Publicaciones Matemáticas del Uruguay

Proceedings of the CIMPA Research School
Hamiltonian and Lagrangian Dynamics

A tribute to *Ricardo Mañé (1948-1995)*

edited by

Ezequiel Maderna
Ludovic Rifford
Jana Rodriguez Hertz

Volumen 16, Julio 2016

Publicaciones Matemáticas del Uruguay

Publicada por:

CMAT — Facultad de Ciencias
IMERL — Facultad de Ingeniería

Universidad de la República

http://pmu.uy/

ISSN: 0797-1443

Créditos:

Diseño de tapa: J. Rodriguez Hertz
Editor LaTeX: G. Tornaría

Publicaciones Matemáticas del Uruguay

Volumen 16 Julio 2016

NOTAS DE CURSOS

CIMPA Research School: Hamiltionian and Lagrangian Dynamics (2015)

Prefacio

El presente volumen de las Publicaciones Matemáticas del Uruguay contiene las notas de cursos de la escuela *CIMPA Research School Hamiltonian and Lagrangian Dynamics* realizada en la ciudad de Salto entre los días 10 al 19 de marzo de 2015, y está dedicado a la memoria del matemático uruguayo **Ricardo Mañé** (1948-1995). La edición del mismo estuvo a cargo de quienes suscriben, habiendo sido los dos primeros los responsables científicos de la escuela CIMPA.

Debemos agradecer muy especialmente al profesor Claude Cibils, que como director del CIMPA brindó un asesoramiento invaluable al comité organizador. Su apoyo fue fundamental tanto en el trabajo previo al evento (que abarcó casi todo el año que lo precedió), como el que brindó personalmente durante la realización del mismo.

Agradecemos también a la Comisión Sectorial de Investigación Científica de la Universidad de la República, al Instituto de Matemática y Estadística *Rafael Laguardia* de la Facultad de Ingeniería, al Centro de Matemática de la Facultad de Ciencias, así como al Área Matemática del Programa de Desarrollo de las Ciencias Básicas, todas instituciones de la Universidad de la República, por el permanente apoyo brindado para la realización del evento, y en particular por haber financiado conjuntamente la impresión de estas actas.

Por último nuestro mayor agradecimiento va dirigido a todos los estudiantes y profesores que participaron. Fue sin lugar a dudas gracias a su enorme entusiasmo y dedicación que la escuela resultó tan provechosa.

Ezequiel Maderna
Ludovic Rifford
Jana Rodriguez Hertz

Montevideo, Julio 2016.

Lista de participantes

Estudiantes

Alfonso Artigue (Uruguay)
Ignacio Bustamante (Uruguay)
O'Bryan Cárdenas (Chile)
Sebastián Decuadro (Uruguay)
Maximiliano Escayola (Uruguay)
Yuriria Estrada (México)
Eddaly Guerra (México)
Connor Jackman (Estados Unidos)
Serginei Liberato (Brasil)
Gonzalo Manzano (Chile)
Diego Marcon Farias (Brasil)
Santiago Martinchich (Uruguay)
Ignacio Monteverde (Uruguay)
Nicolás Oliva (Uruguay)
Facundo Oliú (Uruguay)
Magdalena Rubio (Uruguay)
Boris Percino (México)
Luiz Perona (Brasil)
Luis Piñeyrúa (Uruguay)
Kenyi Josué Ramírez (México)
Rúbya Ramos (Brasil)
Israel Ramos (Mexico)
Carlos Salazar (Brasil)
Emiliano Sequeira (Uruguay)
Mario Shannon (Uruguay)

Profesores

Marie-Claude Arnaud (Francia)
Gabriele Benedetti (Alemania)
Jorge Delgado (Brasil)
Mario Jorge Dias Carneiro (Brasil)
Claude Cibils (Francia)
Gonzalo Contreras (México)
Nicolas Gourmelon (Francia)
Marcel Guardia (España)
Ezequiel Maderna (Uruguay)
Roberto Markarian (Uruguay)
Marco Mazzucchelli (Francia)
Rafael Potrie (Uruguay)
Ludovic Rifford (Francia)
Alvaro Rovella (Uruguay)
Rafael Oswaldo Ruggiero (Brasil)
Martín Sambarino (Uruguay)
Tere M. Seara (España)
Alfonso Sorrentino (Italia)
Susanna Terracini (Italia)
Andrea Venturelli (Francia)
José Vieitez (Uruguay)
Juliana Xavier (Uruguay)

Secretaría : Lydia Tappa

Las escuelas CIMPA

CIMPA (Centre International de Mathématiques Pures et Appliquées) es una asociación internacional, creada en Niza (Francia) en 1978. Su objetivo es promover la cooperación internacional en beneficio de los países en desarrollo en el campo de la educación superior y la investigación en matemática y disciplinas relacionadas, incluyendo la informática.

La organización de escuelas de investigación es la tarea principal del CIMPA. Sus metas son las contribuir en la formación *mediante la investigación* de nuevas generaciones de matemáticos. Todos los años se abren llamados con la finalidad de organizar aproximadamente una docena de escuelas de investigación en lugares en los que la matemática se encuentra en desarrollo. Estos proyectos son evaluados por el Consejo Científico del CIMPA en el respeto de tres grandes equilibrios: *geográfico*, *temático*, y de *género*.

Más información: `http://www.cimpa-icpam.org/`.

Publicaciones Matemáticas del Uruguay
Volumen 16, Julio 2016, Pages ix–xii
S 0797-1443

SOBRE RICARDO MAÑÉ

EZEQUIEL MADERNA

El día 9 de marzo de 2015, previo al inicio de nuestra escuela de investigación CIMPA "Hamiltonian and Lagrangian Dynamics" realizada en la ciudad de Salto, se cumplían veinte años desde la desaparición física del gran matemático uruguayo *Ricardo Mañé*. Algunos de los participantes y organizadores de este evento tuvimos la suerte y el agrado de conocerlo personalmente; en particular Gonzalo Contreras, Jorge Delgado, Miguel Paternain y Alvaro Rovella realizaron estudios de doctorado bajo su orientación en el Instituto de Matemática Pura y Aplicada (IMPA, Brasil). No fue sorpresa constatar que todos los participantes – incluso entre los estudiantes más jóvenes – conocían en mayor o menor grado de profundidad, pero de forma ineludible, sobre los importantes aportes de Mañé a la matemática. Tampoco es sorprendente que en otras reuniones científicas, congresos o seminarios dedicados a los sistemas dinámicos y en diversos lugares del mundo, se citen frecuentemente sus resultados o los problemas que dejó planteados. Según Mathematical Reviews de la American Mathematical Society, al día de hoy las cincuenta publicaciones indexadas de su pluma cuentan con más de dos mil citaciones, aunque esta información es absolutamente insuficiente si queremos transmitir la importancia de su legado científico. Tampoco es nuestro objetivo hacerlo en este breve artículo – el lector interesado encontrará abundante literatura sobre la vida y obra de Ricardo Mañé – salvo sobre algunos aspectos de sus últimos trabajos que abordaremos más adelante. Un artículo que a mi juicio sintetiza fielmente muchas características personales de Mañé es el que fuera publicado en Revista Matemática Universitaria de la Sociedad Brasilera de Matemática en el número 18 correspondiente al mes de junio de 1995 (pp.1-18) con el título *Triálogo sobre Ricardo Mañé*. Consiste en una entrevista simultánea a Welington De Melo, Jacob Palis y Marcelo Viana.

Ricardo Mañé Ramírez nació en Montevideo el día 14 de enero de 1948. A finales de la década del sesenta realizaba estudios en la Facultad de Ingeniería, donde su padre *Edelmiro Mañé* era profesor de termodinámica. Su madre, *María Adelaida Ramírez* era una conocida artista lírica uruguaya. Mientras estudiaba los fundamentos de la carrera de ingeniero electricista, se incorporó a un grupo de estudiantes de matemática de dicha facultad, y se interesó particularmente por los problemas que planteaba el profesor Lewowicz sobre la teoría de los sistemas dinámicos. En 1971 solicitó con éxito la admisión en el programa de doctorado del IMPA, Río de Janeiro, donde se doctoró bajo la orientación de Jacob Palis, y posteriormente desarrolló su brillante carrera, formó a decenas de matemáticos e influenció a muchos más con su profunda visión de la matemática.

Reinstaurada la democracia en Uruguay y liberado de prisión José Luis Massera en 1985, se inicia en Uruguay un importante proceso de reconstrucción académica que condujo al desarrollo y la consolidación de su escuela matemática. A fines de los ochenta los matemáticos que realizaban investigaciones en el país se contaban con los dedos de las manos y casi todos eran retornados del exterior. Actualmente esa cifra es aproximadamente diez veces mayor. Ricardo Mañé no fue ajeno a esa reconstrucción, visitaba esporádicamente los grupos de matemáticos que crecían en Montevideo, y realizaba también un puente importante entre estos equipos y el IMPA, en el cual se formaron una gran cantidad de los actuales matemáticos uruguayos. Recuerdo vivamente mi primer encuentro con Ricardo Mañé a principios de 1993. Siendo yo un estudiante de la licenciatura en matemática en la nueva facultad de Ciencias, me había interesado en la conjetura de Aizerman sobre estabilidad asintótica en grande. Había logrado probar la conjetura con ciertas hipótesis adicionales. Nuestro encuentro se produjo una tarde en el lugar inevitable: corredor del Instituto de Matemática y Estadística "Rafael Laguardia" de la Facultad de Ingeniería. Se presentó diciéndome que venía de Brasil y que alguien le había comentado algo de mi trabajo, sobre el cual conversamos un momento, al tiempo que me solicitó una copia de las notas que había redactado. En ese momento, ignoraba por completo quien era esa extraña persona y de hecho me olvidé por completo de ese encuentro hasta la tarde del día siguiente, en que volvemos a vernos, esta vez en el Centro de Matemática de la Facultad de Ciencias. Había leído todo minuciosamente, me sugirió ciertas mejoras y me indicó posibles caminos para poder continuar trabajando en el problema. Meses más tarde, en mayo, recibí la noticia de que el problema había sido resuelto completamente por Carlos Gutiérrez. Gracias a una invitación que me extendió Mañé para visitar el IMPA durante enero y febrero de 1994 pude hablar personalmente con Gutiérrez. Durante mi primer estadía en ese instituto pude comprender la importancia que tenía Mañé para la comunidad que lo integraba. Asistía casi siempre por las tardes, y era consultado permanentemente por colegas y estudiantes, tanto en los corredores como en su oficina o en la sala del café. Se percibía claramente la gran admiración que allí todos le profesaban, y era notable ver como se deleitaba colaborando con sus ideas en las diferentes problemáticas que le planteaban. Nadie dejaba la conversación con Mañé con las mismas ideas que había llegado. Tampoco se perdía la oportunidad de hablar de temas polémicos, de criticar a diestra y siniestra de forma tan aguda y sarcástica que causaba generalmente la risa de todos quienes lo escuchaban hablar. Recuerdo también su enorme conocimiento en materia de ópera y música clásica, y en especial recuerdo largas conversaciones que mantuvimos sobre la obra sinfónica de Mahler – sobre la cual opinaba que debía reducirse en duración a la mitad, sin afectar en lo más mínimo la primera de ellas – y las diferentes interpretaciones que conocíamos.

Dedicó los últimos años de su carrera científica al estudio de los sistemas dinámicos lagrangianos, realizando notables descubrimientos en esta disciplina. Motivado primero por la lectura de algunos trabajos de Sergey Bolotin, y luego por los artículos de John Mather sobre las medidas minimizantes para sistemas lagrangianos autónomos, logra desentrañar un concepto que resultó fundamental para todos los desarrollos posteriores de esta teoría: el del valor de energía crítico de un sistema lagrangiano. Podemos describirlo groso modo como un valor peculiar de la energía del sistema, cuyo correspondiente conjunto de nivel contiene necesariamente ciertos conjuntos invariantes caracterizados por propiedades variacionales globales. Estos conjuntos son esenciales para la comprensión de la dinámica global: de alguna forma articulan el sistema tal como las órbitas parabólicas lo hacen en las ecuaciones de Kepler. Por otra parte, percibió que la complejidad dinámica de estos conjuntos invariantes no admite a priori limitación alguna y que, sin embargo, la descripción de los mismos debía ser factible para sistemas genéricos. No es difícil constatar que desde entonces, se han publicado una gran cantidad de artículos de investigación inextricablemente relacionados a este importante concepto, que lo desarrollan, o que lo vinculan con otras teorías. Por ejemplo, resulta imposible distinguir hoy las fronteras entre los resultados obtenidos originalmente por Mañé, y los que forman parte de la teoría de Aubry-Mather, o la teoría weak KAM iniciada por Albert Fathi poco más tarde.

Mañé viajó a Montevideo a fines de noviembre de 1994, entre otras cosas para votar en las elecciones nacionales del día domingo 27 de ese mes, en las cuales su primo *Juan Andrés Ramírez* era candidato a la presidencia de la República. Su visita debía extenderse por algunas semanas y de hecho, a mediados del mes de diciembre, dictó una conferencia en el Centro de Matemática que en aquel entonces funcionaba en su local propio en la calle Eduardo Acevedo. Describió entonces sus más recientes trabajos sobre dinámica lagrangiana, explicando el fuerte vínculo que tenían con la mencionada teoría de Aubry-Mather. Fue aclamado por el público presente, que al igual que el resto de la comunidad científica, en aquella época ya reconocía en Mañé a uno de los matemáticos del mayor relieve internacional. Tenía previsto su regreso a Río de Janeiro, donde vivía, para poco días después de las fiestas tradicionales de fin de año, las cuales pasaría en compañía de su familia. Ocurre entonces algo absolutamente inesperado y trágico, que lo lleva a permanecer en Montevideo hasta el final de sus días, tres meses más tarde: los médicos le detectan el inicio de una metástasis, por causa de un cáncer de pulmón. Mientras su salud comienza a deteriorarse rápidamente, él se ocupa con gran empeño de poner sobre papel todas las ideas matemáticas que proyectaba desarrollar. Recuerdo haber llevado en ese entonces a la casa de su madre, donde recibía todas las tardes la visita de familiares y amigos, media docena de libros que me encargó pedir en préstamo de nuestra biblioteca. Muchas tardes nos encontrábamos todos en la vereda de su casa, esperando que termine la entrevista sistemática que Ricardo mantenía con un sacerdote. El 8 de marzo de 1995, internado en el sanatorio español en la calle Garibaldi, Mañé continuaba escribiendo sus extensas notas. Finalmente, las mismas dieron lugar a su célebre publicación póstuma *Lagrangian flows: the dynamics of globally minimizing orbits* en el Boletín de la Sociedad Matemática Brasilera del año 1997. Obviamente, muchos detalles quedaron inconclusos, pero en los años que siguieron dichas omisiones fueron subsanadas, sus observaciones fueron ampliadas o corregidas, y muchas de las demostraciones fueron finalmente establecidas con el mayor rigor que caracteriza el trabajo matemático (ver por ejemplo [1]).

En esa línea de investigación, una importante pregunta subsiste hasta el día de hoy a pesar de los importantes avances recientes: los especialistas se refieren a ella como *conjetura de Mañé*, a pesar de que Mañé nunca fue muy explícito en su formulación. El problema es decidir si es cierto o no que para cualquier sistema lagrangiano (convexo, superlineal, sobre una variedad cerrada) son genéricas las perturbaciones que hacen que el conjunto invariante crítico (conjunto de Aubry) consista exclusivamente de una órbita periódica hiperbólica o un punto de equilibrio hiperbólico. Mañé logró probar que, sumando a un tal lagrangiano una función genérica de las posiciones, es posible lograr que el sistema perturbado admita una única medida minimizante. Recientemente, los trabajos de Contreras, Figalli y Rifford [2, 3] permitieron establecer que en dimensión dos, es decir en superficies, genéricamente el soporte de la única medida minimizante consiste en una órbita periódica o un punto de equilibrio. En variedades de dimensión tres o superior el problema se mantiene abierto.

Es para mi un placer y un gran honor presentar este volumen de las Publicaciones Matemáticas del Uruguay, dedicado a la memoria de este notable amigo, maestro y matemático.

Montevideo, 9 de marzo de 2016.

Referencias

[1] G. Contreras, R. Iturriaga, G. Paternain, M. Paternain, *Lagrangian Graphs, Minimizing Measures and Mañé's Critical Values*, GAFA Geom. funct. anal. **8** n.5 (1998), pp.788–809.

[2] G. Contreras, A. Figalli, L. Rifford, *Generic hyperbolicity of Aubry sets on surfaces*, Invent. math. **200** n.1 (2015), pp.201–261.

[3] G. Contreras, *Ground states are generically a periodic orbit*, Invent. math. **205** n.2 (2016), pp.383–412.

Ricardo Mañé, 1948–1995.

Ilustración de Horacio Cassinelli

NOTAS DE CURSOS

CIMPA Research School

Hamiltionian and Lagrangian Dynamics (2015)

Publicaciones Matemáticas del Uruguay
Volumen 16, Julio 2016, Páginas 1–39
S 0797-1443

HYPERBOLICITY FOR CONSERVATIVE TWIST MAPS OF THE 2-DIMENSIONAL ANNULUS

MARIE-CLAUDE ARNAUD

Abstract. These are notes for a minicourse given at Regional Norte UdelaR in Salto, Uruguay for the conference CIMPA Research School *Hamiltonian and Lagrangian Dynamics*. We will present Birkhoff and Aubry-Mather theory for the conservative twist maps of the 2-dimensional annulus and focus on what happens close to the Aubry-Mather sets: definition of the Green bundles, link between hyperbolicity and shape of the Aubry-Mather sets, behaviour close to the boundaries of the instability zones. We will also give some open questions. This course is the second part of a minicourse that was begun by R. Potrie. Some topics of the part of R. Potrie will be useful for this part.

Many thanks to E. Maderna and L. Rifford for the invitation to give the mini-course and to R. Potrie for accepting to share the course with me.

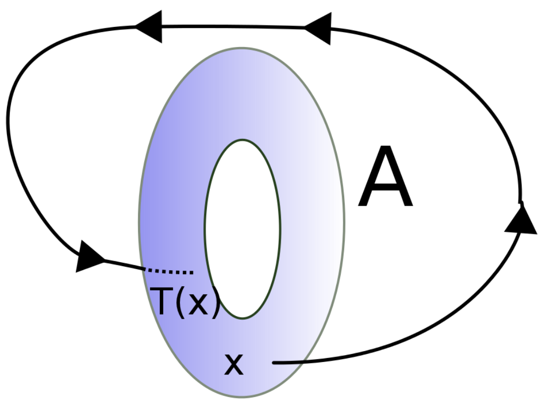
Contents

The author is member of *Institut Universitaire de France* and was partially supported by the ANR project 12-BLAN-WKBHJ (France).

1. Introduction to conservative twist maps

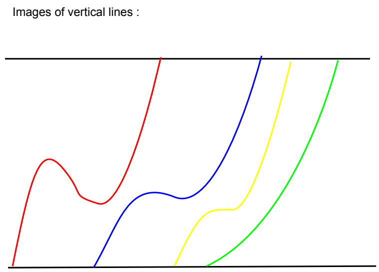
Notations 1.1.
- $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ is the circle; $\mathbb{A} = \mathbb{T} \times \mathbb{R}$ is the annulus and $(\theta, r) \in \mathbb{A}$ refers to a point of $\mathbb{A}$;
- $\mathbb{A}$ is endowed with its symplectic form $\omega = dr \wedge d\theta = d\lambda$ where $\lambda = rd\theta$ is the Liouville 1-form;
- $p : \mathbb{R}^2 \to \mathbb{A}$ is the universal covering;
- $\pi : \mathbb{A} \to \mathbb{T}$ is the first projection: $\pi(\theta, r) = \theta$ and $\pi : \mathbb{R}^2 \to \mathbb{R}$ is its lift, which is also a projection: $\pi(\theta, r) = \theta$;
- for every point $x = (\theta, r)$, the vertical line at x is $\mathcal{V}(x) = \{\theta\} \times \mathbb{R} \subset \mathbb{R}^2$ or $\mathcal{V}(x) = \{\theta\} \times \mathbb{R} \subset \mathbb{A}$;
- the vertical subspace is the tangent subspace to the vertical line: $V(x) = T_x\mathcal{V}(x)$;
- all the measures we will deal with are assumed to be Borel probabilities. The support of μ is denoted by $\operatorname{supp}\mu$.

If $x \in M$ is an elliptic periodic point of a Hamiltonian flow that is defined on a 4-dimensional symplectic manifold M, using symplectic polar coordinates in an annular Poincaré section contained in the energy level of x, we obtain in general a first return map $T : \mathcal{A} \to \mathbb{A}$ that is defined on some bounded sub-annulus $\mathcal{A}$ of $\mathbb{A}$ by $T(\theta, r) = (\theta + \alpha + \beta r, r) + o(r)$ with $\beta \neq 0$. This is locally a *conservative twist map*.



Definition 1.2. A *positive (resp. negative) twist map* is a C^1-diffeomorphism $f : \mathbb{A} \to \mathbb{A}$ such that

(1) f is isotopic to the identity map $\mathrm{Id}_{\mathbb{A}}$ (i.e. f preserve the orientation and the two boundaries of the annulus);
(2) f satisfies the *twist condition* i.e. there exists $\varepsilon > 0$ such that for any $x \in \mathbb{A}$, we have: $\frac{1}{\varepsilon} > D(\pi \circ f)(x)(0,1) > \varepsilon$ (resp. $-\frac{1}{\varepsilon} < D(\pi \circ f)(x)(0,1) < -\varepsilon$). In the first case the twist is positive, in the second case it is negative.



The twist map is *conservative* (or *exact symplectic*) is $f^*\lambda - \lambda$ is an exact 1-form.

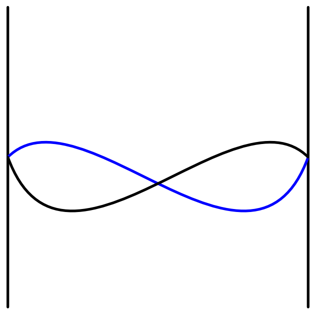
Remarks 1.3. (1) Saying that the diffeomorphism f is isotopic to identity means that:

- f preserves the orientation;
- f fixes the two ends $\mathbb{T} \times \{-\infty\}$ and $\mathbb{T} \times \{+\infty\}$ of the annulus.

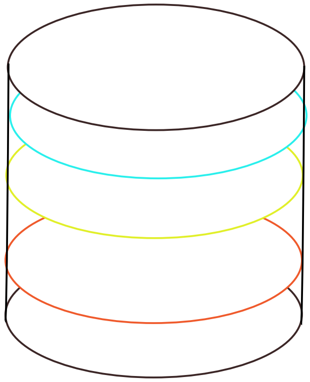
(2) The reader can ask why we don't just ask that f preserves the area form (symplectic form) ω, i.e. $0 = f^*\omega - \omega = d(f^*\lambda - \lambda)$. We ask not only that $f^*\lambda - \lambda$ is closed, we ask that it is exact. Indeed, we want to avoid symplectic twist maps as $(\theta, r) \mapsto (\theta + r, r + 1)$: all the orbits come from $\mathbb{T}\times\{-\infty\}$ and go to $\mathbb{T}\times\{+\infty\}$ and there is no non-empty compact invariant set for such a map. We will see in section 3 that this never happens for exact symplectic twist maps;

(3) Note that f is a positive conservative twist map if and only if f^{-1} is a negative conservative twist map. Hence from now we will assume that all the considered conservative twist maps are positive.

Exercise 1.4. Let $f : \mathbb{A} \to \mathbb{A}$ be a conservative twist map. Using Stokes formula, prove that if $\gamma : \mathbb{T} \to \mathbb{A}$ is a C^1-embedding, then the (algebraic) area of the domain that is between γ and $f(\gamma)$ is zero.



Example 1.5. Consider the map we introduced by using polar coordinates for a first return map $T(\theta, r) = (\theta + \alpha + \beta r, r)$ and assume that $\beta > 0$ (or replace T by T^{-1}). Then $D(\pi \circ T)\begin{pmatrix}0\\1\end{pmatrix} = \beta > 0$ hence T is a (positive) twist map. Moreover, $T^*(rd\theta) - rd\theta = \beta r dr = d\left(\frac{\beta}{2}r^2\right)$ hence T is a conservative twist map.

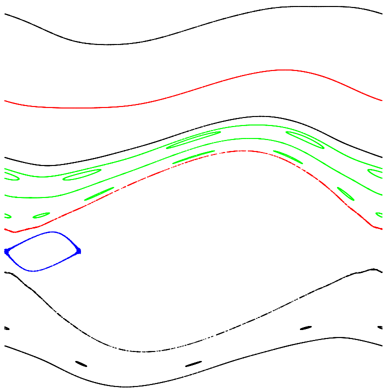


Note that the dynamics is very simple: the annulus is foliated by invariant circles $\mathbb{T} \times \{r\}$ and the restriction of T to every such circle is a rotation.

Example 1.6. The *standard family* depends on a parameter $\lambda \in \mathbb{R}$. It is defined by

$$f_\lambda(\theta, r) = (\theta + r + \lambda \sin 2\pi\theta, r + \lambda \sin 2\pi\theta).$$

Note that for $\lambda = 0$, the map is just the map $T = f_0$ of Example 1.5. When λ increases from 0 to $+\infty$, we observe fewer and fewer invariant graphs.



J. Mather and S. Aubry even proved that for $2\pi\lambda > 4/3$, f_λ has no continuous invariant graph.

Exercise 1.7. (1) *Check that the functions f_λ are all conservative twist maps. Assume that the graph of a continuous map $\psi : \mathbb{T} \to \mathbb{R}$ is invariant by a map f_λ.*

(2) *Prove that $g_\lambda(\theta) = \theta + \lambda \sin(2\pi\theta) + \psi(\theta)$ is an orientation preserving homeomorphism of $\mathbb{T}$.*
Hint*: note that $\pi \circ f_\lambda(\theta, \psi(\theta)) = g_\lambda(\theta)$.*

(3) *Prove that $g_\lambda^{-1}(\theta) = \theta - \psi(\theta)$.*
Hint*: prove that $f^{-1}(\theta, r) = (\theta - r, r - \lambda \sin 2\pi(\theta - r))$.*

(4) *Check that $g_\lambda(\theta) + g_\lambda^{-1}(\theta) = 2\theta + \lambda \sin 2\pi\theta$. Deduce that for $\lambda > \frac{1}{\pi}$, f_λ has no continuous invariant graph.*

We can characterize the conservative twist maps by their *generating functions*.

Proposition 1.8. *Let $F : \mathbb{R}^2 \to \mathbb{R}^2$ be a C^1 map. Then F is a lift of a conservative twist map $f : A \to \mathbb{A}$ if and only if there exists a C^2 function such that*

- *$\forall \theta, \Theta \in \mathbb{R}, S(\theta+1, \Theta+1) = S(\theta, \Theta)$;*
- *there exists $\varepsilon > 0$ so that for all $\theta, \Theta \in \mathbb{R}$, we have*
$$\varepsilon < -\frac{\partial^2 S}{\partial\theta\partial\Theta}(\theta, \Theta) < \frac{1}{\varepsilon};$$
- $F(\theta, r) = (\Theta, R) \Longleftrightarrow R = \frac{\partial S}{\partial \Theta}(\theta, \Theta)$ and $r = -\frac{\partial S}{\partial \theta}(\theta, \Theta)$.

In this case, we say that S is a *generating function* for F (or f). The proof of Proposition 1.8 is given in subsection 5.1.

Exercise 1.9. Check that a generating function of the standard map f_λ is
$$S_\lambda(\theta, \Theta) = \frac{1}{2}(\Theta - \theta)^2 - \frac{\lambda}{2\pi}\cos 2\pi\theta\,.$$

Remark 1.10. Generating functions are very useful to construct new examples or perturbations of known examples of conservative twist maps. Indeed, we only need a function to define a 2-dimensional conservative twist map.
Using generating functions, we can for example prove that for every $k \in [1, \infty]$, there is a dense G_δ subset $\mathcal{G}$ of the set of C^k conservative twist maps such that at every periodic point x of $f \in \mathcal{G}$ with period n, $Df^n(x)$ has two distinct eigenvalues (and then these eigenvalues are different from ± 1). A similar dense G_δ subset $\mathcal{G}$ exists such that the intersections of the stable and unstable submanifolds of every pair of periodic hyperbolic points transversely intersect (when they intersect).

2. The invariant curves

2.1. **Invariant continuous graphs and first Birkhoff theorem.**

In the '20s, G. D. Birkhoff proved (see [11]) that the invariant continuous graphs by a twist map are locally uniformly Lipschitz.

Theorem 1. (G. D. Birkhoff) *Let $f : \mathbb{A} \to \mathbb{A}$ be a conservative twist map and $x \in \mathbb{A}$. Then there exists a C^1-neighborhood $\mathcal{U}$ of f, a neighborhood U of x in $\mathbb{A}$ and a constant $C > 0$ such that if the graph of a continuous map $\psi : \mathbb{T} \to \mathbb{R}$ meets U and is invariant by a $g \in \mathcal{U}$, then ψ is C-Lipschitz.*

Theorem 1 is a consequence of a result that concerns all the Aubry-Mather sets and that we will prove later: Proposition 3.24.

Corollary 2.1. *Let $f : \mathbb{A} \to \mathbb{A}$ be a conservative twist map and let $K \subset \mathbb{A}$ be a compact subset of $\mathbb{A}$. Then there exists a C^1-neighborhood $\mathcal{U}$ of f and a constant $C > 0$ such that if the graph of a continuous map $\psi : \mathbb{T} \to \mathbb{R}$ meets K and is invariant by a $g \in \mathcal{U}$, then ψ is C-Lipschitz.*

Exercise 2.2. Prove Corollary 2.1.

From Theorem 1 and Ascoli theorem, we deduce

Corollary 2.3. *Let f be a conservative twist map of $\mathbb{A}$. The the union $\mathcal{I}(f)$ of all its invariant continuous graphs is a closed invariant subset of f.*

Exercise 2.4. Prove Corollary 2.3.

Remarks 2.5. (1) The set $\mathcal{I}(f)$ can be empty: this is the case for the standard map f_λ with $\lambda > \frac{2}{3\pi}$.

(2) Using the connecting lemma that was proved by S. Hayashi in 2006 (see [17]) and more specifically some related results that are contained in [7], Marie Girard proved (in her non-published PhD thesis) that there is dense G_δ subset $\mathcal{G}$ of the set of C^1 conservative twist maps such that every $f \in \mathcal{G}$ has no continuous invariant graph.

(3) Don't deduce that having an invariant graph rarely happens for the conservative twist maps: it depends on their regularity ($C^1, C^3, \dots, C^\infty$). Indeed, the famous theorems K.A.M. (for Kolmogorov-Arnol'd-Moser, see [8], [22], [28]) tell us that if a C^∞ conservative twist map f has a C^∞ invariant graph $\mathcal{C}$ such that the restriction $f_{|\mathcal{C}}$ is C^∞ conjugated to a Diophantine rotation $\theta \mapsto \theta + \alpha$ (i.e. α is Diophantine: there exist $\gamma, \delta > 0$ so that for every $p \in \mathbb{Z}$ and $q \in \mathbb{N}^*$, we have $|\alpha - \frac{p}{q}| \geq \frac{\gamma}{q^{1+\delta}}$), there exists a neighborhood $\mathcal{U}$ of f in C^∞-topology such that every $g \in \mathcal{U}$ has a C^∞ invariant graph Γ such that $g_{|\Gamma}$ is C^∞-conjugated to $f_{|\mathcal{C}}$.

As the completely integrable standard map f_0 has a lot of such invariant graphs, we deduce that for λ small enough, f_λ has many C^∞ invariant graphs.

Remark 2.6. We will see that even when a conservative twist map has no continuous invariant graph, it has a lot of compact invariant subsets: periodic orbits, and even invariant Cantor sets (these are the Aubry-Mather sets, see section 3).

2.2. **Circle homeomorphisms and dynamics on $\mathcal{I}(f)$.** Now let us explain how is the dynamics restricted to $\mathcal{I}(f)$.
The dynamics restricted to every invariant graph is Lipschitz conjugated (via π) to an orientation preserving bi-Lipschitz homeomorphism of $\mathbb{T}$. The classification of the orientation preserving homeomorphisms of the circle is due to H. Poincaré and given in [21] (see [18] for more results). Let us recall quickly the main results. We assume that $h : \mathbb{T} \to \mathbb{T}$ is an orientation preserving homeomorphism and that $H_1, H_2 : \mathbb{R} \to \mathbb{R}$ are some lifts of h (then $H_2 - H_1 = k$ is an integer). Then

- the sequence $\left(\frac{H_i^n - Id}{n}\right)_{n\in\mathbb{N}}$ uniformly converge to a real number $\rho(H_i)$ that is called the *rotation number* of H_i; note that $\rho(H_2) - \rho(H_1) = k$; then the class of $\rho(H_i)$ modulo $\mathbb{Z}$ defines a unique number $\rho(h) \in \mathbb{T}$ and is called the rotation number of h;
- $\rho(H_i) = \frac{m}{n} \in \mathbb{Q}$ (with m and n relatively prime) if and only if there exists a point $t \in \mathbb{R}$ so that $H_i^n(t) = t + m$; in this case a point t of $\mathbb{T}$ is either periodic for h or such that there exist two periodic points t_-, t_+ with period n for h such that
$$\lim_{\ell \to +\infty} d(h^{-\ell}t, h^{-\ell}t_-) = \lim_{\ell \to +\infty} d(h^\ell t, h^\ell t_+) = 0.$$
In this last case, t is negatively heteroclinic to t_- and positively heteroclinic to t_+.
- when $\rho(h) \notin \mathbb{Q}/\mathbb{Z}$, h has no periodic points and either the dynamics is minimal and C^0-conjugated to the rotation $t \mapsto t + \rho(h)$ or the non wandering set of h is a Cantor subset (i.e. non-empty compact totally disconnected with no isolated point) Ω, $h_{|\Omega}$ is minimal and all the orbits in $\mathbb{T}\backslash\Omega$ are wandering and homoclinic to Ω (this means that $\lim\limits_{\ell \to \pm\infty} d(h^\ell t, \Omega) = 0$). Moreover, f has a unique invariant measure, and its support is Ω.

Moreover, if $q \in \mathbb{Z}^*$, $p \in \mathbb{Z}$ are such that $\rho(H_i) < \frac{p}{q}$ (resp. $\rho(H_i) > \frac{p}{q}$), then we have $H_i^q(t) - t - p < 0$ (resp. $H_i^q(t) - t - p > 0$). We deduce that

$$\forall k \in \mathbb{Z}, |H_i^k(t) - t - k\rho(H_i)| \leq 1.$$

Definition 2.7. When an invariant graph has an irrational (resp. rational) rotation number, we will say that the graph is *irrational* (resp. *rational*).
When the rotation number is irrational and the dynamics is not minimal, we have a *Denjoy counter-example.*

2.3. Lyapunov exponents of the invariant curves.

Definition 2.8. Let $\mathcal{C} \subset \mathbb{A}$ be a set that is invariant by a map $f : \mathbb{A} \to \mathbb{A}$. Then its stable and unstable sets are defined by

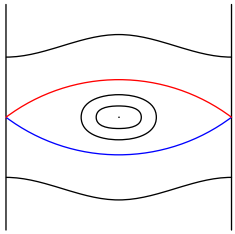
$$W^s(\mathcal{C}, f) = \{x \in \mathbb{A}; \lim_{k \to +\infty} d(f^k x, \mathcal{C}) = 0\}$$

and

$$W^u(\mathcal{C}, f) = \{x \in \mathbb{A}; \lim_{k \to +\infty} d(f^{-k} x, \mathcal{C}) = 0\}.$$

One of these two sets is *trivial* if it is equal to $\mathcal{C}$.

Example 2.9. We consider the Hamiltonian flow of the pendulum. In other words, we define $H : \mathbb{A} \to \mathbb{R}$ by $H(\theta, r) = \frac{1}{2}r^2 + \cos 2\pi\theta$ and its Hamiltonian flow (φ_t) is determined by the Hamilton equations: $\dot{\theta} = \frac{\partial H}{\partial r} = r$ and $\dot{r} = -\frac{\partial H}{\partial \theta} = 2\pi \sin 2\pi\theta$. For $t > 0$ small enough, the time t map $f = \varphi_t$ is a conservative twist map, and as H is constant along the orbits we can find a lot of invariant curves.



Note on this picture that there exists two Lipschitz but non C^1 invariant graphs, that are the separatrices of the hyperbolic fixed point.
Such a separatrix carries only one invariant ergodic measure, the Dirac mass at the hyperbolic fixed point, and then the Lyapunov exponents of this measure are non zero, and there are non-trivial stable and unstable sets for this separatrix (that is the union of the two separatrices).

Hence this is an example of a rational invariant graph that carries an hyperbolic invariant measure. What happens in the irrational case? It is not hard to prove that if the graph of a C^1-map is invariant by a conservative twist map and irrational, then the unique ergodic measure supported in the curve has zero Lyapunov exponents. When the invariant curve is just assumed to be Lipschitz, this is less easy to prove but also true as we will see in Theorem 2.

Remark 2.10. There exist examples of C^2 conservative twist maps that have an irrational invariant Lipschitz graph that is not C^1. Such an example is built in [2]. We don't know if such an example exists when the twist map in C^∞ or when the dynamics restricted to the graph is not Denjoy (i.e. has a dense orbit).

Question 2.11. Does there exist a C^∞ conservative twist map that has an invariant continuous graph on which the dynamics is Denjoy?

Question 2.12. Does there exist a C^∞ conservative twist map that has an invariant irrational continuous graph that is not C^1?

Question 2.13. If a conservative twist map has an invariant irrational continuous graph on which the restricted dynamics has a dense orbit, is the invariant curve necessarily C^1?

Remarks 2.14.
(1) From Theorem 2 and Theorem 9 that we will prove later, it is not hard to deduce that if a conservative twist map has an invariant irrational graph γ that carries the invariant probability measure μ, then γ is C^1-regular μ-almost everywhere (see Definition 4.16).
(2) In fact, I proved in [1] that any graph that is invariant by a conservative twist map is C^1 above a G_δ subset of $\mathbb{T}$ that has full Lebesgue measure.

With P. Berger, we proved the following result (see [6]).

Theorem 2. (M.-C. Arnaud & P. Berger) *Let γ be an irrational invariant graph by a $C^{1+\alpha}$ conservative twist map. Then the Lyapunov exponents of the unique invariant probability with support in γ are zero. Hence*

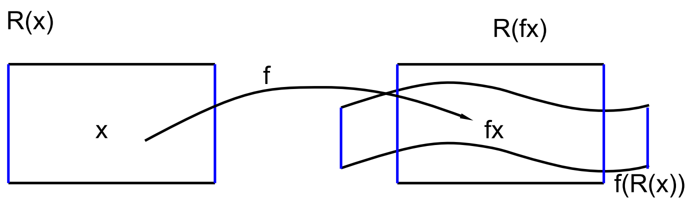
$$\forall \varepsilon > 0, \forall x \in W^s(\gamma, f)\backslash\gamma, \lim_{n\to+\infty} e^{n\varepsilon} d(f^n x, \gamma) = +\infty.$$

The convergence to an irrational invariant curve is slower than exponential. We will explain in subsection 2.4 that a lot of conservative twist maps have an irrational invariant curve with a non trivial stable set.

Proof We begin by proving the first part of the theorem.

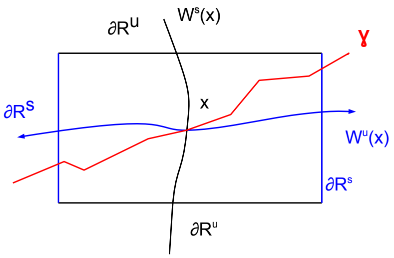
Assume that γ is an invariant continuous graph by a $C^{1+\alpha}$ conservative twist map f and that some ergodic invariant probability μ with support in γ is hyperbolic, i.e. has two Lyapunov exponents such that $\lambda_1 < 0 < \lambda_2$. As f is symplectic, then $\lambda_2 = -\lambda_1 = \lambda$.

We use Pesin theory and Lyapunov charts (rectangles $R(f^k x)$) along a generic orbit $(f^k x)$ for μ: in such a chart, the dynamics is almost linear and hyperbolic



We will prove that μ-almost x is periodic. The curve γ is endowed with some orientation. Note that $f_{|\gamma}$ is orientation preserving.

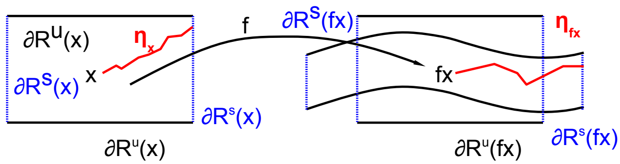
We decompose the boundary ∂R of the domain of a chart R into $\partial^s R = \{-\rho, \rho\} \times [-\rho, \rho]$ and $\partial^u R = [-\rho, \rho] \times \{-\rho, \rho\}$



Let γ_x be the connected components of $\gamma \cap R(x)$ that contains x and let η_x be the set of the points of γ_x that are after x (for the orientation of γ_x).

We will prove that μ-almost x is periodic and $\eta_x \subset W^s(x)$ or $\eta_x \subset W^u(x)$.

Lemma 2.15. *We have either for μ almost every x, $\eta_x(1) \in \partial R^s(x)$ or for μ almost every x, $\eta_x(1) \notin \partial R^s(x)$.*



Proof If $\eta_x(1) \in \partial R^s(x)$, then for all $n \geq 1$, we have $\eta_{f^n x}(1) \in \partial R^s(f^n x)$. Then the map $\mathcal{I}$ defined by $\mathcal{I}(x) = 1$ if $\eta_x(1) \in \partial R^s(x)$ and $\mathcal{I}(x) = 0$ if not is non-decreasing along the orbits and then constant almost everywhere.

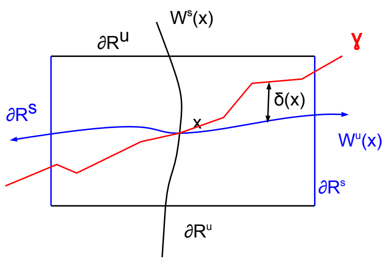
We have indeed $\int (\mathcal{I} \circ f - \mathcal{I}) d\mu = 0$ and $\mathcal{I} \circ f \geq \mathcal{I}$. Hence $\mathcal{I} \circ f = \mathcal{I}$ μ- a.e. and then as μ is ergodic $\mathcal{I}$ is constant μ-almost everywhere. $\square$

Assume for example that we have almost everywhere $\eta_x(1) \in \partial^s R(x)$. Hence we have $\eta_{fx} \subset f(\eta_x)$.

The local unstable manifold at x is the graph of a continuous function g_x^u.

If $\eta_x = (\eta_x^1, \eta_x^2)$ we introduce the notation:

$$\delta(x) = \max_{t \in [0,1]} |\eta_x^2(t) - g_x^u(\eta_x^1(t))|.$$



Using hyperbolicity, we obtain $\delta(fx) \leq e^{-\frac{\lambda}{2}} \delta(x)$, and then $\int \delta d\mu \leq e^{-\frac{\lambda}{2}} \int \delta d\mu$ and then $\delta = 0$ μ almost everywhere.

We deduce that the corresponding branch of $W^u(x)$ is contained in γ for μ-almost every x.

Assume that γ is irrational. Then $f_{|\gamma}$ has to be Denjoy (because for some points we have $\lim\limits_{n\to+\infty} d(f^{-n}x, f^{-n}y) = 0$).

In this case, the only points $x \in \text{supp}\mu$ such that $W^u(x) \neq \{x\}$ are the endpoints of the wandering intervals and there are only countably many such points: their set has μ-measure 0.

Finally, γ cannot be irrational.

The second part of Theorem 2 is a consequence of the following theorem that we will prove.

Theorem 3. *Let $f : M \to M$ be a C^1-diffeomorphism of a manifold M. Let $K \subset M$ be a compact set that is invariant by f. We assume that $f_{|K}$ is uniquely ergodic and we denote the unique Borel invariant probability with support in K by μ. We assume that all the Lyapunov exponents of μ are zero. Let $x_0 \in W^s(K, f)\backslash K$. Then we have:*

$$\forall \varepsilon > 0, \lim_{n\to+\infty} e^{\varepsilon n} d(f^n(x_0), K) = +\infty.$$

Let us now prove this theorem.

Proof. By hypothesis, we have for μ-almost every point :

$$\lim_{n\to\pm\infty} \frac{1}{n} \log \|Df^n(x)\| = 0.$$

We can use a refinement Kingman's subadditive ergodic theorem that is due to A. Furman (see Theorem 12 of subsection 5.5) that implies that we have

$$\limsup_{n\to\pm\infty} \max_{x\in K} \frac{1}{n} \log \|Df^n(x)\| \leq 0.$$

In particular, for any $\varepsilon > 0$, there exists $N \geq 1$ such that:

$$\forall x \in K, \forall n \geq N, \frac{1}{n} \log \|Df^{-n}(x)\| \leq \frac{\varepsilon}{8}. \tag{1}$$

Observe that the following norm with $k \geq N$ large:

$$\|u\|'_x = \sum_{n=0}^{k} e^{-n\varepsilon/4} \|Df^{-n}(x)u\|_x,$$

satisfies uniformly on x for $u \neq 0$:

$$\frac{\|Df^{-1}(x)u\|'_{f^{-1}(x)}}{\|u\|'_x} = e^{\varepsilon/4} + \frac{e^{-k\varepsilon/4}\|Df^{-k-1}(x)u\|_x - e^{\varepsilon/4}\|u\|}{\|u\|'_x}$$

$$\leq e^{\varepsilon/4} + \frac{e^{-k\varepsilon/4}\|Df^{-k-1}(x)u\|_x}{\|u\|'_x} \leq e^{\varepsilon/4} + e^{-k\varepsilon/8}$$

Hence by changing the Riemannian metric by the latter one, we can assume that the norm of $D_x f^{-1}$ is smaller than $e^{\varepsilon/3}$ for every $x \in K$.

Consequently, on a η-neighborhood N_η of K, it holds for every $x \in N_\eta$ that:

$$\|D_x f^{-1}\|' \leq e^{\varepsilon/2}$$

Let $x_0 \in M$ be such that $x_n := f^n(x_0) \to K$, we want to show that

$$\liminf \frac{1}{n} \log d(x_n, K) \geq -\varepsilon.$$

We suppose that $\liminf \frac{1}{n} \log d(x_n, K) < -\varepsilon$ for the sake of a contradiction. Hence there exists n arbitrarily large so that x_n belongs to the $e^{-n\varepsilon}\eta$-neighborhood of K. Let γ be a C^1-curve connecting x_n to K and of length at most $e^{-n\varepsilon}\eta$. By induction on $k \leq n$, we notice that $f^{-k}(\gamma)$ is a curve that connects x_{n-k} to K, and has length at most $e^{-n\varepsilon+k\varepsilon/2}\eta$, and so is included in N_η. Thus the point x_0 is at most $e^{-n\varepsilon/2}\eta$-distant from K. Taking n large, we obtain that x_0 belongs to K. A contradiction. $\square$

2.4. **Instability zones and the second Birkhoff theorem.** As now we know how the dynamics restricted to $\mathcal{I}(f)$ is, we will look to the complement $\mathcal{U}(f)$ of $\mathcal{I}(f)$.

Definition 2.16. An *essential curve* is a C^0-embedded circle in $\mathbb{A}$ that is not homotopic to a point, i.e. a loop that winds around the annulus.
An *essential subannulus* of $\mathbb{A}$ is a subset of $\mathbb{A}$ that is homeomorphic to $\mathbb{A}$ and that contains an essential curve of $\mathbb{A}$.

Proposition 2.17. *Let f be a conservative twist map. Every connected components of $\mathcal{U}(f)$ is either a bounded disc or an essential sub-annulus of $\mathbb{A}$.*

- *When such a component is a disc $\mathcal{D}$, then this disc is periodic i.e. there exists $N \geq 1$ such that $f^N(\mathcal{D}) = \mathcal{D}$. Moreover, the boundary of $\mathcal{D}$ is the union of parts of two invariant continuous graphs that have the same rational rotation number.*
- *When such a component is an essential sub-annulus, then it is invariant by f, and each of the two components of its boundary is either $\mathbb{T} \times \{\pm\infty\}$ or an invariant continuous graph.*

Proof Let U be a connected component of $\mathcal{U}(f)$. Then there is a partition of the set of the invariant continuous graphs in two parts: the set $\mathcal{S}_+$ of such curves that are above U and the set $\mathcal{S}_-$ of those that are under U. Let us differentiate which cases can occur

(1) if $\mathcal{S}_- = \mathcal{S}_+ = \emptyset$, then $U = \mathbb{A}$ is an essential annulus;
(2) if $\mathcal{S}_- = \emptyset$ and $\mathcal{S}_+ \neq \emptyset$ (resp. $\mathcal{S}_+ = \emptyset$ and $\mathcal{S}_- \neq \emptyset$), let us denote by γ_+ (resp. γ_-) the smallest element in $\mathcal{S}_+$ (resp. the largest element in $\mathcal{S}_-$). Then U is the component under γ_+ (resp. above γ_-), that is an essential sub-anulus, and its boundary is γ_+ (resp. γ_-);
(3) if $\mathcal{S}_- \neq \emptyset$ and $\mathcal{S}_+ \neq \emptyset$, let us denote by γ_+ (resp. γ_-) the smallest element in $\mathcal{S}_+$ (resp. the largest element in $\mathcal{S}_-$). Then U is a connected component of the points that are between γ_- and γ_+. If $\gamma_- \cap \gamma_+ \neq \emptyset$, it is a disc $\mathcal{D}$ such that $\partial\mathcal{D} \subset \gamma_- \cup \gamma_+$; moreover, as γ_- meets γ_+, this two curve have the same rotation number and $\gamma_- \cap \gamma_+$ contains exactly two points of $\partial\mathcal{D}$ and they are periodic: the rotation number is rational . If $\gamma_- \cap \gamma_+ = \emptyset$, then U is an essential sub annulus with boundary $\gamma_- \cup \gamma_+$.

From the fact that the invariant curves are invariant, we deduce that the the annular components of $\mathcal{U}(f)$ are invariant. The components U that are homeomorphic to a disc are between two invariant curves, hence contained in an invariant domain with finite Lebesgue measure. This implies that for some $N \geq 1$, we have $f^N(U) \cap U \neq \emptyset$ and then $f^N(U) = U$. $\square$

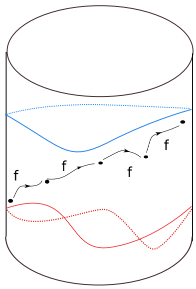
Definition 2.18. If f is a conservative twist map, an annular component of $\mathcal{U}(f)$ is called an *instability zone.*

The following result, which was proved independently by J. Mather (see [27] where the author uses variational methods) and P. Le Calvez (see [23] where the author uses topological methods), explains why these regions are called instability zones.

Theorem 4. (P. Le Calvez; J. N. Mather) *Let $\mathcal{A}$ be an instability zone of a conservative twist map f of the annulus. We choose boundaries $\mathcal{C}_-, \mathcal{C}_+$ of $\mathcal{A}$. Then there exists $x \in \mathcal{A}$ so that $\lim\limits_{k\to\pm\infty} d(f^k x, \mathcal{C}_\pm) = 0$.*

Remarks 2.19. (1) Note that we can choose $\mathcal{C}_- = \mathcal{C}_+$.

(2) Theorem 4 tells us that $W^u(\mathcal{C}_-) \cap W^s(\mathcal{C}_+) \cap \mathcal{A} \neq \emptyset$



Ideas of proof Let us explain in a few words what are the ideas to prove a weaker but related result due to Birkhoff: assume $\mathcal{C}_- \neq \mathcal{C}_+$, fix a neighborhood $\mathcal{U}_-$ of $\mathcal{C}_-$ and $\mathcal{U}_+$ of $\mathcal{C}_+$ in $\bar{\mathcal{A}}$, then there exists $x \in \mathcal{U}_-$ and $N \geq 0$ so that $f^N x \in \mathcal{U}_+$.

The main argument is a theorem due to Birkhoff.

Theorem 5. (G. D. Birkhoff) *Let $\mathcal{A} \subset \mathbb{A}$ be an essential sub-annulus that is invariant by a conservative twist map of the annulus and that is equal to the interior of its closure. Then every bounded connected component of $\partial\mathcal{A}$ is the graph of a Lipschitz map.*

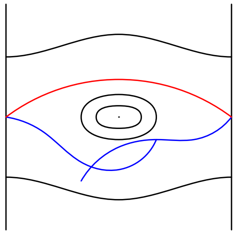
A complete proof of Theorem 5 can be found in the appendix of the first chapter of [18] (in French).

Then assume that $\mathcal{U}_-$ is annular and that the result we want to prove is false. For every $n \in \mathbb{N}$, let V be the connected component of the complement in $\bar{\mathcal{A}}$ of $\overline{\bigcup\limits_{n\in\mathbb{N}} f^n(\mathcal{U}_-)}$ that contains $\mathcal{C}_+$. One can check that the interior of $\bar{V}$ satisfies the hypothesis of Theorem 5, hence we find an invariant continuous graph that is in $\mathcal{A}$ (the boundary of V), that is incompatible with the definition of an instability zone. $\square$

Note an important corollary of theorem 5.

Corollary 2.20. *Let γ be an essential curve that is invariant by a conservative twist map. Then γ is the graph of a Lipschitz map.*

Example 2.21. This example was introduced by Birkhoff in [12]. We consider the Hamiltonian flow f of the pendulum for a small enough time. Using a perturbation of the generating function of f, we can create a transverse intersection between the lower stable branch and the lower unstable branch of the hyperbolic fixed point:



Then the remaining separatrix is the upper boundary of an instability zone.

Exercise 2.22. Prove the last assertion in Example 2.21.

Michel Herman proved in [19] that for a general conservative twist map, there is no essential invariant curve that contains a periodic point. More precisely:
Let $k \in [1, +\infty]$ be a positive integer or ∞. There exists a dense G_δ-subset $\mathcal{G}$ of the set of the C^k PSTM such that every $f \in \mathcal{G}$ has no invariant essential curve that contains a periodic point.
The proof of this result is proposed in Exercice 4.10.

Question 2.23. For which parameters λ does the standard map f_λ satisfy this property?

Question 2.24. How is a "general" boundary of an instability zone? Is it the boundary of one or two intability zone(s)? Is it smooth? How is its rotation number: Diophantine, Liouville?

Remark 2.25. This result of Michel Herman joined to the fact that there exist open sets of C^∞ conservative twist maps that have a lot of (Diophantine) invariant graphs, allows us to state :

Proposition 2.26. *There exists a dense G_δ-subset $\mathcal{G}$ (for the C^∞-topology) in a non-empty open set of conservative C^∞ twist map such that every $f \in \mathcal{G}$ has a bounded instability zone with irrational boundaries.*

Then the stable set of such an irrational boundary is not empty (because of Theorem 4) but the convergence to such a boundary is slower than exponential (because of Theorem 2).

Exercise 2.27. Prove Proposition 2.26.

Question 2.28. For which parameters λ has the standard map f_λ an irrational boundary of instability zone?

3. Aubry-Mather theory

3.1. **Action functional and minimizing orbits.** In this section, we assume that $S : \mathbb{R}^2 \to \mathbb{R}$ is a generating function of a lift $F : \mathbb{R}^2 \to \mathbb{R}^2$ of a conservative twist map $f : \mathbb{A} \to \mathbb{A}$.

Definition 3.1. If $k \geq 1$, one defines the *action functional* $\mathcal{F}_{k+1} : \mathbb{R}^{k+1} \to \mathbb{R}$ by

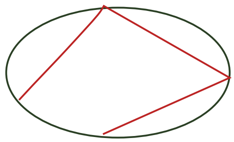
$$\mathcal{F}(\theta_0, \dots, \theta_k) = \sum_{j=1}^{k} S(\theta_{j-1}, \theta_j).$$

For every $k \geq 2$ and every $\theta_b, \theta_e \in \mathbb{R}^n$, the function $\mathcal{F}_{k+1}$ (or $\mathcal{F}$) restricted to the set $\mathcal{E}(k+1, \theta_b, \theta_e)$ of $(k+1)$-uples $(\theta_0, \dots, \theta_k)$ beginning at θ_b and ending at θ_e, i.e. such that $\theta_0 = \theta_e$ and $\theta_k = \theta_e$, has a minimimum and at every critical point for $\mathcal{F}_{k+1|\mathcal{E}(k+1,\theta_b,\theta_e)}$, the following sequence is a piece of orbit for F:

$$(\theta_0, -\frac{\partial S}{\partial \theta}(\theta_0, \theta_1)), (\theta_1, \frac{\partial S}{\partial \Theta}(\theta_0, \theta_1)), (\theta_2, \frac{\partial S}{\partial \Theta}(\theta_1, \theta_2)), \dots, (\theta_k, \frac{\partial S}{\partial \Theta}(\theta_{k-1}, \theta_k)).$$

Observe that for such a critical point, we have $\frac{\partial S}{\partial \Theta}(\theta_{i-1}, \theta_i) + \frac{\partial S}{\partial \theta}(\theta_i, \theta_{i+1}) = 0$ for every $0 < i < k$.

Example 3.2. To illustrate the notion of generating function, let us introduce a very classical example of twist map that is due to G.D. Birkhoff: the so-called Birkhoff billiard. Play billiard on a planar billiard table with a C^2 and convex boundary with non-vanishing curvature. Then we can choose symplectic coordinates (angular coordinate for the point of bounce and radial coordinate that is the sinus of the angle of reflection) in such a way that the dynamical system becomes a conservative twist map (see [29] for details).



In these coordinates, if $\theta_0, \dots, \theta_n \in \mathbb{R}^{n+1}$, then $\mathcal{F}(\theta_0, \dots, \theta_n)$ is just the length of the polygonal line that joins the successive points with angular coordinates $\theta_0, \dots, \theta_n$.

Definition 3.3. A finite or infinite sequence of real numbers $(\theta_n)_{n \in J}$ is a *minimizer* if for every segment $[\ell, k] \subset J$, $(\theta_n)_{\ell \leq n \leq k}$ is a global minimizer of

$$\mathcal{F}_{k-\ell+1|\mathcal{E}(k-\ell+1,\theta_\ell,\theta_k)}\,.$$

When $J = \mathbb{Z}$, we say that (θ_n) is a *minimizing sequence*; we denote the set of minimizing sequences by $\mathcal{M} \subset \mathbb{R}^{\mathbb{Z}}$.

An orbit (θ_n, r_n) of F (and by extension its projection on $\mathbb{A}$) is minimizing *if its projection (θ_n) is a minimizing sequence.*

Remark 3.4. Observe that a minimizer is always the projection of a piece of orbit. From Lemma 3.15, we can deduce

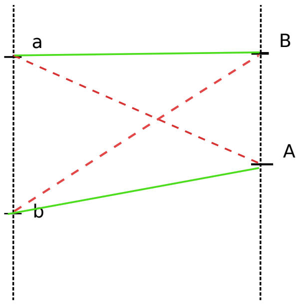
- in every $\mathcal{E} = \mathcal{E}(k+1, \theta_b, \theta_e)$, there exists a minimizer of $\mathcal{F}_{|\mathcal{E}}$; such a minimizer is a segment of the projection of an (non necessarily minimizing) orbit;
- if $(q, p) \in \mathbb{Z}^* \times \mathbb{Z}$, the restriction of $\mathcal{F}_{q+1}$ to the set $\{(\theta_k); \theta_{k+q} = \theta_k + p\}$ has a global minimizer. Any such minimizer is the projection of an orbit and we will even see in Proposition 3.10 that it is a minimizing sequence.

The following theorem is due to J. Mather and proved in subsection 5.2.

Theorem 6. (J. N. Mather) *Assume that the graph of a continuous map $\psi : \mathbb{T} \to \mathbb{R}$ is invariant by a conservative twist map f. Then for any generating function associated to f, all the orbits contained in the graph of ψ are minimizing.*

Now we will give some properties of the minimizers and prove the existence of some periodic minimizers.

Proposition 3.5. **(Aubry & Le Daeron non-crossing lemma)** *Assume*
$$(b-a)(B-A)\le 0.$$
Then
$$S(a,A)+S(b,B)-S(a,B)-S(b,A)\ge 0$$
and equality occurs if and only if $(b-a)(B-A)=0$.

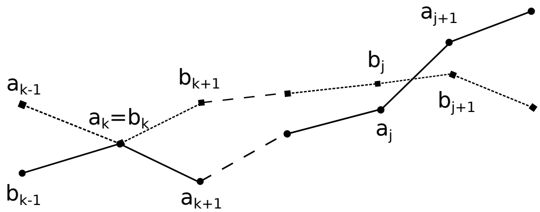


Proof Let us use the notation $A_t=A+t(B-A)$ and $a_t=a+t(b-a)$. We have:
$$S(a,A)+S(b,B)-S(a,B)-S(b,A)=(S(b,B)-S(b,A))-(S(a,B)-S(a,A))$$
$$= (B-A)\int_0^1\left(\frac{\partial S}{\partial\Theta}(b,A_t)-\frac{\partial S}{\partial\Theta}(a,A_t)\right)dt$$
$$= (b-a)(B-A)\int_0^1\int_0^1\frac{\partial^2 S}{\partial\theta\partial\Theta}(a_s,A_t)ds\,dt.$$
From $\frac{\partial^2 S}{\partial\theta\partial\Theta}<0$, we deduce the wanted result. ∎

Definition 3.6. If (θ_k) is a finite or infinite sequence of real numbers, its *Aubry diagram* is the graph of the function obtained when interpolating linearly the sequence (k,θ_k).

Two sequences $(a_k)_{k\in I}$ and $(b_k)_{k\in I}$ *cross* if for some k, j: $(a_k-b_k)(a_j-b_j)<0$.

Remark 3.7. They are two types of crossing: at an integer or at a non-integer:



Note that if two distinct minimizers are such that for a k we have $a_k=b_k$, then we have $a_{k-1}\neq b_{k-1}$ and $a_{k+1}\neq b_{k+1}$; indeed, if two successive terms coincide, then they correspond to a same orbit and then to the same minimizer.

Proposition 3.8. **(Aubry fundamental lemma)** *Two distinct minimizers cross at most once.*

Proof Assume that the minimizers (a_k) and (b_k) cross at two different times t_1 and t_2. Let us introduce the notation $k_i=[t_i]$. We consider the the following finite segments:

- $A=(a_k)_{k_1\le k\le k_2+1}$;
- $B=(b_k)_{k_1\le k\le k_2+1}$;
- $\alpha=(a_{k_1},b_{k_1+1},\dots,b_{k_2},a_{k_2+1})$;
- $\beta=(b_{k_1},a_{k_1+1},\dots,a_{k_2},b_{k_2+1})$.

If t_1 or t_2 is not an integer, we deduce from Proposition 3.5 that

$$\mathcal{F}(A)+\mathcal{F}(B)-\mathcal{F}(\alpha)-\mathcal{F}(\beta)=\\ \sum_{i=1}^{2}\left(S(a_{k_i},a_{k_i+1})+S(b_{k_i},b_{k_i+1})-S(a_{k_i},b_{k_i+1})-S(b_{k_i},a_{k_i+1})\right)>0.$$

As A and α (resp. B and β) have same endpoints, we deduce that A or B is not minimizing, and this is a contradicton.

If $t_i=k_i$ are both integers, then we obtain $\mathcal{F}(A)+\mathcal{F}(B)-\mathcal{F}(\alpha)-\mathcal{F}(\beta)=0$. As $\mathcal{F}(A)\le\mathcal{F}(\alpha)$ and $\mathcal{F}(B)\le\mathcal{F}(\beta)$, we deduce that α and β are also minimizers. But α and A coincides for integers k_2 and k_2+1, hence $\alpha=A$ and then $A=B$. □

Definition 3.9. If $(q,p)\in\mathbb{N}^*\times\mathbb{Z}$, a sequence $(\theta_n)_{n\in\mathbb{Z}}$ is a (q,p)-*minimizer* if

(1) $\forall n,\theta_{n+q}=\theta_n+p$;

(2) $(\theta_n)_{0\le n\le q-1}$ is a minimizer of the function $(\alpha_n)_{0\le n\le q-1}\mapsto\displaystyle\sum_{n=0}^{q}S(\alpha_n,\alpha_{n+1})$ (with the convention $\alpha_q=\alpha_0+p$).

Observe that a (q,p)-minimizer is the projection of an orbit (θ_n,r_n) for F such that $(\theta_{n+q},r_{n+q})=(\theta_n,r_n)+(p,0)$. Hence it corresponds to a q-periodic orbit for f.

Proposition 3.10. *Any (q,p)-minimizer is a minimizing sequence.*

Exercise 3.11. The goal of the exercise is to prove Proposition 3.10.
(a) Using Proposition 3.8, prove that for every $(q,p)\in\mathbb{N}^*\times\mathbb{Z}$ and $k\ge1$, two distinct (q,p)-minimizers cannot cross.
Hint: prove that if they cross, they cross two times within a period.
(b) Deduce that for every $(q,p)\in\mathbb{N}^*\times\mathbb{Z}$ and $k\ge1$, every (kq,kp)-minimizer is in fact a (q,p)-minimizer.
(c) Deduce that being a (q,p)-minimizer is equivalent to be a (kq,kp)-minimizer.
(d) Deduce Proposition 3.10.

Notation 3.12. If $(q,p)\in\mathbb{Z}^2$, we denote by $T_{q,p}:\mathbb{R}^\mathbb{Z}\to\mathbb{R}^\mathbb{Z}$ the map defined by $T_{q,p}((x_k)_{k\in\mathbb{Z}})=(x_{k-q}+p)_{k\in\mathbb{Z}}$.

Note that if $(\theta_k)_{k\in\mathbb{Z}}$ is a (q,p) minimizer, then $T_{q,p}\left((\theta_k)_{k\in\mathbb{Z}}\right)=(\theta_k)_{k\in\mathbb{Z}}$.

Corollary 3.13. *If $(\theta_k)_{k\in\mathbb{Z}}$ and $(\alpha_k)_{k\in\mathbb{Z}}$ are two (q,p)-minimizers, then they don't cross. In particular, $(\theta_k)_{k\in\mathbb{Z}}$ and $T_{a,b}\left((\theta_k)_{k\in\mathbb{Z}}\right)$ do not cross.*

Proposition 3.14. *For every $q\in\mathbb{N}^*$, $p\in\mathbb{Z}$, there exists at least one (q,p)-minimizer.*

Proof We assume that S is a generating function of a lift F of the conservative twist map f.

Lemma 3.15. *We have* $\lim_{|\Theta-\theta|\to+\infty}\dfrac{S(\theta,\Theta)}{|\Theta-\theta|}=+\infty.$

Proof Using the notation $\theta_t=\theta+t(\Theta-\theta)$, we have

$$\begin{aligned} S(\theta,\Theta) &= S(\theta,\theta)+\int_0^1\frac{\partial S}{\partial\Theta}(\theta,\theta_t)(\Theta-\theta)dt\\ &= S(\theta,\theta)+\int_0^1\frac{\partial S}{\partial\Theta}(\theta_t,\theta_t)(\Theta-\theta)dt-\int_0^1\int_0^t\frac{\partial^2 S}{\partial\theta\partial\Theta}(\theta_s,\theta_t)(\Theta-\theta)^2dsdt\\ &\geq m-M|\Theta-\theta|+\frac{\varepsilon}{2}(\Theta-\theta)^2\end{aligned}$$

where $m=\min_{\theta\in[0,1]}S(\theta,\theta)$ and $M=\max_{\theta\in[0,1]}\left|\dfrac{\partial S}{\partial\Theta}(\theta,\theta)\right|$. ∎

We know consider the set

$$\mathcal{E}(q,p)=\{(\theta_k)_{k\in\mathbb{Z}};\forall k\in\mathbb{Z},\theta_{k+q}=\theta_k+p\}$$

and define $\mathcal{W}:\mathcal{E}(q,p)\to\mathbb{R}$ by

$$\mathcal{W}((\theta_k)_{k\in\mathbb{Z}})=\sum_{k=0}^{q-1}S(x_k,x_{k+1}).$$

Note that if $\ell\in\mathbb{Z}$, then $\mathcal{W}((\theta_k)_{k\in\mathbb{Z}})=\mathcal{W}((\theta_k+\ell)_{k\in\mathbb{Z}})$. Hence we can define $\mathcal{W}$ on the quotient of $\mathcal{E}(q,p)$ by the diagonal action of $\mathbb{Z}$. On this space, $\mathcal{W}$ is coercive and has then a global minimimum. Then this global minimum is attained at a (q,p)-minimizer.

∎

Exercise 3.16. Write the details in the proof of Proposition 3.14.

3.2. F-ordered sets.

Definition 3.17. We say that a subset $E\subset\mathbb{R}^2$ is F-ordered if it is invariant by F and every integer translations $(\theta,r)\mapsto(\theta+k,r)$ with $k\in\mathbb{Z}$ and if

$$\forall x,x'\in E,\pi(x)<\pi(x')\Rightarrow\pi\circ F(x)<\pi\circ F(x').$$

Remark 3.18. We deduce from Corollary 3.13 that if $q\in\mathbb{Z}^*$ and $p\in\mathbb{Z}$, the union of the (q,p)-minimizing orbits is an F-ordered set.

Exercise 3.19. Let $\psi:\mathbb{T}\to\mathbb{R}$ be a continuous map such that the graph of ψ is invariant by a conservative twist map f. Prove for any lift F of f, the graph of ψ is F-ordered.

The following proposition explains how we can construct other F-ordered sets.

Proposition 3.20. *Let F be a lift of a conservative twist map.*

1. *The closure of every F-ordered set is F-ordered;*
2. *Let $(E_n)_{n\in\mathbb{N}}$ be a sequence of F-ordered sets. Let $E\in\mathbb{R}^2$ be the set of points $x\in\mathbb{R}^2$ so that there exist $(x_n)\in\mathbb{R}^2$ satisfying $x_n\in E_n$ and $\lim_{n\to\infty}x_n=x$.*

Then E is F-ordered.

Remark 3.21. The main remark that is useful to prove Proposition 3.20 is the following one.
Assume that $E \subset \mathbb{R}^2$ is invariant by F and all maps $(\theta, r) \mapsto (\theta + k, r)$ with $k \in \mathbb{Z}$. Then E is F-ordered if and only if

$$\forall x, x' \in E, \pi(x) < \pi(x') \Rightarrow \pi \circ F(x) \leq \pi \circ F(x') \quad \text{and} \quad \pi \circ F^2(x) \leq \pi \circ F^2(x').$$

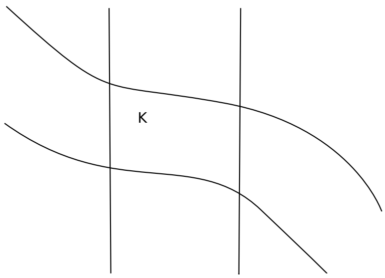
To prove that, observe that if $\pi \circ F(x) = \pi \circ F(x')$ for some $x \neq x'$ in $\mathbb{R}^2$, then $(\pi \circ F^{-1}(x), \pi \circ F^{-1}(x'))$ and $(\pi \circ F(x), \pi \circ F(x'))$ are not in the same order.

Proposition 3.22. *Let F be a lift of a conservative twist map and let $E \subset \mathbb{R}^2$ be a non-empty and closed F-ordered set. Then π maps E homeomorphically onto a closed subset of $\mathbb{R}$ that is invariant by the map $t \in \mathbb{R} \mapsto t + 1$.*

Proof The map π is continuous and open. Assume that there exist two points $x \neq y$ of E such that $\pi(x) = \pi(y)$. Because of the twist condition, we have $x_- = \pi \circ F^{-1}(x) \neq \pi \circ F^{-1}(y) = y_-$ and this contradicts the fact that E is F-ordered.

We just have to prove that $\pi(E)$ is closed. Assume that (x_n) is a sequence of points of E such that $(\pi(x_n))$ converges to some $\theta \in \mathbb{R}$. Then there exists $a, b \in \mathbb{Z}$ so that $\forall n \in \mathbb{N}, \pi(x_0) + a < \pi(x_n) < \pi(x_0) + b$. Because E is F-ordered, we have then $\forall n \in \mathbb{N}, \pi \circ F(x_0) + a < \pi \circ F(x_n) < \pi \circ F(x_0) + b$. Hence

$$x_n \in \pi^{-1}([\pi(x_0) + a, \pi(x_0) + b]) \cap F^{-1}(\pi^{-1}([\pi \circ F(x_0) + a, \pi \circ F(x_0) + b])) = K.$$



Because of the twist condition, K is compact. Hence we can extract a convergent subsequence from (x_n). Because E is closed, $x = \lim x_n \in E$ and then $\theta = \pi(x) \in \pi(E)$. □

We deduce the following statement.

Proposition 3.23. *Let F be the lift of a conservative twist map and let $E \subset \mathbb{R}^2$ be a non-empty and closed F-ordered set. Then there exists an increasing homeomophism $H : \mathbb{R} \to \mathbb{R}$ such that*

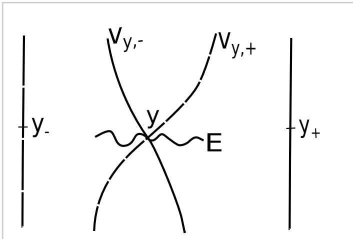
- $\forall t \in \mathbb{R}, H(t + 1) = H(t) + 1$;
- $\forall x \in E, H \circ \pi(x) = \pi \circ F(x)$.

Hence the dynamics F restricted to E is conjugated (via π) to the one of a lift of a circle homeomorphism. We even deduce from Proposition 3.24 that H is bi-Lipschitz. We can then associate to every F-ordered set a rotation number.

Proposition 3.24. *Let $f : \mathbb{A} \to \mathbb{A}$ be a conservative twist map and $x \in \mathbb{A}$. Then there exists a C^1-neighborhood $\mathcal{U}$ of f, a neighborhood U of x in $\mathbb{A}$ and a constant $C > 0$ such that* *if $E \subset \mathbb{R}^2$ is a G-ordered set for a lift G of some $g \in \mathcal{U}$ that meets $U + \mathbb{Z} \times \{0\}$, then E is the graph of some C-Lipschitz map $\psi : \pi(E) \to \mathbb{R}$.*

Note that this proposition is similar to Theorem 1 (in fact, we can deduce Theorem 1 from Proposition 3.24).

Proof Let F be a lift of the conservative twist map $f = (f_1, f_2)$, let $\varepsilon > 0$ be so that $\frac{\partial f_1}{\partial r} \in (\varepsilon, \frac{1}{\varepsilon})$ and let $x = (\theta, r)$ be a point of $\mathbb{R}^2$. Let us choose a compact neighbourhood B of x.
Then for every $y = (\alpha, \rho) \in B$, if we use the notation $y_- = F^{-1}(y) = (\alpha_-, \rho_-)$ and $y_+ = F(y) = (\alpha_+, \rho_+)$, the curves $F^{-1}(\{\alpha_+\} \times [r_+ - \frac{1}{\varepsilon}, r_+ + \frac{1}{\varepsilon}])$ and $F(\{\alpha_-\} \times [r_- - \frac{1}{\varepsilon}, r_- + \frac{1}{\varepsilon}])$ are graphs of some C^1 functions $v_{y,-}, v_{y,+}$ whose domains contain $[\alpha - 1, \alpha + 1]$.



Because $F(F^{-1}(B) + \{0\} \times [-\frac{1}{\varepsilon}, \frac{1}{\varepsilon}])$ and $F^{-1}(F(B) + \{0\} \times [-\frac{1}{\varepsilon}, \frac{1}{\varepsilon}])$ are compact, there exists $K > 0$ such that $\mathbb{R} \times [-K, K]$ contains these two sets.

We define now $\mathcal{U}$ as being the set of conservative twist maps $g = (g_1, g_2)$ with a lift G such that

- $\forall x \in \left(G^{-1}(B) + \{0\} \times [-\frac{1}{\varepsilon}, \frac{1}{\varepsilon}]\right) \cup G^{-1}\left(G(B) + \{0\} \times [-\frac{1}{\varepsilon}, \frac{1}{\varepsilon}]\right)$
$$\frac{\partial g_1}{\partial r}(x) \in (\varepsilon, \frac{1}{\varepsilon})$$
- $G(G^{-1}(B) + \{0\} \times [-\frac{1}{\varepsilon}, \frac{1}{\varepsilon}]) \cup G^{-1}(G(B) + \{0\} \times [-\frac{1}{\varepsilon}, \frac{1}{\varepsilon}]) \subset \mathbb{R} \times [-K, K]$.

Assume that G is such a lift of $g \in \mathcal{U}$. Let E be a G-ordered set that meets B at some y. We deduce from Proposition 3.22 that E is the graph of a map $\psi : \pi(E) \to \mathbb{R}$ and then $y = (\alpha, \psi(\alpha))$ for some $\alpha \in \pi(E) \subset \mathbb{R}$. Because $g \in \mathcal{U}$, we know that $G(G^{-1}(y) + \{0\} \times [-\frac{1}{\varepsilon}, \frac{1}{\varepsilon}])$ and $G^{-1}(G(y) + \{0\} \times [-\frac{1}{\varepsilon}, \frac{1}{\varepsilon}])$ are some subsets of $\mathbb{R} \times [-K, K]$ and are graphs of some C^1 maps v_-, v_+ whose domains contain $[\alpha - 1, \alpha + 1]$. We can even extend these functions to $\mathbb{R}$ by asking that v_- (resp. v_+) is the graph of $G^{-1}(\mathcal{V}(G(y)))$ (resp. $G(\mathcal{V}(G^{-1}(y)))$).
Because E is G-ordered, we have

$$G^{-1}(\{z \in E, \pi(z) \leq \pi \circ G(y)\}) = \{z \in E; \pi(z) \leq \alpha\}.$$

Hence $\{z \in E; \pi(z) < \alpha\}$ is in the connected component of $\mathbb{R}^2 \backslash G^{-1}(\mathcal{V}(G(y)))$ that is under v_-. Using some similar arguments, we finally obtain

$$\forall t \in (-\infty, \alpha) \cap \pi(E), v_+(t) < \psi(t) < v_-(t)$$

and

$$\forall t \in (\alpha, +\infty) \cap \pi(E), v_-(t) < \psi(t) < v_+(t).$$

Using the invariance by integer translation of E (i.e. $E + (1, 0) = E$) and the fact that the graphs of v_- and v_+ restricted to $[\alpha-1, \alpha+1]$ are contained in $\mathbb{R}\times[-K, K]$, we deduce that $E \subset \mathbb{R} \times [-K, K]$.

We will now add a condition to define $\mathcal{U}$. Let $L > \frac{1}{\varepsilon}$ be a real number such that

$$\forall x \in F^{-1}(\mathbb{R} \times [-K - \frac{1}{\varepsilon}, K + \frac{1}{\varepsilon}]) \cup (\mathbb{R} \times [-K - \frac{1}{\varepsilon}, K + \frac{1}{\varepsilon}])$$

we have,

$$\max\{\left|\frac{\partial f_2}{\partial r}(x)\right|, \left|\frac{\partial f_1}{\partial \theta}(x)\right|\} < L.$$

Then we ask that every lift G of an element $g = (g_1, g_2)$ of $\mathcal{U}$ (in addition to the other conditions we gave before that) satisfies

- $\forall x \in G^{-1}(\mathbb{R} \times [-K - \frac{1}{\varepsilon}, K + \frac{1}{\varepsilon}]) \cup (\mathbb{R} \times [-K - \frac{1}{\varepsilon}, K + \frac{1}{\varepsilon}])$
$$\max\left\{\left|\frac{\partial g_2}{\partial r}(x)\right|, \left|\frac{\partial g_1}{\partial \theta}(x)\right|\right\} < L;$$
- and
$$\forall x \in G^{-1}(\mathbb{R} \times [-K - \frac{1}{\varepsilon}, K + \frac{1}{\varepsilon}]) \cup (\mathbb{R} \times [-K - \frac{1}{\varepsilon}, K + \frac{1}{\varepsilon}]), \frac{\partial g_1}{\partial r}(x) > \varepsilon.$$

Let us now consider $y = (\alpha, \psi(\alpha)) \in E$. Repeating the same argument than before, we know that

$$\forall t \in \pi(E), \min\{v_-(t), v_+(t)\} \le \psi(t) \le \max\{v_-(t), v_+(t)\}.$$

Note that $v_-'(t) = -\frac{\partial g_1}{\partial \theta}(t, v_-(t))\left(\frac{\partial g_1}{\partial r}(t, v_-(t))\right)^{-1}$ and then for every $t \in [\alpha - 1, \alpha + 1]$, $|v_-'(t)| < \frac{L}{\varepsilon}$.

Moreover, we have $v_+'(t) = \frac{\partial g_2}{\partial r}(G^{-1}(t, v_+(t)))\left(\frac{\partial g_1}{\partial r}(G^{-1}(t, v_+(t)))\right)^{-1}$ and then for every $t \in [\alpha - 1, \alpha + 1]$, $|v_+'(t)| < \frac{L}{\varepsilon}$.

We introduce the notation $C = \frac{L}{\varepsilon}$. We have then: $\forall t \in [\alpha-1, \alpha+1], |\psi(t)-\psi(\alpha)| \le \max\{|v_-(t) - \psi(\alpha)|, |v_+(t) - \psi(\alpha)|\} \le C|t - \alpha|$.

This implies that ψ is C-Lipschitz. ∎

3.3. **Aubry-Mather sets.**

Definition 3.25. Let F be a lift of a conservative twist map f. An *Aubry Mather set* for F is a closed F-ordered set.

The Aubry-Mather set is *minimizing* if every orbit contained in it is minimizing.

We noticed that any F-ordered set has a rotation number.

Notation 3.26. If E is an Aubry-Mather set, we denote by $\rho(E)$ its rotation number. The Aubry-Mather set E is said to be *rational* (resp. *irrational*) if $\rho(E)$ is rational (resp. irrational).

Proposition 3.27. *Let E be an Aubry-Mather set. For every $\varepsilon > 0$, there exists a neighborhood U of E that is invariant by the integer translations $(\theta, r) \mapsto (\theta + k, r)$ for $k \in \mathbb{Z}$ and such that every Aubry-Mather set $\mathcal{E}$ that meets U satisfies: $|\rho(E) - \rho(\mathcal{E})| < \varepsilon$.*

Proof We deduce from Proposition 3.24 that E is contained in some strip $\mathcal{K} = \mathbb{R} \times [-K, K]$. On such a strip, every DF^k is uniformly bounded.

Let $\mathcal{E}$ be an Aubry-Mather set that meets the same strip $\mathcal{K}$. Let (θ_k, r_k) be an orbit in E and (α_k, β_k) be an orbit in $\mathcal{E}$. We deduce from proposition 3.23 that for every $k \in \mathbb{Z}$, we have:

$$|\theta_k - \theta_0 + k\rho(E)| \leq 1 \quad \text{and} \quad |\alpha_k - \alpha_0 - k\rho(\mathcal{E})| \leq 1.$$

We deduce

$$|\rho(E) - \rho(\mathcal{E})| \leq \frac{2}{k} + \frac{|\theta_k - \alpha_k|}{k} + \frac{|\theta_0 - \alpha_0|}{k}.$$

Fixing $k > \frac{4}{\varepsilon}$ large enough, we choose a neighborhood U of E that is invariant by the integer translations $(\theta, r) \mapsto (\theta + k, r)$ for $k \in \mathbb{Z}$ and such that for every $y = (\alpha, \beta) \in U$, there exists $x = (\theta, r) \in E$ that satisfies $|\theta - \alpha| < \frac{\varepsilon}{4}$ and $\|F^k(x) - F^k(y)\| < \frac{\varepsilon}{4}$. Then for every Aubry-Mather set $\mathcal{E}$ that meets U, we have $|\rho(E) - \rho(\mathcal{E})| < \varepsilon$. □

Proposition 3.28. *Let F be a lift of a conservative twist map f. Then for every $\alpha \in \mathbb{R}$, there exists at least one minimizing Aubry-Mather set with rotation number α.*

Proof If $\alpha = \frac{p}{q} \in \mathbb{Q}$ is rational, we have proved in Proposition 3.14 the existence of a (q, p)-minimizer (θ_k). Then the corresponding F-orbit (θ_k, r_k) is minimizing and we deduce from Corollary 3.13 that $E = \{(\theta_k, r_k)\} + \mathbb{Z} \times \{0\}$ is a minimizing Aubry-Mather set with rotation number $\frac{p}{q}$.

If $\alpha \in \mathbb{R} \backslash \mathbb{Q}$ is irrational, we consider a sequence $(\frac{p_n}{q_n})$ of rational numbers that converge to α and for every n a (q_n, p_n)-minimizing orbit $(\theta_k^n, r_k^n)_{k \in \mathbb{Z}}$. As $\theta_{q_n}^n = \theta_0^n + p_n$, there exists $k_n \in [0, p_n - 1]$ such that $\theta_{k_n+1}^n - \theta_{k_n}^n \in [0, \frac{p_n}{q_n}]$. Replacing $(\theta_k^n, r_k^n)_{k \in \mathbb{Z}}$ by $(\alpha_k^n, \beta_k^n) = (\theta_{k+k_n}^n - [\theta_{k_n}^n], r_k^n)_{k \in \mathbb{Z}}$ that is also a (q_n, p_n)-minimizer, we obtain a sequence of minimizers so that:

- $\alpha_0^n \in [0, 1]$;
- $(\alpha_1^n - \alpha_0^n)_{n \in \mathbb{N}}$ is bounded and then $(\alpha_0^n, \beta_0^n)_{n \in \mathbb{N}}$ is also bounded;
- the rotation number of the (q_n, p_n)-minimizer $(\alpha_k^n, \beta_k^n)_{k \in \mathbb{Z}}$ is $\frac{p_n}{q_n}$.

We then extract a subsequence so that $(\alpha_0^n, \beta_0^n)_{n \in \mathbb{N}}$ converges to some (θ, r). Then the orbit of (θ, r) is also minimizing. If $E = \text{Closure}\big(\{F^k(\theta, r) + (j, 0); k, j \in \mathbb{Z}\}\big)$, then we deduce from Proposition 3.20 that E is F ordered and then E is a minimizing Aubry-Mather set. We deduce from Proposition 3.27 that $\rho(E) = \alpha$. □

3.4. **Further results on Aubry-Mather sets.** In [15], it is proved that the closure of the union of the $\mathbb{Z} \times \{0\}$-translated sets of every minimizing orbit is an Aubry-Mather set (hence every minimizing orbit has a rotation number).

In [10], more precise results concerning the minimizing Aubry-Mather sets are proved. Let us explain them.

We denote the set of points $(\theta, r) \in \mathbb{R}^2$ having a minimizing orbit by $\mathcal{M}(F)$. Then it is closed and $p(\mathcal{M}(F)) \subset \mathbb{A}$ is closed too. The rotation number $\rho : \mathcal{M}(F) \to \mathbb{R}$ is continuous and for every $\alpha \in \mathbb{R}$, the set

$$\mathcal{M}_\alpha(F) = \{x \in \mathcal{M}(F), \rho(x) = \alpha\}$$

is non-empty.

If α is irrational, then $K_\alpha = p(\mathcal{M}_\alpha(F)) \subset \mathbb{A}$ is the graph of a Lipschitz map above a compact subset of $\mathbb{T}$. Moreover, there exists a bi-Lipschitz orientation preserving homeomorphims $h : \mathbb{T} \to \mathbb{T}$ such that

$$\forall x \in K_\alpha, h(\pi(x)) = \pi(f(x)).$$

Hence K_α is:

- either not contained in an invariant loop and then is the union of a Cantor set C_α on which the dynamics is minimal and some homoclinic orbits to C_α;
- or is an invariant loop. In this case the dynamics restricted to K_α can be minimal or Denjoy.

If $\alpha = \frac{p}{q}$ is rational, then $\mathcal{M}_\alpha(F)$ is the disjoint union of 3 invariant sets:

- $\mathcal{M}_\alpha^{\rm per}(F) = \{x \in \mathcal{M}_\alpha(F), \pi \circ F^q(x) = \pi(x) + p\}$;
- $\mathcal{M}_\alpha^{+}(F) = \{x \in \mathcal{M}_\alpha(F), \pi \circ F^q(x) > \pi(x) + p\}$;
- $\mathcal{M}_\alpha^{-}(F) = \{x \in \mathcal{M}_\alpha(F), \pi \circ F^q(x) < \pi(x) + p\}$.

The two sets $K_\alpha^+ = p(\mathcal{M}_\alpha^{\rm per}(F) \cup \mathcal{M}_\alpha^+(F))$ and $K_\alpha^- = p(\mathcal{M}_\alpha^{\rm per}(F) \cup \mathcal{M}_\alpha^-(F))$ are then invariant Lipschitz graphs above a compact part of $\mathbb{T}$. The points of $p(\mathcal{M}_\alpha^+(F) \cup \mathcal{M}_\alpha^-(F))$ are heteroclinic orbits to some periodic points contained in $p(\mathcal{M}_\alpha^{\rm per}(F))$.

4. Ergodic theory for minimizing measures

4.1. **Green bundles.** We fix a lift F of a conservative twist map. As before $\mathcal{M}(F)$ is the set of points whose orbit is minimizing. We use some new notations.

Notations 4.1.

- $V(x) = T_x\mathcal{V}(x) = \{0\} \times \mathbb{R} \subset T_x\mathbb{R}^2$, and for $k \neq 0$, we have
$$G_k(x) = Df^k(f^{-k}x)V(f^{-k}x)\,;$$
- the slope of G_k (when defined) is denoted by s_k:
$$G_k(x) = \{(\delta\theta, s_k(x)\delta\theta); \delta\theta \in \mathbb{R}\}\,;$$
- if γ is a real Lipschitz function defined on $\mathbb{T}$ or $\mathbb{R}$, then
$$\gamma'_+(x) = \limsup_{y,z\to x, y\neq z} \frac{\gamma(y) - \gamma(z)}{y - z} \quad \text{and} \quad \gamma'_-(t) = \liminf_{y,z\to x, y\neq z} \frac{\gamma(y) - \gamma(z)}{y - z}.$$

We introduce now a set, called Green(f). We will see very soon that we can define two natural invariant sub bundles in tangent lines at every point of Green(f), that will be very useful in our further study. An important result (see Corollary 4.7) is that all the minimizing Aubry-Mather sets are contained in Green(f).

Notation 4.2. We denote by Green(f) the set of the points of $\mathbb{A}$ such that *along the whole orbit of these points*, we have

$$\forall n \geq 1, s_{-n}(x) < s_{-n-1}(x) < s_{n+1}(x) < s_n(x).$$

Definition 4.3. If $x \in$ Green(f), the two *Green bundles* at x are $G_+(x), G_-(x) \subset T_x(\mathbb{R}^2)$ with slopes s_-, s_+ where $s_+(x) = \lim_{n\to+\infty} s_n(x)$ and $s_-(x) = \lim_{n\to+\infty} s_{-n}(x)$.

The two Green bundles satisfy the following properties

Proposition 4.4. *Let f be a conservative twist map.*

- *Then the two Green bundles defined on* Green(f) *are invariant under Df: $Df(G_\pm) = G_\pm \circ f$;*

- *we have* $s_+ \geq s_-$;
- *the map* $s_- : \mathrm{Green}(f) \to \mathbb{R}$ *is lower semi-continuous and the map* $s_+ : \mathrm{Green}(f) \to \mathbb{R}$ *is upper semi-continuous;*
- *hence* $\{G_- = G_+\}$ *is a* G_δ *subset of* $\mathrm{Green}(f)$ *and* $s_- = s_+$ *is continuous at every point of this set.*

Exercise 4.5. Prove Proposition 4.4.

Theorem 7. *Let* $f : \mathbb{A} \to \mathbb{A}$ *be a conservative twist map and let* $(x_n)_{n\in\mathbb{Z}}$ *be the orbit of a point* $x = x_0$. *The following assertions are equivalent:*

(0) $x \in \mathrm{Green}(f)$;
(1) *the projection of every finite segment of the orbit of* x *is locally minimizing among the segments of points (of* $\mathbb{R}$*) that have same length and same endpoints;*
(2) *along the orbit of* x*, we have for every* $k \geq 1$, $s_k > s_{-1}$;
(3) *along the orbit of* x*, we have for every* $k \geq 1$, $s_{-k} < s_1$;
(4) *there exists a field of half-lines* $\delta_+ \subset T\mathbb{A}$ *along the orbit of* x *such that:*
  - δ_+ *is invariant by* Df: $Df(\delta_+) = \delta_+ \circ f$;
  - $D\pi \circ \delta_+ = \mathbb{R}_+$ *(*δ_+ *is oriented to the right).*

Remarks 4.6. (1) Observe that in the point *(4)*, you cannot replace 'field of half-lines' by 'field of lines'. Indeed, along the orbit of every point that is not periodic you can find an invariant field of lines that is transverse to the vertical.
(2) in fact, in the proof, we will see that if we denote by d_+ the slope of δ_+, we necessarily have $s_- \leq d_+ \leq s_+$.

We postpone the proof of Theorem 7 to subsection 5.3.

Corollary 4.7. *Let* f *be a conservative twist map. Then*

- *every accumulation point of an Aubry-Mather set is in* $\mathrm{Green}(f)$*;*
- *every minimizing orbit is in* $\mathrm{Green}(f)$.

Proof • Assume that x is an accumulation point of an invariant Aubry-Mather set E. We look at the action of DF on the half-lines $\mathbb{R}_+v$ that are in the tangent space to $\mathbb{R}^2$ along the orbit of x. As E is the graph of a Lipschitz map $\gamma : F \to \mathbb{R}$ and x is an accumulation point of E, we have for every $k \in \mathbb{Z}$:

$$\gamma'_+(\pi(F^k x)) = \limsup_{y,z\to\pi(F^k x), y,z\in E, y\neq z} \frac{\gamma(y) - \gamma(z)}{y - z}.$$

This bundle $\Gamma_+ = \mathbb{R}_+(1, \gamma'_+)$ in half-lines is transverse to the vertical bundle and invariant by DF. We use the characterization *(4)* of Theorem 7 to conclude.
• The second point of the corollary is a direct consequence of the characterization *(1)* of $\mathrm{Green}(f)$. □

An interesting consequence of the characterization *(1)* of $\mathrm{Green}(f)$ is

Corollary 4.8. *The set* $\mathrm{Green}(f)$ *is closed.*

When x is a generic point in the support of some hyperbolic measure, G_- is the stable bundle and G_+ is the unstable one:

Proposition 4.9. (Dynamical criterion) *Assume that $x \in \mathrm{Green}(F)$ has its orbit contained in some strip $\mathbb{R} \times [-K, K]$ (for example $x \in \mathcal{M}(F)$ or x is in some Aubry-Mather set) and that $v \in T_x\mathbb{A}$. Then*

- *if* $\liminf\limits_{n\to+\infty} |D(\pi \circ F^n)(x)v| < +\infty$, *then* $v \in G_-(x)$;
- *if* $\liminf\limits_{n\to+\infty} |D(\pi \circ F^{-n})(x)v| < +\infty$, *then* $v \in G_+(x)$.

Proof We use a symplectic change of linear coordinates along the orbit of x in such a way that the horizontal subspace is now G_- and the vertical subspace doesn't change.

As the orbit of x is contained in some strip $\mathbb{R} \times [-K, K]$, the slopes s_{-1} and s_1 of $G_{-1} = DF^{-1}(V \circ F)$ and $G_1 = DF(V \circ F^{-1})$ are uniformly bounded along the orbit of x. Hence $s_- \in [s_{-1}, s_1]$ is also uniformly bounded and so the changes of basis $P = \begin{pmatrix} 1 & 0 \\ s_- & 1 \end{pmatrix}$ as $P^{-1} = \begin{pmatrix} 1 & 0 \\ -s_- & 1 \end{pmatrix}$ are also uniformly bounded. Then the matrix of $DF^n(x)$ in this new basis is

$$\begin{pmatrix} b_n(x)(s_-(x) - s_{-n}(x)) & b_n(x) \\ 0 & (s_n(F^n x) - s_-(F^n x))b_n(x) \end{pmatrix}$$

We know that the determinant is $1 = (b_n(x))^2(s_-(x)-s_{-n}(x))(s_n(F^n x)-s_-(F^n x))$, that $|s_n(F^n x) - s_-(F^n x)| \le (s_1(F^n x) - s_{-1}(F^n x))$ is uniformly bounded and that $\lim\limits_{n\to+\infty}(s_-(x) - s_{-n}(x)) = 0$; hence $\lim\limits_{n\to+\infty}|b_n(x)| = +\infty$.

Let now v be a vector in $T_x\mathbb{A}$. We denote by (v_1, v_2) its coordinates in the new base we defined just before. Then we have: $|D(\pi \circ F^n)(x)v| = |b_n(x)|.|(s_-(x) - s_{-n}(x))v_1 + v_2|$. As $\lim\limits_{n\to+\infty}(s_-(x) - s_{-n}(x)) = 0$ and $\lim\limits_{n\to+\infty}|b_n(x)| = +\infty$, we deduce that if $v_2 \neq 0$ (i.e. if $v \notin G_-(x)$), then $\lim\limits_{n\to+\infty}|D(\pi \circ F^n)(x)v| = +\infty$. □

Exercise 4.10. Let $k \in [1, \infty]$. Let us admit that there exists a dense G_δ subset $\mathcal{G}$ of the set of the C^k conservative twist maps such that for every $f \in \mathcal{G}$, for every periodic point x for f, if we denote by N the period of x, then we have:

- the eigenvalues of $Df^N(x)$ are distinct;
- every heteroclinic intersection of two hyperbolic periodic orbits is transverse.

Prove that every $f \in \mathcal{G}$ has no rational invariant graph.
Hint: using the invariance of the Green bundle G_-, prove that every periodic point contained in such a rational invariant graph has to be hyperbolic.

4.2. **Lyapunov exponents and Green bundles.** We have noticed that if a measure μ with support contained in $\mathrm{Green}(f) \cap (\mathbb{R} \times [-K, K])$ is hyperbolic, then we have μ-almost everywhere: $E^s = G_-$ and $E^u = G_+$. In this case, we have $G_- \neq G_+$ μ-almost everywhere.
We will prove the reverse implication.

Theorem 8. (M.-C. Arnaud) *Let f be a conservative twist map and let μ be a measure that is ergodic for f, with compact support and such that* $\mathrm{supp}\mu \subset \mathrm{Green}(f)$. *Then $d = \dim(G_- \cap G_+)$ is constant μ-almost everywhere and*

- *if $d = 0$, the measure μ is hyperbolic with Lyapunov exponents $-\lambda(\mu) < \lambda(\mu)$ given by:* $\lambda(\mu) = \frac{1}{2}\int \log\left(\frac{s_+ - s_{-1}}{s_- - s_{-1}}\right) d\mu = \frac{1}{2}\int \log\left(1 + \frac{s_+ - s_-}{s_- - s_{-1}}\right) d\mu$;

- *if $d = 1$, the Lyapunov exponents of μ are zero.*

Remark 4.11. Observe that the first part of Theorem 8 says to us that the more distant the Green bundles are, the greater the positive Lyapunov exponent is.
A general result for hyperbolic measures of smooth dynamics is that when the stable and unstable bundles are close together, the Lyapunov exponents are close to zero (see for example [5]).
The reverse result is not true in general and what we prove is then specific to the case of the twist maps. Consider for example the Dirac measure at $(0,0)$ that is invariant by the linear map of $\mathbb{R}^2$ with matrix $\begin{pmatrix} e^{\varepsilon} & 0 \\ 0 & e^{-\varepsilon} \end{pmatrix}$. Then the unstable and stable bundles are $\mathbb{R} \times \{0\}$ and $\{0\} \times \mathbb{R}$ that are far from each other. But the Lyapunov exponents ε, $-\varepsilon$, can be very close to 0.

Proof As the dynamics is symplectic, the sum of the Lyapunov exponents is $\int \log(\det(Df))d\mu = 0$, hence there are two Lyapunov exponents $-\lambda(\mu) \leq \lambda(\mu)$. Either these two Lyapunov exponents are zero or the measure is hyperbolic.
We have noticed that when μ is hyperbolic, then $G_- = E^s \neq G_+ = E^u$ μ-almost everywhere. Hence when $d = 1$, the Lyapunov exponents are zero. Assume now that $d = 0$. Using a bounded change of basis along a generic point for μ as in the proof of the dynamical criterion, we obtain that $Df(x)_{|G_-(x)}$ is represented by $b_1(x)(s_-(x)-s_{-1}(x))$ and that $Df(x)_{|G_+(x)}$ is represented by $b_1(x)(s_+(x)-s_{-1}(x))$. Hence if $v_\pm$ is a base of $G_\pm$, we have:

$$\begin{aligned} \lambda(v_\pm) &= \lim_{n\to+\infty} \frac{1}{n} \log\left(\|Df^n(x)v_\pm\|\right) \\ &= \lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} \log\left(b_1(f^jx)(s_\pm(f^jx) - s_{-1}(f^jx))\right) \end{aligned}$$

and then by Birkhoff ergodic theorem

$$\lambda(v_+) - \lambda(v_-) = \lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} \log\left(\frac{s_+(f^jx) - s_{-1}(f^jx)}{s_-(f^jx) - s_{-1}(f^jx)}\right) = \int \log\left(\frac{s_+ - s_{-1}}{s_- - s_{-1}}\right) d\mu.$$

As $s_+ > s_-$ μ-almost everywhere, we have then $\lambda(v_+) > \lambda(v_-)$. Hence we are in the case of an hyperbolic measure. Then $G_+ = E^u$ and $G_- = E^s$ and $\lambda(v_+) = \lambda(\mu)$, $\lambda(v_-) = -\lambda(\mu)$ and thus $2\lambda(\mu) = \int \log\left(\frac{s_+ - s_{-1}}{s_- - s_{-1}}\right) d\mu$.

$\square$

We have seen in subsection 2.3 that the Lyapunov exponents of the measures that are on the irrational invariant curves are zero. But Patrice Le Calvez proved that for general conservative twist maps, many Aubry-Mather sets are (uniformly) hyperbolic, and then are not curves.

Proposition 4.12. (P. Le Calvez, [24]**)** *Let $k \in [1, \infty]$. There exists a dense G_δ subset $\mathcal{G}_k$ of the set of the C^k conservative twist maps such that for any $f \in \mathcal{G}_k$, there exists an open and dense subset $U(f) \subset \mathbb{R}$ such that the minimizing Aubry-Mather sets having their rotation number in $U(f)$ are uniformly hyperbolic.*

It may even happen that all the minimizing Aubry-Mather sets are hyperbolic (see [16]).

Proposition 4.13. (D. L. Goroff) *For $|\lambda| > \frac{\sqrt{1+\pi^2}}{\pi}$, the union of the minimizing Aubry-Mather sets for the standard map f_λ is uniformly hyperbolic.*

Proof We assume that $|\lambda| > \frac{\sqrt{1+\pi^2}}{\pi}$.
The standard map with parameter λ is defined by $f_\lambda(\theta, r) = (\theta + r + \lambda \sin 2\pi\theta, r + \lambda \sin 2\pi\theta)$ and has the generating function $S_\lambda(\theta, \Theta) = \frac{1}{2}(\Theta - \theta)^2 - \frac{\lambda}{2\pi}\cos 2\pi\theta$.
Let E be a minimizing Aubry-Mather set for f_λ. Observe that $F_\lambda(\theta, r+1) = F_\lambda(\theta, r) + (1, 1)$. Hence we can assume that the rotation number of E is in $(-1, +1)$. Then by the inequalities that we recalled in subsection 2.2 for circle homeomorphisms, we have for every orbit (θ_n, r_n) in E: $\theta_n - \theta_{n+1} \in (-1, 1)$ and $\theta_n - \theta_{n-1} \in (-1, 1)$ have opposite signs.
As $0 = \frac{\partial S_\lambda}{\partial \theta}(\theta_n, \theta_{n+1}) + \frac{\partial S_\lambda}{\partial \Theta}(\theta_{n-1}, \theta_n) = \theta_n - \theta_{n-1} + \lambda \sin 2\pi\theta_n + \theta_n - \theta_{n-1}$, we deduce that $\lambda \sin 2\pi\theta_n \in (-1, 1)$ i.e. $|\sin 2\pi\theta_n| < \frac{1}{|\lambda|}$. This implies that $|\cos 2\pi\theta_n| > \sqrt{1 - \frac{1}{\lambda^2}}$.
Moreover, as the orbit is minimizing, we have

$$0 \leq \frac{\partial^2 S_\lambda}{\partial \theta^2}(\theta_n, \theta_{n+1}) + \frac{\partial^2 S_\lambda}{\partial \Theta^2}(\theta_{n-1}, \theta_n) = 2 + 2\pi\lambda \cos 2\pi\theta_n$$

and then $2 \geq -2\pi\lambda \cos 2\pi\theta_n$. As $2\pi|\lambda||\cos 2\pi\theta_n| \geq 2\pi|\lambda|\sqrt{1 - \frac{1}{\lambda^2}} = 2\pi\sqrt{\lambda^2 - 1} > \frac{2\pi}{\pi} = 2$, we have $2\pi\lambda \cos 2\pi\theta_n > 0$ and then $2\pi\lambda \cos 2\pi\theta_n > 2$.

We can now compute $Df(\theta, r) = \begin{pmatrix} 1 + 2\pi\lambda\cos 2\pi\theta & 1 \\ 2\pi\lambda \cos 2\pi\theta & 1 \end{pmatrix}$. Observe that $1 + 2\pi\lambda \cos 2\pi\theta_n > 3$ and $2\pi\lambda \cos 2\pi\theta_n > 2$. Hence if $C = \{(v_1, v_2) \in \mathbb{R}^2; v_1.v_2 \geq 0\}$, we have $Df(C) \subset C$ and $\forall v \in C, \|Df(v)\| \geq \sqrt{2}\|v\|$ along the orbit (θ_n, r_n).
We have too $(Df(\theta, r))^{-1} = \begin{pmatrix} 1 & -2\pi\lambda\cos 2\pi\theta \\ -1 & 1 + 2\pi\lambda \cos 2\pi\theta \end{pmatrix}$. Hence if $C' = \{(v_1, v_2) \in \mathbb{R}^2; v_1.v_2 \leq 0\}$, we have along the orbit (θ_n, r_n): $Df^{-1}(C') \subset C'$ and $\forall v \in C', \|Df^{-1}(v)\| \geq \sqrt{2}\|v\|$.

This implies the wanted result.

□

A. Katok proved that the union of the hyperbolic Aubry-Mather sets has zero Lebesgue measure (see [20]). This can be compared to K.A.M. theory that gives in general a union of invariant circles with positive Lebesgue measure.

Theorem 8 can be more precise in the case of uniform hyperbolicity.

Proposition 4.14. (M.-C. Arnaud) *Let M be a compact invariant set by a conservative twist map that is contained in* Green(f). *Then E is uniformly hyperbolic if and only if at every point of M, the two Green bundles G_- and G_+ are transverse.*

We postpone the proof of Proposition 4.14 to subsection 5.4. We don't know if there exist examples of Aubry-Mather sets that are non-uniformly hyperbolic.

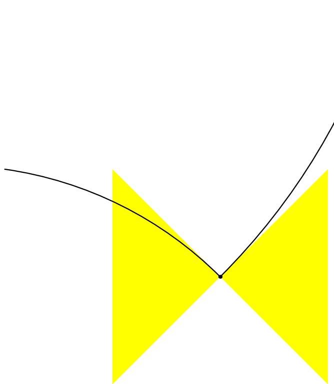
Question 4.15. Does there exist a conservative twist map that has a non-uniformly hyperbolic Aubry-Mather set?

4.3. **Lyapunov exponents and shape of the Aubry-Mather sets.** In the previous subsection, we compared the size of the Lyapunov exponents for the ergodic measures with support in Green(f) with the distance between the two Green bundles. We ask now if we can see a link between the shape of the support of such a measure and the Lyapunov exponents.

Definition 4.16. Let $M \subset \mathbb{A}$ be a subset of $\mathbb{A}$ and $x \in M$ a point of M. The *paratangent cone* to M at x is the cone of $T_x\mathbb{A}$ denoted by $P_M(x)$ whose elements are the limits

$$v = \lim_{n\to\infty} \frac{x_n - y_n}{\tau_n}$$

where (x_n) and (y_n) are sequences of elements of M converging to x, (τ_n) is a sequence of elements of $\mathbb{R}_+^*$ converging to 0, and $x_n - y_n \in \mathbb{R}^2$, refers to the unique lift of this element of $\mathbb{A}$ that belongs to $[-\frac{1}{2}, \frac{1}{2}[^2$.
Here we draw the paratangent cone to a curve at a corner:



We will say that M is C^1*-regular* at x if there exists a line D of $T_x\mathbb{A}$ such that $P_M(x) \subset D$.
If M is not C^1-regular at x, we say that M is C^1*-irregular* at x.

Remark 4.17. Observe that the graph of a Lipschitz map γ is C^1-regular if and only if γ is C^1.

Notation 4.18. We denote the set of the slopes of the elements of $P_M(x)$ by $p_M(x)$.

Theorem 9. (M.-C. Arnaud) *Let μ be an ergodic measure for a conservative twist map with support in some irrational Aubry-Mather set. Then*

- *either the Lyapunov exponents of μ are zero and* suppμ *is C^1-regular μ-almost everywhere;*
- *or μ is hyperbolic and* suppμ *is C^1-irregular μ-almost everywhere.*

Corollary 4.19. *If the support of an ergodic measure has an irrational rotation number and is contained in some (non necessarily invariant) C^1 curve, then its Lyapunov exponents are zero.*

Proof Let M be an irrational Aubry-Mather set and let μ be the unique ergodic measure with support in M. Looking at the proof of Corollary 4.7 (see also Remark 4.6), we deduce easily that for μ-almost every point $x \in \text{supp}\mu$, we have

$$s_-(x) \leq p_M(x) \leq s_+(x).$$

Assume that the Lyapunov exponents of μ are zero. Then, by theorem 8, we have μ-almost everywhere $G_- = G_+$ i.e. $s_- = s_+$ and then $P_M(x)$ is contained in a line. This exactly means that $\mathrm{supp}(\mu)$ is C^1-regular μ-almost everywhere.

Now we assume that the Lyapunov exponents of μ are non zero: $-\lambda(\mu) < \lambda(\mu)$. The set where $\mathrm{supp}(\mu)$ is C^1-regular is measurable and invariant by f. Hence either μ is C^1-regular μ-almost everywhere or μ is C^1-irregular μ-almost everywhere. Assume that μ is C^1-regular μ-almost everywhere.

We will prove the following result (we use for $h'_\pm$ the notations 4.1) in subsection 5.5.

Proposition 4.20. (M.-C. Arnaud) *Let $h : \mathbb{T} \to \mathbb{T}$ be a bi-Lipschitz orientation preserving homeomorphism with irrational rotation number. We denote by μ its unique invariant measure and assume that h is C^1-regular μ-almost everywhere. Then uniformly in $\theta \in \mathbb{T}$, we have*

$$\lim_{n\to+\infty} \frac{1}{n} \log{(h^n)'_+} = \lim_{n\to+\infty} \frac{1}{n} \log{(h^n)'_-} = 0.$$

Let us explain how we deduce the wanted result. As M is an Aubry-Mather set, it is the graph of a Lipschitz map $\gamma : \pi(M) \to \mathbb{R}$. We consider the projected-restricted dynamics to M, which is $h : \pi(M) \to \pi(M)$ that is defined by $h(\theta) = \pi \circ f(\theta, \gamma(\theta))$. We denote again by μ the projected measure $\pi_*\mu$ of μ, that is the unique invariant measure by h. We extend h linearly in its gaps in such a way we obtain a bi-Lipschitz homeomorphism h of $\mathbb{T}$. Because μ is C^1-regular μ-almost everywhere, h is also C^1-regular μ-almost everywhere and we deduce from Proposition 4.20 that uniformly in $\theta \in \mathbb{T}$, we have

$$\lim_{n\to+\infty} \frac{1}{n} \log{(h^n)'_+} = \lim_{n\to+\infty} \frac{1}{n} \log{(h^n)'_-} = 0.$$

Observe that $Df^n(\theta, \gamma(\theta)).(1, \gamma'_+(\theta)) = \log\left((h^n)'_+(\theta)\right)(1, \gamma'_+(h^n\theta))$. We deduce that the Lyapunov exponent associated to the vector $(1, \gamma'_+(\theta))$ is zero, which is impossible if the measure is hyperbolic.

$\square$

Theorem 10. *Let M be an irrational Aubry-Mather set of a conservative twist map f of $\mathbb{A}$. Then M is uniformly hyperbolic if and only if at every $x \in M$, M is C^1-irregular.*

Proof As $s_- \leq p_M \leq s_+$, if M is C^1-irregular everywhere, then $G_- \neq G_+$ at every point of M and by Proposition 4.14, M is uniformly hyperbolic.

Assume now that M is uniformly hyperbolic. At first, let us notice that such a M cannot be a curve because of Theorem 2.

Hence M is a Cantor and the dynamics on M is Lipschitz conjugate to the one of a Denjoy counter-example on its minimal invariant set. Then we consider two points $x \neq y$ of M such that there exists an open interval $I \subset \mathbb{T}$ whose ends are $\pi(x)$ and $\pi(y)$ and which doesn't meet $\pi(M)$: $I \cap \pi(M) = \emptyset$. We deduce from the dynamics of the Denjoy counter-examples (see [18]) that:

- the positive and negative orbits of x and y under f are dense in M;
- $\lim\limits_{n\to+\infty} d(f^n x, f^n y) = \lim\limits_{n\to+\infty} d(f^{-n}x, f^{-n}y) = 0$.

As M is uniformly hyperbolic, we can define a local stable and unstable laminations containing M, W^s_{loc} and W^u_{loc}. Then for large enough n, $f^n x$ and $f^n y$ belong to the same local stable leaf, and $f^{-n}x$ and $f^{-n}y$ belong to the same local unstable leaf. Hence, because

$$\lim_{n\to+\infty} d(f^n x, f^n y) = \lim_{n\to+\infty} d(f^{-n}x, f^{-n}y) = 0,$$

for large enough n, the vector joining $f^n x$ to $f^n y$ (resp. $f^{-n}x$ to $f^{-n}y$) is close the stable bundle E^s (resp. the unstable bundle E^u).

Let now $z \in M$ be any point. Then there exist two sequences (i_n) and (j_n) of integers which tends to $+\infty$ and are such that:

$$\lim_{n\to+\infty} f^{i_n}x = \lim_{n\to+\infty} f^{i_n}y = \lim_{n\to+\infty} f^{-j_n}x = \lim_{n\to+\infty} f^{-j_n}y = z.$$

The direction of the "vector" joining $f^{i_n}x$ to $f^{i_n}y$ tends to $E^s(z)$ and the direction of the vector joining $f^{-j_n}x$ to $f^{-j_n}y$ tends to $E^u(z)$. Hence: $E^u(z)\cup E^s(z) \subset P_M(z)$ and M is C^1-irregular at z. □

When drawing irrational Aubry-Mather sets that are Cantor sets with the help of a computer, we never observe some angles on these sets. That is why we raise the question:

Question 4.21. Is it possible to draw (with a computer) some irrational Aubry-Mather sets that have some "corners"?

Remark 4.22. There is a difficulty in 'seing' these corners. On the K.A.M. invariant graphs, the two Green bundles coincide. As $s_+ - s_-$ is non-negative and upper-semicontinuous, we deduce that close to the KAM curves, the paratangent cones are very thin, and thus very hard to detect.

5. Complements

5.1. **Proof of the equivalent definition of a conservative twist map.** We recall the statement.

Proposition. *Let $F : \mathbb{R}^2 \to \mathbb{R}^2$ be a C^1 map. Then F is a lift of a conservative twist map $f : A \to \mathbb{A}$ if and only if there exists a C^2 function such that*

- $\forall \theta, \Theta \in \mathbb{R}, S(\theta+1, \Theta+1) = S(\theta, \Theta)$;
- *there exists $\varepsilon > 0$ so that for all $\theta, \Theta \in \mathbb{R}$, we have*
$$\varepsilon < -\frac{\partial^2 S}{\partial\theta\partial\Theta}(\theta,\Theta) < \frac{1}{\varepsilon};$$
- $F(\theta, r) = (\Theta, R) \iff R = \frac{\partial S}{\partial \Theta}(\theta,\Theta)$ and $r = -\frac{\partial S}{\partial \theta}(\theta,\Theta)$.

Proof ($\Rightarrow$) Assume that $F : \mathbb{R}^2 \to \mathbb{R}^2$ is the lift of a conservative twist map f such that $\forall x \in \mathbb{A}, \frac{1}{\varepsilon} > D(\pi\circ f)(x)(0,1) > \varepsilon$. Then for every $\theta \in \mathbb{R}$, the map $F_\theta : \mathbb{R} \to \mathbb{R}$ defined by $F_\theta(r) = \pi \circ F(\theta, r)$ satisfies $\frac{1}{\varepsilon} > F'_\theta > \varepsilon$. Hence every map F_θ is a C^1-diffeomorphism of $\mathbb{R}$ and $G : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $G(\theta,\Theta) = (\theta, F_\theta^{-1}(\Theta))$ is a C^1 diffeomorphism.

We introduce the notation $F(\theta,r) = (\Theta(\theta,r), R(\theta,r))$. Note that $G(\theta,\Theta(\theta,r)) = (\theta, r)$ i.e. $F_\theta(r) = \Theta(\theta,r)$. As f is an exact symplectic twist map, we have: $G^*(f^*\lambda - \lambda)$ is exact.

Hence there exists a function $S : \mathbb{R}^2 \to \mathbb{R}$ such that $DS(\theta, \Theta) = R \circ G(\theta, \Theta)d\Theta - F_\theta^{-1}(\Theta)d\theta$. This means exactly that

$$\frac{\partial S}{\partial \Theta}(\theta, \Theta) = R \circ G(\theta, \Theta) \quad \text{and} \quad -\frac{\partial S}{\partial \theta} = F_\theta^{-1}(\Theta);$$

and implies that S is C^2. Thus we have proved the third point of Proposition 1.8.

Let us fix $(\theta, r) \in \mathbb{A}$. We denote by γ the loop of $\mathbb{A}$ defined by $\gamma(t) = (\theta + t, r)$ and by Γ its lift $\Gamma(t) = (\theta + t, r)$. As f is exact symplectic, we have $\int_\gamma f^*\lambda = \int_\gamma \lambda$. Let us use the notation $F \circ \Gamma(t) = (\Theta_t, R_t)$. As f is isotopic to identity, we have $\Theta_1 = \Theta_0 + 1$. Moreover:

$$\begin{aligned} 0 &= \int_\gamma (f^*\lambda - \lambda) = \int_{G\circ\gamma} G^*(f^*\lambda - \lambda) = \int_{G\circ\gamma} dS = \\ &= \int_{(\theta+t,\Theta_t)} dS = S(\theta + 1, \Theta_0 + 1) - S(\theta, \Theta_0)\,; \end{aligned}$$

this gives the first point of Proposition 1.8.

From $\frac{\partial S}{\partial \theta}(\theta, \Theta(\theta, r)) = -r$ we deduce that $\frac{\partial^2 S}{\partial \Theta \partial \theta}(\theta, \Theta(\theta, r)).\frac{\partial \Theta}{\partial r}(\theta, r) = -1$. As $\frac{1}{\varepsilon} > \frac{\partial \Theta}{\partial r}(\theta, r) = D(\pi \circ f)(x)(0, 1) > \varepsilon$, we deduce the second point of Proposition 1.8.

($\Leftarrow$) Assume that S satisfies the conclusions of Proposition 1.8. Because of the second point, the maps $\frac{\partial S}{\partial \theta}(\theta, .)$ and $\frac{\partial S}{\partial \Theta}(., \Theta)$ are C^1-diffeomorphisms of $\mathbb{R}$. Hence the third point allows us to define a diffeomorphism $F : \mathbb{R}^2 \to \mathbb{R}^2$.
From the first point we deduce that $F(\theta+1, r) = F(\theta, r) + (1, 0)$ hence F is the lift of a C^1-diffeomorphism $f : \mathbb{A} \to \mathbb{A}$.

Let us prove that f is a conservative twist map. We use as before the notation $F(\theta, r) = (\Theta(\theta, r), R(\theta, r))$.

From $\frac{\partial S}{\partial \theta}(\theta, \Theta(\theta, r)) = -r$ we deduce that $\frac{\partial^2 S}{\partial \Theta \partial \theta}(\theta, \Theta(\theta, r)).\frac{\partial \Theta}{\partial r}(\theta, r) = -1$ and then we have the twist condition $\varepsilon < \frac{\partial \Theta}{\partial r}(\theta, r) < \frac{1}{\varepsilon}$.

Because $S(\theta + 1, \Theta + 1) = S(\theta, \Theta)$, we can define a C^2-function $s : \mathbb{A} \to \mathbb{R}$ such that for any lift $\tilde{\theta} \in \mathbb{R}$ of θ, we have: $s(\theta, r) = S(\tilde{\theta}, \Theta(\tilde{\theta}, r))$. Then $f^*\lambda - \lambda = ds$ is exact. In particular, f preserves the orientation. As moreover $F(\theta + 1, r) = F(\theta, r) + (1, 0)$, we deduce that f is isotopic to identity. Finally, f is conservative. □

5.2. **Proof that every invariant continuous graph is minimizing.** Let us recall the result due to J. Mather.

Theorem. *Assume that the graph of a continuous map $\psi : \mathbb{T} \to \mathbb{R}$ is invariant by a conservative twist map f. Then for any generating function associated to f, all the orbits contained in the graph of ψ are minimizing.*

Proof Let us introduce the constant $c = \int_0^1 \psi(t)dt$ and let us define the $\mathbb{Z}$-periodic C^1-function η by $\eta(\theta) = \int_0^\theta \psi(t)dt - c\theta$. If S is a generating function of the lift F of f such that $\frac{\partial^2 S}{\partial\theta\partial\Theta} < -\varepsilon$, then we define $W(\theta, \Theta) = S(\theta, \Theta) + c(\theta - \Theta) + \eta(\theta) - \eta(\Theta)$. Observe that $W(\theta + 1, \Theta + 1) = W(\theta, \Theta)$. Moreover, we have proved in Lemma

3.15 that:
$$\lim_{|\Theta-\theta|\to+\infty}\frac{S(\theta,\Theta)}{|\Theta-\theta|}=+\infty.$$
Hence $\displaystyle\lim_{|\Theta-\theta|\to+\infty}\frac{W(\theta,\Theta)}{|\Theta-\theta|}=+\infty$ and hence W has a global minimimum μ. The minimizers of W being critical points, let us look after the critical points of W. We have
$$\frac{\partial W}{\partial\theta}(\theta,\Theta)=\frac{\partial S}{\partial\theta}(\theta,\Theta)+c+\eta'(\theta)=\frac{\partial S}{\partial\theta}(\theta,\Theta)+\psi(\theta);$$
$$\frac{\partial W}{\partial\Theta}(\theta,\Theta)=\frac{\partial S}{\partial\Theta}(\theta,\Theta)-c-\eta'(\Theta)=\frac{\partial S}{\partial\theta}(\theta,\Theta)-\psi(\Theta).$$
Hence (θ,Θ) is a critical point if and only $\Theta=\pi\circ F(\theta,\psi(\theta))$. The set of the critical points of W is then a 1-dimensional connected submanifold of $\mathbb{R}^2$ that corresponds to the graph of ψ. We deduce that the minimum μ of W is attained exactly on this set.

Let now $(\theta_k,r_k)_{k\in\mathbb{Z}}$ be the orbit of a point $(\theta,\psi(\theta))$ that is on the invariant graph of ψ. Assume that $(\alpha_n)_{\ell\le n\le k}$ is a sequence of real numbers so that $\alpha_\ell=\theta_\ell$ and $\alpha_k=\theta_k$. Then
$$\begin{aligned}(k-\ell+1)\mu &= \sum_{n=\ell+1}^{k}W(\theta_{n-1},\theta_n)\\ &= \sum_{n=\ell+1}^{k}\left(S(\theta_{n-1},\theta_n)+c(\theta_n-\theta_{n-1})+\eta(\theta_{n-1})-\eta(\theta_n)\right)\end{aligned}$$
is less or equal than
$$\sum_{n=\ell+1}^{k}W(\alpha_{n-1},\alpha_n)=\sum_{n=\ell+1}^{k}\left(S(\alpha_{n-1},\alpha_n)+c(\alpha_n-\alpha_{n-1})+\eta(\alpha_{n-1})-\eta(\alpha_n)\right);$$
i.e.
$$\left(\sum_{n=\ell+1}^{k}S(\theta_{n-1},\theta_n)\right)+c(\theta_k-\theta_\ell)+\eta(\theta_\ell)-\eta(\theta_k)\le$$
$$\le\left(\sum_{n=\ell+1}^{k}S(\alpha_{n-1},\alpha_n)\right)+c(\alpha_k-\alpha_\ell)+\eta(\alpha_\ell)-\eta(\alpha_k).$$
As $\alpha_\ell=\theta_\ell$ and $\theta_k=\alpha_k$, we obtain $\displaystyle\sum_{n=\ell+1}^{k}S(\theta_{n-1},\theta_n)\le\sum_{n=\ell+1}^{k}S(\alpha_{n-1},\alpha_n)$ i.e. the orbit of $(\theta,\psi(\theta))$ is minimizing. ∎

5.3. **Proof of the equivalence of different definitions of** Green(f)**.** The result that we will prove is

Theorem. *Let $f:\mathbb{A}\to\mathbb{A}$ be a conservative twist map and let $(x_n)_{n\in\mathbb{Z}}$ be the orbit of a point $x=x_0$. The following assertions are equivalent:*

(0) $x\in$ Green(f)*;*

(1) *the projection of every finite segment of the orbit of x is locally minimizing among the segments of points (of $\mathbb{R}$) that have same length and same endpoints;*

(2) *along the orbit of x, we have for every $k \geq 1$, $s_k > s_{-1}$;*
(3) *along the orbit of x, we have for every $k \geq 1$, $s_{-k} < s_1$;*
(4) *there exists a field of half-lines $\delta_+ \subset T\mathbb{A}$ along the orbit of x such that:*
  - *δ_+ is invariant by Df: $Df(\delta_+) = \delta_+ \circ f$;*
  - *$D\pi \circ \delta_+ = \mathbb{R}_+$ (δ_+ is oriented to the right).*

We will use the following notations.

Notations 5.1.
- F being a lift of f, we note:

$$DF^k(y) = \begin{pmatrix} a_k(y) & b_k(y) \\ c_k(y) & d_k(y) \end{pmatrix};$$

- an *infinitesimal orbit* along (x_n) is

$$(\delta\theta_n, \delta r_n) = (Df^n(x)(\delta\theta, \delta r))_{n\in\mathbb{Z}};$$

- a Jacobi field is then the projection $(\delta\theta_n)_{n\in\mathbb{N}}$ of an infinitesimal orbit;
- if $x_k = (\theta_k, r_k)$, we use the notation

$$\beta_k = \frac{\partial^2 S}{\partial\theta\partial\Theta}(\theta_k, \theta_{k+1}), \quad \alpha_k = \frac{\partial^2 S}{\partial\theta^2}(\theta_k, \theta_{k+1}) + \frac{\partial^2 S}{\partial\Theta^2}(\theta_{k-1}, \theta_k).$$

Remark 5.2. A Jacobi field with two successive zeroes is the zero field.

Let us begin the proof of the theorem.

(1)$\Longrightarrow$(2) We deduce from the definition of the generating functions that

$$Df(x_k) = \begin{pmatrix} -\frac{1}{\beta_k}\frac{\partial^2 S}{\partial\theta^2}(\theta_k, \theta_{k+1}) & -\frac{1}{\beta_k} \\ \beta_k - \frac{1}{\beta_k}\frac{\partial^2 S}{\partial\theta^2}(\theta_k, \theta_{k+1})\frac{\partial^2 S}{\partial\Theta^2}(\theta_k, \theta_{k+1}) & -\frac{1}{\beta_k}\frac{\partial^2 S}{\partial\Theta^2}(\theta_k, \theta_{k+1}) \end{pmatrix}.$$

Observe too that $(\delta\theta_k)$ is a Jacobi field if and only if for every k, we have

$$(*)\beta_{k-1}\delta\theta_{k-1} + \alpha_k\delta\theta_k + \beta_k\delta\theta_{k+1} = 0.$$

As we assume that the orbit is locally minimizing, every matrix $H_{n,m}$ is positive semi-definite if:

$$H_{n,m} = \begin{pmatrix} \alpha_{n+1} & \beta_{n+1} & 0 & . & . & . & 0 \\ \beta_{n+1} & \alpha_{n+2} & \beta_{n+2} & . & . & . & 0 \\ 0 & \beta_{n+2} & \alpha_{n+3} & . & . & . & 0 \\ . & . & . & . & . & . & 0 \\ 0 & . & . & . & 0 & \alpha_{m-2} & \beta_{m-2} \\ 0 & . & . & . & 0 & \beta_{m-2} & \alpha_{m-1} \end{pmatrix}$$

Lemma 5.3. *Every matrix $H_{n,m}$ is positive defnite.*

Proof Let us assume that $(\delta\theta_k)_{k\in[n+1,m-1]}$ is in the kernel of $H_{n,m}$. Using $(*)$ and the fact that $\beta_k \neq 0$ (that is the twist condition), we extend $(\delta\theta_k)$ in a Jacobi field such that $\delta\theta_n = \delta\theta_m = 0$.
Then, $\delta Q = (0, 0, \delta\theta_{n+1}, \delta\theta_{n+2}, \dots, \delta\theta_{m-2}, \delta\theta_{m-1}, 0, 0)$ is in the isotropic cone of $H_{n-2,m+2}$, and then in its kernel because the matrix is positive semi-definite. Hence we have a Jacobi field with two successive zeroes, it is the zero field.

□

Lemma 5.4. *If $k \geq 1$, we have along the orbit of x: $s_k > s_{-1}$.*

Proof Let $(\Delta_j)_{j\in[n-k+1,n]}$ be the image by the matrix $H_{n-k,n+1}$ of the Jacobi field $(\delta\theta_j)_{j\in[n-k+1,n]}$ that corresponds to an infinitesimal orbit $(\delta x_j)_{j\in[n-k+1,n]}$ of a vector $\delta x_{n-k}\in V(x_{n-k})$. Then we have

- $\Delta_{n-k+1}=0$ because $\delta\theta_{n-k}=0$;
- for every $j\in[n-k+2,n-2]$, we have $\Delta_j=0$ because we have a Jacobi field;
- as $\delta x_n=\begin{pmatrix}\delta\theta_n\\ s_k(x_n)\delta\theta_n\end{pmatrix}$, we have

$$\Delta_n=\beta_{n-1}\delta\theta_{n-1}+\alpha_n\delta\theta_n=-\beta_n\delta\theta_{n+1}=-\beta_n D(\pi\circ F)\begin{pmatrix}\delta\theta_n\\ s_k(x_n)\delta\theta_n\end{pmatrix}$$

and then

$$\Delta_n=-\beta_n(\beta_n^{-1}(\frac{\partial^2 S}{\partial\theta^2}(\theta_n\theta_{n+1})+s_k(x_n)))\delta\theta_n=(s_k(x_n)-s_{-1}(x_n))\delta\theta_n.$$

Finally, we obtain $H_{n-k,n+1}((\delta\theta_j),(\delta\theta_j))=(s_k(x_n)-s_{-1}(x_n))\delta\theta_n^2>0$. □

(2)$\Longrightarrow$(3)

Lemma 5.5. *Assume that we have along the orbit of x and for all $k\geq 1$: $s_k>s_{-1}$. Then we have too along the orbit of x: $s_k>s_{k+1}>s_{-1}$.*

Proof We have

$$Df(x_n)\begin{pmatrix}1\\ s_k(x_n)\end{pmatrix}=\begin{pmatrix}-\beta_n^{-1}(s_k(x_n)-s_{-1}(x_n))\\ \beta_n-\beta_n^{-1}s_1(x_{n+1})(s_k(x_n)-s_{-1}(x_n))\end{pmatrix}$$

hence $s_{k+1}(x_{n+1})=-\beta_n^2(s_k(x_n)-s_{-1}(x_n))^{-1}+s_1(x_{n+1})$
i.e. $(s_{k+1}-s_{-1})(x_{n+1})=(s_1-s_{-1})(x_{n+1})-\beta_n^2(s_k(x_n)-s_{-1}(x_n))^{-1}$
and in particular
$(s_2-s_{-1})(x_{n+1})=(s_1-s_{-1})(x_{n+1})-\beta_n^2(s_1(x_n)-s_{-1}(x_n))^{-1}$
where $-\beta_n^2(s_1(x_n)-s_{-1}(x_n))^{-1}<0$. Hence $s_2<s_1$.
Substracting what happens for s_k from what happens for s_{k+1} we obtain:

$$(s_{k+1}-s_k)(x_{n+1})=\beta_n^2\left((s_{k-1}(x_n)-s_{-1}(x_n))^{-1}-(s_k(x_n)-s_{-1}(x_n))^{-1}\right)$$

and by recurrence the fact that (s_k) is strictly decreasing. □

Lemma 5.6. *If along the orbit of x we have $s_k>s_{k+1}>s_{-1}$ for every k, then we have too for every k: $s_1>s_{-k}$.*

Proof We assume that $k\geq 2$. We work on the projective space of $\mathbb{R}^2$ that is nothing else than a circle. On this circle, the lines G_{-1}, G_{k+1}, G_k, G_{k-1} are ordered in the direct sense. As Df^{1-k} is symplectic, its projective action preserves the orientation on the circle and then G_{-k}, G_2, G_1 and V are oriented in the direct sense. This means that $s_{-k}<s_2<s_1$. □

(3)$\Longrightarrow$(0) Applying results that are analogous to Lemmata 5.5 and 5.6, we deduce that if (3) is satisfied, then we have along the orbit of x for every $k\geq 1$: $s_{-1}<s_{k+1}<s_k$ and $s_{-k}<s_{-(k+1)}<s_1$.

Lemma 5.7. *Assume that we have $s_1>s_{-k}$ for every k along the orbit of x. Then for every $n,k\geq 1$, we have: $s_{-k}<s_n$.*

PROOF We assume that $k, n \geq 2$. As in the proof of Lemma 5.6, we work in the projective space. We know that G_{-1}, G_{n+k}, G_{n+k-1} and G_{k-1} are in the direct sense. Hence their image by Df^{1-k} that are G_{-k}, G_{n+1}, G_n and V are in the direct sense too, and then $s_{-k} < s_n$. □

(0)$\Longrightarrow$(1) We fix a point along the orbit of x (that is denoted by x too) and we go along its orbit until it becomes non strictly minimizing. The matrix $H_{0,n}$ is then positive definite but the matrix $H_{0,n+1}$ is not positive definite:

$$H_{0,n+1} = \begin{pmatrix} \alpha_1 & \beta_1 & 0 & \dots & . & 0 \\ \beta_1 & \alpha_2 & . & \dots & . & 0 \\ . & \dots & . & \dots & . & 0 \\ . & \dots & . & \dots & . & \beta_{n-1} \\ 0 & \dots & \dots & 0 & \beta_{n-1} & \alpha_n \end{pmatrix}.$$

A vector $(\eta_1, \dots, \eta_n)$ is in the orthogonal subspace to $\mathbb{R}^{n-1} \times \{0\}$ for $H_{0,n+1}$ if and only if we have $\alpha_1\eta_1 + \beta_1\eta_2 = 0$ and for every $j \in [2, n-1]$: $\beta_{j-1}\eta_{j-1} + \alpha_j\eta_j + \beta_j\eta_{j+1} = 0$, i.e. if (η_j) is the projection of an orbit of $V(x)$.

Hence if $H_{0,n+1}$ is not positive definite, there exists $\eta_0, \dots, \eta_n$ that is the projection of the orbit of a point of $V(x)\backslash\{0\}$ such that:

$$0 \geq \eta_n(\beta_{n-1}\eta_{n-1} + \alpha_n\eta_n) = -\beta_n\eta_n\eta_{n+1}.$$

Note that $Df(x_n) = \begin{pmatrix} -b_n s_{-1} & b_n \\ * & * \end{pmatrix}$ hence $\eta_{n+1} = D(\pi \circ f)(x_n) \begin{pmatrix} \eta_n \\ s_n(x_n)\eta_n \end{pmatrix} = b_n(s_n(x_n) - s_{-1}(x_n))\eta_n = -\beta_n^{-1}(s_n(x_n) - s_{-1}(x-n))\eta_n$. We obtain finally $(s_n - s_{-1})(x_n)\eta_n^2 \leq 0$. As $x \in \text{Green}(f)$, we know that $\eta_n \neq 0$. We deduce that $s_n \leq s_{-1}$, a contradiction with the fact that $x \in \text{Green}(f)$.
We deduce that all the matrices $H_{n,m}$ are positive definite and then (1).

(4)$\Longrightarrow$(0) Now we work on the set of half-lines. We denote by $V_+ = \mathbb{R}_+ \times \{0\}$ the upper vertical and $V_- = -V_+$, $\delta_- = -\delta_+$. This set is a circle and V_-, δ_+, V_+ and δ_- are in the direct sense.
Because Df preserves the orientation, their images are in the direct sense too, i.e. δ_+, $\mathbb{R}_+(1, s_1)$, δ_- and $\mathbb{R}_+(-1, -s_1)$ are in the direct sense too. This implies that δ_+, $\mathbb{R}_+(1, s_1)$, V_+, δ_-, $\mathbb{R}_+(-1, -s_1)$ and V_- are in the direct sense. Taking the images by Df, we find that δ_+, $\mathbb{R}_+(1, s_2)$, $\mathbb{R}_+(1, s_1)$, δ_-, $\mathbb{R}_+(-1, -s_1)$ and $\mathbb{R}_+(-1, -s_2)$ are in the direct sense and so $\delta_+ < s_2 < s_1$. Iterating the method, we obtain: $\delta_+ < s_{n+1} < s_n$. Replacing f by f^{-1} we obtain too $s_{-n} < s_{-n-1} < \delta_+$.

(0)$\Longrightarrow$(4) The idea is to use $\delta_+ = \mathbb{R}_+(1, s_+)$.

5.4. **Proof of a criterion for uniform hyperbolicity.** We want to prove Proposition 4.14:

Proposition. (M.-C. Arnaud) *Let M be a compact invariant set by a conservative twist map that is contained in* Green(f). *Then M is uniformly hyperbolic if and only if at every point of M, the two Green bundles G_- and G_+ are transverse.*

We have noticed that when M is uniformly hyperbolic, we have $G_- = E^s$ and $G_+ = E^u$ on M. Hence G_- and G_+ are transverse at every point of M.

Now we assume that G_- and G_+ are transverse at every point of M.

Definition 5.8. Let $(F_k)_{k\in\mathbb{Z}}$ be a continuous cocycle on a linear normed bundle $P : E \to K$ above a compact metric space K. We say that the cocycle is *quasi-hyperbolic* if

$$\forall v \in E, v \neq 0 \Rightarrow \sup_{k\in\mathbb{Z}} \|F_k v\| = +\infty.$$

A consequence of the dynamical criterion (Proposition 4.9) is that if $K \subset \text{Green}(f)$ is a compact invariant subset of $\text{Green}(f)$ such that for every $x \in K$, $G_+(x)$ and $G_-(x)$ are transverse, then $(Df^k_{|K})_{k\in\mathbb{Z}}$ is a quasi-hyperbolic cocycle. Hence, we only have to prove the following statement to deduce Proposition 4.14.

Theorem 11. *Let (F_k) be a continuous, symplectic and quasi-hyperbolic cocycle on a linear and symplectic (finite dimensional) bundle $P : E \to K$ above a compact metric space K. Then $(F_k)_{k\in\mathbb{Z}}$ is hyperbolic.*

We will deduce Theorem 11 from two lemmata that we will now state and prove. The ideas of the two lemmata and their proofs are similar to the ideas contained in [25] in the setting of the so-called "quasi-Anosov diffeomorphisms".

Lemma 5.9. *Let $(F_k)_{k\in\mathbb{Z}}$ be a continuous and quasi-hyperbolic cocycle on a linear normed bundle $P : E \to K$ above a compact metric space K. Let us define*

- $E^s = \{v \in E; \sup_{k\geq 0} \|F_k v\| < \infty\}$;
- $E^u = \{v \in E; \sup_{k\leq 0} \|F_k v\| < \infty\}$.

Then $(F_{n|E^s})_{n\geq 0}$ and $(F_{-n|E^u})_{n\geq 0}$ are uniformly contracting.

Lemma 5.10. *Let $(F_k)_{k\in\mathbb{Z}}$ be a continuous and quasi-hyperbolic cocycle on a linear normed bundle $P : E \to K$ above a compact metric space K. We denote by $f_k : K \to K$ the underlying dynamics such that $f_k \circ P = P \circ F_k$. If (x_n) is a sequence of points of K tending to x and (k_n) a sequence of integers tending to $+\infty$ such that $\lim_{n\to\infty} f_{k_n}(x_n) = y \in K$, then $\dim E^u(y) \geq \text{codim} E^s(x)$.*

Let us explain how to deduce Theorem 11 from these lemmata:

Proof of theorem 11: If the dimension of E is $2d$, we only have to prove that: $\forall x \in K, \dim E^u(x) = \dim E^s(x) = d$. Let us prove for example that $\dim E^u(x) = d$. By lemma 5.9, $(F_{n|E^s})_{n\geq 0}$ and $(F_{-n|E^u})_{n\geq 0}$ are uniformly contracting. As the cocycle is symplectic, we deduce that every $E^s(x)$ and $E^u(x)$ is isotropic for the symplectic form and then $\dim E^s(x) \leq d$ and $\dim E^u(x) \leq d$.
Let us now consider $x \in K$. As K is compact, we can find a sequence $(k_n)_{n\in\mathbb{N}}$ of integers tending to $+\infty$ such that the sequence $(f_{k_n}(x))_{n\in\mathbb{N}}$ converges to a point $y \in K$. Then, by Lemma 5.10, we have: $\dim E^u(y) \geq \text{codim} E^s(x)$. But we know that $\dim E^u(y) \leq d$, hence $2d - \dim E^s(x) \leq \dim E^u(y) \leq d$ and $\dim E^s(x) = d$. $\square$

Let us now prove the two lemmata.
Proof of lemma 5.9: We will only prove the result for E^s.
Let us assume that we know that:

$$(*) \quad \forall C > 1, \exists N_C \geq 1, \forall v \in E^s, \forall n \geq N_C, \|F_n v\| \leq \frac{\sup\{\|F_k v\|; k \geq 0\}}{C}.$$

We choose $C > 1$. Then $\sup\{\|F_k v\|; k \geq 0\} = \sup\{\|F_k v\|; k \in |[0, N_C]|\}$. We define: $M = \sup\{\|F_k(x)\|; x \in K, k \in |[0, N_C]|\}$. Then, if $j \in |[0, N_C - 1]|$ and $n \in \mathbb{N}$:

$$\|F_{nN_c+j}v\| \leq \frac{1}{C}\sup\{\|F_{(n-1)N_C+j+k}v\|; k \geq 0\} \leq \frac{1}{C^2}\sup\{\|F_{(n-2)N_C+j+k}v\|; k \geq 0\}$$

$$\cdots \leq \frac{1}{C^n}\sup\{\|F_{j+k}v\|; k\geq 0\} \leq \frac{1}{C^n}\sup\{\|F_k v\|; k\geq 0\} \leq \frac{M}{C^n}\|v\|.$$

This proves exponential contraction.

Let us now prove $(*)$. If $(*)$ is not true, there exists $C>1$, a sequence (k_n) in $\mathbb{N}$ tending to $+\infty$ and $v_n \in E^s$ with $\|v_n\|=1$ such that:

$$\forall n\in\mathbb{N}, \|F_{k_n}v_n\| \geq \frac{\sup\{\|F_k v_n\|; k\geq 0\}}{C}.$$

We define: $w_n = \frac{F_{k_n}(v_n)}{\|F_{k_n}(v_n)\|}$. Taking a subsequence, we can assume that the sequence (w_n) converges to a limit $w\in E$. Then we have:

$$\forall n\in\mathbb{N}, \forall k\in[-k_n,+\infty[, \|F_k w_n\| = \frac{\|F_{k+k_n}(v_n)\|}{\|F_{k_n}v_n\|} \leq \frac{\sup\{\|F_j v_n\|; j\geq 0\}}{\|F_{k_n}v_n\|} \leq C.$$

Hence, $\forall k\in\mathbb{Z}, \|F_k w\|\leq C$. This is impossible because $\|w\|=1$ and the cocycle is quasi-hyperbolic. ∎

Proof of lemma 5.10: With the notation of this lemma, we choose a linear subspace $V\subset E_x$ such that V is transverse to $E^s(x)$. We want to prove that $\dim E^u(y)\geq \dim V$.
We choose $V_n\subset E_{x_n}$ such that $\lim\limits_{n\to\infty} V_n = V$. Extracting a subsequence, we have: $\lim\limits_{n\to\infty} F_{k_n}(V_n) = V'\subset E_y$. Then we will prove that $V'\subset E^u(y)$.

Let us assume that we have proved that there exists $C>0$ such that

$$(*)\quad \forall n, \forall 0\leq k\leq k_n, \|F_{-k|F_{k_n}(V_n)}\|\leq C.$$

Then, $\forall w\in V', \forall k\in\mathbb{Z}_-, \|F_k w\|\leq C\|w\|$ and $w\in E^u(y)$.

Let us now assume that $(*)$ is not true. Replacing (k_n) by a subsequence, we find for all $n\in N$ an integer i_n between 0 and k_n such that $\|F_{-i_n|F_{k_n}(V_n)}\|\geq n$. We choose $w_n\in F_{k_n}(V_n)$ such that $\|w_n\|=1$ and $\|F_{-i_n}(w_n)\| = \|F_{-i_n|F_{k_n}(V_n)}\|$. We may even assume that: $\|F_{-i_n}(w_n)\| = \sup\{\|F_k(w_n)\|; -k_n\leq k\leq 0\}\geq n$.
Then $\lim\limits_{n\to+\infty} i_n = +\infty$. If $v_n = \frac{F_{-i_n}(w_n)}{\|F_{-i_n}(w_n)\|}$, we may extract a subsequence and assume that: $\lim\limits_{n\to\infty} v_n = v$, with $\|v\|=1$.
Then we have $\forall k\in|[0,i_n]|, \|F_k v_n\|\leq\|v_n\|$ for all $k=0,\dots,i_n$, and therefore $\|F_k v\|\leq \|v\|$ for all $k\in\mathbb{N}$ and $v\in E^s$.
Now, we have two cases:

- either (k_n-i_n) doesn't tend to $+\infty$; we may extract a subsequence and assume that $\lim\limits_{n\to+\infty}(k_n-i_n) = N\geq 0$; then:

$$F_{-N}v = \lim_{n\to\infty} F_{i_n-k_n}(v_n) = \lim_{n\to\infty}\frac{F_{-k_n}(w_n)}{\|F_{-i_n}(w_n)\|}$$

  We have:

$$\frac{F_{-k_n}(w_n)}{\|F_{-i_n}(w_n)\|}\in V_n$$

  and then $F_{-N}v\in V$. Moreover, $F_{-N}v\in F_{-N}E^s = E^s$. As $\|v\|=1$ and V is transverse to E^s_x, we obtain a contradiction.

- or $\lim_{n\to\infty}(k_n - i_n) = +\infty$. In this case, for every $k = -k_n + i_n, \dots, i_n$, we have $-k_n \le k - i_n \le 0$ and therefore $\|F_k v_n\| = \frac{\|F_{k-i_n} w_n\|}{\|F_{-i_n} w_n\|} \le 1 = \|v_n\|$. Hence, since $v_n \to v$, $i_n \to +\infty$, and $-k_n + i_n \to -\infty$, when $n \to +\infty$, we obtain $\|F_k v\| \le \|v\| = 1$, for all $k \in \mathbb{Z}$. This implies $v \in E^s \cap E^u$. This contradicts $\|v\| = 1$ and the fact that the cocycle is quasi-hyperbolic.

$\square$

5.5. **Proof of Proposition 4.20.** We will prove

Proposition. (M.-C. Arnaud) *Let $h : \mathbb{T} \to \mathbb{T}$ be a bi-Lipschitz orientation preserving homeomorphism with irrational rotation number. We denote by μ its unique invariant measure and assume that h is C^1-regular μ-almost everywhere. Then uniformly in $\theta \in \mathbb{T}$, we have*

$$\lim_{n\to+\infty} \frac{1}{n} \log (h^n)'_+ = \lim_{n\to+\infty} \frac{1}{n} \log (h^n)'_- = 0.$$

Proof A fundamental argument of the proof is a result proved by A. Furman in [14] that is an improvement of Kingman subadditive theorem in the case of a unique ergodic measure.

Theorem 12. (A. Furman) *Let (X, μ) be a Borel probability space, T be a continuous measure preserving transformation of (X, μ) such that μ is uniquely ergodic for T and let $(f_n) \in L^1(X, \mu)$ be a T-sub-additive sequence of upper semi-continuous functions. Let $\Lambda((f_k)) = \lim_{n\to\infty} \frac{1}{n} \int f_n d\mu$ be the constant associated to f via the sub-additive ergodic theorem. Then:*

$$\forall \varepsilon > 0, \exists N \ge 0, \forall n \ge N, \forall x \in X, \frac{1}{n} f_n(x) \le \Lambda((f_k)) + \varepsilon.$$

We apply Theorem 12 for $(X, \mu) = (\mathbb{T}, \mu)$, $T = h$ (resp. $T = h^{-1}$) and $f_n = -\log\left((h^n)'_-\right)$ (resp. $f_n = -\log\left((h^{-n})'_-\right)$). Fixing $\varepsilon > 0$, we find $N \ge 0$ such that for every $n \ge N$ and every $\theta \in \mathbb{T}$, we have

$$-\frac{1}{n} \log\left((h^n)'_-(\theta)\right) \le \Lambda((f_k)) + \varepsilon.$$

We denote by $d\theta$ the Lebesgue measure on $\mathbb{T}$. Because of Jensen inequality for the convex function $-\log$, we have

$$-\log\left(\int \left((h^n)'_-\right) d\theta\right) \le -\int \log\left((h^n)'_-\right) d\theta.$$

Moreover, if H is a lift of h,

$$\int \left((h^n)'_-\right) d\theta \le \int (h^n)' \, d\theta = [H^n]_0^1 = 1.$$

We deduce

$$\Lambda((f_k)) + \varepsilon \ge -\frac{1}{n} \int \log (h^n)'_- \, d\theta \ge -\log 1 = 0$$

and then $\Lambda((f_k)) \ge 0$.

Finally, we obtain in particular:

$$\Lambda\left(-\log((h^n)'_-)\right) \ge 0$$

and

$$\Lambda\left(-\log((h^{-n})'_-)\right) \geq 0\,.$$

Observe that $(h^{-n})'_-(\theta) = \frac{1}{(h^n)'_+(h^{-n}\theta)}$ hence

$$\int \log\left(h^{-n}\right)'_- d\mu = -\int \log\left(h^n\right)'_+ d\mu.$$

Because h is C^1-regular μ-almost everywhere we have μ-almost everywhere

$$\prod_{j=0}^{n-1} h'_+(h^j\theta) = \prod_{j=0}^{n-1} h'_-(h^j\theta)\,.$$

Because $(h^n)'_-$ and $(h^n)'_+$ are between these two numbers, we deduce that we have μ-almost everywhere $(h^n)'_-(\theta) = (h^n)'_+(\theta)$ and then

$$\frac{1}{n}\int \log\left(h^{-n}\right)'_- d\mu = -\frac{1}{n}\int \log\left(h^n\right)'_- d\mu$$

and

$$\Lambda\left(-\log((h^n)'_-)\right) = -\Lambda\left(-\log((h^{-n})'_-)\right) = 0.$$

We deduce then from Theorem 12 that for every $\varepsilon > 0$, there exists $N \geq 0$ such that for every $n \geq N$ and every $\theta \in \mathbb{T}$, we have

$$-\frac{1}{n}\log\left((h^n)'_-(\theta)\right) \leq \varepsilon \quad \text{and} \quad \frac{1}{n}\log\left((h^n)'_+\theta)\right) = -\frac{1}{n}\log\left((h^n)'_-(h^n\theta)\right) \leq \varepsilon$$

then

$$-\varepsilon \leq \frac{1}{n}\log\left((h^n)'_-(\theta)\right) \leq \frac{1}{n}\log\left((h^n)'_+\theta)\right) \leq \varepsilon.$$

□

References

[1] M.-C. Arnaud, *Three results on the regularity of the curves that are invariant by an exact symplectic twist map*, Publ. Math. Inst. Hautes Etudes Sci. **109**, 1-17(2009)

[2] M.-C. Arnaud, *A non-differentiable essential irrational invariant curve for a C^1 symplectic twist map*, Journal of Modern Dynamics, **5**, No. 3, 583 – 591 (2011).

[3] M.-C. Arnaud, *The link between the shape of the Aubry-Mather sets and their Lyapunov exponents*, Annals of Mathematics, **174-3** , p 1571-1601 (2011).

[4] M.-C. Arnaud, *Boundaries of instability zones for symplectic twist maps*, Journal of the Institute of Mathematics of Jussieu, **13-1** , 19-41 (2013).

[5] M.-C. Arnaud, *Lower and upper bounds for the Lyapunov exponents of twisting dynamics: a relationship between the exponents and the angle of the Oseledet's splitting*, Ergodic Theory and Dynamical Systems, volume **33**, issue 03, (2013), pp. 693-712

[6] M. -C. Arnaud & P. Berger. *The non-hyperbolicity of irrational invariant curves for twist maps and all that follows*, preprint november 2014, preprint HAL 01087349

[7] M.-C. Arnaud, C. Bonatti, & S. Crovisier, *Dynamiques symplectiques génériques,* Ergodic Theory Dynam. Systems **25** , no. 5, 1401-1436 (2005).

[8] V Arnol'd. *Proof of a theorem of A. N. Kolmogorov on the preservation of conditionally periodic motions under a small perturbation of the Hamiltonian,* (Russian) Uspehi Mat. Nauk **18** no. 5 (113), 13–40 (1963).

[9] S. Aubry & P. Y. Le Daeron. *The discrete Frenkel-Kontorova model and its extensions. I. Exact results for the ground-states*, Phys. D 8 (1983), no. 3, 381–422.

[10] V. Bangert, *Mather sets for twist maps and geodesics on tori*, Dynamics reported, Vol. 1, 1–56, Dynam. Report. Ser. Dynam. Systems Appl., 1, Wiley, Chichester, 1988.

[11] G. D. Birkhoff, *Surface transformations and their dynamical application*, Acta Math. **43** (1920) 1-119.

[12] G. D. Birkhoff, *Sur l'existence de régions d'instabilité en Dynamique*, Ann. Inst. H. Poincaré **2** , no. 4, 369-386 (1932).

[13] A. Chenciner, *La dynamique au voisinage d'un point fixe elliptique conservatif: de Poincaré et Birkhoff à Aubry et Mather,* (French) [The dynamics at the neighborhood of a conservative elliptic fixed point: from Poincaré and Birkhoff to Aubry and Mather] Seminar Bourbaki, Vol. 1983/84. *Astérisque* No. **121-122** , 147–170 (1985)

[14] A. Furman *On the multiplicative ergodic theorem for uniquely ergodic systems,* Ann. Inst. H. Poincaré Probab. Statist. **33** , no. 6, 797–815 (1997).

[15] C. Golé, Symplectic twist maps, Global variational techniques. Advanced Series in Nonlinear Dynamics, 18. World Scientific Publishing Co., Inc., River Edge, NJ, 2001. xviii+305 pp

[16] D. L. Goroff, *Hyperbolic sets for twist maps*, Ergodic Theory Dynam. Systems **5** (1985), no. 3, 337–339.

[17] S. Hayashi, *Connecting invariant manifolds and the solution of the C^1 stability and Ω-stability conjectures for flows*, Ann. of Math. (2) **145** , no. 1, 81-137 (1997).

[18] M. Herman, *Sur la conjugaison différentiable des difféomorphismes du cercle à des rotations*, Inst. Hautes Études Sci. Publ. Math. No. **49** (1979), 5-233

[19] M. Herman, *Sur les courbes invariantes par les difféomorphismes de l'anneau*, Vol. 1, Asterisque **103-104** (1983).

[20] A. Katok, *Some remarks of Birkhoff and Mather twist map theorems*, Ergodic Theory Dynamical Systems **2** (1982), no. 2, 185194 .

[21] A. Katok & B. Hasselblatt, *Introduction to the modern theory of dynamical systems. With a supplementary chapter by Katok and Leonardo Mendoza,* Encyclopedia of Mathematics and its Applications, 54. Cambridge University Press, Cambridge, 1995. xviii+802 pp

[22] A. N. Kolmogorov. *On conservation of conditionally periodic motions for a small change in Hamilton's function,* (Russian) Dokl. Akad. Nauk SSSR (N.S.) 98, 527–530 (1954).

[23] P. Le Calvez, *Propriétés dynamiques des régions d'instabilité,* (French) [Dynamical properties of regions of instability] Ann. Sci. École Norm. Sup. (4) 20 , no. 3, 443-464 (1987).

[24] P. Le Calvez, *Les ensembles d'Aubry-Mather d'un difféomorphisme conservatif de l'anneau déviant la verticale sont en général hyperboliques*, (French) [The Aubry-Mather sets of a conservative diffeomorphism of the annulus twisting the vertical are hyperbolic in general] C. R. Acad. Sci. Paris Sér. I Math. 306 (1988), no. 1, 51–54.

[25] R. Mañé. *Quasi-Anosov diffeomorphisms and hyperbolic manifolds*, Trans. Amer. Math. Soc. **229** , 351-370 (1977).

[26] J. N. Mather. *Existence of quasiperiodic orbits for twist homeomorphisms of the annulus*, Topology 21 (1982), no. 4, 457–467.

[27] J. N. Mather, *Variational construction of orbits of twist diffeomorphisms,* J. Amer. Math. Soc. **4** no. 2, 207-263 (1991).

[28] H. Rüssman. *On the existence of invariant curves of twist mappings of an annulus,* Geometric dynamics (Rio de Janeiro, 1981), 677–718, Lecture Notes in Math., **1007**, Springer, Berlin, (1983).

[29] K.-F. Siburg, *The principle of least action in geometry and dynamics.* Lecture Notes in Mathematics, **1844** Springer-Verlag, Berlin, 2004. xii+128 pp

Laboratoire de Mathématiques d'Avignon (EA 2151)
Université d'Avignon et des Pays de Vaucluse
F-84 018 Avignon, France
E-mail address: Marie-Claude.Arnaud@univ-avignon.fr

Publicaciones Matemáticas del Uruguay
Volumen 16, Julio 2016, Páginas 41–79
S 0797-1443

ON CLOSED ORBITS FOR TWISTED AUTONOMOUS TONELLI LAGRANGIAN FLOWS

GABRIELE BENEDETTI

Abstract. These lecture notes were prepared in occasion of a mini-course given by the author at the "CIMPA Research School - Hamiltonian and Lagrangian Dynamics" (10–19 March 2015 - Salto, Uruguay). In this series of talk we illustrated some techniques to prove the existence of periodic orbits of prescribed energy for autonomous Tonelli Lagrangian systems on the twisted cotangent bundle of a closed manifold.

Contents

1. Introduction

The study of invariant sets plays a crucial role in the understanding of the properties of a dynamical system: it can be used to obtain information on the dynamics both at a local scale, such as the existence of nearby stable motions, and at a global one, such as the presence of chaos (see [Mos73]). In the realm of continuous flows periodic orbits are the simplest example of invariant sets and, therefore, they usually represent the first object of study. For systems admitting a Lagrangian formulation closed orbits received special consideration in the past years, in particular for the cases having geometrical or physical significance, such as geodesic flows [Kli78] or mechanical flows in phase space [Koz85]. In [Con06] Contreras formulated a very general theorem about the existence of periodic motions for autonomous Lagrangian systems over compact configuration spaces. This result was later analysed in detail by Abbondandolo, who discussed it in a series of lecture notes [Abb13]. It is the purpose of the present paper to give a generalization of such theorem to systems which admit only a *local* Lagrangian description (Theorem 1.6 below). Among these we find the important example of magnetic flows on surfaces, which we introduce in Section 1.6. We look at them in detail in the last part of this note: we will sketch a different method, devised by Taĭmanov in [Taĭ93], to find periodic orbits with low energy and we will study the *stability* of the energy levels, a purely symplectic property, which has important consequences for the existence of periodic orbits.

Let us start now our study by making precise the general setting in which we work.

1.1. Twisted Lagrangian flows over closed manifolds.

Let M be a closed connected n-dimensional manifold and denote by

$$\begin{array}{rcl} \pi : TM & \longrightarrow & M \\ (q,v) & \longmapsto & q \end{array} \qquad\qquad \begin{array}{rcl} \pi : T^*M & \longrightarrow & M \\ (q,p) & \longmapsto & q \end{array}$$

the tangent and the cotangent bundle projection of M. Let us fix also an auxiliary Riemannian metric g on M and let $|\cdot|$ denote the associated norm.

Let $\sigma \in \Omega^2(M)$ be a closed 2-form on M which we refer to as the *magnetic form*. We call *twisted cotangent bundle* the symplectic manifold (T^*M, ω_σ), where $\omega_\sigma := d\lambda - \pi^*\sigma$. Here λ is the canonical 1-form defined by

$$\lambda_{(q,p)} \;=\; p \circ d_{(q,p)}\pi\,, \qquad \forall\,(q,p) \in T^*M\,.$$

If $K : T^*M \to \mathbb{R}$ is a smooth function, we denote by $t \mapsto \Phi_t^{(K,\sigma)}$ the Hamiltonian flow of K. It is generated by the vector field $X_{(K,\sigma)}$ defined by

$$\omega_\sigma(X_{(K,\sigma)}, \cdot\,) \;=\; -dK\,.$$

In local coordinates on T^*M such flow is obtained by integrating the equations

$$\begin{cases} \dot q = \dfrac{\partial K}{\partial p}\,, \\[2mm] \dot p = -\dfrac{\partial K}{\partial q} \;+\; \sigma\left(\dfrac{\partial K}{\partial p}, \cdot\,\right). \end{cases} \tag{1}$$

The function K is an integral of motion for $\Phi^{(K,\sigma)}$. Moreover, if k is a regular value for K, then the flow lines lying on $\{K = k\}$ are tangent to the 1-dimensional distribution $\ker \omega_\sigma|_{\{K=k\}}$. This means that if $K' : T^*M \to \mathbb{R}$ is another Hamiltonian with a regular value k' such that $\{K' = k'\} = \{K = k\}$, then $\Phi^{(K',\sigma)}$ and $\Phi^{(K,\sigma)}$ are the same up to a *time reparametrization* on the common hypersurface. In other words, there exists a smooth family of diffeomorphisms $\tau_z : \mathbb{R} \to \mathbb{R}$ parametrized by $z \in \{K' = k'\} = \{K = k\}$ such that

$$\tau_z(0) \;=\; 0 \quad\text{ and }\quad \Phi_t^{(K,\sigma)}(z) \;=\; \Phi^{(K',\sigma)}_{\tau_z(t)}(z)\,.$$

Hence, there is a bijection between the closed orbits of the two flows on the hypersurface.

Let $L : TM \to \mathbb{R}$ be a *Tonelli Lagrangian.* This means that for every $q \in M$, the restriction $L|_{T_qM}$ is superlinear and strictly convex (see [Abb13]):

$$\begin{aligned} \lim_{|v|\to+\infty} \frac{L(q,v)}{|v|} \;&=\; +\infty\,, \quad \forall\, q \in M\,, \\ \frac{\partial^2 L}{\partial v^2}(q,v) \;&>\; 0\,, \quad \forall\,(q,v) \in TM\,, \end{aligned} \tag{2}$$

where $\frac{\partial^2 L}{\partial v^2}(q,v)$ is the Hessian of $L|_{T_qM}$ at $v \in T_qM$. The *Legendre transform* associated to L is the fibrewise diffeomorphism

$$\begin{aligned} \mathcal{L} : TM \;&\longrightarrow\; T^*M \\ (q,v) \;&\longmapsto\; \frac{\partial L}{\partial v}(q,v)\,. \end{aligned}$$

The *Legendre dual* of L is the *Tonelli Hamiltonian*

$$\begin{aligned} H : T^*M \;&\longrightarrow\; \mathbb{R} \\ (q,p) \;&\longmapsto\; p\Big(\mathcal{L}^{-1}(q,p)\Big) - L\big(\mathcal{L}^{-1}(q,p)\big)\,, \end{aligned}$$

which satisfies the analogue of (2) on T^*M. For every $k \in \mathbb{R}$, let $\Sigma_k^* := \{H = k\}$. These sets are compact and invariant for $\Phi^{(H,\sigma)}$. As a consequence such a flow is complete. We can use $\mathcal{L}$ to pull back to TM the Hamiltonian flow of H.

Definition 1.1. *Let $\Phi^{(L,\sigma)}$ be the flow on TM defined by conjugation*

$$\mathcal{L} \circ \Phi^{(L,\sigma)} \;=\; \Phi^{(H,\sigma)} \circ \mathcal{L}\,.$$

We call $\Phi^{(L,\sigma)}$ a twisted Lagrangian flow *and we write $X_{(L,\sigma)}$ for its generating vector field. Since $\Phi^{(H,\sigma)}$ is complete, $\Phi^{(L,\sigma)}$ is complete as well.*

The next proposition shows that the flow $\Phi^{(L,\sigma)}$ is locally a standard Lagrangian flow.

Proposition 1.2. *Let $U \subset M$ be an open set such that $\sigma|_U = d\theta$ for some $\theta \in \Omega^1(U)$. There holds*

$$X_{(L-\theta,0)} \;=\; X_{(L,\sigma)}|_U\,,$$

where $L-\theta : TU \to \mathbb{R}$ is the Tonelli Lagrangian defined by $(L-\theta)(q,v) = L(q,v) - \theta_q(v)$ and $X_{(L-\theta,0)}$ is the standard Lagrange vector field of $L-\theta$.

The proof of this result follows from the next exercise.

Exercise 1. *Prove the following generalization of the Euler-Lagrange equations. Consider a smooth curve $\gamma : [0,T] \to M$. Then, the curve $(\gamma,\dot\gamma)$ is a flow line of $X_{(L,\sigma)}$ if and only if for every open set $W \subset M$ and every linear symmetric connection ∇ on W,*

$$\left(\nabla_{\dot\gamma}\frac{\partial L}{\partial v}\right)(\gamma,\dot\gamma) \;=\; \frac{\partial L}{\partial q}(\gamma,\dot\gamma) \;+\; \sigma_\gamma(\dot\gamma,\cdot) \tag{3}$$

at every time $t \in [0,T]$ such that $\gamma(t) \in W$. In the above formula $\frac{\partial L}{\partial q} \in T^*M$ denotes the restriction of the differential of L to the horizontal distribution given by ∇.

1.2. **The magnetic form.** Let $[\sigma] \in H^2(M;\mathbb{R})$ denote the cohomology class of σ. We observe that for any $\theta \in \Omega^1(M)$, there holds

$$X_{(L+\theta,\sigma+d\theta)} = X_{(L,\sigma)}\,.$$

Since $L+\theta$ is still a Tonelli Lagrangian, we expect that general properties of the dynamics depend on σ only via $[\sigma]$. Moreover, if $\theta \in \Omega^1(M)$ is defined by $\theta_q := -\frac{\partial L}{\partial v}(q,0)$, then

$$\min_{v \in T_qM}\Big(L(q,v) \;+\; \theta_q(v)\Big) \;=\; L(q,0) \;+\; \theta_q(0)\,, \quad \forall\, q \in M\,.$$

Therefore, without loss of generality we assume from now on that $L|_{T_qM}$ attains its minimum at $(q,0)$, for every $q \in M$.

We can refine the classification of σ given by $[\sigma]$ by looking at the cohomological properties of its lift to the universal cover. Let $\widetilde\sigma$ be the pull-back of σ to the universal cover $\widetilde M \to M$. We say that σ is *weakly exact* if $[\widetilde\sigma] = 0$. This is equivalent to asking that

$$\int_{S^2} u^*\sigma \;=\; 0\,, \quad \forall\, u : S^2 \longrightarrow M\,.$$

We say that σ admits a *bounded weak primitive* if there is $\widetilde\theta \in \Omega^1(\widetilde M)$ such that $d\widetilde\theta = \widetilde\sigma$ and

$$\sup_{\widetilde q \in \widetilde M} |\widetilde\theta_{\widetilde q}| \;<\; +\infty\,.$$

In this case we write $[\widetilde\sigma]_b = 0$. Notice that both notions that we just introduced depend on σ only via $[\sigma]$.

Exercise 2. *If M is a surface and $[\sigma] \neq 0$, show that*

- *if $M = S^2$, then $[\widetilde\sigma] \neq 0$;*

- *if* $M = \mathbb{T}^2$, *then* $[\widetilde{\sigma}] = 0$, *but* $[\widetilde{\sigma}]_b \neq 0$;
- *if* $M \notin \{S^2, \mathbb{T}^2\}$, *then* $[\widetilde{\sigma}]_b = 0$.

Using the second point, prove that

- *if* $M = \mathbb{T}^n$ *and* $[\sigma] \neq 0$, *then* $[\widetilde{\sigma}] = 0$, *but* $[\widetilde{\sigma}]_b \neq 0$;
- *if* M *is any manifold and* $[\widetilde{\sigma}]_b = 0$, *then*

$$\int_{\mathbb{T}^2} u^*\sigma \;=\; 0\,, \quad \forall\, u : \mathbb{T}^2 \longrightarrow M\,.$$

1.3. **Energy.** As twisted Lagrangian flows are described by an autonomous Hamiltonian on the twisted cotangent bundle, they possess a natural first integral. It is the Tonelli function $E : TM \to \mathbb{R}$ given by $E := H \circ \mathcal{L}$. We call it the *energy* of the system and we write $\Sigma_k := \{E = k\}$, for every $k \in \mathbb{R}$. Let $V : M \to \mathbb{R}$ denote the restriction of E to the zero section and let $e_m(L)$ and $e_0(L)$ denote the minimum and maximum of V, respectively.

Proposition 1.3. *The energy can be written as*

$$E(q,v) \;=\; \frac{\partial L}{\partial v}(q,v)(v) \;-\; L(q,v)$$

and, for every $q \in M$, *we have*

$$\min_{v \in T_qM} E(q,v) \;=\; E(q,0) \;=\; V(q) \;=\; -L(q,0)\,.$$

Moreover,

- $k > e_0(L)$ *if and only if* $\pi : \Sigma_k \to M$ *is an* S^{n-1}*-bundle (isomorphic to the unit tangent bundle of* M*).*
- $k < e_m(L)$ *if and only if* $\Sigma_k = \emptyset$.

Exercise 3. *If* $q_0 \in M$ *is a critical point of* V, *then* $(q_0, 0)$ *is a constant periodic orbit of* $\Phi^{(L,\sigma)}$ *with energy* $V(q_0)$.

1.4. **The Mañé critical value of the universal cover.** When σ is weakly exact we define the *Mañé critical value* of the universal cover as

$$c(L,\sigma) \;:=\; \inf_{d\widetilde{\theta} = \widetilde{\sigma}} \left(\sup_{\widetilde{q} \in \widetilde{M}} \widetilde{H}(\widetilde{q}, \widetilde{\theta}_{\widetilde{q}}) \right) \;\in\; \mathbb{R} \cup \{+\infty\}\,, \tag{4}$$

where $\widetilde{H} : T^*\widetilde{M} \to \mathbb{R}$ is the lift of H to $\widetilde{M}$. This number plays an important role, since as it will be apparent from Theorem 1.6 and the examples in Section 1.6 the dynamics on Σ_k changes dramatically when k crosses $c(L,\sigma)$.

Proposition 1.4. *If* σ *is weakly exact, then*

- $c(L,\sigma) < +\infty$ *if and only if* $[\widetilde{\sigma}]_b = 0$;
- $c(L,\sigma) \geq e_0(L)$;
- *if* $\sigma = d\theta_0$, *where* $\theta_0(\cdot) = \mathcal{L}(\cdot, 0)$, *then* $c(L,\sigma) = e_0(L)$ *and the converse is true, provided* $e_0(L) = e_m(L)$;
- *given two Tonelli Lagrangians* L_1 *and* L_2 *and two real numbers* k_1 *and* k_2 *such that* $\{H_1 = k_1\} = \{H_2 = k_2\}$, *then*

$$c(L_1,\sigma) \geq k_1 \iff c(L_2,\sigma) \geq k_2 \quad \text{and} \quad c(L_1,\sigma) \leq k_1 \iff c(L_2,\sigma) \leq k_2\,.$$

1.5. **Example I: electromagnetic Lagrangians.** Let g be a Riemannian metric on M and $V : M \to \mathbb{R}$ be a function. Suppose that the Lagrangian is of *mechanical type*, namely it has the form

$$L(q,v) \;=\; \frac{1}{2}|v|^2 \;-\; V(q)\,,$$

where $|\cdot|$ is the norm associated to g. In this case we refer to $\Phi^{(L,\sigma)}$ as a *magnetic flow* since we have the following physical interpretation of this system: it models the motion of a charged particle γ moving in M under the influence of a potential V and a stationary magnetic field σ. Using Exercise 1, the equation of motion reads

$$\nabla_{\dot\gamma}\dot\gamma \;=\; -\nabla V(\gamma) \;+\; Y_\gamma(\dot\gamma)\,, \tag{5}$$

where ∇V is the gradient of V and, for every $q \in M$, $Y_q : T_qM \to T_qM$ is defined by

$$g_q(Y_q(v_1), v_2) \;=\; \sigma_q(v_1, v_2)\,, \quad \forall\, v_1, v_2 \in T_qM\,.$$

Exercise 4. *Prove that, if $k > \max V$, $\Phi^{(L,\sigma)}|_{\Sigma_k}$ can be described in terms of a purely kinetic system. Namely, define the Jacobi metric $g_k := \frac{k-V}{k}g$ and the Lagrangian $L_k(q,v) := \frac{1}{2}|v|_k^2$, where $|\cdot|_k$ is the norm induced by g_k. Using the Hamiltonian formulation, show that $\Phi^{(L,\sigma)}|_{\{E=k\}}$ is conjugated (up to time reparametrization) to $\Phi^{(L_k,\sigma)}|_{\{E_k=k\}}$, where E_k is the energy function of L_k.*

In the particular case $M = S^2$, magnetic flows describe yet another interesting mechanical system. Consider a rigid body in $\mathbb{R}^3$ with a fixed point and moving under the influence of a potential V. Suppose that V is invariant under rotations around the axis $\hat z$. We identify the rigid body as an element $\psi \in SO(3)$. Since $SO(3)$ is a Lie group, we use left multiplications to get $TSO(3) \simeq SO(3) \times \mathbb{R}^3 \ni (\psi, \Omega)$, where Ω is the angular speed of the body. Thus, we have a Lagrangian system on $SO(3)$ with $L = \frac{1}{2}|\Omega|^2 - V(\psi)$ and $\sigma = 0$. Here $|\cdot|$ denote the metric induced by the tensor of inertia of the body.

The quotient of $SO(3)$ by the action of the group of rotations around $\hat z$ is a two-sphere. The quotient map $q : SO(3) \to S^2$ sends ψ to the unit vector in $\mathbb{R}^3$, whose entries are the coordinates of $\hat z$ in the basis determined by ψ.

By the rotational symmetry, the quantity $\Omega \cdot \hat z$ is an integral of motion. Hence, for every $\omega \in \mathbb{R}$, the set $\{\Omega \cdot \hat z = \omega\} \subset TSO(3)$ is invariant under the flow and we have the commutative diagram

$$\begin{array}{ccc}
\big(\{\Omega\cdot\hat z = \omega\}, X_{(L,0)}\big) & \xrightarrow{\;dq\;} & \big(TS^2, X_{(L_\omega,\sigma_\omega)}\big) \\
\Big\downarrow{\scriptstyle \pi} & & \Big\downarrow{\scriptstyle \pi} \\
SO(3) & \xrightarrow{\quad q \quad} & S^2\,.
\end{array}$$

The resulting twisted Lagrangian system $(L_\omega, \sigma_\omega)$ on S^2 can be described as follows:

- $L_\omega(q,v) \;=\; \frac{1}{2}|v|^2 - V_\omega(q)$, where $|\cdot|$ is the norm associated to a *convex* metric g on S^2 (independent of ω) and V_ω is a potential (depending on ω);
- $\sigma_\omega = \omega \cdot \kappa$, where κ is the curvature form of g (in particular σ_ω has integral $4\pi\omega$ and, if $\omega \neq 0$, it is a symplectic form on S^2).

The rigid body model presented in this subsection is described in detail in [Kha79]. We refer the reader to [Nov82], for other relevant problems in classical mechanics that can be described in terms of twisted Lagrangian systems.

1.6. **Example II: magnetic flows on surfaces.** We now specialize further the example of electromagnetic Lagrangians that we discussed in the previous subsection and we consider purely kinetic systems on a closed oriented Riemannian surface (M, g). In this case

$$L(q, v) \;:=\; \frac{1}{2}|v|^2\,, \tag{6}$$

and $\sigma = f \cdot \mu$, where μ is the metric area form and $f : M \to \mathbb{R}$. The magnetic endomorphism can be written as $Y = f \cdot \imath$, where $\imath : TM \to TM$ is the fibrewise rotation by $\pi/2$.

Remark 1.5. *If the surface is isometrically embedded in the Euclidean space $\mathbb{R}^3$, Y is the classical Lorentz force. Namely, we have $Y_q(v) = v \times B(q)$, where $\times$ is the outer product of vectors in $\mathbb{R}^3$ and B is the vector field $B : M \to \mathbb{R}^3$ perpendicular to M and determined by the equation* $\mathrm{vol}_{\mathbb{R}^3}(B, \cdot, \cdot) = \sigma$*, where* $\mathrm{vol}_{\mathbb{R}^3}$ *is the Euclidean volume.*

For purely kinetic systems $E = L$ and, therefore, the solutions of the twisted Euler-Lagrange equations are parametrized by a multiple of the arc length. More precisely, if $(\gamma, \dot\gamma) \subset \Sigma_k$, then $|\dot\gamma| = \sqrt{2k}$. In particular, the solutions with $k = 0$ are exactly the constant curves. To characterise the solutions with $k > 0$ we write down explicitly the twisted Euler-Lagrange equation (5):

$$\nabla_{\dot\gamma}\dot\gamma = f(\gamma) \cdot \imath\dot\gamma\,. \tag{7}$$

We see that γ satisfies (7) if and only if $|\dot\gamma| = \sqrt{2k}$ and

$$\kappa_\gamma = s \cdot f(\gamma)\,, \qquad s \;:=\; \frac{1}{\sqrt{2k}}\,, \tag{8}$$

where κ_γ is the geodesic curvature of γ. The advantage of working with Equation (8) is that it is invariant under orientation-preserving reparametrizations.

Let us do some explicit computations when the data are homogeneous. Thus, let g be a metric of constant curvature on M and let $\sigma = \mu$. When $M \neq \mathbb{T}^2$ we assume, furthermore, that the absolute value of the Gaussian curvature is 1. By (8), in order to find the trajectories of $\Phi^{(L,\sigma)}$ we need to solve the equation $\kappa_\gamma = s$ for all $s > 0$.

Denote by $\widetilde{M}$ the universal cover of M. Then, $\widetilde{S^2} = S^2$, $\widetilde{\mathbb{T}^2} = \mathbb{R}^2$ and, if M has genus larger than one, $\widetilde{M} = \mathbb{H}$, where $\mathbb{H}$ is the hyperbolic plane. Our strategy will be to study the trajectories of the lifted flow and then project them down to M. Working on the universal cover is easier since there the problem has a bigger symmetry group. Notice, indeed, that the lifted flow is invariant under the group of orientation preserving isometries $\mathrm{Iso}_+(\widetilde{M})$.

1.6.1. *The two-sphere.* Let us fix geodesic polar coordinates $(r, \varphi) \in (0, \pi) \times \mathbb{R}/2\pi\mathbb{Z}$ around a point $q \in S^2$ corresponding to $r = 0$. The metric takes the form $dr^2 + (\sin r)^2 d\varphi^2$. Let $C_r(q)$ be the boundary of the geodesic ball of radius r oriented in the counter-clockwise sense. We compute $\kappa_{C_r(q)} = \frac{1}{\tan r}$. Observe that $\tan r$ takes every positive value exactly once for $r \in (0, \pi/2)$. Therefore, if $s > 0$, the trajectories of the flow are all supported on $C_{r(s)}(q)$, where q varies in S^2 and

$$r(s) \;=\; \arctan\frac{1}{s} \;\in\; (0, \pi/2)\,. \tag{9}$$

In particular, all orbits are closed and their period is

$$T(s) \;=\; \frac{2\pi s}{\sqrt{s^2+1}}\,.$$

1.6.2. *The two-torus.* In this case we readily see that the trajectories of the lifted flow are circles of radius $r(s) = 1/s$. In particular, all the orbits are closed and contractible. Their period is $T(s) = 2\pi$, hence it is independent of s (or k).

1.6.3. *The hyperbolic surface.* We fix geodesic polar coordinates $(r, \varphi) \in (0, +\infty) \times \mathbb{R}/2\pi\mathbb{Z}$ around a point $q \in \mathbb{H}$ corresponding to $r = 0$. The metric takes the form $dr^2 + (\sinh r)^2 d\varphi^2$. Defining $C_r(q)$ as in the case of S^2, we find $\kappa_{C_r(q)} = \frac{1}{\tanh r}$. Observe that $\tanh r$ takes all the values in $(0, 1)$ exactly once, for $r \in (0, +\infty)$. Therefore, if $s \in (1, +\infty)$, the trajectories of the flow are the closed curves $C_{r(s)}(q)$, where q varies in $\mathbb{H}$ and

$$r(s) \;=\; \operatorname{arctanh} \frac{1}{s} \;\in\; (0, +\infty) \tag{10}$$

In particular, for s in this range all periodic orbits are contractible. The formula for the periods now reads

$$T(s) = \frac{2\pi s}{\sqrt{s^2-1}}\,.$$

To understand what happens, when $s \leq 1$ we take the upper half-plane as a model for the hyperbolic plane. Thus, let $\mathbb{H} = \{\, z = (x, y) \in \mathbb{C} \mid y > 0 \,\}$. In these coordinates, the hyperbolic metric has the form $\frac{dx^2+dy^2}{y^2}$ and

$$\mathrm{Iso}_+(\mathbb{H}) \;=\; \Big\{\, z \mapsto \frac{az+b}{cz+d} \;\Big|\; a, b, c, d \in \mathbb{R}\,,\; ad - bc \;=\; 1 \Big\}\,.$$

We readily see that the affine transformations $z \mapsto az$, with $a > 0$ form a subgroup of $\mathrm{Iso}_+(\mathbb{H})$. This subgroup preserves all the Euclidean rays from the origin and acts transitively on each of them. Hence, we conclude that such curves have constant geodesic curvatures. If $\varphi \in (0, \pi)$ is the angle made by such ray with the x-axis, we find that the geodesic curvature of such ray is $\cos\varphi$. In order to do such computation one has to write the metric using Euclidean polar coordinates centered at the origin. Using the whole isometry group, we see that all the segments of circle intersecting $\partial\mathbb{H}$ with angle φ have geodesic curvature $\cos\varphi$.

We claim that if $s \in (0, 1)$ and $\nu \neq 0$ is a free homotopy class of loops of M, there is a unique closed curve $\gamma_{s,\nu}$ in the class ν, which has geodesic curvature s. The class ν correspond to a conjugacy class in $\pi_1(M)$. We identify $\pi_1(M)$ with the set of deck transformations and we let $F : \mathbb{H} \to \mathbb{H}$ be a deck transformation belonging to the given conjugacy class. By a standard result in hyperbolic geometry, F has two fixed points on $\partial\mathbb{H}$ (remember, for example, that there exists a geodesic in $\mathbb{H}$ invariant under F). Then, $\gamma_{s,\nu}$ is the projection to M of the unique segment of circle connecting the fixed points of F and making an angle $\varphi = \arccos s$ with $\partial\mathbb{H}$. The uniqueness of $\gamma_{s,\nu}$ stems form the uniqueness of such segment of circle.

In a similar fashion, we consider the subgroup of $\mathrm{Iso}_+(\mathbb{H})$ made by the maps $z \mapsto z + b$, with $b \in \mathbb{R}$. It preserves the horizontal line $\{y = 1\}$ and act transitively on it. Hence, such curve has constant geodesic curvature. A computation shows that it is equal to 1, if it is oriented by ∂_x. Using the whole isometry group, we see that all the circles tangent to $\partial\mathbb{H}$ have geodesic curvature equal to 1. Following [Gin96] we see that there is no closed curve in M with such geodesic curvature. By contradiction, if such curve exist, then its lift would be preserved by a non-constant deck transformation. We can assume without loss of generality that such lift is the line $\{y = 1\}$. We readily see that the only elements in $\mathrm{Iso}_+(\mathbb{H})$ which preserve $\{y = 1\}$ are the horizontal translation. However, no such transformation can be a deck transformation, since it has only one fixed point on $\partial\mathbb{H}$.

Exercise 5. *Show that in this case* $c(L, \sigma) = \frac{1}{2}$.

1.7. **The Main Theorem.** We are now ready to state the central result of this mini-course.

Theorem 1.6. *The following four statements hold.*

(1) *Suppose* $[\widetilde{\sigma}]_b = 0$. *For every* $k > c(L, \sigma)$,
    (a) *there exists a closed orbit on* Σ_k *in any non-trivial free homotopy class;*
    (b) *if* $\pi_{d+1}(M) \neq 0$ *for some* $d \geq 1$, *there exists a contractible orbit on* Σ_k.
(2) *Suppose* $[\widetilde{\sigma}] = 0$. *There exists a contractible orbit on* Σ_k, *for almost every energy* $k \in (e_0(L), c(L, \sigma))$.
(3) *Suppose* $[\widetilde{\sigma}] \neq 0$. *There exists a contractible orbit on* Σ_k, *for almost every energy* $k \in (e_0(L), +\infty)$.
(4) *There exists a contractible orbit on* Σ_k, *for almost every* $k \in (e_m(L), e_0(L))$.

The set for which existence holds in (2), (3) and (4) contains all the k'*s for which* Σ_k^* *is a stable hypersurface in* (T^*M, ω_σ) *(see* [HZ94, *page 122*]*).*

In these notes, we will prove (1), (2) and (3) above by relating closed orbits of the flow to the zeros of a closed 1-form η_k on the space of loops on M. We introduce such form and prove some of its general properties in Section 2. In Section 3 we describe an abstract minimax method that we apply in Section 4 to obtain zeros of η_k in the specific cases listed in the theorem. A proof of (4) relies on different methods and it can be found in [AB14].

Remark 1.7. *When* $[\sigma] = 0$, *the theorem was proven by Contreras* [Con06]. *Point (1) and (2), with the additional hypothesis* $[\widetilde{\sigma}]_b = 0$, *were proven by Osuna* [Osu05]. *Point (2) was proven in* [Mer10], *for electromagnetic Lagrangians, and in* [AB14] *for general systems. A sketch of the proof of point (3) was given in* [Nov82, Section 3] *and in* [Koz85, Section 3.2]. *It was rigorously established in* [AB14]. *Point (4) follows by employing tools in symplectic geometry. For the weakly exact case it can also be proven using a variational approach as shown in* [Abb13, Section 7]. *For Lagrangians of mechanical type and vanishing magnetic form the existence problem in such interval has historically received much attention (see* [Koz85, Section 2] *and references therein).*

We end up this introduction by defining the notion of stability mentioned in the theorem.

1.8. **Stable hypersurfaces.** In general, the dynamics on Σ_k^* may exhibit very different behaviours as k changes. However, given a regular energy level $\Sigma_{k_0}^*$, in some special cases we can find a new Hamiltonian $H' : T^*M \to \mathbb{R}$ such that $\{H' = k_0'\} = \Sigma_{k_0}^*$ and such that $\Phi^{(H',\sigma)}|_{\{H'=k_0'\}}$ and $\Phi^{(H',\sigma)}|_{\{H'=k'\}}$ are conjugated, up to a time reparametrization, provided k' is sufficiently close to k_0'.

Definition 1.8. *We say that an embedded hypersurface $\imath : \Sigma^* \longrightarrow T^*M$ is* stable *in the symplectic manifold (T^*M, ω_σ) if there exists an open neighbourhood W of Σ^* and a diffeomorphism $\Psi_W : \Sigma^* \times (-\varepsilon_0, \varepsilon_0) \to W$ with the property that:*

- $\Psi_W|_{\Sigma^* \times \{0\}} = \imath$;
- *the function $H^W : W \to \mathbb{R}$ defined through the commutative diagram*

$$\begin{tikzcd} \Sigma^* \times (-\varepsilon_0, \varepsilon_0) \arrow[r, "\Psi_W"] \arrow[d, "\mathrm{pr}_2"'] & W \arrow[dl, "H^W"] \\ (-\varepsilon_0, \varepsilon_0) & \end{tikzcd},$$

  is such that, for every $k \in (-\varepsilon_0, \varepsilon_0)$,

$$\Phi^{(H^W,\sigma)}|_{\{H^W=0\}} \quad \text{and} \quad \Phi^{(H^W,\sigma)}|_{\{H^W=k\}}$$

  are conjugated by the diffeomorphism

$$w \mapsto \Psi_W(\imath^{-1}(w), k)$$

  up to time reparametrization. In this case, the reparametrizing maps $\tau_{(z,k)}$ vary smoothly with $(z,k) \in \Sigma^* \times (-\varepsilon_0, \varepsilon_0)$ and satisfy $\tau_{(z,0)} = \mathrm{Id}_{\mathbb{R}}$, for all $z \in \Sigma^*$.

This implies that there is a bijection between the periodic orbits on $\Sigma^* = \{H^W = 0\}$ and those on $\{H^W = k\}$.

Thanks to a result of Macarini and G. Paternain [MP10], if Σ^* is the energy level of some Tonelli Hamiltonian, the function H^W can be taken to be Tonelli as well.

Proposition 1.9. *Suppose that for some $k > e_0(L)$, Σ^*_k is stable with stabilizing neighbourhood W. Up to shrinking W, there exists a Tonelli Hamiltonian $H_k : T^*M \to \mathbb{R}$ such that $H^W = H_k$ on W.*

In order to check whether an energy level is stable or not, we give the following necessary and sufficient criterion that can be found in [CM05, Lemma 2.3].

Proposition 1.10. *A hypersurface Σ^*_k is stable if and only if there exists $\alpha \in \Omega^1(\Sigma^*_k)$ such that*

$$\text{(a)}\ \ d\alpha(X_{(H,\sigma)}, \cdot\,) \;=\; 0\,, \qquad \text{(b)}\ \ \alpha(X_{(H,\sigma)})(z) \;\neq\; 0\,, \quad \forall\, z \in \Sigma^*_k\,.$$

In this case α is called a stabilizing form*. The first condition is implied by the following stronger assumption*

$$\text{(a')}\ \ d\alpha \;=\; \omega_\sigma|_{\Sigma^*_k}\,.$$

If (a') and (b) are satisfied we say that Σ^*_k is of contact type *and we call α a contact form. We distinguish between* positive *and* negative *contact forms according to the sign of the function $\alpha(X_{(H,\sigma)})$.*

In Section 6, we give some sufficient criteria for stability for magnetic flows on surfaces.

2. The free period action form

For the proof of the Main Theorem we need to characterize the periodic orbits on Σ_k via a variational principle on a space of loops. To this purpose we have first to adjust L.

2.1. Adapting the Lagrangian.

Let us introduce a class of Tonelli Lagrangians whose fibrewise growth is quadratic. In this class we will be enabled to define the action functional on the space of loops with square-integrable velocity.

Definition 2.1. *We say that L is* quadratic at infinity *if there exists a metric g_∞ and a potential $V_\infty : M \to \mathbb{R}$ such that $L(q,v) = \frac{1}{2}|v|^2_\infty - V_\infty(q)$ outside a compact set.*

The next result tells us that, if we look at the dynamics on a fixed energy level, it is not restrictive to assume that the Lagrangian is quadratic at infinity.

Proposition 2.2. *For any fixed $k \in \mathbb{R}$, there exists a Tonelli Lagrangian $L_k : TM \to \mathbb{R}$ which is quadratic at infinity and such that $L_k = L$ on $\{E \le k_0\}$, for some $k_0 > k$. By choosing k_0 sufficiently large, we can obtain $e_0(L) = e_0(L_k)$ and, if $[\widetilde{\sigma}] = 0$, also $c(L,\sigma) = c(L_k,\sigma)$.*

From now on, we assume that L is quadratic at infinity. In this case there exist positive constants C_0 and C_1 such that

$$C_1|v|^2 \;-\; C_0 \;\le\; L(q,v) \;\le\; C_1|v|^2 \;+\; C_0\,, \quad \forall\,(q,v) \in TM\,. \tag{11}$$

An analogous statement holds for the energy.

2.2. The space of loops.

We define the space of loops where the variational principle will be defined. Given $T > 0$, we call $W^{1,2}(\mathbb{R}/T\mathbb{Z}, M)$ the set

$$\Big\{\, \gamma : \mathbb{R}/T\mathbb{Z} \to M \;\Big|\; \gamma \text{ is absolutely continuous}\,,\; \int_0^T |\dot\gamma|^2\,\mathrm{d}t < \infty \,\Big\}\,.$$

Since we look for periodic orbits of arbitrary period, we want to let T vary among all the positive real numbers $\mathbb{R}^+$. This is the same as fixing the parametrization space to $\mathbb{T} := \mathbb{R}/\mathbb{Z}$ and keeping track of the period as an additional variable. Namely, we have the identification

$$\begin{aligned} \bigsqcup_{T>0} W^{1,2}(\mathbb{R}/T\mathbb{Z}, M) \;&\longrightarrow\; \Lambda \,:=\, W^{1,2}(\mathbb{T}, M) \times \mathbb{R}^+ \\ \gamma(t) \;&\longmapsto\; \big(x(s) := \gamma(sT), T\big)\,. \end{aligned}$$

Given a free homotopy class $\nu \in [\mathbb{T}, M]$, we denote by $W^{1,2}_\nu \subset W^{1,2}(\mathbb{T}, M)$ and $\Lambda_\nu \subset \Lambda$ the loops belonging to such class. We use the symbol 0 for the class of contractible loops.

Proposition 2.3. *The set Λ is a Hilbert manifold with $T_{(x,T)}\Lambda \simeq T_x W^{1,2} \times \mathbb{R}$, where $T_x W^{1,2} \simeq W^{1,2}(\mathbb{T}, x^*(TM))$ is the space of absolutely continuous vector fields along x with square-integrable covariant derivative. The metric on Λ is given by $g_\Lambda = g_{W^{1,2}} + \mathrm{d}T^2$, where*

$$(g_{W^{1,2}})_x(\xi_1,\xi_2) \;:=\; \int_0^1 g_{x(s)}(\xi_1(s),\xi_2(s))\,\mathrm{d}s \;+\; \int_0^1 g_{x(s)}(\xi_1'(s),\xi_2'(s))\,\mathrm{d}s\,.$$

For any $T_- > 0$, $W^{1,2} \times [T_-, +\infty) \subset \Lambda$ is a complete metric space.

For more details on the space of loops we refer to [Abb13, Section 2] and [Kli78]. We end this subsection with two more definitions, which will be useful later on. First, we let

$$\frac{\partial}{\partial T} \;\in\; \Gamma(\Lambda)$$

denote the coordinate vector associated with the variable T. Then, if $x \in W^{1,2}$, we let

$$e(x) \;:=\; \int_0^1 |x'|^2 \mathrm{d}s \qquad \text{and} \qquad \ell(x) \;:=\; \int_0^1 |x'|\, \mathrm{d}s$$

be the L^2-*energy* and the *length* of x, respectively. We define analogous quantities for $\gamma \in \Lambda$. We readily see that $\ell(x) = \ell(\gamma)$ and $e(x) = Te(\gamma)$. Moreover, $\ell(x)^2 \le e(x)$ holds.

2.3. **The action form.** In this subsection, for every $k \in \mathbb{R}$, we construct $\eta_k \in \Omega^1(\Lambda)$, which vanishes exactly at the set of periodic orbits on Σ_k. Such 1-form will be made of two pieces: one depending only on L and k and one depending only on σ. The first piece will be the differential of the function

$$\begin{aligned} A_k : \Lambda &\longrightarrow \mathbb{R} \\ \gamma &\longmapsto \int_0^T \Big[L(\gamma, \dot\gamma) + k \Big]\, \mathrm{d}t \;=\; T \cdot \int_0^1 \Big[L\Big(x, \frac{x'}{T} \Big) + k \Big]\, \mathrm{d}s\,. \end{aligned}$$

Such function is well-defined since L is quadratic at infinity (see (11)). It was proven in [AS09] that A_k is a $C^{1,1}$ function (namely, A_k is differentiable and its differential is locally uniformly Lipschitz-continuous).

In order to define the part of η_k depending on σ, we first introduce a differential form $\tau^\sigma \in \Omega^1(W^{1,2})$ called the *transgression* of σ. It is given by

$$\tau^\sigma_x(\xi) \;:=\; \int_0^1 \sigma_{x(s)}(\xi(s), x'(s))\, \mathrm{d}s\,, \quad \forall\, (x,\xi) \in TW^{1,2}\,.$$

By writing τ^σ in local coordinates, it follows that it is locally uniformly Lipschitz.

If $u : [0,1] \to W^{1,2}$ is a path of class C^1, then

$$\int_0^1 u^*\tau^\sigma \;=\; \int_{[0,1]\times\mathbb{T}} \hat u^*\sigma\,, \tag{12}$$

where $\hat u : [0,1]\times\mathbb{T} \to M$ is the cylinder given by $\hat u(r,t) = u(r)(t)$. If $u_a : \mathbb{T} \to W^{1,2}$ is a homotopy of closed paths with parameter $a \in [0,1]$, then we get a corresponding homotopy of tori $\hat u_a$. Since σ is closed, the integral of $\hat u_a^*\sigma$ on $\mathbb{T}^2$ is independent of a. We conclude that the integral of τ^σ on u_a does not depend on a either. Namely, τ^σ is a *closed form.*

Definition 2.4. *The* free period action form *at energy k is $\eta_k \in \Omega^1(\Lambda)$ defined as*

$$\eta_k \;:=\; dA_k \;-\; \mathrm{pr}^*_{W^{1,2}}\,\tau^\sigma\,, \tag{13}$$

where $\mathrm{pr}^*_{W^{1,2}} : \Lambda \to W^{1,2}$ is the natural projection $(x,T) \mapsto x$.

Proposition 2.5. *The free period action form is closed and its zeros correspond to the periodic orbits of $\Phi^{(L,\sigma)}$ on Σ_k.*

The correspondence with periodic orbits follows by computing η_k explicitly on $TW^{1,2} \times 0$ and on $\frac{\partial}{\partial T}$. If $\xi \in TW^{1,2}$, then

$$(\eta_k)_\gamma(\xi, 0) \;=\; \int_0^T \Big[\frac{\partial L}{\partial q}(\gamma,\dot\gamma)\cdot \xi_T \;+\; \frac{\partial L}{\partial v}(\gamma,\dot\gamma)\cdot \dot\xi_T \;+\; \sigma_\gamma(\dot\gamma,\xi_T)\Big]\mathrm{d}t\,, \tag{14}$$

where ξ_T is the reparametrization of ξ on $\mathbb{R}/T\mathbb{Z}$. In the direction of the period we have

$$\begin{aligned}(\eta_k)_\gamma\left(\frac{\partial}{\partial T}\right) \;=\; d_\gamma A_k\left(\frac{\partial}{\partial T}\right) \;&=\; \int_0^1 L\left(x, \frac{x'}{T}\right)\mathrm{d}s + \\ &\quad+\; k \;-\; T\cdot\int_0^1 \frac{\partial L}{\partial v}\left(x,\frac{x'}{T}\right)\cdot\frac{x'}{T^2}\,\mathrm{d}s \\ &=\; k - \int_0^1 E\left(x,\frac{x'}{T}\right)\mathrm{d}s \\ &=\; k - \frac{1}{T}\int_0^T E(\gamma,\dot\gamma)\,\mathrm{d}t\,.\end{aligned} \tag{15}$$

2.4. **Vanishing sequences.** Our strategy to prove existence of periodic orbits will be to construct zeros of η_k by approximation.

Definition 2.6. *Let $\nu \in [\mathbb{T}, M]$ be a free homotopy class. A sequence $(\gamma_m) \subset \Lambda_\nu$ is called a* vanishing sequence *(at level k), if*

$$\lim_{m\to\infty} |\eta_k|_{\gamma_m} \;=\; 0\,.$$

A limit point of a vanishing sequence is a zero of η_k. Thus, the crucial question is: when does a vanishing sequence admit limit points? Clearly, if $T_m \to 0$ or $T_m \to +\infty$ the set of limit points is empty. We now see that the opposite implication also holds.

Lemma 2.7. *If (γ_m) is a vanishing sequence, there exists $C > 0$ such that*

$$e(x_m) \;\le\; C\cdot T_m^2\,. \tag{16}$$

Proof. We compute

$$C_1\cdot\frac{e(x_m)}{T_m^2} - C_0 \;\overset{(\star)}{\le}\; \int_0^1 E\left(x_m,\frac{x_m'}{T_m}\right)\mathrm{d}s \;=\; k - \eta^k_{\gamma_m}\left(\frac{\partial}{\partial T}\right) \;\overset{(\star\star)}{\le}\; k + \sup_m |\eta_k|_{\gamma_m}\,.$$

where in $(\star)$ we used (11) applied to E, and in $(\star\star)$ we used that

$$\left|\frac{\partial}{\partial T}\right| \;=\; 1\,.$$

The desired estimate follows by observing that, since the sequence $\big(|\eta_k|_{\gamma_m}\big) \subset [0,+\infty)$ is infinitesimal, it is also bounded from above. $\square$

Proposition 2.8. *If (γ_m) is a vanishing sequence and $0 < T_- \le T_m \le T_+ < +\infty$ for some T_- and T_+, then (γ_m) has a limit point.*

Proof. By compactness of $[T_-, T_+]$, up to subsequences, $T_m \to T_\infty > 0$. By (16), the L^2-energy of x_m is uniformly bounded. Thus, (x_m) is uniformly 1/2-Hölder continuous. By the Arzelà-Ascoli theorem, up to subsequences, (x_m) converges uniformly to a continuous $x_\infty : \mathbb{T} \to M$. Therefore, x_m eventually belongs to a local chart $\mathcal{U}$ of $W^{1,2}$. In $\mathcal{U}$, η_k can be written as the differential of a standard action functional depending on time (see [AB14]) and the same argument contained in [Abb13, Lemma 5.3] when $\sigma = 0$ implies that (γ_m) has a limit point. □

In order to construct vanishing sequences we will exploit some geometric properties of η_k. One of the main ingredients to achieve this goal will be to define a vector field on Λ generalizing the negative gradient vector field of the function A_k. We introduce it in the next subsection.

2.5. **The flow of steepest descent.** Let X_k denote the vector field on Λ defined by

$$X_k := -\frac{\sharp\, \eta_k}{\sqrt{1 + |\eta_k|^2}}$$

where $\sharp$ denote the duality between 1-forms and vector fields induced by g_Λ. Since X_k is locally uniformly Lipschitz, it gives rise to a flow which we denote by $r \mapsto \Phi^k_r$. For every $\gamma \in \Lambda$, we denote by $u_\gamma : [0, R_\gamma) \to \Lambda$ the maximal positive flow line starting at γ. We say that Φ^k is *positively complete* on a subset $Y \subset \Lambda$ if, for all $\gamma \in \Lambda$, either $R_\gamma = +\infty$ or there exists $R_{\gamma,Y} \in [0, R_\gamma)$ such that $u_\gamma(R_{\gamma,Y}) \notin Y$.

Except for the scaling factor $1/\sqrt{1 + |\eta_k|^2}$, the vector field X_k is the natural generalization of $-\nabla A_k = -\sharp(dA_k)$ to the case of non-vanishing magnetic form. We introduce such scaling so that $|X_k| \leq 1$ and we can give the following characterization of the flow lines u_γ with $R_\gamma < +\infty$.

Proposition 2.9. *Let $u : [0, R) \to \Lambda$ be a maximal positive flow line of X_k and for all $r \in [0, R)$ set $u(r) := \gamma(r) = (x(r), T(r))$. If $R < +\infty$, then there exists a sequence $(r_m)_{m\in\mathbb{N}} \subset [0, R)$ and a constant C such that*

$$\lim_{m\to\infty} r_m = R\,, \qquad \lim_{m\to\infty} T(r_m) = 0\,, \qquad e(x(r_m)) \;\leq\; C \cdot T(r_m)^2\,, \quad \forall\, m \in \mathbb{N}\,. \tag{17}$$

Proof. By contradiction, we suppose that $0 < T_- := \inf_{[0,R)} T(r)$. Since $|X_k| \leq 1$, u_γ is uniformly continuous and, by the completeness of $W^{1,2} \times [T_-, +\infty)$, there exists

$$\gamma_\infty := \lim_{r\to R} u(r)\,.$$

By the existence theorem of solutions of ODE's, there exists a neighbourhood $\mathcal{B}$ of γ_∞ and $R_{\mathcal{B}} > 0$ such that

$$\forall\, \gamma \in \mathcal{B}\,, \quad r \longmapsto \Phi^k_r(\gamma) \text{ exists in } [0, R_{\mathcal{B}}]\,.$$

This contradicts the fact that R is finite as soon as $r \in [0, R)$ is such that $\gamma(r) \in \mathcal{B}$ and $R - r < R_{\mathcal{B}}$. Therefore, $\inf T = 0$. Hence, we find a sequence $r_m \to R$ such that $T(r_m) \to 0$ and, for every m, $\frac{dT}{dr}(r_m) \leq 0$. The last property implies that

$$0 \;\geq\; \frac{dT}{dr}(r_m) \;=\; d_{u(r_m)}T(X_k) \;=\; -\frac{\eta_k\left(\frac{\partial}{\partial T}\right)}{\sqrt{1 + |\eta_k|^2}}(u(r_m))\,. \tag{18}$$

Finally, using Equation (15) and the estimates in (11), we have

$$\begin{aligned} 0 \leq (\eta_k)_{u(r(m))}\left(\frac{\partial}{\partial T}\right) &= k - \int_0^1 E\left(x(r_m), \frac{x'(r_m)}{T(r_m)}\right) \mathrm{d}s \\ &\leq k - C_1 \int_0^1 \frac{|x'(r_m)|^2}{T(r_m)^2} \mathrm{d}s + C_0 \,. \end{aligned}$$

□

The above proposition shows that flow lines whose interval of definition is finite come closer and closer to the subset of constant loops. As we saw in Lemma 2.7 the same is true for vanishing sequences with infinitesimal period. For these reasons in the next subsection we study the behaviour of the action form on the set of loops with short length.

2.6. **The subset of short loops.** We now define a local primitive for η_k close to the subset of constant loops. For $k > e_0(L)$, such primitive will enjoy some properties that will enable us to apply the minimax theorem of Section 3 to prove the Main Theorem. For our arguments we will need estimates which hold uniformly on a compact interval of energies. Hence, for the rest of this subsection we will suppose that a compact interval $I \subset (e_0(L), +\infty)$ is fixed.

Let $M_0 \subset W_0^{1,2}$ be the constant loops parametrized by $\mathbb{T}$ and $M_0 \times \mathbb{R}^+ \subset \Lambda_0$ the constant loops with arbitrary period. We readily see that $\tau^\sigma|_{M_0} = 0$. Thus, $\eta_k = dA_k|_{M_0 \times \mathbb{R}^+}$ and

$$A_k(x,T) \;=\; T\,(k - V(x))\,, \quad \forall\,(x,T) \in M_0 \times \mathbb{R}^+\,. \tag{19}$$

Now that we have described η_k on constant loops, let us see what happens nearby. First, we need the following lemma.

Lemma 2.10. *There exists $\delta_* > 0$ such that $\{\ell < \delta\} \subset W^{1,2}$ retracts with deformation on M_0, for all $\delta \leq \delta_*$. Thus, we have $\tau^\sigma|_{\{\ell<\delta_*\}} = dP^\sigma$, where*

$$\begin{aligned} P^\sigma : \{\ell < \delta_*\} &\longrightarrow \mathbb{R} \\ x &\longmapsto \int_{B^2} \hat{u}_x^* \sigma\,, \end{aligned} \tag{20}$$

where $\hat{u}_x : B^2 \to M$ is the disc traced by x under the action of the deformation retraction. Furthermore, there exists $C > 0$ such that

$$|P^\sigma(x)| \;\leq\; C \cdot \ell(x)^2\,. \tag{21}$$

Proof. Choose $\delta < 2\rho(g)$, where $\rho(g)$ is the injectivity radius of g. With this choice, for each $x \in \{\ell < \delta\}$ and each $s \in \mathbb{T}$, there exists a unique geodesic $y_s : [0,1] \to M$ joining $x(0)$ to $x(s)$. For each $a \in [0,1]$ define $x_a : \mathbb{T} \to M$ by $x_a(s) := y_s(a)$. Taking a smaller δ if necessary, one can prove that $a \mapsto |x_a'|$ is a non-decreasing family of functions (use normal coordinates at $x(0)$). Thus, $a \mapsto \ell(x_a)$ is non-decreasing as well and

$$\begin{aligned} [0,1] \times \{\ell < \delta\} &\longrightarrow \{\ell < \delta\} \\ (a,x) &\longmapsto x_a \end{aligned}$$

yields the desired deformation. In order to estimate P^σ is enough to bound the area of the deformation disc $\hat{u}_x$:

$$\begin{aligned} \mathrm{area}(\hat{u}_x) &\leq \int_0^1 \mathrm{d}a \int_0^1 \left|\frac{dy_s}{da}(a)\right| \cdot |x_a'(s)|\,\mathrm{d}s \\ &\leq \int_0^1 \mathrm{d}a \int_0^1 d(x(0),x(s))|x'(s)|\,\mathrm{d}s \leq \frac{\ell(x)}{2}\ell(x)\,. \end{aligned}$$

□

In view of this lemma, for all $\delta \in (0,\delta_*]$, we define the set

$$\mathcal{V}^\delta \;:=\; \{\ell < \delta\} \times \mathbb{R}^+ \;\subset\; \Lambda_0 \tag{22}$$

and the function $S_k : \mathcal{V}^{\delta_*} \longrightarrow \mathbb{R}$ given by

$$S_k \;:=\; A_k \;-\; P^\sigma \circ \mathrm{pr}_{W^{1,2}}\;. \tag{23}$$

Such a function is a primitive of η_k on $\mathcal{V}^{\delta_*}$. By (11), it admits the following upper bound.

Proposition 2.11. *There exists $C > 0$ such that, for every $\gamma \in \mathcal{V}^{\delta_*}$, there holds*

$$S_k(\gamma) \;\leq\; C \cdot \left(\frac{e(x)}{T} \;+\; T \;+\; \ell(x)^2\right), \qquad \forall\, k \in I\,. \tag{24}$$

This result has an immediate consequence on vanishing sequences and flow lines of Φ^k.

Corollary 2.12. *Let $b > 0$ and $k \in I$ be fixed. The following two statements hold:*
1. *if (γ_m) is a vanishing sequence such that $\gamma_m \notin \{S_k < b\}$ for all $m \in \mathbb{N}$, then T_m is bounded away from zero;*
2. *the flow Φ^k is positively complete on the set $\Lambda \setminus \{S_k < b\}$.*

We conclude this section by showing that the infimum of S_k on short loops is zero and it is approximately achieved on constant loops with small period. Furthermore, S_k is bounded away from zero on the set of loops having some fixed positive length.

Proposition 2.13. *There exist $\delta_I \leq \delta_*$ and positive numbers b_I, T_I such that, for all $k \in I$,*

$$\text{(a)}\;\; \inf_{\mathcal{V}^{\delta_I}} S_k \;=\; 0\,, \qquad \text{(b)}\;\; \inf_{\partial\mathcal{V}^{\delta_I}} S_k \;\geq\; b_I\,, \qquad \text{(c)}\;\; \sup_{M_0\times\{T_I\}} S_k \;<\; \frac{b_I}{2}\,. \tag{25}$$

Proof. Since for all $q \in M$, the function $L|_{T_qM}$ attains its minimum at $(q,0)$, the estimate from below on L obtained in (11) can be refined to

$$L(q,v) \;\geq\; C_1|v|^2 + \min_{q\in M} L(q,0) \;=\; C_1|v|^2 \;-\; e_0(L)\,.$$

From this inequality and (21), we can bound from below $S_k(\gamma)$:

$$\begin{aligned} S_k(\gamma) \;&\geq\; T \cdot \int_0^1 \Big[C_1 \cdot \frac{|x'|^2}{T^2} - e_0(L) + k\Big]\mathrm{d}s \;-\; C\cdot\ell(x)^2 \\ &\geq\; C_1 \cdot \frac{e(x)}{T} + (k - e_0(L))\cdot T \;-\; C\cdot\ell(x)^2 \\ &\overset{(\star)}{\geq}\; 2\sqrt{C_1(\min I - e_0(L))}\cdot\ell(x) \;-\; C\cdot\ell(x)^2\,. \end{aligned}$$

where in $(\star)$ we made use of the inequality between arithmetic and geometric mean. Hence, there exists $\delta_I > 0$ sufficiently small, such that the last quantity is positive if $\ell(x) < \delta_I$ and bounded from below by

$$b_I \ := \ 2\sqrt{C_1(\min I - e_0(L))} \cdot \delta_I \ - \ C \cdot \delta_I^2 \ > 0$$

if $\ell(x) = \delta_I$. This implies Inequality *(b)* in (25) and that $\inf_{V^{\delta_I}} S_k \geq 0$. To prove that $\inf_{V^{\delta_I}} S_k \leq 0$ and that there exists T_I such that Inequality *(c)* in (25) holds, we just recall from (19) that

$$\lim_{T\to 0} \ \sup_{M_0\times\{T\}} \ S_k \ = \ 0\,. \qquad \square$$

In the next section we will prove a minimax theorem for a class of closed 1-form on abstract Hilbert manifolds. Such a class will satisfy a general version of the properties we have proved so far for η_k.

3. The minimax technique

In this section we present an abstract minimax technique which represents the core of the proof of the Main Theorem. We formulate it in a very general form on a non-empty Hilbert manifold $\mathscr{H}$.

3.1. **An abstract theorem.** We start by setting some notation for homotopy classes of maps from Euclidean balls into $\mathscr{H}$. Let $d \in \mathbb{N}$ and $\mathscr{U}$ be a subset of $\mathscr{H}$. Define $\big[(B^d, \partial B^d), (\mathscr{H}, \mathscr{U})\big]$ as the set of homotopy classes of maps $\gamma : (B^d, \partial B^d) \to (\mathscr{H}, \mathscr{U})$. By this we mean that the maps send B^d to $\mathscr{H}$ and ∂B^d to $\mathscr{U}$, and that the homotopies do the same. The classes $[\gamma]$, where γ is such that $\gamma(B^d) \subset \mathscr{U}$ are called *trivial.* If $\mathscr{U}' \subset \mathscr{U}$, we have a map

$$i_{\mathscr{U}}^{\mathscr{U}'} : \big[(B^d, \partial B^d), (\mathscr{H}, \mathscr{U}')\big] \longrightarrow \big[(B^d, \partial B^d), (\mathscr{H}, \mathscr{U})\big]$$

We are now ready to state the main result of this section.

Theorem 3.1. *Let $\mathscr{H}$ be a non-empty Hilbert manifold, $\mathscr{I} = [k_0, k_1]$ be a compact interval and $d \geq 1$ an integer. Let $\alpha_k \in \Omega^1(\mathscr{H})$ be a family of Lipschitz-continuous forms parametrized by $k \in \mathscr{I}$ and such that*

- *the integral of α_k over contractible loops vanishes;*
- *$\alpha_k = \alpha_{k_0} + (k - k_0)d\mathscr{T}$, where $\mathscr{T} : \mathscr{H} \to (0, +\infty)$ is a $C^{1,1}$ function such that*

$$\sup_{\mathscr{H}} |d\mathscr{T}| \ < \ +\infty\,. \tag{26}$$

Define the vector field

$$\mathscr{X}_k \ := \ -\frac{\sharp\,\alpha_k}{\sqrt{1 + |\alpha_k|^2}}\,, \tag{27}$$

where $\sharp$ is the metric duality, and suppose that there exists an open set $\mathscr{V} \subset \mathscr{H}$ such that:

- *there exists $\mathscr{S}_k : \overline{\mathscr{V}} \to \mathbb{R}$ satisfying*

$$d\mathscr{S}_k \ = \ \alpha_k\,, \qquad \mathscr{S}_k \ = \ \mathscr{S}_{k_0} \ + \ (k - k_0)\,\mathscr{T}\,; \tag{28}$$

- *there exists a real number*

$$\beta_0 \ < \ \inf_{\partial\mathscr{V}} \mathscr{S}_{k_0} \ =: \ \beta_{\partial\mathscr{V}} \tag{29}$$

such that the flow $r \mapsto \Phi_r^{\mathscr{X}_k}$ is positively complete on the set $\mathscr{H}\setminus\{\mathscr{S}_k < \beta_0\}$;
- *there exists a set $\mathscr{M} \subset \{\mathscr{S}_{k_1} < \beta_0\}$ and a class $\mathscr{G} \in [(B^d, \partial B^d), (\mathscr{H}, \mathscr{M})]$ such that $i_{\mathscr{V}}^{\mathscr{M}}(\mathscr{G})$ is non-trivial.*

Then, the following two statements hold true. First, for all $k \in \mathscr{I}$, there exists a sequence $(h_m^k)_{m\in\mathbb{N}} \subset \mathscr{H} \setminus \{\mathscr{S}_k < \beta_0\}$ such that

$$\lim_{m\to\infty} |\alpha_k|_{h_m^k} \ = \ 0\,.$$

Second, there exists a subset $\mathscr{I}_* \subset \mathscr{I}$ such that

- *$\mathscr{I} \setminus \mathscr{I}_*$ is negligible with respect to the 1-dimensional Lebesgue measure;*
- *for all $k \in \mathscr{I}_*$ we have*

$$\sup_{m\in\mathbb{N}} \mathscr{T}(h_m^k) \ < \ +\infty\,.$$

Moreover, if there exists a $C^{1,1}$-function $\widehat{\mathscr{S}_k} : \mathscr{H} \to \mathbb{R}$ which extends $\mathscr{S}_k$ and satisfies (28) on the whole $\mathscr{H}$, we also have that

$$\lim_{m\to\infty} \widehat{\mathscr{S}_k}(h_m^k) \ = \ \inf_{\gamma\in\mathscr{G}} \sup_{\xi\in B^d} \ \widehat{\mathscr{S}_k} \circ \gamma\,(\xi) \ \geq \ \beta_{\partial\mathscr{V}}\,. \tag{30}$$

To prove Theorem 1.6*(1a)* we will also need a version of the minimax theorem for $d = 0$, namely when the maps are simply points in $\mathscr{H}$. We state it here for a single function and not for a 1-parameter family since this will be enough for the intended application. For a proof we refer to [Abb13, Remark 1.10].

Theorem 3.2. *Let $\mathscr{H}$ be a non-empty Hilbert manifold and let $\widehat{\mathscr{S}} : \mathscr{H} \to \mathbb{R}$ be a $C^{1,1}$-function bounded from below. Suppose that the flow of the vector field*

$$\mathscr{X} \ := \ -\frac{\nabla\widehat{\mathscr{S}}}{\sqrt{1 + |\nabla\widehat{\mathscr{S}}|^2}}$$

is positively complete on some non-empty sublevel set of $\widehat{\mathscr{S}}$. Then, there exists a sequence $(h_m)_{m\in\mathbb{N}} \subset \mathscr{H}$ such that

$$\lim_{m\to+\infty} |d_{h_m}\widehat{\mathscr{S}}| \ = \ 0\,, \qquad \lim_{m\to+\infty} \widehat{\mathscr{S}}(h_m) \ = \ \inf_{\mathscr{H}} \widehat{\mathscr{S}}\,. \tag{31}$$

In the next two subsections we prove Theorem 3.1. First, we introduce some preliminary definitions and lemmas and then we present the core of the argument.

3.2. **Preliminary results.** We start by defining the *variation* of the 1-form α_k along any path $u : [a_0, a_1] \to \mathscr{H}$. It is the real number

$$\alpha_k(u) \ := \ \int_{a_0}^{a_1} \alpha_k\left(\frac{du}{da}\right)(u(a))\,\mathrm{d}a\,. \tag{32}$$

We collect the properties of the variation along a path in a lemma.

Lemma 3.3. *If u is a path in $\mathscr{H}$ and $\overline{u}$ is the inverse path, we have*

$$\alpha_k(\overline{u}) \ = \ -\,\alpha_k(u)\,. \tag{33}$$

If u_1 and u_2 are two paths in $\mathscr{H}$ such that the ending point of u_1 coincides with the starting point of u_2, we denote by $u_1 * u_2$ the concatenation of the two paths and we have

$$\alpha_k(u_1 * u_2) \;=\; \alpha_k(u_1) \;+\; \alpha_k(u_2)\,, \tag{34}$$

If u is a contractible closed path in $\mathscr{H}$, we have

$$\alpha_k(u) \;=\; 0\,. \tag{35}$$

Finally, let $\gamma : Z \to \mathscr{H}$ be any smooth map from a Hilbert manifold Z such that there exists a function $\mathscr{S}_k^\gamma : Z \to \mathscr{H}$ with the property that

$$d\,\mathscr{S}_k^\gamma \;=\; \gamma^*\alpha_k\,. \tag{36}$$

Then, for all paths $z : [a_0, a_1] \to Z$ we have

$$\alpha_k(\gamma \circ z) \;=\; \mathscr{S}_k^\gamma(z(a_1)) \;-\; \mathscr{S}_k^\gamma(z(a_0))\,. \tag{37}$$

Let us come back to the statement of Theorem 3.1. Fix a point $\xi_* \in \partial B^d$ and for every $\gamma \in \mathscr{G}$ define the unique $\mathscr{S}_k^\gamma : B^d \to \mathscr{H}$ such that

$$d\,\mathscr{S}_k^\gamma \;=\; \gamma^*\alpha_k\,, \qquad \mathscr{S}_k^\gamma(\xi_*) \;=\; \mathscr{S}_k(\gamma(\xi_*))\,. \tag{38}$$

We observe that this is a good definition since B^d is simply connected and $\gamma(\xi_*)$ belongs to the domain of definition of $\mathscr{S}_k$ as $\gamma \in \mathscr{G}$. Moreover, if α_k admits a global primitive $\widehat{\mathscr{S}_k}$ on $\mathscr{H}$ extending $\mathscr{S}_k$, then clearly we have $\mathscr{S}_k^\gamma = \widehat{\mathscr{S}_k} \circ \gamma$. Finally, thanks to the previous lemma, for every $\xi \in B^d$ we have the formula

$$\mathscr{S}_k^\gamma(\xi) \;=\; \mathscr{S}_k(\gamma(\xi_*)) \;+\; \alpha_k(\gamma \circ z_\xi)\,, \tag{39}$$

where $z_\xi : [0,1] \to B^d$ is any path connecting ξ_* and ξ.

Remark 3.4. *If $d \neq 1$, then $\mathscr{S}_k^\gamma$ does not depend on the choice of the point $\xi_* \in \partial B^d$ as $S^{d-1} = \partial B^d$ is connected. On the other hand, if $d = 1$ there are two possible choices for ξ_* and the two corresponding primitives of $\gamma^*\eta_k$ differ by a constant, which depends only on the class $\mathscr{G}$ and not on γ.*

Definition 3.5. *We define the* minimax function *$c_{\mathscr{G}} : \mathscr{I} \to \mathbb{R} \cup \{-\infty\}$ by*

$$c_{\mathscr{G}}(k) := \inf_{\gamma \in \mathscr{G}} \sup_{\xi \in B^d} \mathscr{S}_k^\gamma(\xi)\,. \tag{40}$$

In the next lemma we show that $c_{\mathscr{G}}(k)$ is finite and that, for each $\gamma \in \mathscr{G}$, the points almost realizing the supremum of the function $\mathscr{S}_k^\gamma$ lie in the complement of the set $\{\mathscr{S}_k < \beta_0\}$.

Lemma 3.6. *Let $k \in \mathscr{I}$ and $\gamma \in \mathscr{G}$. There holds*

$$\sup_{B^d} \mathscr{S}_k^\gamma \;\geq\; \beta_{\partial\mathscr{V}}\,. \tag{41}$$

Moreover, if $\beta_1 < \beta_{\partial\mathscr{V}}$, then $\forall\, \xi \in B^d$ the following implication holds

$$\mathscr{S}_k^\gamma(\xi) \;\geq\; \sup_{B^d} \mathscr{S}_k^\gamma \;-\; (\beta_{\partial\mathscr{V}} - \beta_1) \quad\Longrightarrow\quad \gamma(\xi) \;\notin\; \{\mathscr{S}_k < \beta_1\}\,. \tag{42}$$

Proof. Since $i^{\mathscr{M}}_{\mathscr{V}}(\mathscr{G})$ is non-trivial, the set $\{\xi \in B^d \,|\, \gamma(\xi) \in \partial\mathscr{V}\}$ is non-empty. Therefore, there exists an element $\widehat{\xi}$ in this set and a path $z_{\widehat{\xi}} : [0,1] \to B^d$ from ξ_* to $\widehat{\xi}$ such that $\gamma \circ z_{\widehat{\xi}}|_{[0,1)} \subset \mathscr{V}$. By (39) and (37) we have

$$\begin{aligned}\mathscr{S}^{\gamma}_k(\widehat{\xi}) \;&=\; \mathscr{S}_k(\gamma(\xi_*)) \,+\, \alpha_k(\gamma \circ z_{\widehat{\xi}}) \\ &=\; \mathscr{S}_k(\gamma(\xi_*)) + \Big(\mathscr{S}_k(\gamma(\widehat{\xi})) - \mathscr{S}_k(\gamma(\xi_*))\Big) \;=\; \mathscr{S}_k(\gamma(\widehat{\xi}))\,,\end{aligned}$$

which implies (41) by (29). In order to prove the second statement we consider $\xi \in B^d$ such that $\gamma(\xi) \in \{\mathscr{S}_k < \beta_1\}$. Without loss of generality there exists a path $z_{\xi,\widehat{\xi}} : [0,1] \to B^d$ from ξ to $\widehat{\xi}$ such that $z_{\xi,\widehat{\xi}}|_{[0,1)} \subset \mathscr{V}$. Using (37) twice, we compute

$$\begin{aligned}\sup_{B_d} \mathscr{S}^{\gamma}_k \geq \mathscr{S}^{\gamma}_k(\widehat{\xi}) \;&=\; \mathscr{S}^{\gamma}_k(\xi) \,+\, \alpha_k(\gamma \circ z_{\xi,\widehat{\xi}}) \\ &=\; \mathscr{S}^{\gamma}_k(\xi) \,+\, \Big(\mathscr{S}_k(\gamma(\widehat{\xi})) - \mathscr{S}_k(\gamma(\xi))\Big) > \mathscr{S}^{\gamma}_k(\xi) \,+\, \big(\beta_{\partial\mathscr{V}} - \beta_1\big)\,,\end{aligned}$$

which yields the contrapositive of the implication we had to show. $\square$

We now see that, since the family $k \mapsto \alpha_k$ is monotone in the parameter k, the same is true for the numbers $c_{\mathscr{G}}(k)$.

Lemma 3.7. *If $k_2 \leq k_3$ and $\gamma \in \mathscr{G}$, we have*

$$\mathscr{S}^{\gamma}_{k_3} \;=\; \mathscr{S}^{\gamma}_{k_2} \;+\; (k_3 - k_2)\,\mathscr{T} \circ \gamma\,. \tag{43}$$

As a consequence, $c_{\mathscr{G}}$ is a non-decreasing function.

Proof. We observe that

- $d\big(\mathscr{S}^{\gamma}_{k_3} - \mathscr{S}^{\gamma}_{k_2}\big) \;=\; \gamma^*\big(\alpha_{k_3} - \alpha_{k_2}\big) \;=\; \gamma^*\big((k_3 - k_2)\, d\mathscr{T}\big)$
- $\mathscr{S}^{\gamma}_{k_3}(\xi_*) \,-\, \mathscr{S}^{\gamma}_{k_2}(\xi_*) \;=\; \mathscr{S}_{k_3}(\gamma(\xi_*)) \,-\, \mathscr{S}_{k_2}(\gamma(\xi_*)) \;=\; (k_3 - k_2)\,\mathscr{T}(\gamma(\xi_*))\,.$

These two equalities imply that the function $\mathscr{S}^{\gamma}_{k_2} + (k_3 - k_2)\mathscr{T} \circ \gamma$ satisfies (38) with $k = k_3$. Since these conditions identify a unique function, equation (43) follows. In particular, we have $\mathscr{S}^{\gamma}_{k_2} \leq \mathscr{S}^{\gamma}_{k_3}$. Taking the inf-sup of this inequality on $\mathscr{G}$, we get $c_{\mathscr{G}}(k_2) \leq c_{\mathscr{G}}(k_3)$. $\square$

We end this subsection by adjusting the vector field $\mathscr{X}_k$ so that its flow becomes positively complete on all $\mathscr{H}$. We fix $\beta_1 \in (\beta_0, \beta_{\partial\mathscr{V}})$ and let $\mathscr{B} : [\beta_0, \beta_1] \to [0,1]$ be a function that is equal to 0 in a neighbourhood of β_0 and equal to 1 in a neighbourhood of β_1. We set

$$\check{\mathscr{X}}_k \;:=\; (\mathscr{B} \circ \mathscr{S}_k) \cdot \mathscr{X}_k \;\in\; \Gamma(\mathscr{H})\,.$$

We observe that

$$\bullet\; \check{\mathscr{X}}_k \;=\; 0 \;\text{ on }\; \{\mathscr{S}_k < \beta_0\}\,, \qquad \bullet\; \check{\mathscr{X}}_k \;=\; \mathscr{X}_k \;\text{ on }\; \mathscr{H} \setminus \{\mathscr{S}_k < \beta_1\}\,,$$

and, hence, the flow $\Phi^{\check{\mathscr{X}}_k}$ is positively complete.

3.3. **Proof of Theorem 3.1.** Let us define $\mathscr{I}_*$ as the set

$$\Big\{ k \in [k_0, k_1) \Big| \, \exists\, C(k_*) \text{ such that } c_{\mathscr{G}}(k) - c_{\mathscr{G}}(k_*) \leq C(k_*)(k - k_*),\ \forall k \in [k_*, k_1] \Big\} .$$

Namely, $\mathscr{I}_*$ is the set of points at which $c_{\mathscr{G}}$ is Lipschitz-continuous on the right. Since $c_{\mathscr{G}}$ is a non-decreasing real function, by Lebesgue Differentiation Theorem, $c_{\mathscr{G}}$ is Lipschitz-continuous at almost every point. In particular, $\mathscr{I} \setminus \mathscr{I}_*$ has measure zero.

We are now ready to show that

(1) for all $k \in \mathscr{I}$, there exists a vanishing sequence $(h_m^k)_{m\in\mathbb{N}} \subset \mathscr{H} \setminus \{\mathscr{S}_k < \beta_0\}$ and that
(2) for all $k_* \in \mathscr{I}_*$, such vanishing sequence can be taken to satisfy

$$\sup_{m\in\mathbb{N}} \mathscr{T}(h_m^{k_*}) \;<\; C(k_*) \;+\; 3\,.$$

We will prove only the statement about the vanishing sequences with parameter in $\mathscr{I}_*$, as the argument can be easily adapted to prove the statement for a general parameter in $\mathscr{I}$.

We assume by contradiction that there exists a positive number ε_0 such that

$$|\alpha_{k_*}| \;\geq\; \varepsilon_0\,, \quad \text{on } \{\mathscr{T} \;<\; C(k_*) + 3\} \setminus \{\mathscr{S}_{k_*} \;<\; \beta_1\}\,. \tag{44}$$

Consider a decreasing sequence $(k_m)_{m\in\mathbb{N}} \subset (k_*, k_1]$ such that $k_m \to k_*$. Set $\delta_m := k_m - k_*$ and take a corresponding sequence $(\gamma_m)_{m\in\mathbb{N}} \subset \mathscr{G}$ such that

$$\sup_{B^d} \mathscr{S}_{k_m}^{\gamma_m} \;<\; c_{\mathscr{G}}(k_m) \;+\; \delta_m\,.$$

For every $\xi \in B^d$ we consider the sequence of flow lines

$$\begin{aligned} u_m^\xi : [0,1] &\longrightarrow \mathscr{H} \\ r &\longmapsto \Phi_r^{\check{\mathscr{X}}_{k_*}}(\gamma_m(\xi))\,. \end{aligned}$$

Conversely, for any time parameter $r \in [0,1]$, we get the map

$$\gamma_m^r \;:=\; \Phi_r^{\check{\mathscr{X}}_{k_*}}(\gamma_m)\,. \tag{45}$$

We readily see that $\gamma_m^r|_{\partial B^d} = \gamma_m|_{\partial B^d}$ and $\gamma_m^r \in \mathscr{G}$. In particular, for every $\xi \in B^d$ and $r \in [0,1]$ the concatenated curve

$$\big(\gamma_m \circ z_\xi\big) \;*\; u_m^\xi|_{[0,r]} \;*\; \big(\overline{\gamma_m^r \circ z_\xi}\big) \tag{46}$$

is contractible. Therefore, Lemma 3.3 and Equation (39) yield

$$\mathscr{S}_{k_*}^{\gamma_m^r}(\xi) \;=\; \mathscr{S}_{k_*}^{\gamma_m}(\xi) \;+\; \alpha_{k_*}(u_m^\xi|_{[0,r]})\,. \tag{47}$$

Finally, since u_m^ξ is a flow line, we have

$$\begin{aligned} \alpha_{k_*}(u_m^\xi|_{[0,r]}) \;&=\; \int_0^r \alpha_{k_*}\left(-\frac{\mathscr{B}\cdot\sharp\alpha_{k_*}}{\sqrt{1+|\alpha_{k_*}|^2}}\right)(u_m^\xi(\rho))\,\mathrm{d}\rho \\ &=\; -\int_0^r \frac{\mathscr{B}\cdot|\alpha_{k_*}|^2}{\sqrt{1+|\alpha_{k_*}|^2}}(u_m^\xi(\rho))\,\mathrm{d}\rho\,. \end{aligned} \tag{48}$$

Therefore $\alpha_{k_*}(u_m^\xi|_{[0,r]}) \leq 0$ and we find that, for every $m \in \mathbb{N}$,

$$r \longmapsto \mathscr{S}_{k_*}^{\gamma_m^r} \quad \text{is a non-increasing family of functions on } B^d\,. \tag{49}$$

Let us estimate the supremum of $\mathscr{S}_{k_*}^{\gamma_m^r}$. When $r = 0$, (43) and the definition of $\mathscr{I}_*$ imply:

$$\sup_{B^d} \mathscr{S}_{k_*}^{\gamma_m} \leq \sup_{B^d} \mathscr{S}_{k_m}^{\gamma_m} < c_{\mathscr{G}}(k_m) + \delta_m \leq c_{\mathscr{G}}(k_*) + (C(k_*) + 1)\,\delta_m \,. \tag{50}$$

Thus, by (49) we get, for every $r \in [0, 1]$,

$$\sup_{B^d} \mathscr{S}_{k_*}^{\gamma_m^r} < c_{\mathscr{G}}(k_*) + (C(k_*) + 1)\,\delta_m \,. \tag{51}$$

If $r \in [0, 1]$, we define the sequence of subsets of B^d

$$J_m^r := \left\{ \mathscr{S}_{k_*}^{\gamma_m^r} > c_{\mathscr{G}}(k_*) - \delta_m \right\}.$$

Let us give a closer look to these sets. First, we observe that if $\xi \in J_m^r$, then (47) and (51) imply that

$$\alpha_{k_*}(u_m^\xi|_{[0,r]}) > c_{\mathscr{G}}(k_*) - \delta_m - \big(c_{\mathscr{G}}(k_*) + (C(k_*) + 1)\,\delta_m \big) = -(C(k_*) + 2)\,\delta_m \,. \tag{52}$$

Then, we claim that for m large enough

$$\xi \in J_m^r \implies \gamma_m^r(\xi) \in \big\{\mathscr{T} < C(k_*) + 3\big\} \setminus \big\{\mathscr{S}_{k_*} < \beta_1\big\}, \quad \forall\, r \in [0, 1]\,. \tag{53}$$

First, we observe that

$$\mathscr{S}_{k_*}^{\gamma_m^r}(\xi) > c_{\mathscr{G}}(k_*) - \delta_m \geq \sup_{B^d} \mathscr{S}_{k_*}^{\gamma_m^r} - (C(k_*) + 2)\,\delta_m \,. \tag{54}$$

If m is large enough, then $(C(k_*)+2)\,\delta_m < (\beta_{\partial\mathscr{V}} - \beta_1)$ and Lemma 3.6 implies that $\gamma_m^r(\xi) \notin \{\mathscr{S}_{k_*} < \beta_1\}$. As a by-product we get that $u_m^\xi|_{[0,r]}$ is a genuine flow line of $\Phi^{\mathscr{X}_{k_*}}$. Then, we estimate $\mathscr{T}(\gamma_m^r(\xi))$. We start by taking $r = 0$. In this case from (43) we get

$$\mathscr{T}(\gamma_m(\xi)) = \frac{\mathscr{S}_{k_m}^{\gamma_m}(\xi) - \mathscr{S}_{k_*}^{\gamma_m}(\xi)}{\delta_m} < \frac{c_{\mathscr{G}}(k_m) + \delta_m - c_{\mathscr{G}}(k_*) + \delta_m}{\delta_m} < C(k_*) + 2\,.$$

To prove the inequality for arbitrary r we bound the variation of $\mathscr{T}$ along $u_m^\xi|_{[0,r]}$ in terms of the action variation:

$$\begin{aligned}
-\alpha_{k_*}(u_m^\xi|_{[0,r]}) &= -\int_0^r \alpha_{k_*}\left(\frac{du_m^\xi}{d\rho}\right) \mathrm{d}\rho \geq \int_0^r \left|\frac{du_m^\xi}{d\rho}\right|^2 \mathrm{d}\rho \\
&\geq \frac{1}{r}\left(\int_0^r \left|\frac{du_m^\xi}{d\rho}\right| \mathrm{d}\rho\right)^2 \\
&\geq \frac{1}{r}\left(\int_0^r \frac{1}{1 + \sup_{\mathscr{H}} |d\mathscr{T}|} \left|\frac{d(\mathscr{T} \circ u_m^\xi)}{d\rho}\right| \mathrm{d}\rho\right)^2 \\
&\geq \frac{1}{r(1 + \sup_{\mathscr{H}} |d\mathscr{T}|)^2} |\mathscr{T}(u_m^\xi(r)) - \mathscr{T}(u_m^\xi(0))|^2 \,.
\end{aligned}$$

Using (52) and rearranging the terms we get for m large enough

$$|\mathscr{T}(\gamma_m^r(\xi)) - \mathscr{T}(\gamma_m(\xi))|^2 \leq r \cdot (1 + \sup_{\mathscr{H}} |d\mathscr{T}|)^2 \cdot (C(k_*) + 2)\,\delta_m < 1\,.$$

Hence, if m is large enough the bound on $\mathscr{T}$ we were looking for follows from

$$\mathscr{T}(\gamma_m^r(\xi)) \leq \mathscr{T}(\gamma_m(\xi)) + |\mathscr{T}(\gamma_m^r(\xi)) - \mathscr{T}(\gamma_m(\xi))| < (C(k_*) + 2) + 1\,. \tag{55}$$

The claim is thus completely established.

The last step to finish the proof of Theorem 3.1 is to show that $J^1_m = \emptyset$ for m large enough. By contradiction, let $\xi \in J^1_m$. Since $\xi \in J^r_m$ for all $r \in [0,1]$, we see that u^ξ_m is a flow line of $\Phi^{\mathscr{X}_{k_*}}$ contained in $\{\mathscr{T} < C(k_*)+3\} \setminus \{\mathscr{S}_{k_*} < \beta_1\}$. Using (52) and continuing the chain of inequalities in (48), we find

$$-\left(C(k_*)+2\right)\delta_m \;<\; \alpha_{k_*}(u^\xi_m) \;\leq\; -\frac{\varepsilon_0^2}{\sqrt{1+\varepsilon_0^2}}$$

(where we used that the real function $w \mapsto \frac{w}{\sqrt{1+w}}$ is increasing). Such inequality cannot be satisfied for m large, proving that the sets J^1_m become eventually empty.

Finally, since $J^1_m = \emptyset$, we obtain that $c_{\mathscr{G}}(k_*) \leq \sup_{B^d} \mathscr{S}^{\gamma^1_m}_{k_*} \leq c_{\mathscr{G}}(k_*) - \delta_m$. This contradiction finishes the proof of Theorem 3.1.

In the next section we will determine when η_k satisfies the hypotheses of the abstract theorem we have just proved.

4. Proof of the Main Theorem

We now move to the proof of points *(1)*, *(2)*, *(3)* of Theorem 1.6. In the first preparatory subsection, we will see when the action form is exact.

4.1. **Primitives for η_k.** We know that η_k is exact if and only if so is τ^σ. The next proposition, whose simple proof we omit, gives necessary and sufficient conditions for the transgression form to be exact.

Proposition 4.1. *If $[\widetilde{\sigma}] \neq 0$, then $\tau^\sigma|_{W^{1,2}_\nu}$ is not exact for any ν.*

If $[\widetilde{\sigma}] = 0$, then

$$\widehat{P}^\sigma : W^{1,2}_0 \longrightarrow \mathbb{R}\,. \qquad x \longmapsto \int_{B^2} \hat{u}_x^* \sigma$$

is a primitive for τ^σ. Here $\hat{u}_x$ is any capping disc for x. This definition extends the primitive P^σ, which we constructed on the subset of short loops.

If $[\widetilde{\sigma}]_b = 0$, then, given ν and a reference loop $x_\nu \in W^{1,2}_\nu$,

$$\widehat{P}^\sigma : W^{1,2}_\nu \longrightarrow \mathbb{R}\,. \qquad x \longmapsto \int_{B^2} \hat{u}_{x_\nu,x}^* \sigma$$

is a primitive for τ^σ. Here $\hat{u}_{x_\nu,x}$ is a connecting cylinder from x_ν to x. If we take x_0 as a constant loop, the two definitions of $\widehat{P}^\sigma$ coincide on $W^{1,2}_0$.

Exercise 6. *Show that if $M = \mathbb{T}^2$ and $[\sigma] \neq 0$, then $\tau^\sigma|_{W^{1,2}_\nu}$ is not exact if $\nu \neq 0$.*

We set $\widehat{S}_k := A_k - \widehat{P}^\sigma \circ \mathrm{pr}_{W^{1,2}}$ in the two cases above where $\widehat{P}^\sigma$ is defined. Theorem A in [CIPP98] tells us when $\widehat{S}_k$ is bounded from below.

Proposition 4.2. *If $[\widetilde{\sigma}] = 0$, then $\widehat{S}_k : \Lambda_0 \to \mathbb{R}$ is bounded from below if and only if $k \geq c(L,\sigma)$. If $[\widetilde{\sigma}]_b = 0$, the same is true for $\widehat{S}_k : \Lambda_\nu \to \mathbb{R}$.*

Remark 4.3. *Originally the critical value was introduced by Mañé as the infimum of the values of k such that $\widehat{S}_k : \Lambda_0 \to \mathbb{R}$ is bounded from below* [Mañ97, CDI97]. *Thus, the proposition above establishes the equivalence between the more geometric definition in* (4) *and the original one.*

Exercise 7. *Prove that $\widehat{S}_k|_{\Lambda_\nu}$ is bounded from below if and only if $\widehat{S}_k|_{\Lambda_0}$ is bounded from below if and only if $\widehat{S}_k|_{\Lambda_0}$ is non-negative.*

As a by-product of Proposition 4.2, we can give a criterion guaranteeing that a vanishing sequence for η_k has bounded periods, provided $k > c(L,\sigma)$.

Corollary 4.4. *Let $\nu \in [\mathbb{T}, M]$ and $[\widetilde{\sigma}]_b = 0$. If $k > c(L,\sigma)$ and $b \in \mathbb{R}$, then there exists a constant $C(\nu,k,b)$ such that*

$$\forall\, \gamma \in \Lambda_\nu\,, \quad \widehat{S}_k(\gamma) \;<\; b \quad \Longrightarrow \quad T \;<\; C(\nu,k,b)\,.$$

Proof. We readily compute

$$T \;=\; \frac{\widehat{S}_k(\gamma) - \widehat{S}_{c(L,\sigma)}(\gamma)}{k - c(L,\sigma)} \;\leq\; \frac{b - \inf_{\Lambda_\nu} \widehat{S}_{c(L,\sigma)}}{k - c(L,\sigma)} \;=:\; C(\nu,k,b)\,. \qquad \square$$

4.2. **Non-contractible orbits.** We now prove the existence of non-contractible orbits as prescribed by the Main Theorem.

Proof of Theorem 1.6.(1*a*). Let $\nu \in [\mathbb{T}, M]$ be a non-trivial class, σ be a magnetic form such that $[\sigma]_b = 0$ and $k > c(L,\sigma)$. Thanks to Proposition 4.2, the infimum of $\widehat{S}_k$ on Λ_ν is finite. Then, we apply Theorem 3.2 with $\mathscr{H} = \Lambda_\nu$ and $\widehat{\mathscr{S}} = \widehat{S}_k$ and we obtain a vanishing sequence $(\gamma_m)_{m\in\mathbb{N}}$ such that $\widehat{S}_k(\gamma_m)$ is uniformly bounded. By Corollary 4.4 the sequence of periods is bounded from above. By Corollary 2.12 the sequence of periods is also bounded away from zero. Therefore, we can apply Proposition 2.8 to get a limit point of the sequence. $\square$

4.3. **Contractible orbits.** We start by recalling a topological lemma.

Proposition 4.5. *If $d \geq 1$ and $\delta \leq \delta_*$ (see Lemma 2.10), there are natural bijections*

$$\begin{tikzcd}
\frac{\pi_{d+1}(M)}{\pi_1(M)} \arrow[r] \arrow[dr, "F"] & {[\, S^{d+1},\, M\,]} \arrow[d] & \\
& {\big[\,\big(B^d,\partial B^d\big)\,,\,\big(W^{1,2}_0, M_0\big)\,\big]} \arrow[r, "i^{M_0}_{\{\ell<\delta\}}"] & {\big[\,\big(B^d,\partial B^d\big)\,,\,\big(W^{1,2}_0, \{\ell<\delta\}\big)\,\big]}
\end{tikzcd}$$

where $\pi_{d+1}(M)/\pi_1(M)$ is the quotient of $\pi_{d+1}(M)$ by the action of $\pi_1(M)$[1]*. The trivial classes on the second line are identified with the class of constant maps in $[S^{d+1}, M]$ and with the class of the zero element in $\pi_{d+1}(M)/\pi_1(M)$.*

Proof. The first horizontal map is $\frac{[\hat{u}]}{\pi_1(M)} \mapsto [\hat{u}]$. We leave as an exercise to the reader to show that is a bijection. The vertical map sends $[\hat{u}]$ to $[u]$, where u is defined as follows. Consider the equivalence relation $\sim$ on $B^d \times \mathbb{T}$:

$$(56) \quad (z_1,s_1) \sim (z_2,s_2) \qquad \Longleftrightarrow \qquad (z_1,s_1) \;=\; (z_2,s_2) \quad \vee \quad z_1 \;=\; z_2 \;\in\; \partial B^d\,.$$

[1]Here a choice of an arbitrary base point $q_0 \in M$ is to be understood: $\pi_{d+1}(M) := \pi_{d+1}(M, q_0)$ and $\pi_1(M) := \pi_1(M, q_0)$

If we interpret B^d as the unit ball in $\mathbb{R}^d$ and S^{d+1} as the unit sphere in $\mathbb{R}^{d+2}$ we can define the homeomorphism

$$Q : \frac{B^d \times \mathbb{T}}{\sim} \longrightarrow S^{d+1}$$
$$[z,s] \longmapsto (z, \sqrt{1-|z|^2}\cdot e^{2\pi i s}),$$

where $e^{2\pi i s}$ belongs to $S^1 \subset \mathbb{R}^2$. We set $u(z)(s) := (\hat{u}\circ Q)([z,s])$. For a proof that the vertical map is well-defined and it is a bijection, we refer the reader to [Kli78, Proposition 2.1.7]. Finally, the second horizontal map is a bijection thanks to Lemma 2.10. □

We can now prove the parts of the Main Theorem dealing with contractible orbits.

Proof of Theorem 1.6.(1b). Let $[\widetilde{\sigma}]_b = 0$, $k > c(L,\sigma)$ and fix some non-zero $\mathfrak{u} \in \pi_{d+1}(M)$, which exists by hypothesis. We apply Proposition 2.13 to the trivial interval $\{k\}$ and get the positive real numbers $\delta_{\{k\}}$, $b_{\{k\}}$ and $T_{\{k\}}$. Let

(57)
$$\Gamma_{\mathfrak{u}} := \Big\{ \gamma = (x,T) : (B^d, \partial B^d) \longrightarrow (\Lambda_0, M_0 \times \{T_{\{k\}}\}) \ \Big|\ [x] \in F(\mathfrak{u}/\pi_1(M)) \Big\}$$

By Proposition 4.5 we see that $\Gamma_{\mathfrak{u}} \in [(B^d,\partial B^d),(\Lambda_0, M_0\times\{T_{\{k\}}\})]$ and that $i^{M_0\times\{T_{\{k\}}\}}_{\mathscr{V}^{\delta_{\{k\}}}}(\Gamma_{\mathfrak{u}})$ is non-trivial. Therefore, we apply Theorem 3.1 with

$$\begin{bmatrix} \mathscr{H} = \Lambda_0 & \mathscr{I} = \{k\} & \widehat{\mathscr{S}_k} = \widehat{S}_k \\ \beta_0 = b_{\{k\}}/2 & \mathscr{V} = \mathcal{V}^{\delta_{\{k\}}} & \mathscr{M} = M_0 \times \{T_{\{k\}}\} \\ \mathscr{G} = \Gamma_{\mathfrak{u}} & & \end{bmatrix}$$

and we obtain a vanishing sequence $(\gamma_m)_{m\in\mathbb{N}}$ such that

$$\lim_{m\to+\infty} \widehat{S}_k(\gamma_m) \ = \ c_{\mathfrak{u}}(k) \ := \ \inf_{\gamma\in\Gamma} \sup_{B^d} \widehat{S}_k \circ \gamma \ \geq \ b_{\{k\}}\,.$$

The sequence of periods (T_m) is bounded from above by Corollary 4.4. The sequence (T_m) is also bounded away from zero by Corollary 2.12, since $\gamma_m \notin \{S_k < b_{\{k\}}/2\}$ for m big enough. Applying Proposition 2.8 we obtain a limit point of (γ_m). □

Proof of Theorem 1.6.(2). Let $[\widetilde{\sigma}] = 0$ and fix $I = [k_0,k_1] \subset (e_0(L), c(L,\sigma))$. Let δ_I, b_I and T_I be as in Proposition 2.13. Fix $\gamma_0 \in M_0 \times \{T_I\}$ and $\gamma_1 \in \Lambda_0$ such that $\widehat{S}_{k_1}(\gamma_1) < 0$. Such element exists thanks to Proposition 4.2. Let $u_* : [0,1] \to \Lambda$ be some path such that $u_*(0) = \gamma_0$ and $u_*(1) = \gamma(1)$ and denote by $[u_*] \in [(B^1,\partial B^1),(\Lambda_0,\{\gamma_0,\gamma_1\})]$ its homotopy class. By Proposition 2.13, γ_0 and γ_1 belong to different components of $\{\widehat{S}_{k_0} < b_I\}$. Thus, $i^{\{\gamma_0,\gamma_1\}}_{\{\widehat{S}_{k_0}<b_I\}}([u_*])$ is non-trivial. Therefore, we apply Theorem 3.1 with

$$\begin{bmatrix} \mathscr{H} = \Lambda_0 & \mathscr{I} = I & \widehat{\mathscr{S}_k} = \widehat{S}_k \\ \beta_0 = b_I/2 & \mathscr{V} = \{\widehat{S}_{k_0} < b_I\} & \mathscr{M} = \{\gamma_0,\gamma_1\} \\ \mathscr{G} = [u_*] & & \end{bmatrix}$$

and we get a vanishing sequence $(\gamma^k_m)_{m\in\mathbb{N}}$ with bounded periods, for almost every $k\in I$. Moreover, we have

$$\lim_{m\to+\infty}\widehat{S}_k(\gamma_m) \;=\; c_{[u_*]}(k) \;:=\; \inf_{u\in[u_*]}\,\sup_{B^1}\widehat{S}_k\circ u \;\geq\; b_I\,.$$

In particular, $\gamma^k_m\notin\{\widehat{S}_k<b_I/2\}$ for m large enough. Hence, the periods are bounded away from zero by Corollary 2.12. Now we apply Proposition 2.8 to get a limit point of (γ^k_m). Taking an exhaustion of $(e_0(L),c(L,\sigma))$ by compact intervals, we get a critical point for almost every energy in $(e_0(L),c(L,\sigma))$. $\square$

Proof of Theorem 1.6.(3). Let $[\widetilde{\sigma}]\neq 0$ and fix $I=[k_0,k_1]\subset(e_0(L),+\infty)$. Let δ_I, b_I and T_I be as in Proposition 2.13. Since $[\widetilde{\sigma}]\neq 0$, there exists a non-zero $\mathfrak{u}\in\pi_2(M)$. We set

$$\text{(58)}\quad \Gamma_{\mathfrak{u}} := \Big\{\,\gamma=(x,T):\big(B^1,\partial B^1\big)\longrightarrow\big(\Lambda_0,M_0\times\{T_I\}\big)\;\Big|\;[x]\in F\big(\mathfrak{u}/\pi_1(M)\big)\,\Big\}$$

By Proposition 4.5 we see that $\Gamma_{\mathfrak{u}}\in\big[(B^1,\partial B^1),(\Lambda_0,M_0\times\{T_I\})\big]$ and that

$$i^{M_0\times\{T_I\}}_{\mathcal{V}^{\delta_I}}(\Gamma_{\mathfrak{u}})$$

is non-trivial. Therefore, we apply Theorem 3.1 with

$$\begin{bmatrix} \mathscr{H}=\Lambda_0 & \mathscr{I}=I & \alpha_k=\eta_k \\ \beta_0=b_I/2 & \mathscr{V}=\mathcal{V}^{\delta_I} & \mathscr{M}=M_0\times\{T_I\} \\ \mathscr{G}=\Gamma_{\mathfrak{u}} & & \end{bmatrix}$$

and we obtain a vanishing sequence $(\gamma^k_m)_{m\in\mathbb{N}}\subset\Lambda_0\setminus\{S_k<b_I/2\}$ with bounded periods, for almost every $k\in I$. Since, the periods are bounded away from zero by Corollary 2.12, Proposition 2.8 yields a limit point of (γ^k_m), for almost every $k\in I$.

Taking an exhaustion of $(e_0(L),+\infty)$ by compact intervals, we get a contractible zero of η_k for almost every $k>e_0(L)$. $\square$

5. Magnetic flows on surfaces I: Taĭmanov minimizers

In this and in the next section we are going to focus on the 2-dimensional case. Therefore, let us assume that M is a closed connected oriented surface. In this case $H^2(M;\mathbb{R})\simeq\mathbb{R}$, where the isomorphism is given by integration and we identify $[\sigma]$ with a real number. Up to changing the orientation on M, we assume that $[\sigma]\geq 0$.

For simplicity, we are going to work in the setting of Section 1.6 and consider only purely kinetic Lagrangians. Namely, we take $L(q,v)=\frac{1}{2}|v|^2$, where $|\cdot|$ is induced by a metric g.

Since L depends only on g, we will use the notation (g,σ) where we previously used (L,σ). We readily see that $e_m(L)=e_0(L)=0$ and that $c(g,\sigma)=0$ if and only if $\sigma=0$ (see Proposition 1.4). We recall that the periodic orbits with positive energy are parametrized by a positive multiple of the arc-length. Thus, they are immersed curve in M.

5.1. **The space of embedded curves.** The space of curves on a 2-dimensional manifold M has a particularly rich geometric structure. Observe, indeed, that for $n \geq 3$ the curves on M are generically embedded. On the other hand, if M is a surface, intersections between curves and self-intersections are generically stable. Therefore, one can refine the existence problem by looking at periodic orbits having a particular shape (see the beginning of Section 1.1 in [HS13] and references therein for a precise notion of the shape of a curve on a surface). For example, we consider the following question.

For which k and ν there exists a *simple* periodic orbit $\gamma \in \Lambda_\nu$ with energy $k > 0$?

Let us start by investigating the case $\nu = 0$. If $\gamma = (x, T)$ is a contractible simple curve, there exists an embedded disc $\hat{u} : B^2 \to M$ such that $\hat{u}(e^{2\pi i s}) = x(s)$. This map yields a path (u, T) in Λ_0 from a constant path (x_0, T), representing the centre of the disc, to (x, T). Integrating η_k along this path and summing the value of S_k at (x_0, T), we get

$$\int_0^1 (u,T)^*\eta_k \;+\; S_k(x_0,T) \;=\; \frac{e(x)}{2T} \;+\; kT \;-\; \int_{B^2} \hat{u}^*\sigma\,. \tag{59}$$

Since $\hat{u}$ is an embedding, $\mathrm{area}(\hat{u}) \leq \mathrm{area}(M)$ and we find a uniform bound from below

$$\int_0^1 (u,T)^*\eta_k + S_k(x_0,T) \;\geq\; 0 + 0 - \sup_M |\sigma|\cdot\mathrm{area}(\hat{u}) \;\geq\; -\sup_M |\sigma|\cdot\mathrm{area}(M)\,. \tag{60}$$

This observation gives us the idea of defining a functional on the space of simple contractible loops and look for its global minima. First, we notice that $\int_{B^2} \hat{u}^*\sigma$ is invariant under an orientation-preserving change of parametrization. In order to make the whole right-hand side of (59) independent of the parametrization, we ask that $(\gamma, \dot{\gamma}) \in \Sigma_k$. This implies that

$$\sqrt{2k}\cdot T \;=\; \ell(x)\,, \qquad e(x) \;=\; \ell(x)^2\,.$$

Substituting in (59), we get

$$\int_0^1 (u,T)^*\eta_k \;+\; S_k(x_0,T) \;=\; \sqrt{2k}\cdot\ell(\partial D) \;-\; \int_D \sigma \;=:\; \mathcal{T}_k(D)\,, \tag{61}$$

where

$$D = [\hat{u}] \in \mathcal{D}(M) := \left\{ \begin{array}{c} \text{embeddings } \hat{u} : B^2 \longrightarrow M\,, \\ \text{up to orientation-preserving reparametrizations} \end{array} \right\}$$

and ∂D represents the boundary of D oriented in the counter-clockwise sense. We readily see that the critical points of this functional correspond to the periodic orbits we are looking for.

Proposition 5.1. *If D is a critical point of $\mathcal{T}_k : \mathcal{D}(M) \to \mathbb{R}$, then ∂D is the support of a simple contractible periodic orbit with energy k.*

In view of this proposition and the fact that $\mathcal{T}_k$ is bounded from below, we consider a minimizing sequence $(D_m)_{m\in\mathbb{N}} \subset \mathcal{D}(M)$. However, the sequence D_m might converge to a disc D_∞ which is not embedded. For example, D_∞ might have a self-tangency at some point q on its boundary (see Figure 1). However, in this case the support of D_∞ in M can be interpreted as an annulus A_∞ whose two boundary components touch exactly at q. Now we can resolve the singularity in the

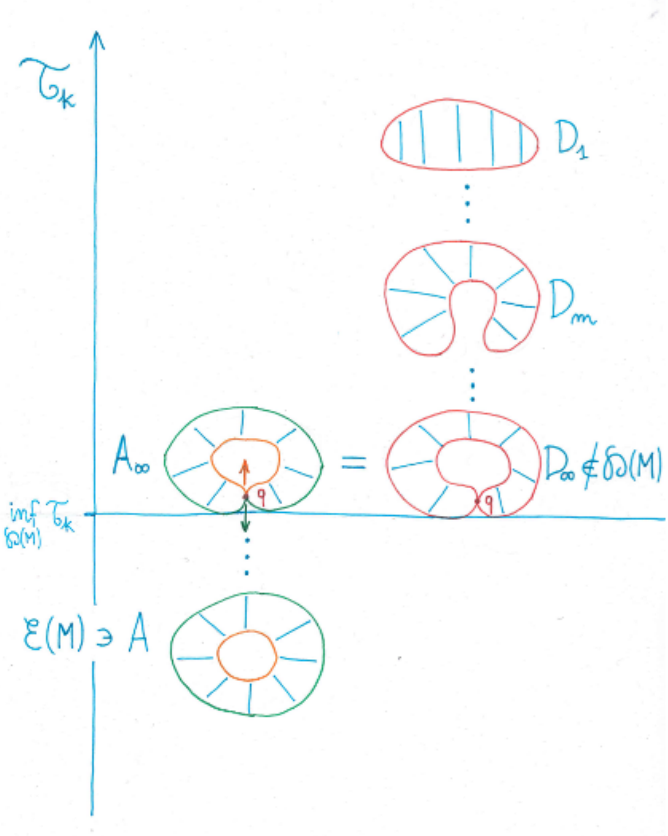


FIGURE 1. Minimizing sequence for $\mathcal{T}_k$ on $\mathcal{D}(M)$

space of annuli and get an embedded annulus A close to A_∞. The key observation is that $\mathcal{T}_k$ can be extended to the space of annuli and that

$$\mathcal{T}_k(D_\infty) \;=\; \mathcal{T}_k(A_\infty) \;>\; \mathcal{T}_k(A)\,. \tag{62}$$

To justify the inequality in the passage above, we observe that $\ell(\partial A) < \ell(\partial A_\infty)$ from classic estimates in Riemannian geometry and that the contribution given by the integral of σ is of higher order. This heuristic argument prompts us to give the following definitions.

Definition 5.2. *Let $\mathcal{E}(M) = \{$oriented embedded surfaces $\Pi \to M\} \cup \{\emptyset\}$ and denote by $\mathcal{E}_+(M)$ and $\mathcal{E}_-(M)$ the surfaces having the same orientation as M and the opposite orientation, respectively. If $\Pi \in \mathcal{E}(M)$, then $\partial\Pi$ denotes the (possibly empty) multi-curve made by the boundary components of Π. If we define the length $\ell(\partial\Pi)$ as the sum of the lengths of the boundary components, we have a natural extension*

$$\begin{aligned} \mathcal{T}_k : \mathcal{E}(M) \;&\longrightarrow\; \mathbb{R} \\ \Pi \;&\longmapsto\; \sqrt{2k}\cdot\ell(\partial\Pi) \;-\; \int_\Pi \sigma\,. \end{aligned}$$

As in (60) we find that $\mathcal{T}_k$ is bounded from below by $-\sup|\sigma|\cdot\mathrm{area}(M)$. Moreover, we observe that there is a bijection

$$\begin{array}{rcl} \mathcal{E}_+(M) & \longrightarrow & \mathcal{E}_-(M) \\ \Pi & \longmapsto & M\setminus\mathring{\Pi} \end{array} \qquad \text{such that} \qquad \mathcal{T}_k(M\setminus\mathring{\Pi}) \;=\; \mathcal{T}_k(\Pi) \;+\; \int_M \sigma\,. \tag{63}$$

Therefore, it is enough to look for a minimizer on $\mathcal{E}_-(M)$. The chain of inequalities (62) hints at the following result.

Proposition 5.3. *For all $k>0$, there exists a minimizer Π^k of $\mathcal{T}_k|_{\mathcal{E}_-(M)}$. If $\partial\Pi^k=\{\gamma_i^k\}_i$, then the γ_i^k are periodic orbits with energy k.*

For a rigorous proof of this proposition we refer to [Taĭ93] and [CMP04]:

- In the former reference, Taĭmanov uses a finite dimensional reduction and works on the space of surfaces $\Pi\in\mathcal{E}(M)$ whose boundary is made by piecewise solutions of the twisted Euler-Lagrange equations with energy k.
- In the latter reference, the authors use a weak formulation of the problem on the space of integral currents $I_2(M)\supset\mathcal{E}(M)$.

In order to use Proposition 5.3 to prove the existence of periodic orbits with energy k, we have to ensure that $\partial\Pi^k\neq\emptyset$. To this purpose, we observe that $\partial\Pi^k=\emptyset$ implies $\Pi^k\in\{\emptyset,\overline{M}\}$, where $\overline{M}$ is M with the opposite orientation. We easily compute $\mathcal{T}_k(\emptyset)=0$ and $\mathcal{T}_k(\overline{M})=\int_M\sigma\geq 0$. Therefore, for every $k>0$ we have

$$\inf_{\mathcal{E}_-(M)}\mathcal{T}_k\;\leq\;0 \qquad\text{and}\qquad \Big(\inf_{\mathcal{E}_-(M)}\mathcal{T}_k\;<\;0 \quad\Longrightarrow\quad \partial\Pi^k\;\neq\;\emptyset\Big).$$

Since the family of functionals $\mathcal{T}_k$ is monotone in k, we are led to define

$$\tau(g,\sigma)\;:=\;\inf\Big\{k\;\Big|\;\inf_{\mathcal{E}_-(M)}\mathcal{T}_k\;=\;0\Big\}. \tag{64}$$

Proposition 5.4. *The value $\tau(g,\sigma)$ is a non-negative real number. Moreover,*

$$\tau(g,\sigma)\;>\;0\qquad\Longleftrightarrow\qquad\sigma_{q_0}\;<\;0\,,\ \textit{for some } q_0\in M\,.$$

If σ is exact, then

$$\tau(g,\sigma)\;=\;c_0(g,\sigma)\;:=\;\inf_{d\theta=\sigma}\sup_{q\in M}|\theta_q|\,. \tag{65}$$

We leave the proof of the first statement of the proposition as an exercise to the reader. The second statement follows from [CMP04]. We can summarize our answer to the question raised at the beginning of this section with the following theorem.

Theorem 5.5. *Suppose that there exists $q_0\in M$ such that $\sigma_{q_0}<0$. Then, we can find a positive real number $\tau(g,\sigma)$, coinciding with $c_0(g,\sigma)$ when σ is exact, such that for every $k\in(0,\tau(g,\sigma))$, there exists a non-empty set of simple periodic orbits $\{\gamma_i^k\}$ having energy k and satisfying*

$$\sum_i\,[\gamma_i^k]\;=\;0\,\in\,H^1(M;\mathbb{Z})\,.$$

6. Magnetic flows on surfaces II: stable energy levels

In this last section we continue the study of twisted Lagrangian flows of kinetic type on surfaces by investigating the stability properties of their energy levels. To have a better geometric intuition, we are going to pull-back the twisted symplectic form to the tangent bundle. Thus, let $\flat : TM \to T^*M$ be the duality isomorphism given by g. We define the twisted tangent bundle as the symplectic manifold $(TM, \omega_{g,\sigma})$, where $\omega_{g,\sigma} := d(\flat^*\lambda) - \pi^*\sigma$. We readily see that $X_{(g,\sigma)}$ is the Hamiltonian flow of E with respect to the symplectic form $\omega_{g,\sigma}$. In this language, our problem is to understand when the hypersurface Σ_k is stable in the twisted tangent bundle. We will summarize the current knowledge on the subject in the following four propositions.

The first one sheds light on the relation between stability and the contact property in the generic case.

Proposition 6.1. *Let $k > 0$. If $[\sigma] \neq 0$ and $M = \mathbb{T}^2$, Σ_k is not of contact type. Moreover, if $X_{(g,\sigma)}|_{\Sigma_k}$ does not admit any non-trivial integral of motion, then:*

(1) *If $[\sigma] = 0$ or $M \neq \mathbb{T}^2$ and $[\sigma] \neq 0$, Σ_k is stable if and only if it is of contact type.*
(2) *If $M = \mathbb{T}^2$ and $[\sigma] \neq 0$, every stabilizing form on Σ_k is closed and it has non-vanishing integral over the fibers of π.*

The second proposition gives obstruction to the contact property.

Proposition 6.2. *The following statements hold true.*

(1) *If $[\sigma] = 0$, then Σ_k is not of negative contact type.*
(2) *If $[\sigma] \neq 0$, then*
    (a) *if $M = S^2$, Σ_k is not of negative contact type;*
    (b) *if M has genus higher than 1, there exists $c_h(g,\sigma) > 0$ such that*
        - *Σ_k is not of negative contact type, when $k > c_h(g,\sigma)$;*
        - *$\Sigma_{c_h(g,\sigma)}$ is not of contact type;*
        - *Σ_k is not of positive contact type, when $k < c_h(g,\sigma)$;*

The third proposition deals with positive results on stability.

Proposition 6.3. *The following statements hold true.*

(1) *If $[\sigma] = 0$, Σ_k is of contact type if $k > c_0(g,\sigma)$. If $M = \mathbb{T}^2$, for every Riemannian metric g there exists an exact form σ_g for which $\Sigma_{c_0(g,\sigma_g)}$ is of contact type.*
(2) *If $[\sigma] \neq 0$ and $M \neq \mathbb{T}^2$, Σ_k is of contact type for k big enough.*
(3) *If σ is a symplectic form on M, then Σ_k is stable for k small enough.*

The last proposition deals with negative results on stability.

Proposition 6.4. *The following statements hold true.*

(1) *If $[\sigma] = 0$ and $M \neq \mathbb{T}^2$, Σ_k is not of contact type, for $k < c_0(g,\sigma)$;*
(2) *If $[\sigma] \neq 0$ and there exists $q \in M$ such that $\sigma_q < 0$, then*
    (a) *when $M \neq \mathbb{T}^2$, Σ_k is not of contact type, for k low enough;*
    (b) *when $M = \mathbb{T}^2$, Σ_k does not admit a closed stabilizing form, for k low enough.*
(3) *If $M = S^2$, there exists an energy level associated to some $\overline{g}$ and some everywhere positive form $\overline{\sigma}$, which is not of contact type.*

Before embarking in the proof of such propositions, we make the following observation.

Lemma 6.5. *Let $k > 0$ and set $s := 1/\sqrt{2k}$. Then, the flows of $\Phi^{(g,\sigma)}|_{\Sigma_k}$ and $\Phi^{(g,s\sigma)}|_{\Sigma_{1/2}}$ are conjugated up to a time reparametrization.*

Proof. By Section 1.6 we know that the projections to M of the trajectories of $\Phi^{(g,\sigma)}|_{\Sigma_k}$ and of $\Phi^{(g,s\sigma)}|_{\Sigma_{1/2}}$ both satisfy the equation $\kappa_\gamma = s \cdot f(\gamma)$. Therefore, if

$$t \;\longmapsto\; \Big(\gamma(t), \frac{d\gamma}{dt}(t)\Big)$$

is a trajectory of the former flow and we set $\gamma_s(t') := \gamma(st')$, then

$$t' \;\longmapsto\; \Big(\gamma_s(t'), \frac{d\gamma_s}{dt'}(t')\Big) \;=\; \Big(\gamma(st'),\, s\cdot\frac{d\gamma}{dt}(st')\Big)$$

is a trajectory of the latter flow. □

Therefore, given (g,σ), instead of studying the flow $\Phi^{(g,\sigma)}$ on each energy level Σ_k, we can study the 1-parameter family of flows $\Phi^{(g,s\sigma)}$ on $SM := \Sigma_{1/2}$ as s varies in $(0,+\infty)$. The advantage of rescaling σ is that now we can work on a fixed three-dimensional manifold: SM. The tangent bundle of SM has a global frame (X,V,H) and corresponding dual co-frame (α,ψ,β), which we now define.

Let $\mathcal{H} \subset SM$ be the horizontal distribution given by the Levi-Civita connection of g. For every $(q,v) \in SM$, $X_{(q,v)}$ and $H_{(q,v)}$ are defined as the unique elements in $\mathcal{H}$ such that

$$d_{(q,v)}\pi\big(X_{(q,v)}\big) \;=\; v\,, \qquad d_{(q,v)}\pi\big(H_{(q,v)}\big) \;=\; \imath\cdot v\,.$$

Analogously, $\alpha_{(q,v)}$ and $\beta_{(q,v)}$ are defined by

$$\alpha_{(q,v)}(\cdot) \;=\; g_q\big(v, d_{(q,v)}\pi(\cdot)\big)\,, \qquad \beta_{(q,v)}(\cdot) \;=\; g_q\big(\imath\cdot v, d_{(q,v)}\pi(\cdot)\big)\,.$$

The vector V is the generator of the rotations along the fibers $\varphi \mapsto (q, \cos\varphi\, v + \sin\varphi\, \imath\cdot v)$. The form ψ is the connection 1-form of the Levi-Civita connection. If $W \in T_{(q,v)}SM$ and $w(t) = (\gamma(t), v(t))$ is a curve such that $w(0) = (q,v)$ and $\dot w(0) = W$, then

$$\psi_{(q,v)}(W) \;=\; g_q\big(\nabla_{\dot\gamma(0)}v, \imath\cdot v\big).$$

Finally, we orient SM using the frame (X,V,H).

The proof of the following proposition giving the structural relations for the co-frame is a particular case of the identities proven in [GK02].

Proposition 6.6. *Let K be the Gaussian curvature of g. We have the relations:*

$$d\alpha \;=\; \psi\wedge\beta\,, \qquad d\psi \;=\; K\beta\wedge\alpha \;=\; -K\pi^*\mu\,, \qquad d\beta \;=\; \alpha\wedge\psi\,. \tag{66}$$

Using the frame (X,V,H) we can write

$$X_s \;:=\; X_{(g,s\sigma)} \;=\; X + sfV\,, \qquad \omega_s \;:=\; \omega_{g,s\sigma}|_{SM} \;=\; d\alpha - s\pi^*\sigma\,.$$

We also use the notation Φ^s for the flow of X_s on SM.

6.1. **Stability of the homogeneous systems.** Let us start by describing the stability properties of the homogeneous examples introduced in Section 1.6.

6.1.1. *The two-sphere.* In this case we have $\sigma = \mu = K\mu$. Hence,

$$\omega_s \;=\; d\alpha - s\pi^*\sigma \;=\; d(\alpha + s\psi) \quad \text{and} \quad (\alpha+s\psi)(X_s) \;=\; (\alpha+s\psi)(X+sV) \;=\; 1+s^2\,.$$

Every energy level is of positive contact type.

6.1.2. *The two-torus.* In this case we compute

$$d\psi \;=\; K\mu \;=\; 0 \quad \text{and} \quad \psi(X_s) \;=\; \psi(X+sV) \;=\; s\,.$$

Every energy level is stable.

6.1.3. *The hyperbolic surface.* In this case we have $\sigma = \mu = -K\mu$. Hence,

$$\omega_s \;=\; d\alpha - s\pi^*\sigma \;=\; d(\alpha - s\psi) \quad \text{and} \quad (\alpha-s\psi)(X_s) \;=\; (\alpha-s\psi)(X+sV) \;=\; 1-s^2\,.$$

Every energy level Σ_k with $k > \frac{1}{2}$ is of positive contact type. Every energy level Σ_k with $k < \frac{1}{2}$ is of negative contact type. As follows from Proposition 6.2, $c_h(g,\sigma) = 1/2$ and $\Sigma_{1/2}$ is not stable.

6.2. **Invariant measures on** SM**.** A fundamental ingredient in the proof of the four propositions is the notion of invariant measure for a flow. In this subsection, we recall this notion and we observe that twisted systems of purely kinetic type always possess a natural invariant measure called the *Liouville measure*.

Definition 6.7. *A Borel measure ξ on SM is Φ^s-*invariant*, if $\xi(\Phi^s_t(A)) = \xi(A)$, for every $t \in \mathbb{R}$ and every Borel set A. This is equivalent to asking*

$$\int_{SM} dh(X_s)\,\xi \;=\; 0\,, \qquad \forall\, h \in C^\infty(SM,\mathbb{R})\,. \tag{67}$$

The rotation vector *of ξ is $\rho(\xi) \in H_1(SM,\mathbb{R})$ defined by duality on $[\tau] \in H^1(SM,\mathbb{R})$:*

$$< [\tau], \rho(\xi) > \;=\; \int_{SM} \tau(X_s)\,\xi\,, \tag{68}$$

where $\tau \in \Omega^1(SM)$ is any closed form representing the class $[\tau]$.

Since X_s is a section of $\ker \omega_s$ and ω_s is nowhere vanishing, we can find a unique volume form Ω_s such that $\imath_{X_s}\Omega_s = \omega_s$. We can write $\Omega_s = \tau_s \wedge \omega_s$, where τ_s is any 1-form such that $\tau_s(X_s) = 1$. We easily see that $\alpha(X + sfV) = 1 + 0$. Hence, $\Omega_s = \alpha \wedge \omega_s = \alpha \wedge d\alpha$. Notice, indeed, that $\alpha \wedge \pi^*\sigma = 0$ since it is annihilated by V.

Definition 6.8. *The* Liouville measure ξ_{SM} *on SM is the Borel measure defined by integration with the differential form $\alpha \wedge d\alpha$. It is an invariant measure for Φ^s for every $s > 0$.*

In order to compute the rotation vector of ξ_{SM}, we need a lemma which tells us when ω_s is exact. The easy proof is left to the reader.

Lemma 6.9. *If σ is exact, then $\pi^*\sigma$ is exact and we have an injection*

$$\begin{array}{rcl} \textit{Primitives of } \sigma & \longrightarrow & \textit{Primitives of } \omega_s \\ \zeta & \longmapsto & \alpha \;-\; s\pi^*\zeta\,. \end{array} \tag{69}$$

If $M \neq \mathbb{T}^2$, then $\pi^*\sigma$ is exact and we have an injection

$$\begin{array}{rrcl} & \textit{Primitives of } \ \sigma - \dfrac{[\sigma]}{2\pi\chi(M)}K\mu & \longrightarrow & \textit{Primitives of } \ \omega_s \\ (70) & & & \\ & \zeta & \longmapsto & \alpha - s\pi^*\zeta + s\dfrac{[\sigma]}{2\pi\chi(M)}\psi\,. \end{array}$$

If $M = \mathbb{T}^2$ and σ is non-exact, then ω_s is non-exact.

We can now state a proposition concerning $\rho(\xi_{SM})$.

Proposition 6.10. *If $[\sigma] \neq 0$ and $M = \mathbb{T}^2$, then there holds $\rho(\xi_{SM}) = s[\sigma]\cdot[S_qM]$, where $[S_qM] \in H_1(SM,\mathbb{Z})$ is the class of a fiber of $SM \to M$ oriented counter-clockwise. Otherwise, $\rho(\xi_{SM}) = 0$.*

Proof. Let $[\tau] \in H^1(SM;\mathbb{R})$. We notice that

$$\tau(X_s)\,\alpha \wedge d\alpha \;=\; \imath_{X_s}\Big(\tau \wedge \alpha \wedge d\alpha\Big) \;+\; \tau \wedge \imath_{X_s}\big(\alpha \wedge d\alpha\big) \;=\; 0 \;+\; \tau \wedge \omega_s\,.$$

Therefore,

$$< [\tau], \rho(\xi_{SM}) > = \int_{SM} \tau \wedge \omega_s \;=\; s\int_{SM} \tau \wedge \pi^*\sigma\,.$$

If $M = \mathbb{T}^2$, then $S\mathbb{T}^2 \simeq S^1 \times \mathbb{T}^2$ and we can use Fubini's theorem to integrate separately in the vertical directions and in the horizontal direction. Observe that since τ is closed, the integral over a fiber $S_q\mathbb{T}^2$ does not depend on q. Thus we find

$$\int_{S\mathbb{T}^2} \tau \wedge \pi^*\sigma \;= < [\tau], [S_q\mathbb{T}^2] > \cdot\,[\sigma]\,.$$

and the proposition is proven for the 2-torus. When $M \neq \mathbb{T}^2$, $\pi^*\sigma$ is exact and, therefore, $\int_{SM} \tau \wedge \pi^*\sigma = 0$. The proposition is proven also in this case. $\square$

We now proceed to the proofs of the four propositions.

6.3. **Proof of Proposition 6.1.** If $M = \mathbb{T}^2$ and $[\sigma] \neq 0$, then ω_s is not exact by Lemma 6.9. In particular, SM cannot be of contact type. This proves the first statement of the proposition. Now let $\tau_s \in \Omega^1(SM)$ be a stabilizing form for ω_s. Since $\ker(d\tau_s) \supset \ker\omega_s$, there exists a function $\rho_s : SM \to \mathbb{R}$ such that $d\tau_s = \rho_s\omega_s$. Taking the exterior differential in this equation, we get $0 = d\rho_s \wedge \omega_s$. Plugging in the vector field X_s we get $0 = d\rho_s(X_s)\omega_s$. Since ω_s is nowhere zero, we conclude that $d\rho_s(X_s) = 0$. Namely, ρ_s is a first integral for the flow. By assumption, ρ_s is equal to a constant. If $\rho_s = 0$, then τ_s is closed, if $\rho_s \neq 0$, then τ_s is a contact form. Suppose the first alternative holds. Since $\tau_s(X_s) \neq 0$ everywhere, we have

$$0 \;\neq\; \int_{SM} \tau_s(X_s)\xi_{SM} \;= < [\tau_s], \rho(\xi_{SM}) > \;.$$

By Proposition 6.10, this can only happen if $M = \mathbb{T}^2$ and $< [\tau_s], [S_q\mathbb{T}^2] > \neq 0$, which is what we had to prove.

6.4. **Proof of Proposition 6.2.** The proof of the second proposition is based on the fact that when ω_s is exact we can associate a number to every invariant measure with zero rotation vector.

Definition 6.11. *Suppose ω_s is exact and that ξ is a Φ^s-invariant measure with $\rho(\xi)=0$. We define the* action *of ξ as the number*

$$\mathcal{S}_s(\xi) \;:=\; \int_{SM} \tau_s(X_s)\,\xi\,, \tag{71}$$

where τ_s is any primitive for ω_s. Such number does not depend on τ_s since $\rho(\xi)=0$.

The action of invariant measures gives an obstruction to being of contact type.

Lemma 6.12. *Suppose ω_s is exact and that ξ is a non-zero Φ^s-invariant measure with $\rho(\xi)=0$. If $\mathcal{S}_s(\xi)\leq 0$, then SM cannot be of positive contact type. If $\mathcal{S}_s(\xi)\geq 0$, then SM cannot be of negative contact type.*

Proof. If SM is of positive contact type, there exists τ_s such that $d\tau_s=\omega_s$ and $\tau_s(X_s)>0$. Therefore,

$$\mathcal{S}_s(\xi) \;=\; \int_{SM}\tau_s(X_s)\,\xi \;\geq\; \inf_{SM}\tau_s(X_s)\cdot\xi(SM) \;>\; 0\,.$$

For the case of negative contact type, we argue in the same way. ∎

Let us now compute the action of the Liouville measure.

Proposition 6.13. *If σ is exact, then*

$$\mathcal{S}_s(\xi_{SM}) \;=\; \xi_{SM}(SM) \;=\; 2\pi[\mu]\,. \tag{72}$$

If $M\neq\mathbb{T}^2$, then

$$\mathcal{S}_s(\xi_{SM}) \;=\; \xi_{SM}(SM) \;+\; s^2\frac{[\sigma]^2}{\chi(M)}\,. \tag{73}$$

Proof. If $\sigma=d\zeta$, then $\alpha-s\pi^*\zeta$ is a primitive of ω_s by Lemma 6.9 and we have

$$(\alpha-s\pi^*\zeta)(X_s)_{(q,v)} \;=\; 1 \;-\; s\zeta_q(v)\,, \quad \forall\,(q,v)\in SM\,. \tag{74}$$

Consider the *flip* $I:SM\to SM$ given by $I(q,v):=(q,-v)$. We see that

$$(I^*\alpha)_{(q,v)} \;=\; \alpha_{I(q,v)}dI \;=\; g_q(-v,d\pi dI\cdot) \;=\; \alpha_{I(q,v)}\,.$$

Hence ξ_{SM} is I-invariant. However, $\zeta\circ I(q,v)=-\zeta(q,v)$. Therefore,

$$\int_{SM}\zeta\,\xi_{SM} \;=\; 0 \tag{75}$$

and from the definition of action given in (71), we see that (72) is satisfied. To prove the second identity, we consider a primitive $\alpha-s\pi^*\zeta+s\frac{[\sigma]}{2\pi\chi(M)}\psi$ for ω_s as prescribed by Lemma 6.9. We compute

$$\left(\alpha \;-\; s\pi^*\zeta \;+\; s\frac{[\sigma]}{2\pi\chi(M)}\psi\right)(X_s)_{(q,v)} \;=\; 1 \;-\; s\zeta_q(v) \;+\; s^2\frac{[\sigma]}{2\pi\chi(M)}f(q)\,. \tag{76}$$

Thus, we need to estimate the integral of $f\circ\pi$ on SM. Let U_i be an open cover of M such that $SU_i\simeq S^1\times U_i$ and let a_i be a partition of unity subordinated to it.

We have

$$\begin{aligned}\int_{SM} f(q)\,\alpha\wedge d\alpha \;=\; \int_{SM} f(q)\,\alpha\wedge\psi\wedge\beta \;&=\; -\int_{SM} f(q)\,\psi\wedge\pi^*\mu\\ &=\; -\sum_i\int_{SU_i} a_i(q)\,\psi\wedge\pi^*\sigma\\ &=\; -\sum_i\int_{S^1\times U_i} a_i(q)\,(-d\varphi\wedge\pi^*\sigma)\\ &=\; \sum_i\int_{U_i} a_i(q)\,\sigma\int_{S^1} d\varphi\\ &=\; 2\pi\sum_i\int_{U_i} a_i(q)\,\sigma\\ &=\; 2\pi[\sigma]\,,\end{aligned}$$

where φ is an angular coordinate on S_qU_i going in the clockwise direction (hence the presence of an additional minus sign in the third line). Putting this computation together with (75), we get the desired identity. □

Proposition 6.2 now follows from Lemma 6.12 and Proposition 6.13 after defining

$$c_h(g,\sigma) \;:=\; -\,\frac{[\sigma]^2}{4\pi\chi(M)[\mu]}\,,\quad \text{when } M \text{ has genus higher than one}\,. \tag{77}$$

Remark 6.14. *We have seen in the homogeneous example above that $c_h(g,\sigma) = c(g,\sigma)$. The relation between c_h and the Mañé critical value was studied in general by G. Paternain in* [Pat09]. *There the author proves that $c_h(g,\sigma)\le c(g,\sigma)$ and that $c_h(g,\sigma) = c(g,\sigma)$ if and only if g is a metric of constant curvature and σ is a multiple of the area form.*

6.5. **Proof of Proposition 6.3.** Suppose that σ is exact and let us consider a primitive $\alpha - s\pi^*\zeta$ given by Lemma 6.9. We have

$$(\alpha \;-\; s\pi^*\zeta)(X_s)_{(q,v)} \;=\; 1 \;-\; s\zeta_q(v) \;\ge\; 1 \;-\; s\sup_M|\zeta|\,.$$

Requiring that the right hand-side is positive is equivalent to saying that

$$k \;=\; \frac{1}{2s^2} \;>\; \sup_M\frac{1}{2}|\zeta|^2\,.$$

Since this holds for every ζ which is a primitive for σ, we have that the last inequality is equivalent to $k > c_0(g,\sigma)$. Contreras, Macarini and G. Paternain also found in [CMP04] examples of exact systems on $\mathbb{T}^2$, which are of contact type for $k = c_0(g,\sigma)$ (see also [Ben14a, Section 4.1.1]). We will not discuss these examples here and we refer the reader to the cited literature for more details.

Let us now deal with the non-exact case. If $M\neq\mathbb{T}^2$, then we consider a primitive of the form $\alpha - s\pi^*\zeta + s\frac{[\sigma]}{2\pi\chi(M)}\psi$ and we compute

$$\Big(\alpha \;-\; s\pi^*\zeta \;+\; s\frac{[\sigma]}{2\pi\chi(M)}\psi\Big)(X_s)_{(q,v)} \;=\; 1 \;-\; s\zeta_q(v) \;+\; s^2\frac{[\sigma]}{2\pi\chi(M)}f(q)\,. \tag{78}$$

We can give the estimate from below

$$1 \;-\; s\zeta_q(v) \;+\; s^2\frac{[\sigma]}{2\pi\chi(M)}f(q) \;\geq\; 1 \;-\; s\sup_M|\zeta| - s^2\left|\frac{[\sigma]}{2\pi\chi(M)}\right|\cdot\sup_M|f|$$

and we see that this quantity is strictly positive for s small enough.

Suppose now that σ is a symplectic form on M. We have three cases.

(1) If $M = S^2$, then the quantity in (78) is bounded from below by

$$1 - s\sup_M|\zeta| + s^2\frac{[\sigma]}{4\pi}\cdot\inf_M f\,.$$

Since $[\sigma] > 0$, we have that $\inf f > 0$ and we see that such quantity is strictly positive for big s.

(2) If M has genus larger than 1, then the quantity in (78) is bounded from above by

$$1 \;+\; s\sup_M|\zeta| \;+\; s^2\frac{[\sigma]}{2\pi\chi(M)}\cdot\inf_M f\,.$$

Since $\chi(M) < 0$ and $\inf f > 0$, such quantity is strictly negative for big s.

(3) If $M = \mathbb{T}^2$, then there exists a closed form $\tau \in \Omega^1(S\mathbb{T}^2)$ such that $\tau(V) = 1$ (prove such statement as an exercise). Thus, we get

$$\tau(X_s) \;=\; \tau(X) \;+\; sf \;\geq\; \inf_{SM}\tau(X) \;+\; s\inf_M f \tag{79}$$

and such quantity is positive provided $\inf f > 0$ and s is big enough.

6.6. **Proof of Proposition 6.4.** If σ is exact and $k < c_0(g,\sigma)$, we can use Theorem 5.5 to find an embedded surface $\Pi \subset M$ with non-empty boundary $\partial\Pi = \{\gamma_i\}$ such that $\mathcal{T}_k(\Pi) < 0$ and the γ_i's are periodic orbits of Φ^s (parametrized by arc-length). Let $(\gamma_i, \dot\gamma_i)$ be the corresponding curve on SM and let ξ_i be the associated invariant measure. Define $\xi_{\partial\Pi} := \sum_i \xi_i$. What is its rotation vector? Call $\pi_* : H_1(SM;\mathbb{R}) \to H_1(M;\mathbb{R})$ the map induced by the projection π in homology and observe that

$$\pi_*(\rho(\xi_{\partial\Pi})) \;=\; \sum_i \pi_*(\rho(\xi_i)) \;=\; \sum_i [\gamma_i] \;=\; [\partial\Pi] \;=\; 0\,. \tag{80}$$

Exercise 8. *The map π_* is an isomorphism if $M \neq \mathbb{T}^2$.*

Thus, we conclude that $\rho(\xi_{\partial\Pi}) = 0$, if $M \neq \mathbb{T}^2$. Let us compute the action in this case. As before, we use a primitive $\alpha - s\pi^*\zeta$:

$$\begin{aligned}\mathcal{S}_s(\xi_{\partial\Pi}) = \sum_i\int_{SM}(1 - s\zeta_q(v))\xi_i \;&=\; \sum_i\int_0^{\ell(\gamma_i)}\big(1 - s\zeta_{\gamma_i}(\dot\gamma_i)\big)\mathrm{d}t\\ &=\; \sum_i\ell(\gamma_i) - s\int_0^{\ell(\gamma_i)}\gamma_i^*\zeta\\ &=\; \ell(\partial\Pi) - s\int_\Pi\sigma = s\mathcal{T}_k(\Pi)\,.\end{aligned} \tag{81}$$

By hypothesis the last quantity is negative and Lemma 6.12 tells us that Σ_k cannot be of positive contact type. Since by Proposition 6.2, Σ_k cannot be of negative contact type either, point *(1)* of the proposition is proved.

We now move to prove point *(2a)* with the aid of a little exercise.

Exercise 9. *We prove a generalization of* (81)*, when $M \neq \mathbb{T}^2$. Let Π be an embedded surface such that $\partial\Pi$ is a union of periodic orbits and let $\xi_{\partial\Pi}$ be the invariant measure constructed as before. Then,*

$$\frac{\mathcal{S}_s(\xi_{\partial\Pi})}{s} \;=\; \mathcal{T}_k(\Pi) \;+\; \frac{\mathfrak{o}(\Pi)\chi(\Pi)[\sigma]}{\chi(M)}\,, \tag{82}$$

where $\mathfrak{o}(\Pi) \in \{+1,-1\}$ record the orientation of Π. To prove such identity one recalls that $\kappa_{\gamma_i} = sf(\gamma_i)$ and then uses the Gauss-Bonnet theorem (taking into account orientations) to express the integral of the geodesic curvature along $\partial\Pi$. What happens if we consider $M \setminus \Pi$? Do the two expressions for $\mathcal{S}_s(\xi_{\partial\Pi})$ agree? Remember relation (63)*.*

The problem with formula (82) is that Theorem 5.5 does not give any information on the Euler characteristic of Π. To circumvent this problem we need the following result by Ginzburg [Gin87] (see also [AB15, Chapter 7]).

Proposition 6.15. *If $\sup f > \varepsilon$ for some $\varepsilon < 0$, there exists a constant $C > 0$ such that for every small enough k we can find a simple periodic orbit γ^k_+ supported on $\{f > \varepsilon\}$ and such that $\ell(\gamma^k_+) \leq \sqrt{2k}C$.*

If $\inf f < -\varepsilon$, for some $\varepsilon > 0$, there exists $C > 0$ such that for every small enough k, there exists a simple periodic orbit γ^k_- supported on $\{f < -\varepsilon\}$ and such that $\ell(\gamma^k_-) \leq \sqrt{2k}C$.

If f is negative at some point, by Proposition 6.15, there exists γ^k_- with the properties listed above, for k small. In particular, γ^k_- bounds a small disc D^k_-. Since the geodesic curvature of γ^k_- is very negative, such disc lies in $\mathcal{E}_-(M)$. When $M \neq \mathbb{T}^2$, we use (82) and find

$$\frac{S_s(\xi_{\partial D^k_-})}{s} \;=\; \mathcal{T}_k(D^k_-) \;-\; \frac{2}{\chi(M)}[\sigma]\,.$$

By the estimate on the length of γ^k_- we get that $|\mathcal{T}_k(D^k_-)| \leq Ck^2$ (see (21)). Therefore, $S_s(\xi_{\partial D^k_-})$ has the opposite sign of $\chi(M)$ for k small enough. Combining Lemma 6.12 and Proposition 6.2, point *(2a)* is proven.

Let us deal now with the case of the 2-torus. Since $[\sigma] > 0$, by Proposition 6.15 there exists also γ^k_+ bounding a disc D^k_+. Let $\Pi^k = D^k_- \cup D^k_+$. We claim that the measure $\xi_{\partial\Pi^k}$ has zero rotation vector.

Exercise 10. *Prove the claim by showing that $(\gamma^k_+, \dot\gamma^k_+)$ is freely homotopic in $S\mathbb{T}^2$ to $[S_q\mathbb{T}^2]$, namely the class of a fiber with orientation given by V. Analogously, prove that $(\gamma^k_-, \dot\gamma^k_-)$ is freely homotopic to a fiber with the opposite orientation.*

If τ_s is a closed stabilizing form, we have that the function $\tau_s(X_s)$ is nowhere zero. Therefore,

$$0 \;\neq\; \int_{S\mathbb{T}^2} \tau_s(X_s)\,\xi_{\partial\Pi^k} \;= <[\tau_s], \rho(\xi_{\partial\Pi^k})> \;=\; 0\,,$$

which is a contradiction.

We omit the proof of point *(3)*, for which we refer the reader to [Ben14b].

Acknowledgements

We would like to express our gratitude to Ezequiel Maderna and Ludovic Rifford for organizing the research school and for the friendly atmosphere they created while we stayed in Uruguay. We also sincerely thank Marco Mazzucchelli and Alfonso Sorrentino for many engaging discussions during our time at the school.

References

[AB14] L. Asselle and G. Benedetti, *The Lusternik-Fet theorem for autonomous Tonelli Hamiltonian systems on twisted cotangent bundles*, J. Topol. Anal., Online version doi:10.1142/S1793525316500205 (2015).

[AB15] ———, *Infinitely many periodic orbits of non-exact oscillating magnetic flows on surfaces with genus at least two for almost every low energy level*, Calc. Var. Partial Diff. Equations, Online version, http://link.springer.com/article/10.1007/s00526-015-0834-1 doi:10.1007/s00526-015-0834-1, (2015).

[Abb13] A. Abbondandolo, *Lectures on the free period Lagrangian action functional*, J. Fixed Point Theory Appl. **13** (2013), no. 2, 397–430.

[AS09] A. Abbondandolo and M. Schwarz, *A smooth pseudo-gradient for the Lagrangian action functional*, Adv. Nonlinear Stud. **9** (2009), no. 4, 597–623.

[Ben14a] G. Benedetti, *The contact property for magnetic flows on surfaces*, Ph.D. thesis, University of Cambridge, 2014.

[Ben14b] ———, *The contact property for symplectic magnetic fields on S^2*, Ergodic Theory Dynam. Systems **36** n. 3 (2014), 682–713.

[CDI97] G. Contreras, J. Delgado, and R. Iturriaga, *Lagrangian flows: the dynamics of globally minimizing orbits. II*, Bol. Soc. Brasil. Mat. (N.S.) **28** (1997), no. 2, 155–196.

[CIPP98] G. Contreras, R. Iturriaga, G. P. Paternain, and M. Paternain, *Lagrangian graphs, minimizing measures and Mañé's critical values*, Geom. Funct. Anal. **8** (1998), no. 5, 788–809.

[CM05] K. Cieliebak and K. Mohnke, *Compactness for punctured holomorphic curves*, J. Symplectic Geom. **3** (2005), no. 4, 589–654, Conference on Symplectic Topology.

[CMP04] G. Contreras, L. Macarini, and G. P. Paternain, *Periodic orbits for exact magnetic flows on surfaces*, Int. Math. Res. Not. (2004), no. 8, 361–387.

[Con06] G. Contreras, *The Palais-Smale condition on contact type energy levels for convex Lagrangian systems*, Calc. Var. Partial Differential Equations **27** (2006), no. 3, 321–395.

[Gin87] V. L. Ginzburg, *New generalizations of Poincaré's geometric theorem*, Funktsional. Anal. i Prilozhen. **21** (1987), no. 2, 16–22, 96.

[Gin96] ———, *On closed trajectories of a charge in a magnetic field. An application of symplectic geometry*, Contact and symplectic geometry (Cambridge, 1994), Publ. Newton Inst., vol. 8, Cambridge Univ. Press, Cambridge, 1996, pp. 131–148.

[GK02] S. Gudmundsson and E. Kappos, *On the geometry of tangent bundles*, Expo. Math. **20** (2002), no. 1, 1–41.

[HS13] U. L. Hryniewicz and P. A. S. Salomão, *Global properties of tight Reeb flows with applications to Finsler geodesic flows on S^2*, Math. Proc. Cambridge Philos. Soc. **154** (2013), no. 1, 1–27.

[HZ94] H. Hofer and E. Zehnder, *Symplectic invariants and Hamiltonian dynamics*, Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks], Birkhäuser Verlag, Basel, 1994.

[Kha79] M. P. Kharlamov, *Some applications of differential geometry in the theory of mechanical systems*, Mekh. Tverd. Tela (1979), no. 11, 37–49, 118.

[Kli78] W. Klingenberg, *Lectures on closed geodesics*, Springer-Verlag, Berlin-New York, 1978, Grundlehren der Mathematischen Wissenschaften, Vol. 230.

[Koz85] V. V. Kozlov, *Calculus of variations in the large and classical mechanics*, Uspekhi Mat. Nauk **40** (1985), no. 2(242), 33–60, 237.

[Mañ97] R. Mañé, *Lagrangian flows: the dynamics of globally minimizing orbits*, Bol. Soc. Brasil. Mat. (N.S.) **28** (1997), no. 2, 141–153.

[Mer10] W. J. Merry, *Closed orbits of a charge in a weakly exact magnetic field*, Pacific J. Math. **247** (2010), no. 1, 189–212.

[Mos73] J. Moser, *Stable and random motions in dynamical systems*, Princeton University Press, Princeton, N. J.; University of Tokyo Press, Tokyo, 1973, With special emphasis on celestial mechanics, Hermann Weyl Lectures, the Institute for Advanced Study, Princeton, N. J, Annals of Mathematics Studies, No. 77.

[MP10] L. Macarini and G. P. Paternain, *On the stability of Mañé critical hypersurfaces*, Calc. Var. Partial Differential Equations **39** (2010), no. 3-4, 579–591.

[Nov82] S. P. Novikov, *The Hamiltonian formalism and a multivalued analogue of Morse theory*, Uspekhi Mat. Nauk **37** (1982), no. 5(227), 3–49, 248.

[Osu05] O. Osuna, *Periodic orbits of weakly exact magnetic flows*, preprint, 2005.

[Pat09] G. P. Paternain, *Helicity and the Mañé critical value*, Algebr. Geom. Topol. **9** (2009), no. 3, 1413–1422.

[Taĭ93] I. A. Taĭmanov, *Closed non-self-intersecting extremals of multivalued functionals*, Siberian Math. J. **33** (1993), no. 4, 686–692.

WWU Münster, Mathematisches Institut
Einsteinstrasse 62, D-48149 Münster, Germany
E-mail address: benedett@uni-muenster.de

Publicaciones Matemáticas del Uruguay
Volumen 16, Julio 2016, Páginas 81–125
S 0797-1443

THE MORSE INDEX OF CHAPERON'S GENERATING FAMILIES

MARCO MAZZUCCHELLI

ABSTRACT. This expository paper is devoted to the Morse index of Chaperon's generating families of Hamiltonian diffeomorphisms. After reviewing the construction of such generating families, we present Bott's iteration theory in this setting: we study how the Morse index of a critical point corresponding to an iterated periodic orbit depends on the order of iteration of the orbit. We also investigate the precise dependence of the Morse index from the choice of the generating family associated to a given Hamiltonian diffeomorphism, which will allow to see the Morse index as a Maslov index for the linearized Hamiltonian flow in the symplectic group. We will conclude the survey with a proof that the classical Morse index from Tonelli Lagrangian dynamics coincides with the Maslov index.

CONTENTS

The author is partially supported by the ANR projects WKBHJ (ANR-12-BS01-0020) and COSPIN (ANR-13-JS01-0008-01).

1. Introduction

1.1. **Chaperon's generating family.** Generating families are classical objects that describe Hamiltonian diffeomorphisms of symplectic Euclidean spaces[1]. Consider a Hamiltonian diffeomorphism ϕ_0 of the standard symplectic $(\mathbb{R}^{2d}, \omega = \mathrm{d}x \wedge \mathrm{d}y)$. The graph of ϕ_0 is a Lagrangian submanifold of the product $\mathbb{R}^{2d} \times \mathbb{R}^{2d}$ equipped with the symplectic form $(-\omega) \oplus \omega$. The graph of the identity diffeomorphism on $\mathbb{R}^{2d}$ is the diagonal subspace of $\mathbb{R}^{2d} \times \mathbb{R}^{2d}$, and the fixed points of ϕ_0 correspond to the intersection points of its graph with the diagonal. Let us translate this picture on the cotangent bundle $\mathrm{T}^*\mathbb{R}^{2d}$, which is equipped with the canonical symplectic structure given by minus the exterior derivative of the Liouville form $\lambda = p\,\mathrm{d}q$ (here q and p are the variables on the base and on the fiber respectively). We choose a symplectomorphism $(\mathbb{R}^{2d} \times \mathbb{R}^{2d}, (-\omega) \oplus \omega) \to (\mathrm{T}^*\mathbb{R}^{2d}, -\mathrm{d}\lambda)$ that sends the diagonal subspace to the zero-section. In this survey, we will employ the following one:

$$(x_0, y_0, x_1, y_1) \mapsto (\underbrace{x_1, y_0}_{q}, \underbrace{y_1 - y_0, x_0 - x_1}_{p}).$$

The image of the graph of ϕ_0 under this symplectomorphism is a Lagrangian submanifold L_0. Assume now that L_0 is a section of the cotangent bundle, that is, the graph of a one-form μ_0 on the base $\mathbb{R}^{2d}$. This is always verified provided ϕ_0 is sufficiently close to the identity in the C^1-topology, or more generally whenever ϕ_0 admits an associated diffeomorphism $\psi_0 : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ such that $\phi_0(x_0, y_0) = (x_1, y_1)$ if and only if $\psi_0(x_1, y_0) = (x_0, y_1)$. Lagrangian sections of cotangent bundles are precisely the graphs of closed one-forms on the base (we refer the reader to [HZ94, MS98] for this and other background results from symplectic geometry). Therefore, the one-form μ_0 must be exact, i.e. $\mu_0 = \mathrm{d}f_0$. We say that $f_0 : \mathbb{R}^{2d} \to \mathbb{R}$ is a **generating function** for the Hamiltonian diffeomorphism ϕ_0. The explicit way f_0 determines ϕ_0 is the following:

$$\phi_0(x_0, y_0) = (x_1, y_1) \qquad \text{if and only if} \qquad \begin{cases} x_1 - x_0 = -\partial_y f_0(x_1, y_0), \\ y_1 - y_0 = \partial_x f_0(x_1, y_0). \end{cases}$$

Not only the function f_0 describes the Hamiltonian diffeomorphism ϕ_0, it also provides a variational principle for the fixed points of ϕ_0: they are precisely the critical points of f_0. Notice that the generating function of a Hamiltonian diffeomorphism is unique up to an additive constant.

A general Hamiltonian diffeomorphism ϕ of $(\mathbb{R}^{2d}, \omega)$ does not necessarily admit a generating function, since its associated Lagrangian submanifold $L \subset \mathrm{T}^*\mathbb{R}^{2d}$ may not be a section. However, the following construction originally due to Chaperon [Cha84, Cha85] allows to draw a similar conclusion provided the behavior of ϕ at infinity is suitably controlled. For instance, assume that ϕ is the time-1 map of a

[1]More generally, generating families describe certain Lagrangian submanifolds, those who are images of the zero-section under a Hamiltonian diffeomorphism, of cotangent bundles. This more general notion originated from the work of Hörmander [Hör71], but was introduced in symplectic topology by Sikorav [Sik87] and further studied by many other authors.

non-autonomous Hamiltonian flow ϕ_H^t whose associated Hamiltonian $H_t : \mathbb{R}^{2d} \to \mathbb{R}$ has C^2-norm uniformly bounded in $t \in [0,1]$ by a finite constant (this condition can be weakened). By means of this flow, we can factorize ϕ as

$$\phi = \phi_{k-1} \circ ... \circ \phi_0,$$

where each factor is given by $\phi_j := \phi_H^{(j+1)/k} \circ (\phi_H^{j/k})^{-1}$. As we increase the number $k \in \mathbb{N}$ of factors, each ϕ_j becomes closer and closer to the identity in the C^1 topology. In particular, for k large enough, each factor ϕ_j is described by a generating function $f_j : \mathbb{R}^{2d} \to \mathbb{R}$ as explained in the previous paragraph. Chaperon's brilliant idea was to combine these functions together in a suitable way, in order to obtain a function defined on a larger space that defines the original ϕ. This function $F : \mathbb{R}^{2d} \times \mathbb{R}^{2d(k-1)} \to \mathbb{R}$ has the form

$$F(x_k, y_0, \boldsymbol{z}) = \sum_{j \in \mathbb{Z}_k} \Big(\langle y_j, x_{j+1} - x_j \rangle + f_j(x_{j+1}, y_j) \Big), \tag{1}$$

where $\boldsymbol{z} = (z_1, ..., z_{k-1})$ and $z_j = (x_j, y_j)$. A straightforward computation shows that

$$\begin{cases} \phi_0(x_0, y_0) = z_1, \\ \phi_1(z_1) = z_2, \\ \vdots \\ \phi_{k-2}(z_{k-2}) = z_{k-1}, \\ \phi_{k-1}(z_{k-1}) = (x_k, y_k), \end{cases}$$

if and only if

$$\begin{cases} x_k - x_0 = -\partial_{y_0} F(x_k, y_0, \boldsymbol{z}), \\ y_k - y_0 = \partial_{x_k} F(x_k, y_0, \boldsymbol{z}), \\ 0 = \partial_{\boldsymbol{z}} F(x_k, y_0, \boldsymbol{z}). \end{cases}$$

As before, the function F provides a variational principle for the fixed points of ϕ: the vector $(x_k, y_0, x_1, y_1, ..., x_{k-1}, y_{k-1})$ is a critical point of F if and only if $\phi_j(x_j, y_j) = (x_{j+1}, y_{j+1})$ for all cyclic indices $j \in \mathbb{Z}_k$. We say that F is a **generating family** for the Hamiltonian diffeomorphism ϕ, associated to its factorization $\phi_{k-1} \circ ... \circ \phi_0$. Notice that a generating family becomes a simple generating function if the parameter k is equal to 1. In the following, since we will employ generating families only in order to use their variational principle, we will write x_0 for x_k in their expression.

Let us have a closer look at Chaperon's construction in the special case where the Hamiltonian diffeomorphism ϕ is linear, that is, when $\phi(z) = Pz$ for some symplectic matrix $P \in \mathrm{Sp}(2d)$. Since the symplectic group $\mathrm{Sp}(2d)$ is connected, we can find a continuous path $\Gamma : [0,1] \to \mathrm{Sp}(2d)$ joining the identity $\Gamma(0) = I$ with $\Gamma(1) = P$. This allows to build a factorization $\phi = \phi_{k-1} \circ ... \circ \phi_0$, where each factor is the linear Hamiltonian diffeomorphism $\phi_j(z) = P_j z$ associated to the symplectic matrix

$$P_j = \Gamma(\tfrac{j+1}{k}) \Gamma(\tfrac{j}{k})^{-1} \in \mathrm{Sp}(2d).$$

Since ϕ_j is linear, there is a canonical way to normalize its generating function $f_j : \mathbb{R}^{2d} \to \mathbb{R}$ so that it becomes a quadratic function of the form

$$f_j(X_{j+1}, Y_j) = \tfrac{1}{2}\langle A_j X_{j+1}, X_{j+1}\rangle + \langle B_j X_{j+1}, Y_j\rangle + \tfrac{1}{2}\langle C_j Y_j, Y_j\rangle,$$

where A_j, B_j, and C_j are (small) $dk\times dk$ matrices, A_j and C_j being symmetric. This readily implies that the generating family $F : \mathbb{R}^{2dk} \to \mathbb{R}$ given by the expression (1) is a quadratic function as well, which we write as

$$F(\boldsymbol{Z}) = \tfrac{1}{2}\langle H\boldsymbol{Z}, \boldsymbol{Z}\rangle$$

for a suitable $2dk \times 2dk$ symmetric matrix H.

1.2. **Morse indices.** Let ϕ be a Hamiltonian diffeomorphism of $\mathbb{R}^{2d}$ described by the generating family F of equation (1). Let z_0 be a fixed point of ϕ, so that, if we set $z_j := \phi_{j-1}(z_{j-1})$ for all $j = 1, ..., k-1$, we have a corresponding critical point $\boldsymbol{z} = (z_0, ..., z_{k-1})$ of the generating family F. We are interested in the Morse indices of F at $\boldsymbol{z}$, which are defined as follows. The **Morse index** $\mathrm{ind}(\boldsymbol{z})$ is the number of negative eigenvalues of the Hessian of F at $\boldsymbol{z}$ counted with multiplicity, that is, the dimension of a maximal subspace of $\mathbb{R}^{2dk}$ where such Hessian is negative definite. Analogously, the **Morse coindex** $\mathrm{coind}(\boldsymbol{z})$ is the number of positive eigenvalues counted with multiplicity, and finally the **nullity** $\mathrm{nul}(\boldsymbol{z})$ is the dimension of the kernel of the Hessian of F at $\boldsymbol{z}$. Notice that

$$\mathrm{ind}(\boldsymbol{z}) + \mathrm{coind}(\boldsymbol{z}) + \mathrm{nul}(\boldsymbol{z}) = 2dkp.$$

In order to study these indices, let us first have a look at the Hessian of F_p at $\boldsymbol{z}$. We denote by $H(\boldsymbol{z})$ the symmetric $2dk \times 2dk$ matrix such that

$$\mathrm{Hess}F(\boldsymbol{z})[\boldsymbol{Z}, \boldsymbol{Z}'] = \langle H(\boldsymbol{z})\boldsymbol{Z}, \boldsymbol{Z}'\rangle, \qquad \forall \boldsymbol{Z}, \boldsymbol{Z}' \in \mathbb{R}^{2dk}.$$

Given any vector $\boldsymbol{Z} = (Z_1, ..., Z_{k-1}) \in \mathbb{R}^{2dk}$, its image $\boldsymbol{Z}' := H(\boldsymbol{z})\boldsymbol{Z}$ is given by

$$\begin{aligned} X_j' &= Y_{j-1} - Y_j + A_{j-1}(\boldsymbol{z})X_j + B_{j-1}(\boldsymbol{z})^T Y_{j-1}, \\ Y_j' &= X_{j+1} - X_j + B_j(\boldsymbol{z})X_{j+1} + C_j(\boldsymbol{z})Y_j. \end{aligned} \tag{2}$$

Here, we have adopted the common notation $Z_j = (X_j, Y_j)$ and $Z_j' = (X_j, Y_j)$. Moreover, as before, the index j must be understood as an element of the cyclic group $\mathbb{Z}_k$, and we have set

$$\begin{aligned} A_j(\boldsymbol{z}) &:= \partial_{xx} f_j(x_{j+1}, y_j), \\ B_j(\boldsymbol{z}) &:= \partial_{xy} f_j(x_{j+1}, y_j), \\ C_j(\boldsymbol{z}) &:= \partial_{yy} f_j(x_{j+1}, y_j). \end{aligned} \tag{3}$$

From now on, we will assume that the parameter k is large enough, so that the norms of the matrices $A_j(\boldsymbol{z})$, $B_j(\boldsymbol{z})$, and $C_j(\boldsymbol{z})$ are bounded from above by some $\epsilon < 1$.

Remark 1.1. The quadratic function

$$\tilde{f}_j(X_{j+1}, Y_j) = \tfrac{1}{2}\langle A_j(\boldsymbol{z})X_{j+1}, X_{j+1}\rangle + \langle B_j(\boldsymbol{z})X_{j+1}, Y_j\rangle + \tfrac{1}{2}\langle C_j(\boldsymbol{z})Y_j, Y_j\rangle$$

is the generating function for the linearized map $\mathrm{d}\phi_j(z_j)$. Therefore, the quadratic function $\tilde{F} : \mathbb{R}^{2dk} \to \mathbb{R}$ given by

$$\tilde{F}(\boldsymbol{Z}) = \tfrac{1}{2}\langle H(\boldsymbol{z})\boldsymbol{Z}, \boldsymbol{Z}\rangle = \sum_{j\in\mathbb{Z}_k} \Big(\langle Y_j, X_{j+1} - X_j\rangle + \tilde{f}_j(X_{j+1}, Y_j)\Big).$$

is the generating family of the linearized map $\mathrm{d}\phi(z_0)$ associated to his factorization $\mathrm{d}\phi_{k-1}(z_{k-1}) \circ ... \circ \mathrm{d}\phi_0(z_0)$. □

In the context of convex Hamiltonian systems, for instance in the study of closed geodesics in Riemannian manifolds, it is well known that the classical Lagrangian action functional has finite Morse indices (we will discuss this further in Section 4). Even more remarkably, there are closed geodesics that have Morse index zero when they are iterated any number of times, for instance in hyperbolic Riemannian manifolds. On the contrary, the Hamiltonian action functional has always infinite Morse indices at his critical points. Since our generating family F can be considered a finite dimensional approximation of the Hamiltonian action functional, the unboundedness of the Hamiltonian Morse indices is reflected by the fact that the Morse indices of F tend to be large. For instance, if the Hamiltonian diffeomorphism ϕ we started with were the identity, we could choose $f_0 = ... = f_{k-1} \equiv 0$; the function F would then be a degenerate quadratic form with Morse index and coindex both equal to $d(k-1)$. In general, we have at least the following lower bounds.

Proposition 1.2. *For all critical points $\boldsymbol{z}$ of F, we have*

$$\min\{\mathrm{ind}(\boldsymbol{z}), \mathrm{coind}(\boldsymbol{z})\} \geq d\lfloor k/2 \rfloor.$$

Proof. Consider the vector subspace of $(\mathbb{R}^{2d})^k$ given by

$$\mathbb{V} := \{\boldsymbol{Z} = (Z_0, ..., Z_{k-1}) \in \mathbb{R}^{2dk} \mid Z_j = 0 \ \forall j \text{ even}, \quad Y_h = X_h \ \forall h \text{ odd}\}.$$

By (2), for all $\boldsymbol{Z} \in \mathbb{V}$ we have

$$\begin{aligned}
\langle H(\boldsymbol{z})\boldsymbol{Z}, \boldsymbol{Z}\rangle &= \sum_{j \text{ odd}} \Big(-|X_j|^2 - |Y_j|^2 + \langle A_{j-1}X_j, X_j\rangle + \langle C_j Y_j, Y_j\rangle \Big) \\
&\leq \sum_{j \text{ odd}} (\epsilon - 1)(|X_j|^2 + |Y_j|^2) \\
&= \underbrace{(\epsilon - 1)}_{<0} |\boldsymbol{Z}|^2.
\end{aligned}$$

This shows that the Hessian of F at $\boldsymbol{z}$ is negative definite on $\mathbb{V}$, and in particular $\mathrm{ind}(\boldsymbol{z}) \geq \dim \mathbb{V} = d\lfloor k/2 \rfloor$. By an analogous computation, the Hessian of F at $\boldsymbol{z}$ is positive definite on

$$\mathbb{W} := \{\boldsymbol{Z} \in \mathbb{R}^{2dk} \mid Z_j = 0 \ \forall j \text{ even}, \quad Y_h = -X_h \ \forall h \text{ odd}\},$$

which implies $\mathrm{coind}(\boldsymbol{z}) \geq \dim \mathbb{W} = d\lfloor k/2 \rfloor$. □

Studying the properties of the Morse indices of generating families is tremendously important for the applications to the existence and multiplicity of periodic orbits of Hamiltonian systems. Indeed, minimax methods from non-linear analysis allow to find critical points of a generating family with almost prescribed indices. More precisely, a minimax scheme of dimension n, such as a minimax over the family of representative of an homology or homotopy class of degree n, may only converge to critical points with Morse index less than or equal to n and Morse index plus nullity larger than or equal to n. Suppose that we are interested in the periodic points of a Hamiltonian diffeomorphism ϕ described by a generating family F. The factorization $\phi = \phi_{k-1} \circ ... \circ \phi_0$ employed to build F can be iterated p times in order to build a generating family F_p for the iterated Hamiltonian diffeomorphism ϕ^p. A fixed point z_0 of ϕ gives a critical point $\boldsymbol{z} = (z_0, ..., z_{k-1})$ of the generating family F, and its p-th fold juxtaposition $\boldsymbol{z}^p = (\boldsymbol{z}, ..., \boldsymbol{z})$ gives a critical point of the generating family F_p.

Now, assume that one can setup a minimax scheme with every function F_p that produces a critical point $\boldsymbol{z}_p$ with Morse index $i_p = \mathrm{ind}(\boldsymbol{z}_p)$, coindex $c_p = \mathrm{coind}(\boldsymbol{z}_p)$, and nullity $n_p = \mathrm{nul}(\boldsymbol{z}_p)$. The natural question to ask is whether the family of critical points $\{\boldsymbol{z}_p \mid p \in \mathbb{N}\}$ corresponds to infinitely many distinct periodic points of ϕ. As we just saw, the answer in general is no: in the worst case, all the critical points $\boldsymbol{z}_p$ may be of the form $\boldsymbol{z}^p = (\boldsymbol{z}, ..., \boldsymbol{z})$ and thus correspond to the same fixed point z_0. One way to address this question is to study the admissible behavior of the function $p \mapsto (\mathrm{ind}(\boldsymbol{z}^p), \mathrm{coind}(\boldsymbol{z}^p), \mathrm{nul}(\boldsymbol{z}^p))$ that associate to a period p the indices of the critical points of F_p corresponding to a fixed point z_0 of ϕ. In the more special setting of Tonelli Lagrangian systems (c.f. Section 4), this idea goes back to the work of Hedlund [Hed32] and Morse-Pitcher [MP34] from the 1930s, and was greatly developed two decades later by Bott in his celebrated paper [Bot56]. If the sequence of indices $\{(i_p, c_p, n_p) \mid p \in \mathbb{N}\}$ provided by the minimax schemes does not have an admissible behavior, one can immediately conclude that the family of critical points $\{\boldsymbol{z}_p \mid p \in \mathbb{N}\}$ cannot correspond to a single fixed point of ϕ. Sometimes, this argument can be pushed further to show that such family of critical points cannot correspond to a finite set of periodic points of ϕ, and thus infer that ϕ possesses infinitely many periodic points.

1.3. **Organization of the paper.** In Section 2 we will present the aforementioned Bott's iteration theory in the general setting of generating families. We will not provide applications of this theory, but we will mention some of them in the last Subsection 2.5. In Section 3 we will discuss the dependence of the Morse index from the specific choice of the generating family. We will show that the Morse index can be seen as a Maslov index, a certain homotopy invariant for continuous paths in the symplectic group. In Section 4 we will consider the special case of Hamiltonian diffeomorphisms generated by a non-autonomous Tonelli Hamiltonian. We will show that, in this case, the Morse indices of the generating family (or the Maslov indices of the associated symplectic paths) are related to the Morse indices of the classical Lagrangian action functional. As the reader will see, throughout the sections we will often be dealing with quadratic forms, which inevitably involves some linear algebra. In the Appendix of the paper we have collected the less standard tools from plain and symplectic linear algebra that we will need. None of the results contained in this survey is original, although some of the proofs are different form the ones available in the literature. Many authors contributed to the theory presented, and it seems almost impossible to provide a complete and precise historical account. We will give the main references to the vast bibliography at the end of each section.

2. Bott's iteration theory for generating families

2.1. **Bott indices.** Consider a Hamiltonian diffeomorphism $\phi \in \mathrm{Ham}(\mathbb{R}^{2d})$. Assume that the behavior of ϕ at infinity is suitably controlled, so that we have a factorization $\phi = \phi_{k-1} \circ ... \circ \phi_0$ where each $\phi_j \in \mathrm{Ham}(\mathbb{R}^{2d})$ is defined by a generating function $f_j : \mathbb{R}^{2d} \to \mathbb{R}$. For each period $p \in \mathbb{N}$, the iterated diffeomorphism

ϕ^p is defined by the generating family $F_p : \mathbb{R}^{2dkp} \to \mathbb{R}$ given by

$$F_p(z_0, ..., z_{kp-1}) = \sum_{j\in\mathbb{Z}_{kp}} \Big(\langle y_j, x_{j+1} - x_j \rangle + f_{j \bmod k}(x_{j+1}, y_j) \Big),$$

where as usual we have adopted the notation $z_j = (x_j, y_j) \in \mathbb{R}^{2d}$. Consider a fixed point z_0 of ϕ, with associated critical point $\boldsymbol{z} = (z_0, ..., z_{k-1})$ of F_1. For all periods $p \in \mathbb{N}$, the critical point of F_p corresponding to the p-periodic orbit of ϕ starting at z_0 is given by $\boldsymbol{z}^p = (\boldsymbol{z}, ..., \boldsymbol{z})$. Let $H_p = H_p(\boldsymbol{z}^p)$ be the $2dkp \times 2dkp$ symmetric matrix associated to the Hessian of F_p at $\boldsymbol{z}^p$, i.e.

$$\mathrm{Hess} F_p(\boldsymbol{z}^p)[\boldsymbol{Z}', \boldsymbol{Z}''] = \langle H_p \boldsymbol{Z}', \boldsymbol{Z}'' \rangle.$$

Due to the special form of our critical point $\boldsymbol{z}^p$, an image $\boldsymbol{Z}' := H_p \boldsymbol{Z}$ is defined by

$$\begin{aligned} X'_j &= Y_{j-1} - Y_j + A_{j-1 \bmod k} X_j + B^T_{j-1 \bmod k} Y_{j-1}, \\ Y'_j &= X_{j+1} - X_j + B_{j \bmod k} X_{j+1} + C_{j \bmod k} Y_j, \end{aligned} \tag{4}$$

where the subscript j belongs to $\mathbb{Z}_{kp}$, and the matrices $A_j = A_j(\boldsymbol{z})$, $B_j = B_j(\boldsymbol{z})$, and $C_j = C_j(\boldsymbol{z})$ are defined as before in (3).

We wish to investigate the behavior of the Morse indices under iteration, that is, the behavior of the functions $p \mapsto \mathrm{ind}(\boldsymbol{z}^p)$, $p \mapsto \mathrm{coind}(\boldsymbol{z}^p)$, and $p \mapsto \mathrm{nul}(\boldsymbol{z}^p)$. For this purpose, let us interpret H_p in an equivalent, but conceptually slightly different, way: we see it as a second order difference operator $\mathcal{H}_p$ acting on the vector space of kp-periodic sequences

$$\mathbb{V}_p := \big\{ (Z_j)_{j\in\mathbb{Z}} \in (\mathbb{R}^{2d})^{\mathbb{Z}} \mid Z_{j+kp} = Z_j \quad \forall j \in \mathbb{Z} \big\}.$$

Following Bott [Bot56], let us complexify the setting by introducing, for every θ in the unit circle $S^1 \subset \mathbb{C}$, the vector space of sequences

$$\mathbb{V}_{p,\theta} := \big\{ (Z_j)_{j\in\mathbb{Z}} \in (\mathbb{C}^{2d})^{\mathbb{Z}} \mid Z_{j+kp} = \theta Z_j \quad \forall j \in \mathbb{Z} \big\}.$$

We equip this vector space with the Hermitian product

$$\langle (Z_j)_{j\in\mathbb{Z}}, (Z'_j)_{j\in\mathbb{Z}} \rangle_{p,\theta} = \sum_{j=0}^{kp-1} \langle Z_j, Z'_j \rangle = \sum_{j=0}^{kp-1} Z_j \overline{Z'_j}.$$

We introduce the linear operator $\mathcal{H}_{p,\theta} : \mathbb{V}_{p,\theta} \to \mathbb{V}_{p,\theta}$ given by $\mathcal{H}_{p,\theta}(Z_j)_{j\in\mathbb{Z}} = (Z'_j)_{j\in\mathbb{Z}}$. Here, we have denoted $Z_j = (X_j, Y_j)$, and defined $Z'_j = (X'_j, Y'_j)$ by the equations (4), where the subscript j is now in $\mathbb{Z}$. The operator $\mathcal{H}_{p,\theta}$ is Hermitian with respect to the product $\langle \cdot, \cdot \rangle_{p,\theta}$, and in particular it has real spectrum. Indeed, the vector space $\mathbb{V}_{p,\theta}$ is isomorphic to $(\mathbb{C}^{2d})^{kp}$ via the map

$$(Z_j)_{j\in\mathbb{Z}} \mapsto (Z_0, ..., Z_{kp-1}),$$

which pulls back the standard Hermitian product on $\mathbb{C}^{2dkp}$ to $\langle \cdot, \cdot \rangle_{p,\theta}$. By means of this isomorphism, we can see $\mathcal{H}_{p,\theta}$ as the complex linear endomorphism $H_{p,\theta}$ of $\mathbb{C}^{2dkp}$ given by $H_{p,\theta} \boldsymbol{Z} = \boldsymbol{Z}'$, where $X'_1, ..., X'_{kp-1}, Y'_0, ..., Y'_{p-2}$ are defined as in (4), while

$$\begin{aligned} X'_0 &= \overline{\theta}\, Y_{kp-1} - Y_0 + A_{k-1} X_0 + \overline{\theta}\, B^T_{k-1} Y_{kp-1}, \\ Y'_{kp-1} &= \theta\, X_0 - X_{kp-1} + \theta\, B_{k-1} X_0 + C_{k-1} Y_{kp-1}. \end{aligned}$$

The difference with respect to (4) is that there are some coefficients θ or $\overline{\theta}$ appearing, according to the fact that the sequences in $\mathbb{V}_{p,\theta}$ are kp-periodic only after "twisting" them by θ. If we see $H_{p,\theta}$ as a $2dkp \times 2dkp$ complex matrix, the above expressions readily imply that $H^*_{p,\theta} = H_{p,\theta}$.

In the following, we will refer to $H_{p,\theta}$ as to the **θ-Hessian** of F_p at $\boldsymbol{z}^p$. We generalize the Morse indices and the nullity by introducing the following **Bott indices**

$$\begin{aligned}
\mathrm{ind}_\theta(\boldsymbol{z}^p) &= \sum_{\lambda<0} \dim_{\mathbb{C}} \ker(\mathcal{H}_{p,\theta} - \lambda I) = \sum_{\lambda<0} \dim_{\mathbb{C}} \ker(H_{p,\theta} - \lambda I),\\
\mathrm{coind}_\theta(\boldsymbol{z}^p) &= \sum_{\lambda>0} \dim_{\mathbb{C}} \ker(\mathcal{H}_{p,\theta} - \lambda I) = \sum_{\lambda>0} \dim_{\mathbb{C}} \ker(H_{p,\theta} - \lambda I),\\
\mathrm{nul}_\theta(\boldsymbol{z}^p) &= \dim_{\mathbb{C}} \ker \mathcal{H}_{p,\theta} = \dim_{\mathbb{C}} \ker H_{p,\theta}.
\end{aligned}$$

The usual Morse indices correspond to the case where $\theta = 1$, that is,

$$\begin{aligned}
\mathrm{ind}(\boldsymbol{z}^p) &= \mathrm{ind}_1(\boldsymbol{z}^p),\\
\mathrm{coind}(\boldsymbol{z}^p) &= \mathrm{coind}_1(\boldsymbol{z}^p),\\
\mathrm{nul}(\boldsymbol{z}^p) &= \mathrm{nul}_1(\boldsymbol{z}^p).
\end{aligned}$$

The first elementary properties of the Bott indices are the following.

Lemma 2.1.

(i) *The functions $\theta \mapsto \mathrm{ind}_\theta(\boldsymbol{z}^p)$, $\theta \mapsto \mathrm{coind}_\theta(\boldsymbol{z}^p)$ and $\theta \mapsto \mathrm{nul}_\theta(\boldsymbol{z}^p)$ are invariant by complex conjugation.*

(ii) $\mathrm{nul}_\theta(\boldsymbol{z}^p) = \dim_{\mathbb{C}} \ker(\mathrm{d}\phi^p(z_0) - \theta I)$.

(iii) *The functions $\theta \mapsto \mathrm{ind}_\theta(\boldsymbol{z}^p)$ and $\theta \mapsto \mathrm{coind}_\theta(\boldsymbol{z}^p)$ are locally constant on $S^1 \setminus \sigma(\mathrm{d}\phi^p(z_0))$, the complement of the set of eigenvalues of $\mathrm{d}\phi^p(z_0)$ on the unit circle. Given an open interval $U \subset S^1$ such that the intersection $U \cap \sigma(\mathrm{d}\phi^p(z_0))$ contains only one point θ, for all $\theta' \in U \setminus \{\theta\}$ we have*

$$\underbrace{\mathrm{ind}_{\theta'}(\boldsymbol{z}^p) - \mathrm{ind}_\theta(\boldsymbol{z}^p)}_{\geq 0} + \underbrace{\mathrm{coind}_{\theta'}(\boldsymbol{z}^p) - \mathrm{coind}_\theta(\boldsymbol{z}^p)}_{\geq 0} = \mathrm{nul}_\theta(\boldsymbol{z}^p).$$

Proof. Point (i) is an immediate consequence of the fact that $\overline{H_{p,\theta}} = H_{p,\overline{\theta}}$.

As for point (ii), notice that a vector $\boldsymbol{Z} = (Z_0, ..., Z_{kp-1})$ belongs to the kernel of $H_{p,\theta}$ if and only if it satisfies, for all $j = 0, ..., kp-2$,

$$\begin{aligned}
X_{j+1} - X_j &= -B_{j \bmod k} X_{j+1} - C_{j \bmod k} Y_j,\\
Y_{j+1} - Y_j &= A_{j \bmod k} X_{j+1} + B^T_{j \bmod k} Y_j,
\end{aligned}$$

and

$$\begin{aligned}
\theta\, X_0 &= X_{kp-1} - B_{k-1}\theta\, X_0 - C_{k-1} Y_{kp-1},\\
\theta\, Y_0 &= Y_{kp-1} + A_{k-1}\theta\, X_0 + B^T_{k-1} Y_{pk-1}.
\end{aligned}$$

We already saw in Remark 1.1 that the quadratic function $\tilde{f}_j$ is the generating function for the linearized map $\mathrm{d}\phi_j(z_j)$. Therefore, we can rephrase the above conditions by saying that a vector $\boldsymbol{Z} = (Z_0, ..., Z_{kp-1})$ belongs to the kernel of $H_{p,\theta}$ if and only if $\mathrm{d}\phi_{j \bmod k}(z_{j \bmod k}) Z_j = Z_{j+1}$ for all $j = 0, ..., kp-2$ and $\mathrm{d}\phi_{k-1}(z_{k-1}) Z_{kp-1} = \theta\, Z_0$. The projection $\boldsymbol{Z} \mapsto Z_0$ is thus a diffeomorphism between the kernel of $H_{p,\theta}$ and the kernel of $\mathrm{d}\phi^p(z_0) - \theta I$.

Point (iii) is a consequence of the continuity of the function that associates to a matrix his set of eigenvalues. Let us explain this in detail. First of all, since the matrix $H_{p,\theta}$ is Hermitian, it is diagonalizable. In particular $\dim_{\mathbb{C}} \ker(H_{p,\theta} - \lambda I)$ is equal to the algebraic multiplicity of λ as an eigenvalue of $H_{p,\theta}$ (which is understood to be zero if λ is not an eigenvalue). Fix an arbitrary $\theta \in S^1$. For an open interval $U \subset S^1$ containing θ, there exist a continuous function

$$\boldsymbol{\lambda} = (\lambda_1, \lambda_2, ..., \lambda_{2dkp}) : U \to \mathbb{R}^{2dkp}$$

such that, for all $\theta' \in U$, the numbers $\lambda_1(\theta'), \lambda_2(\theta'), ..., \lambda_{2dkp}(\theta')$ are the eigenvalues of $H_{p,\theta'}$ repetead according to their algebraic multiplicity. In particular, we have

$$\begin{aligned}
\mathrm{ind}_{\theta'}(\boldsymbol{z}^p) &= \#\{j \mid \lambda_j(\theta') < 0\}, \\
\mathrm{coind}_{\theta'}(\boldsymbol{z}^p) &= \#\{j \mid \lambda_j(\theta') > 0\}, \\
\mathrm{nul}_{\theta'}(\boldsymbol{z}^p) &= \#\{j \mid \lambda_j(\theta') = 0\}.
\end{aligned}$$

This immediately implies that, if $\mathrm{nul}_\theta(\boldsymbol{z}^p) = 0$, the function $\theta' \mapsto \mathrm{ind}_{\theta'}(\boldsymbol{z}^p)$ is constant in a neighborhood of θ. Assume now that $\mathrm{nul}_\theta(\boldsymbol{z}^p) > 0$, and shrink U around θ so that it does not contains other eigenvalues of $\mathrm{d}\phi^p(z_0)$. In particular, the sign of each function λ_j is locally constant on $U \setminus \{\theta\}$. Therefore, the difference $\mathrm{ind}_{\theta'}(\boldsymbol{z}^p) - \mathrm{ind}_\theta(\boldsymbol{z}^p)$ is precisely the number of subscripts j such that $\lambda_j(\theta') < 0$ and $\lambda_j(\theta) = 0$. Analogously, $\mathrm{coind}_{\theta'}(\boldsymbol{z}^p) - \mathrm{coind}_\theta(\boldsymbol{z}^p)$ is the number of subscripts j such that $\lambda_j(\theta') > 0$ and $\lambda_j(\theta) = 0$. Finally, $\mathrm{nul}_{\theta'}(\boldsymbol{z}^p) = 0$ for all $\theta' \in U \setminus \{\theta\}$. This proves point (iii). □

As we mentioned earlier, the reason for introducing the Bott indices is that the function $\theta \mapsto \mathrm{ind}_\theta(\boldsymbol{z})$ alone determines the iterated index $\mathrm{ind}(\boldsymbol{z}^p)$ for all periods $p \in \mathbb{N}$, and the same property holds for the coindices and the nullities. The precise way this works is explained by the following lemma.

Lemma 2.2 (Bott's formulae). *For all $p \in \mathbb{N}$ and $\theta \in S^1$, we have*

$$\begin{aligned}
\mathrm{nul}_\theta(\boldsymbol{z}^p) &= \sum_{\mu \in \sqrt[p]{\theta}} \mathrm{nul}_\mu(\boldsymbol{z}), \\
\mathrm{ind}_\theta(\boldsymbol{z}^p) &= \sum_{\mu \in \sqrt[p]{\theta}} \mathrm{ind}_\mu(\boldsymbol{z}), \\
\mathrm{coind}_\theta(\boldsymbol{z}^p) &= \sum_{\mu \in \sqrt[p]{\theta}} \mathrm{coind}_\mu(\boldsymbol{z}).
\end{aligned}$$

Proof. The first equality follows from a general property of matrices. Indeed, by Lemma 2.1(ii), such an equality can be rewritten as

$$\dim_{\mathbb{C}} \ker(\mathrm{d}\phi^p(z_0) - \theta I) = \sum_{\mu \in \sqrt[p]{\theta}} \dim_{\mathbb{C}} \ker(\mathrm{d}\phi(z_0) - \mu I),$$

which follows from Proposition A.1.

Now, we are going to provide an argument that proves the three equalities of the lemma at once. Indeed, we will show that

$$\dim_{\mathbb{C}} \ker(\mathcal{H}_{p,\theta} - \lambda I) = \sum_{\mu \in \sqrt[p]{\theta}} \dim_{\mathbb{C}} \ker(\mathcal{H}_{1,\mu} - \lambda I), \qquad \forall \lambda \in \mathbb{R}. \tag{5}$$

For this, we need an ingredient from elementary Fourier analysis. Notice first that $\mathbb{V}_{1,\mu}$ is a vector subspace of $\mathbb{V}_{p,\theta}$ whenever $\mu^p = \theta$. Any sequence of complex vectors $\boldsymbol{Z} = (Z_j)_{j\in\mathbb{Z}} \in \mathbb{V}_{p,\theta}$ can be decomposed as

$$\boldsymbol{Z} = \sum_{\mu \in \sqrt[p]{\theta}} \boldsymbol{Z}_\mu, \tag{6}$$

where $\boldsymbol{Z}_\mu = (Z_{\mu,j})_{j\in\mathbb{Z}} \in \mathbb{V}_{1,\mu}$ is given by

$$Z_{\mu,j} := \frac{1}{kp} \sum_{h=0}^{kp-1} \mu^{1-h} Z_{h+j}.$$

Given two distinct roots $\mu, \sigma \in \sqrt[p]{\theta}$, the corresponding vector spaces $\mathbb{V}_{1,\mu}$ and $\mathbb{V}_{1,\sigma}$ are orthogonal with respect to the Hermitian product $\langle\cdot,\cdot\rangle_{p,\theta}$. Indeed, if $\boldsymbol{Z}' \in \mathbb{V}_{1,\mu}$ and $\boldsymbol{Z}'' \in \mathbb{V}_{1,\sigma}$, we have

$$\langle \boldsymbol{Z}', \boldsymbol{Z}''\rangle_{p,\theta} = \sum_{j=0}^{kp-1} Z_j' \overline{Z_j''} = \sum_{j=0}^{k-1} Z_j' \overline{Z_j''} \underbrace{\sum_{h=0}^{p-1} (\mu\overline{\sigma})^h}_{=0} = 0.$$

This readily implies that the decomposition (6) is unique, and defines a $\langle\cdot,\cdot\rangle_{p,\theta}$-orthogonal splitting

$$\mathbb{V}_{p,\theta} = \bigoplus_{\mu \in \sqrt[p]{\theta}} \mathbb{V}_{1,\mu}.$$

Actually, this splitting turns out to be orthogonal also with respect to the Hermitian form $\langle \mathcal{H}_{p,\theta}\cdot,\cdot\rangle_{p,\theta}$. Indeed, $\mathcal{H}_{p,\theta}|_{\mathbb{V}_{1,\mu}} = \mathcal{H}_{1,\mu}$ and, if $\boldsymbol{Z}'$ and $\boldsymbol{Z}''$ are as above, we have

$$\langle \mathcal{H}_{p,\theta}\boldsymbol{Z}', \boldsymbol{Z}''\rangle_{p,\theta} = \langle \underbrace{\mathcal{H}_{1,\mu}\boldsymbol{Z}'}_{\in \mathbb{V}_{1,\mu}}, \boldsymbol{Z}''\rangle_{p,\theta} = 0.$$

In particular, the λ-eigenspace of $\mathcal{H}_{p,\theta}$ is the direct sum of the λ-eigenspaces of the operators $\mathcal{H}_{1,\mu}$, for all $\mu \in \sqrt[p]{\theta}$, and equation (5) follows. □

Lemmata 2.1 and 2.2 give a clear picture of the qualitative behavior of the functions $p \mapsto \mathrm{ind}(\boldsymbol{z}^p)$ and $p \mapsto \mathrm{coind}(\boldsymbol{z}^p)$. In particular, they imply that the quantities

$$\begin{aligned} \overline{\mathrm{ind}}(\boldsymbol{z}) &:= \frac{1}{2\pi}\int_0^{2\pi} \mathrm{ind}_{e^{it}}(\boldsymbol{z})\,\mathrm{d}t, \\ \overline{\mathrm{coind}}(\boldsymbol{z}) &:= \frac{1}{2\pi}\int_0^{2\pi} \mathrm{coind}_{e^{it}}(\boldsymbol{z})\,\mathrm{d}t \end{aligned} \tag{7}$$

are always finite, and we have

$$\overline{\mathrm{ind}}(\boldsymbol{z}) = \lim_{p\to\infty} \frac{\mathrm{ind}(\boldsymbol{z}^p)}{p}, \tag{8}$$

$$\overline{\mathrm{coind}}(\boldsymbol{z}) = \lim_{p\to\infty} \frac{\mathrm{coind}(\boldsymbol{z}^p)}{p}. \tag{9}$$

In the following, we will refer to $\overline{\mathrm{ind}}(\boldsymbol{z})$ and $\overline{\mathrm{coind}}(\boldsymbol{z})$ respectively as to the **average Morse index** and **coindex** of the critical point $\boldsymbol{z}$. Notice that, by the conjugacy-invariance of the function $\theta \mapsto \mathrm{ind}_\theta(\boldsymbol{z})$, in the above expressions (7) we can replace

2π by π, that is, we can equivalently average the index functions on the upper semi-circle. Equations (8) and Proposition 1.2 imply that

$$dk/2 \leq \overline{\mathrm{ind}}(\boldsymbol{z}) \leq 2dk,$$
$$dk/2 \leq \overline{\mathrm{coind}}(\boldsymbol{z}) \leq 2dk.$$

Since $\mathrm{ind}(\boldsymbol{z}^p) + \mathrm{coind}(\boldsymbol{z}^p) + \mathrm{nul}(\boldsymbol{z}^p) = 2dkp$, we further have

$$\overline{\mathrm{ind}}(\boldsymbol{z}) + \overline{\mathrm{coind}}(\boldsymbol{z}) = 2dk.$$

Another property of the average indices that follows immediately from their definitions is that

$$\overline{\mathrm{ind}}(\boldsymbol{z}^p) = p\,\overline{\mathrm{ind}}(\boldsymbol{z}),$$
$$\overline{\mathrm{coind}}(\boldsymbol{z}^p) = p\,\overline{\mathrm{coind}}(\boldsymbol{z}).$$

Now, we are going to find optimal bounds from the gap between the average and the actual Morse indices. Such bounds plays an essential role in the multiplicity problem for periodic points of Hamiltonian diffeomorphisms (see Section 2.5). For now, we can only deal with the non-degenerate situation (Theorem 2.3 will be superseded by the general Theorem 2.10). We recall that d is the half-dimension of the domain of our Hamiltonian diffeomorphism ϕ.

Theorem 2.3. *Assume that $\boldsymbol{z}$ is a non-degenerate critical point of F_1, i.e.* $\mathrm{nul}(\boldsymbol{z}) = 0$. *Then* $|\overline{\mathrm{ind}}(\boldsymbol{z}) - \mathrm{ind}(\boldsymbol{z})| < d$ *and* $|\overline{\mathrm{coind}}(\boldsymbol{z}) - \mathrm{coind}(\boldsymbol{z})| < d$.

Proof. We will provide the proof for the Morse index, the one for the coindex being identical. For any eigenvalue on the unit circle $\theta \in \sigma(\mathrm{d}\phi(z_0)) \cap S^1$, let $\epsilon > 0$ be a small enough quantity so that $\sigma(\mathrm{d}\phi(z_0)) \cap S^1$ does not contain other eigenvalues with arguments in the interval $[\arg(\theta) - \epsilon, \arg(\theta) + \epsilon]$. We set $\theta^\pm := \theta e^{\pm i\epsilon}$. For all $\mu \in S^1$ with $\mathrm{Im}(\mu) > 0$, we denote by σ_μ the (possibly empty) set of eigenvalues of $\mathrm{d}\phi(z_0)$ on the unit circle with argument in the open interval $(0, \arg(\mu))$, and we define

$$f(\mu) := \sum_{\theta \in \sigma_\mu} \big(\mathrm{ind}_{\theta^+}(\boldsymbol{z}) - \mathrm{ind}_{\theta^-}(\boldsymbol{z})\big).$$

By its definition, the function f is piecewise constant. By Lemma 2.1(iii), if μ is not an eigenvalue of $\mathrm{d}\phi(z_0)$, we have

$$\mathrm{ind}_\mu(\boldsymbol{z}) - \mathrm{ind}(\boldsymbol{z}) = f(\mu).$$

By integrating this equality in μ on the upper semi-circle, we obtain

$$\begin{aligned}\overline{\mathrm{ind}}(\boldsymbol{z}) - \mathrm{ind}(\boldsymbol{z}) &= \frac{1}{\pi}\int_0^\pi \big(\mathrm{ind}_{e^{it}}(\boldsymbol{z}) - \mathrm{ind}(\boldsymbol{z})\big)\,\mathrm{d}t \\ &= \frac{1}{\pi}\int_0^\pi f(e^{it})\,\mathrm{d}t,\end{aligned}$$

By the equality in Lemma 2.1(iii), for all $t \in (0, \pi)$ we can estimate

$$|f(e^{it})| \leq \sum_{\theta \in \sigma_{\exp(it)}} \mathrm{nul}_\theta(\boldsymbol{z}) \leq \sum_{\theta \in S^1 \cap \{\mathrm{Im} > 0\}} \mathrm{nul}_\theta(\boldsymbol{z}) \leq \frac{1}{2}\sum_{\theta \in S^1} \mathrm{nul}_\theta(\boldsymbol{z}) \leq d.$$

Let $\delta > 0$ be such that there is no eigenvalue of $\mathrm{d}\phi(z_0)$ on the unit circle with argument in $[0, \delta]$. In particular, the function $t \mapsto f(e^{it})$ is zero on the interval $[0, \delta]$. Therefore, we conclude

$$\begin{aligned}|\overline{\mathrm{ind}}(\boldsymbol{z}) - \mathrm{ind}(\boldsymbol{z})| &= \left|\frac{1}{\pi}\int_\delta^\pi f(e^{it})\,\mathrm{d}t\right| \\ &\leq \frac{1}{\pi}\int_\delta^\pi |f(e^{it})|\,\mathrm{d}t \\ &\leq \frac{\pi-\delta}{\pi}\,d \\ &< d.\end{aligned}$$

□

2.2. **Splitting numbers.** The generalization of Theorem 2.3 to the degenerate situation requires new ingredients, which incidentally will shed some light on the dependence of the Morse index of the critical point associated to a fixed point $z_0 \in \mathrm{fix}(\phi)$ from the specific generating family employed (this dependence will be explored further in Section 3).

Since in this section we will work in the fixed period $p = 1$, in order to ease the notation we will drop it from all appearing symbols, thus writing H_θ for the θ-Hessian $H_{1,\theta}$. We will denote by $h_\theta : \mathbb{C}^{2dk} \times \mathbb{C}^{2dk} \to \mathbb{C}$ the Hermitian bilinear form associated to H_θ, i.e.

$$h_\theta(\boldsymbol{Z}, \boldsymbol{Z}') = \langle H_\theta \boldsymbol{Z}, \boldsymbol{Z}'\rangle.$$

We consider the vector subspace

$$\mathbb{V} := \big\{\boldsymbol{Z} = (Z_0, ..., Z_{k-1}) \in \mathbb{C}^{2dk} \ \big|\ X_0 = 0\big\},$$

where, as before, we write $Z_j = (X_j, Y_j)$. We will reduce the computation of the inertia of h_θ to the inertia of its restrictions to $\mathbb{V}$ and to its h_θ-orthogonal space $\mathbb{V}^{h_\theta}$ by means of Propositions A.2 and A.3, which give

$$\begin{aligned}\mathrm{ind}(h_\theta) &= \mathrm{ind}(h_\theta|_{\mathbb{V}\times\mathbb{V}}) + \mathrm{ind}(h_\theta|_{\mathbb{V}^{h_\theta}\times\mathbb{V}^{h_\theta}}) \\ &\quad + \dim_{\mathbb{C}}(\mathbb{V}\cap\mathbb{V}^{h_\theta}) - \dim_{\mathbb{C}}(\mathbb{V}\cap\ker(H_\theta)),\end{aligned} \tag{10}$$

$$\begin{aligned}\mathrm{coind}(h_\theta) &= \mathrm{coind}(h_\theta|_{\mathbb{V}\times\mathbb{V}}) + \mathrm{coind}(h_\theta|_{\mathbb{V}^{h_\theta}\times\mathbb{V}^{h_\theta}}) \\ &\quad + \dim_{\mathbb{C}}(\mathbb{V}\cap\mathbb{V}^{h_\theta}) - \dim_{\mathbb{C}}(\mathbb{V}\cap\ker(H_\theta)),\end{aligned} \tag{11}$$

$$\mathrm{nul}(h_\theta) = \mathrm{nul}(h_\theta|_{\mathbb{V}^{h_\theta}\times\mathbb{V}^{h_\theta}}) - \dim_{\mathbb{C}}(\mathbb{V}\cap\mathbb{V}^{h_\theta}) + \dim_{\mathbb{C}}(\mathbb{V}\cap\ker(H_\theta)). \tag{12}$$

We refer the reader to Appendix A.2 for the terminology and the notation concerning Hermitian forms. The restriction of h_θ to $\mathbb{V}$ is independent of θ. Indeed, for all

$\boldsymbol{Z}, \boldsymbol{Z}' \in \mathbb{V}$, we have

$$
\begin{aligned}
h_\theta(\boldsymbol{Z}, \boldsymbol{Z}') =& \langle \overline{\theta}\, Y_{k-1} - Y_0 + A_{k-1} X_0 + \overline{\theta}\, B_{k-1}^T Y_{k-1}, X_0' \rangle \\
&+ \langle \theta\, X_0 - X_{k-1} + \theta\, B_{k-1} X_0 + C_{k-1} Y_{k-1}, Y_{k-1}' \rangle \\
&+ \sum_{j=1}^{k-1} \langle Y_{j-1} - Y_j + A_{j-1} X_j + B_{j-1}^T Y_{j-1}, X_j' \rangle \\
&+ \sum_{j=0}^{k-2} \langle X_{j+1} - X_j + B_j X_{j+1} + C_j Y_j, Y_j' \rangle \\
=& \langle -X_{k-1} + C_{k-1} Y_{k-1}, Y_{k-1}' \rangle + \langle X_1 + B_0 X_1 + C_0 Y_0, Y_0' \rangle \\
&+ \sum_{j=1}^{k-1} \langle Y_{j-1} - Y_j + A_{j-1} X_j + B_{j-1}^T Y_{j-1}, X_j' \rangle \\
&+ \sum_{j=1}^{k-2} \langle X_{j+1} - X_j + B_j X_{j+1} + C_j Y_j, Y_j' \rangle.
\end{aligned}
$$

In particular, the inertia indices $\mathrm{ind}(h_\theta|_{\mathbb{V}\times\mathbb{V}})$ and $\mathrm{coind}(h_\theta|_{\mathbb{V}\times\mathbb{V}})$ are independent of $\theta \in S^1$. The orthogonal space $\mathbb{V}^{h_\theta}$ contains precisely the vectors $\boldsymbol{Z} \in \mathbb{C}^{2dk}$ such that

$$
\begin{aligned}
\theta\, X_0 - X_{k-1} + \theta\, B_{k-1} X_0 + C_{k-1} Y_{k-1} &= 0, \\
X_{j+1} - X_j + B_j X_{j+1} + C_j Y_j &= 0, \qquad && \forall j = 0, ..., k-2, \\
Y_{j-1} - Y_j + A_{j-1} X_j + B_{j-1}^T Y_{j-1} &= 0, && \forall j = 1, ..., k-1.
\end{aligned}
$$

This means that, if we set $P_j := \mathrm{d}\phi_j(z_j)$ for all $j = 0, ..., k-1$,

$$
\mathbb{V}^{h_\theta} = \left\{ (Z_0, ..., Z_{k-1}) \in \mathbb{C}^{2dk} \;\middle|\; \begin{array}{l} P_j Z_j = Z_{j+1} \quad \forall j = 0, ..., k-2 \\ P_{k-1} Z_{k-1} = (\theta X_0, \tilde{Y}_k) \text{ for some } \tilde{Y}_k \in \mathbb{C}^d \end{array} \right\}.
$$

In particular, $\mathbb{V}^{h_\theta}$ is isomorphic to $(\mathrm{d}\phi(z_0) - \theta I)^{-1}(\{0\} \times \mathbb{C}^d)$ via the isomorphism $\boldsymbol{Z} \mapsto Z_0$. Therefore, its dimension is bounded as

$$
\dim_{\mathbb{C}} \mathbb{V}^{h_\theta} \leq d + \dim_{\mathbb{C}} \ker(\mathrm{d}\phi(z_0) - \theta I).
$$

The intersection $\mathbb{V} \cap \mathbb{V}^{h_\theta}$ is equal to

$$
\mathbb{V} \cap \mathbb{V}^{h_\theta} = \left\{ (Z_0, ..., Z_{k-1}) \in \mathbb{C}^{2dk} \;\middle|\; \begin{array}{l} X_0 = 0 \\ P_j Z_j = Z_{j+1} \quad \forall j = 0, ..., k-2 \\ P_{k-1} Z_{k-1} = (0, \tilde{Y}_k) \text{ for some } \tilde{Y}_k \in \mathbb{C}^d \end{array} \right\}.
$$

In particular, it is independent of θ. The intersection $\mathbb{V} \cap \ker H_\theta$ is equal to

$$
\mathbb{V} \cap \ker H_\theta = \left\{ (Z_0, ..., Z_{k-1}) \in \mathbb{C}^{2dk} \;\middle|\; \begin{array}{l} X_0 = 0 \\ P_j Z_j = Z_{j+1} \quad \forall j = 0, ..., k-2 \\ P_{k-1} Z_{k-1} = (0, \theta Y_0) \end{array} \right\}.
$$

Therefore, the map $\boldsymbol{Z} \mapsto Z_0 = (0, Y_0)$ is an isomorphism between $\mathbb{V} \cap \ker H_\theta$ and $\ker(\mathrm{d}\phi(z_0) - \theta I) \cap (\{0\} \times \mathbb{C}^d)$. In particular

$$
\dim_{\mathbb{C}}(\mathbb{V} \cap \ker H_\theta) = \dim_{\mathbb{C}} \big(\ker(\mathrm{d}\phi(z_0) - \theta I) \cap (\{0\} \times \mathbb{C}^d) \big).
$$

Let us now have a look at the restriction of the Hermitian form h_θ to $\mathbb{V}^{h_\theta}$. For all $\boldsymbol{Z}, \boldsymbol{Z}' \in \mathbb{V}^{h_\theta}$, we have

$$\begin{aligned} h_\theta(\boldsymbol{Z}, \boldsymbol{Z}') &= \langle \overline{\theta} Y_{k-1} - Y_0 + \underbrace{A_{k-1} X_0 + \overline{\theta} B_{k-1}^T Y_{k-1}}_{\overline{\theta}\tilde{Y}_k - \overline{\theta} Y_{k-1}}, X_0' \rangle \\ &= \langle \overline{\theta} \tilde{Y}_k - Y_0, X_0' \rangle \\ &= \omega((I - \overline{\theta}\, \mathrm{d}\phi(z_0)) Z_0, Z_0'), \end{aligned}$$

where ω denotes the Hermitian extension of the standard symplectic form on $\mathbb{R}^{2d}$, given by $\omega(Z, Z') = \langle X, Y' \rangle - \langle Y, X' \rangle$. Summing up, we have shown that $\mathrm{ind}(h_\theta|_{\mathbb{V}\times\mathbb{V}})$ and $\dim_{\mathbb{C}}(\mathbb{V} \cap \mathbb{V}^{h_\theta})$ are independent of θ, while $\dim_{\mathbb{C}}(\mathbb{V} \cap \ker(H_\theta))$ and $\mathrm{ind}(h_\theta|_{\mathbb{V}^{h_\theta}\times\mathbb{V}^{h_\theta}})$ are completely determined by the linearized map $\mathrm{d}\phi(z_0)$. This, together with equations (10) and (11), implies the following.

Lemma 2.4. *The functions $\theta \mapsto \mathrm{ind}_\theta(\boldsymbol{z}) = \mathrm{ind}(h_\theta)$ and $\theta \mapsto \mathrm{coind}_\theta(\boldsymbol{z}) = \mathrm{coind}(h_\theta)$ are completely determined by the linearized map $P := \mathrm{d}\phi(z_0) \in \mathrm{Sp}(2d)$ up to additive constants.* □

We call **splitting numbers** of the linearized map P at $\theta \in S^1$ the two quantities

$$\begin{aligned} \mathrm{S}_P^+(\theta) &= \mathrm{ind}_{\theta^+}(\boldsymbol{z}) - \mathrm{ind}_\theta(\boldsymbol{z}) = \mathrm{ind}(h_{\theta^+}) - \mathrm{ind}(h_\theta), \\ \mathrm{S}_P^-(\theta) &= \mathrm{ind}_{\theta^-}(\boldsymbol{z}) - \mathrm{ind}_\theta(\boldsymbol{z}) = \mathrm{ind}(h_{\theta^-}) - \mathrm{ind}(h_\theta), \end{aligned} \tag{13}$$

where $\theta^\pm = \theta e^{\pm i\epsilon}$, and $\epsilon > 0$ is sufficiently small so that $\sigma(P) \cap S^1$ does not contain eigenvalues with arguments in $[\arg(\theta) - \epsilon, \arg(\theta)) \cup (\arg(\theta), \arg(\theta) + \epsilon]$. Lemma 2.4 guarantees that $\mathrm{S}_P^\pm$ is a good notation: the splitting numbers only depend on the linearized map $P \in \mathrm{Sp}(2d)$. Namely, if ϕ' is another Hamiltonian diffeomorphism of $\mathbb{R}^{2d}$ with a fixed point z_0' and the same linearized map $P = \mathrm{d}\phi(z_0) = \mathrm{d}\phi'(z_0')$, given a generating family F' for ϕ', the splitting numbers functions associated to the θ-Hessian of F' at the critical point corresponding to z_0' are still $\mathrm{S}_P^\pm$. By replacing indices with coindices in (13), we can define the **cosplitting numbers**

$$\begin{aligned} \mathrm{coS}_P^+(\theta) &= \mathrm{coind}_{\theta^+}(\boldsymbol{z}) - \mathrm{coind}_\theta(\boldsymbol{z}) = \mathrm{coind}(h_{\theta^+}) - \mathrm{coind}(h_\theta), \\ \mathrm{coS}_P^-(\theta) &= \mathrm{coind}_{\theta^-}(\boldsymbol{z}) - \mathrm{coind}_\theta(\boldsymbol{z}) = \mathrm{coind}(h_{\theta^-}) - \mathrm{coind}(h_\theta), \end{aligned}$$

which possess analogous properties. The equality in Lemma 2.1(iii) can be rewritten as

$$\underbrace{\mathrm{S}^\pm(\theta)}_{\geq 0} + \underbrace{\mathrm{coS}^\pm(\theta)}_{\geq 0} = \dim_{\mathbb{C}} \ker(P - \theta I), \qquad \forall \theta \in S^1. \tag{14}$$

Warning 2.5. Many authors in symplectic topology use a different sign convention, and thus call splitting numbers what we call cosplitting numbers. The convention adopted in a paper can be easily checked on Example 2.12. See also Warning 3.2 in the next section. □

Remark 2.6. The splitting and cosplitting numbers can be defined for any symplectic matrix $P \in \mathrm{Sp}(2d)$. Indeed, the symplectic group $\mathrm{Sp}(2d)$ is connected, and therefore the map $\phi(z) = Pz$ is a Hamiltonian diffeomorphism such that $\mathrm{d}\phi(0) = P$. □

We will now strengthen Lemma 2.4 as follows.

Lemma 2.7. *The splitting and cosplitting numbers* $\mathrm{S}^{\pm}_{P}(\theta)$ *and* $\mathrm{coS}^{\pm}_{P}(\theta)$ *only depend on the conjugacy class of* P *in the symplectic group: for all* $Q \in \mathrm{Sp}(2d)$*, we have*

$$\mathrm{S}^{\pm}_{P}(\theta) = \mathrm{S}^{\pm}_{QPQ^{-1}}(\theta),$$
$$\mathrm{coS}^{\pm}_{P}(\theta) = \mathrm{coS}^{\pm}_{QPQ^{-1}}(\theta).$$

Proof. Since the symplectic group is connected, there exists a smooth path of symplectic matrices $Q_t \in \mathrm{Sp}(2d)$ such that $Q_0 = I$ and $Q_1 = Q$. We set $\phi^t(z) := Q_t P Q_t^{-1} z$, and we consider $t \mapsto \phi^t$ as a smooth path of Hamiltonian diffeomorphisms of $\mathbb{R}^{2d}$. For $k \in \mathbb{N}$ large enough, there exists a smooth homotopy

$$F^t : \mathbb{R}^{2dk} \to \mathbb{R}, \qquad t \in [0,1],$$

F^t being the quadratic generating family of ϕ^t. The origin $0 \in \mathbb{R}^{2dk}$ is the critical point of F^t corresponding to the fixed point $0 \in \mathbb{R}^{2d}$ of ϕ^t. For all $\theta \in S^1$, we denote by H^t_θ the θ-Hessian of F^t at the origin, and by h^t_θ the associated Hermitian form

$$h^t_\theta(\boldsymbol{Z}, \boldsymbol{Z}') = \langle H^t_\theta \boldsymbol{Z}, \boldsymbol{Z}' \rangle.$$

Notice that H^t_θ depends smoothly on $(t,\theta) \in [0,1] \times S^1$, and that

$$\ker H^t_\theta = \ker(Q_t P Q_t^{-1} - \theta I) = \ker(Q_t (P - \theta I) Q_t^{-1}).$$

In particular, the function $t \mapsto \dim_{\mathbb{C}} \ker H^t_\theta$ is constant. This readily implies that the functions $t \mapsto \mathrm{ind}(h^t_\theta)$ and $t \mapsto \mathrm{coind}(h^t_\theta)$ are constant as well, and therefore

$$\mathrm{S}^{\pm}_{P}(\theta) = \mathrm{ind}(h^0_{\theta\pm}) - \mathrm{ind}(h^0_\theta) = \mathrm{ind}(h^1_{\theta\pm}) - \mathrm{ind}(h^1_\theta) = \mathrm{S}^{\pm}_{QPQ^{-1}}(\theta),$$
$$\mathrm{coS}^{\pm}_{P}(\theta) = \mathrm{coind}(h^0_{\theta\pm}) - \mathrm{coind}(h^0_\theta) = \mathrm{coind}(h^1_{\theta\pm}) - \mathrm{coind}(h^1_\theta) = \mathrm{coS}^{\pm}_{QPQ^{-1}}(\theta).$$

□

Consider two positive integers d', d'', and set $d := d' + d''$. We identify $\mathbb{R}^{2d'}$ with the symplectic subspace $\mathbb{R}^{2d'} \times \{0\} \subset \mathbb{R}^{2d}$, and $\mathbb{R}^{2d''}$ with the symplectic subspace $\{0\} \times \mathbb{R}^{2d''} \subset \mathbb{R}^{2d}$. Given two symplectic matrices $P' \in \mathrm{Sp}(2d')$ and $P'' \in \mathrm{Sp}(2d'')$, their direct sum is the symplectic matrix $P = P' \oplus P'' \in \mathrm{Sp}(2d)$ given by $P(\boldsymbol{z}', \boldsymbol{z}'') = (P'\boldsymbol{z}', P''\boldsymbol{z}'')$. The next lemma shows that the splitting and cosplitting numbers behave naturally with respect to the direct sum operation.

Lemma 2.8. *For all* $P' \in \mathrm{Sp}(2d')$ *and* $P'' \in \mathrm{Sp}(2d'')$*, we have*

$$\mathrm{S}^{\pm}_{P' \oplus P''}(\theta) = \mathrm{S}^{\pm}_{P'}(\theta) + \mathrm{S}^{\pm}_{P''}(\theta),$$
$$\mathrm{coS}^{\pm}_{P' \oplus P''}(\theta) = \mathrm{coS}^{\pm}_{P'}(\theta) + \mathrm{coS}^{\pm}_{P''}(\theta).$$

Proof. For an integer $k > 0$ large enough, we can find quadratic generating families $F' : \mathbb{R}^{2d'k} \to \mathbb{R}$ and $F'' : \mathbb{R}^{2d''k} \to \mathbb{R}$ for the matrices P' and P'' (seen as Hamiltonian diffeomorphisms of $\mathbb{R}^{2d'}$ qnd $\mathbb{R}^{2d''}$ respectively). For each $\theta \in S^1$, we denote by H'_θ and H''_θ the θ-Hessians of F' and F'' at the origin, and by h'_θ and h''_θ the associated Hermitian bilinear forms. The function $F : \mathbb{R}^{2dk} \to \mathbb{R}$ given by $F(\boldsymbol{z}', \boldsymbol{z}'') = F'(\boldsymbol{z}') + F''(\boldsymbol{z}'')$ is a quadratic generating function for the matrix $P' \oplus P''$. Its θ-Hessian at the origin is $H_\theta = H'_\theta \oplus H''_\theta$. In particular, index and coindex of the associated Hermitian form h_θ satisfy

$$\mathrm{ind}(h_\theta) = \mathrm{ind}(h'_\theta) + \mathrm{ind}(h''_\theta),$$
$$\mathrm{coind}(h_\theta) = \mathrm{coind}(h'_\theta) + \mathrm{coind}(h''_\theta).$$

This implies the lemma. □

The following statement is the last ingredient that we need in order to prove the generalization of Theorem 2.3.

Lemma 2.9. *For all $P \in \mathrm{Sp}(2d)$ and $\theta \in S^1$, we have*

$$0 \leq \mathrm{S}_P^{\pm}(\theta) \leq \min\{\dim_{\mathbb{C}} \ker(P - \theta I), d\},$$
$$0 \leq \mathrm{coS}_P^{\pm}(\theta) \leq \min\{\dim_{\mathbb{C}} \ker(P - \theta I), d\}.$$

Proof. Notice that

$$\dim_{\mathbb{C}} \ker(P - \theta I) \leq d, \qquad \forall \theta \in S^1 \setminus \{1, -1\}. \tag{15}$$

Indeed $\ker(P - \theta I)$ and $\ker(P - \overline{\theta} I)$ are vector subspaces of the same dimension (one is the complex conjugate of the other), and they have trivial intersection since $\theta \neq \overline{\theta}$. This, together with (14), implies the bound of the lemma for $\theta \notin \{1, -1\}$.

The inequality (15) does not hold for $\theta = \pm 1$ (consider, for instance, the counterexample given by $P = I$ and $\theta = 1$). The remaining bounds on the splitting numbers will be proved by equations (10) and (12), which imply

$$\mathrm{ind}(h_\theta) = \mathrm{ind}(h_\theta|_{\mathbb{V}\times\mathbb{V}}) + \mathrm{ind}(h_\theta|_{\mathbb{V}^{h_\theta}\times\mathbb{V}^{h_\theta}}) + \mathrm{nul}(h_\theta|_{\mathbb{V}^{h_\theta}\times\mathbb{V}^{h_\theta}}) - \mathrm{nul}(h_\theta).$$

We already remarked that the restricted form $h_\theta|_{\mathbb{V}\times\mathbb{V}}$ is independent of θ, and therefore so is the summand $\mathrm{ind}(h_\theta|_{\mathbb{V}\times\mathbb{V}})$ in the above equation. Clearly

$$\mathrm{ind}(h_\theta) - \mathrm{ind}(h_\theta|_{\mathbb{V}\times\mathbb{V}}) \geq 0.$$

Moreover

$$\begin{aligned} \mathrm{ind}(h_\theta|_{\mathbb{V}^{h_\theta}\times\mathbb{V}^{h_\theta}}) + \mathrm{nul}(h_\theta|_{\mathbb{V}^{h_\theta}\times\mathbb{V}^{h_\theta}}) &\leq \dim_{\mathbb{C}} \mathbb{V}^{h_\theta} \\ &\leq d + \dim_{\mathbb{C}} \ker(P - \theta I) \\ &= d + \mathrm{nul}(h_\theta). \end{aligned}$$

Therefore

$$\begin{aligned} S_P^{\pm}(\theta) &= \mathrm{ind}(h_{\theta^{\pm}}) - \mathrm{ind}(h_\theta) \\ &\leq \mathrm{ind}(h_{\theta^{\pm}}|_{\mathbb{V}^{h_{\theta^{\pm}}}\times\mathbb{V}^{h_{\theta^{\pm}}}}) + \mathrm{nul}(h_{\theta^{\pm}}|_{\mathbb{V}^{h_{\theta^{\pm}}}\times\mathbb{V}^{h_{\theta^{\pm}}}}) - \mathrm{nul}(h_{\theta^{\pm}}) \\ &\leq d. \end{aligned}$$

This completes the proof of the bound for the splitting numbers. The one for the cosplitting numbers is proved by the same argument, with indices replaced by coindices. □

2.3. **The iteration inequality.** We can finally state and prove the general iteration inequality for the Morse index of generating families. We will adopt the notation of Section 2.1, so that $F_p : \mathbb{R}^{2dkp} \to \mathbb{R}$ denotes the generating function of the iterated Hamiltonian diffeomorphism $\phi^p \in \mathrm{Ham}(\mathbb{R}^{2d})$.

Theorem 2.10 (Iteration inequalities)**.** *Let $\boldsymbol{z} = (z_0, ..., z_{k-1})$ be a critical point of the generating function F_1, and let $p \in \mathbb{N}$. Then*

$$\begin{aligned} p\,\overline{\mathrm{ind}}(\boldsymbol{z}) - d &\leq \mathrm{ind}(\boldsymbol{z}^p), \\ \mathrm{ind}(\boldsymbol{z}^p) + \mathrm{nul}(\boldsymbol{z}^p) &\leq p\,\overline{\mathrm{ind}}(\boldsymbol{z}) + d. \end{aligned} \tag{16}$$

If at least one of the above inequalities is an equality, then $\sigma(\mathrm{d}\phi(z_0)) = \{1\}$ and $\mathrm{nul}(\boldsymbol{z}^p) \geq d$. Both inequalities are equalities if and only if $\mathrm{d}\phi(z_0)^p = I$.

Remark 2.11. Since $\mathrm{ind}(\boldsymbol{z}^p)+\mathrm{coind}(\boldsymbol{z}^p)+\mathrm{nul}(\boldsymbol{z}^p)=2dkp$, the iteration inequalities (16) can be rewritten for the Morse coindex as

$$\begin{aligned} p\,\overline{\mathrm{coind}}(\boldsymbol{z})-d &\leq \mathrm{coind}(\boldsymbol{z}^p),\\ \mathrm{coind}(\boldsymbol{z}^p)+\mathrm{nul}(\boldsymbol{z}^p) &\leq p\,\overline{\mathrm{coind}}(\boldsymbol{z})+d. \end{aligned}$$ □

Proof of Theorem 2.10. Let $\mathbb{E}:=\ker(P-I)^{2d}\subset\mathbb{R}^{2d}$ be the generalized eigenspace of the eigenvalue 1 of the symplectic matrix $P:=\mathrm{d}\phi(z_0)$. This vector subspace is symplectic (by Lemma A.5) and clearly invariant by P. Let $\mathbb{E}^\omega$ be its symplectic orthogonal, that is

$$\mathbb{E}^\omega=\big\{Z\in\mathbb{R}^{2d}\;\big|\;\omega(Z,\cdot)|_{\mathbb{E}}=0\big\}.$$

The space $\mathbb{E}^\omega$ is also invariant by P. Indeed, if $Z\in\mathbb{E}^\omega$, we have

$$\omega(PZ,Z')=\omega(Z,\underbrace{P^{-1}Z'}_{\in\mathbb{E}})=0,\qquad\forall Z'\in\mathbb{E}.$$

Hence, by decomposing $\mathbb{R}^{2d}$ as the symplectic direct sum $\mathbb{E}\oplus\mathbb{E}^\omega$, the matrix P takes the form $P'\oplus P''$, where $P'=P|_{\mathbb{E}}$ and $P''=P|_{\mathbb{E}^\omega}$. Notice that P' is a unipotent matrix, i.e. $\sigma(P')=\{1\}$, while $\sigma(P'')$ does not contain 1. Therefore, by Lemmata 2.8 and 2.9, we have

$$\mathrm{S}^\pm_P(1)=\mathrm{S}^\pm_{P'}(1)\leq\tfrac{1}{2}\dim_{\mathbb{C}}\mathbb{E},\tag{17}$$

$$\mathrm{S}^\pm_P(\theta)=\mathrm{S}^\pm_{P''}(\theta)\leq\dim_{\mathbb{C}}(P''-\theta I),\qquad\forall\theta\in S^1\setminus\{1\}.\tag{18}$$

Analogously

$$\begin{aligned}&\mathrm{coS}^\pm_P(1)=\mathrm{coS}^\pm_{P'}(1)\leq\tfrac{1}{2}\dim_{\mathbb{C}}\mathbb{E},\\ &\mathrm{coS}^\pm_P(\theta)=\mathrm{coS}^\pm_{P''}(\theta)\leq\dim_{\mathbb{C}}(P''-\theta I),\qquad\forall\theta\in S^1\setminus\{1\}.\end{aligned}$$

We now proceed in a similar fashion as in Theorem 2.3 (the argument will indeed reduce to that of Theorem 2.3 if $\dim_{\mathbb{C}}\mathbb{E}=0$). For all $\mu\in S^1$ with $\mathrm{Im}(\mu)>0$, we denote by σ_μ the set of eigenvalues of P'' on the unit circle S^1 with argument in the open interval $(0,\arg(\mu))$, and we set

$$\begin{aligned} f(\mu)&:=\sum_{\theta\in\sigma_\mu}\big(\mathrm{S}^+_{P''}(\theta)-\mathrm{S}^-_{P''}(\theta)\big),\\ g(\mu)&:=\sum_{\theta\in\sigma_\mu}\big(\mathrm{coS}^+_{P''}(\theta)-\mathrm{coS}^-_{P''}(\theta)\big).\end{aligned}$$

These functions are piecewise constant, with possible jumps only at the eigenvalues of P''. By Lemma 2.1(iii) and (17), if μ is not an eigenvalue of P'', we have

$$\begin{aligned}\mathrm{ind}_\mu(\boldsymbol{z})-\mathrm{ind}(\boldsymbol{z})&=\mathrm{S}^+_P(1)+\sum_{\theta\in\sigma_\mu}\big(\mathrm{S}^+_P(\theta)-\mathrm{S}^-_P(\theta)\big)\\ &=\mathrm{S}^+_{P'}(1)+f(\mu)\\ &\leq\frac{1}{2}\dim_{\mathbb{C}}\mathbb{E}+f(\mu).\end{aligned}\tag{19}$$

By (18) we have

$$\begin{aligned} f(\mu) &\leq \sum_{\theta\in\sigma_\mu} \dim_{\mathbb{C}}(P''-\theta I) \\ &\leq \frac{1}{2}\sum_{\theta\in S^1} \dim_{\mathbb{C}}(P''-\theta I) \\ &\leq \frac{1}{2}\dim_{\mathbb{C}} \mathbb{E}^\omega. \end{aligned}$$

Analogously, we have

$$\begin{aligned} \operatorname{coind}_\mu(\boldsymbol{z}) - \operatorname{coind}(\boldsymbol{z}) &\leq \frac{1}{2}\dim_{\mathbb{C}}\mathbb{E} + g(\mu), \\ g(\mu) &\leq \frac{1}{2}\dim_{\mathbb{C}}\mathbb{E}^\omega. \end{aligned} \tag{20}$$

Notice that $\operatorname{coind}(\boldsymbol{z}) = 2dk - \operatorname{ind}(\boldsymbol{z}) - \operatorname{nul}(\boldsymbol{z})$ and, since μ is not an eigenvalue of P, $\operatorname{coind}_\mu(\boldsymbol{z}) = 2dk - \operatorname{ind}_\mu(\boldsymbol{z})$. Therefore, the inequality (20) can be rewritten as

$$\operatorname{ind}(\boldsymbol{z}) + \operatorname{nul}(\boldsymbol{z}) - \operatorname{ind}_\mu(\boldsymbol{z}) \leq \frac{1}{2}\dim_{\mathbb{C}}\mathbb{E} + g(\mu). \tag{21}$$

Let $\delta > 0$ be such that there is no eigenvalue of P'' on the unit circle with argument in $[0,\delta]$. In particular, the functions $t \mapsto f(e^{it})$ and $t \mapsto g(e^{it})$ vanish on the interval $[0,\delta]$. By integrating (19) in μ on the upper semi-circle, we obtain

$$\begin{aligned} \overline{\operatorname{ind}}(\boldsymbol{z}) - \operatorname{ind}(\boldsymbol{z}) &= \frac{1}{\pi}\int_0^\pi \big(\operatorname{ind}_{e^{it}}(\boldsymbol{z}) - \operatorname{ind}(\boldsymbol{z})\big)\,\mathrm{d}t \\ &\leq \frac{1}{2}\dim_{\mathbb{C}}\mathbb{E} + \frac{1}{\pi}\int_\delta^\pi f(e^{it})\,\mathrm{d}t \\ &\leq \frac{1}{2}\left(\dim_{\mathbb{C}}\mathbb{E} + \frac{\pi-\delta}{\pi}\dim_{\mathbb{C}}\mathbb{E}^\omega\right). \end{aligned}$$

In particular

$$\overline{\operatorname{ind}}(\boldsymbol{z}) - \operatorname{ind}(\boldsymbol{z}) \leq \tfrac{1}{2}(\dim_{\mathbb{C}}\mathbb{E} + \dim_{\mathbb{C}}\mathbb{E}^\omega) = d. \tag{22}$$

If this inequality is not strict, then $\dim_{\mathbb{C}}\mathbb{E}^\omega = 0$, that is, $\sigma(P) = \{1\}$. Moreover, in this case we have $\operatorname{ind}_\theta(\boldsymbol{z}) = \overline{\operatorname{ind}}(\boldsymbol{z})$ for all $\theta \neq 1$, so that $\mathrm{S}^{\pm}_P(1) = \overline{\operatorname{ind}}(\boldsymbol{z}) - \operatorname{ind}(\boldsymbol{z}) = d$, and by (14) we conclude that $\operatorname{nul}(\boldsymbol{z}) \geq \mathrm{S}^{\pm}_P(1) = d$.

If we now integrate (21), we obtain

$$\begin{aligned} \operatorname{ind}(\boldsymbol{z}) + \operatorname{nul}(\boldsymbol{z}) - \overline{\operatorname{ind}}(\boldsymbol{z}) &= \frac{1}{\pi}\int_0^\pi \big(\operatorname{ind}(\boldsymbol{z}) + \operatorname{nul}(\boldsymbol{z}) - \operatorname{ind}_{e^{it}}(\boldsymbol{z})\big)\mathrm{d}t \\ &\leq \frac{1}{2}\dim_{\mathbb{C}}\mathbb{E} + \frac{1}{\pi}\int_\delta^\pi g(e^{it})\,\mathrm{d}t \\ &\leq \frac{1}{2}\left(\dim_{\mathbb{C}}\mathbb{E} + \frac{\pi-\delta}{\pi}\dim_{\mathbb{C}}\mathbb{E}^\omega\right). \end{aligned}$$

Therefore

$$\operatorname{ind}(\boldsymbol{z}) + \operatorname{nul}(\boldsymbol{z}) - \overline{\operatorname{ind}}(\boldsymbol{z}) \leq \tfrac{1}{2}(\dim_{\mathbb{C}}\mathbb{E} + \dim_{\mathbb{C}}\mathbb{E}^\omega) = d. \tag{23}$$

As before, if this inequality is not strict, then $\dim_{\mathbb{C}}\mathbb{E}^\omega = 0$, that is, $\sigma(P) = \{1\}$. Moreover, in this case we have $\operatorname{coind}_\theta(\boldsymbol{z}) = \overline{\operatorname{coind}}(\boldsymbol{z})$ for all $\theta \neq 1$, so that

$$\mathrm{coS}^{\pm}_P(1) = \overline{\operatorname{coind}}(\boldsymbol{z}) - \operatorname{coind}(\boldsymbol{z}) = \operatorname{ind}(\boldsymbol{z}) + \operatorname{nul}(\boldsymbol{z}) - \overline{\operatorname{ind}}(\boldsymbol{z}) = d,$$

and by (14) we conclude that $\mathrm{nul}(\boldsymbol{z}) \geq \mathrm{coS}^{\pm}_{P}(1) = d$.

Both inequalities in (22) and (23) are simultaneously equalities if and only if $\mathrm{nul}(\boldsymbol{z}) = 2d$, that is, if and only if P is the identity. This completes the proof of the theorem for period $p = 1$. The case of an arbitrary period $p \in \mathbb{N}$ readily follows by recalling that $\overline{\mathrm{ind}}(\boldsymbol{z}^p) = p\,\overline{\mathrm{ind}}(\boldsymbol{z})$. □

2.4. **Computation of splitting numbers.** We close this section by providing a recipe for computing the splitting numbers of a symplectic matrix $P \in \mathrm{Sp}(2d)$. We consider a quadratic generating family $F : \mathbb{R}^{2dk} \to \mathbb{R}$ for the linear Hamiltonian diffeomorphism $\phi(z) = Pz$. We denote by H_θ the θ-Hessian of F, and by $h_\theta : \mathbb{C}^{2dk} \times \mathbb{C}^{2dk} \to \mathbb{C}$ the associated Hermitian bilinear form, so that in particular

$$F(\boldsymbol{Z}) = \tfrac{1}{2}\langle H_1 \boldsymbol{Z}, \boldsymbol{Z}\rangle = \tfrac{1}{2} h_1(\boldsymbol{Z}, \boldsymbol{Z}), \qquad \forall \boldsymbol{Z} \in \mathbb{R}^{2dk}.$$

In Section 2.2, we studied the inertia of the restriction of h_θ to the vector subspace $\mathbb{V}$ and to its h_θ-orthogonal in order to show that the splitting and cosplitting numbers depend only on the considered symplectic matrix P. The choice of the vector space $\mathbb{V}$ was suitable in order to establish the bounds of Lemma 2.9, but is not convenient for the numeric computation of the splitting and cosplitting numbers. For this purpose, we rather choose the vector space

$$\mathbb{W} := \big\{\boldsymbol{Z} = (Z_0, ..., Z_{k-1}) \in \mathbb{C}^{2dk} \;\big|\; Z_0 = 0\big\}.$$

As in the case of $\mathbb{V}$, the restriction of the Hermitian form h_θ to $\mathbb{W}$ is independent of the parameter θ, since for all $\boldsymbol{Z}, \boldsymbol{Z}' \in \mathbb{C}^{2dk}$ we have

$$\begin{aligned}
h_\theta(\boldsymbol{Z}, \boldsymbol{Z}') &= \langle \overline{\theta}\, Y_{k-1} - Y_0 + A_{k-1}X_0 + \overline{\theta}\, B_{k-1}^T Y_{k-1}, X_0'\rangle \\
&\quad + \langle \theta\, X_0 - X_{k-1} + \theta\, B_{k-1}X_0 + C_{k-1}Y_{k-1}, Y_{k-1}'\rangle \\
&\quad + \sum_{j=1}^{k-1} \langle Y_{j-1} - Y_j + A_{j-1}X_j + B_{j-1}^T Y_{j-1}, X_j'\rangle \\
&\quad + \sum_{j=0}^{k-2} \langle X_{j+1} - X_j + B_j X_{j+1} + C_j Y_j, Y_j'\rangle \\
&= \langle -X_{k-1} + C_{k-1}Y_{k-1}, Y_{k-1}'\rangle + \langle -Y_1 + A_0 X_1, X_1'\rangle \\
&\quad + \sum_{j=2}^{k-1} \langle Y_{j-1} - Y_j + A_{j-1}X_j + B_{j-1}^T Y_{j-1}, X_j'\rangle \\
&\quad + \sum_{j=1}^{k-2} \langle X_{j+1} - X_j + B_j X_{j+1} + C_j Y_j, Y_j'\rangle.
\end{aligned}$$

In particular the functions $\theta \mapsto \mathrm{ind}(h_\theta|_{\mathbb{W}\times\mathbb{W}})$ and $\theta \mapsto \mathrm{coind}(h_\theta|_{\mathbb{W}\times\mathbb{W}})$ are independent of θ. We recall that the kernel of H_θ is the space of vectors $\boldsymbol{Z} = (Z_0, ..., Z_{k-1}) \in \mathbb{C}^{2dk}$ such that $\phi_j Z_j = Z_{j+1}$ for all $j = 0, ..., k-2$, and $\phi_{k-1} Z_{k-1} = \theta\, Z_0$. Therefore

$$\mathbb{W} \cap \ker H_\theta = \{0\}.$$

The orthogonal vector space $\mathbb{W}^{h_\theta}$ is given by the solutions $\boldsymbol{Z} \in \mathbb{C}^{2dk}$ of the following linear system

$$\begin{aligned}
\theta\, X_0 - X_{k-1} + \theta\, B_{k-1} X_0 + C_{k-1} Y_{k-1} &= 0, \\
Y_{j-1} - Y_j + A_{j-1} X_j + B_{j-1}^T Y_{j-1} &= 0, \qquad \forall j = 1, ..., k-1, \\
X_{j+1} - X_j + B_j X_{j+1} + C_j Y_j &= 0, \qquad \forall j = 1, ..., k-2.
\end{aligned}$$

Namely,

$$\mathbb{W}^{h_\theta} = \left\{ (Z_0, ..., Z_{k-1}) \in \mathbb{C}^{2dk} \;\middle|\; \begin{array}{l} \phi_0(\tilde{X}_0, Y_0) = Z_1 \text{ for some } \tilde{X}_0 \in \mathbb{C}^d \\ \phi_j Z_j = Z_{j+1} \quad \forall j = 1, ..., k-2 \\ \phi_{k-1} Z_{k-1} = (\theta X_0, \tilde{Y}_k) \text{ for some } \tilde{Y}_k \in \mathbb{C}^d \end{array} \right\}.$$

We denote by $\Psi_\theta : \mathbb{C}^{2d} \to \mathbb{W}^{h_\theta}$ the isomorphism given by $\Psi_\theta^{-1} \boldsymbol{Z} = (\tilde{X}_0, Y_0)$. Notice that $P \circ \Psi_\theta^{-1}(\boldsymbol{Z}) = (\theta X_0, \tilde{Y}_k)$.

The intersection

$$\mathbb{W} \cap \mathbb{W}^{h_\theta} = \left\{ (0, Z_1, ..., Z_{k-1}) \in \mathbb{C}^{2dk} \;\middle|\; \begin{array}{l} \phi_0(\tilde{X}_0, 0) = Z_1 \text{ for some } \tilde{X}_0 \in \mathbb{C}^d \\ \phi_j Z_j = Z_{j+1} \quad \forall j = 1, ..., k-2 \\ \phi_{k-1} Z_{k-1} = (0, \tilde{Y}_k) \text{ for some } \tilde{Y}_k \in \mathbb{C}^d \end{array} \right\}$$

is independent of θ. Proposition A.3 gives

$$\begin{aligned}
\mathrm{ind}(h_\theta) &= \mathrm{ind}(h_\theta|_{\mathbb{W}\times\mathbb{W}}) + \mathrm{ind}(h_\theta|_{\mathbb{W}^{h_\theta}\times\mathbb{W}^{h_\theta}}) + \dim_{\mathbb{C}}(\mathbb{W} \cap \mathbb{W}^{h_\theta}), \\
\mathrm{coind}(h_\theta) &= \mathrm{coind}(h_\theta|_{\mathbb{W}\times\mathbb{W}}) + \mathrm{coind}(h_\theta|_{\mathbb{W}^{h_\theta}\times\mathbb{W}^{h_\theta}}) + \dim_{\mathbb{C}}(\mathbb{W} \cap \mathbb{W}^{h_\theta}).
\end{aligned}$$

Only the second summand in the right-hand sides of these two equations depends on θ. Therefore, the splitting and cosplitting numbers are given by

$$\begin{aligned}
\mathrm{S}_P^{\pm}(\theta) &= \mathrm{ind}(h_{\theta^\pm}|_{\mathbb{W}^{h_{\theta^\pm}}\times\mathbb{W}^{h_{\theta^\pm}}}) - \mathrm{ind}(h_\theta|_{\mathbb{W}^{h_\theta}\times\mathbb{W}^{h_\theta}}), \\
\mathrm{coS}_P^{\pm}(\theta) &= \mathrm{coind}(h_{\theta^\pm}|_{\mathbb{W}^{h_{\theta^\pm}}\times\mathbb{W}^{h_{\theta^\pm}}}) - \mathrm{coind}(h_\theta|_{\mathbb{W}^{h_\theta}\times\mathbb{W}^{h_\theta}}),
\end{aligned}$$

Let us compute the restriction of the Hermitian form h_θ to $\mathbb{W}^{h_\theta}$. For all pair of vectors $\boldsymbol{Z}, \boldsymbol{Z}' \in \mathbb{W}^{h_\theta}$, we have

$$\begin{aligned}
h_\theta(\boldsymbol{Z}, \boldsymbol{Z}') &= \langle \overline{\theta}\, Y_{k-1} - Y_0 + A_{k-1} X_0 + \overline{\theta}\, B_{k-1}^T Y_{k-1}, X_0' \rangle \\
&\quad + \langle X_1 - X_0 + B_0 X_1 + C_0 Y_0, Y_0' \rangle \\
&= \langle \overline{\theta}\, Y_{k-1} - Y_0 + \overline{\theta}\, \underbrace{(A_{k-1} \theta\, X_0 + B_{k-1}^T Y_{k-1})}_{\tilde{Y}_k - Y_{k-1}}, X_0' \rangle \\
&\quad + \langle X_1 - X_0 + \underbrace{B_0 X_1 + C_0 Y_0}_{\tilde{X}_0 - X_1}, Y_0' \rangle \\
&= \langle \tilde{Y}_k - \theta\, Y_0, \theta\, X_0' \rangle + \langle \tilde{X}_0 - \overline{\theta}\theta\, X_0, Y_0' \rangle,
\end{aligned}$$

where $\tilde{X}_0$ and $\tilde{Y}_0$ depends on $\boldsymbol{Z}$ as in the above characterization of $\mathbb{W}^{h_\theta}$. Let us choose the more convenient coordinates given by the isomorphism Ψ_θ. Namely, we consider the Hermitian form $g_\theta : \mathbb{C}^{2d} \times \mathbb{C}^{2d} \to \mathbb{C}$ given by

$$g_\theta(Z, Z') := h_\theta(\Psi_\theta Z, \Psi_\theta Z').$$

If we write $(\tilde X,\tilde Y):=P(X,Y)$ and $(\tilde X',\tilde Y'):=P(X',Y')$, the Hermitian form g_θ can be written as

$$g_\theta((X,Y),(X',Y'))=\langle \tilde Y-\theta\,Y,\tilde X'\rangle+\langle X-\overline{\theta}\,\tilde X,Y'\rangle \tag{24}$$

The splitting and cosplitting numbers can be conveniently computed as

$$\begin{aligned}\mathrm{S}_P^{\pm}(\theta)&=\mathrm{ind}(g_{\theta^\pm})-\mathrm{ind}(g_\theta),\\ \mathrm{coS}_P^{\pm}(\theta)&=\mathrm{coind}(g_{\theta^\pm})-\mathrm{coind}(g_\theta).\end{aligned}$$

Example 2.12 (Splitting numbers of a shear)**.** For $r\in\mathbb{R}$, consider the unipotent symplectic matrix

$$P=\begin{pmatrix}1 & r\\ 0 & 1\end{pmatrix}.$$

For each $\theta\in S^1$, the associated Hermitian form g_θ is given by

$$\begin{aligned}g_\theta(Z,Z')&=\langle(1-\theta)Y,X'+rY'\rangle+\langle(1-\overline{\theta})X-\overline{\theta}r\,Y,Y'\rangle\\ &=(1-\theta)\langle Y,X'\rangle+(1-\overline{\theta})\langle X,Y'\rangle+r(1-2\,\mathrm{Re}(\theta))\langle Y,Y'\rangle\end{aligned}$$

The Hermitian matrix associated to g_θ is given by

$$\begin{pmatrix}0 & 1-\theta\\ 1-\overline{\theta} & r(1-2\,\mathrm{Re}(\theta))\end{pmatrix},$$

whose eigenvalues $\lambda_1,\lambda_2\in\mathbb{R}$ satisfy $\lambda_1\lambda_2=-|1-\theta|^2$ and $\lambda_1+\lambda_2=r(1-2\,\mathrm{Re}(\theta))$. Therefore

$$\begin{aligned}\mathrm{ind}(g_\theta)&=\begin{cases}1 & \text{if } \theta\neq 1 \text{ or } r>0,\\ 0 & \text{if } \theta=1 \text{ and } r\leq 0,\end{cases}\\ \mathrm{coind}(g_\theta)&=\begin{cases}1 & \text{if } \theta\neq 1 \text{ or } r<0,\\ 0 & \text{if } \theta=1 \text{ and } r\geq 0,\end{cases}\end{aligned}$$

which implies

$$\begin{aligned}\mathrm{S}_P^{\pm}(1)&=\begin{cases}1 & \text{if } r\leq 0,\\ 0 & \text{if } r>0,\end{cases}\\ \mathrm{coS}_P^{\pm}(1)&=\begin{cases}1 & \text{if } r\geq 0,\\ 0 & \text{if } r<0.\end{cases}\end{aligned}$$ □

Example 2.13 (Splitting numbers of a $\pi/2$-rotation)**.** Consider now the symplectic matrix of the standard complex structure of $(\mathbb{R}^{2d},\omega)$, which is

$$J=\begin{pmatrix}0 & -I\\ I & 0\end{pmatrix}.$$

The eigenvalues of J are i and $-i$. The associated Hermitian forms g_θ are given by

$$g_\theta(Z,Z')=\langle X-\theta Y,-Y'\rangle+\langle X+\overline{\theta}Y,Y'\rangle,$$

with associated Hermitian matrices

$$\begin{pmatrix}0 & 0\\ 0 & (\theta+\overline{\theta})I\end{pmatrix},$$

This readily implies

$$\operatorname{ind}(g_\theta) = \begin{cases} 0 & \text{if } \operatorname{Re}(\theta) \geq 0, \\ d & \text{if } \operatorname{Re}(\theta) < 0, \end{cases}$$
$$\operatorname{coind}(g_\theta) = \begin{cases} d & \text{if } \operatorname{Re}(\theta) > 0, \\ 0 & \text{if } \operatorname{Re}(\theta) \leq 0, \end{cases}$$

and therefore

$$\mathrm{S}^+_J(i) = \mathrm{S}^-_J(-i) = \mathrm{coS}^-_J(i) = \mathrm{coS}^+_J(-i) = d,$$
$$\mathrm{S}^-_J(i) = \mathrm{S}^+_J(-i) = \mathrm{coS}^+_J(i) = \mathrm{coS}^-_J(-i) = 0. \qquad \square$$

2.5. **Bibliographical remarks.** The iteration theory for the Morse indices of periodic orbits was introduced in the setting of Tonelli Lagrangian systems by Bott [Bot56], who developed ideas introduced earlier by Hedlund [Hed32] and Morse-Pitcher [MP34]. The setting of this section is more general than Bott's one, as we will discuss in Section 4. A special Morse index theory for the Hamiltonian action functional was first studied by Conley and Zehnder in their papers [CZ84a, CZ84b]. As we already mentioned before Proposition 1.2, the critical points of the Hamiltonian action functional always have infinite Morse index and coindex. The index that Conley and Zehnder defined coincides with the Maslov index, which we will introduce in Section 3. Theorem 2.3 is the translation, in the finite dimensional setting of Chaperon's generating families, of Conley-Zehnder's iteration inequality for the Maslov index of non-degenerate symplectic paths. This inequality is the crucial ingredient in the proof of one of Conley-Zehnder's famous theorems from [CZ84b] (see also the author's [Maz13] for a proof using Chaperon's generating families, and Salamon-Zehnder's [SZ92] for a generalization to all closed symplectically aspherical manifolds): a generic Hamiltonian diffeomorphism on a standard symplectic $2d$-torus possesses periodic points of arbitrarily large minimal period. The general iteration inequalities, or more precisely their translation in terms of the Maslov indices (Theorem 3.6), are due to Liu and Long [LL98, LL00]. One of their most remarkable application is to the non-generic version of Conley-Zehnder's Theorem, which was a long standing conjecture due to Conley and established by Hingston [Hin09]: any Hamiltonian diffeomorphism on a standard symplectic $2d$-torus with finitely many fixed points possesses periodic points of arbitrarily large minimal period. Generalizations of Hingston's Theorem to larger and larger classes of closed symplectic manifolds were established by Ginzburg [Gin10], Ginzburg-Gürel [GG10, GG12] and Hein [Hei12]. We refer the reader to Long's monograph [Lon02] for other applications of the iteration inequalities. Many proofs that we provided in this section, as well as the recipe for computing the splitting numbers of symplectic matrices, were inspired by Ballmann-Thorbergsson-Ziller's [BTZ82].

3. The Maslov index

3.1. **Behavior of the inertia indices under stabilization.** Consider a symplectic matrix $P \in \mathrm{Sp}(2d)$. Choose a factorization

$$P = P_{k-1} \circ ... \circ P_0 \tag{25}$$

such that each P_j is sufficiently close to the identity in $\mathrm{Sp}(2d)$, and therefore it is described by a quadratic generating function $f_j : \mathbb{R}^{2d} \to \mathbb{R}$. As before, we write this function as

$$f_j(X_{j+1}, Y_j) = \tfrac{1}{2}\langle A_j X_{j+1}, X_{j+1}\rangle + \langle B_j X_{j+1}, Y_j\rangle + \tfrac{1}{2}\langle C_j Y_j, Y_j\rangle,$$

where A_j, B_j, and C_j are (small) $dk \times dk$ real matrices, A_j and C_j being symmetric. The factorization (25) singles out a path in the symplectic group joining the identity to P. Indeed, for all $t \in [0,1]$, let P_j^t be the symplectic matrix defined by the generating function $t\, f_j$, i.e.

$$P_j^t Z_j = Z_{j+1} \qquad \text{if and only if} \qquad \begin{cases} X_{j+1} - X_j = -t(B_j X_{j+1} + C_j Y_j), \\ Y_{j+1} - Y_j = t(A_j X_{j+1} + B_j^T Y_j). \end{cases}$$

Notice that $P_j^0 = I$ and $P_j^1 = P_j$. For all $t \in [0,1]$, we set

$$P^t = P_j^s \circ P_{j-1} \circ ... \circ P_0, \qquad \text{where } j = \lfloor kt \rfloor,\ s = kt - j.$$

The continuous path $t \mapsto P^t$ in the symplectic group $\mathrm{Sp}(2d)$ joins $P^0 = I$ and $P^1 = P$.

On the other hand, if we started with a continuous path $\Gamma : [0,1] \to \mathrm{Sp}(2d)$ such that $\Gamma(0) = I$ and $\Gamma(1) = P$, up to choosing k large enough, for all $|t_1 - t_2| \leq 1/k$ the symplectic matrix $\Gamma(t_2)\Gamma(t_1)^{-1}$ becomes as close to the identity as we wish, and in particular close enough to being described by a quadratic generating function. If we now set

$$P_j := \Gamma(\tfrac{j+1}{k})\Gamma(\tfrac{j}{k})^{-1}, \qquad \forall j = 0, ..., k-1, \tag{26}$$

and denote by $t \mapsto P^t$ the symplectic path associated to the factorization (25) as above, the paths Γ and $t \mapsto P^t$ are homotopic (via a homotopy that fixes the endpoints). Indeed, their restrictions to any time interval of the form $[j/k, (j+1)/k]$ are homotopic with fixed endpoints.

Let $F : \mathbb{R}^{2dk} \to \mathbb{R}$ be the quadratic generating family associated to the factorization (25) of P, that is,

$$F(\boldsymbol{Z}) = \tfrac{1}{2}\langle H\boldsymbol{Z}, \boldsymbol{Z}\rangle = \sum_{j \in \mathbb{Z}_k} \Big(\langle Y_j, X_{j+1} - X_j\rangle + f_j(X_{j+1}, Y_j) \Big). \tag{27}$$

We denote by $h(\boldsymbol{Z}, \boldsymbol{Z}') = \langle H\boldsymbol{Z}, \boldsymbol{Z}'\rangle$ the Hessian bilinear form associated to F. We recall that an image vector $\boldsymbol{Z}' = H(\boldsymbol{Z})$ is defined by

$$\begin{aligned} X_j' &= Y_{j-1} - Y_j + A_{j-1} X_j + B_{j-1}^T Y_{j-1}, \\ Y_j' &= X_{j+1} - X_j + B_j X_{j+1} + C_j Y_j. \end{aligned}$$

From Lemma 2.1(ii), we know that $\mathrm{nul}(h) = \dim\ker(P - I)$. However, it is not hard to convince ourselves that the data of P alone is not enough to determine the other inertia indices of h (see for instance the example of the identity mentioned before Proposition 1.2). In this section we are going to show that the index $\mathrm{ind}(h)$ and the coindex $\mathrm{coind}(h)$ are completely determined by the number of factors k in the factorization (25) and by the homotopy class of the path $t \mapsto P^t$ in the symplectic group $\mathrm{Sp}(2d)$.

Let us begin by studying how the inertia indices change if we increase k by adding trivial factors in (25). For some $l > k$, let us set $P_k = P_{k+1} = ... = P_{l-1} := I$, and consider the generating function $F' : \mathbb{R}^{2dl} \to \mathbb{R}$ associated to the factorization

$$P = P_{l-1} \circ ... \circ P_k \circ P_{k-1} \circ ... \circ P_0.$$

We denote by $h'(\boldsymbol{Z},\boldsymbol{Z}') = \langle H'\boldsymbol{Z},\boldsymbol{Z}'\rangle$ the Hessian bilinear form associated to F'. The following lemma shows that h' is essentially a stabilization of h.

Lemma 3.1. *The inertia indices of h and h' are related by*

$$\begin{aligned}\mathrm{nul}(h') &= \mathrm{nul}(h),\\ \mathrm{ind}(h') &= \mathrm{ind}(h) + d(l-k),\\ \mathrm{coind}(h') &= \mathrm{coind}(h) + d(l-k).\end{aligned}$$

Proof. We already know the claim about the nullities, so let us focus on the other two. Consider the vector space

$$\mathbb{V} = \big\{\boldsymbol{Z} \in \mathbb{R}^{2dl} \bigm| Z_k = Z_{k+1} = ... = Z_{l-1} = Z_0\big\}.$$

Let $\pi : \mathbb{R}^{2dl} \to \mathbb{R}^{2dk}$ be the projection

$$\pi(X_0, Y_0, ..., X_{l-1}, Y_{l-1}) = (X_k, Y_0, X_1, Y_1, ..., X_{k-1}, Y_{k-1}),$$

and $\iota : \mathbb{R}^{2dk} \to \mathbb{V}$ the isomorphism

$$\iota(Z_0, ..., Z_{k-1}) = \iota(Z_0, ..., Z_{k-1}, Z_0, Z_0, ..., Z_0).$$

Notice that

$$h'(\boldsymbol{Z}, \iota(\boldsymbol{Z}')) = h(\pi(\boldsymbol{Z}), \boldsymbol{Z}'), \qquad \forall \boldsymbol{Z} \in \mathbb{R}^{2dl}, \boldsymbol{Z}' \in \mathbb{R}^{2dk}. \tag{28}$$

Since the inverse of ι is given by $\pi|_{\mathbb{V}}$, the restriction of h' to $\mathbb{V}$ coincides with h, in the sense that $h'(\iota(\cdot), \iota(\cdot)) = h$. Therefore

$$\mathrm{ind}(h'|_{\mathbb{V}\times\mathbb{V}}) = \mathrm{ind}(h), \qquad \mathrm{coind}(h'|_{\mathbb{V}\times\mathbb{V}}) = \mathrm{coind}(h).$$

Now, we need to study the h'-orthogonal to $\mathbb{V}$. By (28), we infer that

$$\mathbb{V}^{h'} = \pi^{-1}(\ker(h)).$$

Namely, $\mathbb{V}^{h'}$ is the vector space of the solutions $\boldsymbol{Z} \in \mathbb{R}^{2dl}$ of the linear system

$$\begin{aligned} X_1 - X_k + B_0 X_1 + C_0 Y_0 &= 0, & \\ X_{j+1} - X_j + B_j X_{j+1} + C_j Y_j &= 0, & \forall j = 1, ..., k-1,\\ Y_{k-1} - Y_0 + A_{k-1} X_k + B_{k-1}^T Y_{k-1} &= 0, & \\ Y_{j-1} - Y_j + A_{j-1} X_j + B_{j-1}^T Y_{j-1} &= 0, & \forall j = 1, ..., k-1.\end{aligned}$$

This means that

$$\mathbb{V}^{h'} = \left\{ (Z_0, ..., Z_{l-1}) \in \mathbb{R}^{2dl} \;\middle|\; \begin{array}{l} P_0(X_k, Y_0) = (X_1, Y_1) \\ P_j(X_j, Y_j) = (X_{j+1}, Y_{j+1}) \quad \forall j = 1, ..., k-2 \\ P_{k-1}(X_{k-1}, Y_{k-1}) = (X_k, Y_0) \end{array} \right\}.$$

Notice that $\mathbb{V} \cap \mathbb{V}^{h'} = \ker H = \mathbb{V} \cap \ker H$. By Proposition A.3 we infer that

$$\begin{aligned}\mathrm{ind}(h') &= \mathrm{ind}(h) + \mathrm{ind}(h'|_{\mathbb{V}^{h'}\times\mathbb{V}^{h'}}),\\ \mathrm{coind}(h') &= \mathrm{coind}(h) + \mathrm{coind}(h'|_{\mathbb{V}^{h'}\times\mathbb{V}^{h'}}).\end{aligned}$$

In order to complete the proof, we only have to show that

$$\mathrm{ind}(h'|_{\mathbb{V}^{h'}\times\mathbb{V}^{h'}}) = \mathrm{coind}(h'|_{\mathbb{V}^{h'}\times\mathbb{V}^{h'}}) = d(l-k). \tag{29}$$

We prove this equality as follows. For all $\boldsymbol{Z}, \boldsymbol{Z}' \in \mathbb{V}^{h'}$, we have

$$\begin{aligned} h'(\boldsymbol{Z}, \boldsymbol{Z}') =& \sum_{j=k}^{l-1} \Big(\langle Y_j - Y_{j+1}, X'_{j+1} \rangle + \langle X_{j+1} - X_j, Y'_j \rangle \Big) \\ &+ \langle -Y_k + \underbrace{Y_{k-1} + A_{k-1} X_k + B_{k-1}^T Y_{k-1}}_{=Y_0}, X'_k \rangle \\ &+ \langle -X_0 + \underbrace{X_1 + B_0 X_1 + C_0 Y_0}_{=X_k}, Y'_0 \rangle. \end{aligned}$$

This expression readily implies that the symmetric bilinear form $h'|_{\mathbb{V}^{h'} \times \mathbb{V}^{h'}}$ is negative definite on the following vector subspace of $\mathbb{V}^{h'}$

$$\mathbb{E}^- = \left\{ (X_0, 0, ..., 0, Y_k, X_{k+1}, Y_{k+1}, ..., X_{l-1}, Y_{l-1}) \in \mathbb{R}^{2dl} \;\middle|\; \begin{array}{r} X_{j+1} = Y_{j+1} - Y_j \\ \forall j = k, ..., l-1 \end{array} \right\}.$$

Analogously, $h'|_{\mathbb{V}^{h'} \times \mathbb{V}^{h'}}$ is positive definite on the following vector subspace of $\mathbb{V}^{h'}$

$$\mathbb{E}^+ = \left\{ (X_0, 0, ..., 0, Y_k, X_{k+1}, Y_{k+1}, ..., X_{l-1}, Y_{l-1}) \in \mathbb{R}^{2dl} \;\middle|\; \begin{array}{r} X_{j+1} = Y_j - Y_{j+1} \\ \forall j = k, ..., l-1 \end{array} \right\}.$$

Notice that $\dim \mathbb{E}^- = \dim \mathbb{E}^+ = d(l-k)$, and obviously the intersection of $\mathbb{E}^-$ with $\mathbb{E}^+$ is trivial. Notice further that the kernel of $h'|_{\mathbb{V}^{h'} \times \mathbb{V}^{h'}}$ is given by

$$\ker(h'|_{\mathbb{V}^{h'} \times \mathbb{V}^{h'}}) = \left\{ (Z_0, ..., Z_{l-1}) \in \mathbb{V}^{h'} \;\middle|\; \begin{array}{l} X_0 = X_k = X_{k+1} = ... = X_{l-1} \\ Y_0 = Y_k = Y_{k+1} = ... = Y_{l-1} \end{array} \right\},$$

whose dimension is $\mathrm{nul}(h'|_{\mathbb{V}^{h'} \times \mathbb{V}^{h'}}) = \dim \ker(P - I)$. Finally, we remark that there is an isomorphism

$$\Psi : \mathbb{V}^{h'} \to \ker(P - I) \times \mathbb{R}^{2d(l-k)}$$

given by

$$\Psi(Z_0, ..., Z_{l-1}) = \big((X_k, Y_0), (X_0, Y_k, X_{k+1}, Y_{k+1}, X_{k+2}, Y_{k+2}, ..., X_{l-1}, Y_{l-1}) \big).$$

In particular

$$\dim \mathbb{V}^{h'} = \dim \ker(P - I) + 2d(l-k) = \mathrm{nul}(h'|_{\mathbb{V}^{h'} \times \mathbb{V}^{h'}}) + \dim(\mathbb{E}^-) + \dim(\mathbb{E}^+).$$

Therefore $\mathbb{E}^-$ and $\mathbb{E}^+$ are maximal vector subspaces of $\mathbb{V}^{h'}$ where the bilinear form $h'|_{\mathbb{V}^{h'} \times \mathbb{V}^{h'}}$ is negative definite and positive definite respectively. This implies our claim in (29). □

3.2. **Morse and Maslov indices.** We now have all the ingredients to introduce the main character of this section: the Maslov index. For $n = 0, ..., 2d$, we introduce the spaces of symplectic paths

$$\mathcal{P}_n(2d) := \big\{ \Gamma : [0,1] \to \mathrm{Sp}(2d) \;\big|\; \Gamma(0) = I,\ \dim \ker(\Gamma(1) - I) = n \big\}$$

endowed with the C^0-topology. This gives a partition of the full space of symplectic paths

$$\mathcal{P}(2d) := \bigcup_{n=0}^{2d} \mathcal{P}_n(2d).$$

Given $\Gamma \in \mathcal{P}$ with $\Gamma(1) =: P$, we choose a parameter k large enough and we consider the symplectic matrices (26), which give the factorization (25) and the associated quadratic generating function with Hessian bilinear form $h : \mathbb{R}^{2dk} \times \mathbb{R}^{2dk} \to \mathbb{R}$. We define the **Maslov index** of Γ as

$$\mathrm{mas}(\Gamma) := \mathrm{ind}(h) - dk \in \mathbb{Z}.$$

Analogously, we defined the **Maslov coindex** of Γ as

$$\mathrm{comas}(\Gamma) := \mathrm{coind}(h) - dk \in \mathbb{Z}.$$

Since the inertia indices are related by $\mathrm{ind}(h) + \mathrm{coind}(h) + \mathrm{nul}(h) = 2dk$, we have

$$\mathrm{mas}(\Gamma) + \mathrm{comas}(\Gamma) + \dim\ker(\Gamma(1) - I) = 0. \tag{30}$$

In particular the Maslov index is equal to minus the Maslov coindex on the subspace $\mathcal{P}_0(2d)$.

Warning 3.2. Many authors in symplectic topology call Maslov index what we call Maslov coindex. This different convention amounts to changing the sign of the generating families. Example 3.5 below can be useful to recognize the sign convention adopted in a paper. □

The next Theorem implies that these are good definitions.

Theorem 3.3. *The Maslov index is a well defined function*

$$\mathrm{mas} : \mathcal{P}(2d) \to \mathbb{Z},$$

i.e. $\mathrm{mas}(\Gamma)$ *is independent of the chosen parameter* k*. Moreover, it is a lower semi-continuous function, and it is locally constant on every subspace* $\mathcal{P}_n(2d)$*. The same properties hold for the Maslov coindex.*

Proof. In order to show that the Maslov index and coindex are well defined, we only have to prove that we obtain the same indices if we replace k by a larger parameter l in the setting above. We proceed as follows. We define a homotopy $\Gamma_s : [0, 1] \to \mathrm{Sp}(2d)$, for $s \in [0, 1]$, such that $\Gamma_0 = \Gamma$, each Γ_s has the same endpoints as Γ, and Γ_1 runs along the whole Γ in the time interval $[0, k/l]$, and stays constant at $\Gamma(1)$ in the remaining time interval $[k/l, 1]$. This homotopy is defined by the formula

$$\Gamma_s(t) := \Gamma\big(\min\big\{1, \tfrac{l}{l+s(k-l)} t\big\}\big).$$

For each $s \in [0, 1]$, we introduce the factorization

$$\Gamma(1) = P_{l-1,s} \circ P_{l-2,s} \circ ... \circ P_{0,s}, \tag{31}$$

where

$$P_{j,s} := \Gamma_s(\tfrac{j+1}{l})\Gamma_s(\tfrac{j}{l})^{-1}, \qquad \forall j = 0, ..., l-1.$$

Since the parameter l is larger than k, each symplectic matrix $P_{j,s}$ is sufficiently close to the identity to be described by a quadratic generating function. We denote by $h_s : \mathbb{R}^{2dl} \times \mathbb{R}^{2dl} \to \mathbb{R}$ the Hessian bilinear form of the quadratic generating family associated to the factorization (31). For $s = 1$, equation (31) gives the factorization of $\Gamma(1)$ corresponding to the parameter l. Hence, all we need to do is to prove that

$$\mathrm{ind}(h_1) - dl = \mathrm{ind}(h) - dk,$$
$$\mathrm{coind}(h_1) - dl = \mathrm{coind}(h) - dk.$$

Notice that h_s depends continuously on $s \in [0,1]$. Moreover, its nullity is constant in s, since

$$\mathrm{nul}(h_s) = \dim\ker(\Gamma(1) - I), \qquad \forall s \in [0,1].$$

This implies that the functions $s \mapsto \mathrm{ind}(h_s)$ and $s \mapsto \mathrm{coind}(h_s)$ are constant in s as well. For $s = 0$, equation (31) gives the following factorization of $\Gamma(1)$

$$\Gamma(1) = \underbrace{I \circ I \circ ... \circ I}_{\times l-k} \circ P_{k-1} \circ P_{k-2} \circ ... \circ P_0.$$

By Lemma 3.1, we have

$$\begin{aligned}\mathrm{ind}(h_0) &= \mathrm{ind}(h) + d(l-k),\\ \mathrm{coind}(h_0) &= \mathrm{coind}(h) + d(l-k).\end{aligned}$$

Therefore

$$\begin{aligned}\mathrm{ind}(h_1) - dl &= \mathrm{ind}(h_0) - dl = \mathrm{ind}(h) + d(l-k) - dl = \mathrm{ind}(h) - dk,\\ \mathrm{coind}(h_1) - dl &= \mathrm{coind}(h_0) - dl = \mathrm{coind}(h) + d(l-k) - dl = \mathrm{coind}(h) - dk.\end{aligned}$$

This completes the proof that the Maslov index and coindex are well defined. Their lower semi-continuity follows immediately by the same property for the inertia index and coindex of symmetric bilinear forms.

Finally, let $s \mapsto \Gamma_s$, $s \in [0,1]$, be a path inside a space $\mathcal{P}_n(2d)$, for some $n \in \mathbb{N}$. Notice that $\Gamma_s(0) = I$ and $\dim\ker(\Gamma_s(1) - I) = n$, but the path of symplectic matrices $s \mapsto \Gamma_s(1)$ does not have to be constant. For $k \in \mathbb{N}$ large enough, let us introduce the factorization

$$\Gamma_s(1) = P_{k-1,s} \circ P_{k-2,s} \circ ... \circ P_{0,s},$$

where $P_{j,s} := \Gamma_s(\frac{j+1}{l})\Gamma_s(\frac{j}{l})^{-1}$. We denote by $h_s : \mathbb{R}^{2dk} \times \mathbb{R}^{2dk} \to \mathbb{R}$ the Hessian bilinear form associated to this factorization of $\Gamma_s(1)$. As before, h_s depends continuously on s, and its nullity is constantly equal to n. This implies that the functions $s \mapsto \mathrm{ind}(h_s)$ and $s \mapsto \mathrm{coind}(h_s)$ are constant. In particular $\mathrm{mas}(\Gamma_0) = \mathrm{mas}(\Gamma_1)$ and $\mathrm{comas}(\Gamma_0) = \mathrm{comas}(\Gamma_1)$. □

3.3. **Bott's iteration theory for the Maslov index.** By combining Sections 2 and 3, we obtain an iteration theory for the Maslov index and coindex. Consider a continuous path $\Gamma : [0,1] \to \mathrm{Sp}(2d)$ with $\Gamma(0) = I$. Fix a parameter k large enough, and consider the factorization $\Gamma(1) = P_{k-1} \circ ... \circ P_0$ whose factors are defined by (26), and the associated quadratic generating family $F : \mathbb{R}^{2dk} \to \mathbb{R}$ given by (27). For $\theta \in S^1$, let H_θ be the θ-Hessian of F, and $h_\theta : \mathbb{C}^{2dk} \times \mathbb{C}^{2dk} \to \mathbb{C}$ the associated Hermitian bilinear form. We defined the **θ-Maslov index** and **coindex** of Γ as

$$\mathrm{mas}_\theta(\Gamma) := \mathrm{ind}(h_\theta) - dk, \qquad \mathrm{comas}_\theta(\Gamma) := \mathrm{coind}(h_\theta) - dk,$$

so that

$$\mathrm{mas}_\theta(\Gamma) + \mathrm{comas}_\theta(\Gamma) + \dim\ker(\Gamma(1) - \theta I) = 0.$$

These indices are well defined independently of the sufficiently large parameter k employed. Indeed, for $\theta = 1$ these are the standard Maslov index and coindex, and the fact that they are independent of k was already proved in the previous subsection. Moreover, Lemma 2.4 implies that the functions $\theta \mapsto \mathrm{mas}_\theta(\Gamma) - \mathrm{mas}_1(\Gamma)$ and $\theta \mapsto \mathrm{comas}_\theta(\Gamma) - \mathrm{comas}_1(\Gamma)$ are completely determined by the symplectic matrix $\Gamma(1)$.

Theorem 3.3 is generalized by the following.

Theorem 3.4. *The θ-Maslov index* $\mathrm{mas}_\theta : \mathcal{P}(2d) \to \mathbb{Z}$ *is a lower semi-continuous function and, for each $n \in \mathbb{N}$, is locally constant on the subspace*

$$\mathcal{P}_{\theta,n}(2d) := \big\{\Gamma : [0,1] \to \mathrm{Sp}(2d) \;\big|\; \Gamma(0) = I,\ \dim\ker(\Gamma(1) - \theta I) = n\big\}.$$

The same properties hold for the Maslov coindex.

Proof. The proof is entirely analogous to the one of Theorem 3.3. Briefly, the lower semi-continuity of the θ-Maslov index follows from the same property for the index of Hermitian bilinear forms. As for the other claim, consider a path $s \mapsto \Gamma_s$ inside a subspace $\mathcal{P}_{\theta,n}(2d)$. For k large enough, there exists a continuous family $h_{s,\theta} : \mathbb{C}^{2dk} \times \mathbb{C}^{2dk} \to \mathbb{C}$ of associated θ-Hessian Hermitian bilinear forms. Since

$$\mathrm{nul}(h_{s,\theta}) = \dim\ker(\Gamma_s(1) - \theta I) = n,$$

we readily have that the functions $s \mapsto \mathrm{ind}(h_{s,\theta})$ and $s \mapsto \mathrm{coind}(h_{s,\theta})$ are constant, and so are the functions $s \mapsto \mathrm{mas}(\Gamma_s)$ and $s \mapsto \mathrm{comas}(\Gamma_s)$. □

Let us provide the motivation for the introduction of such generalized Maslov indices. We define the **p-th iteration** of Γ as the continuous path $\Gamma_p : [0,1] \to \mathrm{Sp}(2d)$ given by

$$\Gamma_p(\tfrac{j+t}{p}) = \Gamma(t)\Gamma(1)^j, \qquad \forall j = 0, ..., p-1,\ t \in [0,1].$$

This notion arises naturally in the context of periodic Hamiltonian systems. Indeed, assume that $H_t : \mathbb{R}^{2d} \to \mathbb{R}$ is a smooth non-autonomous Hamiltonian that is 1-periodic in time, i.e. $H_{t+1} = H_t$ for all $t \in \mathbb{R}$. If H defines a global Hamiltonian flow ϕ_t, this verifies $\phi_{t+1} = \phi_t \circ \phi_1$ for all $t \in \mathbb{R}$. If now z is a fixed point of ϕ_1, we can linearize the flow at z, thus obtaining the symplectic path $\Gamma : \mathbb{R} \to \mathrm{Sp}(2d)$ given by $\Gamma(t) = \mathrm{d}\phi_t(z)$. The p-th iteration of the path $\Gamma|_{[0,1]}$ is the reparametrization of the path $\Gamma|_{[0,p]}$ given by $\Gamma_p(t) = \Gamma(pt)$, for $t \in [0,1]$.

The iterated path Γ_p defines a factorization of $\Gamma_p(1) = \Gamma(1)^p$ which is precisely the p-th fold juxtaposition of the original factorization $P_{k-1} \circ ... \circ P_0$ of $\Gamma(1)$. This puts ourselves in the setting of Section 2. In particular, Bott's formulae of Lemma 2.2 can be stated for the Maslov index and coindex as

$$\begin{aligned} \mathrm{mas}_\theta(\Gamma_p) &= \sum_{\mu \in \sqrt[p]{\theta}} \mathrm{mas}_\mu(\Gamma), \\ \mathrm{comas}_\theta(\Gamma_p) &= \sum_{\mu \in \sqrt[p]{\theta}} \mathrm{comas}_\mu(\Gamma). \end{aligned} \tag{32}$$

In particular, the Maslov index and coindex of any p-th iterate of Γ are completely determined by the functions $\theta \mapsto \mathrm{mas}_\theta(\Gamma)$ and $\theta \mapsto \mathrm{comas}_\theta(\Gamma)$ respectively.

Example 3.5. Let us compute the Maslov index and coindex of the symplectic path $\Gamma : [0,1] \to \mathrm{Sp}(2)$ given by rigid rotations from angle 0 to some angle $\beta > 0$, i.e.

$$\Gamma(t) = \begin{pmatrix} \cos(t\beta) & -\sin(t\beta) \\ \sin(t\beta) & \cos(t\beta) \end{pmatrix}.$$

Let $p \in \mathbb{N}$ be large enough so that, for all $t \in [0, 1/p]$, the symplectic matrix $\Gamma(t)$ is described by a generating function. This is verified precisely when the angle $\alpha := \beta/p$ lies in the interval $(0, \pi/2)$. We denote by $\Upsilon : [0,1] \to \mathrm{Sp}(2)$ the continuous path $\Upsilon(t) := \Gamma(pt)$, so that Γ is the p-th iteration of Υ. The matrix

$\Upsilon(1)$ is described by the quadratic generating function $F(z) = \frac{1}{2}h(z,z) = \frac{1}{2}\langle Hz, z\rangle$ whose Hessian matrix is

$$H = \begin{pmatrix} \tan(\alpha) & \cos(\alpha)^{-1} - 1 \\ \cos(\alpha)^{-1} - 1 & \tan(\alpha) \end{pmatrix}.$$

The eigenvalues $\lambda_1, \lambda_2 \in \mathbb{R}$ of this matrix satisfy $\lambda_1\lambda_2 = (1 - \cos(\alpha))\cos(\alpha)^{-2}$ and $\lambda_1 + \lambda_2 = 2\tan(\alpha)$. Therefore

$$\begin{aligned} \mathrm{mas}(\Upsilon) &= \mathrm{ind}(h) - 1 = -1, \\ \mathrm{comas}(\Upsilon) &= \mathrm{coind}(h) - 1 = 1. \end{aligned}$$

In order to compute the θ-Maslov indices we can make the same computation with the θ-Hessian of the generating function F, or equivalently apply the recipe from Section 2.4. Let us choose the second option. For all $\theta \in S^1$, the Hermitian bilinear form g_θ associated to the symplectic matrix $P = \Upsilon(1)$ as in (24) is given by

$$\begin{aligned} g_\theta((X,Y),(X',Y')) = {} & \sin(\alpha)\cos(\alpha)\langle X, X'\rangle \\ & + \cos(\alpha)(\cos(\alpha) - \theta)\langle Y, X'\rangle \\ & + \cos(\alpha)(\cos(\alpha) - \overline{\theta})\langle X, Y'\rangle \\ & + \sin(\alpha)(2\,\mathrm{Re}(\theta) - \cos(\alpha))\langle Y, Y'\rangle. \end{aligned}$$

The eigenvalues $\kappa_1, \kappa_2 \in \mathbb{R}$ of the associated Hermitian matrix

$$\begin{pmatrix} \sin(\alpha)\cos(\alpha) & \cos(\alpha)(\cos(\alpha) - \theta) \\ \cos(\alpha)(\cos(\alpha) - \overline{\theta}) & \sin(\alpha)(2\,\mathrm{Re}(\theta) - \cos(\alpha)) \end{pmatrix}.$$

satisfy $\kappa_1 + \kappa_2 = 2\,\mathrm{Re}(\theta)\sin(\alpha)$ and $\kappa_1\kappa_2 = 2\cos(\alpha)(\mathrm{Re}(\theta) - \cos(\alpha))$. In particular

$$\begin{aligned} \mathrm{ind}(g_{e^{i\alpha+}}) &= \mathrm{coind}(g_{e^{i\alpha+}}) = 1, \\ \mathrm{ind}(g_{e^{i\alpha}}) &= \mathrm{ind}(g_{e^{i\alpha-}}) = 0, \\ \mathrm{coind}(g_{e^{i\alpha}}) &= 1, \\ \mathrm{coind}(g_{e^{i\alpha-}}) &= 2. \end{aligned}$$

The splitting and cosplitting numbers of $\Upsilon(1)$ at the eigenvalue $e^{i\alpha}$ are given by

$$\begin{aligned} \mathrm{S}^+_{\Upsilon(1)}(e^{i\alpha}) &= \mathrm{coS}^-_{\Upsilon(1)}(e^{i\alpha}) = 1, \\ \mathrm{S}^-_{\Upsilon(1)}(e^{i\alpha}) &= \mathrm{coS}^+_{\Upsilon(1)}(e^{i\alpha}) = 0. \end{aligned}$$

Recall that

$$\begin{aligned} \mathrm{S}^+_{\Upsilon(1)}(\theta) &= \mathrm{mas}_{\theta+}(\Upsilon) - \mathrm{mas}_\theta(\Upsilon), \\ \mathrm{coS}^+_{\Upsilon(1)}(\theta) &= \mathrm{comas}_{\theta+}(\Upsilon) - \mathrm{comas}_\theta(\Upsilon). \end{aligned}$$

Therefore

$$\begin{aligned} \mathrm{mas}_\theta(\Upsilon) &= \begin{cases} -1 & \text{if } \arg(\theta) \in [-\alpha, \alpha], \\ 0 & \text{otherwise}, \end{cases} \\ \mathrm{comas}_\theta(\Upsilon) &= \begin{cases} 1 & \text{if } \arg(\theta) \in (-\alpha, \alpha), \\ 0 & \text{otherwise}. \end{cases} \end{aligned}$$

This computation, together with Bott's formulae (32), allows us to compute the Maslov index and coindex of the original path $\Gamma = \Upsilon_p$. Indeed, consider the subsets of complex p-th roots of unity

$$\mathbb{I} := \{\theta \in \sqrt[p]{1} \mid \arg(\theta) \in [-\alpha, \alpha]\}.$$

Its cardinality is given by

$$|\mathbb{I}| = 2\left\lfloor \frac{\beta}{2\pi} \right\rfloor + 1.$$

Bott's formulae for the Maslov index give

$$\mathrm{mas}(\Gamma) = \sum_{\theta \in \sqrt[p]{1}} \mathrm{mas}_\theta(\Upsilon) = |\mathbb{I}|\,\mathrm{mas}(\Upsilon) = -2\left\lfloor \frac{\beta}{2\pi} \right\rfloor - 1.$$

If β is not a multiple of 2π, Bott's formulae for the Maslov index give

$$\mathrm{comas}(\Gamma) = \sum_{\theta \in \sqrt[p]{1}} \mathrm{comas}_\theta(\Upsilon) = |\mathbb{I}|\,\mathrm{comas}(\Upsilon) = 2\left\lfloor \frac{\beta}{2\pi} \right\rfloor + 1,$$

whereas if β is a multiple of 2π, i.e. $\Gamma(1) = I$, they give

$$\begin{aligned}\mathrm{comas}(\Gamma) &= \sum_{\theta \in \sqrt[p]{1}} \mathrm{comas}_\theta(\Upsilon)\\ &= (|\mathbb{I}| - 2)\,\mathrm{comas}(\Upsilon) + 2\,\mathrm{comas}_{e^{i\alpha}}(\Upsilon)\\ &= 2\left\lfloor \frac{\beta}{2\pi} \right\rfloor - 1\\ &= \frac{\beta}{\pi} - 1. \qquad \square\end{aligned}$$

We conclude this section by rephrasing the iteration inequalities of Theorem 2.10 in the language of the Maslov index and coindex. In this form, the theorem is due to Liu and Long [LL98, LL00].

Theorem 3.6 (Iteration inequalities for the Maslov indices)**.** *Let* $\Gamma : [0,1] \to \mathrm{Sp}(2d)$ *be a continuous path such that* $\Gamma(0) = I$, *and let* $p \in \mathbb{N}$. *Then*

$$\begin{aligned} p\,\overline{\mathrm{mas}}(\boldsymbol{z}) - d &\leq \mathrm{mas}(\Gamma_p),\\ \mathrm{mas}(\Gamma_p) + \dim\ker(\Gamma(1)^p - I) &\leq p\,\overline{\mathrm{mas}}(\Gamma) + d, \end{aligned} \tag{33}$$

where $\overline{\mathrm{mas}}(\Gamma)$ *denotes the* **average Maslov index**, *given by*

$$\overline{\mathrm{mas}}(\Gamma) := \frac{1}{2\pi}\int_0^{2\pi} \mathrm{mas}_{e^{it}}(\Gamma)\,\mathrm{d}t = \lim_{p\to\infty} \frac{\mathrm{mas}(\Gamma^p)}{p} \in \mathbb{R}.$$

If at least one of the inequalities (33) *is an equality, then* $\sigma(\Gamma(1)^p) = \{1\}$ *and* $\dim\ker(\Gamma(1)^p - I) \geq d$. *Both inequalities are equalities if and only if* $\Gamma(1)^p = I$. $\square$

Remark 3.7. By (30), the inequalities (33) are equivalent to

$$\begin{aligned} p\,\overline{\mathrm{comas}}(\boldsymbol{z}) - d &\leq \mathrm{comas}(\Gamma_p),\\ \mathrm{comas}(\Gamma_p) + \dim\ker(\Gamma(1)^p - I) &\leq p\,\overline{\mathrm{comas}}(\Gamma) + d, \end{aligned}$$

where $\overline{\mathrm{comas}}(\Gamma)$ denotes the **average Maslov coindex**, given by

$$\overline{\mathrm{comas}}(\Gamma) := \frac{1}{2\pi}\int_0^{2\pi} \mathrm{comas}_{e^{it}}(\Gamma)\,\mathrm{d}t = \lim_{p\to\infty} \frac{\mathrm{comas}(\Gamma^p)}{p} \in \mathbb{R}. \qquad \square$$

3.4. **Bibliographical remarks.** The Maslov index has quite a long history. It was first introduced by Gel'fand and Lidskiĭ [GL58] as an index for the connected components of the space of strongly stable linear periodic Hamiltonian systems. It was later rediscovered by Maslov [Mas72] as an intersection number of a loop of Lagrangian subspaces with the so called Maslov cycle, a singular hypersurface in the Lagrangian Grassmannian. Conley and Zehnder reinterpreted the Maslov index as a relative Morse index in [CZ84a], and for this reason many authors in symplectic topology prefer the terminology **Conley-Zehnder index**. Our presentation of the Maslov index as a renormalized Morse index of Chaperon's generating families is analogous to Conley and Zehnder's one. This approach was already followed by Théret [Thé96, Chapter IV] for more general generating families of Lagrangian submanifolds of cotangent bundles. Théret inferred that the Maslov index is well defined (which is part of Theorem 3.3 above) as a consequence of Viterbo's uniqueness Theorem for generating families [Vit92, Thé99]. An alternative proof of the relation between Maslov and Morse indices was provided by Robbin and Salamon [RS93b]. In the references given so far, the Maslov index was considered only for "non-degenerate" paths, that is, for paths in $\mathcal{P}_0(2d)$. The first author who defined the Maslov index on the whole space of symplectic paths $\mathcal{P}(2d)$ was Long [Lon90], who later on also defined the θ-Maslov index and established its iteration theory à la Bott [Lon99]. Building on previous work of Conley and Zehnder, Long proved that the θ-Maslov index classifies the path-connected components of the space $\mathcal{P}_{\theta,0}(2d)$: two paths Γ_1 and Γ_2 belong to the same path-connected component of $\mathcal{P}_{\theta,0}(2d)$ if and only if $\mathrm{mas}_\theta(\Gamma_1) = \mathrm{mas}_\theta(\Gamma_2)$ or, equivalently, $\mathrm{comas}_\theta(\Gamma_1) = \mathrm{comas}_\theta(\Gamma_2)$. We refer the reader to the monograph [Lon02] for a comprehensive account of the θ-Maslov index and for the many applications. A different extension of the Maslov index to degenerate paths, which is widely employed in symplectic topology, was given by Robbin and Salamon in [RS93a].

4. The Lagrangian Morse index

4.1. **Tonelli Lagrangian and Hamiltonian systems.** In this section, we focus on a special class of Hamiltonian systems, for which the Maslov index can be described as a traditional Morse index of an action (without need of renormalization by a constant). This class can be described in the Lagrangian formulation as follows (we refer the reader to, e.g., [AM78, Arn78, Maz12] for a comprehensive treatment of Lagrangian dynamics). Let M be a manifold equipped with an auxiliary Riemannian metric. A **Tonelli Lagrangian** is a smooth time-dependent function $L_t : \mathrm{T}M \to \mathbb{R}$ such that $L_t = L_{t+1}$ and each function $v \mapsto L_t(q, v)$ has everywhere positive-definite Hessian and superlinear growth, i.e.

$$\begin{aligned} &\partial^2_{vv} L_t(q,v)[w,w] > 0, && \forall t \in \mathbb{R},\ (q,v) \in \mathrm{T}M,\ w \in \mathrm{T}_q M \setminus \{0\}, \\ &\lim_{|v|_q \to \infty} L_t(q,v)/|v|_q = \infty, && \forall t \in \mathbb{R},\ q \in M. \end{aligned}$$

A Tonelli Lagrangian defines a second-order partial flow on M, that is, a flow on the tangent bundle $\mathrm{T}M$ whose integral lines are velocity vectors of curves on M. These curves $\gamma : (T_0, T_1) \to M$ are solution of the Euler-Lagrange equation

$$\tfrac{\mathrm{d}}{\mathrm{d}t} \partial_v L_t(\gamma(t), \dot\gamma(t)) - \partial_q L_t(\gamma(t), \dot\gamma(t)) = 0.$$

Let us assume for simplicity that the solutions of this equation are defined for all time. This is always true if M is a closed manifold and the Lagrangian is autonomous, or more generally if its dependence on time is suitably controlled.

The fiberwise derivative $\partial_v L$ is a diffeomorphism of the tangent bundle $\mathrm{T}M$ onto the cotangent bundle T^*M. The dual **Tonelli Hamiltonian** $H_t : \mathrm{T}^*M \to \mathbb{R}$ is defined by

$$H_t(q,p) = \max_{v \in \mathrm{T}_q M} \{pv - L_t(q,v)\}.$$

This function still enjoys the Tonelli properties listed above: it is fiberwise convex and superlinear. Its fiberwise derivative $\partial_p H$ is the diffeomorphism inverse to $\partial_v L$, and we have $L(q,v) + H(q,p) = pv$, where $p = \partial_v L(q,v)$ and $v = \partial_p H(q,p)$. The velocity curve $t \mapsto (\gamma(t), \dot\gamma(t))$ of a solution of the Euler-Lagrange equation is mapped by $\partial_v L$ to an integral curve $t \mapsto (\gamma(t), \partial_v L(\gamma(t), \dot\gamma(t)))$ of the Hamiltonian flow of H. We recall that the Hamiltonian flow ϕ_H^t is the integral of the non-autonomous Hamiltonian vector field X_H, which with our convention is defined by $\omega(X_{H_t}, \cdot) = \mathrm{d}H_t$, where $\omega = \mathrm{d}q \wedge \mathrm{d}p$ is the canonical symplectic form on the cotangent bundle T^*M.

4.2. **The Lagrangian action functional.** A classical computation in calculus of variations shows that a smooth 1-periodic curve $\gamma : \mathbb{R} \to M$ is a solution of the Euler-Lagrange equation if and only if it is a critical point of the action functional $A : C^\infty(\mathbb{R}/\mathbb{Z}; M) \to \mathbb{R}$ given by

$$A(\gamma) = \int_0^1 L_t(\gamma(t), \dot\gamma(t))\, \mathrm{d}t.$$

We wish to investigate the properties of the Morse index $\mathrm{ind}(\gamma)$ of this functional at a critical point γ. For calligraphic convenience, let us assume that M is the Euclidean space $\mathbb{R}^d$, so that the dual Hamiltonian H defines a Hamiltonian flow on the standard symplectic $(\mathbb{R}^{2d}, \omega)$. We associate to γ the continuous path of symplectic matrices $\Gamma : [0,1] \to \mathrm{Sp}(2d)$ given by

$$\Gamma(t) := \mathrm{d}\phi_H^t(\gamma(0), \partial_v L(\gamma(0), \dot\gamma(0))).$$

Namely, Γ is the path that begins at the identity matrix $\Gamma(0) = I$ and follows the linearized Hamiltonian flow at the starting point of the Hamiltonian periodic orbit corresponding to γ.

A priori, we do not know whether the Morse index $\mathrm{ind}(\gamma)$ is finite. This is a consequence of the following theorem, whose proof will be given at the end of Section 4.3, after several preliminaries.

Theorem 4.1. *The Morse index of A at a critical point γ coincides with the Maslov index of the associated symplectic path Γ, i.e.* $\mathrm{ind}(\gamma) = \mathrm{mas}(\Gamma)$.

The Hessian of A at γ is the bilinear form on the infinite dimensional Fréchet space $C^\infty(\mathbb{R}/\mathbb{Z}; \mathbb{R}^d)$ given by

$$\mathrm{Hess}A(\gamma)[\xi,\eta] = \int_0^1 \Big(\langle \alpha\, \dot\xi, \dot\eta \rangle + \langle \beta\, \xi, \dot\eta \rangle + \langle \dot\xi, \beta\, \eta \rangle + \langle \delta\, \xi, \eta \rangle \Big) \mathrm{d}t,$$

where

$$\alpha_t := \partial_{vv} L_t(\gamma(t), \dot\gamma(t)), \quad \beta_t := \partial_{qv} L_t(\gamma(t), \dot\gamma(t)), \quad \delta_t := \partial_{qq} L_t(\gamma(t), \dot\gamma(t)).$$

Since we are only interested in this Hessian form and not in the action functional A itself, we can assume without loss of generality that the Lagrangian L has the form

$$L_t(q,v) = \tfrac{1}{2}\langle \alpha_t v, v\rangle + \langle \beta_t q, v\rangle + \tfrac{1}{2}\langle \delta_t q, q\rangle, \tag{34}$$

and that γ is the constant curve at origin. In this way, the Euler-Lagrange equation becomes linear of the form

$$\alpha\,\ddot{\xi} + (\dot{\alpha} + \beta - \beta^T)\dot{\xi} + (\dot{\beta} - \delta)\xi = 0, \tag{35}$$

and the action A becomes a quadratic function, i.e. $A(\xi) = \frac{1}{2}\mathrm{Hess}A(\gamma)[\xi,\xi]$. From now on, we will simply write $\mathrm{Hess}A$ for $\mathrm{Hess}A(\gamma)$.

Let us extend $\mathrm{Hess}A$ as a bilinear form on the Sobolev space $W^{1,2}(\mathbb{R}/\mathbb{Z};\mathbb{R}^d)$ of absolutely continuous curves with squared-integrable first derivative. One can show that the self-adjoint operator associated to this extension is Fredholm, and that the inertia index of the bilinear form is finite. Indeed, if the matrices β_t and δ_t were identically zero, the Hessian form would clearly be semi-positive definite, since it would reduce to the integral

$$\int_0^1 \langle \alpha\,\dot{\xi}, \dot{\eta}\rangle\,\mathrm{d}t,$$

and $\alpha(t)$ is a positive definite matrix; the kernel of this bilinear form is given by the constant curves $\xi \equiv \xi(0)$, in particular it has finite dimension d. The general case, when β_t and δ_t do not necessarily vanish identically, is a compact perturbation of this special one. When we add a compact perturbation to a semi-positive definite Fredholm bilinear form, the index of the resulting form is finite (see e.g. [Maz12, Lemma 2.1.2 and errata corrige]). Therefore, $\mathrm{Hess}A$ has finite index. Since $C^\infty(\mathbb{R}/\mathbb{Z};\mathbb{R}^d)$ is dense in $W^{1,2}(\mathbb{R}/\mathbb{Z};\mathbb{R}^d)$, one can show that the Morse index of the Hessian is the same whether we consider it as a bilinear form on $C^\infty(\mathbb{R}/\mathbb{Z};\mathbb{R}^d)$ or on $W^{1,2}(\mathbb{R}/\mathbb{Z};\mathbb{R}^d)$.

Consider now, for each integer $k \geq 2$, the vector space

$$\mathbb{E}_k := \big\{\xi \in C^0(\mathbb{R}/\mathbb{Z};\mathbb{R}^d) \;\big|\; \xi|_{[j/k,(j+1)/k]} \text{ is a solution of (35) } \forall j \in \mathbb{Z}_k\big\}.$$

Notice that $\mathbb{E}_k$ has finite dimension dk, and the evaluation map

$$\xi \mapsto (\xi(0), \xi(1/k), ..., \xi((k-1)/k))$$

is an isomorphism of $\mathbb{E}_k$ onto $\mathbb{R}^{dk}$. Moreover $\mathbb{E}_k \subset \mathbb{E}_{2k}$. As k increases, $\mathbb{E}_k$ contains finer and finer approximations of any given smooth 1-periodic curve. Actually, one can show that the union of all the $\mathbb{E}_k$'s is dense in the Sobolev space $W^{1,2}(\mathbb{R}/\mathbb{Z};\mathbb{R}^d)$, and therefore that

$$\mathrm{ind}(\mathrm{Hess}A) = \mathrm{ind}(\mathrm{Hess}A|_{\mathbb{E}_k\times\mathbb{E}_k}), \qquad \forall k \geq 2 \text{ large enough.}$$

We recall that the action A is assumed to be a quadratic function. In particular, a curve ξ is in the kernel of $\mathrm{Hess}A$ if and only if it is a critical point of A, that is, if and only if it is a solution of the Euler-Lagrange equation (35). Therefore

$$\ker(\mathrm{Hess}A) = \ker(\mathrm{Hess}A|_{\mathbb{E}_k\times\mathbb{E}_k}), \qquad \forall k \geq 2.$$

For more details on this, we refer the reader to [Maz12, Section 4.4].

Let us have a look at the expression of the Hessian of A on the space $\mathbb{E}_k$. For all $\xi,\eta\in\mathbb{E}_k$, we have

$$\begin{aligned}
\mathrm{Hess}A[\xi,\eta] &= \int_0^1 \Big(\langle\alpha\,\dot\xi,\dot\eta\rangle + \langle\beta\,\xi,\dot\eta\rangle + \langle\dot\xi,\beta\,\eta\rangle + \langle\delta\,\xi,\eta\rangle\Big)\,\mathrm{d}t\\
&= \sum_{j=0}^{k-1}\int_{j/k}^{(j+1)/k} \langle \underbrace{-\,\alpha\,\ddot\xi - (\dot\alpha+\beta-\beta^T)\dot\xi - (\dot\beta-\delta)\xi}_{=0},\eta\rangle\,\mathrm{d}t\\
&\quad + \sum_{j=0}^{k-1}\langle\alpha\dot\xi+\beta\xi,\sigma\rangle\Big|_{j^+/k}^{(j+1)^-/k}\\
&= \sum_{j=0}^{k-1}\langle\alpha_{j/k}\big(\dot\xi(\tfrac{j}{k}^-)-\dot\xi(\tfrac{j}{k}^+)\big),\sigma(\tfrac{j}{k})\rangle
\end{aligned}$$

It will be more convenient to write down this Hessian in a slightly different way as follows. We denote by ϕ_L^t the Euler-Lagrange flow on the tangent bundle $\mathrm{T}\mathbb{R}^d=\mathbb{R}^{2d}$, which is defined by $\phi_L^t(\xi(0),\dot\xi(0))=(\xi(t),\dot\xi(t))$ if $\xi:[0,t]\to\mathbb{R}^d$ is a solution of the Euler-Lagrange equation. We set

$$Q_j := \phi_L^{(j+1)/k}\circ(\phi_L^{j/k})^{-1},\qquad \forall j=0,...,k-1,$$

so that $\phi_L^{j/k}=Q_j\circ...\circ Q_0$, and we denote by $\pi_1:\mathbb{R}^{2d}\to\mathbb{R}^d$ the projection $\pi(X,V)=X$. We introduce the vector space

$$\mathbb{V} := \big\{(X_0,V_0,...,X_{k-1},V_{k-1})\in\mathbb{R}^{2dk}\ \big|\ \pi_1\circ Q_j(X_j,V_j)=X_{j+1}\quad\forall j\in\mathbb{Z}_k\big\}\,.$$

Notice that there is an isomorphism $\Psi:\mathbb{E}_k\to\mathbb{V}$ given by

$$\Psi(\xi)=\big(\xi(0),\dot\xi(0^+),\xi(\tfrac1k),\dot\xi(\tfrac1k^+),...,\xi(\tfrac{k-1}{k}),\dot\xi(\tfrac{k-1}{k}^+)\big).$$

If we pull-back the Hessian of the action A by the isomorphism Ψ^{-1}, we obtain the simmetric bilinear form $h_L:\mathbb{V}\times\mathbb{V}\to\mathbb{R}$ that reads

$$h_L(\boldsymbol{Z},\boldsymbol{Z}')=\mathrm{Hess}A[\Psi^{-1}\boldsymbol{Z},\Psi^{-1}\boldsymbol{Z}']=\sum_{j\in\mathbb{Z}_k}\langle\alpha_{j/k}(\tilde V_j-V_j),X_j'\rangle,$$

where $\boldsymbol{Z}=(X_0,V_0,...,X_{k-1},V_{k-1})$, $\boldsymbol{Z}'=(X_0',V_0',...,X_{k-1}',V_{k-1}')$, and we have adopted the notation $(X_{j+1},\tilde V_{j+1})=Q_j(X_j,V_j)$. Summing up, in order to prove Theorem 4.1, we have to show that

$$\mathrm{ind}(h_L)=\mathrm{mas}(\Gamma). \tag{36}$$

4.3. **The generating family of a Tonelli Hamiltonian flow.** Let us now focus on the linear Hamiltonian flow ϕ_H^t, which we discretize by setting

$$P_j := \phi_H^{(j+1)/k}\circ(\phi_H^{j/k})^{-1},\qquad\forall j=0,...,k-1.$$

Notice that the matrices P_j are related to the matrices Q_j of the previous subsection by

$$P_j\circ\partial_v L_{j/k}=\partial_v L_{(j+1)/k}\circ Q_j, \tag{37}$$

and $\partial_v L_t(x,v) = (x, \alpha_t v + \beta_t q)$. Since our parameter k is assumed to be large enough, each symplectic matrix $P_j \in \mathrm{Sp}(2d)$ is close to the identity, and therefore admits a quadratic generating function

$$f_j(X_{j+1}, Y_j) = \tfrac{1}{2}\langle A_j X_{j+1}, X_{j+1}\rangle + \langle B_j X_{j+1}, Y_j\rangle + \tfrac{1}{2}\langle C_j Y_j, Y_j\rangle,$$

where A_j, B_j, and C_j are (small) $dk \times dk$ real matrices, A_j and C_j being symmetric. As we know, this means that

$$P_j Z_j = Z_{j+1} \qquad \text{if and only if} \qquad \begin{cases} X_{j+1} - X_j = -B_j X_{j+1} - C_j Y_j, \\ Y_{j+1} - Y_j = A_j X_{j+1} + B_j^T Y_j. \end{cases}$$

Let us show the precise relationship between the Hamiltonian H and the generating functions f_j.

Lemma 4.2. *If $(X(t), Y(t)) := \phi_H^t(X(0), Y(0))$ is an orbit of the Hamiltonian flow and we set $(X_j, Y_j) := (X(j/k), Y(j/k))$, we have*

$$f_j(X_{j+1}, Y_j) = \langle Y_j, X_j - X_{j+1}\rangle + \int_{j/k}^{(j+1)/k} \Big(\langle Y(t), \dot X(t)\rangle - H_t(X(t), Y(t))\Big)\,\mathrm{d}t.$$

Proof. For syntactic convenience, let us focus on the case $j = 0$, the other cases being completely analogous. Consider the primitive $-y\,\mathrm{d}x$ of the symplectic form $\omega = \mathrm{d}x \wedge \mathrm{d}y$. Since the Hamiltonian flow ϕ_H^t is symplectic, $(\phi_H^t)^* y\,\mathrm{d}x - y\,\mathrm{d}x$ is a closed 1-form, hence exact by the Poincaré Lemma. Let $g_0 : \mathbb{R}^{2d} \to \mathbb{R}$ be a function defined up to an additive constant by

$$\mathrm{d}g_0 = (\phi_H^{1/k})^* y\,\mathrm{d}x - y\,\mathrm{d}x. \tag{38}$$

By applying the Fundamental Theorem of Calculus to the right-hand side of this equation, we obtain

$$\begin{aligned} \mathrm{d}g_0 &= (\phi_H^{1/k})^* y\,\mathrm{d}x - y\,\mathrm{d}x \\ &= \int_0^{1/k} (\phi_H^t)^* \mathcal{L}_{X_{H_t}}(y\,\mathrm{d}x)\,\mathrm{d}t \\ &= \int_0^{1/k} (\phi_H^t)^* (\mathrm{d}(y\,\mathrm{d}x(X_{H_t})) - \omega(X_{H_t}, \cdot))\,\mathrm{d}t \\ &= \mathrm{d}\left(\int_0^{1/k} (\phi_H^t)^* (y\,\mathrm{d}x(X_{H_t}) - H_t)\,\mathrm{d}t\right) \end{aligned}$$

If we normalize g_0 by setting $g_0(0) = 0$, we have

$$g_0 = \int_0^{1/k} (\phi_H^t)^* (y\,\mathrm{d}x(X_{H_t}) - H_t)\,\mathrm{d}t.$$

By evaluating this expression at the starting point (X_0, Y_0) of our orbit, we obtain

$$g_0(X_0, Y_0) = \int_0^{1/k} \Big(\langle Y(t), \dot X(t)\rangle - H_t(X(t), Y(t))\Big)\,\mathrm{d}t.$$

Notice that g_0 is a quadratic function.

Now, let us consider X_1 and Y_0 as independent variables, while $X_0 = X_0(X_1, Y_0)$ and $Y_1 = Y_1(X_1, Y_0)$. More precisely, we denote by $R_0 : \mathbb{R}^2 \to \mathbb{R}^2$ the linear isomorphism such that $P_0(X_0,Y_0)=(X_1,Y_1)$ if and only if $R_0(X_1,Y_0)=(X_0,Y_0)$. Equation (38) becomes

$$\begin{aligned}
\mathrm{d}(g_0\circ R_0) &= R_0^*(\mathrm{d}g_0)\\
&= Y_1\,\mathrm{d}X_1 - Y_0\,\mathrm{d}X_0\\
&= (Y_1-Y_0)\,\mathrm{d}X_1 - Y_0\,(\mathrm{d}X_0-\mathrm{d}X_1)\\
&= \underbrace{(Y_1-Y_0)}_{\partial_{X_1}f_0}\,\mathrm{d}X_1 + \underbrace{(X_0-X_1)}_{\partial_{Y_0}f_0}\,\mathrm{d}Y_0 - \mathrm{d}(\langle Y_0, X_0-X_1\rangle)\\
&= \mathrm{d}f_0 - \mathrm{d}(\langle Y_0,X_0-X_1\rangle).
\end{aligned}$$

This defines the generating function f_0 up to a constant. Since f_0 is a quadratic function, it vanishes at the origin, and therefore we conclude

$$f_0 = \langle Y_0, X_0-X_1\rangle + g_0\circ Q_0. \qquad \square$$

Remark 4.3. The above proof works with any (not necessarily linear) Hamiltonian flow, except that the functions f_0 and g_0 are not quadratic anymore and therefore can only be defined up to an additive constant.

Lemma 4.2 allows us to translate the Tonelli fiberwise convexity property of the Hamiltonian H to a concavity property for the generating functions f_j.

Lemma 4.4. *If the parameter k is large enough, each matrix C_j is negative definite.*

Proof. Let us compute the explicit expression of our Hamiltonian H dual to the quadratic Lagrangian (34). Given $(q,v)\in\mathbb{R}^{2d}$, the dual moment variable p is given by

$$(q,p)=\partial_v L_t(q,v) = (q, \alpha_t v+\beta_t q).$$

Therefore

$$\begin{aligned}
H_t(q,p) &= pv - L(q,v)\\
&= \langle p, \alpha^{-1}(p-\beta_t q)\rangle - L(q,\alpha^{-1}(p-\beta_t q))\\
&= \tfrac12\langle\alpha^{-1}p,p\rangle - \langle\alpha_t^{-1}\beta_t q,p\rangle + \tfrac12\langle(\beta_t^T\alpha_t^{-1}\beta_t-\delta_t)q,q\rangle.
\end{aligned}$$

Let $Y_j\in\mathbb{R}^d$ and $X_j:=C_jY_j$, so that $P_j(X_j,Y_j)=(0,Y_{j+1})$. By Lemma 4.2, we have

$$\begin{aligned}
\langle C_jY_j,Y_j\rangle &= 2f_j(0,Y_j)\\
&= 2\langle Y_j,X_j\rangle + 2\int_{j/k}^{(j+1)/k}\Big(\langle Y(t),\dot X(t)\rangle - H_t(X(t),Y(t))\Big)\,\mathrm{d}t\\
&= 2\int_{j/k}^{(j+1)/k}\Big(-\langle \dot Y(t),X(t)\rangle - H_t(X(t),Y(t))\Big)\,\mathrm{d}t\\
&= 2\int_{j/k}^{(j+1)/k}\Big(\partial_q H(X(t),Y(t))\,X(t) - H_t(X(t),Y(t))\Big)\,\mathrm{d}t\\
&= \int_{j/k}^{(j+1)/k}\Big(-\langle\alpha_t^{-1}Y(t),Y(t)\rangle + \langle(\beta_t^T\alpha_t^{-1}\beta_t-\delta_t)X(t),X(t)\rangle\Big)\,\mathrm{d}t\\
&\leq -a\int_{j/k}^{(j+1)/k}|Y(t)|^2\,\mathrm{d}t + b\int_{j/k}^{(j+1)/k}|X(t)|^2\,\mathrm{d}t,
\end{aligned}$$

where

$$a := \min_{t\in\mathbb{R}/\mathbb{Z}} \left|\alpha_t^{-1}\right| > 0, \qquad b := \max_{t\in\mathbb{R}/\mathbb{Z}} \left|\beta_t^T\alpha_t^{-1}\beta_t - \delta_t\right|.$$

We recall that the Hamiltonian flow ϕ_H^t is linear. For all $\epsilon > 0$ there exists $k \in \mathbb{N}$ large enough such that, for all $t_1, t_2 \in [0,1]$ with $|t_1 - t_2| \leq 1/k$, we have

$$|\phi_H^{t_1} \circ (\phi_H^{t_2})^{-1} - I| < \epsilon.$$

In other words, if $t \mapsto Z(t) = (X(t), Y(t))$ is a non-zero integral curve of the Hamiltonian flow ϕ_H^t, we have

$$\frac{|Z(t_1) - Z(t_2)|}{|Z(t_2)|} < \epsilon, \qquad \forall t_1, t_2 \in [0,1] \text{ with } |t_1 - t_2| \leq 1/k,$$

and if we further assume that $X(t_2) = 0$, we infer

$$|X(t_1)| < \epsilon|Y(t_2)|, \qquad |Y(t_1)| > (1-\epsilon)|Y(t_2)|.$$

By plugging these inequalities into the estimate for $\langle C_jY_j, Y_j\rangle$ above, we obtain

$$\langle C_jY_j, Y_j\rangle \leq -a\frac{(1-\epsilon)^2}{k}|Y_{j+1}|^2 + b\frac{\epsilon^2}{k}|Y_{j+1}|^2 = \underbrace{(-a(1-\epsilon)^2 + b\epsilon^2)}_{(*)}\frac{|Y_{j+1}|^2}{k},$$

and the term $(*)$ is negative provided ϵ is small enough. $\square$

Proof of Theorem 4.1. Consider the quadratic generating family $F : \mathbb{R}^{2dk} \to \mathbb{R}$ associated to the factorization $\phi_H^1 = P_{k-1} \circ ... \circ P_0$. We recall that the Hessian bilinear form $h : \mathbb{R}^{2dk} \times \mathbb{R}^{2dk} \to \mathbb{R}$ of F is given by

$$\begin{aligned} h(\boldsymbol{Z}, \boldsymbol{Z}') = &\sum_{j\in\mathbb{Z}_k} \langle Y_{j-1} - Y_j + A_{j-1}X_j + B_{j-1}^TY_{j-1}, X_j'\rangle \\ &+ \sum_{j\in\mathbb{Z}_k} \langle X_{j+1} - X_j + B_jX_{j+1} + C_jY_j, Y_j'\rangle. \end{aligned}$$

We wish to take advantage of the fact that the matrices C_j are negative definite (Lemma 4.4) in order to compute the Morse index of h and, a fortiori, the Maslov index of Γ. We introduce the vector subspace

$$\mathbb{W} := \big\{(X_0, Y_0, ..., X_{k-1}, Y_{k-1}) \in \mathbb{R}^{2dk} \bigm| X_j = 0 \quad \forall j = 0, ..., k-1\big\},$$

and its h-orthogonal

$$\mathbb{W}^h = \left\{ (X_0, Y_0, ..., X_{k-1}, Y_{k-1}) \in \mathbb{R}^{2dk} \;\middle|\; \begin{array}{r} X_{j+1} - X_j + B_jX_{j+1} + C_jY_j = 0 \\ \forall j = 0, ..., k-1 \end{array} \right\},$$

For all $\boldsymbol{Z}, \boldsymbol{Z}' \in \mathbb{W}$, we have

$$h(\boldsymbol{Z}, \boldsymbol{Z}') = \sum_{j\in\mathbb{Z}_k} \langle C_jY_j, Y_j'\rangle.$$

Therefore, $h|_{\mathbb{W}\times\mathbb{W}}$ is a negative definite bilinear form, and in particular

$$\mathrm{ind}(h|_{\mathbb{W}\times\mathbb{W}}) = \dim \mathbb{W} = dk.$$

The intersection $\mathbb{W} \cap \mathbb{W}^h$ is given by those vectors $\boldsymbol{Z} \in \mathbb{R}^{2dk}$ such that $X_j = C_jY_j = 0$ for all $j = 0, ..., k-1$. Since the matrices C_j are negative definite, they are invertible, and therefore the intersection $\mathbb{W} \cap \mathbb{W}^h$ is trivial. Since $\mathbb{W} \cap \ker(h)$ is contained in $\mathbb{W} \cap \mathbb{W}^h$, it is trivial as well. By applying Proposition A.3, we obtain

$$\mathrm{ind}(h) = \mathrm{ind}(h|_{\mathbb{W}\times\mathbb{W}}) + \mathrm{ind}(h|_{\mathbb{W}^h\times\mathbb{W}^h}) = dk + \mathrm{ind}(h|_{\mathbb{W}^h\times\mathbb{W}^h}).$$

By rephrasing in terms of the Maslov index of the path Γ, we have

$$\mathrm{mas}(\Gamma) = \mathrm{ind}(h) - dk = \mathrm{ind}(h|_{\mathbb{W}^h\times\mathbb{W}^h}).$$

Let us now focus on the form $h|_{\mathbb{W}^h\times\mathbb{W}^h}$. We denote by $\pi_1 : \mathbb{R}^{2d} \to \mathbb{R}^d$ the projection $\pi_1(X, Y) = X$. Notice that the vector space $\mathbb{W}^h$ can be characterized as

$$\mathbb{W}^h := \left\{(X_0, Y_0, ..., X_{k-1}, Y_{k-1}) \in \mathbb{R}^{2dk} \,\middle|\, \pi_1 \circ P_j(X_j, Y_j) = X_{j+1} \quad \forall j \in \mathbb{Z}_k\right\}.$$

In particular, $\mathbb{W}^h$ is isomorphic to the vector space $\mathbb{V}$ of the previous subsection via the isomorphism $\Omega : \mathbb{V} \to \mathbb{W}^h$ given by

$$\Omega(X_0, V_0, ..., X_{k-1}, V_{k-1}) = (X_0, Y_0, ..., X_{k-1}, Y_{k-1}),$$

where

$$(X_j, Y_j) = \partial_v L_{j/k}(X_j, V_j) = (X_j, \alpha_{j/k} V_j + \beta_{j/k} X_j).$$

We also set

$$\tilde{Y}_j := Y_{j-1} + A_{j-1} X_j + B^T_{j-1} Y_{j-1}, \qquad \forall j \in \mathbb{Z}_k,$$

so that $P_j(X_j, Y_j) = (X_{j+1}, \tilde{Y}_{j+1})$. We recall the notation of the previous subsection: we write $Q_j(X_j, V_j) = (X_{j+1}, \tilde{V}_{j+1})$. Equation (37) implies that the vectors $\tilde{V}_j$ and $\tilde{Y}_j$ are related by the usual duality

$$(X_j, \tilde{Y}_j) = \partial_v L_{j/k}(X_j, \tilde{V}_j) = (X_j, \alpha_{j/k} \tilde{V}_j + \beta_{j/k} X_j).$$

For all $\boldsymbol{Z}, \boldsymbol{Z}' \in \mathbb{W}^h$, we have

$$\begin{aligned} h(\boldsymbol{Z}, \boldsymbol{Z}') &= \sum_{j\in\mathbb{Z}_k} \langle Y_{j-1} - Y_j + A_{j-1} X_j + B^T_{j-1} Y_{j-1}, X'_j\rangle \\ &= \sum_{j\in\mathbb{Z}_k} \langle \tilde{Y}_j - Y_j, X'_j\rangle. \end{aligned}$$

If we pull-back $h|_{\mathbb{W}^h\times\mathbb{W}^h}$ by the isomorphism Ω, we obtain

$$\begin{aligned} h(\Omega\boldsymbol{Z}, \Omega\boldsymbol{Z}') &= \sum_{j\in\mathbb{Z}_k} \langle \alpha_{j/k}\tilde{V}_j + \beta_{j/k} X_j - \alpha_{j/k} V_j - \beta_{j/k} X_j, X'_j\rangle \\ &= \sum_{j\in\mathbb{Z}_k} \langle \alpha_{j/k}(\tilde{V}_j - V_j), X'_j\rangle \\ &= h_L(\boldsymbol{Z}, \boldsymbol{Z}'). \end{aligned}$$

In particular $\mathrm{ind}(h|_{\mathbb{W}^h\times\mathbb{W}^h}) = \mathrm{ind}(h_L)$. This completes the proof of (36), and thus of Theorem 4.1. □

4.4. **Bibliographical remarks.** Historically, the first statement of the kind of Theorem 4.1 above is the Index Theorem from Riemannian geometry [Mil63, Section 15], asserting that the Morse index of a geodesic with prescribed endpoints is given by its number of conjugate points counted with multiplicity. Indeed, this count corresponds to the Maslov index of an associated path of Lagrangian subspaces. The periodic orbit case for Tonelli Lagrangian systems was first established by Duistermaat [Dui76]. The proof that we provided in this section is conceptually similar to the one given by Abbondandolo in [Abb03]. See also [Vit87, AL98, Abb01, Lon02] for other proofs and related results.

Appendix A. Some linear algebra

A.1. **Eigenspaces of power matrices.** A non-diagonalizable squared complex matrix M must have an eigenvalue λ whose algebraic multiplicity is strictly larger than its geometric one. If $\lambda = 0$ with algebraic multiplicity n, then the algebraic multiplicity of the eigenvalue λ becomes equal to its geometric one for the power matrix M^n. The next proposition shows that this never occurs for non-zero eigenvalues.

Proposition A.1. *For every squared complex matrix M we have*

$$\dim_{\mathbb{C}} \ker(M^n - \theta I) = \sum_{\mu \in \sqrt[n]{\theta}} \dim_{\mathbb{C}} \ker(M - \mu I), \qquad \forall n \in \mathbb{N},\ \theta \in \mathbb{C} \setminus \{0\}.$$

Proof. Assume without loss of generality that M is in Jordan normal form, with Jordan blocks $M_1, ..., M_r$. Hence, its n-th power M^n is a block-diagonal matrix with blocks $M_1^n, ..., M_r^n$, and since

$$\dim_{\mathbb{C}} \ker(M^n - \theta I) = \sum_{j=1}^{r} \dim_{\mathbb{C}} \ker(M_j^n - \theta I),$$

it suffices to prove the proposition for the case in which $M = M_1$ is a single Jordan block with eigenvalue $\mu \neq 0$, i.e.

$$M = \begin{pmatrix} \mu & 1 & & & \\ & \mu & 1 & & \\ & & \ddots & \ddots & \\ & & & \mu & 1 \\ & & & & \mu \end{pmatrix}.$$

In this case, the claim of the proposition reduces to

$$\dim_{\mathbb{C}} \ker(M^n - \mu^n I) = \dim_{\mathbb{C}} \ker(M - \mu I) = 1.$$

By a straightforward computation, we can verify that the power matrix M^n is still upper-triangular, where the entries in the diagonal are all equal to μ^n, while the entries in the super-diagonal are all equal to $n\,\mu^{n-1}$. The matrix $M^n - \mu^n I$ is upper-triangular, with entries in the diagonal all equal to zero, and entries in the super-diagonal all equal to $n\,\mu^{n-1}$. In particular, the first column of $M^n - \mu^n I$ is the zero one, while the other columns are linearly independent. This proves that the kernel of $M^n - \mu^n I$ is one-dimensional. $\square$

A.2. **Inertia of restricted Hermitian forms.** Let H be a Hermitian $d\times d$ matrix, and $h:\mathbb{C}^d\times\mathbb{C}^d\to\mathbb{C}$ the associated Hermitian form $h(v,w)=\langle Hv,w\rangle$. We recall the definition of the **inertia** triple of h: the **index** $\mathrm{ind}(h)$ equal to the maximal dimension of a vector subspace over which h is negative definite, the **coindex** $\mathrm{coind}(h)=\mathrm{ind}(-h)$ equal to the maximal dimension of a vector subspace over which h is positive definite, and the **nullity** $\mathrm{nul}(h)$ equal to the dimension of the kernel of h, that is, the kernel of the matrix H. Of course, in a Hermitian setting, dimension will always stand for complex dimension. If the matrix H is real, the exact same results of this section hold for the real simmetric bilinear form $h|_{\mathbb{R}^d\times\mathbb{R}^d}$ by replacing complex dimension with real dimension in all the formulae (as well as in the definition of index, coindex, and nullity of $h|_{\mathbb{R}^d\times\mathbb{R}^d}$).

Given a complex vector subspace $\mathbb{V}\subseteq\mathbb{C}^d$, its h-orthogonal is the complex vector subspace defined by

$$\mathbb{V}^h=\big\{w\in\mathbb{C}^d\ \big|\ h(w,v)=0\quad\forall v\in\mathbb{V}\big\}.$$

It readily follows from its definition that

$$\mathbb{V}^h=(H\mathbb{V})^\perp=H^{-1}(\mathbb{V}^\perp),$$

where $\perp$ denotes the orthogonal with respect to the Hermitian inner product $\langle\cdot,\cdot\rangle$. Moreover,

$$(\mathbb{V}^h)^h=H^{-1}H\mathbb{V}=\mathbb{V}+\ker(H).$$

The inertia of h is related to the one of the restricted forms $h|_{\mathbb{V}\times\mathbb{V}}$ and $h|_{\mathbb{V}^h\times\mathbb{V}^h}$ according to the following statements.

Proposition A.2. $\mathrm{nul}(h)=\mathrm{nul}(h|_{\mathbb{V}^h\times\mathbb{V}^h})-\dim_{\mathbb{C}}(\mathbb{V}\cap\mathbb{V}^h)+\dim_{\mathbb{C}}(\mathbb{V}\cap\ker(H))$.

Proof. The kernel of the Hermitian matrix associated to the restricted form $h|_{\mathbb{V}^h\times\mathbb{V}^h}$ is given by

$$\begin{aligned}\ker(h|_{\mathbb{V}^h\times\mathbb{V}^h})&=\big\{v\in\mathbb{V}^h\ \big|\ Hv\in((H\mathbb{V})^\perp)^\perp\big\}\\&=\big\{v\in\mathbb{V}^h\ \big|\ Hv\in H\mathbb{V}\big\}\\&=\big\{v\in\mathbb{V}^h\ \big|\ v\in H^{-1}H\mathbb{V}\big\}\\&=(\mathbb{V}+\ker(H))\cap\mathbb{V}^h.\end{aligned}$$

Notice that $\ker(H)\subset\mathbb{V}^h$. Moreover

$$\begin{aligned}\dim_{\mathbb{C}}((\mathbb{V}+\ker(H))\cap\mathbb{V}^h)&=\dim_{\mathbb{C}}(\mathbb{V}+\ker(H))+\dim_{\mathbb{C}}(\mathbb{V}^h)\\&\quad-\dim_{\mathbb{C}}(\mathbb{V}+\ker(H)+\mathbb{V}^h)\\&=\dim_{\mathbb{C}}(\mathbb{V})+\dim_{\mathbb{C}}\ker(H)-\dim_{\mathbb{C}}(\mathbb{V}\cap\ker(H))\\&\quad+\dim_{\mathbb{C}}(\mathbb{V}^h)-\dim_{\mathbb{C}}(\mathbb{V}+\mathbb{V}^h)\\&=\dim_{\mathbb{C}}(\mathbb{V}\cap\mathbb{V}^h)+\mathrm{nul}(h)-\dim_{\mathbb{C}}(\mathbb{V}\cap\ker(H)).\end{aligned}$$

These two equations prove the proposition. $\square$

Proposition A.3.

$$\mathrm{ind}(h)=\mathrm{ind}(h|_{\mathbb{V}\times\mathbb{V}})+\mathrm{ind}(h|_{\mathbb{V}^h\times\mathbb{V}^h})+\dim_{\mathbb{C}}(\mathbb{V}\cap\mathbb{V}^h)-\dim_{\mathbb{C}}(\mathbb{V}\cap\ker(H)),$$

$$\mathrm{coind}(h)=\mathrm{coind}(h|_{\mathbb{V}\times\mathbb{V}})+\mathrm{coind}(h|_{\mathbb{V}^h\times\mathbb{V}^h})+\dim_{\mathbb{C}}(\mathbb{V}\cap\mathbb{V}^h)-\dim_{\mathbb{C}}(\mathbb{V}\cap\ker(H)).$$

Proof. Since the coindex of a quadratic form is equal to the index of minus the same quadratic form, it is enough to prove the equality for the index. We first give a proof in case h is a non-degenerate bilinear form, that is, in case the associated Hermitian matrix H is invertible. Under this assumption, the last summand on the right-hand side of the equality that we want to prove is zero. Moreover

$$\mathbb{V} \cap \mathbb{V}^h = \ker(h|_{\mathbb{V}\times\mathbb{V}}) = \ker(h|_{\mathbb{V}^h\times\mathbb{V}^h}).$$

The restricted Hermitian form $h|_{\mathbb{V}\times\mathbb{V}}$ can be written as

$$h|_{\mathbb{V}\times\mathbb{V}}(v,w) = \langle P_{\mathbb{V}} \circ H|_{\mathbb{V}} v, w\rangle,$$

where $P_{\mathbb{V}} : \mathbb{C}^d \to \mathbb{V}$ is the orthogonal projector onto $\mathbb{V}$, which is an Hermitian linear map. Notice that $P_{\mathbb{V}} \circ H|_{\mathbb{V}}$ is Hermitian. In particular it is diagonalizable and has only real eigenvalues. Therefore, the vector subspace $\mathbb{V}$ splits as the direct sum

$$\mathbb{V} = \mathbb{E}^- \oplus \mathbb{E}^+ \oplus (\mathbb{V} \cap \mathbb{V}^h),$$

where $\mathbb{E}^-$ is the direct sum of the eigenspaces of $P_{\mathbb{V}} \circ H|_{\mathbb{V}}$ corresponding to negative eigenvalues, while $\mathbb{E}^+$ is the direct sum of the eigenspaces corresponding to positive eigenvalues. These three vector spaces in the direct-sum decomposition of $\mathbb{V}$ are orthogonal with respect to both the Hermitian inner product $\langle\cdot,\cdot\rangle$ and the Hermitian form $h|_{\mathbb{V}\times\mathbb{V}}$. The inertia of this latter form is precisely

$$\begin{aligned}
\operatorname{ind}(h|_{\mathbb{V}\times\mathbb{V}}) &= \dim_{\mathbb{C}} \mathbb{E}^-,\\
\operatorname{coind}(h|_{\mathbb{V}\times\mathbb{V}}) &= \dim_{\mathbb{C}} \mathbb{E}^+,\\
\operatorname{nul}(h|_{\mathbb{V}\times\mathbb{V}}) &= \mathbb{V} \cap \mathbb{V}^h.
\end{aligned}$$

Since $\mathbb{E}^+$ and $\mathbb{E}^-$ are invariant by the linear map $P_{\mathbb{V}} \circ H$, we have that $H(\mathbb{E}^\pm) \subset \mathbb{E}^\pm + \mathbb{V}^\perp$. Let us introduce an analogous splitting

$$\mathbb{V}^h = \mathbb{F}^+ \oplus \mathbb{F}^- \oplus (\mathbb{V} \cap \mathbb{V}^h),$$

where $\mathbb{F}^-$ and $\mathbb{F}^+$ are the direct sum of the eigenspaces of $P_{\mathbb{V}^h} \circ H|_{\mathbb{V}^h}$ corresponding to the negative eigenvalues and to the positive eigenvalues respectively. As before, we have

$$\begin{aligned}
\operatorname{ind}(h|_{\mathbb{V}^h\times\mathbb{V}^h}) &= \dim_{\mathbb{C}} \mathbb{F}^-,\\
\operatorname{coind}(h|_{\mathbb{V}^h\times\mathbb{V}^h}) &= \dim_{\mathbb{C}} \mathbb{F}^+,\\
\operatorname{nul}(h|_{\mathbb{V}^h\times\mathbb{V}^h}) &= \mathbb{V} \cap \mathbb{V}^h.
\end{aligned}$$

Notice that the vector subspaces $\mathbb{E}^-$ and $\mathbb{F}^-$ are h-orthogonal, and in particular the form h is negative definite on the subspace $\mathbb{E}^- \oplus \mathbb{F}^-$. The analogous consideration holds for the vector subspaces $\mathbb{E}^+$ and $\mathbb{F}^+$. The matrix H maps the intersection $\mathbb{V} \cap \mathbb{V}^h$ isomorphically onto $\mathbb{V}^\perp \cap (\mathbb{V}^h)^\perp = (\mathbb{V} + \mathbb{V}^h)^\perp$. We fix a real constant $\lambda \in (0, 2/\|H^{-1}\|)$, and we introduce the vector subspaces

$$\begin{aligned}
\mathbb{G}^- &:= \{v - \lambda H^{-1} v \mid v \in \mathbb{V} \cap \mathbb{V}^h\},\\
\mathbb{G}^+ &:= \{v + \lambda H^{-1} v \mid v \in \mathbb{V} \cap \mathbb{V}^h\}.
\end{aligned}$$

The form h is negative definite on $\mathbb{G}^-$. Indeed, for all $v \in \mathbb{V} \cap \mathbb{V}^h$,

$$\begin{aligned} h(v - \lambda H^{-1}v, v - \lambda H^{-1}v) &= h(v,v) - 2\lambda h(H^{-1}v, v) + \lambda^2 h(H^{-1}v, H^{-1}v) \\ &= -2\lambda\|v\|^2 + \lambda^2 \langle v, H^{-1}v\rangle \\ &\leq \lambda\|v\|^2 \underbrace{\big(- 2 + \lambda\|H^{-1}\|\big)}_{<0}. \end{aligned}$$

Analogously, h is positive definite on $\mathbb{G}^+$. The vector spaces $\mathbb{G}^\pm$ are h-orthogonal to $\mathbb{E}^\pm \oplus \mathbb{F}^\pm$, since for all $v \in \mathbb{V} \cap \mathbb{V}^h$ and $w \oplus z \in \mathbb{E}^\pm \oplus \mathbb{F}^\pm$ we have

$$\begin{aligned} h(w + z, v \pm \lambda H^{-1}v) &= h(w, v) + h(z, v) \pm \lambda\, h(w, H^{-1}v) \pm \lambda\, h(z, H^{-1}v) \\ &= \pm\lambda\langle w, v\rangle \pm \lambda\langle z, v\rangle \\ &= 0. \end{aligned}$$

We conclude that h is negative definite on $\mathbb{E}^- \oplus \mathbb{F}^- \oplus \mathbb{G}^-$ and positive definite on $\mathbb{E}^+ \oplus \mathbb{F}^+ \oplus \mathbb{G}^+$. Since the direct sum of these two vector subspaces is the whole $\mathbb{C}^d$, we have that

$$\begin{aligned} \operatorname{ind}(h) &= \dim_{\mathbb{C}}(\mathbb{E}^-) + \dim_{\mathbb{C}}(\mathbb{F}^-) + \dim_{\mathbb{C}}(\mathbb{G}^-) \\ &= \operatorname{ind}(h|_{\mathbb{V}\times\mathbb{V}}) + \operatorname{ind}(h|_{\mathbb{V}^h\times\mathbb{V}^h}) + \dim_{\mathbb{C}}(\mathbb{V} \cap \mathbb{V}^h), \end{aligned} \tag{39}$$

which is the identity that we wanted to prove.

Let us now relax the assumption that h is non-degenerate, and call $\mathbb{K} := \ker(h) = \ker(H)$. The form h induces a non-degenerate bilinear form h' on the quotient $\mathbb{C}^d/\mathbb{K}$ simply by

$$h'(v + \mathbb{K}, w + \mathbb{K}) = h(v, w).$$

Any vector subspace of $\mathbb{C}^d/\mathbb{K}$ is of the form $\mathbb{V}/\mathbb{K}$, for some vector subspace $\mathbb{V} \subseteq \mathbb{C}^d$, and this correspondence behaves naturally with respect to the passage to the h-orthogonal, i.e.

$$(\mathbb{V}/\mathbb{K})^{h'} = \mathbb{V}^h/\mathbb{K}.$$

By applying (39) to the non-degenerate Hermitian form h', we obtain

$$\operatorname{ind}(h') = \operatorname{ind}(h'|_{\mathbb{V}/\mathbb{K}\times\mathbb{V}/\mathbb{K}}) + \operatorname{ind}(h'|_{\mathbb{V}^h/\mathbb{K}\times\mathbb{V}^h/\mathbb{K}}) + \dim_{\mathbb{C}}(\mathbb{V}/\mathbb{K} \cap \mathbb{V}^h/\mathbb{K}).$$

Notice that

$$\begin{aligned} \operatorname{ind}(h') &= \operatorname{ind}(h), \\ \operatorname{ind}(h'|_{\mathbb{V}/\mathbb{K}\times\mathbb{V}/\mathbb{K}}) &= \operatorname{ind}(h|_{\mathbb{V}\times\mathbb{V}}), \\ \operatorname{ind}(h'|_{\mathbb{V}^h/\mathbb{K}\times\mathbb{V}^h/\mathbb{K}}) &= \operatorname{ind}(h|_{\mathbb{V}^h\times\mathbb{V}^h}). \end{aligned}$$

Finally, since $\mathbb{K} = \ker(H) \subset \mathbb{V}^h$,

$$\dim_{\mathbb{C}}(\mathbb{V}/\mathbb{K} \cap \mathbb{V}^h/\mathbb{K}) = \dim_{\mathbb{C}}((\mathbb{V} \cap \mathbb{V}^h)/\mathbb{K}) = \dim_{\mathbb{C}}(\mathbb{V} \cap \mathbb{V}^h) - \dim_{\mathbb{C}}(\mathbb{V} \cap \ker(H)).$$

This completes the proof. ∎

A.3. **Generalized eigenspaces of symplectic matrices.** Consider the standard symplectic vector space $(\mathbb{R}^{2d}, \omega)$. The symplectic form ω can be extended to a non-degenerate skew-Hermitian form on $\mathbb{C}^{2d}$ by setting

$$\omega(\lambda z, z') = \lambda\, \omega(z, z') = \omega(z, \overline{\lambda} z'), \qquad \forall z, z' \in \mathbb{R}^{2d},\ \lambda \in \mathbb{C}.$$

We denote by $\mathrm{Sp}(2d, \mathbb{C})$ the complex symplectic group, which is given by the $2d \times 2d$ complex matrices P such that

$$\omega(z, z') = \omega(Pz, Pz'), \qquad \forall z, z' \in \mathbb{C}^{2d}.$$

Notice that $\mathrm{Sp}(2d, \mathbb{C})$ contains the real symplectic group $\mathrm{Sp}(2d) = \mathrm{Sp}(2d, \mathbb{R})$, which is given by the matrices as above having zero imaginary part. Given a complex symplectic matrix $P \in \mathrm{Sp}(2d, \mathbb{C})$, we are interested in its generalized eigenspaces

$$\mathbb{F}_\lambda := \ker(P - \lambda I)^{2d}, \qquad \lambda \in \mathbb{C}.$$

Notice that the generalized eigenspaces span $\mathbb{C}^{2d}$, i.e.

$$\mathbb{C}^{2d} = \bigoplus_{\lambda \in \sigma(P)} \mathbb{F}_\lambda.$$

Lemma A.4. *Given a pair of eigenvalues $\lambda, \theta \in \mathbb{C}$ of a complex symplectic matrix $P \in \mathrm{Sp}(2d, \mathbb{C})$ such that $\lambda\overline{\theta} \neq 1$, the generalized eigenspaces $\mathbb{F}_\lambda$ and $\mathbb{F}_\theta$ are ω-orthogonal, i.e. $\omega(z, z') = 0$ for all $z \in \mathbb{F}_\lambda$ and $z' \in \mathbb{F}_\theta$.*

Proof. Consider two arbitrary generalized eigenvectors $z \in \mathbb{F}_\lambda$ and $z' \in \mathbb{F}_\theta$. We will prove the lemma by induction on the sum of the ranks of z and z'. If $(P - \lambda I)^n z = (P - \theta I)^m z' = 0$ with $n = m = 1$, we have

$$\omega(z, z') = \omega(Pz, Pz') = \lambda\overline{\theta}\, \omega(z, z'),$$

which implies that $\omega(z, z') = 0$ since $\lambda\overline{\theta} \neq 1$. Let us make the inductive hypothesis that $\omega(z, z') = 0$ holds whenever $n + m \leq k$.

Consider z and z' such that $n + m = k + 1$, and set $w := (P - \lambda I)z$ and $w' := (P - \theta I)z'$. The generalized eigenvectors w and w' have rank $n-1$ and $m-1$ respectively. By the inductive hypothesis, we have

$$\omega(z, w') = \omega(w, z') = \omega(w, w') = 0,$$

which implies

$$\omega(z, Pz') = \overline{\theta}\, \omega(z, z'),$$
$$\omega(Pz, z') = \lambda\, \omega(z, z'),$$

and

$$\omega(Pz, Pz') = \lambda\, \omega(z, Pz') + \overline{\theta}\, \omega(Pz, z') - \lambda\overline{\theta}\, \omega(z, z') = \lambda\overline{\theta}\, \omega(z, z').$$

Since P is a symplectic matrix, this latter equality becomes $\omega(z, z') = \lambda\overline{\theta}\, \omega(z, z')$, and as before this implies $\omega(z, z') = 0$. $\square$

Consider now a real symplectic matrix $P \in \mathrm{Sp}(2d)$, and the real generalized eigenspace

$$\mathbb{E}_1 := \ker(P - I)^{2d} \subset \mathbb{R}^{2d}.$$

Lemma A.5. *The space $\mathbb{E}_1$ is a (possibly zero dimensional) symplectic vector subspace of $(\mathbb{R}^{2d}, \omega)$.*

Proof. Consider the complex generalized eigenspaces of P, which give the direct sum decomposition $\mathbb{C}^{2d} = \mathbb{F}_1 \oplus \mathbb{F}'$, where

$$\mathbb{F}' = \bigoplus_{\lambda \neq 1} \mathbb{F}_\lambda.$$

By Lemma A.4, the vector subspaces $\mathbb{F}_1$ and $\mathbb{F}'$ are ω-orthogonal. Since ω is a non-degenerate skew-Hermitian form on $\mathbb{C}^{2d}$, this implies that its restriction to $\mathbb{F}_1$ is non-degenerate. Since ω is a real bilinear form, its restriction to the real part of $\mathbb{F}_1$ must be non-degenerate as well. But the real part of $\mathbb{F}_1$ is precisely $\mathbb{E}_1$. $\square$

References

[Abb01] A. Abbondandolo, *Morse theory for Hamiltonian systems*, Chapman & Hall/CRC Research Notes in Mathematics, vol. 425, Chapman & Hall/CRC, Boca Raton, FL, 2001.

[Abb03] ———, *On the Morse index of Lagrangian systems*, Nonlinear Anal. **53** (2003), no. 3-4, 551–566.

[AL98] T. An and Y. Long, *On the index theories for second order hamiltonian systems*, Nonlinear Anal. **34** (1998), no. 4, 585–592.

[AM78] R. Abraham and J. E. Marsden, *Foundations of mechanics*, Benjamin/Cummings Publishing Co. Inc. Advanced Book Program, 1978, Second edition.

[Arn78] V. I. Arnold, *Mathematical methods of classical mechanics*, Springer-Verlag, New York, 1978.

[Bot56] R. Bott, *On the iteration of closed geodesics and the Sturm intersection theory*, Comm. Pure Appl. Math. **9** (1956), 171–206.

[BTZ82] W. Ballmann, G. Thorbergsson, and W. Ziller, *Closed geodesics on positively curved manifolds*, Ann. of Math. **116** (1982), no. 2, 213–247.

[Cha84] M. Chaperon, *Une idée du type "géodésiques brisées" pour les systèmes hamiltoniens*, C. R. Acad. Sci. Paris Sér. I Math. **298** (1984), no. 13, 293–296.

[Cha85] ———, *An elementary proof of the Conley-Zehnder theorem in symplectic geometry*, Dynamical systems and bifurcations (Groningen, 1984), Lecture Notes in Math., vol. 1125, Springer, Berlin, 1985, pp. 1–8.

[CZ84a] C. C. Conley and E. Zehnder, *Morse-type index theory for flows and periodic solutions for Hamiltonian equations*, Comm. Pure Appl. Math. **37** (1984), no. 2, 207–253.

[CZ84b] ———, *Subharmonic solutions and Morse theory*, Phys. A **124** (1984), no. 1-3, 649–657.

[Dui76] J. J. Duistermaat, *On the Morse index in variational calculus*, Advances in Math. **21** (1976), no. 2, 173–195.

[GG10] V. L. Ginzburg and B. Z. Gürel, *Local Floer homology and the action gap*, J. Symplectic Geom. **8** (2010), no. 3, 323–357.

[GG12] ———, *Conley conjecture for negative monotone symplectic manifolds*, Int. Math. Res. Not. (2012), no. 8, 1748–1767.

[Gin10] V. L. Ginzburg, *The Conley conjecture*, Ann. Math. **172** (2010), no. 2, 1127–1180.

[GL58] I. M. Gel'fand and V. B. Lidskiĭ, *On the structure of the regions of stability of linear canonical systems of differential equations with periodic coefficients*, Amer. Math. Soc. Transl. (2) **8** (1958), 143–181.

[Hed32] G. A. Hedlund, *Poincaré's rotation number and morse's type number*, Trans. Amer. Math. Soc. **34** (1932), 75–77.

[Hei12] D. Hein, *The Conley conjecture for irrational symplectic manifolds*, J. Symplectic Geom. **10** (2012), no. 2, 183–202.

[Hin09] N. Hingston, *Subharmonic solutions of Hamiltonian equations on tori*, Ann. Math. **170** (2009), no. 2, 529–560.

[Hör71] L. Hörmander, *Fourier integral operators. I*, Acta Math. **127** (1971), no. 1-2, 79–183.

[HZ94] H. Hofer and E. Zehnder, *Symplectic invariants and Hamiltonian dynamics*, Birkhäuser Advanced Texts: Basler Lehrbücher, Birkhäuser Verlag, Basel, 1994.

[LL98] C. Liu and Y. Long, *An optimal increasing estimate of the iterated Maslov-type indices*, Chinese Sci. Bull. **43** (1998), no. 13, 1063–1066.

[LL00] ———, *Iteration inequalities of the Maslov-type index theory with applications*, J. Differential Equations **165** (2000), no. 2, 355–376.

[Lon90] Y. Long, *Maslov-type index, degenerate critical points, and asymptotically linear Hamiltonian systems*, Sci. China Ser. A **33** (1990), no. 12, 1409–1419.

[Lon99] ______, *Bott formula of the maslov-type index theory*, Pacific J. Math. **187** (1999), no. 1, 113–149.

[Lon02] ______, *Index theory for symplectic paths with applications*, Progress in Mathematics, vol. 207, Birkhäuser Verlag, Basel, 2002.

[Mas72] V. P. Maslov, *Théorie des perturbations et méthodes asymptotiques*, Dunod, Paris, 1972.

[Maz12] M. Mazzucchelli, *Critical point theory for Lagrangian systems*, Progress in Mathematics, vol. 293, Birkhäuser/Springer Basel AG, Basel, 2012.

[Maz13] ______, *Symplectically degenerate maxima via generating functions*, Math. Z. **275** (2013), no. 3-4, 715–739.

[Mil63] J. Milnor, *Morse theory*, Based on lecture notes by M. Spivak and R. Wells. Annals of Mathematics Studies, No. 51, Princeton University Press, Princeton, N.J., 1963.

[MP34] M. Morse and E. Pitcher, *On certain invariants of closed extremals*, Proc. Nat. Acad. Sci. **20** (1934), 282–287.

[MS98] D. McDuff and D. Salamon, *Introduction to symplectic topology*, second ed., Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, New York, 1998.

[RS93a] J. Robbin and D. Salamon, *The Maslov index for paths*, Topology **32** (1993), no. 4, 827–844.

[RS93b] ______, *Phase functions and path integrals*, Symplectic geometry, London Math. Soc. Lecture Note Ser., vol. 192, Cambridge Univ. Press, Cambridge, 1993, pp. 203–226.

[Sik87] J.-C. Sikorav, *Problèmes d'intersections et de points fixes en géométrie hamiltonienne*, Comment. Math. Helv. **62** (1987), no. 1, 62–73.

[SZ92] D. Salamon and E. Zehnder, *Morse theory for periodic solutions of Hamiltonian systems and the Maslov index*, Comm. Pure Appl. Math. **45** (1992), no. 10, 1303–1360.

[Thé96] D. Théret, *Utilisation des fonctions génératrices en géométrie symplectique globale*, Ph.D. Thesis, Université Paris-VII, 1996.

[Thé99] ______, *A complete proof of Viterbo's uniqueness theorem on generating functions*, Topology Appl. **96** (1999), 249–266.

[Vit87] C. Viterbo, *Intersection de sous-variétés lagrangiennes, fonctionnelles d'action et indice des systèmes hamiltoniens*, Bull. Soc. Math. France **115** (1987), no. 3, 361–390.

[Vit92] ______, *Symplectic topology as the geometry of generating functions*, Math. Ann. **292** (1992), no. 4, 685–710.

CNRS and École Normale Supérieure de Lyon, UMPA
69364 Lyon Cedex 07, France
E-mail address: marco.mazzucchelli@ens-lyon.fr

Publicaciones Matemáticas del Uruguay
Volumen 16, Julio 2016, Páginas 127–167
S 0797-1443

INTRODUCTION TO NON-UNIFORM AND PARTIAL HYPERBOLICITY

RAFAEL POTRIE

Abstract. These are notes for a minicourse given at Regional Norte UdelaR in Salto, Uruguay for the conference "CIMPA Research School - Hamiltonian and Langrangian Dynamics". The purpose of the notes is to present the theory of non-uniformly hyperbolic diffeomorphisms trying to concentrate in some simplified contexts and explain some of the main techniques in the field. Some of the topics include: Lyapunov exponents, Invariant manifolds (Pesin theory and persistence properties) and dynamical consequences. The topics will help introduce some concepts for the second part of the minicourse given by M.C. Arnaud but will also cover some topics of independent interest.

1. Introduction

The dynamics of uniformly hyperbolic systems is by now quite well understood in many aspects; for example: the spectral decomposition theorem allows one to decompose the dynamics in basic pieces which admit a quite precise coding (via Markov partitions) and the thermodynamical formalism provides information on the ergodic properties of invariant measures which have relevant dynamical or geometric meaning (see [Sh, KH], for example).

Of course, the understanding of uniformly hyperbolic systems is not complete, but there are many reasons for considering weaker forms of hyperbolicity. An important reason is that conservative dynamics are rarely uniformly hyperbolic[1].

There are essentially two ways to weaken uniform hyperbolicity: one consists on weakening the uniformity, by allowing to see hyperbolicity in almost every orbit but so that to see the hyperbolicity one has to "wait" a different amount of time depending on the point (this is called *non-uniform hyperbolicity*); the other consists in retaining the uniformity, but weakening the hyperbolicity by allowing certain bundles to be neutral yet "dominated" by the uniformly hyperbolic ones (this is called *partial hyperbolicity*).

In this notes, we pretend to give a unified view of this two generalizations by trying to study the dynamics from a local point of view, building charts around each point and considering the dynamics of sequences of diffeomorphisms of an Euclidean space. The main results we present have to do with the construction of invariant manifolds and the point of view is to try first to explain the (easier) case of periodic points and then try to convince the reader that the arguments go through

[1]In the conservative setting, being uniformly hyperbolic is the same as being Anosov, and it is well know that this imposes several strong restrictions on the topology of the manifold and isotopy class of the diffeomorphism (see [KH]). Moreover, there are also some local obstructions (such as possessing totally elliptic periodic points).

in these more general settings albeit some heavier notation and some adjustments on the statements.

This text has a strong subjective selection of topics and it is by no means a survey of the subject. It is intended as a first introduction to these topics which should be then complemented and deepened by the use of the standard references such as [KH, Sh, HPS] or others. Even if the text lacks a complete presentation of results, we have tried to provide at least a glimpse on further developments and ramifications of the subject. This choice has been even more subjective and depends heavily on the taste of the author.

1.1. **Organization of the notes.** In section 2 we give some preliminaries on ergodic theory which are relevant to what follows; in particular we provide a sketch of the proof of Oseledet's theorem in dimension 2. In this section we start to show the analogies between periodic orbits and ergodic measures.

In section 3 we show how one can pass the information on the tangent dynamics back to the manifold. This is probably the most important section of the notes and where the proof of the stable manifold theorem for periodic points is done in quite some detail and then the study of Pesin's charts and manifolds is explained. In section 4 we give a glimpse on the classical theory of non-uniform hyperbolicity and in section 5 we do the same with partial hyperbolicity and dominated splittings.

Finally, we end in section 6 presenting some applications of the previous result and explaining a recent result joint with Sylvain Crovisier and Martín Sambarino dealing with the geometry of partially hyperbolic attractors.

1.2. **Acknowledgments.** Thanks to E. Maderna and L. Rifford for the invitation to give the mini-course and to M.C. Arnaud for accepting to share the course with me. Discussions with S. Crovisier were important in the preparation of this notes (as well as in learning the material) and I thank him also for sharing with me some preliminary notes ([Cr_2]) he wrote about similar subjects. Thanks also to A. Passeggi for discussions and comments on the writting and G. Contreras for interesting input on the contributions of R. Mañé. Finally, I wish to dedicate this text to Ricardo Mañé, I did not know him personally, but his work has been a true inspiration for the whole mathematical community in Uruguay. The author was partially supported by Grupo CSIC 618.

2. Basics in differentiable ergodic theory

This section is devoted to present some basic results of ergodic theory which will be needed in the rest of the text. We shall restrict to the specific context we are interested in: M will denote a closed manifold and $f : M \to M$ a diffeomorphism of M. We refer the reader to [M_4] or [KH, Chapters 4 and 5] for a more complete account.

2.1. **Invariant and ergodic measures.** A probability measure μ in M will be said to be f-*invariant* if for every measurable set $A \subset M$ one has $\mu(f^{-1}(A)) = \mu(A)$.

We denote as $\mathcal{M}(f)$ the set of f-invariant probability measures. It is a standard fact that it is a compact convex subset of the space of measures with the weak-$*$ topology.

Exercise. Show that $\mathcal{M}(f)$ is non empty. (Hint: Consider the *empirical measures* $\mu_{n,x} = \frac{1}{n}\sum_{i=0}^{n-1}\delta_{f^i(x)}$ which are not invariant but as n grows, the defect of invariance decreases to 0).

There is a special important class of invariant measures which are called *ergodic*. A measure μ is called *ergodic* if every f-invariant set A verifies that either $\mu(A) = 0$ or $\mu(A^c) = 0$. We denote by $\mathcal{M}_{erg}(f)$ the subset of $\mathcal{M}(f)$ consisting of ergodic measures.

Exercise. Show that a f-invariant probability measure μ is ergodic if and only if for every f-invariant function φ one has that φ is constant μ-a.e.

One has that $\mathcal{M}_{erg}(f)$ is precisely the set of extremal points of $\mathcal{M}(f)$ (see [M$_4$]).

2.2. **Ergodic theorems.** We say that a sequence $\varphi_n : M \to \mathbb{R}$ is *subaditive* with respect to $f : M \to M$ if $\varphi_{n+m}(x) \leq \varphi_n(f^m(x)) + \varphi_m(x)$. The following result is by now classical:

Theorem 2.1 (Kingman)**.** *Let $f : M \to M$ preserving a measure μ and $\varphi_n : M \to M$ a subaditive sequence of functions such that $\varphi_1 \in L^1(\mu)$. Then, the sequence $\frac{1}{n}\varphi_n(x)$ converges μ-a.e. and in $L^1(\mu)$ to a f-invariant function $\tilde{\varphi} : M \to \mathbb{R}$ in $L^1(\mu)$ such that:*

$$\int \tilde{\varphi} d\mu = \inf_n \frac{1}{n}\int \varphi_n d\mu$$

A particularly concise proof of the pointwise convergence can be found in [AvB$_2$] (a proof which is in turn based on a proof of T. Kamae of Birkhoff's ergodic theorem which we partially reproduce below).

Given a function $\varphi : M \to \mathbb{R}$ we denote its n-th *Birkhoff sum* as:

$$S_n\varphi(x) = \sum_{i=0}^{n-1} \varphi(f^i(x))$$

It follows directly that the sequence $S_n\varphi$ is subaditive (in fact, additive) so the following is a direct consequence of Kingman's Theorem.

Theorem 2.2 (Birkhoff)**.** *Let $f : M \to M$ preserving a measure μ and $\varphi \in L^1(\mu)$. Then, the sequence $\frac{1}{n}S_n\varphi$ converges μ-ae and in $L^1(\mu)$ to a f-invariant function $\tilde{\varphi} \in L^1(\mu)$ and it follows that:*

$$\int \tilde{\varphi} d\mu = \int \varphi d\mu$$

In particular, if μ is ergodic then $\tilde{\varphi}(x) = \int \varphi d\mu$ for μ-ae x.

We give below a proof of the theorem for the particular (and important) case where φ is the characteristic function of a measurable subset $A \subset M$.

Proof of Theorem 2.2 for characteristic functions. (This should be skipped in a first reading.) Let $A \subset M$ be a measurable set and denote as $\varphi_n(x) = S_n\chi_A(x)$. Consider the following functions:

$$\underline{\tau}_A(x) = \liminf_n \frac{1}{n}\varphi_n(x) \quad ; \quad \overline{\tau}_A(x) = \limsup_n \frac{1}{n}\varphi_n(x)$$

Notice that one has that

$$\underline{\tau}_A(x) = \liminf_n \frac{1}{n}\varphi_n(x) = \liminf_n \frac{1}{n}(\chi_A(x) + \varphi_{n-1}(f(x))) = \underline{\tau}_A(f(x))$$

and therefore $\underline{\tau}_A$ is f-invariant. A symmetric argument shows that $\overline{\tau}_A$ is also f-invariant.

We want to show that for μ-almost every $x \in M$, one has that $\underline{\tau}_A(x) = \overline{\tau}_A(x) = \mu(A)$. Since one has obviously that $\underline{\tau}_A(x) \leq \overline{\tau}_A(x)$ for every x, it is enough to show that:

$$\int_M \underline{\tau}_A \geq \mu(A) \geq \int_M \overline{\tau}_A$$

The proofs are symmetric, so we shall only show that $\int_M \underline{\tau}_A \geq \mu(A)$.

Exercise. Use Fatou's lemma to show that $\int_M \underline{\tau}_A \leq \mu(A)$.

To show the inequality, fix $\varepsilon > 0$ and consider the sets

$$E_k = \{x \in M \ : \ \exists 1 \leq j \leq k \text{ such that } \frac{1}{j}\varphi_j(x) \leq \underline{\tau}_A(x) + \varepsilon\}$$

One has that $M = \bigcup_k E_k$ modulo a set of μ-measure zero.

We consider the functions $\psi_k : M \to [0,1]$ defined as follows: if $x \in E_k$ then $\psi_k(x) = \underline{\tau}_A(x) + \varepsilon$ and if $x \notin E_k^c$ then $\psi_k(x) = 1 + \varepsilon$. One has that the sequence ψ_k decreases to $\underline{\tau}_A(x) + \varepsilon$ as $k \to \infty$ (note that $\underline{\tau}_A(x) \leq 1$ for every $x \in M$).

By how we have defined ψ_k, whenever $n \geq k$ and $x \in E_k$, there is $j > 0$ such that $\varphi_n(x) = \varphi_{n-j}(f^j(x)) + \varphi_j(x)$ and such that $\varphi_j \leq j(\underline{\tau}_A(x) + \varepsilon)$. Since $\underline{\tau}_A(x)$ is invariant, one can write this as $\varphi_j(x) \leq \sum_{i=0}^{j-1} \psi_k(f^i(x))$. If $x \notin E_k$ it follows that $\varphi_n(x) = \varphi_{n-1}(f(x)) + \varphi_1(x)$ and $\varphi_1(x) \leq \psi_k(x)$ since $\varphi_1(x) \leq 1 < 1 + \varepsilon = \psi_k(x)$. Using this fact inductively, we know that for every x and $n \geq k$:

$$\varphi_n(x) = \sum_{i=0}^{n-1} \chi_A(f^i(x)) \leq k + \sum_{i=0}^{n-k-1} \psi_k(f^i(x)),$$

integrating in all M and using f-invariance of μ, one obtains:

$$\int_M \varphi_n \leq k + (n-k)\int_M \psi_k$$

Again by invariance of μ, one has that $\int_M \varphi_n(x) = n\mu(A)$, one deduces that:

$$n\mu(A) \leq k + (n-k)\int_M \psi_k$$

dividing by n and letting $n \to \infty$ one deduces:

$$\mu(A) \leq \int_M \psi_k$$

By monotone convergence one deduces that $\mu(A) \leq \int_M \underline{\tau}_A + \varepsilon$. Since ε was arbitrary, we deduce that $\mu(A) \geq \int_M \overline{\tau}_A$. Using a symmetric argument one obtains the other inequality and this concludes the proof of pointwise convergence of $\frac{1}{n}\varphi_n$. Since the functions are bounded by an integrable function, dominated convergence implies the L^1-convergence.

$\square$

Exercise. Show that if a function $\phi : M \to \mathbb{R}$ verifies that $\phi \circ f - \phi$ is integrable, then $\lim_n \frac{1}{n}\phi(f^n(x)) = 0$ for μ-almost every $x \in M$.

2.3. **Periodic orbits and their splittings.** Let p be a fixed point of a C^1-diffeomorphism $f : M \to M$, that is, such that $f(p) = p$. It follows that $Df_p : T_pM \to T_pM$ induces a linear transformation of T_pM which is a finite dimensional linear space. As a consequence of the Jordan decomposition well known in linear algebra, one deduces that there exists a Df_p-invariant decomposition $T_pM = E_1 \oplus \ldots \oplus E_k$ associated to the[2] eigenvalues $\lambda_1, \ldots, \lambda_k$ of the linear transformation Df_p. One has that if a vector $v \in E_i \setminus \{0\}$ then the following is verified:

$$\lim_{n\to\pm\infty} \frac{1}{n} \log \|Df_p^n v\| = \log |\lambda_i|$$

Exercise. Let A be a matrix such that all eigenvalues have the same modulus equal to λ. Show that for every non-zero vector one has that $\lim_{n\to\pm\infty} \frac{1}{n} \log \|A^n v\| = \log |\lambda|$.

Once we have chosen to split the space in the eigenspaces corresponding to the eigenvalues of the same modulus, it is clear that the decomposition is unique.

A similar situation occurs when one has a periodic point p for f, i.e. $f^m(p) = p$ for some $m \geq 1$. Then, one obtains that p is a fixed point of f^m and therefore the splitting $T_pM = E_1(p) \oplus \ldots \oplus E_k(p)$ is Df_p^m-invariant and verifies that if $v \in E_i(p) \setminus \{0\}$:

$$\lim_{n\to\pm\infty} \frac{1}{n} \log \|Df_p^n v\| = \frac{1}{m} \log |\lambda_i|$$

If one considers $f^j(p)$ for some j, it is also a fixed point for f^m and therefore one can define a Df^m-invariant splitting $T_{f^j(p)}M = E_1(f^j(p)) \oplus \ldots \oplus E_k(f^j(p))$. Notice that k is independent of the iterate $f^j(p)$ since the linear transformations Df_p^m and $Df^m_{f^j(p)}$ are conjugate:

$$Df_p^m = Df^{-j}_{f^j(p)} Df^m_{f^j(p)} Df_p^j = (Df_p^j)^{-1} Df^m_{f^j(p)} Df_p^j$$

It follows from uniqueness that the relation: $Df^i_{f^j(p)} E_\ell(f^j(p)) = E_\ell(f^{i+j}(p))$ for every i,j and ℓ is verified.

Notice that eigenvalues can be defined regardless of the choice of a norm in T_pM since this is a well defined notion for vector spaces.

2.4. **Lyapunov exponents.** Invariant ergodic measures can be thought of as a generalization of periodic orbits.

Theorem 2.3 (Oseledets). *Let $f : M \to M$ be a C^1-diffeomorphism and μ an ergodic measure. Then, there exists $k \in \mathbb{Z}^+$, real numbers $\chi_1 < \chi_2 < \ldots < \chi_k$ and for x in a f-invariant full measure set $R^\mu(f)$ a splitting $T_xM = E_1(x) \oplus \ldots \oplus E_k(x)$ with the following properties:*

- *(**Measurability**) The functions $x \mapsto E_i(x)$ are measurable.*
- *(**Invariance**) $Df_x E_i(x) = E_i(f(x))$ for every $x \in R^\mu(f)$.*

[2]In the case where there are complex eigenvalues, we consider them in pairs $\lambda, \overline{\lambda}$ and the subspace corresponds to the real part of the sum of the spaces when considered as a complex linear transformation.

- *(***Lyapunov exponents***) For every $x \in R^\mu(f)$ and $v \in E_i(x) \setminus \{0\}$ one has*
$$\lim_{n\to\pm\infty} \frac{1}{n} \log \|Df_x^n v\| = \chi_i$$
- *(***Subexponential angles***) For every $x \in R^\mu(f)$ and vectors $v_i \in E_i(x)$ and $v_j \in E_j(x)$ one has that:*
$$\lim_{n\to\pm\infty} \frac{1}{n} \log \sin \measuredangle \left(\frac{Df_x^n v_i}{\|Df_x^n v_i\|}, \frac{Df_x^n v_j}{\|Df_x^n v_j\|} \right) = 0$$

Some explainations are in order:

2.4.1. *Lyapunov exponents.* The numbers χ_i appearing in the statement of Theorem 2.3 are usually called *Lyapunov exponents* of μ.

In general, for any diffeomorphism f a point $x \in M$ is called *regular* (or *Lyapunov regular*) if there exists a splitting $T_xM = E_1(x) \oplus \ldots \oplus E_{k(x)}(x)$ and numbers $\chi_1(x) < \chi_2(x) < \ldots < \chi_{k(x)}(x)$ such that for any vector $v \in E_i \setminus \{0\}$ one has that
$$\lim_{n\to\pm\infty} \frac{1}{n} \log \|Df_x^n v\| = \chi_i(x).$$

Exercise. Show that if $x \in M$ is a regular point and $v \in \bigoplus_{j=1}^{i} E_j(x) \setminus \bigoplus_{j=1}^{i-1} E_j(x)$ then
$$\lim_{n\to+\infty} \frac{1}{n} \log \|Df_x^n v\| = \chi_i(x).$$

In particular, every regular point verifies that every vector has a well defined Lyapunov exponent for the future (and the past). The bundles E_i are the ones on which both coincide.

The set of regular points $R(f)$ is f-invariant and Oseledets theorem implies that it has measure 1 for every f-invariant probability measure (one sometimes calls these sets *full measure sets*). It also holds that all the involved functions are measurable with respect to any invariant measure.

Notice that every periodic point has positive measure for an invariant measure (namely the one that gives equal weight to each point in the orbit) and therefore must be regular. Of course, one does not need Oseledets theorem to prove this, this follows exactly from the considerations in the previous section. Notice that if $f^n(p) = p$, then the Lyapunov exponents of p are the logarithms of the modulus of the eigenvalues of Df_p^n divided by n.

The *Pesin set* of f is the set of regular points for which all Lyapunov exponents are different from 0, that is, the set of points $x \in R(f)$ such that $\chi_i(x) \neq 0$ for all $1 \leq i \leq k(x)$. We shall see later why these points are relevant. A measure μ is called *(non-uniformly) hyperbolic* if all its Lyapunov exponents are non-zero: One should be careful with this name, the *non* applies to the uniformity and not to the hyperbolicity and it should be understood as "*not necessarily uniformly hyperbolic but still with a non-uniform form of hyperbolicity*".

For an ergodic (non-uniformly) hyperbolic measure μ for which one has Lyapunov exponents $\chi_1 < \ldots < \chi_i < 0 < \chi_{i+1} < \ldots < \chi_k$ one can group the bundles depending on the sign of the Lyapunov exponent. In this case, we denote $E^s(x) = E_1(x) \oplus \ldots \oplus E_i(x)$ and $E^u(x) = E_{i+1}(x) \oplus \ldots \oplus E_k(x)$. One has that if $v^s \in E^s(x) \setminus \{0\}$ and $v^u \in E^u(x)$ then:

$$\lim_{n\to\infty}\frac{1}{n}\log\|Df_x^n v^s\| < 0 < \lim_{n\to\infty}\frac{1}{n}\log\|Df_x^n v^u\|$$

So that vectors in $E^s(x)$ are the ones which are exponentially contracted in the future by Df and vectors in E^u are exponentially contracted in the past by Df.

2.4.2. *Angles and measurability.* We remark that, differently from the case of periodic orbits, the concept of norm and angle are essential in this setting as they provide a way to compare vectors which do not belong to the same vector space. However:

Exercise. The values of the Lyapunov exponents are independent of the choice of the Riemannian metric in TM.

The Riemannian metric also provides a way to compute angles between vectors and this is the sense one has to give to the last part of the statement of Theorem 2.3. It is possible to show that this last part is a consequence of the rest, but it is so important that it merits to appear explicitly in the statement.

Another relevant comment is about the notion of measurability of the functions $x \mapsto E_i(x)$. This should be understood in the following way: the arrow defines a function from M to the space of subspaces of TM. This can be thought of as a fiber bundle over M in the following way, for a given $j \leq d = \dim M$ one considers $G_j(M)$ to be the fiber bundle over M such that the fiber in each point is the Grasmannian space of T_xM of subspaces of dimension j. This is well known to have a manifold structure and provide a fiber bundle structure over M([3]). This gives a sense to measurable maps from M to some of these Grasmannian bundles, and since one does not a priori require that the bundles have constant dimension one can think of the function E_i to be a function from M to the union of all these bundles and then the measurability of the function makes sense as both the domain and the target of the function are topological spaces.

2.4.3. *Non-ergodic measures.* There is a statement for non-ergodic measures which is very much like the one we stated but for which the constants k and χ_i become functions of the points and some other parts become more tedious. Look [KH, Supplement] or [M_4, Chapter IV.10] for more information.

2.5. **Sketch of the proof of Oseledets theorem in dimension 2.** This section should be skipped in a first reading. For more details, see [AvB].

Consider $f : M \to M$ a C^1-diffeomorphism of a closed surface M. Let μ be an ergodic invariant measure.

Consider the sequence of functions $\varphi_n : M \to \mathbb{R}$ defined as $\varphi_n(x) = \log\|Df_x^n\|$. The chain rule together with the fact that the norm of a product of matrices is less than or equal to the product of their norms implies that the sequence $\varphi_n(x)$ is subaditive and thus Theorem 2.1 applies. Therefore, there exists $\chi_2 = \lim_n \frac{1}{n}\log\|Df_x^n\|$ for μ-almost every $x \in M$.

The same argument applied to f^{-1} implies the existence of

$$\chi_1 = -\lim_n \frac{1}{n}\log\|Df_x^{-n}\|$$

for μ-almost every $x \in M$. Since $\|Df^{-1}_{f(x)}\|^{-1} \leq \|Df_x\|$, one has that $\chi_2 \geq \chi_1$.

[3]For example, if $j = 1$ this is the projective bundle over M.

Exercise. Show that if $\chi = \chi_1 = \chi_2$ then for μ-almost every $x \in M$ and every $v \in T_xM \setminus \{0\}$ one has that

$$\lim_{n\to\pm\infty} \frac{1}{n} \log \|Df_x^n v\| = \chi.$$

We shall then concentrate on the case $\chi_2 > \chi_1$. The first remark is the following:

Exercise. Show that if $A : \mathbb{R}^2 \to \mathbb{R}^2$ is an invertible linear transformation verifying $\|A\| \neq \|A^{-1}\|^{-1}$, where $\|\cdot\|$ is associated to a given Euclidean metric. Then there exists orthogonal unit vectors $s \perp u$ such that $As \perp Au$ and

$$\|Au\| = \|A\| \quad ; \quad \|As\| = \|A^{-1}\|^{-1}.$$

The key to the proof is then to consider, for $x \in M$ such that the limits $\lim_n \frac{1}{n} \log \|Df_x^n\|$ and $\lim_n \frac{1}{n} \log \|Df_x^{-n}\|$ exist[4], the sequence of unit vectors s_n, u_n in T_xM defined such that $s_n \perp u_n$, $Df_x^n s_n \perp Df_x^n u_n$ and such that

$$\|Df_x^n u_n\| = \|Df_x^n\| \quad ; \quad \|Df_x^n s_n\| = \|(Df_x^n)^{-1}\|^{-1}.$$

One shows that the angle between s_n and s_{n+1} converges exponentially to 0 by using the fact that the limits above exist and the fact that $\|Df\|$ is uniformly bounded. Therefore there exists a limit $s = \lim s_n$ which verifies that

$$\lim_n \frac{1}{n} \log \|Df_x^n s\| = \chi_1.$$

It also follows that, for every unit vector v different from s one has that

$$\lim_{n\to\infty} \frac{1}{n} \log \|Df_x^n v\| = \chi_2.$$

The same argument for the past[5] gives the existence of $u \in T_xM$ such that

$$\lim_{n\to-\infty} \frac{1}{n} \log \|Df_x^n u\| = \chi_2.$$

One must then show that $s \neq u$. Then one can easily show that the angle between $s(f^n(x))$ and $u(f^n(x))$ decreases at subexponential rate with n because one has that for $v \neq w \in T_xM \setminus \{0\}$:

$$\|Df_x^{-1}\|^{-2} \leq \frac{\sin\measuredangle(Df_x v, Df_x w)}{\sin\measuredangle(v,w)} \leq \|Df_x\|^2$$

and therefore the function $x \mapsto \log\sin\measuredangle(s(f(x)), u(f(x))) - \log\sin\measuredangle(s(x), u(x))$ is bounded (and thus integrable). The details can be found in [AvB] in the more general case of linear $SL(2,\mathbb{R})$ cocycles.

2.6. **Pesin's reduction.** Oseledets Theorem 2.3 can be thought of as giving the "eigenvalues" of the derivative over an ergodic measure. We shall now present a result, due to Pesin, which can be then compared to "diagonalizing" the derivative over the measure (or putting it in Jordan form). Again, we treat a special case in dimension 2 for simplicity. See [KH, Supplement S] for more general versions.

[4]Notice that this is an f-invariant set.

[5]Notice that the limit of u_n exists and is orthogonal to s. However, this is not the vector we are interested in, since it might grow also for the past. We have to make a symmetric argument for f^{-1} to find the correct subspace.

Theorem 2.4 (Pesin's ν-reduction). *Let $f : M \to M$ be a C^1-surface diffeomorphism and let μ be a ergodic measure with Lyapunov exponents $\chi_1 < \chi_2$. Then, for every $\nu > 0$ there exists a measurable function C_ν such that*[6] $C_\nu(x) \in GL(\mathbb{R}^2, T_xM)$ *and:*

- *(***Diagonalization***) There exists measurable functions functions $a_\nu : M \to (\exp(\chi_1 - \nu), \exp(\chi_1 + \nu))$ and $b_\nu : M \to (\exp(\chi_2 - \nu), \exp(\chi_2 + \nu))$ such that for μ-almost every point $x \in M$ one has that:*
$$C_\nu(f(x))^{-1} \cdot Df_x \cdot C_\nu(x) = \begin{pmatrix} a_\nu(x) & 0 \\ 0 & b_\nu(x) \end{pmatrix}$$
- *(***Subexponential decay of coordinate size:***) One has that for μ-almost every $x \in M$*
$$\lim_{n\to\pm\infty} \log(\|C_\nu(f^n(x))\| + \|(C_\nu(f^n(x)))^{-1}\|) = 0$$

The key part of the Theorem, which follows from the subexponential decay of the angles given by Oseledets theorem, is the fact that the norm of the matrices $C_\nu(f^n(x))$ and $(C_\nu(f^n(x)))^{-1}$ cannot grow to much along the orbit of generic points.

Sketch Let E_1 and E_2 be the measurable bundles given by Oseledets theorem associated to the exponents χ_1 and χ_2.

For a μ-generic point $x \in M$ one defines the vectors v_i as vectors in $E_i(x)$ of norm:

$$\left(\sum_{n\in\mathbb{Z}} \|Df_x^n|_{E_i(x)}\|^2 e^{-2n\chi_i} e^{-\nu|n|} \right)^{\frac{1}{2}}$$

The series converges for μ-almost every point thanks to the existence of Lyapunov exponents (and the extra term $e^{-\nu|n|}$). If one considers the linear transformation that sends the canonical base of $\mathbb{R}^2$ to v_1, v_2 one sees that the diagonalization hypothesis is easily verified.

Since $\|v_i\|$ is bounded from below, one has that the norm of $C(x)$ is uniformly bounded. On the other hand, the subexponential decay of of the angles given by Oseledets theorem as well as the fact that the Lyapunov exponents are the desired ones implies that the norm of $C(f^n(x))^{-1}$ is subexponential. See [KH, Theorem S.2.10] for more details.

□

3. Passing the information to the manifold

We shall restrict to dimension 2 for simplicity. So, in this section M will be a closed surface and $f : M \to M$ a diffeomorphism of M.

One can look at [KH, Section 6 and Supplement S] for more general statements. We remark that the proofs are quite similar in the higher dimensional context albeit more tedious in notation. The reader will notice that the calculations are already quite tedious in dimension 2.

[6]As above, one can define the function as a function from M to the bundle of linear maps from $\mathbb{R}^2$ to T_xM to make sense to the measurability. Alternatively, one can trivialize the tangent bundle of M up to a zero measure subset and then C_ν becomes a function from M to the space of 2×2 matrices.

The main point of this section is to show how one can recover the behavior seen at the level of the derivative in the dynamics in the manifold itself. The most detailed part will be the easiest one: the case of fixed points. Then, we shall try to explain how the other cases are simply complicated versions of the first one.

3.1. **Fixed points.** We shall work with $p \in M$ such that $f(p) = p$. Since we are in dimension two, we have the following possibilities:

- Both eigenvalues have modulus < 1 or both have modulus > 1.
- One eigenvalue has modulus < 1 and the other has modulus ≥ 1 or one eigenvalue has modulus > 1 and the other ≤ 1.
- Both eigenvalues have modulus 1.

The first case is the easiest to treat:

Exercise. Show that if both eigenvalues have modulus < 1 then p is a *sink*, i.e. there is a neighborhood U of p such that $f(\overline{U}) \subset U$ and for every $x \in U$ one has that $f^n(x) \to p$ exponentially fast. Symmetrically, if both eigenvalues have modulus > 1 the point p is a *source* (i.e. a sink for f^{-1}).

When both eigenvalues have modulus 1 less can be said. However, in dimension 2 there exist some results of topological flavor when the fixed point is isolated (see for example [LeR]).

When one non-zero Lyapunov exists, it is possible to reduce the dimension of the study via the following classical result:

Theorem 3.1 (Stable Manifold Theorem I)**.** *Let p be a fixed point of a diffeomorphism $f : M \to M$ such that Df_p has one eigenvalue of modulus < 1 and the other has modulus ≥ 1. Then, there exists an embedded C^1 curve $\mathcal{W}^s_{loc}(p)$ with the following properties:*

- *(***Invariance***) One has $f(\mathcal{W}^s_{loc}(p)) \subset \mathcal{W}^s_{loc}(p)$*
- *(***Convergence***) For every $x \in \mathcal{W}^s_{loc}(p)$ one has that $d(f^n(x), p) \to 0$.*
- *(***Tangency***) The curve $\mathcal{W}^s_{loc}(p)$ is tangent to the subspace of T_pM corresponding to the eigenvalue of modulus < 1.*
- *(***Uniqueness***) If a point $x \in M$ satisfies that $d(f^n(x), p) \to 0$ exponentially fast, then there exists n_0 such that $f^{n_0}(x) \in \mathcal{W}^s_{loc}(p)$.*

The curve $\mathcal{W}^s_{loc}(p)$ is called the *local stable manifold* at p. One can consider the following:

$$\mathcal{W}^s(p) = \bigcup_{n>0} f^{-n}(\mathcal{W}^s_{loc}(p))$$

which we call the *stable manifold* of p.

Exercise. Show that $\mathcal{W}^s(p)$ is an injectively immersed curve diffeomorphic to $\mathbb{R}$. Give an example on which the manifold $\mathcal{W}^s(p)$ has finite length and an example where it has infinite length.

We shall give a quite detailed proof of Theorem 3.1 since many of the ideas will re-appear plenty of times later.

PROOF. Consider a small neighborhood U of p and a chart $\varphi : U \to \mathbb{R}^2$ such that $\varphi(p) = 0$. By composing with a linear transformation, one can assume that $D\varphi$ sends the eigenspaces of Df_p to the axes of $\mathbb{R}^2$. Assume that the eigenvalue of modulus < 1 is sent to the horizontal axis.

Since there exists another neighborhood V of p such that $V \subset U$ and $f(V) \subset U$ we get that in $\varphi(V)$ one can define: $\hat{f} = \varphi \circ f \circ \varphi^{-1} : \varphi(V) \to \mathbb{R}^2$.

We can therefore write $\hat{f}$ in $\varphi(V)$ as:

$$\hat{f}(x,y) = (\lambda_1 x + \alpha(x,y), \lambda_2 y + \beta(x,y))$$

where $\lambda_1 < 1 \leq \lambda_2$ are the eigenvalues of Df_p and one has that $\alpha(0,0) = \beta(0,0) = \nabla\alpha(0,0) = \nabla\beta(0,0) = 0$. The functions α and β are C^1 on $\varphi(V)$ and therefore, given $\varepsilon > 0$ there exists $\delta > 0$ such that the C^1-size of α and β is smaller than ε in $B(0,\delta)$. Here the C^1 size is the maximum value between the images of the function and the norm of its partial derivatives. Notice that $D\hat{f}_0 = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$.

Consider a smooth bump function $\eta : \mathbb{R}^2 \to [0,1]$ with the following properties:

- $\eta(x,y) = 1$ if $\|(x,y)\| \leq \frac{\delta}{2}$.
- $\eta(x,y) = 0$ if $\|(x,y)\| \geq \delta$
- $\|\nabla\eta(x,y)\| \leq \frac{4}{\delta}$ for every (x,y).

We consider then the function $\bar{f} : \mathbb{R}^2 \to \mathbb{R}^2$ defined as $\bar{f} = \eta\hat{f} + (1-\eta)D\hat{f}_0$, i.e.:

$$\bar{f}(x,y) = \eta(x,y)\hat{f}(x,y) + (1-\eta(x,y))(\lambda_1 x, \lambda_2 y)$$

One can thus write:

$$\bar{f}(x,y) = (\lambda_1 x + \bar{\alpha}(x,y), \lambda_2 y + \bar{\beta}(x,y))$$

with $|\bar{\alpha}(x,y) - \bar{\alpha}(z,w)| \leq \bar{\varepsilon}\min\{\delta, \|(x,y)-(z,w)\|\}$ and $|\bar{\beta}(x,y) - \bar{\beta}(z,w)| \leq \bar{\varepsilon}\min\{\delta, \|(x,y)-(z,w)\|\}$. The value of $\bar{\varepsilon}$ can be chosen to be as small as desired by choosing δ, and ε correctly[7]. The advantage is that now we have a globally defined diffeomorphism of $\mathbb{R}^2$. Notice however that we can only say that the orbits by $\bar{f}$ represent orbits of $\hat{f}$ (or of f) while the point remains in $B(0,\frac{\delta}{2})$.

One can write $\bar{f}^{-1} : \mathbb{R}^2 \to \mathbb{R}^2$ as:

$$\bar{f}^{-1}(x,y) = (\lambda_1^{-1}x + \theta(x,y), \lambda_2^{-1}y + \vartheta(x,y))$$

again (maybe after re-choosing δ and ε) with the C^1-size of both θ and ϑ bounded by $\bar{\varepsilon}$.

Now, let us consider first the existence of a (unique) Lipschitz invariant curve for $\bar{f}$ tangent to the x-axis which is contracting.

Consider then the following complete metric space:

$$\mathrm{Lip}_1 = \{\varphi : \mathbb{R} \to \mathbb{R} \ : \ |\varphi(t) - \varphi(s)| \leq |t-s| \ , \ \forall t,s \ ; \ \varphi(0) = 0\}$$

endowed with the metric $d(\varphi,\varphi') = \sup_{t\neq 0 \in \mathbb{R}} \frac{|\varphi(t)-\varphi'(t)|}{t}$.

For a given $\varphi \in \mathrm{Lip}_1$ one can define a new function $\bar{f}_*\varphi$ as the function whose graph is the preimage by $\bar{f}$ of the graph of φ, i.e. $\operatorname{graph}\bar{f}_*\varphi = \bar{f}^{-1}(\operatorname{graph}\varphi)$.

Let us precise the construction of $\bar{f}_*\varphi$ a little further. Let $G_\varphi : \mathbb{R} \to \mathbb{R}$ the function defined by $G_\varphi(t) = \lambda_1 t + \bar{\alpha}(t,\varphi(t))$. One has:

Claim. *If $\bar{\varepsilon}$ is small enough, the function G_φ is an increasing homeomorphism of $\mathbb{R}$ which verifies* $(\lambda_1 - \sqrt{2}\bar{\varepsilon})|t-s| \leq |G_\varphi(t) - G_\varphi(s)| \leq (\lambda_1 + \sqrt{2}\bar{\varepsilon})|t-s|$.

[7]This is the well known fact that the C^1-topology is invariant under rescaling. Given $\bar{\varepsilon}$ there exists δ such that $\|\alpha(x,y)\|_{C^1} + \|\beta(x,y)\|_{C^1} \leq \frac{\bar{\varepsilon}\delta}{8}$ whenever $\|(x,y)\| \leq \delta$. Now, one has that the C^1-distance of $\bar{f}$ and $D\bar{f}_0$ is the C^1 size of $\eta(\bar{f} - D\bar{f}_0)$ which smaller than $\bar{\varepsilon}$ as desired.

Proof. Assume that $\sqrt{2}\bar{\varepsilon} < (1-\lambda_1)$. One computes:

$$|\lambda_1 t + \bar{\alpha}(t,\varphi(t)) - \lambda_1 s - \bar{\alpha}(s,\varphi(s))| \geq \lambda_1|t-s| - \sqrt{2}\bar{\varepsilon}|t-s| \geq (\lambda_1 - \sqrt{2}\bar{\varepsilon})|t-s|$$

this follows from the fact that $|\bar{\alpha}(t,\varphi(t)) - \bar{\alpha}(s,\varphi(s))| \leq \bar{\varepsilon}\sqrt{|t-s|^2 + |\varphi(s)-\varphi(t)|^2}$ and that φ is 1-Lipschitz.

On the other hand, it is easy to see that $|\lambda_1 t + \bar{\alpha}(t,\varphi(t)) - \lambda_1 s - \bar{\alpha}(s,\varphi(s))| \leq (\lambda_1 + \sqrt{2}\bar{\varepsilon})|t-s|$.

$\diamondsuit$

Then, the function $\bar{f}_*\varphi$ verifies $(t, \bar{f}_*\varphi(t)) = \bar{f}^{-1}(G_\varphi(t), \varphi(G_\varphi(t)))$ (see figure 1) and therefore:

$$\bar{f}_*\varphi(t) = \lambda_2^{-1}\varphi(G_\varphi(t)) + \vartheta(G_\varphi(t), \varphi(G_\varphi(t)))$$

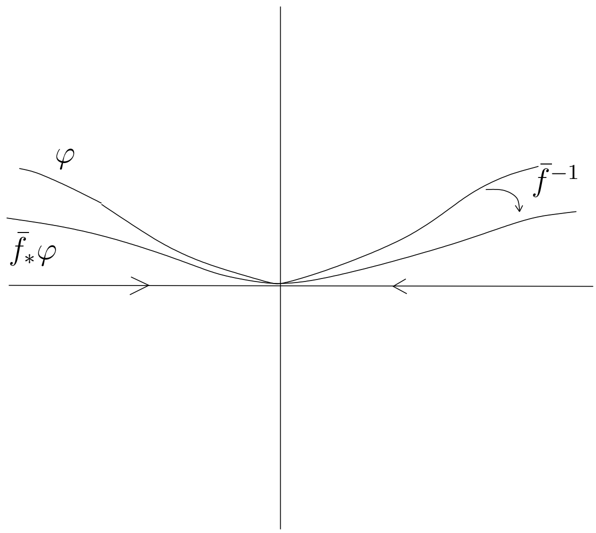


Figure 1. The graph transform of φ.

We have the following properties:

Claim. *If $\bar{\varepsilon}$ is small enough, for $\varphi \in \mathrm{Lip}_1$, the function $\bar{f}_*\varphi \in \mathrm{Lip}_1$.*

Proof. Intuitively, this follows directly from the fact that $D\bar{f}^{-1}$ contracts horizontal cones. Let us do the calculations (which should be skipped in a first reading). First notice that $\bar{f}_*\varphi(0) = 0$ from its definition.

Recall that for $t, s \in \mathbb{R}$ one has $|G_\varphi(t) - G_\varphi(s)| \leq (\lambda_1 + \sqrt{2}\bar{\varepsilon})^{-1}|t-s|$

Given $t, s \in \mathbb{R}$ one has that:

$$\begin{aligned}
|\bar f_*\varphi(t)-\bar f_*\varphi(s)| &= |\lambda_2^{-1}\varphi(G_\varphi(t))+\vartheta(G_\varphi(t),\varphi(G_\varphi(t)))-\\
&\qquad -\lambda_2^{-1}\varphi(G_\varphi(s))+\vartheta(G_\varphi(s),\varphi(G_\varphi(s)))|\\
&\leq \lambda_2^{-1}|\varphi(G_\varphi(t))-\varphi(G_\varphi(s))|+|\vartheta(t,G_\varphi(t))-\vartheta(s,\varphi(G_\varphi(s))|\\
&\leq \lambda_2^{-1}|G_\varphi(t)-G_\varphi(s)|+\bar\varepsilon\,\|(t,\varphi(G_\varphi(t)))-(s,\varphi(G_\varphi(s)))\|\\
&\leq \left(\lambda_2^{-1}(\lambda_1-\sqrt{2}\bar\varepsilon)+\sqrt{2}\bar\varepsilon\right)|t-s|
\end{aligned}$$

and if $\bar\varepsilon$ is small enough, one gets that $\lambda_2^{-1}(\lambda_1-\sqrt{2}\bar\varepsilon)+\sqrt{2}\bar\varepsilon<1$ as desired[8].

♢

Claim. *For sufficiently small $\bar\varepsilon$, there exists $\gamma\in(0,1)$ such that if $\varphi,\varphi'\in \mathrm{Lip}_1$ then $d(\bar f_*\varphi,\bar f_*\varphi')\leq\gamma d(\varphi,\varphi')$.*

Proof. Again, this is a consequence of the contraction of horizontal cones by $D\bar f^{-1}$. Let us perform the computations (the reader should skip them in a first reading).

$$\begin{aligned}
|\bar f_*\varphi(t)-\bar f_*\varphi'(t)| &= |\lambda_2^{-1}\varphi(G_\varphi(t))+\vartheta(G_\varphi(t),\varphi(G_\varphi(t)))-\\
&\qquad -\lambda_2^{-1}\varphi'(G_{\varphi'}(t))+\vartheta(G_{\varphi'}(t),\varphi'(G_{\varphi'}(t)))|\\
&\leq \lambda_2^{-1}|\varphi(G_\varphi(t))-\varphi'(G_{\varphi'}(t))|+\\
&\qquad +|\vartheta(G_\varphi(t),\varphi(G_\varphi(t)))-\vartheta(G_{\varphi'}(t),\varphi'(G_{\varphi'}(t)))|\,.
\end{aligned}$$

Now, one has that

$$\begin{aligned}
|\varphi(G_\varphi(t))-\varphi'(G_{\varphi'}(t))| &\leq |\varphi(G_\varphi(t))-\varphi'(G_\varphi(t))|+|\varphi'(G_\varphi(t))-\varphi'(G_{\varphi'}(t))|\\
&\leq d(\varphi,\varphi')|G_\varphi(t)|+|G_\varphi(t)-G_{\varphi'}(t)|\\
&\leq (\lambda_1+\sqrt{2}\bar\varepsilon)d(\varphi,\varphi')|t|+|\bar\alpha(t,\varphi(t))-\bar\alpha(t,\varphi'(t))|\\
&\leq (\lambda_1+\sqrt{2}\bar\varepsilon)d(\varphi,\varphi')|t|+\sqrt{2}\bar\varepsilon d(\varphi,\varphi')|t|\\
&\qquad =(\lambda_1+2\sqrt{2}\bar\varepsilon)d(\varphi,\varphi')|t|
\end{aligned}$$

Moreover, one has that

$$\begin{aligned}
|\,\vartheta(G_\varphi(t),\varphi(G_\varphi(t))) &- \vartheta(G_{\varphi'}(t),\varphi'(G_{\varphi'}(t)))\,|\leq\\
&\leq |\vartheta(G_\varphi(t),\varphi(G_\varphi(t)))-\vartheta(G_\varphi(t),\varphi'(G_\varphi(t)))|+\\
&\qquad +|\vartheta(G_\varphi(t),\varphi'(G_\varphi(t)))-\vartheta(G_{\varphi'}(t),\varphi'(G_{\varphi'}(t)))|\\
&\leq \bar\varepsilon d(\varphi,\varphi')|t|+\sqrt{2}\bar\varepsilon|G_\varphi(t)-G_{\varphi'}(t)|\\
&\leq (2\sqrt{2}+1)\bar\varepsilon d(\varphi,\varphi')|t|
\end{aligned}$$

Putting all the estimates together it follows that if $\lambda_2^{-1}(\lambda_1+2\sqrt{2}\bar\varepsilon)+(2\sqrt{2}+1)\bar\varepsilon=\gamma<1$ one has the desired statement.

♢

We deduce that there exists a unique function $\bar\varphi$ in Lip_1 whose graph is $\bar f$-invariant. We call $\mathcal{W}^s_{loc}$ to the restriction of the graph to $B(0,\frac{\delta}{2})$. The rest of the

[8]Notice that indeed, one does not need that $\lambda_2\geq 1$ but rather, it is enough that $\lambda_1<\lambda_2$ as long as one chooses $\bar\varepsilon$ correctly.

proof is devoted to showing that this graph (which is identified with a curve in M) verifies the conclusions of the theorem.

Invariance and convergence: This follows quite easily from the fact that if $|y| < t$ the map $t \mapsto \lambda_1 t + \bar{\alpha}(t, y)$ is contracting if $\bar{\varepsilon} < (1 - \lambda_1)$, therefore, since $\bar{\varphi} \in \mathrm{Lip}_1$ one gets contraction for the map $t \mapsto \lambda_1 t + \bar{\alpha}(t, \bar{\varphi}(t))$ is contracting. This also implies that for every $(t, \bar{\varphi}(t))$ one has that $\bar{f}^n(t, \bar{\varphi}(t)) \to 0$ exponentially fast and that if $(t, \bar{\varphi}(t)) \in B(0, \frac{\delta}{2})$ (i.e. $(t, \bar{\varphi}(t)) \in \mathcal{W}^s_{loc}$) then[9] the same holds for $\bar{f}^n(t, \bar{\varphi}(t))$.

Smoothness: We must show that the curve $\mathcal{W}^s_{loc}$ is C^1 and tangent to the x-axis in $(0,0)$. To do so, notice that at each $t_0 \in \mathbb{R}$ one has that the set of accumulation points of

$$\frac{\bar{\varphi}(t) - \bar{\varphi}(t_0)}{t - t_0} \qquad \text{as} \quad t \to t_0$$

is an interval contained in $[-1, 1]$ because $\bar{\varphi} \in \mathrm{Lip}_1$. This is equivalent to say that at each point $(t_0, \bar{\varphi}(t_0))$ the graph of $\bar{\varphi}$ is tangent to a cone of bounded width and transverse to the y-axis. The form of $\bar{f}$ implies that the angle of such a cone is contracted by an uniform amount by $D\bar{f}$. Using the fact that the graph of $\bar{\varphi}$ is $\bar{f}$-invariant, one deduces that the cones must degenerate at each point, or equivalently, the function $\bar{\varphi}$ is everywhere differentiable. A similar argument shows that these tangent spaces have to vary continuously with the point since otherwise one would obtain another invariant cone (by comparing the limits of different subsequences) by $D\bar{f}$ of positive width.

In $(0,0)$ it is clear that the unique direction transverse to the y-axis which is $D\bar{f}$-invariant is the x-axis and therefore the derivative of $\bar{\varphi}$ at 0 is 0 or equivalently, the curve $\mathcal{W}^s_{loc}$ is tangent to the x-axis at $(0,0)$ as desired.

Uniqueness: Assume that there is a point which converges exponentially fast to p for f. Then, one can construct a point $(t_0, s_0) \in B(0, \frac{\delta}{2})$ which converges exponentially fast to $(0,0)$ for $\bar{f}$. Since $\lambda_2 \geq 1$ one has that if $(t_n, s_n) = \bar{f}^n(t_0, s_0)$ then $\frac{s_n}{t_n}$ converges to zero since otherwise, the rate of convergence of (t_n, s_n) to zero is governed by λ_2 at first order[10]. Then, it is possible to construct a subfamily of Lip_1 consisting of functions such that $\varphi(t_n) = s_n$ for every n and one gets that it is a closed $\bar{f}_*$-invariant subset of Lip_1 and therefore, it contains the (unique) fixed point of the contraction. This proves the uniqueness.

□

Remark 3.2. Notice that the only place where we used that $\lambda_1 < 1$ is to show uniqueness (on the other hand, we used $\lambda_1 < \lambda_2$ everywhere). Otherwise, we would get a *locally invariant* curve which depends on the way we choose the extension (which is not canonical) and uniqueness only holds for points whose forward orbit remains in $B(0, \frac{\delta}{2})$. This is the content of the well known *center manifold* theorems ([Sh, HPS]).

Exercise (Chapter 5.III of [Sh])**.** Consider the time one map of the differential equation $\dot{x} = -x$ and $\dot{y} = y^2$ in $\mathbb{R}^2$. Show that $\mathcal{W}^s_{loc}(0,0)$ is the horizontal axis but

[9]The reader might be worried with the fact that the ball is round and therefore the contraction of the x-coordinate might not imply that the point remains in the ball. However, we notice that the intersection of a 1-Lipschitz graph through $(0,0)$ with a ball must be connected.

[10]Indeed, it is enough to show that $\frac{s_n}{t_n} \leq 1$. Notice that if $\lambda_2 > 1$ then the argument is simpler since points (t_0, s_0) such that $s_0 \geq t_0$ verify that its iterates by $\bar{f}$ leave $B(0, \frac{\delta}{2})$.

that uniqueness of the manifold tangent to the other direction is not ensured in the place where the dynamics tangent to the y-axis is contracting.

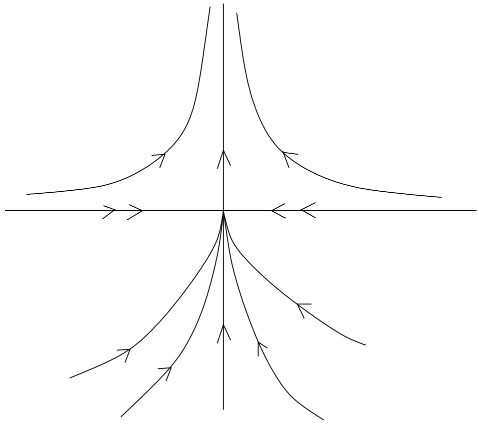


FIGURE 2. The flow of the equation $\dot{x} = -x$ and $\dot{y} = y^2$ in $\mathbb{R}^2$.

3.2. **The case of all Lyapunov exponents negative.** It is easy to pass from the information we gathered for fixed points to periodic points. The motivation from now on is to try to understand what kind of behavior is forced for general ergodic measures. It is natural to expect that zero-Lyapunov exponents will not provide much information, but when the measure is hyperbolic, one expects to obtain some information on the local dynamics for generic points of the measure.

This is an easier version of what follows, we shall see the first relatively easy consequence of a measure having non-zero Lyapunov exponents.

Theorem 3.3. *Let μ be an ergodic measure of a C^1-diffeomorphism f of a surface M such that both Lyapunov exponents are negative. Then, μ is supported in a periodic sink.*

Of course a symmetric statement holds for measures having all Lyapunov exponents positive where one obtains a periodic source applying the previous result to f^{-1}.

PROOF. One has that there exists $\chi < 0$ such that for μ-almost every $x \in M$ and every $v \in T_xM \setminus \{0\}$ one has that:

$$\limsup_n \frac{1}{n} \log \|Df_x^n v\| < \chi < 0.$$

Claim. *There exists $N_0 > 0$ such that for μ-almost every $x \in M$ and $N \geq N_0$ one has that*

$$\frac{1}{kN} \sum_{i=0}^{k-1} \log \|Df^N(f^{iN}(x))\| \to \hat{\chi}(x) \leq \chi$$

PROOF. We assume that μ is ergodic for f^N for every $N > 0$. Notice that it might be that it has (finitely) many ergodic components, the proof in this more general case is a little bit more tedious (see [AbBC, Lemma 8.4]).

The fact that all Lyapunov exponents are smaller than χ implies that
$$\frac{1}{n}\int \log \|Df^n\| d\mu \to \hat{\chi} \leq \chi$$
as $n\to\infty$. In particular, for sufficiently large N_0 one has that if $N \geq N_0$ then
$$\frac{1}{N}\int \log \|Df^N\| d\mu \leq \chi \, .$$

Now, the result follows from applying Birkhoff's theorem to the dynamics f^N and the function $x \mapsto \log \|Df^N(x)\|$.

$\diamondsuit$

Let us fix $\varepsilon \leq \frac{-\chi}{10}$, a value of $N \geq N_0$ as given by the previous claim and let $\Delta_f \geq \max_x \|Df(x)\|$.

There exists $\delta_0 > 0$ such that if $d(x,y) \leq \delta_0$ then for every vector $v \in T_yM$ one has that

$$\|Df^N(y)v\| \leq e^{N\varepsilon}\|Df^N(x)\|\|v\|$$

Let $R : M \to \mathbb{R}$ be defined as[11]:

$$R(x) = \max_{k\geq 0}\left\{ e^{-kN(\chi+\varepsilon)} \prod_{i=0}^{k-1} \|Df^N(f^{iN}(x))\| \right\} \geq 1$$

Notice that the previous claim implies that the value of $R(x)$ is well defined[12] on generic points with respect to μ since for sufficiently large k the value of $\prod_{i=0}^{k-1}\|Df^N(f^{iN}(x))\| \leq e^{kN(\chi+\varepsilon)}$.

Now consider $\delta_1 < \Delta_f^{-N}\delta_0$ and for μ-almost every $x \in M$ consider $\rho(x) = \frac{\delta_1}{R(x)}$. We have the following (compare with [AbBC, Lemma 8.10]):

Claim. *For μ-almost every $x \in M$ and $n \geq 0$ one has $f^n(B(x,\rho(x))) \subset B(x,\delta_0)$. Moreover, the diameter of $f^n(B(x,\rho(x)))$ converges to zero exponentially fast as $n \to \infty$.*

Proof. Let us first prove that $f^n(B(x,\rho(x)) \subset B(x,\delta_0)$. Assume that this is the case for $k \leq n-1$. Consider $\ell \geq 0$ the largest integer for which $\ell N \leq n$. By induction and noticing that the derivative in $f^i(x)$ is approximately the same in points in $B(f^i(x),\delta_0)$ one shows that:

$$\mathrm{diam}(f^n(B(x,\rho(x)))) \leq \left(\Delta_f^N e^{\ell N\varepsilon} \prod_{i=0}^{\ell-1} \|Df^N(f^{iN}(x)\| \right)\rho(x) \leq \delta_0$$

One deduces using the definition of $R(x)$ that:

$$\mathrm{diam}(f^n(B(x,\rho(x)))) \leq \Delta_f^N e^{\ell N(\chi+2\varepsilon)} R(x)\frac{\delta_1}{R(x)} \leq e^{\ell N(\chi+2\varepsilon)}\Delta_f^N\delta_1 \leq \delta_0$$

But we have also established that
$$\mathrm{diam}(f^n(B(x,\rho(x))) \leq e^{\ell N(\chi+2\varepsilon)}\Delta_f^N\delta_1$$

[11]We make the convention that $\prod_{i=0}^{0} a_i = 1$.

[12]Indeed, standard arguments give that the sequence $R(f^n(x))$ is subexponential. See [AbBC, Lemma 8.7].

for every $n \geq 0$ which implies that the diameter goes to zero exponentially fast.

$\diamondsuit$

Consider a generic point x for μ, which is recurrent, i.e. there exists $n_j \to \infty$ such that $f^{n_j}(x) \in \Lambda_\varepsilon$ and $f^{n_j}(x) \to x$ and verifies the conditions of the previous claim. Such a point exists thanks to Poincare's recurrence theorem.

For large enough j, one has that $d(f^{n_j}(x), x) \ll \rho(x)$ and therefore

$$f^{n_j}(B(x, \rho(x))) \subset B(x, \rho(x))$$

and distances are contracted uniformly. This implies that $f^{n_j}|_{B(x,\rho(x))}$ has a unique (attracting) fixed point p and that $f^{kn_j}(y) \to p$ for every $y \in B(x, \rho(x))$. Since x was recurrent, this implies that $x = p$, which must be a periodic sink and this concludes the proof.

$\square$

Exercise. Prove Poincare's recurrence theorem (the statement used in the proof of Theorem 3.3) using Birkhoff's ergodic theorem.

3.3. **A result on sequences of diffeomorphisms.** We treat in this section a situation similar to the one we reduced in the fixed point case. Instead of dealing with a unique global diffeomorphism of $\mathbb{R}^2$ which is C^1-close to a linear transformation, we shall deal with a sequence of such maps and "notice" that we never really used the exact properties of the global diffeomorphism but instead we used the fact that the bounds were uniform. The reader can try to predict what purpose the result in this subsection will serve: one will consider charts around each point and extend the maps to global diffeomorphisms by using a bump function to glue the map with its derivative.

Let us introduce the context on which we shall work: A sequence $\{f_n\}_{n\in\mathbb{Z}}$ of diffeomorphisms of $\mathbb{R}^2$ is called a $(\lambda_1, \lambda_2, \varepsilon)$-*hyperbolic sequence of diffeomorphism* if it satisfies the following properties:

- $f_n(x, y) = (a_n x + \alpha_n(x, y), b_n y + \beta_n(x, y))$ where $0 < a_n < \lambda_1 < 1 < \lambda_2 < b_n$ and $\alpha_n(0, 0) = \beta_n(0, 0) = \nabla\alpha_n(0, 0) = \nabla\beta_n(0, 0) = 0$.
- The maps $\alpha_n : \mathbb{R}^2 \to \mathbb{R}$ and $\beta_n : \mathbb{R}^2 \to \mathbb{R}$ are C^1 and their C^1-distance to 0 is $< \varepsilon$. That is, for every $(x, y) \in \mathbb{R}^2$ one has that

$$|\alpha_n(x, y)|, \ \ |\beta_n(x, y)|, \ \ \|\nabla\alpha_n(x, y)\| \ \text{ and } \ \|\nabla\beta_n(x, y)\|$$

  are all smaller than ε.

The main result of this subsection is:

Theorem 3.4 (Stable Manifold Theorem for Hyperbolic Sequences)**.** *Given $\lambda_1 < 1 < \lambda_2$, there exists $\varepsilon > 0$ such that if $\{f_n\}_{n\in\mathbb{Z}}$ is a $(\lambda_1, \lambda_2, \varepsilon)$-hyperbolic sequence of diffeomorphisms, then, there exists a family of C^1 functions $\varphi_n : \mathbb{R} \to \mathbb{R}$ such that:*

- *(**Invariance**) The graphs are f_n-invariant, i.e. for every $n \in \mathbb{Z}$ and $t \in \mathbb{R}$ there exists $s \in \mathbb{R}$ such that $f_n(t, \varphi_n(t)) = (s, \varphi_{n+1}(s))$.*
- *(**Convergence**) For every $n \in \mathbb{Z}$ and $t \in \mathbb{R}$ one has that*

$$\lim_{m\to\infty} \|f_{n+m} \circ \ldots \circ f_n(t, \varphi_n(t))\| = 0\,.$$

- *(**Tangency**) The derivative $\varphi_n'(0) = 0$ for every $n \in \mathbb{Z}$.*
- *(**Uniqueness**) The family is the unique family with the first two properties.*

Indeed, the proof of this Theorem follows exactly the same lines as the proof we did in subsection 3.1. When looking at the proof of Theorem 3.1 one can identify two stages:

- First, one fixes a small chart around the fixed point where one can construct a global diffeomorphism of $\mathbb{R}^2$ which is C^1-close to a linear diagonal matrix with an eigenvalue smaller than one in the x-axis and larger than one in the y-axis.
- Then, one proves a result which is equivalent to Theorem 3.4 for a constant sequence $(\lambda_1, \lambda_2, \varepsilon)$-hyperbolic sequence of diffeomorphisms $f_n = \bar{f}$ for all $n \in \mathbb{Z}$ (to keep the notation of the proof of Theorem 3.1).

There is a minor difference on how to implement the proof. Instead of working with the space of Lipschitz functions (which has no longer much sense since the iterative process has to take place in "different" $\mathbb{R}^2$s), one has to work with sequences of Lipschitz functions. That is, one works with the space:

$$\mathrm{Lip}_1^{seq} = \{\{\varphi_n\}_n \ : \ \varphi_n(0) = 0 \ , \ |\varphi_n(t) - \varphi_n(s)| \leq |t-s|\}$$

endowed with the metric $d(\{\varphi_n\}, \{\varphi_n'\}) = \sup_{n \in \mathbb{Z}} d(\varphi_n, \varphi_n')$ which is also a complete metric space. One defines a *graph transform* of the form $\{f_n\}_*\{\varphi_n\} = \{\psi_n\}$ so that ψ_n is the function whose graph is the graph of $(f_n)^{-1}(\varphi_{n+1})$. The rest of the proof follows more or less verbatim as this graph transform preserves the space Lip_1^{seq} and contracts its metric giving a unique fixed point which will satisfy all the desired properties.

Exercise. Try to implement the same proof as in Theorem 3.1 to recover Theorem 3.4.

Theorem 3.4 is known as Hadamard-Perron's theorem. See [KH, Section 6.2] for more information and a complete proof in any dimension.

3.4. **Pesin charts.** The following result is the place where the $C^{1+\alpha}$ hypothesis appears in Pesin's theory. It allows to lift the dynamics to a subexponential neighborhood of a generic point for a hyperbolic measure μ and therefore obtain a hyperbolic sequence of diffeomorphisms. This allows to construct stable and unstable manifolds for those points using Theorems 2.4 and 3.4. By inspection of the proof one can see that the key place where the Hölder continuity of the derivative is used is to control the fact that angles can be very small (i.e. the norm of C_ν or C_ν^{-1} of Theorem 2.4 can be very large).

Theorem 3.5 (Pesin-Lyapunov Charts)**.** *Let $f : M \to M$ be a $C^{1+\alpha}$ diffeomorphism of a closed surface M. Let μ be an ergodic measure with Lyapunov exponents $\chi_1 > \chi_2$. Then, for every $\rho_0 > 0$ and $\nu > 0$ there exists a measurable function $\rho : M \to (0, \rho_0)$ and a family of smooth charts $\xi_z : B(0, \rho(p)) \subset \mathbb{R}^2 \to M$ indexed in a full μ-measure set of $z \in M$ with the following properties:*

- *(***Lift of the dynamics:***) The map $\tilde{f}_z : B(0, \frac{\rho(z)}{\|Df\|}) \to B(0, \rho(f(z)))$ defined as $\tilde{f}_z = \xi_{f(z)}^{-1} \circ f \circ \xi_z$ is well defined and can be extended to a diffeomorphism $\hat{f}_z : \mathbb{R}^2 \to \mathbb{R}^2$ of the form:*

$$\hat{f}_z(x,y) = (a_z x + \alpha(x,y), b_z y + \beta(x,y))$$

where $\log a_z \in [\chi_2 - \nu, \chi_2 + \nu]$ and $\log b_z \in [\chi_1 - \nu, \chi_1 + \nu]$ and $\alpha(0,0) = \beta(0,0) = \nabla\alpha(0,0) = \nabla\beta(0,0) = 0$.

- (**Extension:**) *One can choose the extension* $\hat{f}_z$ *in such a way that the maps* $\alpha : \mathbb{R}^2 \to \mathbb{R}$ *and* $\beta : \mathbb{R}^2 \to \mathbb{R}$ *are* C^1 *and their* C^1*-distance to* 0 *is less than* ν*. That is, for every* $w \in \mathbb{R}^2$ *one has that* $|\alpha(w)|, |\beta(w)|, \|\nabla\alpha(w)\|$ *and* $\|\nabla\beta(w)\|$ *are all smaller than* ν*.*
- (**Subexponential decay of size:**) *The function* $\rho : M \to (0, \rho_0)$ *satisfies that* $\rho(f(z)) \in (e^{-\nu}\rho(z), e^{\nu}\rho(z))$ *and* $\lim_n \frac{1}{n} \log \rho(f^n(z)) = 0$*.*

Proof. Let $\exp : TM \to M$ be the exponential mapping with respect to a given Riemannian metric. We know that $\exp_z : T_zM \to M$ verifies that $\exp_z(0) = z$ and $D(\exp_z)_0 = Id$. Using compactness of M we know that there exists $R_0 > 0$ such that $\exp_z : B(0, R_0) \to M$ is a diffeomorphism verifying that

$$\|(D\exp_z)_w\|, \|((D\exp_z)_w)^{-1}\|^{-1} \leq 2$$

for every $w \in B(0, R_0) \subset T_zM$.

Consider the linear change of coordinates $C_\nu(z) \in GL(\mathbb{R}^2, T_zM)$ given by 2.4 such that for μ-almost every point $z \in M$ one has that:

$$C_\nu(f(z))^{-1} \cdot Df_z \cdot C_\nu(z) = \begin{pmatrix} a_\nu(z) & 0 \\ 0 & b_\nu(z) \end{pmatrix}$$

and $a_\nu : M \to (\exp(\chi_1 - \nu), \exp(\chi_1 + \nu))$ and $b_\nu : M \to (\exp(\chi_2 - \nu), \exp(\chi_2 + \nu))$. One can choose $C_\nu(z)$ so that $\lim_{n\to\pm\infty} \log(\|C_\nu(f^n(z))\| + \|(C_\nu(f^n(z)))^{-1}\|) = 0$ for μ-almost every $z \in M$.

The function ξ_z will be the restriction of $\tilde{\xi}_z := (\exp_z \circ C_\nu(z)) : \mathbb{R}^2 \to M$ to a convenient neighborhood of 0.

First we shall define $\rho_1 : M \to (0, 1]$ to be a function verifying that for $z \in M$ the value $\rho_1(z)$ is the maximal value ≤ 1 such that:

- $C_\nu(z)(B(0, \rho_1(z))) \subset B(0, R_0)$ and
- $C_\nu(f(z))^{-1} \circ \exp_{f(z)}^{-1} \circ f \circ \exp_z \circ C_\nu(z)(B(0, \rho_1(z))) \subset B(0, R_0)$.

Technically, the function ρ_1 is only defined in points where C_ν is defined, but these form a full μ-measure set, so it is no problem for our purposes. In these points, the function is clearly positive and well defined. Moreover, the function $\tilde{\xi}_z$ is a diffeomorphism when restricted to $B(0, \rho_1(z))$ and we can therefore define:

$$\tilde{f}_z : B(0, \rho_1(z)) \to \mathbb{R}^2 \ , \quad \tilde{f}_z = \tilde{\xi}_{f(z)}^{-1} \circ f \circ \tilde{\xi}_z$$

The key difficulty is to obtain that the lift of f is C^1-close to its linear part $D(\tilde{f}_z)_0 = C_\nu(f(z))^{-1} \cdot Df_z \cdot C_\nu(z)$ (i.e. that the functions α and β are C^1-close to 0). It is for this that we shall restrict ρ further and use the $C^{1+\alpha}$ hypothesis.

We write $\tilde{f}_z = D(\tilde{f}_z)_0 + h_z$ where the function $h_z = (\alpha(z), \beta(z))$ and α and β verify $\alpha(0,0) = \beta(0,0) = \nabla\alpha(0,0) = \nabla\beta(0,0) = 0$.

In $B(0, R_0/\|Df_z\|)$ we can write $\exp_{f(z)}^{-1} \circ f \circ \exp_z = Df_z + g_z$ where g_z is $C^{1+\alpha}$ with similar constant as f (notice that exp is C^∞ and R_0 is chosen so that the derivative is controlled). So, there exists $c > 0$ such that for $w \in B(0, R_0/\|Df_z\|)$ one has:

$$\|(Dg_z)_w\| \leq c\|w\|^\alpha$$

Since $h_z = C_\nu(f(z))^{-1} \circ g_z \circ C_\nu(z)$, we have that

$$D(h_z)_w = D(C_\nu(f(z))^{-1} \circ g_z \circ C_\nu(z))_w = C_\nu(f(z))^{-1} \circ D(g_z)_{C_\nu(z)w}$$

so

$$\|D(h_z)_w\| \leq \|C_\nu(f(z))^{-1}\| \|D(g_z)_{C_\nu(z)w}\| \leq c\|C_\nu(f(z))^{-1}\| \|C_\nu(z)\|^\alpha \|w\|^\alpha$$

Notice that from the hypothesis on C_ν we know that

$$k(z) = c\|C_\nu(f(z))^{-1}\| \|C_\nu(z)\|^\alpha$$

has subexponential decay (i.e. $\lim_n \frac{1}{n} \log k(f^n(z)) = 0$) and if we choose $\rho_2 : M \to (0, \rho_0)$ small enough so that the norm of $\|D(h_z)_w\|$ is smaller than ν it is not hard to extend the functions $\tilde{f}_z$ to satisfy the extension property.

It remains to show that one can now choose $\rho : M \to (0, \rho_0)$ such that:

- $\rho(z) \leq \rho_2(z)$ for every $z \in M$,
- $\rho(f(z)) \in (e^{-\nu}\rho(z), e^\nu \rho(z))$
- $\lim_n \frac{1}{n} \log \rho(f^n(z)) = 0$.

The third condition follows immediately from the second. To construct ρ verifying the first two properties, it is enough to consider

$$\rho(z) := \inf_{n \in \mathbb{Z}} e^{\frac{\nu|n|}{4}} \rho_2(f^n(z))$$

which is well defined since $\lim_n \frac{1}{n} \log \rho_2(f^n(z)) = 0$ and verifies the desired properties. This concludes the proof of the Theorem.

□

Remark 3.6. By construction, one sees that there exists a measurable $K : M \to [1, \infty)$ such that if $w, w' \in B(0, \rho(z))$ then

$$d(\xi_z(w), \xi_z(w')) \leq \|w - w'\| \leq K(z) d(\xi_z(w), \xi_z(w'))$$

and such that $\lim_n \frac{1}{n} \log K(f^n(z)) = 0$.

One obtains the following result applying Theorems 3.4 and 3.5 (and Remark 3.6) which provides the so called Pesin's stable and unstable manifolds.

Theorem 3.7 (Pesin stable manifold theorem)**.** *Let $f : M \to M$ be a $C^{1+\alpha}$ diffeomorphism of a closed surface and μ a hyperbolic measure for f with Lyapunov exponents $\chi^s < 0 < \chi^u$. Then, there exists an f-invariant subset $R \subset M$ such that $\mu(R) = 1$ and:*

- *(**Existence:**) for every $x \in R$ there exists a C^1-curve $\mathcal{W}^s_{Pes}(x)$ centered at x and tangent to $E^s(x)$ with length $2\rho(x)$,*
- *(**Invariance:**) one has that $f(\mathcal{W}^s_{Pes}(x)) \subset \mathcal{W}^s_{Pes}(f(x))$,*
- *(**Convergence:**) for $y \in \mathcal{W}^s_{Pes}(x)$ one has that $\frac{1}{n} \log(d(f^n(x), f^n(y))) \to \chi^s$ for $n \to +\infty$,*
- *(**Uniqueness:**) if a point $y \in M$ verifies that $\frac{1}{n} \log(d(f^n(x), f^n(y))) \to \chi^s$ as $n \to +\infty$ then there exists n_y such that $f^{n_y}(y) \in \mathcal{W}^s_{Pes}(f^{n_y}(x))$,*
- *(**Size:**) the function $\rho : M \to (0, \rho_0)$ verifies that*

$$\rho(f(x)) \in (e^{-\nu}\rho(x), e^\nu \rho(x))$$

for $\nu \ll \min\{|\chi^s|, \chi^u\}$ and therefore $\lim \frac{1}{n} \log \rho(f^n(x)) = 0$.

3.5. **The $C^{1+\alpha}$ hypothesis.** In [BoCS] an example is constructed showing the importance of the Hölder continuity of the derivative in order to construct the stable and unstable manifolds for generic points with respect to a hyperbolic measure. In their example, the measure is hyperbolic but every point in the support of the measure verifies that their stable (resp. unstable) manifold is reduced to a point. We refer the reader to that paper to see the construction which in higher dimensions (≥ 3) gives also open sets of diffeomorphisms where C^1-generic diffeomorphisms in those open subsets have these pathological type of hyperbolic measures. We remark that their examples verify that the measures have zero entropy, and it is possible in principle that positive entropy allows to recover some of the Pesin theory in the C^1-context. We also refer the reader to [AbBC] for other contexts where Pesin theory holds for C^1-diffeomorphisms.

4. Entropy and horseshoes in the presence of hyperbolic measures

4.1. **Shadowing for hyperbolic sequences of diffeomorphisms.** Again, for simplicity, we shall restrict to the case of surface diffeomorphisms.

Consider a hyperbolic sequence of diffeomorphism $\{f_n : \mathbb{R}^2 \to \mathbb{R}^2\}_n$ as defined in section 3.3. It is not hard to see that for every $R > 0$ and $z \in \mathbb{R}^2 \setminus \{0\}$ there exists n such that either $f_n \circ \ldots \circ f_0(z) \notin B(0, R)$ or $f_{-n+1}^{-1} \circ \ldots \circ f_0^{-1}(z) \notin B(0, R)$. Indeed, the only points for which is necessary to consider "negative iterates" are the points lying in the stable manifold of 0.

Here we shall consider small perturbations of the diffeomorphisms f_n and try to show that the existence of a bounded orbit remains true. This is usually called *shadowing* (at least, its applications as we shall see in subsection 4.3).

Let $\{f_n\}_n$ be a $(\lambda_1, \lambda_2, \varepsilon)$-hyperbolic sequence of diffeomorphisms and let $\{v_n = (x_n, y_n)\}_n \subset \mathbb{R}^2$ be a sequence of vectors. We consider the following sequence of diffeomorphism $\{f_n^v\}_n$ defined as:

$$\begin{aligned} f_n^v(x, y) &= (a_n x + \alpha_n(x, y), b_n y + \beta_n(x, y)) + (x_n, y_n) \\ &= (a_n x + x_n + \alpha_n(x, y), b_n y + y_n + \beta_n(x, y)) \end{aligned}$$

We say that a sequence $\{z_n\} \subset \mathbb{R}^2$ is an *orbit* of f_n^v if one has that $f_n^v(z_n) = z_{n+1}$ for every $n \in \mathbb{Z}$. An orbit $\{z_n\}_n$ is *bounded* if $\sup_{n\in\mathbb{Z}} \|z_n\| < \infty$.

Theorem 4.1 (Exponential Shadowing for hyperbolic sequences)**.** *Let f_n be a $(\lambda_1, \lambda_2, \varepsilon)$-hyperbolic sequence of diffeomorphisms (with $0 < \lambda_1 < 1 < \lambda_2$ and $\varepsilon \ll \min\{\lambda_2 - 1, 1 - \lambda_1\}$). Then, there exists $R_0 := R_0(\lambda, \mu, \varepsilon)$ such that for every $\delta > 0$ one has that if $\{v_n\}_n \subset \mathbb{R}^2$ is a sequence of vectors satisfying $\sup_n \|v_n\| \leq \delta$ then there exists a unique bounded orbit $\{z_n\}_n$ of $\{f_n^v\}_n$ which verifies the following:*

- $\sup_{n\in\mathbb{Z}} \|z_n\| \leq R_0\delta$ *and,*
- *the orbit $\{z_n\}_n$ is* uniformly hyperbolic*, that is, one has that for every $m \in \mathbb{Z}$ there exist subspaces E^s_m and E^u_m in $T_{z_m}\mathbb{R}^2$ such that $Df^v_m E^\sigma_m = E^\sigma_{m+1}$ (for $\sigma = s, u$) such that for every $n \geq 1$ one has*

$$\|D((f^v_m)^n)_{z_m} E^s_m\| \leq (\lambda_1 + 5\varepsilon)^n \ , \ \|D((f^v_m)^n)_{z_m} E^u_m\| \geq (\lambda_2 - 5\varepsilon)^n$$

Moreover, if $\{\tilde{f}_n\}_n$ is another $(\lambda_1, \lambda_2, \varepsilon)$-hyperbolic sequence of diffeomorphisms and $\{\tilde{v}_n\}_n$ another sequence of vectors such that $\sup_n \|\tilde{v}_n\| \leq \delta$ *verifying that $\tilde{f}_k = f_k$*

and $\tilde{v}_k = v_k$ *for every* $-M \leq k \leq M$ *then one has that if* $\{\tilde{z}_n\}_n$ *is the bounded orbit of* $\{\tilde{f}_n^{\tilde{v}}\}$ *then*

$$\|z_0 - \tilde{z}_0\| \leq R_0\delta((\lambda_1 + 5\varepsilon)^M + (\lambda_2 - 5\varepsilon)^{-M})$$

Remark 4.2. It follows from uniqueness that if for some $m \geq 1$ one has $f_{n+m}^v = f_n^v$ for every $n \in \mathbb{Z}$ then $z_{n+m} = z_n$ for every $n \in \mathbb{Z}$.

Proof. Consider R_0 large enough so that if one considers the square Q of side $2\delta R_0$ around $(0,0)$ one has that $f_n^v(Q)$ is a rectangle which traverses Q (see figure 3). It is clear that this can be done and the value of R_0 is independent of δ.

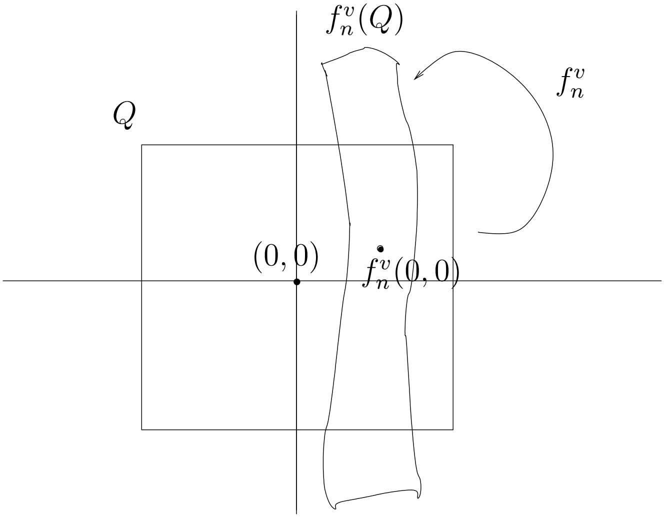


Figure 3. The image of Q by f_n^v.

As in the proof of Theorem 3.1 the square Q can be foliated by horizontal and vertical Lipschitz curves which allow to define a width of $f_n^v(Q)$. This width is contracted by a factor of $\lambda_1 + \varepsilon$. Moreover, the same argument for $(f_n^v)^{-1}$ implies that the height of $(f_n^v)^{-1}(Q) \cap Q$ is contracted by $(\lambda_2 - \varepsilon)^{-1}$. By an inductive argument one can show the existence of the desired orbit $\{z_n\}_n$ whose orbit stays always in Q (and therefore $\sup_n \|z_n\| \leq R_0 d$). Moreover, its localization depends on the intersection of the iterates of the square, so the precision on which we know the location of z_n is exponential in M if we know the form of f_k^v for $-M \leq k \leq M$.

To show uniform hyperbolicity of the orbit $\{z_n\}$ it is enough to make a cone-field argument which is similar to the one that it is possible to make to construct stable and unstable manifolds for the orbit $\{z_n\}_n$ as in Theorem 3.4.

Finally, to show uniqueness, consider two different orbits $\{w_n\}_n$ and $\{w'_n\}_n$. If w_0 differs from w'_0 in the second coordinate more than in the first, it is not hard to see using the form of f_n that $\|w_n - w'_n\| \to \infty$ as $n \to \infty$. Similarly, if the first coordinate differs more than the second, then $\|w_n - w'_n\| \to \infty$ as $n \to -\infty$. This implies that $\sup_n \|w_n - w'_n\| = \infty$ so that only one orbit can be bounded.

$\square$

4.2. **Metric entropy and Lyapunov exponents.** There exists an important relation between the entropy of an ergodic measure and its positive Lyapunov exponents. In a nutshell, entropy measures the exponential growth of the number of different segments of orbits of length n at a given precision as n goes to infinity. Lyapunov exponents measure the exponential speed at which points get separated.

It is therefore natural to expect that the entropy of a measure is bounded from above by the sum of its positive Lyapunov exponents (this is usually called Ruelle's inequality). Also, one expects that if a measure has positive Lyapunov exponents and the measure can "see" the unstable manifolds, then its entropy will be positive (this is also a well known general principle which can be attributed among others to Pesin and Ledrappier-Young). We shall only briefly review a small part of this rich theory and we refer the reader to [BaP] and references therein for a more complete account on this theory.

Before we define entropy of a measure and topological entropy we need some preliminaires. Let $f : M \to M$ be a diffeomorphism of a closed manifold M endowed with a distance d. We consider the *dynamical* or *Bowen balls* defined as:

$$B_n(x,\varepsilon) = \{y \in M \ : \ d(f^j(x), f^j(y)) \leq \varepsilon \ , \ 0 \leq j \leq n\}$$

One defines the *topological entropy* $h_{top}(f)$ of f as:

$$h_{top}(f) = \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log N_f(n,\varepsilon)$$

where $N_f(n,\varepsilon)$ is the smallest number $k > 0$ such that there exist points $x_1, \ldots, x_k$ verifying that $M = \bigcup_{i=1}^k B_n(x_i,\varepsilon)$.

For an ergodic f-invariant measure μ one defines[13] the *entropy* $h_\mu(f)$ as:

$$h_\mu(f) = -\lim_{\varepsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log \mu(B_n(x,\varepsilon)) \ , \quad \text{for } \mu\text{-almost every } x \in M$$

One can check that all the involved limits are well defined, etc (see [KH] or [M_4]).

It is a well known fact, known as the *Variational Principle* (see [M_4] for a proof) that the topological entropy is the supremum of the values of the entropies of the ergodic measures invariant under f, i.e.:

$$h_{top}(f) = \sup_{\mu \ ergodic} h_\mu(f)$$

This will be used in the two senses:

- if one knows that a diffeomorphism has positive topological entropy (for example, by knowing that the action in the first homology group is hyperbolic) then there exist measures with positive entropy,
- if there exists a measure with positive entropy, then the topological entropy is positive.

Entropy of a measure is related to Lyapunov exponents via the following results which we state for diffeomorphisms of surfaces for simplicity. See [BaP] for more general statements and proofs.

[13]This definition is due to Brin and Katok.

Theorem 4.3 (Ruelle's inequality). *Let $f : M \to M$ be a C^1-diffeomorphism of a surface M. For an ergodic f-invariant measure μ one has that if $h_\mu(f) > 0$ then μ is hyperbolic with exponents $\chi^s < 0 < \chi^u$ and moreover:*

$$h_\mu(f) \leq \min\{|\chi^s|, \chi^u\}.$$

In general, the inequality can be strict. For example, if μ is a Dirac delta measure on a hyperbolic fixed saddle p then it is ergodic, invariant and clearly hyperbolic. On the other hand, it is easy to see that $h_\mu(f) = 0$ since $\mu(B_n(p,\varepsilon)) = 1$ for every n and ε.

To obtain entropy of a measure one needs that the measure "sees" the expansion. This can be formulated in the following form for surfaces (again, this is far from being optimal, see [BaP] for more general statements). The following result will follow from Katok's theorem which we shall review in the next section, but it admits more quantitative versions (which depend on desintegration of measures along a lamination and that is why we refer the interested reader to read this elsewhere, for example [BaP]).

Theorem 4.4. *Let μ be a hyperbolic measure of a $C^{1+\alpha}$ diffeomorphism whose support is not finite. Then $h_\mu(f) > 0$.*

There is however one case where the desintegration of the measure can be excluded from the statement and which is important in some contexts. We recall that a diffeomorphism is conservative if it preserves a volume form *vol*. In general, it is too restrictive to assume that *vol* is ergodic, so, in general, the Lyapunov exponents of *vol* are f-invariant functions instead of constants.

Theorem 4.5 (Pesin's entropy formula). *Let $f : M \to M$ be a conservative $C^{1+\alpha}$-diffeomorphism of a surface M. Then one has that*

$$h_{vol}(f) = \int \chi^u dvol = -\int \chi^s dvol.$$

See [M_2] for a hands on proof which does not rely (explicitely) on the absolute continuity of the unstable lamination of the measure. In particular, this proof is one of the first instances where the use of the subexponential size of Pesin charts is used to study the dynamics without the need to construct the invariant manifolds first.

4.3. **Katok's theorem on the existence of horseshoes.** In this subsection we shall explain a stronger version of Theorem 4.4. It is by now a classical result due to Katok (see [KH, Supplement], [Ge] or the appendix of [AvCW] for more general versions and some improvements) that the existence of a hyperbolic measure which is not periodic implies the existence of horseshoes in the $C^{1+\alpha}$ context.

The following is the precise statement in the surface case:

Theorem 4.6 (Katok). *Let $f : M \to M$ be a $C^{1+\alpha}$ diffeomorphism of a closed surface M and μ an ergodic f-invariant measure whose support is not finite. Then, $h_\mu(f) > 0$ and for every $\varepsilon > 0$ there exists a compact f-invariant subset Λ contained in the support of μ such that:*

- *the set Λ is* uniformly hyperbolic*,*
- *the topological entropy $h_{top}(f|_\Lambda)$ of f restricted to Λ is close to $h_\mu(f)$, i.e.*

$$h_{top}(f|_\Lambda) > h_\mu(f) - \varepsilon$$

For the definition of uniform hyperbolicity we refer the reader to section 5. In a uniformly hyperbolic set with positive topological entropy there are infinitely many hyperbolic periodic orbits and it is not hard to show the existence of a *transverse homoclinic intersection* which is well known to produce a *horseshoe* (regardless of the definition of horseshoe that we have not given). We refer the reader to [KH, Chapter 6.5] for a more complete account. We shall explain now the main ingredients of the proof of Theorem 4.6.

The key point is to work on what are sometimes called *Pesin blocks* or *uniformity blocks*. These are subsets on which the constants are uniform (the choice of what constants are chosen to be uniform vary in the literature). For example, in this case, one can consider, for a given $K > 0$ the set Λ_K of points $x \in M$ such that the value of $\rho(x) \geq \frac{1}{K}$ of Theorem 3.5 as well as $\|C_\nu(x)\| + \|C_\nu(x)^{-1}\|^{-1} \leq K$ of 2.4. By reducing Λ_K up to an arbitrarily small measure subset, one can assume that Λ_K is compact. Moreover, for given $\varepsilon > 0$ there exists K such that Λ_K verifies that $\mu(\Lambda_K) \geq 1 - \varepsilon$. Also, it is a standard fact that one can assume that all the involved functions (ρ, C_ν, etc) are continuous on Λ_K (this is the classical Luisin's theorem).

If the set Λ_K were f-invariant this would conclude since it can be chosen as to have entropy as close to $h_\mu(f)$ as one desires. However, in general there is no reason to expect that Λ_K will be f-invariant.

One proceeds then as follows: one considers first a finite covering of the support of μ by Pesin charts which has a Lebesgue number and covers Λ_K. In Λ_K one has uniform charts where the constants of C_ν are bounded, and there remains a finite number of transition charts where points go when the iterates do not belong to Λ_K. In such a way one can construct a large number of "adapted orbits" of f which "see" almost all the entropy of μ and return to Λ_K quite frequently.

This allows to construct enough periodic pseudo-orbits which will be shadowed by periodic orbits of f which are uniformly hyperbolic because they belong to Λ_K. The way to "shadow" these orbits is using Theorem 4.1 by looking at orbits that remain always in the places where the lift of the dynamics given by Theorem 3.5 coincides with f. This step is the most delicate since to ensure that the orbits remain in the Pesin charts at all times one has to carefully choose the orbit one wishes to shadow in order to control its orbit. See [KH, Lemma S.4.10] for more details.

The fact that the periodic orbits one construct are hyperbolic and sufficiently close to each other allows one to show that they are all homoclinically related and therefore belong to the same f-invariant compact subset Λ which is transitive and uniformly hyperbolic. The entropy is as close to $h_\mu(f)$ as desired.

Recently, in [Sar], these arguments have been improved to construct sets which see *all* the entropy of μ. The idea is consider a Pesin chart at each point and consider a countable subcovering. Then, one considers the possible itineraries that orbits make through those charts and constructs a countable Markov partition with similar ideas as those of Bowen for constructing Markov partitions. The sets constructed in [Sar] cease to be uniformly hyperbolic but enjoy coding properties for which much information is known. See [Sar] for more details on this and the previous construction.

4.4. **When is a measure hyperbolic?** Clearly, because of Ruelle's inequality (Theorem 4.3) positive entropy is a sufficient condition for a measure to be hyperbolic in dimension 2. There are some times where one does not have enough information on the invariant measure in order to compute its entropy. It is therefore important to have other methods to guarantee existence of positive Lyapunov exponents. Here it is important to remark that in some (very important) applications one deals with non-ergodic (or a priori non-ergodic) measures where positive entropy only guarantees some ergodic component to have positive Lyapunov exponents. I would like to mention three known methods for establishing hyperbolicity of a measure.

The first one has been developed independently by Lewowicz and Wojkowski (see [Pot$_3$, Section 2.1] and references therein) and it is the method of measurable cone-fields or quadratic forms. This has been quite useful to establish hyperbolicity (and more recently the Bernoulli property) to a large class of billiards (see [DeM]).

The second is also related to cone-fields but it deals more with the notion of *critical points* and extends the ideas which were developed in the setting of one-dimensional dynamics. This was done famously by Benedicks-Carlesson ([BeC]) to study the parameters for which the Hénon family admits non-uniformly hyperbolic attractors and has been used largely since then (see [Ber$_2$] for improvements of that result as well as a panorama of the works related to this).

More recently, in the setting of conservative twist maps, Arnaud has developed some techniques to compute Lyapunov exponents for such maps and discovered some interesting relations of these with the shape of the so called Aubry-Mather set and Green bundles. Explaining this is possibly the objective of Marie-Claude's minicourse, but we also refer to her lecture notes [Arn] (and references therein) for more information.

5. Uniform estimates

In this section we shall briefly present the concepts of dominated splitting, partial hyperbolicity, normal hyperbolicity, etc. Essentially, one can think these concepts as uniform versions of the hyperbolicity of measures: If a compact set admits a continuous splitting of the tangent bundle such that for every measure supported in the compact set, there is a positive gap between the Lyapunov exponents along each of the bundles, then, the compact set is said to admit a dominated splitting. Under this conditions, it is no longer needed to have control on the modulus of continuity of the derivative in order to perform the graph-transform arguments. Moreover, under some assumptions of hyperbolicity, one obtains results of persistence of invariant manifolds which are quite useful in many applications; particularly (in view of the interests of this conference) we mention the use of normally hyperbolic cylinders in the recent proofs of Arnold diffusion ([BeKZ, GK, KZ] and references therein).

M will denote a d-dimensional manifold and $f : M \to M$ a C^1-diffeomorphism.

5.1. **Dominated splittings.** Let $\Lambda \subset M$ be a compact f-invariant set. We say that it admits a *dominated splitting of index i* if there is a continuous splitting $T_\Lambda M = E \oplus F$ (i.e. for every $x \in \Lambda$ one has $T_x M = E(x) \oplus F(x)$ and E and F are continuous functions) such that the bundle $E(x)$ has dimension i and verifies the following properties:

- (**Invariance**) The bundles are Df-invariant, that is, for every $x \in \Lambda$ one has $Df_x(E(x)) = E(f(x))$ and $Df_x(F(x)) = F(f(x))$.
- (**Domination**) There exists $N > 0$ such that for every $x \in \Lambda$ and vectors $v_E \in E(x) \setminus \{0\}$ and $v_F \in F(x) \setminus \{0\}$ one has that:

$$\frac{\|Df_x^N v_E\|}{\|v_E\|} < \frac{\|Df_x^N v_F\|}{\|v_F\|}$$

It follows from compactness that there exists $\lambda \in (0,1)$ such that

$$\frac{\|Df_x^N v_E\|}{\|v_E\|} < \lambda \frac{\|Df_x^N v_F\|}{\|v_F\|}$$

It is possible to choose an *adapted metric* for which the value of N is equal to 1 (see [Gou]). The fact that the splitting is *dominated* is independent of the choice of the Riemannian metric.

Exercise. Show that a continuous Df-invariant decomposition $T_\Lambda M = E \oplus F$ is dominated if and only if there exists $\nu > 0$ such that for every ergodic measure $\mu \in \mathcal{M}_{erg}(f)$ supported on Λ one has that the largest Lyapunov exponent $\chi_E^+(\mu)$ of μ along E and the smallest Lyapunov exponent $\chi_F^-(\mu)$ of μ along F verify:

$$\chi_E^+(\mu) \leq \chi_F^-(\mu) - \nu.$$

In particular, show that if the splitting is dominated, then the Oseledets splitting respects (and refines) the splitting $E \oplus F$.

It is not hard to show that when there is a dominated splitting on a subset Λ, the angle between the subbundles of the domination is uniformly bounded from below, this follows directly by continuity of the bundles and compactness of Λ. We remark here that the continuity of the bundles is not essential in the definition of domination and it follows from the rest of the properties (see [BoDV, Appendix B]).

A key property of dominated splitting is that it is *robust*:

Proposition 5.1. *Let $\Lambda \subset M$ be a compact set admitting a dominated splitting of the form $T_\Lambda M = E \oplus F$ for a diffeomorphism $f : M \to M$ of class C^1. Then, there exists a compact neighborhood U of Λ in M and a neighborhood $\mathcal{U}$ of f in the C^1-topology such that every $g \in \mathcal{U}$ verifies that the set $\Lambda_g = \bigcap_{n\in\mathbb{Z}} g^n(U)$ admits a dominated splitting $T_{\Lambda_g} M = E_g \oplus F_g$ with* $\dim E_g = \dim E$.

We shall not prove this result since it is not in the spirit of this notes, the proof is not hard, see for example [BoDV, Appendix B].

If Df preserves a continuous subbundle $E \subset T_\Lambda M$ we say that E is *uniformly contracted* (resp. *uniformly expanded*) if there exists $N > 0$ such that for every $x \in \Lambda$ and every unit vector $v \in E(x)$ one has that

$$\|Df_x^N v\| \leq \frac{1}{2} \quad (\text{resp.} \quad \geq 2)$$

Exercise. Show that a continuous Df-invariant subbundle $E \subset T_\Lambda M$ is uniformly contracted if and only if every ergodic measure μ supported on Λ has all Lyapunov exponents corresponding to vectors in E negative.

5.2. **Plaque families.** When one has a dominated splitting on a compact subset $\Lambda \subset M$, as we mentioned, one can assume that for every f-invariant ergodic measure verifies that its Oseledets splitting respects the splitting given by the domination. So, in a sense, this means that when considering linear change of coordinates which make the bundles orthogonal these changes of coordinates become uniformly bounded. It is in a sense as if the norm of the maps C_ν and C_ν^{-1} of Theorem 2.4 are uniformly bounded. This is not exactly true since the norm of C_ν and C_ν^{-1} also depend on how quickly the derivative starts behaving as its "limit behaviour".

Let us state another result ([HPS, Theorem 5.5]) which is still in the spirit of the graph transform argument. We shall only sketch its proof. We denote as $\mathbb{D}^k$ to the k-dimensional disk of unit radius in $\mathbb{R}^k$ and $\mathrm{Emb}^1(\mathbb{D}^k, M)$ to the space of C^1-embeddings of $\mathbb{D}^k$ into M. We denote as $\mathbb{D}^k_r \subset \mathbb{D}^k$ to the disk of radius $r \leq 1$

Theorem 5.2 (Plaque Families). *Let $f : M \to M$ be a C^1-diffeomorphism and $\Lambda \subset M$ a compact f-invariant subset admitting a dominated splitting of the form $T_\Lambda M = E \oplus F$. Then, there exists a continuous[14] family $\mathcal{D}_E : \Lambda \to \mathrm{Emb}^1(\mathbb{D}^{\dim E}, M)$ with the following properties:*

- *(**Tangency:**) for every $x \in \Lambda$ one has that $\mathcal{D}_E(x)(0) = x$ and the image of $\mathcal{D}_E(x)$ is tangent to $E(x)$ at x.*
- *(**Local invariance:**) there exists $r_0 < 1$ such that for every $x \in \Lambda$ one has that $f(\mathcal{D}_E(x)(\mathbb{D}^{\dim E}_{r_0})) \subset \mathcal{D}_E(f(x))(\mathbb{D}^{\dim E})$.*

SKETCH Using continuity of the bundles one can choose[15] a continuous linear change of coordinates $C(x) : \mathbb{R}^d \to T_xM$ (recall that $d = \dim M$) such that $C(x)(\mathbb{R}^{\dim E} \times \{0\}^{\dim F}) = E(x)$ and $C(x)(\{0\}^{\dim E} \times \mathbb{R}^{\dim F}) = F(x)$. Using the exponential map $\exp : TM \to M$ one can construct uniform charts around each point $x \in \Lambda$ of the form $\xi_x := \exp_x \circ C(x) : B(0,R) \to M$ verifying that for $y, y' \in B(0,R)$ one has that $\frac{1}{K} d(\xi_x(y), \xi_x(y') \leq \|y - y'\| \leq K d(\xi_x(y), \xi_x(y'))$. Here $R > 0$ and $K > 0$ are fixed constant independent of x.

One can, by using the same technique as in the proof of Theorem 3.1 lift the dynamics by extending the map $\tilde{f}_x := \xi_{f(x)}^{-1} \circ f \circ \xi_x : B(0, R/K\|Df_x\|) \to B(0,R)$ to a diffeomorphism $\hat{f}_x : \mathbb{R}^d \to \mathbb{R}^d$ which in coordinates $\mathbb{R}^d = \mathbb{R}^{\dim E} \oplus \mathbb{R}^{\dim F}$ can be expressed as:

$$\hat{f}_x(v,w) = (A_x v + \alpha_x(v,w), B_x w + \beta_x(v,w))$$

where $A_x : \mathbb{R}^{\dim E} \to \mathbb{R}^{\dim E}$ and $B_x : \mathbb{R}^{\dim F} \to \mathbb{R}^{\dim F}$ are linear transformation which by the domination[16] condition satisfy $\|A_x\| < \lambda \|B_x^{-1}\|^{-1}$ for some $\lambda \in (0,1)$ and such that the C^1 size of α_x and β_x is smaller than $\varepsilon \ll 1 - \lambda$.

Now, for a given $x \in \Lambda$ we can consider the sequence $\{f_n\}_n$ of diffeomorphisms of $\mathbb{R}^d$ defined as $f_n = \hat{f}_{f^n(x)}$. For this sequence it is possible to consider the

[14]Notice that this has only sense when the bundle $E \subset T_\Lambda M$ is trivializable (for example, when Λ is totally disconnected). Technically, it would be more correct to write that $\mathcal{D}_E : \Lambda \to \mathrm{Emb}^1(E_1, M)$ such that $\mathcal{D}_E(x)$ is an embedding of $E_1(x)$ in M where $E_1(x)$ denotes the disk of radius 1 in $E(x)$.

[15]This is not strictly true if the bundle is not trivializable over Λ. But we shall ignore this technical (and unimportant) issue.

[16]We remark that here we are assuming for simplicity that the dominated splitting comes with an adapted norm. This is no loss of generality, but the same argument can be adapted not to use it.

space of graphs of Lipschitz functions from $\mathbb{R}^{\dim E}$ to $\mathbb{R}^{\dim F}$. As in the proof of Theorem 3.4 one shows that the graph transform induced by the sequence f_n is a contraction for a suitable metric and so there exists a unique sequence of graphs which is invariant under the sequence $\{f_n\}_n$ and it is indeed by C^1-graphs which are tangent to $\mathbb{R}^{\dim E} \times \{0\}^{\dim F}$ at $(0,0)$.

Sending the intersection of the graphs with $B(0,R)$ by ξ_x to M one obtains the desired embedding and notice that the intersection with $B(0,R/K\|Df_x\|)$ is sent to the next graph since it remains in the place where $\hat{f}_x$ coincides with $\tilde{f}_x$. This concludes the proof.

□

Remark 5.3. This result does not provide uniqueness of the plaque families since there is no natural way to lift f to the functions $\hat{f}_x$. This means that for each choice of lift $\{\hat{f}_x\}_{x\in\Lambda}$ of the dynamics one obtains an a priori different plaque family. However, if there are dynamical conditions, for example if $y \in \mathcal{D}_E(x)(\mathbb{D}^{\dim E})$ and $f^n(y) \in \xi_{f^n(x)}(B(0,R))$ for every $n \geq 0$ then the point y will belong to every sufficiently large plaque family.

Notice that one can perform the graph transform argument by starting with a foliation of a neighborhood of x and obtain locally invariant local foliations which are almost tangent to E (or F). This is done in [BuW$_2$, Section 3] where *fake foliations* are constructed. Those fake foliations have some technical applications (notably to the study of stable ergodicity when the center direction is not integrable). One should not be confused by the existence of these local foliations almost tangent to E since it is possible that the bundle E is not locally integrable at any point of the manifold (see [BuW] for examples).

Exercise. By combining Theorem 5.2 and the ideas used for Theorem 3.3 try to show that Theorem 3.7 is valid for C^1-diffeomorphisms of surfaces if the support of the measure admits a dominated splitting.

The previous exercise is a particular case of a more general result which states that much of Pesin's theory works in the C^1-setting if one assumes domination on the support of the invariant measures. See [AbBC] for precise statements and proofs.

5.3. **Uniform hyperbolicity and partial hyperbolicity.** Consider a compact f-invariant set Λ and assume that Df-preserves a continuous splitting of $T_\Lambda M$ into three bundles of the form:

$$T_\Lambda M = E^s \oplus E^c \oplus E^u$$

where E^s is uniformly contracted, E^u is uniformly expanded and the splittings $E^s \oplus (E^c \oplus E^u)$ and $(E^s \oplus E^c) \oplus E^u$ are dominated. We say that:

- Λ is *uniformly hyperbolic* if $E^c = 0$.
- Λ is *partially hyperbolic* if either E^s or E^u is non-zero.
- Λ is *strongly partially hyperbolic* if both E^s and E^u are non-zero.

The study of diffeomorphisms for which their limit set (more precisely their chain-recurrent set) is uniformly hyperbolic is one of the milestones of study of dynamical systems from the pioneering work of Anosov and Smale in the 60's to the present. Its study has interacted with the study of geometry and topology as well as it has been the starting point to many advances in different areas of mathematics. One of the main tools of its study is the following classical result.

Theorem 5.4 (Shadowing Theorem). *Let $f : M \to M$ be a C^1-diffeomorphism and $\Lambda \subset M$ a compact f-invariant hyperbolic subset. For every $\varepsilon > 0$ there exists $\delta > 0$ such that if $\{z_n\}_n \subset \Lambda$ is an δ-pseudo orbit (i.e. a sequence such that $d(z_{n+1}, f(z_n)) \leq \delta$) there exists a point $y \in M$ such that its orbit ε-shadows $\{z_n\}_n$ (i.e. one has $d(f^n(y), z_n) \leq \varepsilon$). Moreover:*

- *one has that $\delta \to 0$ as $\varepsilon \to 0$,*
- *if ε is small enough, the point y whose orbit shadows $\{z_n\}_n$ is unique,*
- *if there exists $m > 0$ such that $z_{n+m} = z_n$ for all $n \in \mathbb{Z}$ one can choose y to be a periodic orbit of period m,*
- *if Λ is* locally maximal *(i.e. if there exists a neighborhood U of Λ such that $\Lambda = \bigcap_n f^n(U)$) then the point y can be chosen to belong to Λ.*

Sketch The proof follows exactly the same lines as the proof of Theorem 4.6 but it is much easier. Indeed, one chooses uniform charts and applies exactly the same argument as in the proof of Theorem 4.1.

$\square$

It has been necessary to understand the global panorama of dynamical systems to consider weaker notions of hyperbolicity. In some cases, non-uniform hyperbolicity has been the right generalization, but in many others, it turns out that dominated splittings or partial hyperbolicity have been more suitable. They verify the following general theorem in the same lines as the results we present in this notes.

Theorem 5.5 (Stable Manifold Theorem). *Let $f : M \to M$ be a C^1-diffeomorphism and let $\Lambda \subset M$ be a compact f-invariant set with a partially hyperbolic splitting of the form $T_\Lambda M = E^s \oplus E^{cu}$ where the bundle E^s is uniformly contracted. Then, there exists a continuous family $\mathcal{W}^s_{loc} : \Lambda \to \mathrm{Emb}^1(\mathbb{D}^{\dim E^s}, M)$ with the following properties:*

- *(**Tangency:**) for every $x \in \Lambda$ one has that $\mathcal{W}^s_{loc}(x)(0) = x$ and the image of $\mathcal{W}^s_{loc}(x)$ is tangent to $E^s(x)$ at x,*
- *(**Invariance:**) for every $x \in \Lambda$ one has that*
$$f(\mathcal{W}^s_{loc}(x)(\mathbb{D}^{\dim E^s})) \subset \mathcal{W}^s_{loc}(f(x))(\mathbb{D}^{\dim E^s}),$$
- *(**Convergence:**) if y is in the image of $\mathcal{W}^s_{loc}(x)$ then $d(f^n(x), f^n(y)) \to 0$ exponentially fast as $n \to +\infty$,*
- *(**Uniqueness:**) if one considers for each $x \in \Lambda$ the* strong stable set
$$\mathcal{W}^{ss}_x = \bigcup_n f^{-n}(\mathcal{W}^s_{loc}(f^n(x))(\mathbb{D}^{\dim E^s}))$$
it follows that for $x, y \in \Lambda$ the sets $\mathcal{W}^{ss}_x$ and $\mathcal{W}^{ss}_y$ are injectively immersed submanifolds which are either disjoint or coincide.

Proof. It follows almost directly from Theorem 5.2 and Remark 5.3. Indeed, one can consider any plaques family given by Theorem 5.2 and use the fact that $\|Df_x|_{E^s}\| < \lambda < 1$ to see that the diameter of the forward iterates of the plaques converges exponentially fast to zero. This gives invariance and convergence. Uniqueness follows from the fact that independently on the choice of lift, the plaque families will coincide up to their size (see Remark 5.3) so that when considering the set of points that eventually lie in a plaque one has uniqueness.

□

Remark 5.6. It is possible to show that graph transform argument varies continuously with the diffeomorphism in compact sets so that if $f_n \to f$ then the strong stable manifolds (resp. strong unstable manifolds) for f_n converge in compact subsets to those of f.

5.4. **Normal hyperbolicity and persistence.** It is sometimes useful to perform the graph transform method in a more global way. This is the case in the proof of persistence of normally hyperbolic submanifolds or foliations. We refer the reader to [HPS] or [Ber] for detailed proofs.

Consider $f : M \to M$ a C^1-diffeomorphism and $\Lambda \subset M$ a compact f-invariant set. We shall assume that Λ is *laminated* by an f-invariant lamination $\mathcal{L}$. This means that for each $x \in \Lambda$ there exists a C^1-injectively immersed submanifold $\mathcal{L}(x) \subset \Lambda$ with the following properties:

- if $\mathcal{L}(x) \cap \mathcal{L}(y) \neq \emptyset$ then $\mathcal{L}(x) = \mathcal{L}(y)$,
- if $x_n \to x$ then $\mathcal{L}(x_n)$ converges to $\mathcal{L}(x)$ uniformly in the C^1-topology in compact subsets,
- the map $x \mapsto T_x\mathcal{L}(x) \subset T_xM$ defines a continuous distribution.

The f-invariance means that $f(\mathcal{L}(x)) = \mathcal{L}(f(x))$.

We say that the lamination $\mathcal{L}$ is *normally hyperbolic* if f admits a partially hyperbolic splitting $T_\Lambda M = E^s \oplus E^c \oplus E^u$ where $E^c(x) = T_x\mathcal{L}(x)$ for every $x \in \Lambda$. Moreover, we say it is *normally expanded* (resp. *normally contracted*) if $E^s = \{0\}$ (resp. $E^u = \{0\}$).

Remark 5.7. Notice that if $\mathcal{L}$ is a lamination by points, normal hyperbolicity of $\mathcal{L}$ is equivalent to have that Λ is uniformly hyperbolic.

Whenever there is a normally hyperbolic lamination, one has the following persistence result:

Theorem 5.8 (Stability of normally hyperbolic laminations)**.** *Let $f : M \to M$ be a C^1-diffeomorphism, leaving invariant a normally hyperbolic lamination $\mathcal{L}$ on a compact set Λ. Then, there exists a C^1-neighborhood $\mathcal{U}$ of f such that for every $g \in \mathcal{U}$ there exists a compact g-invariant set Λ_g close to Λ such that:*

- *(**Continuation of leaves:**) for every $x \in \Lambda$ there exists a manifold L^x_g diffeomorphic to $\mathcal{L}(x)$ such that if one considers an immersion $i_x : \mathcal{L}(x) \hookrightarrow M$ there is an immersion $i^g_x : L^x_g \to M$ (possibly no longer injective) such that i_x and i^g_x are C^1-close and L^x_g is everywhere tangent to E^c_g (the continuation of the bundle E^c of f for g in Λ_g),*
- *(**Invariance:**) one has that $f(i^g_x(L^x_g)) = i^g_{f(x)}(L^{f(x)}_g)$,*
- *(**Continuity:**) The leafs $i^g_x(L^x_g)$ with $x \in \Lambda$ saturate Λ_g and vary continuously in the C^1-topology in compact subsets.*

The idea of the proof is to perform a graph transform argument in an entire neighborhood of the immersion. This involves unwrapping the immersion to an abstract immersion into a neighborhood of the leaf which depends on the point and then applying arguments very similar to the ones we have already done albeit more technical.

This result is not completely satisfactory since in principle leafs of the new "lamination" could merge. One sometimes calls this *branching laminations* (sometimes, they are useful for some purposes, see [BuI, Pot_5, HP, HP_2] for use of this notion).

Under a technical condition (which is always satisfied in case the lamination can be extended to a neighborhood into a C^1-foliation) it is possible to improve Theorem 5.8 to have a true lamination for diffeomorphisms close to f. This condition is known by the name of *plaque-expansiveness* and we refer the reader to [HPS] and [Ber] for more information about it. We also refer the reader to [BuW] for information on the related notion of *dynamical coherence*.

We make the following remarks on Theorem 5.8 since we shall not enter in the details of its proof. The first remark is that to be able to perform a global graph transform one uses strongly the fact that the dynamics are C^0-close (not only that the invariant bundles are close), this can be noticed by the fact that (a strong version of Theorem 5.8, the one appearing in [HPS]) implies that after C^1-perturbation, the f-invariant foliation remains homeomorphic to the initial one while there might be very different topological type of foliations which are tangent to closeby distributions (just think about linear foliations on tori). The other remark is that even in the simplest case of a closed submanifold $N \subset M$ which is normally hyperbolic, the graph transform must be performed with some care since does not have a priori a fixed point on which to "center" the graph transform argument. We refer the reader to [BerB] for a short proof in this particular and easier case.

5.5. **Reducing the dimension.** Possibly, the most important information given by the existence of a dominated splitting or of the existence of a partially hyperbolic splitting comes with the fact that Theorem 5.2 allows one to "reduce the dimension" of the study. In general, if one has a strong partially hyperbolic splitting, one can use Theorem 5.2 to reduce the situation to a kind of skew-product over a hyperbolic set, at least, one can think the skew-product over a hyperbolic set as a *toy model* for the general situation. This approach has been very successful when $\dim E^c = 1$ (see [Cr]).

However, there are some cases where the reduction of dimension is even more drastic, instead of obtaining a sequence of maps of a lower dimensional manifolds, one can in some cases deal with a unique one. This is the case when the dynamics one is interested in lives in a normally hyperbolic submanifold. As we have seen, this is a robust property, and we shall quickly review in this subsection a result due to Bonatti and Crovisier ([BoC2]) which allows to detect this situation.

Let us state their result.

Theorem 5.9 (Bonatti-Crovisier). *Let Λ be a compact f-invariant set admitting a partially hyperbolic splitting of the form $T_\Lambda M = E^{cs} \oplus E^u$. Assume moreover that for every $x \in \Lambda$ one has that $\mathcal{W}^{uu}(x) \cap \Lambda = \{x\}$. Then, there exists a C^1-submanifold $\Sigma \subset M$ containing Λ and tangent to E^{cs} at every point of Λ such that it is locally invariant (i.e. one has that $f(\Sigma) \cap \Sigma$ is a neighborhood of Λ relative to Σ).*

Exercise. Show that if Λ is partially hyperbolic and it is contained in such a submanifold, then one has that $\mathcal{W}^{uu}(x) \cap \Lambda = \{x\}$ for every $x \in \Lambda$.

5.6. **When can you guarantee the existence of a dominated splitting?** One has the classical cone-field criteria (see [BoGo]) which ensures domination and can be checked with only finitely many iterates (notice that the explicit bundles depend on the complete orbit of the point). This criteria is the one used to prove Proposition 5.1.

There are also criteria to ensure domination when one has information on certain robust properties of the diffeomorphism, a classical result in this line is [BoDP] (see also [BoDV, Chapter 7]). Also, in the lines of a celebrated conjecture due to Jacob Palis, one knows that far from homoclinic tangencies, the dynamics is *partially hyperbolic* (see [CrSY] and references therein).

In the same spirit as the critical points of Benedicks-Carleson, for surface dynamics there exists the critical point criteria to admit dominated splitting first introduced in [PuRH] and further improved by Crovisier and Pujals [CrPu] (see also [Va] for developments in the holomorphic setting).

6. Attractors and the geometry of unstable laminations

In this section we give a glimpse in further topics which use the tools developed in this notes. They represent a very biased choice based on the author's interests.

The main point is to show some of the results of the notes in "action". First, we shall explain how (with help of some results we will just cite) the ideas in the text allow to show that in dimension 2 there is an open and dense subset of diffeomorphisms in the C^1 topology admitting a hyperbolic attractor. This result is part of Araujo's thesis [Ara]. His proof had a gap, and the result became folklore after the results of Pujals-Sambarino ([PuS]). We shall present the proof that appeared in [Pot] (which uses [PuS] but also some other recent results, notably [BoC]).

After we have presented the proof of this result in dimension 2, we shall try to present quickly (with much less details) a recent joint result with Sylvain Crovisier and Martín Sambarino on finiteness of attractors for certain differentiable dynamics which explores the geometry of the strong unstable manifolds.

We refer the reader to [Pot_4, Chapter 3] for a wider panorama on attractors for differentiable dynamics. We also strongly recommend [Cr_3] for a more global point of view of differentiable dynamics on manifolds with plenty of pertinent references. At this point we wish to point out the important influence of the work of Mañé in this type of results, we mention in particular two landmark papers of his [M_1, M_3]. Also, recently we have written some notes with S. Crovisier which complement and extend the material presented here [CrPo].

6.1. Some preliminaries.

We start by introducing some preliminaries in the study of differentiable dynamics. We consider C^1-diffeomorphism f of a closed d-dimensional manifold M.

A *topological attractor* is a compact invariant set Λ such that there exists an open set U verifying $f(\overline{U}) \subset U$ and $\Lambda = \bigcap_{n\geq 0} f^n(\overline{U})$.

Exercise. Assume that Λ is a partially hyperbolic topological attractor with splitting $T_\Lambda M = E^{cs} \oplus E^u$ where E^u is uniformly expanded. Show that:

- The set Λ is saturated by strong unstable manifolds $\mathcal{W}^{uu}$ (i.e. the strong stable manifolds for f^{-1}, c.f. Theorem 5.5).
- There exists an ergodic invariant measure μ such that the sum of its Lyapunov exponents is ≤ 0. As a consequence, μ has at least one strictly negative Lyapunov exponent.

In fact, one can see that the first assertion of the second item does not need the fact that Λ is partially hyperbolic.

Topological attractors are not completely satisfying, for example, it always holds that the whole manifold M is a topological attractor. In general, one adds some sort of indecomposability hypothesis to the definition of attractor[17]. We say that Λ is an *attractor* for f if it is a topological attractor and $f|_\Lambda$ is transitive. The *basin* of Λ is the set of points whose omega-limit set is contained in Λ. In the case where Λ is an attractor it is $\bigcup_n f^{-n}(U)$.

Exercise. Show that if $\Lambda \subset M$ is an uniformly hyperbolic attractor of f then:

- its basin is an open set of M,
- there exists a neighborhood $\mathcal{U}$ of f and a neighborhood U of Λ such that for $g \in \mathcal{U}$ one has that $g(\overline{U}) \subset U$ and $\Lambda_g = \bigcap_n g^n(U)$ is a uniformly hyperbolic attractor.

Attractors do not always exists (even for C^r-generic dynamics, see [BoLY, Pot$_2$]). So, one sometimes uses the notion of *quasi-attractors*. We say that a compact f-invariant set is a *quasi-attractor* if:

- (**Intersection of topological attractors:**) there exist a basis of neighborhoods U_n of Λ such that $f(\overline{U_n}) \subset U_n$,
- (**Indecomposability:**) if U is an open set such that $f(\overline{U}) \subset U$ and $\Lambda \cap U \neq \emptyset$ then $\Lambda \subset U$.

The second hypothesis is equivalent to Λ being chain-transitive which we shall not define here. A remarkable result due to Bonatti and Crovisier states the following:

Theorem 6.1 (Bonatti-Crovisier [BoC])**.** *There exist a residual (i.e. G_δ-dense) subset $\mathcal{G} \subset \mathrm{Diff}^1(M)$ such that if $f \in \mathcal{G}$ then:*

- *There exists a residual subset $R_f \subset M$ such that for every $x \in R_f$ the omega-limit set of x for f is contained in a quasi-attractor.*
- *If a quasi-attractor Λ contains a periodic point p then it coincides with its* homoclinic class *$H(p)$ (i.e. the closure of the transverse intersections between the orbits of $W^s(p)$ and $W^u(p)$).*

Exercise. Let p be a hyperbolic saddle and $H(p)$ its homoclinic class. Show that $f|_{H(p)}$ is transitive.

6.2. **Attractors in surfaces.** The point of this subsection is to explain the following result:

Theorem 6.2 (Araujo [Ara])**.** *For a given closed surface M, there exists a residual subset $\mathcal{G}_A$ of $\mathrm{Diff}^1(M)$ such that if $f \in \mathcal{G}_A$ then:*

- *either there are infinitely many attracting periodic points (sinks),*
- *either there are finitely many uniformly hyperbolic attractors whose basins cover an open and dense subset (of full Lebesgue measure) of M.*

By the robustness properties of hyperbolic attractors one deduces as a consequence that there exists an open and dense subset of $\mathrm{Diff}^1(M)$ for which there exists hyperbolic attractors. This contrasts with the situation in higher dimensions ([BoLY]).

[17]We warn the reader that there are plenty of possible definitions of attractors, the one we choose, even if quite common, is far from being the unique.

Let us explain the main ideas of the proof. Since having infinitely many sinks is a G_δ-property, we shall assume that f cannot be approximated by a diffeomorphism with infinitely many sinks. Moreover, we can without loss of generality assume that $f \in \mathcal{G}$ of Theorem 6.1 and that the number of sinks that f has is constant[18] in a neighborhood of f.

Now let Λ be a quasi-attractor for f. Notice that since $f \in \mathcal{G}$ such a quasi-attractor exists. We shall assume that Λ is not a sink, otherwise there is nothing to prove. We must show that Λ is uniformly hyperbolic. The key point is to show that Λ admits a dominated splitting since in that case it follows as a consequence of the results of [PuS] that it is uniformly hyperbolic.

So let us show:

Proposition 6.3. Λ *admits a dominated splitting.*

PROOF. This proof is in the lines of what is discussed in subsection 5.6.

Since there exists a neighborhood $\mathcal{U}$ of f such that for $g \in \mathcal{U}$ one has that g has finitely many sinks, one can assume that one cannot create a sink in a neighborhood U of Λ by perturbing f.

We use first that there must exist a measure μ supported in Λ whose sum of Lyapunov exponents is ≤ 0. To show that f admits a dominated splitting in the support of μ we use classical arguments (see for example [AbBC]) which imply that otherwise one can create a sink in an arbitrarily small neighborhood of the support of μ by a small perturbation of f.

Notice that in principle, the support of μ may be smaller than Λ itself. However, we know that f admits a dominated splitting on the support of μ and that the sum of Lyapunov exponents is ≤ 0. This implies that μ is hyperbolic, since otherwise, using Theorem 3.3 one would get a sink[19]. Even if f is only C^1, since it admits a dominated splitting in the support of μ one deduces that one can apply Theorem 4.6 to show that Λ contains periodic points and thus, using again that $f \in \mathcal{G}$ conclude that Λ is the homoclinic class of a periodic point such that the sum of its Lyapunov exponents is ≤ 0.

A classical argument of *transitions* (see [BoDV, Chapter 7]) implies that there is a dense set of periodic points in Λ such that the sum of Lyapunov exponents is ≤ 0. If there were not a dominated splitting in Λ, then a small perturbation (see again [BoDV, Chapter 7]) allows one to construct a sink. This concludes.

□

As we mentioned, by [PuS] this implies that Λ is a hyperbolic attractor. It remains to show that there are finitely many. But the argument is very similar, if this were not the case, one would obtain a sequence $\Lambda_n \to \Gamma$ of hyperbolic attractors. One can take the measures μ_n with Lyapunov exponents adding ≤ 0 and obtain a similar measure μ in Γ. This allows to show that Γ admits a dominated splitting (one needs to use a stronger version of [BoC] which deals with classes which are not necessarily quasi-attractors) and therefore is uniformly hyperbolic according to [PuS]. This is a contradiction and one obtains finiteness.

[18]The number of sinks of a diffeomorphism is a semicontinuous function to the natural numbers and ∞. Therefore, it is continuous (and therefore locally constant) in a residual subset.

[19]There could be a zero Lyapunov exponent, but again using Mañé's ergodic closing lemma and Franks Lemma ([AbBC]) one would create a sink by small perturbation.

The fact that the basin is open and dense is direct from the fact that the basins of hyperbolic attractors is open and Theorem 6.1. To show that the basins cover a full Lebesgue measure subset one has to use a semicontinuity argument on the size of basins and use the fact that for C^2-diffeomorphisms hyperbolic sets have zero Lebesgue measure. See [Ara] or [San] for details.

6.3. **Partially hyperbolic attractors with one dimensional center.** We explain here part of a work in progress joint with S. Crovisier and M. Sambarino which studies the geometry of partially hyperbolic sets saturated by strong unstable manifolds. This study is motivated by the fact that attractors are saturated by strong unstable manifolds.

Together with recent results of [CrPuS] and [CrSY] (which use completely different techniques) our main result gives as a consequence the following result which is a step towards the understanding of dynamics far from homoclinic tangencies. It also improves (in dimension 3) a result announced[20] in [BoGLY] (though their result holds in any dimension).

Theorem 6.4 ([CrPoS], [CrPuS], [CrSY]). *Let M be a 3-dimensional manifold. Then, there exists an open and dense subset $\mathcal{U}$ of $\mathrm{Diff}^1(M)$ such that if $f \in \mathcal{U}$ then:*

- *either f has robustly finitely many quasi-attractors,*
- *or f can be C^1-approximated by a diffeomorphism g which has a hyperbolic periodic point p whose stable and unstable manifolds intersect non-transversally (i.e. g has a* homoclinic tangency*).*

Let us call $\mathrm{HT}^1(M)$ the set of diffeomorphisms of M admitting a homoclinic tangency. Putting together the results of [CrPuS] and [CrSY] one can show[21] the following:

Theorem 6.5 (Crovisier,Pujals,Sambarino,D.Yang). *There exists a residual subset $\mathcal{G}_{CPSY}$ of $\mathrm{Diff}^1(M) \setminus \overline{\mathrm{HT}^1(M)}$ such that if $f \in \mathcal{G}_{CPSY}$ one has the following property:*

- *there exists a* filtration $\emptyset = U_0 \subset U_1 \subset \ldots \subset U_{k-1} \subset U_k = M$ *of open subsets such that $f(\overline{U_i}) \subset U_i$ and such that for every i, if $\Lambda_i = \bigcap_{n\in\mathbb{Z}} f^n(\overline{U_i} \setminus U_{i-1})$ then Λ_i verifies one of the following three possibilities:*
  - *Λ_i is a sink,*
  - *Λ_i is a source, or,*
  - *Λ_i admits a strongly partially hyperbolic splitting $T_{\Lambda_i}M = E^s \oplus E^c \oplus E^u$ where both E^s is uniformly contracted and non-zero, E^u is uniformly expanded and non-zero and E^c admits a subdominated splitting into one-dimensional bundles.*

The advantage of working in dimension 3 is that we always know that the dimension of E^c is at most 1. The strategy of the proof is showing that each Λ_i can contain at most finitely many quasi-attractors, so, we are reduced to showing that in a compact f-invariant subset with a strong partially hyperbolic splitting with

[20]In [BoGLY] they show that there exists a residual subset $\mathcal{G}$ of diffeomorphisms far away from homoclinic tangencies such that if $f \in \mathcal{G}$ then all quasi-attractors of f are isolated from each other (but might in principle accumulate in a set which is not a quasi-attractor). They call *essential attractors* to such quasi-attractors since it can be shown that their basin contains a residual subset of a neighborhood.

[21]These result also rely on other results, see [Cr$_3$] for a more complete account.

$\dim E^c = 1$ one can have at most finitely many quasi-attractors. This is also a consequence of the results of [CrPoS] and what we shall try to briefly explain in what follows.

6.3.1. *Minimal $\mathcal{W}^{uu}$-saturated sets.* Let $f : M \to M$ be a C^1-diffeomorphism. Consider a set Λ which is of the form $\Lambda = \bigcap_{n\in\mathbb{Z}} f^n(\overline{U}\setminus V)$ where U and V are open subsets of M verifying that $f(\overline{U}) \subset U$ and $f(\overline{V}) \subset V$.

The open set $U\setminus\overline{V}$ is a neighborhood of Λ and we know that if a point $x \in U\setminus\overline{V}$ verifies that $f(x) \notin U\setminus\overline{V}$ (resp. $f^{-1}(x)$) then $f^n(x) \notin U\setminus\overline{V}$ for all $n \geq 1$ (resp. $n \leq -1$). This allows one to prove the following.

Exercise. Show that if a quasi-attractor Q intersects $U\setminus\overline{V}$ then $Q \subset \Lambda$.

We shall assume moreover that Λ admits a partially hyperbolic splitting $T_\Lambda M = E^s \oplus E^c \oplus E^u$ where both E^s and E^u are non-zero and such that $\dim E^c = 1$. Our goal is to show that there are finitely many quasi-attractors in Λ. Recall that quasi-attractors are $\mathcal{W}^{uu}$-saturated.

Exercise. Use Theorem 5.5 to show that if $\{Q_n\}$ is a sequence of quasi-attractors in Λ converging to Θ in the Hausdorff topology, then $\Theta \subset \Lambda$ is $\mathcal{W}^{uu}$-saturated.

We say that a (non-empty) compact f-invariant and $\mathcal{W}^{uu}$-saturated subset Γ of Λ is a *minimal $\mathcal{W}^{uu}$-saturated set* if for every Γ' strictly contained in Γ which is compact f-invariant and $\mathcal{W}^{uu}$-saturated one has that $\Gamma' = \emptyset$. That is, a subset $\Gamma \subset \Lambda$ is a minimal $\mathcal{W}^{uu}$-saturated set if it is minimal for being compact, f-invariant and $\mathcal{W}^{uu}$-saturated.

Exercise. Show that if Λ contains a non-empty compact $\mathcal{W}^{uu}$-saturated subset Λ', then there exist minimal $\mathcal{W}^{uu}$-saturated sets. Moreover, show that every quasi-attractor $Q \subset \Lambda$ contains at least one minimal $\mathcal{W}^{uu}$-saturated set.

Notice that if Q and Q' are two different quasi-attractors then $Q\cap Q' = \emptyset$. Therefore, there are fewer quasi-attractors in Λ than there are minimal $\mathcal{W}^{uu}$-saturated sets. The main result on [CrPoS] is the following.

Theorem 6.6 ([CrPoS])**.** *There is an open and dense subset $\mathcal{O}$ of $\mathrm{Diff}^1(M)$ such that if $f \in \mathcal{O}$ and Λ a compact f-invariant set admitting a strong partially hyperbolic splitting $T_\Lambda M = E^s \oplus E^c \oplus E^u$ with $\dim E^c = 1$ then Λ contains at most finitely many $\mathcal{W}^{uu}$-saturated sets.*

The proof of this Theorem has two stages. First, a perturbation result which provides a geometric property of $\mathcal{W}^{uu}$-saturated laminations for diffeomorphisms in a C^1-open and dense subset of diffeomorphisms. The second stage is to show that this geometric property forbids the minimal $\mathcal{W}^{uu}$-saturated sets to accumulate and since (unlike quasi-attractors) minimal $\mathcal{W}^{uu}$-saturated sets are closed in the Hausdorff topology, this concludes.

6.3.2. *Geometry of strong connections.* We start by explaining the consequence of our perturbation result. The proof of this result is the most delicate part of [CrPoS] and it is the part which we shall omit in this notes. It is not (only) because of laziness but because the techniques are farther away from the interests of this notes.

The statement is the following:

Theorem 6.7. *There exists a G_δ-dense subset $\mathcal{G}$ of* $\mathrm{Diff}^1(M)$ *such that for every $f \in \mathcal{G}$ and $\Lambda' \subset M$ a compact f-invariant partially hyperbolic set which is $\mathcal{W}^{uu}$-saturated and for every $r, r't, \gamma > 0$ sufficiently small, there exists $\delta > 0$ with the following property.*

If $x, y \in \Lambda'$ satisfy $y \in \mathcal{W}^{ss}(x)$ and $d_s(x,y) \in (r, r')$, then there is $x' \in \mathcal{W}^{uu}_t(x)$ such that:

$$d(\mathcal{W}^{ss}_\gamma(x'), \mathcal{W}^{uu}_\gamma(y)) > \delta$$

By d_s we refer to the distance inside $\mathcal{W}^{ss}$ and $\mathcal{W}^\sigma_\varepsilon(x)$ $(\sigma = ss, uu)$ to denote the ε-ball around x in $\mathcal{W}^\sigma(x)$ with the intrinsic metric. We could have used $\mathcal{W}^\sigma_{loc}(x)$ in each of the places, but the way we have formulated is a bit more explicit.

Using the continuity of the strong manifolds with respect to the diffeomorphism (see Remark 5.6), one sees that at a given scale (i.e. if one fixes the values of r, r', t and γ), this property holds for small perturbations of $f \in \mathcal{G}$ and therefore, in an open and dense subset of $\mathrm{Diff}^1(M)$.

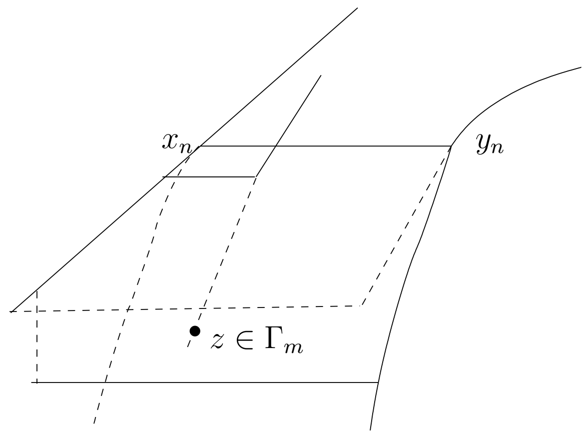


FIGURE 4. The stable manifolds of the minimal sets must intersect.

6.3.3. *Finiteness of minimal saturated sets.* Now, we use Theorem 6.7 as well as the results in the previous sections of this notes to conclude the proof of Theorem 6.6.

Let $\Lambda' \subset \Lambda$ be a compact f-invariant and $\mathcal{W}^{uu}$-saturated set. We must show that there are finitely many minimal $\mathcal{W}^{uu}$-saturated sets in Λ'.

Exercise. Show that if Λ has infinitely many minimal $\mathcal{W}^{uu}$-saturated sets, then there exists $\Lambda' \subset \Lambda$ compact, f-invariant and $\mathcal{W}^{uu}$-saturated containing infinitely many minimal $\mathcal{W}^{uu}$-saturated sets.

The following remark will be important in the proof.

Exercise. Show that if $\Gamma, \Gamma' \subset \Lambda'$ are different minimal $\mathcal{W}^{uu}$-saturated sets, then their stable manifolds are disjoint. That is, if Γ and Γ' are minimal $\mathcal{W}^{uu}$-saturated sets and there exists $x \in \Gamma$ such that there exists $y \in \Gamma'$ such that $d(f^n(x), f^n(y)) \to 0$ as $n \to +\infty$ then $\Gamma = \Gamma'$.

First, we shall show the following.

Proposition 6.8. *In Λ' there are at most finitely many minimal $\mathcal{W}^{uu}$-saturated sets Γ with the property that for every $x \in \Gamma$ one has that $\mathcal{W}^{ss}(x) \cap \Gamma = \{x\}$.*

Proof. Assume by contradiction that there are infinitely many such subsets and denote them as $\{\Gamma_n\}_n$. Notice that thanks to Theorem 5.9 we know that for each n one has that Γ_n is contained in a locally invariant submanifold Σ_n tangent to $E^c \oplus E^u$ at each point of Γ_n.

Notice moreover that there exists $h > 0$ such that $h_{top}(f|_{\Gamma_n}) > h$ for every n. This follows from the following argument: consider a finite covering of Γ_n by balls of radius ε where ε is small enough (independent on n) so that any disk tangent to a small cone around E^u of diameter 1 contains at least two disks of radius ε contained in different balls of the covering. Now, we know that given any disk tangent to a small cone around E^u its iterates grow so that the internal radius multiplies by an uniform amount (independent of n). We can choose such a disk D to be contained in Γ_n (since it is $\mathcal{W}^{uu}$-saturated). We get that for some k_0 (independent of n), the image $f^{k_0}(D)$ contains two such disks. Therefore, inside D one has that in k_0 iterates we duplicate the number of "different" orbits and therefore the entropy of f in Γ_n is larger than $\frac{1}{k_0}\log 2$ (independent of n).

Now, using the variational principle and Ruelle's inequality (Theorem 4.3) for f^{-1} we obtain that Γ_n has a measure μ_n whose Lyapunov exponent for f^{-1} along E^c (recall that Γ_n "lives" in Σ_n) is larger than h. This means that Γ_n has points whose stable manifold has uniform size[22] along $E^s \oplus E^c$. This implies that for every n, there is an open ball B_n of uniform volume such that no other Γ_m can intersect for $m \neq n$. This is impossible if there are infinitely many Γ_n.

□

Now we are in conditions to complete the proof of Theorem 6.6. Consider $f \in \mathcal{G}$ given by Theorem 6.7 (or in a small neighborhood so that the same properties hold at a given scale).

Assume by contradiction that there are infinitely many different minimal $\mathcal{W}^{uu}$-saturated sets $\{\Gamma_n\}_n$ in Λ'. By Proposition 6.8 we can assume that for every n there exists $x_n \in \Gamma_n$ such that $\mathcal{W}^{ss}(x_n) \cap \Gamma_n \neq \{x_n\}$.

By iteration, we can assume that we have points $x_n, y_n \in \Gamma_n$ such that $y_n \in \mathcal{W}^{ss}(x_n)$ and $d_s(x_n, y_n) \in (r, r')$ for some $r' > \Delta r$ where $\Delta \geq \max_x\{\|D_x f^{\pm 1}\|\}$.

Then, these pairs of points converge to points $x, y \in \Lambda'$ which belong to the same local stable manifold ($y \in \mathcal{W}^{ss}(x)$ and $d_s(x,y) \in (r, r')$). Since the strong unstable manifolds get separated by projection by stable holonomy, it is possible to show[23] that this configuration forces the strong stable manifold of one of the Γ_n to intersect some other Γ_m (see Figure 4) contradicting the fact that the minimal sets where different. This concludes.

References

[AbBC] F. Abdenur, C. Bonatti, S.Crovisier, Non-uniform hyperbolicity for C^1-generic diffeomorphisms, *Israel J. of Math.* **183** (2011), 1-60.

[Ara] A. Araujo, Existencia de atratores hiperbolicos para difeomorfismos de superficies, Ph.D. Thesis, IMPA, (1987).

---

[22] We have not proved this explicitely, but it follows from the arguments we have done along the notes.

[23] Details to appear soon in [CrPoS].

[Arn] M. C. Arnaud, Hyperbolicity for conservative twist maps of the 2 - dimensional annulus , *Publicaciones Matemáticas del Uruguay* **16** (2016), 1-39.

[AvB] A. Avila, J. Bochi, Trieste lecture notes on Lyapunov exponents, (2008) available at http://www.mat.uc.cl/ jairo.bochi/docs/trieste.pdf

[AvB$_2$] A. Avila, J. Bochi, On the subaditive ergodic theorem, *unpublished* (2009) available at http://www.mat.uc.cl/ jairo.bochi/docs/kingbirk.pdf

[AvCW] A. Avila, S. Crovisier, A. Wilkinson, Diffeomorphisms with positive metric entropy,*Preprint* arXiv:1408.4252.

[BaP] L. Barreira, Y. Pesin, *Nonuniform hyperbolicity*, Cambridge University Press (2007).

[BeC] M. Benedicks and L. Carleson. The dynamics of the Hénon map. *Ann. Math.* **133** (1991) 73Â–169.

[Ber] P. Berger, Persistence of laminations, *Bull Braz Math Soc, New series* **41** 2, (2010) p. 259-319.

[Ber$_2$] P. Berger, Abundance of non-uniformly hyperbolic Hénon like endomorphisms, *Preprint* arXiv:0903.1473.

[BerB] P. Berger, A. Bounemoura, A geometrical proof of the persistence of normally hyperbolic submanifolds, *Dynamical systems, An International Journal* **28** 4 (2013).

[BeKZ] P. Bernard, V. Kaloshin, K. Zhang, Arnold diffusion in arbitrary degrees of freedom and crumpled 3-dimensional normally hyperbolic invariant cylinders, *Preprint* arXiv:1112.2773

[Bo] J. Bochi, The multiplicative ergodic theorem, *unpublished* (2008) available at http://www.mat.uc.cl/ jairo.bochi/docs/oseledets.pdf

[BoGo] J. Bochi and N. Gourmelon, Some characterizations of domination, *Mathematische Zeitschrift* **263** (2009), no. 1, pp. 221-231.

[BoC] C. Bonatti, S. Crovisier, Récurrence et généricité, *Inventiones Math.* **158** (2004), 33Â–104.

[BoC2] C. Bonatti, S. Crovisier, Center manifolds for partially hyperbolic set without strong unstable connections, to appear in *Journal of the IMJ.*

[BoCS] C. Bonatti, S. Crovisier, K. Shinohara, The $C^{1+\alpha}$ hypothesis in Pesin theory revisited, *Journal of Modern Dynamics* **7** (2013), 605-618.

[BoDP] C. Bonatti, L. Díaz and E. Pujals, A C^1-generic dichotomy for diffeomorphisms: Weak forms of hyperbolicity or infinitely many sinks or sources, *Annals of Mathematics* **158**(2), (2003) 355-418.

[BoDV] C. Bonatti, L. Díaz and M. Viana, *Dynamics Beyond Uniform Hyperbolicity. A global geometric and probabilistic perspective*, Encyclopaedia of Mathematical Sciences **102**. Mathematical Physics III. Springer-Verlag (2005).

[BoGLY] C. Bonatti, S. Gan, M. Li and D. Yang, *In preparation.*

[BoLY] C. Bonatti, M. Li, D. Yang, On the existence of attractors, *Transactions of the AMS* **365** 3 (2013) 1369Â–1391.

[Br] M. Brin, On dynamical coherence, *Ergodic Theory and Dynamical Systems* **23** (2003) 395–401.

[BrP] M. Brin, Y. Pesin, Partially hyperbolic dynamical systems, *Izv. Akad. Nauk SSSR Ser. Mat.* **38** (1974), 170Â–212.

[BuI] D. Burago and S. Ivanov, Partially hyperbolic diffeomorphisms of 3-manifolds with abelian fundamental groups. *Journal of Modern Dynamics* **2** (2008) 541Â–580.

[BuW] K.Burns and A. Wilkinson, Dynamical coherence and center bunching, *Discrete and Continuous Dynamical Systems A* (Pesin birthday issue) **22** (2008) 89-100.

[BuW$_2$] K. Burns and A. Wilkinson, On the ergodicity of partially hyperbolic systems, *Ann. of Math.* **171** (2010) 451-489.

[Cr] S. Crovisier, Birth of homoclinic intersections: a model for the central dynamics of partially hyperbolic systems, *Annals of Math.* **172** (2010), 1641-1677.

[Cr$_2$] S. Crovisier, Dynamics of surface diffeomorphisms, *Course Draft Notes* (2013).

[Cr$_3$] S. Crovisier, Dynamics of C^1-diffeomorphisms: global description and prospects for classification, to appear in Proceedings of the ICM (2014) arXiv:1405.0305.

[CrPo] S. Crovisier, R. Potrie, Introduction to partially hyperbolic dynamics, *Lecture Notes* for a minicourse at ICTP School and Workshop on Dynamical Systems (2015) available in the web page of the conference.

[CrPoS] S. Crovisier, R. Potrie and M. Sambarino, Geometric properties of partially hyperbolic attractors, *In preparation.*

[CrPu] S. Crovisier, E. Pujals, *In preparation.*

[CrPuS] S. Crovisier, E. Pujals, M. Sambarino, Hyperbolicity of extreme bundles, *In preparation.*

[CrSY] S. Crovisier, M. Sambarino, D. Yang, Partial hyperbolicity and homoclinic tangencies, to appear in *Journal of the EMS* arXiv:1103.0869.

[DeM] G. Del Magno, R. Markarian, On the Bernoulli property of planar hyperbolic billiards, *Preprint* http://premat.fing.edu.uy/ 2014/177.

[Ge] K. Gelfert, Horseshoes for diffeomorphisms preserving hyperbolic measures, to appear in *Math. Z.*

[Gou] N. Gourmelon, Adapted metrics for dominated splittings, *Ergodic Theory and Dynamical Systems* **27** (2007), pp 1839-1849.

[GK] M. Guardia, V. Kaloshin, Orbits of nearly integrable systems accumulating to KAM tori, *Preprint* arXiv:1412.7088.

[HP] A. Hammerlindl and R. Potrie, Pointwise partial hyperbolicity in 3-dimensional nilmanifolds, *Journal of the London Math. Society* **89** (3) (2014): 853-875.

[HP_2] A. Hammerlindl and R. Potrie, Partial hyperbolicity in 3-dimensional manifolds with solvable fundamental group, *Journal of Topology* **8** (3), (2015) 842-870 .

[HPS] M. Hirsch, C. Pugh and M. Shub, Invariant Manifolds, *Springer Lecture Notes in Math.*, **583** (1977).

[KZ] V. Kaloshin, K.Zhang, Normally hyperbolic invariant manifolds near strong double resonance, *Preprint* arXiv:1202.1032

[KH] A. Katok and B. Hasselblatt, *Introduction to the modern theory of dynamical systems*, Cambridge University Press (1995).

[LeR] F. Le Roux, L'ensemble de rotation autour d'un point fixe d'homÃ©omorphisme de surface, *Astérisque* **350** (2013).

[M_1] R. Mañé, Contributions to the stability conjecture, *Topology* **17** (1978), 383–396.

[M_2] R. Mañé, A proof of Pesin's entropy formula, *Ergodic Theory and Dynamical Systems* **1** (1981) 95-102.

[M_3] R. Mañé, An ergodic closing lemma, *Annals of Math.* **116** (1982), 503–540.

[M_4] R. Mañé, *Ergodic theory and differentiable dynamics*, Springer-Verlag (1983).

[Pot] R. Potrie, A proof of the existence of attractors in dimension 2, *Preprint* (2009). Not intended for publication, available in my web page.

[Pot_2] R. Potrie, Wild Milnor attractors accumulated by lower dimensional dynamics, *Ergodic Theory and Dynamical Systems* **31** 1 (2014) 236-262.

[Pot_3] R. Potrie, On the work of Jorge Lewowicz on expansive systems, *Publicaciones Matematicas del Uruguay* **14** (2014) 6-22.

[Pot_4] R. Potrie, Partial hyperbolicity and attracting regions in 3-dimensional manifolds, Thesis (2012) arXiv:1207.1822.

[Pot_5] R. Potrie, Partial hyperbolicity and foliations in $\mathbb{T}^3$, *Journal of Modern Dynamics* **9** (2015) 81–121.

[PuRH] E. Pujals, F. Rodriguez Hertz, Critical points for surface diffeomorphisms, *Journal of Modern Dynamics* **1** (4) (2007) 615–648.

[PuS] E. Pujals, M. Sambarino, Homoclinic tangencies and hyperbolicity for surface diffeomorphisms, *Annals of Math.* **151** (2000), 961Â–1023.

[San] B. Santiago, Hyperbolicidade essencial em superficies, Dissertacao de Mestrado UFRJ (2011).

[Sar] O. Sarig, Symbolic dynamics for surface diffeomorphisms with positive entropy, *Journal of the AMS* **26** (2013) 341–426.

[Sh] M. Shub, *Global stability of dynamical systems*, Springer-Verlag (1987).

[Va] F. Valenzuela, On critical points for two dimensional holomorphic systems, to appear in *Ergodic Theory and Dynamical Systems.*

CMAT, Facultad de Ciencias
Universidad de la República, Uruguay
E-mail address: `rpotrie@cmat.edu.uy`
URL: `www.cmat.edu.uy/~rpotrie`

Publicaciones Matemáticas del Uruguay
Volumen 16, Julio 2016, Páginas 169–192
S 0797-1443

LECTURE NOTES ON MATHER'S THEORY FOR LAGRANGIAN SYSTEMS

ALFONSO SORRENTINO

ABSTRACT. These notes are based on a series of lectures that the author gave at the CIMPA Research School *Hamiltonian and Lagrangian Dynamics*, which was held in Salto (Uruguay) in March 2015.

To the memory of Ricardo Mañé
(1948 – 1995)

CONTENTS

1. INTRODUCTION

In these lecture notes we provide a brief introduction to John Mather's variational approach to the study of convex and superlinear Hamiltonian systems, what is generally called *Aubry-Mather theory*. Starting from the observation that invariant Lagrangian graphs can be characterized in terms of their "action-minimizing" properties, we then describe how analogue features can be traced in a more general setting, namely the so-called *Tonelli Hamiltonian systems*. This approach brings to light a plethora of compact invariant subsets for the system, which, under many points of view, can be considered as generalization of invariant Lagrangian graphs, despite not being in general either submanifolds or regular. Besides being very significant from a dinamical systems point of view, these objects also appear and play an important role in many other different contexts: PDEs (e.g., Hamilton-Jacobi equation and weak KAM theory), Symplectic geometry, etc...

Since this notes[1] are meant to be a short introduction and a guide to this theory, we will omit most of the proofs. We refer interested readers to [23] for a more systematic and comprehensive presentation of this and other topics.

Acknowledgement. The author wishes to thank CIMPA and, more specifically, Ezequiel Maderna and Ludovic Rifford, for the organization of this very interesting meeting and for their kind invitation. The author also wishes to thank Princeton University Press for agreeing on the use herein of some of the material from [23].

2. From KAM theory to Aubry-Mather (AM) theory

The celebrated Kolmogorov-Arnol′d -Moser (or KAM) theorem finally settled the old question concerning the existence of *quasi-periodic* motions for *nearly-integrable* Hamiltonian systems, *i.e.,* Hamiltonian systems that are slight perturbation of an integrable one. In the integrable case, in fact, the whole phase space is foliated by invariant Lagrangian submanifolds that are diffeomorphic to tori, and on which the dynamics is conjugate to a rigid rotation. More specifically, let $H : T^*\mathbb{T}^n \longrightarrow \mathbb{R}$ be an integrable Tonelli Hamiltonian in action-angle coordinates, *i.e.*, $H(x,p) = \mathfrak{h}(p)$ with the Hamiltonian depending only on the action variables (see [2])[2]. Let us denote by $\phi_t^{\mathfrak{h}}$ the associated Hamiltonian flow and identify $T^*\mathbb{T}^n$ with $\mathbb{T}^n \times \mathbb{R}^n$, where $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$.

The Hamiltonian flow in this case is very easy to study. Hamilton's equations are:

$$\begin{cases} \dot{x} = \frac{\partial \mathfrak{h}}{\partial p}(p) =: \rho(p) \\ \dot{p} = -\frac{\partial \mathfrak{h}}{\partial x}(p) = 0, \end{cases}$$

therefore $\Phi_t^{\mathfrak{h}}(x_0, p_0) = (x_0 + t\rho(p_0) \bmod \mathbb{Z}^n,\, p_0)$. In particular, p is an integral of motion, that is, it remains constant along the orbits. The phase space $T^*\mathbb{T}^n$ is hence foliated by invariant tori $\Lambda^*_{p_0} = \mathbb{T}^n \times \{p_0\}$ on which the motion is a rigid rotation with rotation vector $\rho(p_0)$ (see figure 1).

On the other hand, it is natural to ask what happens to such a foliation and to these stable motions once the system is perturbed. In 1954 Kolmogorov [11] — and later Arnol′d [1] and Moser [21] in different contexts — proved that, in spite of the generic disappearance of the invariant tori filled by periodic orbits (already pointed out by Henri Poincaré), for small perturbations of an integrable system it is still possible to find invariant Lagrangian tori corresponding to certain rotation vectors (the so-called *diophantine* rotation vectors). This result is commonly referred to as *KAM theorem*, from the initials of the three main pioneers. In addition to open the way to a new understanding of the nature of Hamiltonian systems and their stable motions, this result contributed to raise new interesting questions, such as: what does it happen to the stable motions that are destroyed by effect of the perturbation? Is it possible to identify something reminiscent of their past presence? What can be said for systems that not close to an integrable one?

---

[1]Portions of this material used with permission from Princeton University Press from "*Action-minimizing Methods in Hamiltonian Dynamics: An Introduction to Aubry-Mather Theory*" by Alfonso Sorrentino, 2015 (see [23]).

[2]In general these coordinates can be defined only locally. For the sake of simplicity, in this example we assume — without affecting its main purpose — that they are defined globally.

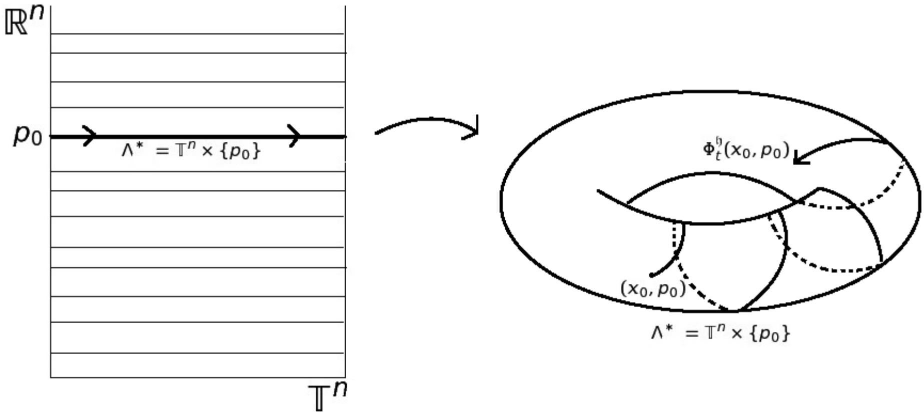


FIGURE 1. The phase space of an integrable system.

Aubry-Mather theory provides answers to these questions. Developed independently by Serge Aubry [3] and John Mather [14] in 1980s, this novel approach to the study of the dynamics of *twist diffeomorphisms of the annulus* (which correspond to Poincaré maps of 1-dimensional non-autonomous Hamiltonian systems) pointed out the existence of many invariant sets, which are obtained by means of variational methods and that always exist, even after rotational curves are destroyed. Besides providing a detailed structure theory for these new sets, this powerful approach yielded to a better understanding of the destiny of invariant rotational curves and to the construction of interesting chaotic orbits as a result of their destruction [15, 17].

Motivated by these achievements, John Mather [18, 19] — and later Ricardo Mañé [13, 12] and Albert Fathi [9] in different ways — developed a generalization of this theory to higher dimensional systems. Positive definite superlinear Lagrangians on compact manifolds, also called *Tonelli Lagrangians* (see Definition 3.1), were the appropriate setting to work in. Under these conditions, in fact, it is possible to prove the existence of interesting invariant sets, known as *Mather*, *Aubry* and *Mañé* sets, which generalize KAM tori and invariant Lagrangian graphs, and which continue to exist beyond the nearly-integrable case.

In the following we will provide a brief overview of Mather's theory. We will first discuss an illustrative example (what happens in the integrable case) and then show how similar ideas can be extended to a more general setting.

3. Tonelli Lagrangians and Hamiltonians on compact manifolds

Before starting, let us introduce the basic setting that we will consider in the following. Let M be a compact and connected smooth manifold without boundary. Denote by TM its tangent bundle and T^*M the cotangent one. A point of TM will be denoted by (x, v), where $x \in M$ and $v \in T_xM$, and a point of T^*M by (x, p), where $p \in T_x^*M$ is a linear form on the vector space T_xM. Let us fix a Riemannian metric g on it and denote by d the induced metric on M; let $\|\cdot\|_x$ be the norm induced by g on T_xM; we will use the same notation for the norm induced on T_x^*M.

We will consider functions $L : TM \longrightarrow \mathbb{R}$ of class C^2, which are called *Lagrangians*. Associated to each Lagrangian, there is a flow on TM called the *Euler-Lagrange flow*, defined as follows. Let us consider the action functional A_L from the space of absolutely continuous curves $\gamma : [a,b] \to M$, with $a \le b$, defined by:

$$A_L(\gamma) := \int_a^b L(\gamma(t), \dot\gamma(t))\, dt.$$

Curves that extremize[3] this functional among all curves with the same end-points (and the same time-length) are solutions of the *Euler-Lagrange equation*:

$$\frac{d}{dt}\frac{\partial L}{\partial v}(\gamma(t),\dot\gamma(t)) = \frac{\partial L}{\partial x}(\gamma(t),\dot\gamma(t)) \qquad \forall\, t \in [a,b]\,.$$

Observe that this equation is equivalent to

$$\frac{\partial^2 L}{\partial v^2}(\gamma(t),\dot\gamma(t))\ddot\gamma(t) = \frac{\partial L}{\partial x}(\gamma(t),\dot\gamma(t)) - \frac{\partial^2 L}{\partial v \partial x}(\gamma(t),\dot\gamma(t))\dot\gamma(t)\,,$$

therefore, if the second partial vertical derivative $\partial^2 L/\partial v^2(x,v)$ is non-degenerate at all points of TM, we can solve for $\ddot\gamma(t)$. This condition

$$\det \frac{\partial^2 L}{\partial v^2} \neq 0$$

is called *Legendre condition* and allows one to define a vector field X_L on TM, such that the solutions of $\ddot\gamma(t) = X_L(\gamma(t),\dot\gamma(t))$ are precisely the curves satisfying the Euler-Lagrange equation. This vector field X_L is called the *Euler-Lagrange vector field* and its flow Φ_t^L is the *Euler-Lagrange flow* associated to L. It turns out that Φ_t^L is C^1 even if L is only C^2 (see Remark 3.3).

Definition 3.1 (Tonelli Lagrangian). *A function $L : TM \longrightarrow \mathbb{R}$ is called a Tonelli Lagrangian if:*

i) *$L \in C^2(TM)$;*

ii) *L is strictly convex in the fibers, in the C^2 sense, i.e., the second partial vertical derivative $\partial^2 L/\partial v^2(x,v)$ is positive definite, as a quadratic form, for all (x,v);*

iii) *L is superlinear in each fiber, i.e.,*

$$\lim_{\|v\|_x \to +\infty} \frac{L(x,v)}{\|v\|_x} = +\infty.$$

This condition is equivalent to ask that for each $A \in \mathbb{R}$ there exists $B(A) \in \mathbb{R}$ such that

$$L(x,v) \ge A\|v\| - B(A) \qquad \forall\,(x,v) \in TM\,.$$

Observe that since the manifold is compact, then condition *iii)* is independent of the choice of the Riemannian metric g.

[3]These extremals are not in general minima. The existence of global minima and the study of the corresponding motions is the core of Aubry-Mather theory; see section 5.

Examples of Tonelli Lagrangians.

- **Riemannian Lagrangians.** Given a Riemannian metric g on TM, the *Riemannian Lagrangian* on (M,g) is given by the *kinetic energy*:
$$L(x,v) = \frac{1}{2}\|v\|_x^2 .$$
Its Euler-Lagrange equation is the equation of the geodesics of g:
$$\frac{D}{dt}\dot{x} \equiv 0 ,$$
and its Euler-Lagrange flow coincides with the geodesic flow.
- **Mechanical Lagrangians.** These Lagrangians play a key-role in the study of classical mechanics. They are given by the sum of the kinetic energy and a *potential* $U : M \longrightarrow \mathbb{R}$:
$$L(x,v) = \frac{1}{2}\|v\|_x^2 + U(x) .$$
The associated Euler-Lagrange equation is given by:
$$\frac{D}{dt}\dot{x} = \nabla U(x) .$$
- **Mañé's Lagrangians.** This is a particular class of Tonelli Lagrangians, introduced by Ricardo Mañé in [12]. If X is a C^k vector field on M, with $k \geq 2$, one can embed its flow φ_t^X into the Euler-Lagrange flow associated to a certain Lagrangian, namely
$$L_X(x,v) = \frac{1}{2}\|v - X(x)\|_x^2 .$$
It is quite easy to check that the integral curves of the vector field X are solutions of the Euler-Lagrange equation. In particular, the Euler-Lagrange flow $\Phi_t^{L_X}$ restricted to $\mathrm{Graph}(X) = \{(x, X(x)),\ x \in M\}$ (which is clearly invariant) is conjugate to the flow of X on M and the conjugacy is given by $\pi|\mathrm{Graph}(X)$, where $\pi : TM \to M$ is the canonical projection. In other words, the following diagram commutes:
$$\begin{tikzcd} \mathrm{Graph}(X) \arrow[r, "\Phi_t^{L_X}"] \arrow[d, "\pi"'] & \mathrm{Graph}(X) \arrow[d, "\pi"] \\ M \arrow[r, "\varphi_t^X"'] & M \end{tikzcd}$$
that is, for every $x \in M$ and every $t \in \mathbb{R}$, $\Phi_t^{L_X}(x, X(x)) = (\gamma_x^X(t), \dot{\gamma}_x^X(t))$, where $\gamma_x^X(t) = \varphi_t^X(x)$.

In the study of classical dynamics it turns often very useful to consider the associated *Hamiltonian system*, which is defined on the cotangent bundle T^*M. Given a Lagrangian L we can define the associated *Hamiltonian* as its *Fenchel transform* (or *Legendre-Fenchel transform*), see [22]:
$$\begin{array}{rccl} H: & T^*M & \longrightarrow & \mathbb{R} \\ & (x,p) & \longmapsto & \sup_{v \in T_xM} \{\langle p,\, v\rangle_x - L(x,v)\} \end{array}$$

where $\langle \cdot, \cdot \rangle_x$ denotes the canonical pairing between the tangent and cotangent bundles.

If L is a Tonelli Lagrangian, one can easily prove that H is finite everywhere (as a consequence of the superlinearity of L), superlinear and strictly convex in each fiber (in the C^2 sense). Observe that H is also C^2. In fact the Euler-Lagrange vector field corresponds, under the Legendre transformation, to a vector field on T^*M given by Hamilton's equation; it is easily seen that this vector field is C^1 (see [6, p. 207]). Such a Hamiltonian is called a *Tonelli* (or *optical*) *Hamiltonian.*

Definition 3.2 (Tonelli Hamiltonian). *A function $H : T^*M \longrightarrow \mathbb{R}$ is called a Tonelli (or optical) Hamiltonian if:*

i) *H is of class C^2;*
ii) *H is strictly convex in each fiber in the C^2 sense, i.e., the second partial vertical derivative $\partial^2 H/\partial p^2(x,p)$ is positive definite, as a quadratic form, for any $(x,p) \in T^*M$;*
iii) *H is superlinear in each fiber, i.e.,*

$$\lim_{\|p\|_x \to +\infty} \frac{H(x,p)}{\|p\|_x} = +\infty\,.$$

Examples of Tonelli Hamiltonians.
Let us see what are the Hamiltonians associated to the Tonelli Lagrangians that we have introduced in the previous examples.

- **Riemannian Hamiltonians.** If $L(x,v) = \frac{1}{2}\|v\|_x^2$ is the Riemannian Lagrangian associated to a Riemannian metric g on M, the corresponding Hamiltonian will be
$$H(x,p) = \frac{1}{2}\|p\|_x^2,$$
where $\|\cdot\|$ represents — in this last expression — the induced norm on the cotangent bundle T^*M.
- **Mechanical Hamiltonians.** If $L(x,v) = \frac{1}{2}\|v\|_x^2 + U(x)$ is a mechanical Lagrangian, the associated Hamiltonian is:
$$H(x,p) = \frac{1}{2}\|p\|_x^2 - U(x).$$
It is sometimes referred to as *mechanical energy.*
- **Mañé's Hamiltonians.** If X is a C^k vector field on M, with $k \geq 2$, and $L_X(x,v) = \|v - X(x)\|_x^2$ is the associated Mañé Lagrangian, one can check that the corresponding Hamiltonian is given by:
$$H(x,p) = \frac{1}{2}\|p\|_x^2 + \langle p, X(x)\rangle\,.$$

Given a Hamiltonian one can consider the associated *Hamiltonian flow* Φ_t^H on T^*M. In local coordinates, this flow can be expressed in terms of the so-called *Hamilton's equations*:

$$\begin{cases} \dot{x}(t) = \frac{\partial H}{\partial p}(x(t),p(t)) \\ \dot{p}(t) = -\frac{\partial H}{\partial x}(x(t),p(t))\,. \end{cases}$$

We will denote by $X_H(x,p) := \left(\frac{\partial H}{\partial p}(x,p), -\frac{\partial H}{\partial x}(x,p)\right)$ the *Hamiltonian vector field* associated to H. This has a more intrinsic (geometric) definition in terms of the canonical symplectic structure ω on T^*M, which in local coordinates can be written as $dx \wedge dp$ (see for example [5]). In fact, X_H is the unique vector field that satisfies

$$\omega\left(X_H(x,p), \cdot\right) = d_x H(\cdot) \qquad \forall (x,p) \in T^*M.$$

For this reason, it is sometime called *symplectic gradient of* H. It is easy to check from both definitions that — only in the autonomous case — the Hamiltonian is a *prime integral of the motion*, *i.e.,* it is constant along the solutions of these equations.

Now, we would like to explain what is the relation between the Euler-Lagrange flow and the Hamiltonian one. It follows easily from the definition of Hamiltonian (and Legendre-Fenchel transform) that for each $(x,v) \in TM$ and $(x,p) \in T^*M$ the following inequality holds:

$$\langle p,\, v\rangle_x \leq L(x,v) + H(x,p)\,. \tag{1}$$

This is called *Fenchel inequality* (or *Legendre-Fenchel inequality*, see [22]) and it plays a crucial role in the study of Lagrangian and Hamiltonian dynamics and in the variational methods that we are going to describe. In particular, equality holds if and only if $p = \partial L/\partial v(x,v)$. One can therefore introduce the following diffeomorphism between TM and T^*M, known as *Legendre transform*:

$$\begin{array}{rcl} \mathcal{L}:\ TM & \longrightarrow & T^*M \\ (x,v) & \longmapsto & \left(x,\, \dfrac{\partial L}{\partial v}(x,v)\right)\,. \end{array} \tag{2}$$

Moreover, the following relation with the Hamiltonian holds:

$$H \circ \mathcal{L}(x,v) = \left\langle \frac{\partial L}{\partial v}(x,v),\, v \right\rangle_x - L(x,v)\,.$$

This diffeomorphism $\mathcal{L}$ represents a conjugacy between the two flows, namely the Euler-Lagrange flow on TM and the Hamiltonian flow on T^*M; in other words, the following diagram commutes:

$$\begin{tikzcd} TM \arrow[r, "\Phi_t^L"] \arrow[d, "\mathcal{L}"'] & TM \arrow[d, "\mathcal{L}"] \\ T^*M \arrow[r, "\Phi_t^H"'] & T^*M \end{tikzcd}$$

Remark 3.3. Since $\mathcal{L}$ and the Hamiltonian flow Φ^H are both C^1, then it follows from the commutative diagram above that the Euler-Lagrange flow is also C^1.

4. Action-minimizing properties of integrable systems

Before entering into the details of Mather's work, we would like to discuss a very easy case: properties of invariant measures of an integrable system (see section 2). This will provide us with a better understanding of the ideas behind Mather's theory and will describe clearer in which sense these *action-minimizing sets* — namely, what we will call *Mather sets* (see section 5) — represent a generalization of KAM tori.

As we have already discussed in section 2, let $H : T^*\mathbb{T}^n \longrightarrow \mathbb{R}$ be an integrable Tonelli Hamiltonian in action-angle coordinates, *i.e.*, $H(x,p) = \mathfrak{h}(p)$ and let $L : T\mathbb{T}^n \longrightarrow \mathbb{R}$, $L(x,v) = \ell(v)$, be the associated Tonelli Lagrangian. We denote by $\Phi^{\mathfrak{h}}$ and Φ^{ℓ} the respective flows, by $\mathcal{L}$ the associated Legendre transform, and identify both $T^*\mathbb{T}^n$ and $T\mathbb{T}^n$ with $\mathbb{T}^n \times \mathbb{R}^n$.

We have recalled in section 3 that the Euler-Lagrange flow can be equivalently defined in terms of a variational principle associated to the *Lagrangian action functional* A_ℓ. We would like to study action-minimizing properties of these invariant manifolds; for, it is much better to work in the Lagrangian setting. Moreover, instead of considering properties of single orbits, it would be more convenient to study "collection" of orbits, in the form of *invariant probability measures*[4] and consider their *average action.* If μ is an invariant probability measure for Φ^ℓ — *i.e.*, $(\Phi^\ell_t)^*\mu = \mu$ for all $t \in \mathbb{R}$, where $(\Phi^\ell_t)^*\mu$ denotes the pull-back of the measure — then we define:

$$A_\ell(\mu) := \int_{T\mathbb{T}^n} \ell(v)\, d\mu.$$

Let us consider any invariant probability measure μ_0 supported on $\tilde{\Lambda}_{p_0} := \mathcal{L}^{-1}(\Lambda_{p_0})$ and compute its action. Observe that on the support of this measure $\ell(v) \equiv \ell(\rho(p_0))$. Then:

$$\begin{aligned} A_\ell(\mu_0) &= \int_{T\mathbb{T}^n} \ell(v)\, d\mu_0 = \int_{T\mathbb{T}^n} \ell(\rho(p_0))\, d\mu_0 = \\ &= \ell(p_0) = p_0 \cdot \rho(p_0) - \mathfrak{h}(p_0), \end{aligned} \tag{3}$$

where in the last step we have used the Legendre-Fenchel duality between h and ℓ.

Let us now consider a general invariant probability measure μ. In this case it is not true anymore that $\ell(v)$ is constant on the support of μ. However, using Legendre-Fenchel inequality (see (1)), we can conclude that $\ell(v) \geq p_0 \cdot v - \mathfrak{h}(p_0)$ for each $v \in \mathbb{R}^n$. Hence:

$$\begin{aligned} A_\ell(\mu) &= \int_{T\mathbb{T}^n} \ell(v)\, d\mu \geq \int_{T\mathbb{T}^n} \left(p_0 \cdot v - \mathfrak{h}(p_0)\right)\, d\mu \\ &= \int_{T\mathbb{T}^n} p_0 \cdot v\, d\mu - \mathfrak{h}(p_0) = p_0 \cdot \left(\int_{T\mathbb{T}^n} v\, d\mu\right) - \mathfrak{h}(p_0). \end{aligned} \tag{4}$$

We would like to compare expressions (3) and (4). However, in the case of a general measure, we do not know how to evaluate the term $\int_{T\mathbb{T}^n} v\, d\mu$. One possible trick to overcome this problem is the following: instead of considering the action of $\ell(v)$, let us consider the action of $\ell(v) - p_0 \cdot v$. It is easy to see that this new

[4]Actually, it is also possible study directly orbits. See Remark 5.8

Lagrangian is also Tonelli (we have subtracted a linear term in v) and that it has the same Euler-Lagrange flow as ℓ. In this way we obtain from (3) and (4) that:

$$A_{\ell-p_0\cdot v}(\mu_0) \quad = \quad -\mathfrak{h}(p_0) \qquad \text{and} \qquad A_{\ell-p_0\cdot v}(\mu) \geq -\mathfrak{h}(p_0),$$

which are now comparable. Hence, we have just showed the following fact:

Fact 1: *Every invariant probability measure supported on $\tilde{\Lambda}_{p_0}$ minimizes the action $A_{\ell-p_0\cdot v}$ amongst all invariant probability measures of Φ^ℓ.*

In particular, we can characterize our invariant tori in a different way:

$$\tilde{\Lambda}_{p_0} = \bigcup \{\operatorname{supp}\mu : \ \mu \text{ minimizes } A_{\ell-p_0\cdot v}\}.$$

Moreover, there is a relation between the energy (Hamiltonian) of the invariant torus and the minimal action of its invariant probability measures:

$$\mathfrak{h}(p_0) = -\min\{A_{\ell-p_0\cdot v}(\mu) : \ \mu \text{ is an inv. prob. measure}\}.$$

Observe that it is somehow expectable that we need to modify the Lagrangian in order to obtain information on a specific invariant torus. In fact, in the case of an integrable system we have a foliation of the space made by these invariant tori and it would be unrealistic to expect that they could all be obtained as extremals of the same action functional. In other words, what we did was to add a *weighting term* to our Lagrangian, in order to magnify some motions rather than others.

Is it possible to distinguish these motions in a different way? Let us go back to (3) and (4). The main problem in comparing these two expression was represented by the term $\int_{T\mathbb{T}^n} v\, d\mu$. This can be interpreted as a sort of average rotation vector of orbits in the support of μ. Hence, let us define the *average rotation vector of* μ as:

$$\rho(\mu) := \int_{T\mathbb{T}^n} v\, d\mu \in \mathbb{R}^n.$$

We will give a more precise definition of it (which is also meaningful on manifolds different from the torus) in section 5.

Let now μ be an invariant probability measure of Φ^ℓ with rotation vector $\rho(\mu) = \rho(p_0)$. It follows from (4) that:

$$\begin{aligned} A_\ell(\mu) &\geq p_0 \cdot \left(\int_{T\mathbb{T}^n} v\, d\mu \right) - \mathfrak{h}(p_0) = p_0 \cdot \rho(\mu) - \mathfrak{h}(p_0) = \\ &= p_0 \cdot \rho(p_0) - \mathfrak{h}(p_0) = \ell(\rho(p_0)). \end{aligned}$$

Therefore, comparing with (3) we obtain another characterization of μ_0:

Fact 2: *Every invariant probability measure supported on $\tilde{\Lambda}_{p_0}$ minimizes the action A_ℓ amongst all invariant probability measures of Φ^ℓ with rotation vector $\rho(p_0)$.*

In particular:

$$\tilde{\Lambda}_{p_0} = \bigcup \{\operatorname{supp}\mu : \ \mu \text{ minimizes } A_\ell \text{ amongst measures with rot. vect. } \rho(p_0)\}.$$

Moreover, there is a relation between the value of the Lagrangian at $\rho(p_0)$ and the minimal action of all invariant probability measures with rotation vector $\rho(p_0)$:

$$\ell(\rho(p_0)) = \min\{A_\ell(\mu) : \ \mu \text{ is an inv. prob. meas. with rot. vect. } \rho(p_0)\}.$$

Remark 4.1. One could also study directly orbits on these tori and try to show that their action minimizes a modified Lagrangian action, in the same spirit as we have just discussed for measures. See [23] and Remark 5.8 for more details.

5. Mather's theory for Tonelli Lagrangian systems

In this section we describe Mather's theory for general Tonelli Lagrangians on compact manifolds. As we have already said before, we refer the reader to [23] for all the proofs and for a more detailed presentation of this theory.

Let $\mathfrak{M}(L)$ be the space of probability measures μ on TM that are invariant under the Euler-Lagrange flow of L and such that $\int_{TM} \|v\|\, d\mu < \infty$. We will hereafter assume that $\mathfrak{M}(L)$ is endowed with the *vague topology*, *i.e.,* the weak*-topology induced by the space C^0_ℓ of continuous functions $f : TM \longrightarrow \mathbb{R}$ having at most linear growth:

$$\sup_{(x,v)\in TM} \frac{|f(x,v)|}{1+\|v\|} < +\infty\,.$$

One can check that $\mathfrak{M}(L) \subset \left(C^0_\ell\right)^*$.

In the case of an autonomous Tonelli Lagrangian, it is easy to see that $\mathfrak{M}(L)$ is non-empty (actually it contains infinitely many measures with distinct supports). In fact, recall that because of the conservation of the energy $E(x,v) := H\circ\mathcal{L}(x,v) = \left\langle \frac{\partial L}{\partial v}(x,v),\, v\right\rangle_x - L(x,v)$ along the orbits, each energy level of E is compact (it follows from the superlinearity condition) and invariant under Φ^L_t. It is a classical result in ergodic theory (sometimes called Kryloff-Bogoliouboff theorem) that a flow on a compact metric space has at least an invariant probability measure, which belongs indeed to $\mathfrak{M}(L)$.

To each $\mu \in \mathfrak{M}(L)$, we may associate its *average action*:

$$A_L(\mu) = \int_{TM} L\, d\mu\,.$$

The action functional $A_L : \mathfrak{M}(L) \longrightarrow \mathbb{R}$ is lower semicontinuous with the vague topology on $\mathfrak{M}(L)$ (this functional might not be necessarily continuous, see [8, Remark 2-3.4]). In particular, this implies that there exists $\mu \in \mathfrak{M}(L)$, which minimizes A_L over $\mathfrak{M}(L)$.

Definition 5.1. *A measure $\mu \in \mathfrak{M}(L)$, such that $A_L(\mu) = \min_{\mathfrak{M}(L)} A_L$, is called an action-minimizing measure of L.*

As we have already seen in section 4, by modifying the Lagrangian (without changing the Euler-Lagrange flow) one can find many other interesting measures besides those found by minimizing A_L. A similar idea can be implemented for a general Tonelli Lagrangian. Observe, in fact, that if η is a 1-form on M, we can interpret it as a function on the tangent bundle (linear on each fiber)

$$\begin{array}{rcl} \hat{\eta} : TM & \longrightarrow & \mathbb{R} \\ (x,v) & \longmapsto & \langle \eta(x),\, v\rangle_x \end{array}$$

and consider a new Tonelli Lagrangian $L_\eta := L - \hat{\eta}$. The associated Hamiltonian will be given by $H_\eta(x,p) = H(x, \eta(x) + p)$.

Observe that:

i) If η is closed, then L and L_η have the same Euler-Lagrange flow on TM. See [18].
ii) If $\mu \in \mathfrak{M}(L)$ and $\eta = df$ is an exact 1-form, then $\int \widehat{df}\, d\mu = 0$. Thus, for a fixed L, the minimizing measures will depend only on the de Rham cohomology class $c = [\eta] \in H^1(M;\mathbb{R})$.

Therefore, instead of studying the action minimizing properties of a single Lagrangian, one can consider a family of such "modified" Lagrangians, parameterized over $H^1(M;\mathbb{R})$. Hereafter, for any given $c \in H^1(M;\mathbb{R})$, we will denote by η_c a closed 1-form with that cohomology class.

Definition 5.2. *Let η_c be a closed 1-form of cohomology class c. Then, if $\mu \in \mathfrak{M}(L)$ minimizes $A_{L_{\eta_c}}$ over $\mathfrak{M}(L)$, we will say that μ is a c-action minimizing measure (or c-minimal measure, or Mather measure with cohomology c).*

Compare with Fact 1 in section 4.

Remark 5.3. Observe that the cohomology class of an action-minimizing invariant probability measure is not intrinsic in the measure itself nor in the dynamics, but it depends on the specific choice of the Lagrangian L. Changing the Lagrangian by a closed 1-form η, *i.e.*, $L \longmapsto L - \eta$, we will change all the cohomology classes of its action minimizing measures by $-[\eta] \in H^1(M;\mathbb{R})$. Compare also with Remark 5.5 (*ii*).

One can consider the following function on $H^1(M;\mathbb{R})$ (the minus sign is introduced for a convention that will probably become clearer later on):

$$\begin{array}{rcl} \alpha : H^1(M;\mathbb{R}) & \longrightarrow & \mathbb{R} \\ c & \longmapsto & -\displaystyle\min_{\mu\in\mathfrak{M}(L)} A_{L_{\eta_c}}(\mu)\,. \end{array}$$

This function α is well-defined (it does not depend on the choice of the representatives of the cohomology classes) and it is easy to see that it is convex. This is generally known as *Mather's α-function.* We have seen in section 4 that for an integrable Hamiltonian $H(x,p) = \mathfrak{h}(p)$, $\alpha(c) = \mathfrak{h}(c)$. For this and several other reasons that we will see later on, this function is sometimes called *effective Hamiltonian.* In particular, it can be proven that $\alpha(c)$ is related to the energy level containing such c-action minimizing measures [7].

We will denote by $\mathfrak{M}_c(L)$ the subset of c-action minimizing measures:

$$\mathfrak{M}_c := \mathfrak{M}_c(L) = \{\mu \in \mathfrak{M}(L) :\ A_{L_{\eta_c}}(\mu) = -\alpha(c)\}.$$

We can now define a first important family of invariant sets: the *Mather sets.*

Definition 5.4. *For a cohomology class $c \in H^1(M;\mathbb{R})$, we define the Mather set of cohomology class c as:*

$$\widetilde{\mathcal{M}}_c := \bigcup_{\mu\in\mathfrak{M}_c} \operatorname{supp}\mu \subset TM\,. \tag{5}$$

The projection on the base manifold $\mathcal{M}_c = \pi\left(\widetilde{\mathcal{M}}_c\right) \subseteq M$ is called projected Mather set (with cohomology class c).

Properties of this set:

i) It is non-empty, compact and invariant [18].

ii) It is contained in the energy level corresponding to $\alpha(c)$ [7].

iii) In [18] Mather proved the celebrated *graph theorem*:
Let $\pi : TM \longrightarrow M$ denote the canonical projection. Then, $\pi|\widetilde{\mathcal{M}}_c$ is an injective mapping of $\widetilde{\mathcal{M}}_c$ into M, and its inverse $\pi^{-1} : \mathcal{M}_c \longrightarrow \widetilde{\mathcal{M}}_c$ is Lipschitz.

Now, we would like to shift our attention to a related problem. As we have seen in section 4, instead of considering different minimizing problems over $\mathfrak{M}(L)$, obtained by modifying the Lagrangian L, one can alternatively try to minimize the Lagrangian L by putting some constraint, such as, for instance, fixing the *rotation vector* of the measures. In order to generalize this to Tonelli Lagrangians on compact manifolds, we first need to define what we mean by rotation vector of an invariant measure.

Let $\mu \in \mathfrak{M}(L)$. Thanks to the superlinearity of L, the integral $\int_{TM} \hat{\eta} d\mu$ is well defined and finite for any closed 1-form η on M. Moreover, if η is exact, then this integral is zero, *i.e.,* $\int_{TM} \hat{\eta} d\mu = 0$. Therefore, one can define a linear functional:

$$\begin{aligned} H^1(M;\mathbb{R}) &\longrightarrow \mathbb{R} \\ c &\longmapsto \int_{TM} \hat{\eta} d\mu\,, \end{aligned}$$

where η is any closed 1-form on M with cohomology class c. By duality, there exists $\rho(\mu) \in H_1(M;\mathbb{R})$ such that

$$\int_{TM} \hat{\eta}\, d\mu = \langle c, \rho(\mu)\rangle \qquad \forall\, c \in H^1(M;\mathbb{R})$$

(the bracket on the right-hand side denotes the canonical pairing between cohomology and homology). We call $\rho(\mu)$ the *rotation vector* of μ. This rotation vector is the same as the Schwartzman's asymptotic cycle of μ (see [24] and [23] for more details).

Remark 5.5. (*i*) It is possible to provide a more geometric interpretation of this. Suppose for the moment that μ is ergodic. Then, it is known that a generic orbit $\gamma(t) := \pi\Phi_t^L(x,v)$, where $\pi : TM \longrightarrow M$ denotes the canonical projection, will return infinitely often close (as close as we like) to its initial point $\gamma(0) = x$. We can therefore consider a sequence of times $T_n \to +\infty$ such that $d(\gamma(T_n), x) \to 0$ as $n \to +\infty$, and consider the closed loops σ_n obtained by closing $\gamma|[0, T_n]$ with the shortest geodesic connecting $\gamma(T_n)$ to x. Denoting by $[\sigma_n]$ the homology class of this loop, one can verify (see [24]) that $\lim_{n\to\infty} \frac{[\sigma_n]}{T_n} = \rho(\mu)$, independently of the chosen sequence $\{T_n\}_n$. In other words, in the case of ergodic measures, the rotation vector tells us how on average a generic orbit winds around TM. If μ is not ergodic, $\rho(\mu)$ loses this neat geometric meaning, yet it may be interpreted as the average of the rotation vectors of its different ergodic components.

(*ii*) It is clear from the discussion above that the rotation vector of an invariant measure depends only on the dynamics of the system (*i.e.,* on the Euler-Lagrange flow) and not on the chosen Lagrangian. Therefore, it does not change when we modify our Lagrangian by adding a closed one form.

Using that the action functional $A_L : \mathfrak{M}(L) \longrightarrow \mathbb{R}$ is lower semicontinuous, one can prove that the map $\rho : \mathfrak{M}(L) \longrightarrow H_1(M;\mathbb{R})$ is continuous and surjective, *i.e.*, for every $h \in H_1(M;\mathbb{R})$ there exists $\mu \in \mathfrak{M}(L)$ with $A_L(\mu) < \infty$ and $\rho(\mu) = h$ (see [18]).

Following Mather [18], let us consider the minimal value of the average action A_L over the probability measures with rotation vector h. Observe that this minimum is actually achieved because of the lower semicontinuity of A_L and the compactness of $\rho^{-1}(h)$ (ρ is continuous and L superlinear). Let us define

$$\begin{array}{rcl} \beta : H_1(M;\mathbb{R}) & \longrightarrow & \mathbb{R} \\ h & \longmapsto & \displaystyle\min_{\mu\in\mathfrak{M}(L):\, \rho(\mu)=h} A_L(\mu)\,. \end{array} \tag{6}$$

This function β is what is generally known as *Mather's β-function* and it is immediate to check that it is convex. We have seen in section 4 that if we have an integrable Tonelli Hamiltonian $H(x,p) = \mathfrak{h}(p)$ and the associated Lagrangian $L(x,v) = \ell(v)$, then $\beta(h) = \ell(h)$. For this and several other reasons, this function is sometime called *effective Lagrangian.*

We can now define what we mean by action minimizing measure with a given rotation vector.

Definition 5.6. *A measure $\mu \in \mathfrak{M}(L)$ realizing the minimum in (6), i.e., such that $A_L(\mu) = \beta(\rho(\mu))$, is called an action minimizing (or minimal, or Mather) measure with rotation vector $\rho(\mu)$.*

Compare with Fact 2 in section 4.

We will denote by $\mathfrak{M}^h(L)$ the subset of action minimizing measures with rotation vector h:

$$\mathfrak{M}^h := \mathfrak{M}^h(L) = \{\mu \in \mathfrak{M}(L) :\ \rho(\mu) = h \text{ and } A_L(\mu) = \beta(h)\}.$$

This allows us to define another important familty of invariant sets.

Definition 5.7. *For a homology class (or rotation vector) $h \in H_1(M;\mathbb{R})$, we define the Mather set corresponding to a rotation vector h as*

$$\widetilde{\mathcal{M}}^h := \bigcup_{\mu\in\mathfrak{M}^h} \operatorname{supp}\mu \subset TM\,, \tag{7}$$

and the projected one as $\mathcal{M}^h = \pi\left(\widetilde{\mathcal{M}}^h\right) \subseteq M$.

Similarly to what we have already seen above, this set satisfies the following properties:

i) It is non-empty, compact and invariant.
ii) It is contained in a given energy level.
iii) It also satisfies the *graph theorem*:
let $\pi : TM \longrightarrow M$ denote the canonical projection. Then, $\pi|\widetilde{\mathcal{M}}^h$ is an injective mapping of $\widetilde{\mathcal{M}}^h$ into M, and its inverse $\pi^{-1} : \mathcal{M}^h \longrightarrow \widetilde{\mathcal{M}}^h$ is Lipschitz.

Remark 5.8. (*i*) In the above discussion we have only discussed properties of invariant probability measures associated to the system. Actually, one could study directly orbits of the systems and look for orbits that globally minimize the action of a modified Lagrangian (in the same spirit as before). This would lead to the definition of two other families of invariant compact sets, the *Aubry sets* $\widetilde{\mathcal{A}}_c$ and the *Mañé sets* $\widetilde{\mathcal{N}}_c$, which are also parameterized by $H^1(M;\mathbb{R})$ (the parameter which describes the modification of the Lagrangian, exactly in the same way as before). For a given $c \in H^1(M;\mathbb{R})$, these sets contain the Mather set $\widetilde{\mathcal{M}}_c$, and this inclusion may be strict. In fact, while the motion on the Mather sets is *recurrent* (it is the union of the supports of invariant probability measures), the Aubry and the Mañé sets may contain non-recurrent orbits as well.
(*ii*) Differently from what happens with invariant probability measures, it will not be always possible to find *action-minimizing orbits* for any given rotation vector (not even possible to define a rotation vector for every action minimizing orbit). For instance, an example due to Hedlund [10] provides the existence of a Riemannian metric on a three-dimensional torus, for which minimal geodesics exist only in three directions. The same construction can be extended to any dimension larger than three.

6. Mather's α and β-functions

The discussion in section 5 led to two equivalent formulations of the minimality of an invariant probability measure μ:

- there exists a homology class $h \in \mathrm{H}_1(M;\mathbb{R})$, namely its rotation vector $\rho(\mu)$, such that μ minimizes A_L amongst all measures in $\mathfrak{M}(L)$ with rotation vector h, *i.e.*, $A_L(\mu) = \beta(h)$.
- There exists a cohomology class $c \in \mathrm{H}^1(M;\mathbb{R})$, such that μ minimizes $A_{L_{\eta_c}}$ amongst all probability measures in $\mathfrak{M}(L)$, *i.e.*, $A_{L_{\eta_c}}(\mu) = -\alpha(c)$.

What is the relation between these two different approaches? Are they equivalent, *i.e.*, $\bigcup_{h\in \mathrm{H}_1(M;\mathbb{R})} \mathfrak{M}^h = \bigcup_{c\in \mathrm{H}^1(M;\mathbb{R})} \mathfrak{M}_c$?

In order to comprehend the relation between these two families of action-minimizing measures, we need to understand better the properties of the these two functions that we have introduced above:

$$\alpha : H^1(M;\mathbb{R}) \longrightarrow \mathbb{R} \quad \text{and} \quad \beta : H_1(M;\mathbb{R}) \longrightarrow \mathbb{R}.$$

Let us start with the following trivial remark.

Remark 6.1. As we have previously pointed out, if we have an integrable Tonelli Hamiltonian $H(x,p) = \mathfrak{h}(p)$ and the associated Lagrangian $L(x,v) = \ell(v)$, then $\alpha(c) = \mathfrak{h}(c)$ and $\beta(h) = \ell(h)$. In this case, the cotangent bundle $T^*\mathbb{T}^n$ is foliated by invariant tori $\mathcal{T}^*_c := \mathbb{T}^n \times \{c\}$ and the tangent bundle $T\mathbb{T}^n$ by invariant tori $\widetilde{\mathcal{T}}^h := \mathbb{T}^n \times \{h\}$. In particular, we proved that

$$\widetilde{\mathcal{M}}_c = \mathcal{L}^{-1}(\mathcal{T}_c) = \widetilde{\mathcal{T}}^h = \widetilde{\mathcal{M}}^h,$$

where h and c are such that $h = \nabla\mathfrak{h}(c) = \nabla\alpha(c)$ and $c = \nabla\ell(h) = \nabla\beta(h)$.

We would like to investigate whether a similar relation linking Mather sets of a certain cohomology class to Mather sets with a certain rotation vector, continues to exist beyond the specificity of this situation. Of course, one main difficulty is that in general the *effective Hamiltonian* α and the *effective Lagrangian* β, although being convex and superlinear (see Proposition 6.2), are not necessarily differentiable.

Before stating the main relation between these two functions, let us recall some definitions and results from classical convex analysis (see [22]). Given a convex function $\varphi : V \longrightarrow \mathbb{R} \cup \{+\infty\}$ on a finite dimensional vector space V, one can consider a *dual* (or *conjugate*) function defined on the dual space V^*, via the so-called *Fenchel transform*: $\varphi^*(p) := \sup_{v \in V} \big(p \cdot v - \varphi(v)\big)$. In our case, the following holds.

Proposition 6.2. *α and β are convex conjugate, i.e., $\alpha^* = \beta$ and $\beta^* = \alpha$. In particular, it follows that α and β have superlinear growth.*

Next proposition will allow us to clarify the relation (and duality) between the two minimizing procedures described above. To state it, recall that, like any convex function on a finite-dimensional space, β admits a subderivative at each point $h \in H_1(M;\mathbb{R})$, *i.e.,* we can find $c \in H^1(M;\mathbb{R})$ such that

$$\forall h' \in H_1(M;\mathbb{R}), \quad \beta(h') - \beta(h) \geq \langle c, h' - h\rangle.$$

As it is usually done, we will denote by $\partial\beta(h)$ the set of $c \in H^1(M;\mathbb{R})$ that are subderivatives of β at h, *i.e.,* the set of c's which satisfy the above inequality. Similarly, we will denote by $\partial\alpha(c)$ the set of subderivatives of α at c. Actually, Fenchel's duality implies an easier characterization of subdifferentials: $c \in \partial\beta(h)$ *if and only if* $\langle c, h\rangle = \alpha(c) + \beta(h)$ (similarly for $h \in \partial\alpha(c)$).

We can now state precisely in which sense what observed in Remark 6.1 continues to hold in the general case

Proposition 6.3. *Let $\mu \in \mathfrak{M}(L)$ be an invariant probability measure. Then:*
(i) *$A_L(\mu) = \beta(\rho(\mu))$ if and only if there exists $c \in \mathrm{H}^1(M;\mathbb{R})$ such that μ minimizes $A_{L_{\eta_c}}$ (i.e., $A_{L_{\eta_c}}(\mu) = -\alpha(c)$).*
(ii) *If μ satisfies $A_L(\mu) = \beta(\rho(\mu))$ and $c \in H^1(M;\mathbb{R})$, then μ minimizes $A_{L_{\eta_c}}$ if and only if $c \in \partial\beta(\rho(\mu))$ (or equivalently $\langle c, h\rangle = \alpha(c) + \beta(\rho(\mu))$).*

Remark 6.4. (*i*) It follows from the above proposition that both minimizing procedures lead to the same sets of invariant probability measures:

$$\bigcup_{h \in \mathrm{H}_1(M;\mathbb{R})} \mathfrak{M}^h = \bigcup_{c \in \mathrm{H}^1(M;\mathbb{R})} \mathfrak{M}_c\,.$$

In other words, minimizing over the set of invariant measures with a fixed rotation vector or globally minimizing the modified Lagrangian (corresponding to a certain cohomology class) are dual problems, as the ones that often appears in linear programming and optimization. In some sense, modifying the Lagrangian by a closed 1-form is analog to the method of Lagrange multipliers for searching constrained critical points of a function.

(*ii*) In particular we have the following inclusions between Mather sets:

$$c \in \partial\beta(h) \quad \Longleftrightarrow \quad h \in \partial\alpha(c) \quad \Longleftrightarrow \quad \widetilde{\mathcal{M}}^h \subseteq \widetilde{\mathcal{M}}_c\,.$$

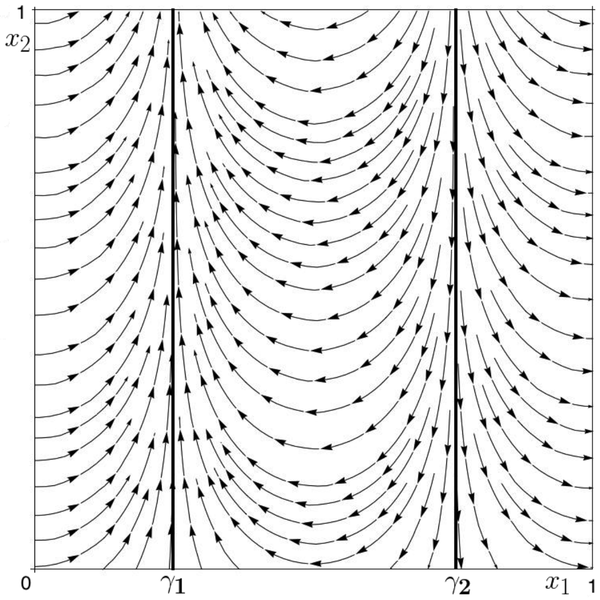


FIGURE 2. Plot of the vector field X.

Moreover, for any $c \in H^1(M;\mathbb{R})$:

$$\widetilde{\mathcal{M}}_c = \bigcup_{h \in \partial\alpha(c)} \widetilde{\mathcal{M}}^h .$$

Observe that the non-differentiability of α at some c produces the presence in $\widetilde{\mathcal{M}}_c$ of (ergodic) invariant probability measures with different rotation vectors. On the other hand, the non-differentiability of β at some h implies that there exist $c \neq c'$ such that $\widetilde{\mathcal{M}}_c \cap \widetilde{\mathcal{M}}_{c'} \neq \emptyset$ (compare with the integrable case discussed in section 4, where these phenomena do not appear).

(*iii*) The minimum of the α-function is sometime called *Mañé's strict critical value.* Observe that if $\alpha(c_0) = \min \alpha(c)$, then $0 \in \partial\alpha(c_0)$ and $\beta(0) = -\alpha(c_0)$. Therefore, the measures with zero homology are contained in the least possible energy level containing Mather sets and $\widetilde{\mathcal{M}}^0 \subseteq \widetilde{\mathcal{M}}_{c_0}$. This inclusion might be strict, unless α is differentiable at c_0; in fact, there may be other action minimizing measures with non-zero rotation vectors corresponding to the other subderivatives of α at c_0.

(*iv)* Note that measures of trivial homology are not necessarily supported on orbits with trivial homology or fixed points. For instance, one can consider the following example. Let $M = \mathbb{T}^2$ equipped with the flat metric and consider a vector field X with norm 1 and such that X has two closed orbits γ_1 and γ_2 in opposite (non-trivial) homology classes and any other orbit asymptotically approaches γ_1 in forward time and γ_2 in backward time; for example one can consider $X(x_1, x_2) = (\cos(2\pi x_1), \sin(2\pi x_1))$, where $(x_1, x_2) \in \mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ (see figure 2).

As we have described in section 3, we can embed this vector field into the Euler-Lagrange vector field given by the Tonelli Lagrangian $L_X(x,v) = \frac{1}{2}\|v - X(x)\|^2$. Let us now consider the probability measure μ_{γ_1} and μ_{γ_2}, uniformly distributed respectively on $(\gamma_1, \dot\gamma_1)$ and $(\gamma_2, \dot\gamma_2)$. Since these two curves have opposite homologies, then $\rho(\mu_{\gamma_1}) = -\rho(\mu_{\gamma_2}) =: h_0 \neq 0$. Moreover, it is easy to see that $A_{L_X}(\mu_{\gamma_1}) = A_{L_X}(\mu_{\gamma_2}) = 0$, since the Lagrangian vanishes on Graph(X). Using the fact that $L_X \geq 0$ (in particular it is strictly positive outside of Graph(X)) and that there are no other invariant ergodic probability measures contained in Graph(X), we can conclude that $\mathcal{M}_0 = \gamma_1 \cup \gamma_2$ and $\alpha(0) = 0$. Moreover, $\mu_0 := \frac{1}{2}\mu_{\gamma_1} + \frac{1}{2}\mu_{\gamma_2}$ has zero homology and its support is contained in $\widetilde{\mathcal{M}}_0$. Therefore (see Proposition 6.3 (i)), μ_0 is action minimizing with rotation vector 0 and $\widetilde{\mathcal{M}}^0 \subseteq \widetilde{\mathcal{M}}_0$; in particular, $\widetilde{\mathcal{M}}^0 = \widetilde{\mathcal{M}}_0$. This also implies that $\beta(0) = 0$ and $\alpha(0) = \min \alpha(c) = 0$.

Observe that α is not differentiable at 0. In fact, reasoning as we have done before for the zero homology class, it is easy to see that for all $t \in [-1,1]$ $\widetilde{\mathcal{M}}^{th_0} = \widetilde{\mathcal{M}}_0$. It is sufficient to consider the convex combination $\mu_\lambda = \lambda\mu_{\gamma_1} + (1-\lambda)\mu_{\gamma_2}$ for any $\lambda \in [0,1]$. Therefore, $\partial\alpha(0) = \{th_0,\ t \in [-1,1]\}$ and $\beta(th_0) = 0$ for all $t \in [-1,1]$.

As we have just seen in item (*iv*) of Remark 6.4, it may happen that the Mather sets corresponding to different homology (resp. cohomology) classes coincide or are included one into the other. This is something that, for instance, cannot happen in the integrable case: in this situation, in fact, these sets form a foliation and are disjoint. The problem in the above mentioned example, seems to be related to a lack of *strict convexity* of β and α. See also the discussion on the simple pendulum in section 7: in this case the Mather sets, corresponding to a non-trivial interval of cohomology classes about 0, coincide.

In the light of this, let us try to understand better what happens when α and β are not strictly convex, *i.e.,* when we are in the presence of *flat* pieces.

Let us first fix some notation. If V is a real vector space and $v_0, v_1 \in V$, we will denote by $\sigma(v_0, v_1)$ the segment joining v_0 to v_1, that is $\sigma(v_0,v_1) := \{tv_0 + (1-t)v_1 : t \in [0,1]\}$. We will say that a function $f: V \longrightarrow \mathbb{R}$ is *affine* on $\sigma(v_0,v_1)$, if there exists $v^* \in V^*$ (the dual of V), such that $f(v) = f(v_0) + \langle v^*, v - v_0\rangle$ for each $v \in \sigma(v_0,v_1)$. Moreover, we will denote by $\mathrm{Int}(\sigma(v_0,v_1))$ the *interior* of $\sigma(v_0,v_1)$, *i.e.,* $\mathrm{Int}(\sigma(v_0,v_1)) := \{tv_0 + (1-t)v_1 :\ t \in (0,1)\}$.

Proposition 6.5. (i) *Let* $h_0, h_1 \in H_1(M;\mathbb{R})$*;* β *is affine on* $\sigma(h_0,h_1)$ *if and only if for any* $h \in \mathrm{Int}(\sigma(h_0,h_1))$ *we have* $\widetilde{\mathcal{M}}^h \supseteq \widetilde{\mathcal{M}}^{h_0} \cup \widetilde{\mathcal{M}}^{h_1}$.
(ii) *Let* $c_0, c_1 \in H^1(M;\mathbb{R})$*;* α *is constant on* $\sigma(c_0,c_1)$ *if and only if for any* $c \in \mathrm{Int}(\sigma(c_0,c_1))$ *we have* $\widetilde{\mathcal{M}}_c \subseteq \widetilde{\mathcal{M}}_{c_0} \cap \widetilde{\mathcal{M}}_{c_1}$.

Remark 6.6. The inclusion in Proposition 6.5 (i) may not be true at the end points of σ. For instance, Remark 6.4 (iv) provides an example in which the inclusion in Proposition 6.5 (i) is not true at the end-points of $\sigma(-h_0,h_0)$.

Remark 6.7. It follows from the previous remarks and Proposition 6.5, that, in general, the action minimizing measures (and consequently the Mather sets $\widetilde{\mathcal{M}}_c$ or $\widetilde{\mathcal{M}}^h$) are not necessarily ergodic. Recall that an invariant probability measure is said to be *ergodic*, if all invariant Borel sets have measure 0 or 1. These measures play a special role in the study of the dynamics of the system, therefore one could ask what are the ergodic action-minimizing measures. It is a well-known result from ergodic theory, that the ergodic measures of a flow correspond to the *extremal points* of the set of invariant probability measures, where by extremal point of a convex set, we mean an element that cannot be obtained as a non-trivial convex combination of other elements of the set. Since β has superlinear growth, its epigraph $\{(h,t) \in H_1(M;\mathbb{R}) \times \mathbb{R} : \ t \geq \beta(h)\}$ has infinitely many extremal points. Let $(h, \beta(h))$ denote one of these extremal points. Then, there exists at least one ergodic action minimizing measure with rotation vector h. It is in fact sufficient to consider any extremal point of the set $\{\mu \in \mathfrak{M}^h(L) : \ A_L(\mu) = \beta(h)\}$: this measure will be an extremal point of $\mathfrak{M}(L)$ and hence ergodic. Moreover, as we have already recalled in Remark 5.5, for such an ergodic measure μ, Birkhoff's ergodic theorem implies that for μ-almost every initial datum, the corresponding trajectory has rotation vector h.

7. An example: the simple pendulum

In this section we would like to describe the Mather sets, the α-function and the β-function, in a specific example: the *simple pendulum*. This system can be described in terms of the Lagrangian:

$$\begin{array}{rcl} L: T\mathbb{T} & \longrightarrow & \mathbb{R} \\ (x,v) & \longmapsto & \frac{1}{2}|v|^2 + \big(1-\cos(2\pi x)\big). \end{array}$$

It is easy to check that the Euler-Lagrange equation provides exactly the equation of the pendulum:

$$\dot v = 2\pi \sin(2\pi x) \qquad \Longleftrightarrow \qquad \left\{ \begin{array}{l} v = \dot x \\ \ddot x - 2\pi\sin(2\pi x) = 0. \end{array}\right.$$

The associated Hamiltonian (or energy) $H : T^*\mathbb{T} \longrightarrow \mathbb{R}$ is given by $H(x,p) := \frac{1}{2}|p|^2 - (1-\cos(2\pi x))$. Observe that in this case the Legendre transform is $(x,p) = \mathcal{L}_L(x,v) = (x,v)$, therefore we can easily identify the tangent and cotangent bundles. In the following we will consider $T\mathbb{T} \simeq T^*\mathbb{T} \simeq \mathbb{T}\times\mathbb{R}$ and identify $H^1(M;\mathbb{R}) \simeq H_1(M;\mathbb{R}) \simeq \mathbb{R}$.

First of all, let us study what are the invariant probability measures of this system.

- Observe that $(0,0)$ and $(\frac{1}{2},0)$ are fixed points for the system (respectively *unstable* and *stable*). Therefore, the Dirac measures concentrated on each of them are invariant probability measures. Hence, we have found two first invariant measures: $\delta_{(0,0)}$ and $\delta_{(\frac{1}{2},0)}$, both with zero rotation vector: $\rho(\delta_{(0,0)}) = \rho(\delta_{(\frac{1}{2},0)}) = 0$. As far as their energy is concerned (*i.e.,* the energy levels in which they are contained), it is easy to check that $E(\delta_{(0,0)}) = H(0,0) = 0$ and $E(\delta_{(\frac{1}{2},0)}) = H(\frac{1}{2},0) = -2$. Observe that these two energy levels cannot contain any other invariant probability measure.

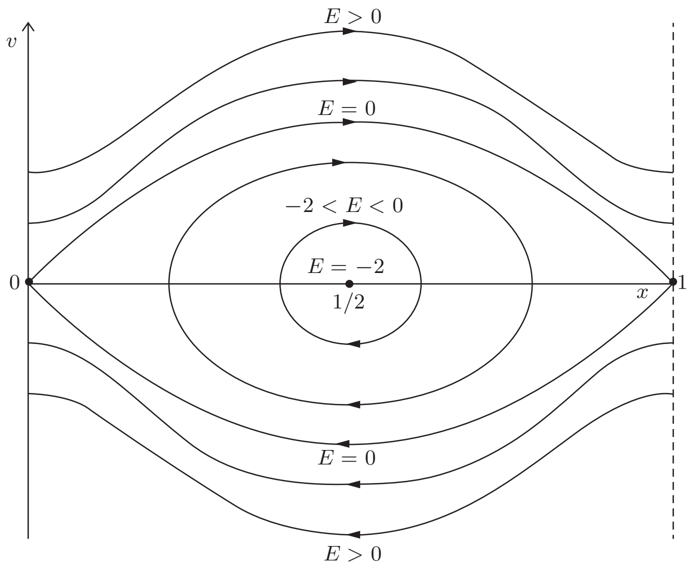


FIGURE 3. The phase space of the simple pendulum.

- If $E > 0$, then the energy level $\{H(x,v) = E\}$ consists of two homotopically non-trivial periodic orbits (*rotation motions*):

$$\mathcal{P}_E^{\pm} := \{(x,v) : \ v = \pm\sqrt{2[(1+E) - \cos(2\pi x)]}, \ \forall\, x \in \mathbb{T}\}.$$

The probability measures evenly distributed along these orbits — which we will denote $\mu_E^{\pm}$ — are invariant probability measures of the system. If we denote by

(8) $$T(E) := \int_0^1 \frac{1}{\sqrt{2[(1+E) - \cos(2\pi x)]}}\, dx$$

the period of such orbits, then it is easy to check that $\rho(\mu_E^{\pm}) = \frac{\pm 1}{T(E)}$ (see Remark 5.5). Observe that this function $T : (0, +\infty) \longrightarrow (0, +\infty)$, which associates to a positive energy E the period of the corresponding periodic orbits $\mathcal{P}_E^{\pm}$, is continuous and strictly decreasing. Moreover, $T(E) \to \infty$ as $E \to 0$ (it is easy to see this, by noticing that motions on the separatrices take an infinitely long time to connect 0 to $1 \equiv 0 \mod \mathbb{Z}$). Therefore, $\rho(\mu_E^{\pm}) \to 0$ as $E \to 0$.
- If $-2 < E < 0$, then the energy level $\{H(x,v) = E\}$ consists of one contractible periodic orbit (*libration motion*):

$$\mathcal{P}_E := \{(x,v) : \ v^2 = 2(1+E) - 2\cos(2\pi x), \quad x \in [x_E, 1 - x_E]\},$$

where $x_E := \frac{1}{2\pi} \arccos(1+E)$. The probability measure evenly distributed along this orbit — which we will denote by μ_E — is an invariant probability measure of the system. Moreover, since this orbit is contractible, its rotation vector is zero: $\rho(\mu_E) = 0$.

The measures above are the only ergodic invariant probability measures of the system. Other invariant measures can be easily obtained as a convex combination of them.

Now we want to understand which of these are action-minimizing for some cohomology class.

Remark 7.1. (*i*) Let us start by remarking that for $-2 < E < 0$ the support of the measure μ_E is not a graph over $\mathbb{T}$, therefore it cannot be action-minimizing for any cohomology class, since otherwise it would violate Mather's graph theorems (see section 5). Therefore all action-minimizing measures will be contained in energy levels corresponding to energy bigger than zero. It follows from what said in sections 5 and 6 that $\alpha(c) \geq 0$ for all $c \in \mathbb{R}$.

(*ii*) Another interesting property of the α-function (in this specific case) is that it is an even function: $\alpha(c) = \alpha(-c)$ for all $c \in \mathbb{R}$. This is a consequence of the particular symmetry of the system, *i.e.,* $L(x,v) = L(x,-v)$. In fact, let us denote $\tau : \mathbb{T} \times \mathbb{R} \longrightarrow \mathbb{T} \times \mathbb{R}$, $(x,v) \longmapsto (x,-v)$ and observe that if μ is an invariant probability measure, then also $\tau^*\mu$ is still an invariant probability measure. Moreover, $\tau^*\mathfrak{M}(L) = \mathfrak{M}(L)$, where $\mathfrak{M}(L)$ denotes the set of all invariant probability measures of L. It is now sufficient to notice that for each $\mu \in \mathfrak{M}(L)$, $\int (L - c\cdot v)\, d\mu = \int (L + c \cdot v) d\tau^*\mu$, and hence conclude that

$$\alpha(c) = -\inf_{\mathfrak{M}(L)} \int (L - c\cdot v)\, d\mu = -\inf_{\mathfrak{M}(L)} \int (L + c\cdot v) d\tau^*\mu = \alpha(-c)\,.$$

(*iii*) It follows from the above symmetry and the convexity of α, that

$$\min_{\mathbb{R}} \alpha(c) = \alpha(0)\,.$$

Let us now start by studying the 0-action minimizing measures, *i.e.,* invariant probability measures that minimize the action of L without any modification. Since $L(x,v) \geq 0$ for each $(x,v) \in \mathbb{T}\times\mathbb{R}$, then $A_L(\mu) \geq 0$ for all $\mu \in \mathfrak{M}(L)$. In particular, $A_L(\delta_{(0,0)}) = 0$, therefore $\delta_{(0,0)}$ is a 0-action minimizing measure and $\alpha(0) = 0$. Since there are not other invariant probability measures supported in the energy level $\{H(x,v) = 0\}$ (*i.e.,* on the separatrices), then we can conclude that:

$$\widetilde{\mathcal{M}}_0 = \{(0,0)\}\,.$$

Moreover, since $\alpha'(0) = 0$ (see Remark 7.1 (*iii*)), then it follows from Remark 6.4 that:

$$\widetilde{\mathcal{M}}^0 = \widetilde{\mathcal{M}}_0 = \{(0,0)\}\,.$$

On the other hand, this could be also deduced from the fact that the only other measures with rotation vector 0, cannot be action minimizing since they do not satisfy the graph theorem (Remark 7.1).

Now let us investigate what happens with other cohomology classes. A naïve observation is that since the α-function is superlinear and continuous, all energy levels for $E \geq 0$ must contain some Mather set; in other words, all energy levels $E \geq 0$ must be obtained as $\alpha(c)$, for some c.

Let $E > 0$ and consider the periodic orbit $\mathcal{P}_E^+$ and the invariant probability measure μ_E^+ evenly distributed on it. The graph of this orbit can be seen as the graph of a closed 1-form $\eta_E^+ := \sqrt{2[(1+E)-\cos(2\pi x)]}\,dx$, whose cohomology class is

$$c^+(E) := [\eta_E^+] = \int_0^1 \sqrt{2[(1+E)-\cos(2\pi x)]}\,dx, \tag{9}$$

which can be interpreted as the (signed) area between the curve and the positive x-semiaxis. This value is clearly continuous and strictly increasing with respect to E (for $E>0$) and as $E \to 0^+$:

$$c^+(E) \longrightarrow \int_0^1 \sqrt{2[1-\cos(2\pi x)]}\,dx = \frac{4}{\pi}\,.$$

Therefore, it defines an invertible function $c^+ : (0,+\infty) \longrightarrow (\frac{4}{\pi},+\infty)$.

We want to prove that μ_E^+ is $c^+(E)$-action minimizing. The proof will be an imitation of what already seen for KAM tori in section 4.

Let us consider the Lagrangian $L_{\eta_E^+}(x,v) := L(x,v) - \eta_E^+(x)\cdot v$. Then, using Legendre-Fenchel inequality (1) (on the support of μ_E^+, because of our choice of η_E^+, this is indeed an equality):

$$\begin{aligned}\int L_{\eta_E^+}(x,v)d\mu_E^+ &= \int \big(L(x,v)-\eta_E^+(x)\cdot v\big)\,d\mu_E^+ = \\ &= \int -H(x,\eta_E^+(x))d\mu_E^+ = -E\,.\end{aligned}$$

Now, let ν be any other invariant probability measure and apply again the same procedure as above (warning: this time Legendre-Fenchel inequality is not an equality anymore!):

$$\begin{aligned}\int L_{\eta_E^+}(x,v)d\nu &= \int \big(L(x,v)-\eta_E^+(x)\cdot v\big)\,d\nu \geq \\ &\geq \int -H(x,\eta_E^+(x))d\nu = -E\,.\end{aligned}$$

Therefore, we can conclude that μ_E^+ is $c^+(E)$-action minimizing. Since it already projects over the whole $\mathbb{T}$, it follows from the graph theorem that it is the only one:

$$\widetilde{\mathcal{M}}_{c^+(E)} = \mathcal{P}_E^+ = \{(x,v):\ v=\sqrt{2[(1+E)-\cos(2\pi x)]},\ \forall\, x\in\mathbb{T}\}.$$

Furthermore, since $\rho(\mu_E^+) = \frac{1}{T(E)}$, then:

$$\widetilde{\mathcal{M}}^{\frac{1}{T(E)}} = \widetilde{\mathcal{M}}_{c^+(E)} = \mathcal{P}_E^+.$$

Similarly, one can consider the periodic orbit $\mathcal{P}_E^-$ and the invariant probability measure μ_E^- evenly distributed on it. The graph of this orbit can be seen as the graph of a closed 1-form $\eta_E^- := -\sqrt{2[(1+E)-\cos(2\pi x)]}\,dx = -\eta_E^+$, whose cohomolgy class is $c^-(E) = -c^+(E)$. Then (see also Remark 7.1 (*ii*)):

$$\widetilde{\mathcal{M}}_{c^-(E)} = \mathcal{P}_E^- = \{(x,v):\ v=-\sqrt{2[(1+E)-\cos(2\pi x)]},\ \forall\, x\in\mathbb{T}\},$$

and

$$\widetilde{\mathcal{M}}^{-\frac{1}{T(E)}} = \widetilde{\mathcal{M}}_{c^-(E)} = \mathcal{P}_E^-.$$

Note that this completes the study of the Mather sets for any given rotation vector, since

$$\rho(\mu_E^{\pm}) = \pm\frac{1}{T(E)} \stackrel{E\to+\infty}{\longrightarrow} \pm\infty \qquad \text{and} \qquad \rho(\mu_E^{\pm}) = \pm\frac{1}{T(E)} \stackrel{E\to 0^+}{\longrightarrow} 0\,.$$

What remains to study is what happens for non-zero cohomology classes in $[-\frac{4}{\pi}, \frac{4}{\pi}]$. The situation turns out to be quite easy. Observe that $\alpha(c^{\pm}(E)) = E$. Thefore, from the continuity of α it follows that (take the limit as $E \to 0$): $\alpha(\pm\frac{4}{\pi}) = 0$. Moreover, since α is convex and $\min \alpha(c) = \alpha(0) = 0$, then: $\alpha(c) \equiv 0$ on $[-\frac{4}{\pi}, \frac{4}{\pi}]$. Therefore, the corresponding Mather sets will lie in the zero energy level. From the above discussion, it follows that in this energy level there is a unique invariant probability measure, namely $\delta_{(0,0)}$, and consequently:

$$\widetilde{\mathcal{M}}_c = \{(0,0)\} \qquad \text{for all } -\frac{4}{\pi} \le c \le \frac{4}{\pi}.$$

Let us summarize what we have found so far. Recall that in (8) and (9) we have introduced these two functions: $T : (0,+\infty) \longrightarrow (0,+\infty)$ and $c^+ : (0,+\infty) \longrightarrow (\frac{4}{\pi}, +\infty)$ representing respectively the period and the cohomology (area below the curve) of the upper periodic orbit of energy E. These functions (for which we have an explicit formula in terms of E) are continuous and strictly monotone (respectively, decreasing and increasing). Therefore, we can define their inverses which provide the energy of the periodic orbit with period T (for all positive periods) or the energy of the periodic orbit with cohomology class c (for $|c| > \frac{4}{\pi}$). We will denote them $E(T)$ and $E(c)$ (observe that this last quantity is exactly $-\alpha(c)$). Then:

$$\widetilde{\mathcal{M}}_c = \begin{cases} \{(0,0)\} & \text{if } -\frac{4}{\pi} \le c \le \frac{4}{\pi} \\ \mathcal{P}^+_{E(c)} & \text{if } c > \frac{4}{\pi} \\ \mathcal{P}^-_{E(-c)} & \text{if } c < -\frac{4}{\pi} \end{cases}$$

and

$$\widetilde{\mathcal{M}}^h = \begin{cases} \{(0,0)\} & \text{if } h = 0 \\ \mathcal{P}^+_{E(\frac{1}{h})} & \text{if } h > 0 \\ \mathcal{P}^-_{E(-\frac{1}{h})} & \text{if } h < 0\,. \end{cases}$$

We can provide an expression for these functions in terms of the quantities introduced above:

$$\alpha(c) = \begin{cases} 0 & \text{if } -\frac{4}{\pi} \le c \le \frac{4}{\pi} \\ E(|c|) & \text{if } |c| > \frac{4}{\pi} \end{cases}$$

and

$$\beta(h) = \begin{cases} 0 & \text{if } h = 0 \\ c(E(\frac{1}{|h|}))|h| - E(\frac{1}{|h|}) & \text{if } h \neq 0\,. \end{cases}$$

Observe that the α-function is C^1. In fact, the only problem might be at $c = \pm\frac{4}{\pi}$, but also there it is differentiable, with derivative 0. If it were not differentiable, then there would exist a subderivative $h \neq 0$ and consequently $\widetilde{\mathcal{M}}^h \subseteq \widetilde{\mathcal{M}}_{\pm\frac{4}{\pi}}$, which is absurd since the set on the right-hand side consists of a single point. However, α is not strictly convex, since there is a flat piece on which it is zero.

As far as β is concerned, it is strictly convex (as a consequence of α being C^1), but it is differentiable everywhere except at the origin. At the origin, in fact, there

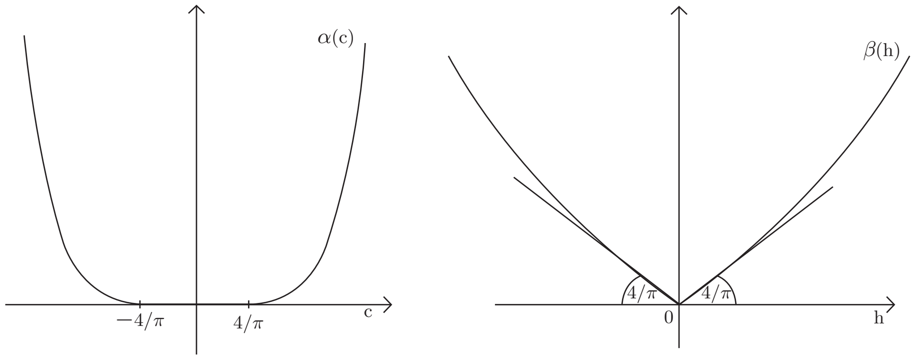


FIGURE 4. Sketch of the graphs of the α and β-functions of the simple pendulum.

is a corner and the set of subderivatives (*i.e.,* the slopes of tangent lines) is given by $\partial\beta(0) = [-\frac{4}{\pi}, \frac{4}{\pi}]$ (this is related to the fact that α has a flat on this interval).

REFERENCES

[1] Vladimir I. Arnol′d. Proof of a theorem of A. N. Kolmogorov on the preservation of conditionally periodic motions under a small perturbation of the Hamiltonian. *Uspehi Mat. Nauk*, 18 (5): 13–40, 1963.

[2] Vladimir I. Arnol′d. Mathematical methods of classical mechanics *Graduate Texts in Mathematics*, 60. Springer-Verlag, New York, 1989.

[3] Serge Aubry and P. Y. Le Daeron. The discrete Frenkel-Kontorova model and its extensions. I. Exact results for the ground-states. *Phys. D*, 8 (3): 381–422, 1983.

[4] Victor Bangert. Mather sets for twist maps and geodesics on tori. *Dyn. Rep.* 1, 1–56, 1988.

[5] Ana Cannas da Silva. Lecture on Symplectic Geometry. *Lecture Notes in Mathematics* 1764, Springer-Verlag, 2001.

[6] Constantin Caratheodory. Variationsrechnung und partielle Differentialgleichung erster Ordnung. Leipzig-Berlin: B.G. Teubner, 1935.

[7] Mario J. Dias Carneiro. On minimizing measures of the action of autonomous Lagrangians. *Nonlinearity*, 8 (6): 1077–1085, 1995.

[8] Gonzalo Contreras and Renato Iturriaga. Global minimizers of autonomous Lagrangians. Preprint, 1999.

[9] Albert Fathi. The Weak KAM theorem in Lagrangian dynamics. Cambridge University Press (to appear).

[10] Gustav A. Hedlund. Geodesics on a two-dimensional Riemannian manifold with periodic coefficients. *Ann. of Math. (2)*, 33 (4): 719–739, 1932.

[11] Andrey N. Kolmogorov. On conservation of conditionally periodic motions for a small change in Hamilton's function. *Dokl. Akad. Nauk SSSR (N.S.)*, 98: 572–530, 1954.

[12] Ricardo Mañé. On the minimizing measures of Lagrangian dynamical systems. *Nonlinearity*, 5(3): 623–638, 1992.

[13] Ricardo Mañé. Lagrangian flows: the dynamics of globally minimizing orbits. *Bol. Soc. Brasil. Mat. (N.S.)*, 28 (2): 141–153, 1997.

[14] John N. Mather. Existence of quasiperiodic orbits for twist homeomorphisms of the annulus. *Topology*, 21 (4): 457–467, 1982.

[15] John N. Mather. A criterion for the nonexistence of invariant circles. *Inst. Hautes Études Sci. Publ. Math.*, 63: 153–204, 1986.

[16] John N. Mather and Giovanni Forni. Action minimizing orbits in Hamiltonian systems. *Transition to chaos in classical and quantum mechanics.* Lectures given at the 3rd session of the Centro Internazionale Matematico Estivo (C.I.M.E.) held in Montecatini Terme, Italy, July 6–13, 1991. Berlin: Springer-Verlag. Lect. Notes Math. 1589, 92–186, 1994.

[17] John N. Mather. Variational construction of orbits of twist diffeomorphisms. *J. Amer. Math. Soc.*, 4 (2): 207–263, 1991.

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

[19] John N. Mather. Variational construction of connecting orbits. *Ann. Inst. Fourier (Grenoble)*, 43 (5): 1349–1386, 1993.

[20] John N. Mather. Order structure on action minimizing orbits. *Symplectic topology and measure preserving dynamical systems.* Papers of the AMS-IMS-SIAM joint summer research conference, Snowbird, UT, USA, July 1–5, 2007. Providence, RI: American Mathematical Society (AMS). Contemporary Mathematics 512, 41–125, 2010.

[21] Jürgen Moser. Convergent series expansions for quasi-periodic motions. *Math. Ann.*, 169: 136–176, 1967.

[22] R. Tyrell Rockafellar. Convex analysis. *Princeton Mathematical Series, No. 28*, Princeton University Press, xviii+451, 1970.

[23] Alfonso Sorrentino. Action-minimizing methods in Hamiltonian dynamics: an introduction to Aubry-Mather theory. *Mathematical Notes* Vol. 50, Princeton and Oxford University Press, pp. xi + 128, May 2015.

[24] Sol Schwartzman. Asymptotic cycles. *Ann. of Math. (2)*, 66: 270–284, 1957.

Dipartimento di Matematica
Università degli Studi di Roma "Tor Vergata"
Viale della Ricerca Scientifica 1, 00133 Rome, Italy
E-mail address: sorrentino@mat.uniroma2.it

Este volumen contiene notas de cursos.

This volume contains course notes.

Publicaciones Matemáticas del Uruguay

Volumen 16 Julio 2016

NOTAS DE CURSOS