An avoiding cones condition for the Poincaré – Birkhoff Theorem

Alessandro Fonda[a,*], Paolo Gidoni[b]

[a] *Dipartimento di Matematica e Geoscienze, Università di Trieste, P.le Europa 1, I-34127 Trieste, Italy*
[b] *SISSA - International School for Advanced Studies, Via Bonomea 265, I-34136 Trieste, Italy*

Abstract

We provide a geometric assumption which unifies and generalizes the conditions proposed in [11, 12], so to obtain a higher dimensional version of the Poincaré–Birkhoff fixed point Theorem for Poincaré maps of Hamiltonian systems.

Keywords: Poincaré–Birkhoff Theorem, avoiding cones condition, Hamiltonian systems, periodic solutions
2010 MSC: 34C25, 37C27

1 Introduction and main result

The seminal work of Henri Poincaré [20] gave rise to a huge amount of research, with the aim of better understanding the far-reaching consequences of the so-called *Poincaré's last geometric Theorem* or *Poincaré–Birkhoff Theorem*. Since then, however, a genuine generalization to higher dimensions of this planar fixed point theorem has never been found. We refer to [1, 15] for a classical introduction, and to [8, 17] for recent reviews on this topic. Recently, however, the first author and Antonio J. Ureña proposed in [11, 12] a higher dimensional version of the Poincaré–Birkhoff Theorem which applies

*Corresponding author
Email addresses: a.fonda@units.it (Alessandro Fonda), pgidoni@sissa.it (Paolo Gidoni)

Version: February 11, 2016

to Poincaré maps of Hamiltonian systems. The aim of this paper is to unify and generalize the geometrical conditions proposed there.

We consider the Hamiltonian system

$$\dot{z} = J\nabla H(t,z)\,, \tag{HS}$$

where $J = \begin{pmatrix} 0 & I_N \\ -I_N & 0 \end{pmatrix}$ denotes the standard $2N \times 2N$ symplectic matrix, and we assume the Hamiltonian function $H\colon \mathbb{R}\times\mathbb{R}^{2N} \to \mathbb{R}$ to be $\mathcal{C}^\infty$-smooth, and T-periodic in its first variable t. (Actually, such a regularity assumption can be considerably weakened, as will be discussed below.) We denote by $\nabla H(t,z)$ the gradient with respect to the variable z.

For every $\zeta \in \mathbb{R}^{2N}$, we denote by $\mathcal{Z}(\cdot,\zeta)$ the unique solution of (HS) satisfying $\mathcal{Z}(0,\zeta) = \zeta$. We assume that these solutions can be continued to the whole time interval $[0,T]$, so that the Poincaré map $\mathcal{P}\colon \mathbb{R}^{2N} \to \mathbb{R}^{2N}$ is well defined, by setting

$$\mathcal{P}(\zeta) = \mathcal{Z}(T,\zeta)\,,$$

and it is a diffeomorphism. The fixed points of $\mathcal{P}$ are associated with the T-periodic solutions of (HS).

For $z \in \mathbb{R}^{2N}$, we use the notation $z = (x,y)$, with $x = (x_1,\dots,x_N) \in \mathbb{R}^N$ and $y = (y_1,\dots,y_N) \in \mathbb{R}^N$, and we assume that $H(t,x,y)$ is 2π-periodic in each of the variables $x_1,\dots,x_N$. Under this setting, T-periodic solutions of (HS) appear in equivalence classes made of those solutions whose components $x_i(t)$ differ by an integer multiple of 2π. We say that two T-periodic solutions are *geometrically distinct* if they do not belong to the same equivalence class. The same will be said for two fixed points of $\mathcal{P}$.

We now describe our geometrical setting, by introducing a family of closed cones associated to a particularly structured vector field.

Let $F\colon \mathbb{R}^N \to \mathbb{R}^N$ be a $\mathcal{C}^\infty$-smooth gradient function, i.e., there is a function $h\colon \mathbb{R}^N \to \mathbb{R}$ such that $F = \nabla h$. We define, for every $y \in \mathbb{R}^N$, the set $\mathcal{A}_F(y)$ as follows: a vector $v \in \mathbb{R}^N$ belongs to $\mathcal{A}_F(y)$ if and only if there exist a sequence $(y_n)_n$ of points in $\mathbb{R}^N$ and a sequence $(\mu_n)_n$ of non-negative real numbers such that

$$y_n \to y\,, \quad \text{and} \quad \mu_n F(y_n) \to v\,.$$

It can be easily seen that $\mathcal{A}_F(y)$ is a closed cone in $\mathbb{R}^N$.

Our main result is the following.

Theorem 1. *Let $F = \nabla h \colon \mathbb{R}^N \to \mathbb{R}^N$ be a $\mathcal{C}^\infty$-smooth function for which there are two constants $K > 0$ and $C > 0$ and a regular symmetric $N \times N$ matrix $\mathbb{S}$ such that*

$$|F(y) - \mathbb{S}y| \leq C\,, \quad \text{when } |y| \geq K\,, \tag{1}$$

and set $D := F^{-1}(0)$. Writing

$$\mathcal{P}(x, y) = (x + \vartheta(x, y), \rho(x, y))\,, \tag{2}$$

suppose that

$$\vartheta(x, y) \notin \mathcal{A}_F(y)\,, \text{ for every } (x, y) \in \mathbb{R}^N \times \partial D\,. \tag{AC}$$

Then, $\mathcal{P}$ has at least $N+1$ geometrically distinct fixed points, all in $\mathbb{R}^N \times D$. Moreover, if all its fixed points are non degenerate, then there are at least 2^N of them.

Assumption (AC) is our *avoiding cones condition.* In other words, for every $(x, y) \in \mathbb{R}^N \times \partial D$, one must have that $\vartheta(x, y) \neq 0$, and that there isn't any sequence $(y_n)_n$ in $\mathbb{R}^N \setminus D$ with

$$y_n \to y\,, \quad \text{and} \quad \frac{F(y_n)}{|F(y_n)|} \to \frac{\vartheta(x, y)}{|\vartheta(x, y)|}\,.$$

The proof of Theorem 1 is provided in Section 3. The geometrical meaning of the avoiding cones condition will be discussed extensively in Section 2, including the study of some substantial cases.

We highlight that Theorem 1 unifies and extends the different boundary twist conditions previously considered, which we now recall in the case of a strongly convex set $D \subseteq \mathbb{R}^N$. The first twist condition, proposed in [11], generalizes an assumption first introduced by Conley and Zehnder in [3].

(T1) There exists a regular symmetric $N \times N$ matrix $\mathbb{B}$ such that

$$\langle \vartheta(x, y), \mathbb{B}\nu_D(y) \rangle > 0\,, \qquad \text{for every } (x, y) \in \mathbb{R}^N \times \partial D\,,$$

where $\nu_D(y)$ denotes the unit outward normal vector to D at y.

Clearly, condition (T1) implies that the vector $\vartheta(x, y)$ has to avoid an entire half-space. Instead, as will be shown in Section 2, our condition (AC) requires to avoid only a half-line almost everywhere, and a two-dimensional half-plane in the remaining points.

The second twist condition in literature was introduced in [19], restricted to the case $\mathbb{B} = I_N$ and requiring a monotone twist of the map $\vartheta(x, y)$. These two assumptions have been dropped in [11].

(T2) There exist an involutory $N \times N$ matrix $\mathbb{B}$ and some point $d_0 \in \operatorname{int} D$ with

$$\langle \vartheta(x, y), \mathbb{B}(y - d_0) \rangle > 0 \,, \qquad \text{for every } (x, y) \in \mathbb{R}^N \times \partial D \,.$$

As in the previous case, the avoiding cones condition (AC) replaces the half-spaces that have to be avoided by $\vartheta(x, y)$ with a half-line almost everywhere, and a half-plane elsewhere.

The third twist condition we want to recall, named *avoiding rays condition*, was introduced in [12], in the general case of sets D whose boundaries are diffeomorphic to a sphere.

(T3) $$\vartheta(x, y) \notin \{\mu \nu_D(y) : \mu \geq 0\} \,, \qquad \text{for every } (x, y) \in \mathbb{R}^N \times \partial D \,.$$

As will be shown in Section 2, our condition (AC) extends (T3) to sets D with a non-smooth boundary. Moreover, (AC) covers also situations of indefinite twist, as those of (T1) and (T2) when an indefinite matrix $\mathbb{B}$ is involved, which are not contemplated by (T3).

Concerning the regularity assumptions, we could assume the Hamiltonian H to be continuous, with a continuous gradient ∇H. However, such a mild regularity requires a lot of technicalities in the proof, which have been carried out in [11]. On the contrary, we prefer here to avoid these difficulties, in order to focus on the different geometric features arising from the avoiding cones condition.

We notice that the cone $\mathcal{A}_F(y)$ associated to a boundary point $y \in \partial D$ generally depends only on the sign of the potential h in a neighbourhood of y (taking $h = 0$ on D), and can be expressed in terms of suitable normal cones, under some mild regularity assumption on the level sets of h. Yet, an abstract and general characterization of the cones $\mathcal{A}_F(y)$ is definitely non-trivial, and has yet to be accomplished.

Another clue in this direction is provided by [6], where a similar avoiding cones condition has been provided and applied in a topological setting, leading to some fixed point theorems. Even if a different construction is adopted there, the strong resemblance between these two families of avoiding cones conditions, together with their links with Conley Index Theory, make us conjecture a common root for the two situations. This would mean a double interpretation of the avoiding cones conditions: as a condition of non-zero degree for continuous maps in $\mathbb{R}^N$, and as a twist condition for symplectic maps is $\mathbb{R}^{2N}$. This duality is well recognizable for $N = 1$, when the avoiding cones condition is reduced to a change of sign of a real valued map between the two endpoints of an interval D. For continuous maps $f : D \to \mathbb{R}$ we recover Bolzano's Theorem; whereas, for symplectic maps $\vartheta : \mathbb{T} \times D \to \mathbb{R}$, as considered in this paper, we get the Poincaré–Birkhoff Theorem with the classical twist condition on the planar annulus.

2 The avoiding cones condition, concretely

We now investigate the nature of our avoiding cones condition. We first present, in Sections 2.1 and 2.2, two particular cases which already include the most relevant features. Later, in Section 2.3, we will show how these two special situations actually have a wider extent. Finally, in Section 2.4, we prove that the twist conditions (T1), (T2) and (T3) are included in (AC) and illustrate how the first two are indeed rather more restrictive.

In the following, we will start from a set $D \subseteq \mathbb{R}^N$ and construct a suitable function $F : \mathbb{R}^N \to \mathbb{R}^N$ satisfying the assumptions of Theorem 1, for which $D = F^{-1}(0)$. Before proceeding in our analysis, a couple of remarks are in order.

It is useful to introduce, in relation to the cone $\mathcal{A}_F(y)$, the set

$$\alpha_F(y) = \{v \in \mathcal{A}_F(y) : \; \|v\| = 1\} ,$$

so that

$$\mathcal{A}_F(y) = \{\mu v : \mu > 0 \, , \, v \in \alpha_F(y)\} \cup \{0\} \, .$$

Notice that, if $y \notin D$, we have

$$\mathcal{A}_F(y) = \{\mu F(y) : \mu \geq 0\} \, , \qquad \qquad \alpha_F(y) = \left\{ \frac{F(y)}{|F(y)|} \right\} ;$$

on the other hand, if y belongs to $\operatorname{int} D$, then

$$\mathcal{A}_F(y) = \{0\} \, , \qquad \qquad \alpha_F(y) = \emptyset \, ,$$

and vice versa. The case when y lies in ∂D is less trivial. We know that $\alpha_F(y) \neq \emptyset$ for every $y \in \partial D$. Indeed, if $y \in \partial D$, there exists a sequence of points $y_n \in \mathbb{R}^N \setminus D$ such that $y_n \to y$ and, consequently, a sequence of vectors $v_n \in \mathbb{R}^N$, such that $|v_n| = 1$ and $\alpha_F(y_n) = \{v_n\}$. By compactness, there exists a subsequence v_{n_k} such that $v_{n_k} \to v$ for some v, with $|v| = 1$, and therefore $v \in \alpha_F(y)$. This shows that, for $y \in \partial D$, the set $\alpha_F(y)$ is non-empty, but, in general, it can be multivalued, as displayed below.

In the following, we illustrate three particular situations which present the key features and provide quite natural tools for applications, minimizing at the same time the required computations. The same techniques and ideas can naturally be applied to more general situations.

In many constructions we will need to consider a $\mathcal{C}^\infty$-smooth function $\gamma\colon \mathbb{R} \to \mathbb{R}$, with

$$\gamma(s) = \begin{cases} 0\,, & \text{if } s \le 0\,, \\ 1\,, & \text{if } s \ge 1\,, \end{cases}$$

and such that, for some $\varepsilon_\gamma > 0$,

$$\gamma'(s) > 0\,, \quad \text{for } s \in]0,1[\,, \qquad \gamma''(s) > 0\,, \quad \text{for } s \in]0,\varepsilon_\gamma[\,.$$

We denote by $\langle \cdot, \cdot \rangle$ the Euclidean scalar product in $\mathbb{R}^N$, with its associated norm $|\cdot|$. We write $\mathcal{B}^N(x_0, r)$ for the open ball in $\mathbb{R}^N$ centred at x_0 with radius $r > 0$, and $\mathcal{B}^N[x_0, r]$ for the closed ball.

2.1 The closed ball

We consider a decomposition of the form $\mathbb{R}^N = \mathbb{R}^{N_1} \times \mathbb{R}^{N_2}$, where N_1 or N_2 may possibly be zero, and we define the matrix

$$\mathbb{B} = \begin{pmatrix} I_{N_1} & 0 \\ 0 & -I_{N_2} \end{pmatrix}\,. \tag{3}$$

Corollary 2. *Let $D = \mathcal{B}^N[0, 1]$ and assume that, for every $(x, y) \in \mathbb{R}^N \times \partial D$,*

$$\vartheta(x, y) \notin \begin{cases} \{\mu y : \mu \ge 0\}\,, & \text{if } \langle \mathbb{B}y, y \rangle > 0\,, \\ \{\mu_1 y + \mu_2 \mathbb{B}y : \mu_1 \in \mathbb{R}, \mu_2 \ge 0\}\,, & \text{if } \langle \mathbb{B}y, y \rangle = 0\,, \\ \{-\mu y : \mu \ge 0\}\,, & \text{if } \langle \mathbb{B}y, y \rangle < 0\,. \end{cases} \tag{4}$$

Then, the same conclusion of Theorem 1 holds.

Proof. We define the function $h\colon \mathbb{R}^N \to \mathbb{R}$ as

$$h(y) = \gamma(|y| - 1)\,\langle \mathbb{B}y, y\rangle\ , \tag{5}$$

and set $F := \nabla h$, i.e.,

$$F(y) = C_1(y)y + C_2(y)\mathbb{B}y\,,$$

with

$$C_1(y) = \frac{\gamma'(|y|-1)}{|y|}\,\langle \mathbb{B}y, y\rangle\ , \qquad C_2(y) = 2\gamma(|y|-1)\,.$$

We observe that $F^{-1}(0) = D$. Indeed, when $y \notin D$, one has $\gamma(|y|-1) > 0$ and, whenever the two vectors $\langle \mathbb{B}y, y\rangle\, y$ and $\mathbb{B}y$ are on the same line, then they have also the same direction. We define, for every $y \notin D$, a rescaling of the coefficients $C_1(y)$ and $C_2(y)$, namely

$$c_1(y) = \frac{C_1(y)}{\sqrt{C_1(y)^2 + C_2(y)^2}}\,, \quad c_2(y) = \frac{C_2(y)}{\sqrt{C_1(y)^2 + C_2(y)^2}}\,. \tag{6}$$

so that, for $y \notin D$, we have $\alpha_F(y) = \{c_1(y)y + c_2(y)\mathbb{B}y\}$. We will prove that, for every $y \in \partial D$,

$$\alpha_F(y) = \begin{cases} \{\operatorname{sgn}(\langle \mathbb{B}y, y\rangle)y\}\,, & \text{if } \langle \mathbb{B}y, y\rangle \neq 0\,, \\[2ex] \big\{\tau y + \sqrt{1-\tau^2}\,\mathbb{B}y : \tau \in [-1,1]\big\}\,, & \text{if } \langle \mathbb{B}y, y\rangle = 0\,. \end{cases} \tag{7}$$

First, let $y \in \partial D$ be such that $\langle \mathbb{B}y, y\rangle \neq 0$. We take a sequence $(Y_n)_n$ of vectors $Y_n \in \mathcal{B}^N(0, 1+\varepsilon_\gamma) \setminus D$, with $Y_n \to y$. For $s \in \,]0, \varepsilon_\gamma[\,$, having assumed $\gamma''(s) > 0$, it follows that $\gamma(s) \le s\gamma'(s)$, hence

$$\lim_n \left|\frac{C_2(Y_n)}{C_1(Y_n)}\right| = \lim_n \frac{2\,|Y_n|\,\gamma(|Y_n|-1)}{\gamma'(|Y_n|-1)\,|\langle \mathbb{B}Y_n, Y_n\rangle|} \le \lim_n \frac{2\,|Y_n|\,(|Y_n|-1)}{|\langle \mathbb{B}Y_n, Y_n\rangle|} = 0\,.$$

This implies (7) in the case $\langle \mathbb{B}y, y\rangle \neq 0$.

Let us now look at the case when $y \in \partial D$ and $\langle \mathbb{B}y, y\rangle = 0$. Since $C_2 \ge 0$, by the properties of the limit we deduce the $\subseteq$ inclusion in (7). To check the $\supseteq$ inclusion, let us take a sequence of positive real numbers $l_n \in \,]0, \varepsilon_\gamma[\,$, with $l_n \to 0$, and consider the two sequences of points

$$P_n = y + l_n y\,, \qquad Q_n = y + l_n \mathbb{B}y\,.$$

We observe that $P_n \to y$ and $Q_n \to y$. We have $C_1(P_n) = 0$, while

$$\lim_n \left| \frac{C_2(Q_n)}{C_1(Q_n)} \right| = \lim_n \left| \frac{2|Q_n|\gamma(|Q_n|-1)}{\gamma'(|Q_n|-1)\langle \mathbb{B}Q_n, Q_n\rangle} \right| \le \lim_n \frac{\sqrt{1+l_n^2}-1}{l_n} = 0\,.$$

Hence,

$$\begin{aligned} c_1(P_n) &= 0\,, & \lim_{n\to+\infty} c_1(Q_n) &= 1\,, \\ c_2(P_n) &= 1\,, & \lim_{n\to+\infty} c_2(Q_n) &= 0\,, \end{aligned}$$

and so both y and $\mathbb{B}y$ belong to $\alpha_F(y)$. By continuity, for every $\tau \in]0,1[$ and every sufficiently large n, there exists $\Lambda_n \in [0,1]$ such that, setting $Y_n = \Lambda_n P_n + (1-\Lambda_n)Q_n$,

$$c_1\big(Y_n\big) = \tau\,, \quad c_2\big(Y_n\big) = \sqrt{1-\tau^2}\,.$$

Since $Y_n \to y$, it follows that

$$\tau y + \sqrt{1-\tau^2}\,\mathbb{B}y \in \alpha_F(y)\,, \qquad \text{for every } \tau \in]0,1[\,.$$

We have thus proved that

$$\alpha_F(y) \supseteq \Big\{ \tau y + \sqrt{1-\tau^2}\,\mathbb{B}y : \tau \in [0,1] \Big\}\,.$$

The remaining part of the proof, i.e. the inclusion with $\tau \in [-1,0]$, can be treated similarly, replacing in the construction above Q_n with

$$Q_n^- = y - l_n\,\mathbb{B}y\,.$$

Hence (7) is established, and proof of the corollary is easily completed. □

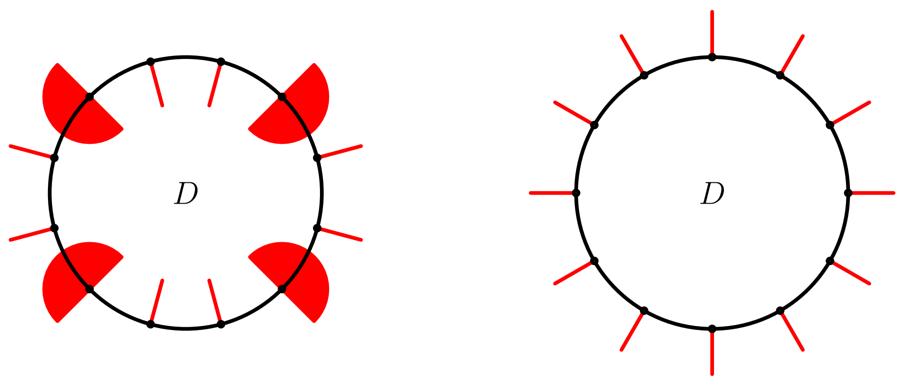
The avoiding cones condition of Corollary 2 is visualized in Figure 1.

Example 3. We take $D = \mathcal{B}^2[0,1]$ and define the Hamiltonian function

$$H(x_1, x_2, y_1, y_2) = y_1^2 + y_2^2 + 2\cos(\pi y_1)\,.$$

The map $\vartheta(x,y) = (\vartheta_1(x,y), \vartheta_2(x,y))$ is given by

$$\begin{aligned} \vartheta_1(x,y) &= T\frac{\partial H}{\partial y_1}(x,y) = 2T[y_1 - \pi\sin(\pi y_1)]\,, \\ \vartheta_2(x,y) &= T\frac{\partial H}{\partial y_2}(x,y) = 2Ty_2\,. \end{aligned}$$



(a) $N_1 = N_2 = 1$.

(b) $N_1 = 2$ *and* $N_2 = 0$.

Figure 1: Visualization of $\mathcal{A}_F(y)$ in the framework of Corollary 2, for different choices of the decomposition $N = N_1 + N_2$.

As illustrated in Figure 2, the avoiding cones condition as in (4) is verified, for $N_1 = 0$ and $N_2 = 2$. The same property is inherited by all the sufficiently small perturbations of H, satisfying the regularity and periodicity assumptions of Theorem 1.

2.2 The product of two closed balls

Let us consider, as before, a decomposition of the type $\mathbb{R}^N = \mathbb{R}^{N_1} \times \mathbb{R}^{N_2}$, where N_1 or N_2 may possibly be zero. For every $y \in \mathbb{R}^{N_1} \times \mathbb{R}^{N_2}$, we write $y = \hat{y}_1 + \hat{y}_2$, with $\hat{y}_1 \in \mathbb{R}^{N_1} \times \{0\}$ and $\hat{y}_2 \in \{0\} \times \mathbb{R}^{N_2}$.

Corollary 4. *Let $D = D_1 \times D_2$, with $D_1 = \mathcal{B}^{N_1}[0,1]$ and $D_2 = \mathcal{B}^{N_2}[0,1]$, and assume that, for every $(x,y) \in \mathbb{R}^N \times \partial D$,*

$$\vartheta(x,y) \notin \begin{cases} \{-\mu\hat{y}_2 : \mu \geq 0\}\,, & \text{if } y \in \operatorname{int} D_1 \times \partial D_2\,, \\ \{\mu_1\hat{y}_1 - \mu_2\hat{y}_2 : \mu_1 \geq 0, \mu_2 \geq 0\}\,, & \text{if } y \in \partial D_1 \times \partial D_2\,, \\ \{\mu\hat{y}_1 : \mu \geq 0\}\,, & \text{if } y \in \partial D_1 \times \operatorname{int} D_2\,. \end{cases} \tag{8}$$

Then, the same conclusion of Theorem 1 holds.

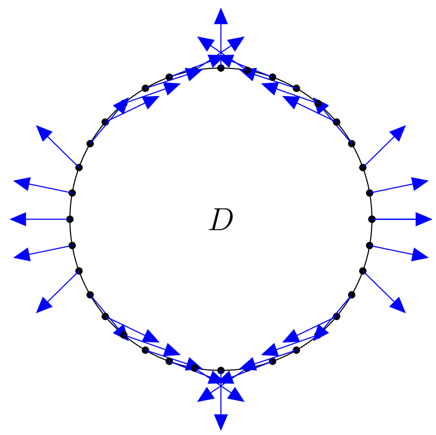


Figure 2: Normalized Poincaré map for the Hamiltonian system of Example 3.

Proof. We define the function $h\colon \mathbb{R}^N \to \mathbb{R}$ as

$$h(y) = \gamma(|\hat{y}_1|)\,|\hat{y}_1|^2 - \gamma(|\hat{y}_2|)\,|\hat{y}_2|^2 \,,$$

and set $F := \nabla h$, i.e.,

$$F(y) = C_1(y)\hat{y}_1 - C_2(y)\hat{y}_2 \,,$$

with

$$C_1(y) = \gamma'(|\hat{y}_1|)\,|\hat{y}_1| + 2\gamma(|\hat{y}_1|)\,, \quad C_2(y) = \gamma'(|\hat{y}_2|)\,|\hat{y}_2| + 2\gamma(|\hat{y}_2|)\,.$$

We observe that $F^{-1}(0) = D$. We will prove that, for every $y \in \partial D$,

$$\alpha_F(y) = \begin{cases} \left\{ -\dfrac{\hat{y}_2}{|\hat{y}_2|} \right\}, & \text{if } y \in \operatorname{int} D_1 \times \partial D_2\,, \\ \left\{ \tau \dfrac{\hat{y}_1}{|\hat{y}_1|} - \sqrt{1-\tau^2}\,\dfrac{\hat{y}_2}{|\hat{y}_2|},\ \tau \in [0,1] \right\}, & \text{if } y \in \partial D_1 \times \partial D_2\,, \\ \left\{ \dfrac{\hat{y}_1}{|\hat{y}_1|} \right\}, & \text{if } y \in \partial D_1 \times \operatorname{int} D_2\,. \end{cases} \tag{9}$$

First of all, we notice that, for every $y \in D_1 \times \mathbb{R}^{N_2}$, the $\mathbb{R}^{N_1}$-component of $F(y)$ is zero, since $C_1(y) = 0$; hence, if $y \in \operatorname{int} D_1 \times \partial D_2$, being $D_1 \times \mathbb{R}^{N_2}$ a neighbourhood y, we deduce that (9) is verified in this case. The case $y \in \partial D_1 \times \operatorname{int} D_2$ is analogous.

Finally, let us consider the case $y \in \partial D_1 \times \partial D_2$. The $\subseteq$ inclusion follows from the fact that the functions c_1 and c_2, defined by a rescaling of C_1 and C_2 as in (6), take values in $[0,1]$ and the sum of their squares is always equal to one. To check the $\supseteq$ inclusion, let us take any sequence of positive real numbers $l_n \to 0$ and consider the two sequences of points

$$P_n = (1+l_n)\hat{y}_1 + \hat{y}_2\,, \qquad Q_n = \hat{y}_1 + (1+l_n)\hat{y}_2\,.$$

We have that $P_n \to y$, $Q_n \to y$ and

$$\begin{aligned} c_1(P_n) &= 1\,, & \qquad c_1(Q_n) &= 0\,, \\ c_2(P_n) &= 0\,, & \qquad c_2(Q_n) &= 1\,. \end{aligned}$$

By continuity, for every $\tau \in [0,1]$ and every sufficiently large n, there exists a $\Lambda_n \in [0,1]$ such that

$$c_1\big(\Lambda_n P_n + (1-\Lambda_n)Q_n\big) = \tau\,, \quad c_2\big(\Lambda_n P_n + (1-\Lambda_n)Q_n\big) = \sqrt{1-\tau^2}\,.$$

Since $\Lambda_n P_n + (1-\Lambda_n)Q_n \to y$, it follows that

$$\tau\frac{\hat{y}_1}{|\hat{y}_1|} - \sqrt{1-\tau^2}\,\frac{\hat{y}_2}{|\hat{y}_2|} \in \alpha_F(y)\,, \qquad \text{for every } \tau \in [0,1]\,.$$

So (9) is verified, and the proof is easily completed. $\square$

The avoiding cones condition (8) of Corollary 4 is visualized in Figure 3(a). It can be restated as

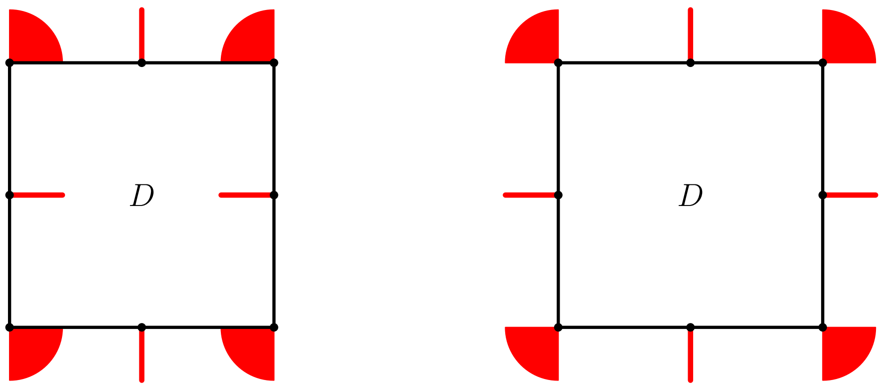
$$\vartheta(x,y) \notin \begin{cases} \{0\} \times -\mathcal{N}_{D_2}(\hat{y}_2)\,, & \text{if } y \in \operatorname{int} D_1 \times \partial D_2\,, \\ \mathcal{N}_{D_1}(\hat{y}_1) \times -\mathcal{N}_{D_2}(\hat{y}_2)\,, & \text{if } y \in \partial D_1 \times \partial D_2\,, \\ \mathcal{N}_{D_1}(\hat{y}_1) \times \{0\}\,, & \text{if } y \in \partial D_1 \times \operatorname{int} D_2\,. \end{cases}$$

Example 5. We take $D = [-1,1] \times [-1,1]$ and define the Hamiltonian function $H\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ as

$$H(x_1,x_2,y_1,y_2) = y_1^2 - y_2^2 - y_2\sin(2\pi y_1)\,.$$

The map $\vartheta(x,y) = (\vartheta_1(x,y), \vartheta_2(x,y))$ is such that

$$\begin{aligned} \vartheta_1(x,y) &= T\frac{\partial H}{\partial y_1}(x,y) = 2T[y_1 - \pi y_2\cos(2\pi y_1)]\,, \\ \vartheta_2(x,y) &= T\frac{\partial H}{\partial y_2}(x,y) = -T[2y_2 + \sin(2\pi y_1)]\,. \end{aligned}$$



(a) *Corollary 4.* (b) *Corollary 6.*

Figure 3: Visualization of $\mathcal{A}_F(y)$ in the framework of Corollaries 4 and 6, for $N_1 = N_2 = 1$.

As illustrated in Figure 4, we see that the avoiding cones condition (8) is satisfied, for $N_1 = N_2 = 1$, cf. also Figure 3(a). The same property is inherited by all the sufficiently small perturbations of H, satisfying the regularity and periodicity assumptions of Theorem 1.

With a similar approach, we can also study the following situation.

Corollary 6. *Let $D = D_1 \times D_2$, with $D_1 = \mathcal{B}^{N_1}[0,1]$ and $D_2 = \mathcal{B}^{N_2}[0,1]$, and assume that*

$$\vartheta(x,y) \notin \mathcal{N}_D(y)\,, \quad \textit{for every } (x,y) \in \mathbb{R}^N \times \partial D\,. \tag{10}$$

Then, the same conclusion of Theorem 1 holds.

Proof. We define the function $h\colon \mathbb{R}^N \to \mathbb{R}$ as

$$h(y) = \gamma(|\hat{y}_1|)\,|\hat{y}_1|^2 + \gamma(|\hat{y}_2|)\,|\hat{y}_2|^2\ .$$

The same arguments used in the proof of Corollary 4 can be successfully applied, simply changing the sign in front of the coefficient C_2. $\square$

We notice that, in Corollary 6, condition (10) can be replaced by

$$\vartheta(x,y) \notin -\mathcal{N}_D(y)\,, \quad \text{for every } (x,y) \in \mathbb{R}^N \times \partial D\,,$$

by simply changing in the proof the sign of the potential h.

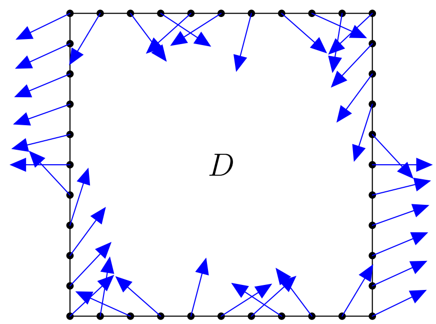


Figure 4: Normalized Poincaré map for the Hamiltonian system of Example 5.

Combining the ideas of the previous two corollaries, let us consider the decomposition $\mathbb{R}^N = \mathbb{R}^{N^+} \times \mathbb{R}^{N^-}$, with $N^+ = N_1^+ + \cdots + N_n^+$ and $N^- = N_1^- + \cdots + N_m^-$, all summands being non-negative integers. For every $y \in \mathbb{R}^{N^+} \times \mathbb{R}^{N^-}$, we write $y = \hat{y}^+ + \hat{y}^-$, with $\hat{y}^+ \in \mathbb{R}^{N^+} \times \{0\}$ and $\hat{y}^- \in \{0\} \times \mathbb{R}^{N^-}$. We thus obtain the following more general result.

Corollary 7. *Let $D = D^+ \times D^-$, with*

$$D^+ = \prod_{i=1}^{n} \mathcal{B}^{N_i^+}[0,1]\,, \qquad\qquad D^- = \prod_{i=1}^{m} \mathcal{B}^{N_i^-}[0,1]\,.$$

Assume that, for every $(x,y) \in \mathbb{R}^N \times \partial D$,

$$\vartheta(x,y) \notin \begin{cases} \{0\} \times -\mathcal{N}_{D^-}(\hat{y}^-)\,, & \text{if } y \in \operatorname{int} D^+ \times \partial D^-\,, \\ \mathcal{N}_{D^+}(\hat{y}^+) \times -\mathcal{N}_{D^-}(\hat{y}^-)\,, & \text{if } y \in \partial D^+ \times \partial D^-\,, \\ \mathcal{N}_{D^+}(\hat{y}^+) \times \{0\}\,, & \text{if } y \in \partial D^+ \times \operatorname{int} D^-\,. \end{cases}$$

Then, the same conclusion of Theorem 1 holds.

2.3 Sets diffeomorphic to a ball

We now show how to apply our results to sets $\mathcal{D}$ which are diffeomorphic to a ball.

Let $\mathcal{D} \subset \mathbb{R}^N$ be a compact set, and let $\mathcal{D}^+$ be a relatively open subset of $\partial\mathcal{D}$. We define $\mathcal{D}^- = \partial\mathcal{D} \setminus \overline{\mathcal{D}^+}$ and $\mathcal{D}^0 = \partial\mathcal{D} \setminus (\mathcal{D}^+ \cup \mathcal{D}^-)$.

Definition 8. We say that the couple $(\mathcal{D}, \mathcal{D}^+)$ is *twist-generating* if there exist two regular symmetric matrices $\mathbb{B}$, $\mathbb{B}_\infty$, with $\mathbb{B}$ of the form (3), and a $\mathcal{C}^\infty$-smooth diffeomorphism $\Psi\colon \mathbb{R}^N \to \mathbb{R}^N$, such that

- $\Psi'(w) = \mathbb{B}_\infty$ for $|w|$ sufficiently large;
- $\Psi(\mathcal{D}) = \mathcal{B}^N[0,1]$;
- $\Psi(\mathcal{D}^+) = \{y : |y| = 1, \langle y, \mathbb{B}y\rangle > 0\}$.

Note that if $(\mathcal{D}, \mathcal{D}^+)$ is twist-generating, then $\mathcal{D}$ has smooth boundary and therefore, for every $w \in \partial\mathcal{D}$, the outer normal cone $\mathcal{N}_\mathcal{D}(w)$ is well defined, and it is the half-line generated by the outer unit normal $\nu_\mathcal{D}(w)$. Moreover, for every point $w \in \mathcal{D}^0$, we can define the vector

$$\sigma(w) = [\Psi'(w)]^T \mathbb{B}\Psi(w)\,.$$

We see that $\sigma(w)$ is orthogonal to $\mathcal{D}^0$ and to $\nu_\mathcal{D}(w)$ (therefore tangent to $\mathcal{D}$).

Corollary 9. *If $(\mathcal{D}, \mathcal{D}^+)$ is twist-generating and, for every $(x, w) \in \mathbb{R}^N \times \partial\mathcal{D}$,*

$$\vartheta(x,w) \notin \begin{cases} \mathcal{N}_\mathcal{D}(w)\,, & \text{if } w \in \mathcal{D}^+, \\ \{\mu_1 \nu_\mathcal{D}(w) + \mu_2 \sigma(w) : \mu_1 \in \mathbb{R}\,, \mu_2 \geq 0\}\,, & \text{if } w \in \mathcal{D}^0, \\ -\mathcal{N}_\mathcal{D}(w)\,, & \text{if } w \in \mathcal{D}^-, \end{cases}$$

then the same conclusion of Theorem 1 holds.

Proof. We consider the function

$$h_A(y) = \gamma(|y| - 1)\,\langle \mathbb{B}y, y\rangle\,,$$

as introduced in Section 2.1, and define $h\colon \mathbb{R}^N \to \mathbb{R}$ as

$$h(w) = h_A(\Psi(w))\,.$$

All the properties required to $F = \nabla h$ are inherited from h_A, and Theorem 1 applies. □

We observe that, in the case $\mathcal{D}^+ = \partial\mathcal{D}$, implying $\mathbb{B} = I$, we have recovered exactly the twist condition (T3).

The same line of reasoning holds if we want to generalize other situations, such as those of Section 2.2, by the use of a diffeomorphism. We omit the details, for briefness.

2.4 Comparison with twist conditions in the previous literature

We now show how the following result obtained in [11, 12] can be proved using Theorem 1.

Corollary 10 (Fonda–Ureña)**.** *Let $D \subset \mathbb{R}^N$ be a $\mathcal{C}^\infty$-smooth strongly convex body, and assume that at least one of the twist conditions* (T1), (T2) *or* (T3) *holds. Then, the same conclusion of Theorem 1 holds.*

Proof. We denote by $\pi_D \colon \mathbb{R}^N \to \mathbb{R}^N$ the projection on the convex set D. Assume that (T1) holds. Let $\widetilde{\mathcal{F}}_1 \colon \mathbb{R}^N \setminus D \to \mathbb{R}^N$ be the map defined as

$$\widetilde{\mathcal{F}}_1(y) = \mathbb{B}\nu_D(\pi_D(y)) .$$

We define $h \colon \mathbb{R}^N \to \mathbb{R}$ by

$$h(y) = \begin{cases} 0 , & \text{if } y \in D , \\ \gamma(|y - \pi_D(y)|)\langle \mathbb{B}(y - \pi_D(y)), y - \pi_D(y)\rangle , & \text{if } y \in \mathbb{R}^N \setminus D . \end{cases}$$

It is clear that h is a $\mathcal{C}^\infty$-smooth function. The function $F = \nabla h$ satisfies (1) with $\mathbb{S} = 2\mathbb{B}$, and $F^{-1}(0) = D$, while

$$\left\langle F(y), \widetilde{\mathcal{F}}_1(y) \right\rangle > 0 , \qquad \text{for every } y \in \mathbb{R}^N \setminus D .$$

(For the details, see [11, Sec. 4].) This implies that

$$\langle \upsilon, \mathbb{B}\nu_D(y)\rangle \geq 0 , \qquad \text{for every } y \in \partial D \text{ and } \upsilon \in \mathcal{A}_F(y) .$$

Combining this with (T1), we have (AC).

Assume now instead that (T2) holds. Without loss of generality, we set $d_0 = 0$ and we define $\widetilde{\mathcal{F}}_2 \colon \mathbb{R}^N \setminus D \to \mathbb{R}^N$ as

$$\widetilde{\mathcal{F}}_2(y) = \mathbb{B}y .$$

When $\mathbb{B}$ is orthogonal, we define $h \colon \mathbb{R}^N \to \mathbb{R}$ by

$$h(y) = \begin{cases} 0 , & \text{if } y \in D , \\ \gamma(|y - \pi_D(y)|)\langle \mathbb{B}y, y - \pi_D(y)\rangle , & \text{if } y \in \mathbb{R}^N \setminus D . \end{cases}$$

The function $F = \nabla h$ satisfies (1) and $F^{-1}(0) = D$, while

$$\left\langle F(y), \widetilde{\mathcal{F}}_2(y) \right\rangle > 0\,, \qquad \text{for every } y \in \mathbb{R}^N \setminus D\,.$$

The conclusion (AC) then follows as above. In the case of a general involutory matrix $\mathbb{B}$, we can reduce to the above situation by a change of basis, since $\mathbb{B}$ is diagonalizable (see [11, Sec. 4]).

Finally, assume that (T3) holds. We define $h\colon \mathbb{R}^N \to \mathbb{R}$ by

$$h(y) = \gamma(|y - \pi_D(y)|)|y - \pi_D(y)|^2\,.$$

The conclusion follows, similarly as above. $\square$

We remark that, in general, assumptions (T1) and (T2) are strictly stronger than the avoiding cones condition (AC), as shown in the following example.

Example 11. Let us set $D = \mathcal{B}^3[0,1]$ and $\mathbb{B} = \mathrm{diag}(1,1,-1)$. We want to compare the avoiding cones condition (AC) induced by $F = \nabla h$, with h as in (5), with the conditions (T1) and (T2), for $d_0 = 0$, which are equivalent, in this situation. For every $y \in \partial D$, if $\langle y, \mathbb{B}y\rangle > 0$ (resp. $\langle y, \mathbb{B}y\rangle < 0$), the avoiding cones condition (AC) requires that $\vartheta(x,y)$ is not contained in the outer (resp. inner) normal cone of D in y, a half-line, whereas (T1) requires that $\vartheta(x,y)$ avoids an entire half-space containing this half-line. If instead $\langle y, \mathbb{B}y\rangle = 0$, then the avoiding cones condition (AC) requires that $\vartheta(x,y)$ avoids the half-plane generated by $\mathbb{B}y$ and $\pm\nu_D(y)$, whereas (T1) requires that $\vartheta(x,y)$ avoids a half-space that includes that half-plane.

3 Proof of Theorem 1

The proof follows the one given in [13].

Let us recall that $\mathcal{Z}\colon \mathbb{R} \times \mathbb{R}^{2N} \to \mathbb{R}^{2N}$ is the $\mathcal{C}^\infty$-map associating to each couple (t,ζ) the value at time t of the unique solution $\mathcal{Z}(\cdot,\zeta)$ of (HS) satisfying $\mathcal{Z}(0,\zeta) = \zeta$. For $\zeta \in \mathbb{R}^{2N}$, we write $\zeta = (\xi,\eta)$, with $\xi = (\xi_1,\ldots,\xi_N) \in \mathbb{R}^N$ and $\eta = (\eta_1,\ldots,\eta_N) \in \mathbb{R}^N$.

Since D is a compact set and the Hamiltonian $H(t,x,y)$ is 2π-periodic in the variables x_i, the continuous image by $\mathcal{Z}$ of $[0,T]\times(\mathbb{R}^N/2\pi\mathbb{Z}^N)\times D$ is contained in $(\mathbb{R}^N/2\pi\mathbb{Z}^N)\times B_r$, for some open ball B_r. Thus, after multiplying H by a smooth cutoff function of y, there is no loss of generality in assuming that there is some $R>r$ for which

$$H(t,x,y)=0\,,\qquad \text{if } |y|\geq R\,.$$

Consequently, there is some constant $c>0$ such that

$$\left|\frac{\partial H}{\partial y}(t,x,y)\right|<c\,,\qquad \text{for every }\ (t,x,y)\in\mathbb{R}\times\mathbb{R}^N\times\mathbb{R}^N\,.$$

As a consequence, we will have that

$$|\vartheta(\xi,\eta)|<cT\,,\qquad \text{for every } \xi,\eta\in\mathbb{R}^N\,. \tag{11}$$

For any t, we write $\mathcal{Z}_t:=\mathcal{Z}(t,\cdot)\colon\mathbb{R}^{2N}\to\mathbb{R}^{2N}$. The following properties hold.

(i) $\mathcal{Z}_0$ *is the identity map in* $\mathbb{R}^{2N}$*;*

(ii) $\mathcal{Z}_t(\zeta+p)=\mathcal{Z}_t(\zeta)+p$, *if* $p\in 2\pi\mathbb{Z}^N\times\{0\}$*;*

(iii) $\mathcal{Z}(t,\xi,\eta)=(\xi,\eta)$, *if* $|\eta|\geq R$*;*

(iv) each $\mathcal{Z}_t$ *is a symplectic* $\mathcal{C}^\infty$*-diffeomorphism of* $\mathbb{R}^{2N}$ *on itself.*

This last property is standard in Hamiltonian dynamics. Nevertheless, let us provide a brief proof of it, for the reader's convenience. For any fixed ζ, let

$$b(t)=\frac{\partial\mathcal{Z}}{\partial\zeta}(t;\zeta)\,,\quad A(t)=\frac{\partial^2 H}{\partial\zeta^2}(t;\mathcal{Z}(t;\zeta))\,.$$

Since $\mathcal{Z}(\cdot;\zeta)$ is a solution of the Hamiltonian system, differentiating we have

$$J\dot b(t)=A(t)b(t)\,,$$

and, since $A(t)^*=A(t)$, $J^*=-J$, and $JJ=-I$,

$$\frac{d}{dt}(b(t)^*Jb(t))=\dot b(t)^*Jb(t)+b(t)^*J\dot b(t)=0\,,$$

for every t. Then, b^*Jb is constant, and since $b(0)=I$, we conclude that $b(t)^*Jb(t)=J$, for every t, thus proving (iv).

By the use of the Ascoli–Arzelà Theorem, we can find some constant $\varepsilon \in]0,1[$ such that

$$\vartheta(\xi,\eta) \notin \{\mu F(\eta) : \mu \geq 0\}\,, \quad \text{if } \; 0 < |F(\eta)| < \varepsilon\,. \tag{12}$$

Recalling that $F = \nabla h$ and that (1) holds, we can assume without loss of generality that

$$h(y) = \tfrac{1}{2}\langle \mathbb{S}y, y\rangle\,, \quad \text{when } |y| \geq K\,.$$

Indeed, choosing $\widetilde{R}$ large enough and defining

$$\widetilde{F}(x) = \begin{cases} F(x)\,, & \text{if } |x| \leq \widetilde{R}\,, \\ F(x) + \gamma(|x| - \widetilde{R})(\mathbb{S}x - F(x))\,, & \text{if } \widetilde{R} \leq |x| \leq \widetilde{R} + 1\,, \\ \mathbb{S}x\,, & \text{if } |x| \geq \widetilde{R} + 1\,, \end{cases}$$

we will have that $D \subseteq \mathcal{B}^N(0,\widetilde{R})$ and $\widetilde{F}^{-1}(0) = D$, while the cones $\mathcal{A}_F(y)$ will not be changed for $y \in \mathcal{B}^N(0,\widetilde{R})$.

We define the function $\mathfrak{R}\colon \mathbb{R}^{2N} \to \mathbb{R}$ as

$$\mathfrak{R}(\xi,\eta) := -\frac{c}{\varepsilon}\, h(\eta)\,,$$

the function $R\colon [0,T] \times \mathbb{R}^{2N} \to \mathbb{R}$ by

$$R(t,z) := \mathfrak{R}(\mathcal{Z}_t^{-1}(z))\,,$$

and the modified Hamiltonian $\widetilde{H}\colon [0,T] \times \mathbb{R}^{2N} \to \mathbb{R}$ as

$$\widetilde{H}(t,z) := H(t,z) + R(t,z)\,.$$

It is a $\mathcal{C}^\infty$-smooth function, and satisfies the following properties:

(j) $\widetilde{H}(t,z+p) = \widetilde{H}(t,z)$, *if* $p \in 2\pi\mathbb{Z}^N \times \{0\}$;

(jj) $\widetilde{H}(t,x,y) = \frac{1}{2}\langle \widetilde{\mathbb{S}}y, y\rangle$, *if* $|y| \geq R$, *where* $\widetilde{\mathbb{S}} = -(c/\varepsilon)\mathbb{S}$;

(jjj) $\widetilde{H}$ *and* H *coincide on the set*

$$\big\{(t, \mathcal{Z}(t,\xi,\eta)) : t \in [0,T],\, \xi \in \mathbb{R}^N, \eta \in D\big\}\,.$$

We consider the modified Hamiltonian system

$$\dot{z} = J\nabla \widetilde{H}(t,z)\,, \tag{$\widetilde{\mathrm{HS}}$}$$

and look for solutions satisfying $z(0) = z(T)$. These will be obtained as critical points of a suitably defined functional.

Consider the Hilbert space $H^{1/2}_T$, whose elements are those functions $z \in L^2(0,T;\mathbb{R}^{2N})$, extended by T-periodicity (in the a.e. sense), with the property that, writing the associated Fourier series

$$z(t) \sim \sum_{k=-\infty}^{+\infty} a_k e^{2\pi k i t/T},$$

one has that

$$\sum_{k=-\infty}^{+\infty} (1+|k|)|a_k|^2 < +\infty\,.$$

We refer to [16, Section 3.3] for the main properties of $H^{1/2}_T$. The functions in $H^{1/2}_T$ are not necessarily continuous, but their restriction to $[0,T]$ belongs to $L^p(0,T;\mathbb{R}^{2N})$, for every $p \in [1,+\infty[$. On the other hand, let H^1_T be the space of those functions $z \in H^{1/2}_T$ for which

$$\sum_{k=-\infty}^{+\infty} (1+|k|^2)|a_k|^2 < +\infty\,.$$

These are absolutely continuous T-periodic functions. In particular, they are such that $z(0)=z(T)$.

We define an auxiliary function $\widehat{H}\colon \mathbb{R}\times\mathbb{R}^{2N} \to \mathbb{R}$ as follows:

$$\widehat{H}(t,z) = \widetilde{H}(\tau,z)\,, \qquad \text{with } \tau \in [0,T[\text{ and } t = \tau + kT, \text{ for some } k \in \mathbb{Z}\,.$$

By construction, $\widehat{H}(t,z)$ is T-periodic in t, but not necessarily continuous. In view of *(j)* and *(jj)* above, it is possible to define the functional $\varphi\colon H^{1/2}_T \to \mathbb{R}$ as

$$\varphi(z) = \int_0^T \Big[\tfrac{1}{2}\langle J\dot z(t), z(t)\rangle + \widehat{H}(t,z(t))\Big]\,dt\,.$$

It can be seen that it is continuously differentiable, and its critical points correspond to the weak T-periodic solutions of

$$\dot z = J\nabla \widehat{H}(t,z)\,. \tag{13}$$

Let z be a critical point of φ. Following [21], we will show that the restriction of z to the closed interval $[0,T]$ is a classical solution of $(\widetilde{\mathrm{HS}})$ satisfying $z(0)=z(T)$.

Since z is a critical point of φ, we have $\langle \nabla\varphi(z), w\rangle = 0$, for every $w \in H_T^{1/2}$. Then, taking w in H_T^1, we have

$$\int_0^T \Big[\langle z(t), J\dot w(t)\rangle + \langle \nabla \widehat H(t, z(t)), w(t)\rangle\Big]\, dt = 0\,. \tag{14}$$

In particular, taking as w the constant functions with all zero components except one of them, we deduce that

$$\int_0^T \nabla \widehat H(t, z(t))\, dt = 0\,.$$

Hence, denoting by $[\cdot]$ the mean of a function defined on $[0,T]$,

$$[J\nabla \widehat H(\cdot, z(\cdot))] = \frac{1}{T}\int_0^T J\nabla \widehat H(t, z(t))\, dt = 0\,. \tag{15}$$

It is known that, for every fixed vector $u \in \mathbb{R}^{2N}$ and every function $g \in L^2(0,T;\mathbb{R}^{2N})$, with $[g]=0$, there is a unique $\upsilon \in H_T^1$ satisfying $[\upsilon] = u$ and $\dot\upsilon = g$ in $L^2(0,T;\mathbb{R}^{2N})$. Hence, from (15) we deduce that there is a unique function $\upsilon \in H_T^1$ such that $[\upsilon] = [z]$ and $\dot\upsilon = J\nabla\widehat H(\cdot, z(\cdot))$ in $L^2(0,T;\mathbb{R}^{2N})$. Therefore, for any $w \in H_T^1$, integrating by parts and using (14),

$$\int_0^T \langle \upsilon, J\dot w\rangle = -\int_0^T \langle \dot\upsilon, Jw\rangle = -\int_0^T \langle \nabla\widehat H(t, z(t)), w(t)\rangle\, dt = \int_0^T \langle z, J\dot w\rangle\,.$$

We deduce that $\upsilon = z$ in H_T^1, and

$$\dot z(t) = J\nabla \widehat H(t, z(t))\,, \tag{16}$$

for almost every $t \in [0,T]$. Moreover, since z belongs to H_T^1, it is continuous, hence $\dot z$ has to be continuous, too, and z satisfies (16) *for every* $t \in [0,T[$. Furthermore, $z(0) = z(T)$. Hence, by continuity, z is a classical solution of $(\widetilde{\mathrm{HS}})$ on $[0,T]$: when restricted to that interval, it belongs to $\mathcal{C}^1([0,T],\mathbb{R}^{2N})$. A bootstrap argument now shows that $z \in \mathcal{C}^\infty([0,T],\mathbb{R}^{2N})$.

For any $z(t) = (x(t), y(t))$ in $H_T^{1/2}$, we write $x(t) = \bar x + \tilde x(t)$, where $\bar x = [x] \in \mathbb{R}^N$. We thus have the decomposition $H_T^{1/2} = \mathbb{R}^N \oplus E$, where E is a Hilbert space. By (j), we can identify $\bar x \in \mathbb{R}^N$ with its projection on the N-torus $\mathbb{T}^N$ and define the functional $\widetilde\varphi\colon \mathbb{T}^N \times E \to \mathbb{R}$ as

$$\widetilde\varphi(\bar x, (\tilde x, y)) = \varphi(\bar x + \tilde x, y)\,.$$

By [22, Theorem 4.2] and [23, Theorem 8.1], the functional $\widetilde{\varphi}$ has of at least $N+1$ critical points, and 2^N of them if all its critical points are nondegenerate. As we saw above, these critical points correspond to geometrically distinct solutions of $(\widetilde{\mathrm{HS}})$ belonging to $\mathcal{C}^\infty([0,T],\mathbb{R}^{2N})$, satisfying $z(0)=z(T)$.

As a consequence of *(jjj)*, the Hamiltonian systems (HS) and $(\widetilde{\mathrm{HS}})$ have the same solutions $z(t)=(x(t),y(t))$, with $t\in[0,T]$, departing with $y(0)\in D$. Thus, in order to complete the proof of Theorem 1, it will suffice to check that $(\widetilde{\mathrm{HS}})$ does not have solutions $z(t)=(x(t),y(t))$, satisfying $z(0)=z(T)$, departing with $y(0)\notin D$.

We argue by contradiction, and assume that such a solution $z(t)$ exists. Let us define the $\mathcal{C}^\infty$-function $\zeta\colon[0,T]\to\mathbb{R}^{2N}$ by

$$\zeta(t):=\mathcal{Z}_t^{-1}(z(t))\,.$$

Differentiating in the equality $z(t)=\mathcal{Z}(t,\zeta(t))$, we find

$$\dot z(t)=\frac{\partial\mathcal{Z}}{\partial t}(t,\zeta(t))+\frac{\partial\mathcal{Z}}{\partial\zeta}(t,\zeta(t))\dot\zeta(t)\,,$$

so that

$$\frac{\partial\mathcal{Z}}{\partial\zeta}(t,\zeta(t))\dot\zeta(t)=J\nabla\widetilde{H}(t,z(t))-J\nabla H(t,z(t))=J\nabla R(t,z(t))\,. \tag{17}$$

By (iv), $\mathcal{Z}_t$ is symplectic, so

$$\frac{\partial\mathcal{Z}}{\partial\zeta}(t,\zeta(t))^*J\,\frac{\partial\mathcal{Z}}{\partial\zeta}(t,\zeta(t))=J\,,\quad\text{for every } t\in\mathbb{R}\,.$$

Hence, if we multiply both sides of (17) by $-J(\partial\mathcal{Z}/\partial\zeta)^*J$, we get

$$\dot\zeta(t)=J\frac{\partial\mathcal{Z}}{\partial\zeta}(t,\zeta(t))^*\,\nabla R(t,z(t))=J\nabla\mathfrak{R}(\zeta(t))\,,$$

the last equality coming from the fact that $R(t,\mathcal{Z}(t,\zeta))=\mathfrak{R}(\zeta)$. Then, recalling that $\zeta(t)=(\xi(t),\eta(t))$,

$$\dot\xi(t)=-\frac{c}{\varepsilon}\,F(\eta(t))\,,\qquad\dot\eta(t)=0\,,$$

and consequently, by *(i)*, writing $z(t)=(x(t),y(t))$,

$$\eta(t)=\eta(0)=y(0)\,,\qquad\xi(t)=x(0)-\frac{ct}{\varepsilon}\,F(y(0))\,,$$

for every $t \in [0,T]$, i.e.,

$$\zeta(t) = \Big(x(0) - \frac{ct}{\varepsilon}\, F(y(0))\,, y(0)\Big)\,.$$

Being $z(t) = \mathcal{Z}(t, \zeta(t))$ and $\mathcal{Z}_T = \mathcal{P}$, we thus have

$$z(T) = \mathcal{P}\Big(x(0) - \frac{cT}{\varepsilon}\, F(y(0))\,, y(0)\Big),$$

and in particular

$$x(T) = x(0) - \frac{cT}{\varepsilon}\, F(y(0)) + \vartheta\Big(x(0) - \frac{cT}{\varepsilon}\, F(y(0)), y(0)\Big)\,.$$

In order to obtain the desired contradiction, we shall show that $x(T) \neq x(0)$, i.e.,

$$\vartheta\Big(x(0) - \frac{cT}{\varepsilon}\, F(y(0)), y(0)\Big) \neq \frac{cT}{\varepsilon}\, F(y(0))\,. \tag{18}$$

We distinguish two situations, according to the initial point of the solution. If $0 < |F(y(0))| < \varepsilon$, by (12) we have

$$\vartheta\Big(x(0) - \frac{cT}{\varepsilon}\, F(y(0)), y(0)\Big) \notin \{\alpha F(y(0)) : \alpha \geq 0\}\,,$$

implying (18). On the other hand, if $|F(y(0))| \geq \varepsilon$, by (11) we get

$$\Big|\vartheta\Big(x(0) - \frac{cT}{\varepsilon}\, F(y(0)), y(0)\Big)\Big| < cT \leq \Big|\frac{cT}{\varepsilon}\, F(y(0))\Big|\,,$$

implying (18), again. The proof is thus completed.

4 Final remarks

With the same strategy adopted for Theorem 1, we can prove the following more general result.

As before, we assume the Hamiltonian function $H\colon \mathbb{R} \times \mathbb{R}^{2N} \to \mathbb{R}$ to be $\mathcal{C}^\infty$-smooth, and T-periodic in its first variable t. Let M be an integer such that $0 \leq M < N$, and assume that $H(t,x,y)$ is 2π-periodic in $x_1, \dots, x_N$ and in $y_1, \dots, y_M$. We still write as in (2) the Poincaré map $\mathcal{P}$ associated to the system (HS), and we define the projection $\pi\colon \mathbb{R}^N \to \mathbb{R}^{N-M}$ as

$$\pi(y_1, \dots, y_N) = (y_{M+1}, \dots, y_N)\,.$$

Theorem 12. *Let* $F = \nabla h\colon \mathbb{R}^{N-M} \to \mathbb{R}^{N-M}$ *be a* $\mathcal{C}^\infty$*-smooth function for which there are two constants* $K > 0$ *and* $C > 0$ *and a regular symmetric* $(N-M) \times (N-M)$ *matrix* $\mathbb{S}$ *such that*

$$|F(w) - \mathbb{S}w| \leq C\,, \quad \textit{when } |w| \geq K,$$

and set $D := F^{-1}(0)$*. If*

$$\pi(\vartheta(x,y)) \notin \mathcal{A}_F(\pi(y))\,, \textit{ for every } (x,y) \in \mathbb{R}^{N+M} \times \partial D\,,$$

then $\mathcal{P}$ *has at least* $N + M + 1$ *geometrically distinct fixed points, all in* $\mathbb{R}^{N+M} \times D$*. Moreover, if all its fixed points are non degenerate, then there are at least* 2^{N+M} *of them.*

Proof. The proof is similar to that of Theorem 1, with the following changes. The construction is based on the $y_{M+1}, \dots, y_N$ coordinates, in the sense that first we assume $H(t,x,y) = 0$ if $|\pi(y)| \geq R$, and later we use the function

$$\mathfrak{R}(\xi, \eta) = -\frac{c}{\varepsilon} h(\pi(\eta))\,,$$

to define the modified Hamiltonian.

Then, when looking for the critical points of the functional φ, we use the decomposition $H_T^{1/2} = \mathbb{R}^{N+M} \oplus \widehat{E}$, where R^{N+M} is the subspace associated to the constant functions with values in $R^{N+M} \times \{0_{\mathbb{R}^{N-M}}\}$, and $\widehat{E}$ is a Hilbert space. The projection $\mathbb{R}^{N+M} \to \mathbb{T}^{N+M}$ will lead to a functional $\widetilde{\varphi}\colon \mathbb{T}^{N+M} \times \widehat{E} \to \mathbb{R}$, having at least $N+M+1$ critical points, or at least 2^{N+M} of them if all critical points are non degenerate. With the same line of reasoning used for Theorem 1, it can be shown that such critical points correspond to geometrically distinct solutions of (HS). □

We notice that, if we extend Theorem 12 to the case $M = N$, no avoiding cones condition is required any longer and we recover a celebrated result on the existence of fixed points for a symplectic map on the torus, as conjectured by Arnold and proved by Conley and Zehnder [3]. Thus Theorem 12 covers the intermediate cases between this result and Theorem 1, corresponding to $M = 0$. We finally notice that we could have assumed the periodicity along a different basis than the usual one in $\mathbb{R}^{N+M}$. Similar situations have also been considered in [2, 4, 7, 14, 18].

We have assumed the Hamiltonian function H to be $\mathcal{C}^\infty$-smooth, not to care about the technical problems arising when less regularity is required. However, as already mentioned in the Introduction, using the methods introduced in [11], it is sufficient to assume $H(t, z)$ to be continuous, with a continuous gradient $\nabla H(t, z)$ with respect to z. Notice also that, slightly modifying the proof, it would have been sufficient to assume the Poincaré map $\mathcal{P}$ to be defined only on the set D.

The results in [11, 12] have already found several applications to some concrete periodic problems: differential systems associated with relativistic or mean-curvature operators were already studied in [11], superlinear systems were treated in [9, 11], systems with singularities in [10], and perturbations of Hamiltonian systems in [5]. Other applications are next to appear. We are confident that our avoiding cones condition, besides its theoretical importance, will be successfully emplemented in the study of more specific problems.

Acknowledgements. The authors have been partially supported by the Gruppo Nazionale per l'Analisi Matematica, la Probabilità e le loro Applicazioni (GNAMPA) of the Istituto Nazionale di Alta Matematica (INdAM).

References

[1] V.I. Arnold, Mathematical Methods in Classical Mechanics, Springer, Berlin, 1978.

[2] K.C. Chang, On the periodic nonlinearity and the multiplicity of solutions, Nonlinear Anal. 13 (1989), 527–537.

[3] C.C. Conley and E.J. Zehnder, The Birkhoff – Lewis fixed point theorem and a conjecture of V.I. Arnold, Invent. Math. 73 (1983), 33–49.

[4] P.L. Felmer, Periodic solutions of spatially periodic Hamiltonian systems, J. Differential Equations 98 (1992), 143–168.

[5] A. Fonda, M. Garrione and P. Gidoni, Periodic perturbations of Hamiltonian systems, Adv. Nonlinear Anal., to appear. (doi: 10.1515/anona-2015-0122)

[6] A. Fonda and P. Gidoni, Generalizing the Poincaré–Miranda Theorem: the avoiding cones condition, Ann. Mat. Pura Appl., to appear. (doi: 10.1007/s10231-015-0519-6)

[7] A. Fonda and J. Mawhin, Multiple periodic solutions of conservative systems with periodic nonlinearity, Differential equations and applications, Vol. I, II (Columbus, OH, 1988), 298–304, Ohio Univ. Press, Athens, OH, 1989.

[8] A. Fonda, M. Sabatini, and F. Zanolin, Periodic solutions of perturbed Hamiltonian systems in the plane by the use of the Poincaré–Birkhoff theorem, Topol. Methods Nonlinear Anal. 40 (2012), 29–52.

[9] A. Fonda and A. Sfecci, Periodic solutions of weakly coupled superlinear systems, J. Differential Equations 260 (2016), 2150–2162.

[10] A. Fonda and A. Sfecci, Multiple periodic solutions of Hamiltonian systems confined in a box, preprint (2014), available at: *www.dmi.units.it/∼fonda/2014_Fonda-Sfecci_box.pdf.*

[11] A. Fonda and A.J. Ureña, On the higher dimensional Poincaré – Birkhoff theorem for Hamiltonian flows. 1. The indefinite twist, preprint (2013), available at: *www.dmi.units.it/∼fonda/2013_Fonda-Urena.pdf.*

[12] A. Fonda and A.J. Ureña, On the higher dimensional Poincaré – Birkhoff theorem for Hamiltonian flows. 2. The avoiding rays condition, preprint (2014), available at: *www.dmi.units.it/∼fonda/2014_Fonda-Urena.pdf.*

[13] A. Fonda and A.J. Ureña, A higher-dimensional Poincaré – Birkhoff theorem without monotone twist, C. R. Math. Acad. Sci. Paris, to appear. (preprint available at: *www.dmi.units.it/∼fonda/2015_Fonda-Urena.pdf*).

[14] G. Fournier, D. Lupo, M. Ramos and M. Willem, Limit relative category and critical point theory, in: Dynamics reported. Expositions in Dynamical Systems, vol. 3, 1–24, Springer, Berlin, 1994.

[15] C. Golé, Symplectic Twist Maps. Global Variational Techniques, Adv. Ser. Nonlin. Dyn. 18, World Scientific, River Edge, 2001.

[16] H. Hofer and E. Zehnder, Symplectic invariants and Hamiltonian dynamics, Birkhäuser, Basel, 1994.

[17] P. Le Calvez, About Poincaré-Birkhoff theorem, Publ. Mat. Urug. 13 (2011), 61–98.

[18] J.Q. Liu, A generalized saddle point theorem, J. Differential Equations 82 (1989), 372–385.

[19] J. Moser and E.J. Zehnder, Notes on Dynamical Systems, Courant Lecture Notes in Mathematics 12, Amer. Math. Soc., Providence, 2005.

[20] H. Poincaré, Sur un théorème de géométrie, Rend. Circ. Mat. Palermo 33 (1912), 375–407.

[21] P.H. Rabinowitz, Minimax methods in Critical Point Theory with Applications to Differential Equations, CBMS 65, Amer. Math. Soc., Providence, 1984.

[22] A. Szulkin, A relative category and applications to critical point theory for strongly indefinite functionals, Nonlinear Anal. 15 (1990), 725–739.

[23] A. Szulkin, Cohomology and Morse theory for strongly indefinite functionals, Math. Z. 209 (1992), 375–418.