another contribution is a combinatorial formula for the characters of near-involutions in the symetric group that allows for the study of those point processes, as well as their asymptotic computation for large degrees.

A later spawn was an expository article [35] that gives a thorough introduction to a 0-dimensional quantum field theory that encodes important information about the representations of symmetric groups of asymptotically infinite order.

References

[1] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows in metric spaces and in the space of probability measures.* Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.

[2] Marie-Claude Arnaud. The link between the shape of the irrational Aubry-Mather sets and their Lyapunov exponents. *Annals of Mathematics*, pages 1571–1601, 2011.

[3] Michel Benaïm, Josef Hofbauer, and Sylvain Sorin. Stochastic approximations and differential inclusions. *SIAM Journal on Control and Optimization*, 44(1):328–348, 2005.

[4] Michel Benaïm, Josef Hofbauer, and Sylvain Sorin. Stochastic approximations and differential inclusions, Part II: Applications. *Mathematics of Operations Research*, 31(4):673–695, 2006.

[5] Patrick Bernard. Existence of $C^{1,1}$ critical sub-solutions of the Hamilton–Jacobi equation on compact manifolds. In *Annales scientifiques de l'Ecole normale supérieure*, volume 40, pages 445–452. Elsevier, 2007.

[6] Pascal Bianchi and Rodolfo Ríos-Zertuche. A closed-measure approach to stochastic approximation. arXiv:2112.05482 [math.OC].

[7] Jérôme Bolte, Aris Daniilidis, Adrian Lewis, and Masahiro Shiota. Clarke subgradients of stratifiable functions. *SIAM J. Optim.*, 18(2):556–572, 2007.

[8] Jérôme Bolte, Edouard Pauwels, and Rodolfo Rios-Zertuche. Long term dynamics of the subgradient method for Lipschitz path differentiable functions. *Journal of the European Mathematical Society*, 2022.

[9] F. H. Clarke. *Optimization and nonsmooth analysis*, volume 5 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 1990.

[10] Gonzalo Contreras and Renato Iturriaga. *Global minimizers of autonomous Lagrangians*. 22º Colóquio Brasileiro de Matemática. [22nd Brazilian Mathematics Colloquium]. Instituto de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 1999.

[11] Michael G. Crandall and Pierre-Louis Lions. Viscosity solutions of Hamilton-Jacobi equations. *Trans. Amer. Math. Soc.*, 277(1):1–42, 1983.

[12] B. Dacorogna. *Direct Methods in the Calculus of Variations*. Springer-Verlag, Berlin, Heidelberg, 1989.

[13] Aris Daniilidis and Dimitriy Drusvyatskiy. Pathological subgradient dynamics. Preprint. arXiv:1910.13604 [math.OC].

[14] Damek Davis, Dmitriy Drusvyatskiy, Sham Kakade, and Jason D. Lee. Stochastic subgradient method converges on tame functions. *Found. Comput. Math.*, 20(1):119–154, 2020.

[15] A. Eskin and A. Okounkov. Pillowcases and quasimodular forms. In *Algebraic geometry and number theory*, volume 253 of *Progr. Math.*, pages 1–25. Birkhäuser Boston, Boston, MA, 2006.

[16] Alex Eskin and Andrei Okounkov. Asymptotics of numbers of branched coverings of a torus and volumes of moduli spaces of holomorphic differentials. *Invent. Math.*, 145(1):59–103, 2001.

[17] Albert Fathi. Weak KAM theorem in Lagrangian dynamics. Preliminary Version Number 10, June 2008.

[18] Albert Fathi and Ezequiel Maderna. Weak KAM theorem on non compact manifolds. *NoDEA Nonlinear Differential Equations Appl.*, 14(1-2):1–27, 2007.

[19] Herbert Federer. *Geometric measure theory*. Die Grundlehren der mathematischen Wissenschaften, Band 153. Springer-Verlag New York Inc., New York, 1969.

[20] Moser Jürgen. Recent developments in the theory of Hamiltonian systems. *SIAM review*, 28(4):459–485, 1986.

[21] M. Kontsevich. Lyapunov exponents and Hodge theory. In *The mathematical beauty of physics (Saclay, 1996)*, volume 24 of *Adv. Ser. Math. Phys.*, pages 318–332. World Sci. Publ., River Edge, NJ, 1997.

[22] M. Kontsevich and A. Zorich. Connected components of the moduli spaces of Abelian differentials with prescribed singularities. *Inventiones Mathematicae*, 153:631–678, 2003. 10.1007/s00222-003-0303-x.

[23] Milan Korda, Didier Henrion, and Jean-Bernard Lasserre. Moments and convex optimization for analysis and control of nonlinear partial differential equations. Preprint. arXiv:1804.07565 [math.OC].

[24] Krzysztof Kurdyka. On gradients of functions definable in o-minimal structures. In *Annales de l'institut Fourier*, volume 48, pages 769–783, 1998.

[25] Monique Laurent and Lucas Slot. An effective version of Schmüdgen's Positivstellensatz for the hypercube. *arXiv preprint arXiv:2109.09528*, 2021.

[26] YanYan Li and Louis Nirenberg. The regularity of the distance function to the boundary. arXiv:math/0510577 [math.AP].

[27] YanYan Li and Louis Nirenberg. The distance function to the boundary, Finsler geometry, and the singular set of viscosity solutions of some Hamilton-Jacobi equations. *Communications on Pure and Applied Mathematics*, 58(1):85–146, 2005.

[28] Ricardo Mañé. *Ergodic theory and differentiable dynamics*, volume 8 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1987. Translated from the Portuguese by Silvio Levy.

[29] John N. Mather. Action minimizing invariant measures for positive definite Lagrangian systems. *Math. Z.*, 207(2):169–207, 1991.

[30] L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze, and E.F. Mishchenko. *The mathematical theory of optimal processes.* Interscience Publishers, John Wiley & Sons Inc. New York-London, 1962. Translated from the Russian by KN Trirogoff.

[31] Rodolfo Ríos-Zertuche. Deformations of closed measures and variational characterization of measures invariant under the Euler-Lagrange flow. Preprint, 2018. arXiv:1810.07838 [math.OC].

[32] Rodolfo Ríos-Zertuche. Existence of differentiable curves in convex sets and the concept of direction of the flow in mass transportation. Preprint. arXiv:1810.05999 [math.OC].

[33] Rodolfo Ríos-Zertuche. Weak KAM theory in higher-dimensional holonomic measure flows. Preprint. arXiv:1901.08181 [math.DS].

[34] Rodolfo Ríos-Zertuche. The pillowcase distribution and near-involutions. *Electronic Journal of Probability*, 19(116), 2014.

[35] Rodolfo Ríos-Zertuche. An introduction to the half-infinite wedge. In *Mexican mathematicians abroad: recent contributions*, volume 657 of *Contemp. Math.*, pages 197–237. Amer. Math. Soc., Providence, RI, 2016.

[36] Rodolfo Ríos-Zertuche. Characterization of minimizable Lagrangian action functionals and a dual Mather theorem. *Discrete & Continuous Dynamical Systems – A*, 40(5):2615–2639, 2020.

[37] Rodolfo Rios-Zertuche. Examples of pathological dynamics of the subgradient method for Lipschitz path-differentiable functions. *Mathematics of Operations Research*, 2022.

[38] Rodolfo Ríos-Zertuche and Milan Korda. The gap between a variational problem and its occupation measure relaxation. Preprint. arXiv:2205.14132 [math.OC].

[39] N.Z. Shor. *Minimization Methods for Non-Differentiable Functions*, volume 3 of *Springer Series in Computational Mathematics*. Springer-Verlag, 1985.

[40] Alfonso Sorrentino. *Action-minimizing Methods in Hamiltonian Dynamics: An Introduction to Aubry-Mather Theory*. Number 50 in Mathematical Notes. Princeton University Press, 2015.