7 or HAM

HAMILTONIAN SYSTEMS VS. TWIST MAPS

March 14, 1999

The last section (elliptic f.p.) also appears in SG. Decide where to put it.

In this chapter, we explore the relationship between Hamiltonian systems and symplectic twist maps on cotangent bundles. In the first part of this chapter, we show how to write Hamiltonian systems as compositions of symplectic twist maps. This is instrumental in setting up a simple variational approach to these systems, which is finite dimensional when one searches for periodic orbits. We start in Section 38 with the geodesic flow, which serves as a reference model for Hamiltonian systems: it plays a role similar to that of the integrable map in the twist map theory. In Section 39, we expend our approach to general Hamiltonian or Lagrangian systems satisfying the Legendre condition (which we see as an analog to the twist condition). In Section 3, we show that, whether or not the Legendre condition is satisfied, the time 1 map of a Hamiltonian system may be decomposed into finitely many symplectic twist maps . This method generalises the classical method of broken geodesics of Riemannian geometry. Our main contribution is to make such a method available for Hamiltonian systems that do not satisfy the Legendre condition.

In Section 41, we see how symplectic twist maps also arise from Hamiltonian systems as Poincaré section maps around elliptic periodic orbits. From an opposite perspective, we show in Section 42 that in many cases, a symplectic twist map may be written as the time 1 of a (time dependant) Hamiltonian system. Most of this last section is courtesy of M. Bialy and L. Polterovitch.

38. Case Study: The Geodesic Flow

A. A Few Facts About Riemannian Geometry

Let (M, g) be a compact *Riemannian manifold*. This means that the tangent fibers $T_{\boldsymbol{q}}M$ are endowed with symmetric, positive definite bilinear forms:

$$(\boldsymbol{v}, \boldsymbol{v}') \mapsto g_{(\boldsymbol{q})}(\boldsymbol{v}, \boldsymbol{v}') \text{ for } \boldsymbol{v}, \boldsymbol{v}' \in T_{\boldsymbol{q}}M$$

varying smoothly with the base point $\boldsymbol{q}$. We will denote the *norm* induced by this metric by $\|\boldsymbol{v}\| := \sqrt{g_{(\boldsymbol{q})}(\boldsymbol{v}, \boldsymbol{v})}$. A curve $\boldsymbol{q}(t)$ in M is a *geodesic* if and only if it is an extremal of the *action* or *energy functional*:

$$A_{t_1}^{t_2}(\gamma) = \int_{t_1}^{t_2} \frac{1}{2} \|\dot{\boldsymbol{q}}\|^2 \, dt.$$

between any two of its points $\boldsymbol{q}(t_1)$ and $\boldsymbol{q}(t_2)$ among all absolutely continuous curves $\beta : [t_1, t_2] \to M$ with same endpoints. Geodesics are usually thought of as length extremals, that is critical points of the functional $\int \frac{1}{2} \|\dot{\boldsymbol{q}}\|\, dt$. But action extremals are length extremals and vice versa (with the difference that action extremals come with a specified parametrization). One usually chooses to compute with the action, since it yields simpler calculations. For more detail on this, as well as a the more abstract definition of geodesic given in terms of a connection see e.g. Milnor (1969) .

The variational problem of finding critical points of A has the Lagrangian

$$L_0(\boldsymbol{q}, \boldsymbol{v}) = \frac{1}{2} g_{(\boldsymbol{q})}(\boldsymbol{v}, \boldsymbol{v}) = \frac{1}{2} \|\dot{\boldsymbol{q}}\|^2 .$$

Following the procedure of Section ??? of Chapter SG, we use the Legendre transform to compute the corresponding Hamiltonian function. In local coordinates $\boldsymbol{q}$ in M, we can write

$$g_{(\boldsymbol{q})}(\boldsymbol{v}, \boldsymbol{v}) = \langle A^{-1}_{(\boldsymbol{q})} \boldsymbol{v}, \boldsymbol{v} \rangle,$$

where $\langle , \rangle$ denotes the dot product in $\mathbb{R}^n$, and $A^{-1}_{(\boldsymbol{q})}$ is a symmetric, positive definite matrix varying smoothly with the base point $\boldsymbol{q}$. With this notation, we have

$$\frac{\partial L_0}{\partial \boldsymbol{v}}(\boldsymbol{q}, \boldsymbol{v}) = A^{-1}_{(\boldsymbol{q})} \boldsymbol{v}, \quad \frac{\partial^2 L_0}{\partial \boldsymbol{v}^2} = A^{-1}_{(\boldsymbol{q})}$$

In particular, $\frac{\partial^2 L_0}{\partial v^2}$ is nondegenerate. Hence the Legendre condition is satisfied and the Legendre transformation is, in coordinates:

$$\mathcal{L} : (\boldsymbol{q}, \boldsymbol{v}) \to (\boldsymbol{q}, \boldsymbol{p}) = (\boldsymbol{q}, A^{-1}_{(\boldsymbol{q})} \boldsymbol{v})$$

which transforms L_0 into a Hamiltonian H_0:

$$H_0(\boldsymbol{q}, \boldsymbol{p}) = \boldsymbol{p}\boldsymbol{v} - L_0(\boldsymbol{q}, \boldsymbol{v}) = \langle \boldsymbol{p}, A_{(\boldsymbol{q})} \boldsymbol{p} \rangle - \frac{1}{2} \langle A^{-1}_{(\boldsymbol{q})} A_{(\boldsymbol{q})} \boldsymbol{p}, A_{(\boldsymbol{q})} \boldsymbol{p} \rangle = \frac{1}{2} \langle A_{(\boldsymbol{q})} \boldsymbol{p}, \boldsymbol{p} \rangle.$$

This Hamiltonian is a *metric on the cotangent bundle*:

$$H_0(\boldsymbol{q}, \boldsymbol{p}) = \frac{1}{2} \langle A_{(\boldsymbol{q})} \boldsymbol{p}, \boldsymbol{p} \rangle \stackrel{\text{def}}{=} \frac{1}{2} g^{\#}_{(\boldsymbol{q})}(\boldsymbol{p}, \boldsymbol{p}).$$

We will also denote the norm associated to this metric by $\|\boldsymbol{p}\| = \sqrt{g^{\#}_{(\boldsymbol{q})}(\boldsymbol{p}, \boldsymbol{p})}$. Note that the Legendre transformation is in this case an isometry between the metrics g and $g^{\#}$: in particular, if $(\boldsymbol{q}, \boldsymbol{p}) = \mathcal{L}(\boldsymbol{q}, \boldsymbol{v})$, then $\|\boldsymbol{p}\| = \|\boldsymbol{v}\|$. Hence the Hamiltonian is half of the speed and we retrieve, from conservation of energy in Hamiltonian systems, the fact well known by geometers that extremals of the action are parametrized at constant speed.

The *geodesic flow* is the Hamiltonian flow h^t_0 generated by H_0 on T^*M. It is not hard to see that the trajectories of the geodesic flow restricted to an energy level project to the same curves on M as the trajectories in any other energy level: the velocities are are just multiplied by a scalar (See Exercise 38.1). For this reason, one often restricts the geodesic flow to the *unit cotangent bundle* $T^*_1 M = \{(\boldsymbol{q}, \boldsymbol{p}) \in T^*M \mid \|\boldsymbol{p}\| = 1\}$. Traditionally, geometers use the term geodesic flow to denote the conjugate $\mathcal{L}^{-1} h^t_0 \mathcal{L}$ on TM of this Hamiltonian flow, as restricted to the unit tangent bundle. Remember that projections of trajectories of a Hamiltonian flow associated to a Lagrangian satisfying the Legendre condition are extremals of the action of the Lagrangian, and vice versa. (See Chapter SG, Section ???). In the present case, if $(\boldsymbol{q}(t), \boldsymbol{p}(t))$ a trajectory of the geodesic flow, then $\boldsymbol{q}(t)$ is a geodesic. Conversely, if $\boldsymbol{q}(t)$ is a geodesic, it is the projection on M of the solution $(\boldsymbol{q}(t), \boldsymbol{p}(t))$ of the geodesic flow with initial condition $(\boldsymbol{q}_0, \boldsymbol{p}_0) = (\boldsymbol{q}(0), A^{-1}\dot{\boldsymbol{q}}(0))$.

We now want to establish a fundamental result of Riemmannian geometry, which we will rephrase in the next subsection by saying that the time t of the geodesic flow is a symplectic twist map . The *exponential map* is defined by:

$$exp_{\boldsymbol{q}_0}(t\boldsymbol{v}) = \boldsymbol{q}(t),$$

where $\boldsymbol{q}(t)$ is the geodesic such that $\dot{\boldsymbol{q}}(0) = \boldsymbol{v}$. Note that any geodesic can be written in this exponential notation. In terms of the geodesic flow, $exp_{\boldsymbol{q}_0}(t\boldsymbol{v}) = \pi \circ h_0^t \circ \mathcal{L}(\boldsymbol{q}_0, \boldsymbol{v})$, where $\pi : T^*M \mapsto M$ is the canonical projection.

Theorem 38.1 *The map* $Exp : TM \to M \times M$

$$(\boldsymbol{q}, \boldsymbol{v}) \mapsto (\boldsymbol{q}, \boldsymbol{Q}) \stackrel{\text{def}}{=} (\boldsymbol{q}, exp_{\boldsymbol{q}}(\boldsymbol{v})) \tag{38.1}$$

defines a diffeomorphism between a neighborhood of the 0–section in TM and some neighborhood of the diagonal in $M \times M$. Moreover, for $(\boldsymbol{q}, \boldsymbol{v})$ in that neighborhood:

$$\text{Dis}(\boldsymbol{q}, exp_{\boldsymbol{q}}(\boldsymbol{v})) = \|\boldsymbol{v}\|$$

One way to paraphrase this theorem is by saying that, any two closeby points are joined by a unique, short enough, geodesic segment.

Proof. By definition, $exp_{\boldsymbol{q}}(0) = \boldsymbol{q}$ and $\frac{d}{ds} exp_{\boldsymbol{q}}(s\boldsymbol{v}) = \boldsymbol{v}$ at $s = 0$,

$$DExp\big|_{(\boldsymbol{q},0)} = \begin{pmatrix} Id & Id \\ 0 & Id \end{pmatrix},$$

whose determinant is 1. Hence, Exp is a local diffeomorphism around each point of a compact neighborhood of the 0-section. In particular we can assume that there is an ϵ such that Exp is a diffeomorphism of an ϵ ball in TM around $(\boldsymbol{q}, 0)$ and a ball in $M \times M$ around $(\boldsymbol{q}, \boldsymbol{q})$, where ϵ is independant of $\boldsymbol{q}$.

We now show that Exp is an embedding when restricted to $U_\epsilon = \{(\boldsymbol{q}, \boldsymbol{v}) \in TM \mid \|\boldsymbol{v}\| \le \epsilon\}$, where ϵ is as above. It is enough to check the injectivity. Let two elements in U_ϵ have the same image under Exp. Since the first factor of Exp gives the base point, this can only occur if they are in the same fiber of U_ϵ. But, by our choice of U_ϵ this implies these elements are the same.

Finally, we show that $\text{Dis}(\boldsymbol{q}, exp_{\boldsymbol{q}}(\boldsymbol{v})) = \|\boldsymbol{v}\|$ whenever $\|\boldsymbol{v}\| \le \epsilon$. We remind the reader that the *distance* $\text{Dis}(\boldsymbol{q}, \boldsymbol{Q})$ between two points $\boldsymbol{q}$, and $\boldsymbol{Q}$ in a compact Riemannian manifold is given by the length of the shortest path between $\boldsymbol{q}$ and $\boldsymbol{Q}$. As a length minimizer, the shortest path is also an action minimizer, and hence a geodesic. Since Exp is an embedding of U_ϵ in $M \times M$, exp is 1 to 1 on $U_\epsilon \cap T_{\boldsymbol{q}}M$ and the unique geodesic that joins $\boldsymbol{q}$ and $exp_{\boldsymbol{q}}(\boldsymbol{v})$ in $exp(U_\epsilon \cap T_{\boldsymbol{q}}M)$ is the curve $t \mapsto \boldsymbol{q}(t) = exp_{\boldsymbol{q}}(t\boldsymbol{v})$. The length of this curve is $\int_0^1 \|\dot{\boldsymbol{q}}\|\, dt = \int_0^1 \|\boldsymbol{v}\|\, dt = \|\boldsymbol{v}\|$ (see Exercise 38.1 c)). The only way our formula may fail is if there were a shorter geodesic joining $\boldsymbol{q}$ and $exp_{\boldsymbol{q}}(\boldsymbol{v})$ not in $exp(U_\epsilon \cap T_{\boldsymbol{q}}M)$). But this is impossible since this geodesic would be of the form $\exp_{\boldsymbol{q}}(t\boldsymbol{w}), t \in [0, 1]$ with length $\|\boldsymbol{w}\| > \epsilon$.

□

Exercise 38.1 a) Check that, in local coordinates, Hamilton's equations for the geodesic flow write:

$$\begin{aligned}\dot{q} &= A_{(\boldsymbol{q})}\boldsymbol{p}\\ \dot{p} &= -\left\langle \frac{\partial A_{(\boldsymbol{q})}}{\partial \boldsymbol{q}}\boldsymbol{p}, \boldsymbol{p}\right\rangle\end{aligned} \tag{38.2}$$

b) Verify that $h_0^{st}(\boldsymbol{q},\boldsymbol{p}) = h_0^t(\boldsymbol{q}, s\boldsymbol{p})$. (*Hint.* if $(\boldsymbol{q}(t),\boldsymbol{p}(t))$ is a trajectory of the geodesic flow, then $(\boldsymbol{q}(st), s\boldsymbol{p}(st))$ is also a trajectory).
c) Show that if $\boldsymbol{q}(t) = exp_{\boldsymbol{q}_0}(t\boldsymbol{v})$, $\|\dot{\boldsymbol{q}}(t)\| = \|\boldsymbol{v}\|$ for all t.

Exercise 38.2 Show that the completely integrable twist map $(x, y) \mapsto (x + y, y)$ is the time 1 map of the geodesic flow on the "flat" circle, *i.e.* the circle given the euclidean metric $g_{(x)}(v, v) = v^2$.

B. The Geodesic Flow As A Twist Map

Theorem 38.1 is the key to the following:

Proposition 38.2 *The time 1 map h_0^1 of the geodesic flow with Hamiltonian $H_0(\boldsymbol{q},\boldsymbol{p}) = \frac{1}{2}\|\boldsymbol{p}\|^2$ is a symplectic twist map on $U_\epsilon = \{(\boldsymbol{q},\boldsymbol{p}) \in T^*M \mid | \|\boldsymbol{p}\| \leq \epsilon\}$, for ϵ small enough. More generally, given any $R > 0$, there is an $t_0 > 0$ (or given any t_0 there is an R) such that $h_0^t, t \in [-t_0, t_0]$, is a symplectic twist map on the set $U_R = \{(\boldsymbol{q},\boldsymbol{p}) \mid | \ \|\boldsymbol{p}\| \leq R\}$. The generating function of h_0^t is given by $S(\boldsymbol{q},\boldsymbol{Q}) = \frac{t}{2}\mathrm{Dis}^2(\boldsymbol{q},\boldsymbol{Q})$.*

Proof. Since h_0^1 is a Hamiltonian map, it is exact symplectic. Define $Exp^{\#} = Exp \circ \mathcal{L}^{-1}$. By Theorem 38.1, $Exp^{\#}$ is a diffeomorphism between $U_\epsilon = \{(\boldsymbol{q},\boldsymbol{p}) \mid \|\boldsymbol{p}\| = \epsilon\}$ and a neighborhood of the diagonal in $M \times M$. But $Exp^{\#}(\boldsymbol{q},\boldsymbol{p}) = (\boldsymbol{q}, \boldsymbol{Q}(\boldsymbol{q},\boldsymbol{p}))$, where $\boldsymbol{Q} = \pi \circ h_0^1(\boldsymbol{q},\boldsymbol{p})$. Hence h_0^1 is a symplectic twist map on U_ϵ, and $\psi_{h_0^1} = Exp^{\#}$. The more general statement derives from the fact that $Exp^{\#}(\boldsymbol{q}, t\boldsymbol{p}) = (\boldsymbol{q}, \boldsymbol{q}(t))$, where $h_0^t(\boldsymbol{q},\boldsymbol{p}) = (\boldsymbol{q}(t), \boldsymbol{p}(t))$.

We now show that $\frac{1}{2}\mathrm{Dis}^2(\boldsymbol{q},\boldsymbol{Q})$ is the generating function of h_0^1 when it is a symplectic twist map on a domain U (the proof for h_0^t is identical). Since h_0^1 is a Hamiltonian map,

$$(h_0^1)^*\boldsymbol{p}d\boldsymbol{q} - \boldsymbol{p}d\boldsymbol{q} = dS, \quad \text{with} \quad S(\boldsymbol{q},\boldsymbol{p}) = \int_\gamma \boldsymbol{p}d\boldsymbol{q} - H_0 dt$$

where γ is the curve $h_0^t(\boldsymbol{q},\boldsymbol{p}),\ t \in [0, 1]$ (see Theorem ??? in Chapter SG). We now need to show that S, expressed as a function of $\boldsymbol{q},\boldsymbol{Q}$ is the one advertised. In this particular case, since $\dot{\boldsymbol{q}} = A_{(\boldsymbol{q})}\boldsymbol{p}$ (see Exercise 38.1) and $H_0 = \frac{1}{2}\langle A_{(\boldsymbol{q})}\boldsymbol{p},\boldsymbol{p}\rangle = \frac{1}{2}\|\boldsymbol{p}\|^2$, the integral simplifies:

$$\int_\gamma \boldsymbol{p}d\boldsymbol{q} - H_0 dt = \int_0^1 \frac{1}{2}\langle A_{(\boldsymbol{q})}\boldsymbol{p},\boldsymbol{p}\rangle - \frac{1}{2}\|\boldsymbol{p}(t)\|^2 dt = \int_0^1 \frac{1}{2}\|\boldsymbol{p}(t)\|^2 dt$$

But the integrand is H_0, which is constant along γ. Hence, using Theorem 38.1 , and the fact that $\mathcal{L}$ is an isometry, we get:

$$S(\boldsymbol{q},\boldsymbol{p}) = \frac{1}{2}\|\boldsymbol{p}\|^2 = \frac{1}{2}\|\dot{\boldsymbol{v}}\|^2 = \frac{1}{2}\mathrm{Dis}^2(\boldsymbol{q}, \boldsymbol{Q}(\boldsymbol{q},\boldsymbol{p})),$$

where $(\boldsymbol{q},\boldsymbol{v}) = \mathcal{L}^{-1}(\boldsymbol{q},\boldsymbol{p})$. This makes S the advertised differentiable function of $\boldsymbol{q}$ and $\boldsymbol{Q}$ whenever $(\boldsymbol{q},\boldsymbol{p}) \mapsto (\boldsymbol{q},\boldsymbol{Q})$ is a diffeomorphism. □

Remark 38.3 As a simple example of what makes h_0^1 cease to be a twist map when the domain U is extended too far, take M to be the unit circle with the arclength metric. In a chart $\theta \in (-\epsilon, 2\pi - \epsilon)$, we have:

$$\mathrm{Dis}(0,\theta) = \begin{cases} \theta & \text{when} \quad \theta \le \pi \\ 2\pi - \theta & \text{when} \quad \theta > \pi \end{cases}$$

As a result, the left derivative of $\frac{1}{2}\mathrm{Dis}^2(0,\theta)$ is π, whereas the right derivative is $-\pi$: the function Dis is not differentiable at this point.

The following will be instrumental in the proof of Theorem 31.1. ???Put it there instead???

Corollary 38.4 *Let $h_0^s(\boldsymbol{q},\boldsymbol{p}) = (\boldsymbol{Q}_s, \boldsymbol{P}_s)$ be the time s of the geodesic flow, then:*

$$\partial_1 \mathrm{Dis}(\boldsymbol{q}, \boldsymbol{Q}_s) = -sign(s).\frac{\boldsymbol{p}}{\|\boldsymbol{p}\|} \quad \textit{and} \quad \partial_2 \mathrm{Dis}(\boldsymbol{q}, \boldsymbol{Q}_s) = sign(s).\frac{\boldsymbol{P}_s}{\|\boldsymbol{P}_s\|} \tag{38.3}$$

Proof. From Proposition 38.2, we get:

$$-\boldsymbol{p} = \partial_1 \frac{1}{2}\mathrm{Dis}^2(\boldsymbol{q}, \boldsymbol{Q}_1) = \mathrm{Dis}(\boldsymbol{q}, \boldsymbol{Q}_1)\partial_1 \mathrm{Dis}(\boldsymbol{q}, \boldsymbol{Q}_1) = \|\boldsymbol{p}\|\, \partial_1 \mathrm{Dis}(\boldsymbol{q}, \boldsymbol{Q}_1)$$

which proves $\partial_1 \mathrm{Dis}(\boldsymbol{q}, \boldsymbol{Q}_1) = -\frac{\boldsymbol{p}}{\|\boldsymbol{p}\|}$. Using $\boldsymbol{Q}_s = \pi \circ h_0^1(\boldsymbol{q}, s\boldsymbol{p})$, one may replace $\boldsymbol{p}$ by $s\boldsymbol{p}$ in the previous computation to prove the first equality. For the second equality, the fact that $\mathrm{Dis}(\boldsymbol{q}, \boldsymbol{Q}_s) = \mathrm{Dis}(\boldsymbol{Q}_s, \boldsymbol{q})$, that $\boldsymbol{q} = \pi \circ h_0^1(\boldsymbol{Q}_s, -s\boldsymbol{P}_s)$ (see Exercise 38.2) and the first equality, enables us to write:

$$\partial_2 \mathrm{Dis}(\boldsymbol{q}, \boldsymbol{Q}_s) = \partial_1 \mathrm{Dis}(\boldsymbol{Q}_s, \boldsymbol{q}) = sign(s).\frac{\boldsymbol{P}_s}{\|\boldsymbol{P}_s\|}$$

□

C. The Method of Broken Geodesics

We now draw the correspondence between the variational methods provided by symplectic twist maps and the classical method of broken geodesics, originally due to Birkhoff (???: check Milnor). As before, let h_0^1 be the time 1 map[7] of the geodesic flow with Hamiltonian H_0. Fix some neighborhood U of the zero section in T^*M. Proposition 38.2implies that if we decompose $h_0^1 = (h_0^{\frac{1}{N}})^N$, then for N big enough each $h_0^{\frac{1}{N}}$ is a symplectic twist map in U. As a result, periodic orbits of period 1 for the geodesic flow H_0, *i.e.* fixed points of h_0^1 are given by the critical points of:

$$W(\overline{\boldsymbol{q}}) = \sum_{k=1}^{N} S(\boldsymbol{q}_k, \boldsymbol{q}_{k+1}), \quad \text{with} \quad \boldsymbol{q}_{N+1} = \boldsymbol{q}_1,$$

where $\overline{\boldsymbol{q}}$ belong to set $X_N(U)$ of sequences in M such that $(\boldsymbol{q}_k, \boldsymbol{q}_{k+1}) \in \psi(U)$, where $\psi = \psi_{h_0^{\frac{1}{N}}}$ We now show that W is the action of a *broken geodesic*. Since $h_0^{\frac{1}{N}}$ is a symplectic twist map, the twist condition implies that, given $(\boldsymbol{q}_k, \boldsymbol{q}_{k+1})$ in $\psi(U)$, there is a unique $(\boldsymbol{p}_k, \boldsymbol{P}_k)$ such that $h_0^{\frac{1}{N}}(\boldsymbol{q}_k, \boldsymbol{p}_k) = (\boldsymbol{q}_{k+1}, \boldsymbol{P}_k)$, i.e., there is exactly one trajectory c_k: $[\frac{k}{N}, \frac{k+1}{N}] \to T^*\tilde{M}$ of the geodesic flow that joins $(\boldsymbol{q}_k, \boldsymbol{p}_k)$ to $(\boldsymbol{q}_{k+1}, \boldsymbol{P}_k)$. The projection $\pi(c_k)$ on M is a geodesic, parametrized at constant speed equal to the norm of $\boldsymbol{p}_k$. As we

[7] The following discussion remains valid if we replace 1 by any time T.

have seen in the proof of Proposition 38.3 , $S(\boldsymbol{q}_k, \boldsymbol{q}_{k+1})$ is the action of c_k: $S(\boldsymbol{q}_k, \boldsymbol{q}_{k+1}) = \int_{c_k} \boldsymbol{p}d\boldsymbol{q} - Hdt$. Hence W, the sum of these actions, is the action of the curve C obtained by the concatenation of the c_k's. C is "broken", $i.e.$ has a corner at the point $\boldsymbol{q}_k$ whenever $\boldsymbol{P}_{k-1} \neq \boldsymbol{p}_k$: via the Legendre transformation, $\boldsymbol{P}_{k-1}$ and $\boldsymbol{p}_k$ correspond to the left derivative and right derivative of the curve C at $\boldsymbol{q}_k$.

If $\overline{\boldsymbol{q}}$ is a critical point of W, $\boldsymbol{P}_k = \boldsymbol{p}_{k+1}$, and thus the left and right derivatives coincide: in this case C is a closed, smooth geodesic.

In conclusion, the function $W(\overline{\boldsymbol{q}})$ can be interpreted as the restriction of the action functional $A(c)$ to a *finite dimensional* subspace (the space of curves C arising from elements of $X_N(U)$, which is homeomorphic to $X_N(U)$) in the (infinite dimensional) loop space of T^*M. One can further justify this method by showing that the finite dimensional space $X_N(U)$ is a deformation retract[8] of a subset of the loop space and that it contains all the critical loops of that subset. This was Morse's way to study the topology of the loop space (see Milnor (1969) , 16). Conversely, and this is the point of view in this book (and more generally that of symplectic topology), knowing the topology of certain subsets of the loop space, one can gain information about the dynamics of the geodesic flow or, as we will see, of many Hamiltonian systems. (Part of this in the Intro???)

39. Decomposition Of Hamiltonian Maps Into Twist Maps

A. Legendre Condition vs. Twist Condition

In this subsection, we generalize Theorem 38.2 by proving that Hamiltonian maps satisfying the Legendre condition are symplectic twist maps , provided appropriate restrictions on the domain of the map. We then reformulate this result in the Lagrangian setting, giving a generalization of the fundamental Theorem 38.1 . In the next subsection, we focus on $T^*\mathbb{T}^n$, where, given further conditions on the Hamiltonian, we extend the domain of these symplectic twist maps to the whole space.

Remember that Hamiltonian maps, which are time t maps of Hamiltonian systems, are exact symplectic (Theorem SGhamexactsymp) and, through the flow, isotopic to Id. Therefore, to show that a certain Hamiltonian map is a symplectic twist map, we need only check the twist condition. Clearly, not all Hamiltonian maps satisfy it. Take $F(\boldsymbol{q}, \boldsymbol{p}) = (\boldsymbol{q} + \boldsymbol{m}, \boldsymbol{p})$ on the cotangent bundle of the torus, for example: it is the time one of $H(\boldsymbol{q}, \boldsymbol{p}) = \boldsymbol{m}.\boldsymbol{p}$, and it is definitely not twist. Here is a heuristic argument, which appeared in Moser (1986) in the context of twist maps, to guide us in our search of the twist condition for Hamiltonian maps. The Taylor series with respect to ϵ of the time ϵ map of a Hamiltonian system with Hamiltonian H is:

$$\begin{aligned} \boldsymbol{q}(\epsilon) &= \boldsymbol{q}(0) + \epsilon.H_{\boldsymbol{p}} + o(\epsilon^2) \\ \boldsymbol{p}(\epsilon) &= \boldsymbol{p}(0) - \epsilon.H_{\boldsymbol{q}} + o(\epsilon^2) \end{aligned}.$$

Thus, up to order ϵ^2, $\partial\boldsymbol{q}(\epsilon)/\partial\boldsymbol{p}(0) = \epsilon.H_{\boldsymbol{pp}}$. This shows that *whenever $H_{\boldsymbol{pp}}$ is nondegenerate, the time ϵ map is a symplectic twist map* in some neighborhood of $\boldsymbol{q}(0), \boldsymbol{p}(0)$. The problem is to extend this argument to given regions of the cotangent bundle: the term $o(\epsilon^2)$ might get large as the initial condition varies.

We now present a rigorous version of this argument, valid on compact sets of cotangent bundles of arbitrary compact manifolds. We say that a Hamiltonian $H : T^*M \times \mathbb{R} \to \mathbb{R}$ satisfies the *global Legendre condition* if the map:

[8] This retraction can be obtained by a piecewise curve shortening method.

$$\text{(39.1)} \qquad \boldsymbol{p} \mapsto H_p(\boldsymbol{q}, \boldsymbol{p}, t)$$

is a diffeomorphism from $T^*_{\boldsymbol{q}}M \mapsto T_{\boldsymbol{q}}M$ for each $\boldsymbol{q}$ and t. We will say that H satisfies the *Legendre embedding condition* if the map $\boldsymbol{p} \mapsto H_{\boldsymbol{p}}$ is an embedding (*i.e.* a 1–1, local diffeomorphism). Note that, although we have written it in a chart of conjugate coordinates in T^*M, this condition is coordinate independant (prove this!).

Examples 39.1 We give two classes of examples. In the first one, the Hamiltonian is *not* assumed to be convex.

Let $H(\boldsymbol{q}, \boldsymbol{p}) = \frac{1}{2}\langle A_{(\boldsymbol{q},t)}\boldsymbol{p}, \boldsymbol{p}\rangle + V(\boldsymbol{q}, t)$ and $\det\ A_{(\boldsymbol{q},t)} \neq 0$, then H satisfies (39.1). This is simply because $\boldsymbol{p} \mapsto H_{\boldsymbol{p}} = A_{(\boldsymbol{q},t)}\boldsymbol{p}$ is linear and nonsingular. Note that no convexity is assumed here, only nondegeneracy of $H_{\boldsymbol{pp}}$ (and its independance of $\boldsymbol{p}$).

Less trivially, if $H_{\boldsymbol{pp}}(\boldsymbol{q}, \boldsymbol{p}, t)$ is definite positive, and its smallest eigenvalue is uniformly bounded below by a strictly positive constant, then H statisfies the global Legendre condition. This is a direct consequence of Lemma STMdiffeo.

If we remove the lower bound on the smallest eigenvalue, one can show (see Exercise 39.1) that the map $\boldsymbol{p} \mapsto H_{\boldsymbol{p}}$ is not necessarily a diffeomorphism any more, but remains an embedding and thus H satisfies the Legendre embedding condition.

Such an embedding condition, and a version of Theorem 39.2 , are also satisfied if $H_{\boldsymbol{pp}}$ is positive on a compact set U invariant under the flow (see Exercise 39.2).

Theorem 39.2 *Let M be a compact, smooth manifold and $H : T^*M \times \mathbb{R}$ be a smooth Hamiltonian function which satisfies either the global Legendre condition (39.1) or the Legendre embedding condition. Then, given any compact neighborhood U in T^*M and starting time a, there exists $\epsilon_0 > 0$ (depending on U) such that, for all $\epsilon < \epsilon_0$ the time ϵ map of the Hamiltonian flow of H is a symplectic twist map on U.*

Proof of Theorem 39.1 Choose a Riemannian metric g on M. Define the compact ball bundles:

$$U(K) = \{(\boldsymbol{q}, \boldsymbol{p}) \in T^*M \mid \|\boldsymbol{p}\| \leq K\}.$$

The nested union of these sets covers T^*M. Hence any compact set U is contained in a $U(K)$ for some K large enough, and we may restrict the proof of the theorem to the case $U = U(K)$. Since the Hamiltonian vector field of H is uniformly Lipschitz on compact sets, there is a time T such that the Hamiltonian flow $h_a^{a+t}(\boldsymbol{z})$ of H is defined on the interval $t \in [0, T]$ whenever $\boldsymbol{z} \in U(K)$.
In the rest of the section, we fix a and abreviate h_a^{a+t} by h^t (the time t of the flow with starting time a.
By continuity of the flow, $h^{[0,T]}(U(K))$ is a compact set. We now show that we can work in appropriately chosen charts of T^*M. Since M is compact, we can find a real $r > 0$ such that T^*M is trivial above each ball of radius $2r$ in M. (Indeed, there exist such a ball around each point. If one had a sequence of points whose corresponding maximum such r converged to zero, a limit point of this sequence would not have a trivializing neighborhood, a contradiction). Take a finite covering $\{B_i\}$ of M by balls of radius r, and let B'_i

be the ball of radius $2r$ with same center as B_i. Choose $\epsilon_3 < T$ such that $\pi \circ h^{[0,\epsilon_3]}(\pi^{-1}(B_i)) \subset B_i'$. Such an ϵ_3 exists since there are finitely many B_i's and since the flow is continuous. From now on, we may work in any of the charts $\pi^{-1}(B_i) \simeq B_i \times \mathbb{R}^n$, and know that for the time interval $[0, \epsilon_3]$, we will remain in the charts $\pi^{-1}(B_i') \simeq B_i' \times \mathbb{R}^n$. We let $(\boldsymbol{q}, \boldsymbol{p})$ denote the conjugate coordinates in these charts.

Let $\epsilon < \epsilon_3$ and write $h^\epsilon(\boldsymbol{q}, \boldsymbol{p}) = (\boldsymbol{q}(\epsilon), \boldsymbol{p}(\epsilon))$. Consider the map $\psi_{h^\epsilon} : (\boldsymbol{q}, \boldsymbol{p}) \mapsto (\boldsymbol{q}, \boldsymbol{q}(\epsilon))$. We need to show that ψ_{h^ϵ} is an embedding of $U(K)$ in $M \times M$. By compactness, it suffices to show that ψ_{h^ϵ} is a local diffeomorphism which is 1–1 on $U(K)$. Write the second order Taylor formula for $\boldsymbol{q}(\epsilon)$ with respect to ϵ (this is a smooth function since the flow is smooth):

$$\boldsymbol{q}(\epsilon) = \boldsymbol{q} + \epsilon H_{\boldsymbol{p}}(\boldsymbol{q}, \boldsymbol{p}, a) + \epsilon^2 R(\boldsymbol{q}, \boldsymbol{p}, \epsilon).$$

The smoothness of the Hamiltonian flow garantees that R is smooth in all its variables. Indeed, its precise expression is (see Lang (1983) , p. 116):

$$R(\boldsymbol{q}, \boldsymbol{p}, \epsilon) = \int_0^1 (1-t) \frac{\partial h^{t\epsilon}(\boldsymbol{q}, \boldsymbol{p})}{\partial t} dt$$

and the integrand is smooth since the flow is. The differential of ψ_{h^ϵ} with respect to $(\boldsymbol{q}, \boldsymbol{p})$ is of the form:

$$D\psi_{h^\epsilon}(\boldsymbol{q}, \boldsymbol{p}) = \begin{pmatrix} Id & 0 \\ * & A \end{pmatrix}, \quad A = \epsilon H_{\boldsymbol{pp}}(\boldsymbol{q}, \boldsymbol{p}, a) + \epsilon^2 R_{\boldsymbol{p}}(\boldsymbol{q}, \boldsymbol{p}, \epsilon).$$

Since $\det\, H_{\boldsymbol{pp}} \neq 0$ by the Legendre condition and since $R_{\boldsymbol{p}}$ is continuous and hence bounded on the compact set $U(K) \times [0, \epsilon_3]$, there exists ϵ_2 in $(0, \epsilon_3]$ such that $\det\, D\psi_{h^\epsilon} = \det\, A \neq 0$ on $U(K) \times (0, \epsilon_2]$ (we have used the fact that there are finitely many of our charts B_i covering $U(K)$). Hence ψ_{h^ϵ} is a local diffeomorphism for all $\epsilon \in (0, \epsilon_2]$. We now show that, by maybe shrinking further the interval of ϵ, ψ_{h^ϵ} is one to one on $U(K)$. Suppose not and $\psi_{h^\epsilon}(\boldsymbol{q}, \boldsymbol{p}) = \psi_{h^\epsilon}(\boldsymbol{q}', \boldsymbol{p}')$ for some $(\boldsymbol{q}, \boldsymbol{p}), (\boldsymbol{q}', \boldsymbol{p}') \in U(K)$. The definition of ψ_{h^ϵ} immediately implies that $\boldsymbol{q} = \boldsymbol{q}'$. Also, since ψ_{h^ϵ} is a local diffeomorphism on $U(K)$, we can assume that $\|\boldsymbol{p} - \boldsymbol{p}'\| > \delta$ for some $\delta > 0$. Using Taylor's formula, we have:

$$\boldsymbol{q}(\epsilon) - \boldsymbol{q}'(\epsilon) = \epsilon(H_{\boldsymbol{p}}(\boldsymbol{q}, \boldsymbol{p}, a) - H_{\boldsymbol{p}}(\boldsymbol{q}, \boldsymbol{p}', a)) + \epsilon^2(R(\boldsymbol{q}, \boldsymbol{p}, \epsilon) - R(\boldsymbol{q}, \boldsymbol{p}', \epsilon)).$$

Define the compact set $P(K) := \{(\boldsymbol{q}, \boldsymbol{p}, \boldsymbol{q}, \boldsymbol{p}\prime) \in U(K) \times U(K) \mid \|\boldsymbol{p} - \boldsymbol{p}\prime\| \geq \delta\}$. Since $\boldsymbol{p} \mapsto H_{\boldsymbol{p}}$ is a diffeomorphism, the continuous function $\|H_{\boldsymbol{p}}(\boldsymbol{q}, \boldsymbol{p}, a) - H_{\boldsymbol{p}}(\boldsymbol{q}, \boldsymbol{p}', a)\|$ is bounded below by some $K_1 > 0$ on $P(K)$. The continuous function $(\boldsymbol{q}, \boldsymbol{p}, \epsilon) \mapsto \|R(\boldsymbol{q}, \boldsymbol{p}, \epsilon) - R(\boldsymbol{q}, \boldsymbol{p}', \epsilon)\|$ is bounded, say by K_2, on $P(K) \times [0, \epsilon_2]$ and hence

$$\|\boldsymbol{q}(\epsilon) - \boldsymbol{q}'(\epsilon)\| \geq (\epsilon K_1 - \epsilon^2 K_2) > 0$$

whenever $\epsilon \in (0, \epsilon_1]$ and ϵ_1 is small enough. Now choosing $\epsilon_0 = \min\{\epsilon_1, \epsilon_2\}$ finishes the proof of the theorem. □

The following proposition, which is a reformulation of Theorem 39.2 in Lagrangian terms, is a generalization of the fundamental Theorem 38.1 . It garantees the existence and *uniqueness* of Euler-Lagrange solutions between any two closeby points. A time that the solution is traversed has to be specified within a compact interval. In Chapter MIN, we will encounter Tonelli's theorem which implies, for fiber convex Lagrangian systems, that these solutions can also be assumed to be action minimizers.

Proposition 39.3 *Let M be a compact manifold and $L : TM \times \mathbb{R} \to \mathbb{R}$ be a Lagrangian function satisfying the global Legendre condition: $\boldsymbol{v} \mapsto L_{\boldsymbol{v}}(\boldsymbol{q}, \boldsymbol{v}, t)$ is a diffeomorphism. Then, for all starting time a and bound on the velocity K there exists an interval of time $[a, a+\epsilon_0]$ such that, for all $\epsilon < \epsilon_0$, there exists a neighborhood $\mathcal{O}$ of the diagonal in $M \times M$ such that whenever $(\boldsymbol{q}, \boldsymbol{Q}) \subset \mathcal{O}$, there exists a unique solution $\boldsymbol{q}(t)$ of the Euler-Lagrange equations such that $\boldsymbol{q} = \boldsymbol{q}(a)$, $\boldsymbol{Q} = \boldsymbol{q}(a+\epsilon)$ and $\|\dot{\boldsymbol{q}}(a)\| \le K$.*

Remark 39.4 Note that, in the case of the geodesic flow, the curves joining the same points $\boldsymbol{q}, \boldsymbol{Q}$ in different time intervals in this proposition are geometrically all the same geodesic, traversed at different speeds. The dependence on the time interval chosen and the speed chosen of the geometric solutions of the Euler-Lagrange equations is one of the main difference, and source of confusion, when trying to generalise notions of Riemannian geometry to Lagrangian mechanics.

Proof. The Legendre condition enables us to define the Legendre transform $\mathcal{L} : (\boldsymbol{q}, \boldsymbol{v}) \to (\boldsymbol{q}, \boldsymbol{p} = L_{\boldsymbol{v}})$ and the Hamiltonian function $H(\boldsymbol{q}, \boldsymbol{p}, t) = \boldsymbol{p}\dot{\boldsymbol{q}} - L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$, where it is understood that $\dot{\boldsymbol{q}} = \dot{\boldsymbol{q}} \circ \mathcal{L}^{-1}(\boldsymbol{q}, \boldsymbol{p})$ (see Section hamsys ??? in SG). H satisfies the global Legendre condition and $\mathcal{L}^{-1}(\boldsymbol{q}, \boldsymbol{p}) = (\boldsymbol{q}, H_{\boldsymbol{p}})$ (see Remark ???), In particular Theorem 39.3applies to the Hamiltonian H. Let

$$U = V(K) = \{(\boldsymbol{q}, \boldsymbol{p}) \mid \|H_{\boldsymbol{p}}(\boldsymbol{q}, \boldsymbol{p}, a)\| \le K\}.$$

This set is compact since it corresponds, under the Legendre transformation, to

$$\mathcal{L}^{-1}(V(K)) = \{(\boldsymbol{q}, \dot{\boldsymbol{q}}) \mid \|\dot{\boldsymbol{q}}(a)\| \le K\}$$

in the tangent bundle. Theorem 39.3 tells us that, for all $\epsilon \in (0, \epsilon_0]$ with ϵ_0 small enough, the map h^ϵ is a symplectic twist map on $V(K)$. Define

$$\mathcal{O} = \psi_{h^\epsilon}(V(K)).$$

We now show, maybe by decreasing ϵ_0, that $\mathcal{O}$ is a neighborhood of the diagonal in $M \times M$. Let $V_{\boldsymbol{q}}(K) = \pi^{-1}(\boldsymbol{q}) \cap V(K)$ and write $h^t(\boldsymbol{q}, \boldsymbol{p}) = (\boldsymbol{q}(t), \boldsymbol{p}(t))$ where, as before, h^t denotes h_a^{a+t}. The curve $\boldsymbol{q}(t)$ is a solution of the Euler-Lagrange equation satisfying $\boldsymbol{q} = \boldsymbol{q}(a)$ and if $(\boldsymbol{q}, \boldsymbol{p}) \in V_{\boldsymbol{q}}(K)$, then $\|\dot{\boldsymbol{q}}(a)\| = \|H_{\boldsymbol{p}}\| \le K$. As in the proof of Theorem 39.2, we write the Taylor approximation of the solution:

$$\pi \circ h^\epsilon(\boldsymbol{q}, \boldsymbol{p}) = \boldsymbol{q}(\epsilon) = \boldsymbol{q} + \epsilon H_{\boldsymbol{p}} + \epsilon^2 R(\boldsymbol{q}, \boldsymbol{p}, \epsilon).$$

At first order in ϵ, the image of $V_{\boldsymbol{q}}(K)$ under $\pi \circ h^\epsilon$ is $\{\boldsymbol{q} + \epsilon H_{\boldsymbol{p}}(\boldsymbol{q}, \boldsymbol{p}) \mid (\boldsymbol{q}, \boldsymbol{p}) \in V_{\boldsymbol{q}}(K)\}$, which is a solid ball centered at $\boldsymbol{q}$. When adding the second order term $\epsilon^2 R$, $\boldsymbol{q}$ will still be in $\pi \circ h^\epsilon(V_{\boldsymbol{q}}(K))$, provided that ϵ is small enough. By compactness ϵ can be chosen to work for all $\boldsymbol{q}$. Thus $(\boldsymbol{q}, \boldsymbol{q}) \in h^\epsilon(V(K)) = \mathcal{O}$ for all $\boldsymbol{q} \in M$, as claimed.

The rest of the proof is a pure translation of the statements of Theorem 39.2: by construction, if $(\boldsymbol{q}, \boldsymbol{Q}) \in \mathcal{O}$, then $(\boldsymbol{q}, \boldsymbol{Q}) = (\boldsymbol{q}, \boldsymbol{q}(\epsilon))$ where $\boldsymbol{q}(t) = \pi \circ h^t(\boldsymbol{q}, \boldsymbol{p})$ and $(\boldsymbol{q}, \boldsymbol{p}) \in V(K)$. Hence $\boldsymbol{q}(t)$ is a solution to the Euler-Lagrange equation starting at $\boldsymbol{q}$ at time a, landing on $\boldsymbol{Q}$ at time $a+\epsilon$. Moreover, since $(\boldsymbol{q}, \boldsymbol{p}) \in V(K)$, $\|\dot{\boldsymbol{q}}(a)\| = \|H_p(\boldsymbol{q}, \boldsymbol{p}, a)\| \le K$. Finally, this solution is unique. Otherwise, by the uniqueness of solutions of O.D.E.'s, there would be $\boldsymbol{p} \ne \boldsymbol{p}'$ such that $\pi \circ h^\epsilon(\boldsymbol{q}, \boldsymbol{p}) = \pi \circ h^\epsilon(\boldsymbol{q}, \boldsymbol{p}')$, a contradiction to the twist condition. □

Exercise 39.1 Show that a C^1 map $f : \mathbb{R}^n \mapsto \mathbb{R}^n$ which satisfies $\langle Df_x \cdot v, v\rangle > 0$ for all $\boldsymbol{v}$ and x in $\mathbb{R}^n$ is an embedding, *i.e.* it is injective with continuous and differentiable inverse. Deduce that a Hamiltonian such that $H_{\boldsymbol{pp}}$ is positive definite satisfies the Legendre embedding condition.

Exercise 39.2 Let U be a compact region which is invariant under the flow of a Hamiltonian H. Assume also that $H_{\boldsymbol{pp}}$ is positive definite on U. Show that the time t map is a symplectic twist map for all $t > 0$ sufficiently small. (*Hint.* First prove, as in the previous exercise, that $\boldsymbol{p} \mapsto H_{\boldsymbol{p}}$ is an embedding of $T^*_{\boldsymbol{q}}M \cap U$ for each $\boldsymbol{q}$. Then adapt the proof of Theorem 39.2).

B. The Case of the Torus

When the configuration manifold is $\mathbb{T}^n$, there is hope to show that the time t maps of a Hamiltonian system is a symplectic twist map on the whole cotangent bundle. We present here some condition under which this is true. No doubt one could find other, even weaker conditions as well.

Assumption 1 (Uniform opticity)

$H(\boldsymbol{q}, \boldsymbol{p}, t) = H_t(\boldsymbol{z})$ is a twice differentiable function on $T^*\mathbb{T}^n \times \mathbb{R}$ and satisfies the following:

(1) $\sup \left\|\nabla^2 H_t\right\| < K$

(2) $C\|\boldsymbol{v}\|^2 < \langle H_{\boldsymbol{pp}}(\boldsymbol{z}, t)\boldsymbol{v}, \boldsymbol{v}\rangle < C^{-1}\|\boldsymbol{v}\|^2$ for some positive C independant of $(\boldsymbol{z}, t)$ and $\boldsymbol{v} \neq 0$.

Sometimes Hamiltonian systems such that $H_{\boldsymbol{pp}}$ is definite positive are called *optical*. This is why we refer to Assumption 1 as one of *uniform opticity*.

Assumption 2 (Assymptotic quadraticity)

$H(\boldsymbol{q}, \boldsymbol{p}, t)$ is a C^2 function on $T^*\mathbb{T}^n$ satisfying the following:

(1) $\det H_{\boldsymbol{pp}} \neq 0$.

(2) For $\|\boldsymbol{p}\| \geq K_1, \quad H(\boldsymbol{q}, \boldsymbol{p}, t) = \langle A\boldsymbol{p}, \boldsymbol{p}\rangle + \boldsymbol{c}.\boldsymbol{p}, \qquad A^t = A, \det A \neq 0.$

Here A denotes a constant matrix, and $\boldsymbol{c}$ a constant in $\mathbb{R}^n$. We stress that A (and hence $H_{\boldsymbol{pp}}$) is *not* necessarily positive definite.

Theorem 39.5 *Let h^ϵ be the time ϵ of a Hamiltonian flow for a Hamiltonian function satisfying any of the Assumptions 1 or 2. Then, for small enough ϵ, h^ϵ is a symplectic twist map of $T^*\mathbb{T}^n$ (or on U, respectively).*

Remark 39.6 Proposition 39.5 holds for $h_t^{t+\epsilon}$ whenever it does for h^ϵ: $h_t^{t+\epsilon}$ is the time ϵ of the Hamiltonian $G(\boldsymbol{z}, s) = H(\boldsymbol{z}, t + s)$, which satisfies all the assumptions H does.

Proof. We prove the proposition with Assumption 1, and indicate how to adapt the proof to the other assumption. We can work in the covering space $\mathbb{R}2n$ of $T^*\mathbb{T}^n$, to which the flow lifts. The differential of h^t at a point $\boldsymbol{z} = (\boldsymbol{q}, \boldsymbol{p})$ is solution of the linear *variational equation* [9]

[9] In general, if ϕ^t is solution of the O.D.E. $\dot{\boldsymbol{z}} = X_t(\boldsymbol{z})$ then $D\phi^t$ is solution of $\dot{U}(t) = DX_t(\phi^t\boldsymbol{z})U(t), U(0) = Id$. Heuristically, this can be seen by differentiating $\frac{d}{dt}\phi^t(\boldsymbol{z}) = X_t(\phi^t(\boldsymbol{z}))$ with respect to $\boldsymbol{z}$ (see e.g. Hirsh & Smale (1974)).

(39.2) $$\dot{U}(t) = J\nabla^2 H(h^t(\boldsymbol{z}))U(t), \quad U(0) = Id, \quad J = \begin{pmatrix} 0 & -Id \\ Id & 0 \end{pmatrix}$$

We first prove that $U(\epsilon)$ is not too far from Id:

Lemma 39.7 *Consider the linear equation:*

$$\dot{U}(t) = A(t)U(t), \quad U(t_0) = U_0$$

where U and A are $n \times n$ matrices and $\|A(t)\| < K, \forall t$. Then :

$$\|U(t) - U_0\| < K\,\|U_0\|\,|t - t_0|e^{K|t-t_0|}.$$

Proof. Let $V(t) = U(t) - U_0$, so that $V(t_0) = 0$. We have:

$$\begin{aligned}\dot{V}(t) &= A(t)\,(U(t) - U_0) + A(t)U_0 \\ &= A(t)V(t) + A(t)U_0\end{aligned}$$

and hence:

$$\|V(t)\| = \|V(t) - V(t_0)\| \leq \int_{t_0}^{t} K\,\|V(s)\|\,ds + |t - t_0|K\,\|U_0\|$$

For all $|t - t_0| \leq \epsilon$, we can apply Gronwall's inequality (see Hirsh & Smale (1974)) to get:

$$\|V(t)\| \leq \epsilon K\,\|U_0\|\,e^{K|t-t_0|}$$

and we get the result by setting $\epsilon = |t - t_0|$. □

We now proceed with the proof of Proposition 39.5. By Lemma 39.7 we can write:

$$U(\epsilon) - Id = \int_0^{\epsilon} J\nabla^2 H(h^s(\boldsymbol{z})).(Id + O_1(s))ds$$

where $\|O_1(s)\| < 2Ks$, for ϵ, and hence s, small enough.

Let $(\boldsymbol{q}(t), \boldsymbol{p}(t)) = h^t(\boldsymbol{q}, \boldsymbol{p}) = h^t(\boldsymbol{z})$. The matrix $\boldsymbol{b}_\epsilon(\boldsymbol{z}) = \partial\boldsymbol{q}(\epsilon)/\partial\boldsymbol{p}$, is the upper right $n \times n$ matrix of $U(\epsilon)$. It is given by:

(39.3) $$\boldsymbol{b}_\epsilon(\boldsymbol{z}) = \int_0^{\epsilon} H_{\boldsymbol{pp}}(h^s(\boldsymbol{z}))ds + \int_0^{\epsilon} O_2(s)ds$$

where $\left|\int_0^\epsilon O_2(s)ds\right| < K^2\epsilon^2$. From this, and the fact that

(39.4) $$C\,\|\boldsymbol{v}\|^2 < \langle H_{\boldsymbol{pp}}(\boldsymbol{z})\boldsymbol{v}, \boldsymbol{v}\rangle < C^{-1}\,\|\boldsymbol{v}\|^2\,,$$

we deduce that:

(39.5) $$(\epsilon C - K^2\epsilon^2)\,\|\boldsymbol{v}\|^2 < \langle \boldsymbol{b}_\epsilon(\boldsymbol{z})\boldsymbol{v}, \boldsymbol{v}\rangle < (\epsilon C^{-1} + K^2\epsilon^2)\,\|\boldsymbol{v}\|^2$$

so that in particular $\boldsymbol{b}_\epsilon(\boldsymbol{z})$ is nondegenerate for small enough ϵ. Since $\boldsymbol{b}_\epsilon(\boldsymbol{z})$ is periodic in $\boldsymbol{z}$, the set of nonsingular matrices $\{\boldsymbol{b}_\epsilon(\boldsymbol{z})\}_{\boldsymbol{z}\in\mathbb{R}2n}$ is included in a compact set and thus:

$$\sup_{z \in \mathbb{R}2n} \left\| b_\epsilon^{-1}(z) \right\| < K', \tag{39.6}$$

for some positive K'. We can now apply Proposition 26.4to show that h^ϵ is a symplectic twist map with a generating function S defined on all of $\mathbb{R}2n$.

Remark 39.8 The above proof shows that h^ϵ satisfies a certain *convexity condition* which can be useful in finding minimal orbits (see Chapter MIN):

$$\langle b_\epsilon^{-1} v, v \rangle = \left\langle \left(\frac{\partial q}{\partial p}(\epsilon) \right)^{-1} v, v \right\rangle \geq a \left\| v \right\|^2, \quad \forall v \in \mathbb{R}^n. \tag{39.7}$$

where a is a positive constant. To see that it is the case, note that, denoting by

$$m = \inf_{\|v\|=1,\ z \in \mathbb{R}2n} \left\| b_\epsilon^{-1}(z) \right\|$$

and M the corresponding sup, (39.5) implies:

$$m(\epsilon C - K^2\epsilon^2)\|v\|^2 < \langle b_\epsilon^{-1}(z)v, v \rangle < M(\epsilon C^{-1} + K^2\epsilon^2) \left\| v \right\|^2.$$

We now adapt the above proof to Assumption 2. Note that under this assumption, we can still derive (39.3) : the boundary condition (2) implies that $\nabla^2 H$ is bounded. Since H is C^2, and $H_{pp} = A$ outside a compact set, $H_{pp}(h^s z)$ is uniformly close to $H_{pp}(z)$ for small s, and thus the first matrix integral in (39.3) is non singular for z and small s. Thus $b_\epsilon(z)$ is also nonsingular for small ϵ. Since $b_\epsilon(z) = \epsilon A$ outside of the compact set $\|p\| \leq K_1$, the set of matrices $\{b_\epsilon(z) \mid |z \in \mathbb{R}^n\}$ is compact and hence (39.5) holds, which proves the proposition in this case. □

C. Decomposition Of Hamiltonian Maps Into Twist Maps

When, as is the case in Theorems 39.2and 39.5, the time ϵ maps of a Hamiltonian system are all symplectic twist maps , one can readily decompose the time 1 map into such twist maps. Take a time dependent Hamiltonian, for example. Its time 1 map h^1 can be written:

$$h^1 = (h^{\frac{1}{N}})^N$$

and, for N large enough, each $h^{\frac{1}{N}}$ is a symplectic twist map . It is only slightly more complicated when H is time dependent. In this case we can write:

$$h^1 = h^1_{\frac{N-1}{N}} \circ (h^{\frac{N-1}{N}}_{\frac{N-2}{N}}) \circ \ldots h^{\frac{k+1}{N}}_{\frac{k}{N}} \circ \ldots h_0^{\frac{1}{N}} \tag{39.8}$$

and each $h^{\frac{k+1}{N}}_{\frac{k}{N}}$ is an symplectic twist map by assumption on our Hamiltonian. as the next Proposition shows.

What may be more surprising, and gives strength to this method, is that there is a large class of Hamiltonian systems which, even though their time ϵ is *not* twist, can be decomposed into a product of symplectic twist maps. This is a generalization of an idea that LeCalvez (astérisque) applied in his variational proof of the Poincaré-Birkhoff Theorem.

This will work with either of the following, very broad, assumptions:

Assumption 3.

H is a C^2 function on $T^*M \times [0,1]$, and the domain U is a compact neighborhood in T^*M.

Assumption 4.

$H(\boldsymbol{z},t) = H_t(\boldsymbol{z})$ is a function on $T^*\mathbb{T}^n \times \mathbb{R}$ satisfying $\sup \|\nabla^2 H_t\| < K$.

Proposition 39.9 (Decomposition) *Let $H(\boldsymbol{z},t)$ be a Hamiltonian function satisfying Asssumptions 3 or 4, or the hypothesis of either Theorem 39.2 or Theorem 39.5. Then the time 1 h^1 of its corresponding Hamiltonian system can be decomposed into a finite product of symplectic twist maps:*

$$h^1 = F_{2N} \circ \ldots \circ F_1.$$

Proof. We have given the trivial proof above for Hamiltonians that satisfies the hypothesis of Theorems 39.2 and 39.5. We now prove the proposition when H satisfies Assumption 3. Pick a ball bundle $U(K) = \{(\boldsymbol{q},\boldsymbol{p}) \mid \|\boldsymbol{p}\| \le K\}$ with K large enough so that $U \subset U(K)$. Let G be the time s of the geodesic flow, where s is chosen so that G is an symplectic twist map on $U(K)$. That such an s exists is proven in Proposition 38.2 . We can write:

$$\begin{aligned} h^1 &= G \circ \left(G^{-1} \circ h^1_{\frac{N-1}{N}}\right) \circ G \circ \ldots \circ \left(G^{-1} \circ h^{\frac{k+1}{N}}_{\frac{k}{N}}\right) \circ \ldots \circ G \circ \left(G^{-1} \circ h_0^{\frac{1}{N}}\right) \\ &= F_{2N} \circ \ldots \circ F_1. \end{aligned} \tag{39.9}$$

One can check that, at each successive step of the decomposition, the points remain in $U(K)$. Our new G is an symplectic twist map , by assumption, and $G^{-1} \circ h^{\frac{k+1}{N}}_{\frac{k}{N}}$ is an symplectic twist map by openess of the set of twist maps on a compact neighborhood (see Exercise STMstmopen).

Suppose now that H satisfies Assumption 4. Let $G(\boldsymbol{q},\boldsymbol{p}) = (\boldsymbol{q}+\boldsymbol{p},\boldsymbol{p})$, our favorite symplectic twist map (see, *eg.* Example STMstandardexample) on $T^*\mathbb{T}^n$. Decompose h^1 as in Equation (39.9) . We now show that $G^{-1} \circ h^{\frac{k+1}{N}}_{\frac{k}{N}}$ is also a symplectic twist map. Lemma 39.4 implies that $h_t^{t+\epsilon}$ satisfies $\|Dh_t^{t+\epsilon} - Id\| < \epsilon K e^{K\epsilon}$. Hence

$$\left\| DG^{-1}.Dh^{\frac{k+1}{N}}_{\frac{k}{N}} - DG^{-1} \right\| < C\frac{1}{N}e^{\frac{K}{N}}$$

for some positive constant C. Thus $G^{-1} \circ h^{\frac{k+1}{N}}_{\frac{k}{N}}$ is twist for N large enough, since the sufficient conditions $\det\, \partial\boldsymbol{Q}/\partial\boldsymbol{p} \neq 0$ and $\|(\partial Q/\partial p)^{-1}\| < \infty$ are both open with respect to the C^1 norm.

□

40. SUSPENSION OF SYMPLECTIC TWIST MAPS BY HAMILTONIAN FLOWS

Moser (1986) showed how to suspend a monotone twist map of the annulus into a time 1 map of a (time dependant) Hamiltonian system satisfying the fiber convexity $H_{pp} > 0$. In subsection A we present a suspension theorem for higher dimensional symplectic twist maps announced by M. Bialy and L. Polterovitch, which implies Moser's theorem in two dimensions. These authors kindly agreed to let their complete proof appear for the first time in this book. In subsection B, we give the proof, due to the author, of a suspension theorem where we let go of a symmetry condition assumed by Bialy and Polterovitch. The price we pay is the loss of the fiber convexity of the suspending Hamiltonian.

A. SUSPENSION WITH FIBER CONVEXITY

Theorem 40.1 (Bialy and Polterovitch) *Let F be a symplectic twist map with generating function S satisfying:*

(40.1) $\partial_{12} S(\boldsymbol{q}, \boldsymbol{Q})$ *is symmetric and negative nondegenerate.*

Then there exists a smooth Hamiltonian function $H(\boldsymbol{q}, \boldsymbol{p}, t)$ on $T^*\mathbb{T}^n \times [0,1]$ convex in the fiber (i.e. $H_{\boldsymbol{pp}}$ is positive definite) such that F is the time 1 map of the Hamiltonian flow generated by H. The Hamiltonian function H can also be made periodic in the time t.

Proof. Following Moser, we will construct a Lagrangian function $L(\boldsymbol{q}, \boldsymbol{v}, t)$ on $\mathbb{R}2n \times [0,1]$ with the following properties:
(40.2) (a) The corresponding solutions of the Euler-Lagrange equations connecting the points $\boldsymbol{q}$ and $\boldsymbol{Q}$ in the covering space $\mathbb{R}^n$ in the time interval $[0,1]$ are straight lines $\boldsymbol{q} + t(\boldsymbol{Q} - \boldsymbol{q})$;
(40.2) (b) $S(\boldsymbol{q}, \boldsymbol{Q}) = \int_0^1 L(\boldsymbol{q} + t(\boldsymbol{Q} - \boldsymbol{q}), \boldsymbol{Q} - \boldsymbol{q}, t)dt$;
(40.2) (c) L is strictly convex with respect to $\boldsymbol{v}$: $\frac{\partial^2 L}{\partial \boldsymbol{v}^2}$ is positive definite.
(40.2) (d) $L(\boldsymbol{q} + \boldsymbol{m}, \boldsymbol{v}, t) = L(\boldsymbol{q}, \boldsymbol{v}, t)$ for all $\boldsymbol{m}$ in $\mathbb{Z}^n$.

If such a function L is constructed, its Legendre transform H satisfies the conclusion of Theorem 40.1: (40.2) (a) and (b) imply that F is the time 1 map of the Hamiltonian H, (40.2) (c) implies that $H_{\boldsymbol{pp}}$ is convex (see Exercise 47.2) and (40.2) (d) that the Euler-Lagrange flow of L takes place on $T\mathbb{T}^n$ and hence the Hamiltonian flow of H is defined on $T^*\mathbb{T}^n$.

Note that if (40.2) (c) is satisfied then (40.2) (a) is equivalent to the following equation:
(40.2) (a') $\frac{\partial^2 L}{\partial \boldsymbol{v} \partial \boldsymbol{q}} \boldsymbol{v} + \frac{\partial^2 L}{\partial \boldsymbol{v} \partial t} - \frac{\partial F}{\partial \boldsymbol{q}} = 0.$

Lemma 40.2 *Set* $R_{ij}(\boldsymbol{q}, \boldsymbol{v}, t) = -\frac{\partial^2 S}{\partial \boldsymbol{q}_i \partial Q_j}(\boldsymbol{q} - t\boldsymbol{v}, \boldsymbol{q} + (1-t)\boldsymbol{v})$. *Then the following holds:*
(40.3) (a) $R_{ij} = R_{ji}$;
(40.3) (b) $\frac{\partial R_{ij}}{\partial v_k} = \frac{\partial R_{ik}}{\partial v_j}$;
(40.3) (c) $\frac{\partial R_{ij}}{\partial q_k} = \frac{\partial R_{ik}}{\partial q_j}$;
(40.3) (d) $\frac{\partial R_{ij}}{\partial t} + \sum_l \frac{\partial R_{lj}}{\partial q_i} v_l = 0$
for all i, j, k.

The proof is straightforward and uses the fact that the matrix $\frac{\partial^2 S}{\partial \boldsymbol{q} \partial \boldsymbol{Q}}$ is symmetric.

Lemma 40.3 *Set* $L(\boldsymbol{q}, \boldsymbol{v}, t) = \int_0^1 (1-\lambda) \sum_{i,j} R_{ij}(\boldsymbol{q}, \lambda \boldsymbol{v}, t) v_i v_j d\lambda$. *Then the following holds:*

(40.4) (a) $\dfrac{\partial L}{\partial v_i} = \int_0^1 \sum_j R_{ij}(\boldsymbol{q}, \tau\boldsymbol{v}, t)d\tau$

(40.4) (b) $\dfrac{\partial^2 L}{\partial v_i \partial v_j} = R_{ij}$

(40.4) (c) *L satisfies Equation (40.2) (a').*

Proof. Rewrite L as follows:

$$
\begin{aligned}
L(\boldsymbol{q}, \boldsymbol{v}, t) &= \int_0^1 \int_\lambda^1 ds \sum_{i,j} R_{ij}(\boldsymbol{q}, \lambda\boldsymbol{v}, t)v_i v_j d\lambda = \int_0^1 ds \int_0^s d\lambda \sum_{i,j} R_{ij}(\boldsymbol{q}, \lambda\boldsymbol{v}, t)v_i v_j \\
&= \int_0^1 ds \int_0^1 s \sum_{i,j} R_{ij}(\boldsymbol{q}, s\tau\boldsymbol{v}, t)v_i v_j d\tau = \int_0^1 \sum_i v_i \alpha_i(\boldsymbol{q}, s\boldsymbol{v}, t)ds,
\end{aligned}
\tag{40.5}
$$

where $\alpha_i(\boldsymbol{q}, \boldsymbol{v}, y) = \int_0^1 \sum_j R_{ij}(\boldsymbol{q}, \tau\boldsymbol{v}, t)v_j d\tau$. We can rewrite the last integral of (40.5) as a path integral:

$$
\int_0^1 \sum_i v_i \alpha_i(\boldsymbol{q}, s\boldsymbol{v}, t)ds = \int_\gamma \sum_i \alpha_i dv_i,
$$

where $\gamma(s) = (\boldsymbol{q}, s\boldsymbol{v}, t)$. Fixin $\boldsymbol{q}$ and t, Equation (40.3) (b) implies that the form $\sum_i \alpha_i dv_i$ is closed, and, because $\boldsymbol{v} \in \mathbb{R}^n$, exact, say $\sum_i \alpha_i dv_i = dA$ for some function $A(\boldsymbol{v})$ on $\mathbb{R}^n$. Then the Fundamental Theorem of Calculus yields:

$$
L(\boldsymbol{q}, \boldsymbol{v}, t) = A(\boldsymbol{v}) - A(0).
$$

Since $\sum_i \alpha_i dv_i = dA = \frac{\partial L}{\partial \boldsymbol{v}} d\boldsymbol{v}$, Equation (40.4) (a) follows. The proof of (40.4) (b) is similar. We now prove (40.4) (c). In view of (40.4) (a), the left hand side I of (40.2) (a)' can be written as follows:

$$
\begin{aligned}
I =& \sum_l v_l \int_0^1 \sum_j \frac{\partial r_{ij}}{\partial q_l}(\boldsymbol{q}, \tau\boldsymbol{v}, t)v_j d\tau + \int_0^1 \sum_j \frac{\partial R_{ij}}{\partial t}(\boldsymbol{q}, \tau\boldsymbol{v}, t)v_j d\tau \\
&+ \int_0^1 (1-\lambda) \sum_{i,j} \frac{\partial R_{lj}}{\partial q_i}(\boldsymbol{q}, \lambda\boldsymbol{v}, t)v_i v_j d\lambda. \\
=& a_1 + a_2 - a_3,
\end{aligned}
$$

where a_k is the k^{th} integral in the above expression. Rewrite a_3 using (40.3) (c) as follows:

$$
a_3 = \int_0^1 \sum_{l,j} \frac{\partial R_{ij}}{\partial q_j} v_l v_j d\tau - \int_0^1 \sum_{l,j} \frac{\partial R_{l,j}}{\partial q_i} v_l v_j \tau d\tau.
$$

The first term is equal to a_1. Therefore:

$$
I = \int_0^1 \sum_j v_j \left\{ \frac{\partial R_{ij}}{\partial t}(\boldsymbol{q}, \tau\boldsymbol{v}, t) + \sum_{l,j} \frac{\partial R_{l,j}}{\partial q_i} \tau v_l \right\} d\tau.
$$

Equation (40.3) implies that the bracket, and hence I, vanish. □

Given any function $L(\boldsymbol{q}, \boldsymbol{v}, t)$, set

$$
\tilde{L}(\boldsymbol{q}, \boldsymbol{Q}) = \int_0^1 L(\boldsymbol{q} + t(\boldsymbol{Q} - \boldsymbol{q}), \boldsymbol{Q} - \boldsymbol{q}, t)dt.
$$

Lemma 40.4 *Assume that L satisfies (40.2) (a'). Then the following holds:*

(40.6) (a) $$\frac{\partial \tilde{L}}{\partial q_i} = -\frac{\partial L}{\partial v_i}(\boldsymbol{q}, \boldsymbol{Q}-\boldsymbol{q}, 0);$$

(40.6) (b) $$\frac{\partial \tilde{L}}{\partial Q_i} = \frac{\partial L}{\partial v_i}(\boldsymbol{Q}, \boldsymbol{Q}-\boldsymbol{q}, 1);$$

(40.6) (c) $$\frac{\partial^2 \tilde{L}}{\partial q_i \partial Q_j} = -\frac{\partial^2 L}{\partial v_i \partial v_j}(\boldsymbol{q}, \boldsymbol{Q}-\boldsymbol{q}, 0).$$

Proof. Equation (40.6) (c) is a consequence of (40.6) (a), which we now prove. The same argument also proves (40.6) (b). It is not hard to check that if L satisfies (40.2) (a)' then:

$$\frac{d}{dt}\left\{\frac{\partial L}{\partial v_i}(\boldsymbol{q}+t(\boldsymbol{Q}-\boldsymbol{q}), \boldsymbol{Q}-\boldsymbol{q}, t)\right\} = \frac{\partial L}{\partial q_i}(\boldsymbol{q}+t(\boldsymbol{Q}-\boldsymbol{q}), \boldsymbol{Q}-\boldsymbol{q}, t).$$

Therefore,

$$\begin{aligned}&\frac{\partial \tilde{L}}{\partial q_i}(\boldsymbol{q}, \boldsymbol{Q}) = \\ &\int_0^1 \left\{-\frac{\partial L}{\partial v_i}(\boldsymbol{q}+t(\boldsymbol{Q}-\boldsymbol{q}), \boldsymbol{Q}-\boldsymbol{q}, t) + (1-t)\frac{d}{dt}\left(\frac{\partial L}{\partial v_i}(\boldsymbol{q}+t(\boldsymbol{Q}-\boldsymbol{q}), \boldsymbol{Q}-\boldsymbol{q}, t)\right)\right\} dt \\ &= \int_0^1 \frac{d}{dt}\left\{(1-t)\frac{\partial L}{\partial v_i}(\boldsymbol{q}+t(\boldsymbol{Q}-\boldsymbol{q}), \boldsymbol{Q}-\boldsymbol{q}, t)\right\} dt = -\frac{\partial L}{\partial vi}(\boldsymbol{q}, \boldsymbol{Q}-\boldsymbol{q}, 0).\end{aligned}$$

□

Given any two differentiable functions $L(\boldsymbol{q}, \boldsymbol{v}, t), f(\boldsymbol{q}, t)$, set:

$$L_f(\boldsymbol{q}, \boldsymbol{v}, t) = F(\boldsymbol{q}, \boldsymbol{v}, t) + \frac{\partial f}{\partial q}(\boldsymbol{q}, t)\boldsymbol{v} + \frac{\partial f}{\partial t}(\boldsymbol{q}, t).$$

Lemma 40.5

(40.7) (a) $\tilde{L}_f(\boldsymbol{q}, \boldsymbol{Q}) = \tilde{L}(\boldsymbol{q}, \boldsymbol{Q}) + f(\boldsymbol{Q}, 1) - f(\boldsymbol{q}, 0)$;

(40.7) (b) *If L satisfies* (40.2) (a') *then L_f satisfies it as well, for all f.*

The proof of this lemma is straightforward. We are now in position to finish the proof of Theorem 40.1. Let L be the function defined in Lemma 40.3. From (40.6) (c) and (40.4) (b), we get:

$$\frac{\partial^2 \tilde{L}}{\partial q_i \partial \boldsymbol{Q}_j}(\boldsymbol{q}, \boldsymbol{Q}) = -\frac{\partial^2 \tilde{L}}{\partial v_i \partial v_j}(\boldsymbol{q}, \boldsymbol{Q}-\boldsymbol{q}, 0) = \frac{\partial^2 S}{\partial q_i \partial Q_j}(\boldsymbol{q}, \boldsymbol{Q}),$$

and therefore

$$\tilde{L}(\boldsymbol{q}, \boldsymbol{Q}) = S(\boldsymbol{q}, \boldsymbol{Q}) + A(\boldsymbol{q}) + b(\boldsymbol{Q})$$

for some differentiable functions a and b. Set

$$f(\boldsymbol{q}, t) = (1-t)A(\boldsymbol{q}) - tb(\boldsymbol{Q}).$$

We claim that the function L_f satisfies (40.2) (a)-(d). We prove these properties one by one.

1. We proved in (40.4) (c) that L satisfies (40.2) (a'), and hence (40.2) (a). Equation (40.7) (b) proves that L_f does as well.

2. From (40.7) (a), we get:

$$\tilde{L}_f(\boldsymbol{q},\boldsymbol{Q}) = \tilde{L}(\boldsymbol{q},\boldsymbol{Q}) - b(\boldsymbol{Q}) - A(\boldsymbol{q}) = S(\boldsymbol{q},\boldsymbol{Q}),$$

which proves (40.2) (b).

3. $\dfrac{\partial^2 \tilde{L}_f}{\partial \boldsymbol{v}^2} = \dfrac{\partial^2 \tilde{L}}{\partial \boldsymbol{v}^2} = (R_{ij}) = -\dfrac{\partial^2 S}{\partial \boldsymbol{q}\partial \boldsymbol{Q}}(\boldsymbol{q} - t\boldsymbol{v}, \boldsymbol{q} + (1-t)\boldsymbol{v})$. Since this last matrix is positive definite by (40.1) , so is the first one.

4. Since $S(\boldsymbol{q}+\boldsymbol{m}, \boldsymbol{Q}+\boldsymbol{m}) = S(\boldsymbol{q},\boldsymbol{Q})$, the function L is periodic in $\boldsymbol{q}$. We need to check that $\frac{\partial f}{\partial t}$ and $\frac{\partial f}{\partial q}$ are also periodic in $\boldsymbol{q}$. Using the definitions and (40.6) (a) and (b), one can easily check that

$$\tilde{L}(\boldsymbol{q},\boldsymbol{q}) = \frac{\partial \tilde{L}}{\partial \boldsymbol{q}}(\boldsymbol{q},\boldsymbol{q}) = \frac{\partial \tilde{L}}{\partial \boldsymbol{Q}}(\boldsymbol{q},\boldsymbol{q}) = 0.$$

From the definitions of the functions a and b we obtain that

$$A(\boldsymbol{q}) + b(\boldsymbol{q}) = -S(\boldsymbol{q},\boldsymbol{q}), \quad \frac{\partial a}{\partial \boldsymbol{q}} = -\frac{\partial S}{\partial \boldsymbol{q}}(\boldsymbol{q},\boldsymbol{q}), \quad \frac{\partial b}{\partial \boldsymbol{q}}(\boldsymbol{q}) = -\frac{\partial S}{\partial \boldsymbol{Q}}(\boldsymbol{q},\boldsymbol{q}).$$

Because of the periodicity of S, all these functions are periodic in $\boldsymbol{q}$. Since

$$\frac{\partial f}{\partial t} = (1-t)\frac{\partial a}{\partial \boldsymbol{q}} - t\frac{\partial b}{\partial \boldsymbol{q}}, \quad \frac{\partial f}{\partial \boldsymbol{q}} = -a - b,$$

both $\frac{\partial f}{\partial t}$ and $\frac{\partial f}{\partial q}$ are periodic. This finishes the proof of our claim, and hence that of Theorem 40.1 . □

B. SUSPENSION WITHOUT CONVEXITY

If we let go of the symmetry of $\frac{\partial^2 S}{\partial q \partial Q}$ (but keep some form of definiteness) in Theorem 40.1 , we can still suspend the twist map F by a Hamiltonian flow. The cost is relatively high however: we can no longer insure that the Hamiltonian is convex in the fiber. The proof, quite different from that of Theorem 40.1 , first appeared in Golé (1994c) .

Theorem 40.6 *Let $F(\boldsymbol{q},\boldsymbol{p}) = (\boldsymbol{Q},\boldsymbol{P})$ be a symplectic twist map of $T^*\mathbb{T}^n$ whose differential $\boldsymbol{b}(\boldsymbol{z}) = \frac{\partial \boldsymbol{Q}(z)}{\partial p}$ satisfies:*

$$\sup_{\boldsymbol{z}\in T^*\mathbb{T}^n} \langle \boldsymbol{b}^{-1}(\boldsymbol{z})\boldsymbol{v}, \boldsymbol{v}\rangle > a\,\|\boldsymbol{v}\|, \quad a > 0,\ \forall \boldsymbol{v} \neq 0 \in \mathbb{R}^n. \tag{40.8}$$

Then F is the time 1 map of a (time dependant) Hamiltonian H.

Remark 40.7 Condition (40.8) tells us that F does not twist infinitely much.

Proof. Let $S(\boldsymbol{q},\boldsymbol{Q})$ be the generating function of F. Since $\boldsymbol{p} = -\partial_1 S(\boldsymbol{q},\boldsymbol{Q})$, we have that $\boldsymbol{b} = \partial Q/\partial \boldsymbol{p} = -\left(\partial_{12} S(\boldsymbol{q},\boldsymbol{Q})\right)^{-1}$. Hence equation (40.8) translates into:

$$\sup_{(\boldsymbol{q},\boldsymbol{Q})\in\mathbb{R}2n} \langle -\partial_{12} S(\boldsymbol{q},\boldsymbol{Q})\boldsymbol{v}, \boldsymbol{v}\rangle > a\,\|\boldsymbol{v}\|, \quad a > 0, \forall \boldsymbol{v} \neq 0 \in \mathbb{R}^n. \tag{40.9}$$

The following lemma show that (40.9) implies the hypothesis of Proposition 26.4, which in turn shows that whenever we have a function on $\mathbb{R}2n$ which is suitably periodic and satisfies (40.9) , it is the generating function for some symplectic twist map.

Lemma 40.7 *Let $\{A_x\}_{x\in\Lambda}$ be a family of $n\times n$ real matrices satisfying:*

$$\sup_{x\in\Lambda} |\langle A_x\boldsymbol{v},\boldsymbol{v}\rangle| > a\|\boldsymbol{v}\|^2, \quad \forall \boldsymbol{v}\neq 0\in \mathbb{R}^n.$$

Then :

$$\sup_{x\in\Lambda}\left\|A_x^{-1}\right\| < a^{-1}.$$

We postpone the proof of this lemma to the end. We now construct a differentiable family $S_t,\ t\in[0,1]$ of generating functions, with $S_1=S$, and then show how to make a Hamiltonian vector field out of it, whose time 1 map is F. Let

$$S_t(\boldsymbol{q},\boldsymbol{Q}) = \begin{cases} \frac{1}{2}af(t)\|\boldsymbol{Q}-\boldsymbol{q}\|^2 & \text{for } 0<t\leq\frac{1}{2} \\ \frac{1}{2}af(t)\|\boldsymbol{Q}-\boldsymbol{q}\|^2 + (1-f(t))S(\boldsymbol{q},\boldsymbol{Q}) & \text{for } \frac{1}{2}\leq t\leq 1. \end{cases}$$

where f is a smooth positive functions, $f(1)=f'(1/2)=0,\ f(1/2)=1$ and $\lim_{t\to 0^+} f(t)=+\infty$. We will ask also that $1/f(t)$, which can be extended continuously to $1/f(0)=0$, be differentiable at 0. The choice of f has been made so that S_t is differentiable with respect to t, for $t\in(0,1]$. Furthermore, it is easy to verify that:

$$\sup_{(\boldsymbol{q},\boldsymbol{Q})\in\mathbb{R}2n} \langle -\partial_{12}S_t(\boldsymbol{q},\boldsymbol{Q})\boldsymbol{v},\boldsymbol{v}\rangle > a\left\|\boldsymbol{v}\right\|^2, \quad a>0, \forall\boldsymbol{v}\neq 0\in\mathbb{R}^n, t\in(0,1].$$

Hence S_t generates a smooth family $F_t,\ t\in(0,1]$ of symplectic twist maps, and in fact $F_t(\boldsymbol{q},\boldsymbol{p}) = (\boldsymbol{q} + (af(t))^{-1}\boldsymbol{p},\boldsymbol{p}), \quad t\leq 1/2)$, so that $\lim_{t\to 0^+}F_t = Id$, in any topology that one desires (on compact sets). Let us write

$$s_t(\boldsymbol{q},\boldsymbol{p}) = S_t\circ\psi_t(\boldsymbol{q},\boldsymbol{p}),$$

where ψ_t is the change of coordinates given by the fact that F_t is twist. It is not hard to verify that $\psi_t(\boldsymbol{q},\boldsymbol{p}) = (\boldsymbol{q},\boldsymbol{q}-(af(t))^{-1}\boldsymbol{p}), \quad t\leq 1/2$. so that:

$$s_t(\boldsymbol{q},\boldsymbol{p}) = \frac{1}{2}(af(t))^{-2}\|\boldsymbol{p}\|^2$$

In particular, by our assumption on $1/f(t)$, s_t can be differentiably continued for all $t\in[0,1]$, with $S_0\equiv 0$. Hence, in the $\boldsymbol{q},\boldsymbol{p}$ coordinates, we can write:

$$F_t^*\boldsymbol{p}d\boldsymbol{q} - \boldsymbol{p}d\boldsymbol{q} = ds_t, \quad t\in[0,1].$$

By Theorem 47.7, F_t is a Hamiltonian isotopy. □

Proof of Lemma 40.7

For all non zero $\boldsymbol{v}\in\mathbb{R}^n$, we have:

$$\inf_{x\in\Lambda}\frac{|\langle A_x\boldsymbol{v},\boldsymbol{v}\rangle|}{\|\boldsymbol{v}\|^2} > a$$

But:

$$\inf_{\boldsymbol{v}\in\mathbb{R}^n}\frac{|\langle A_x\boldsymbol{v},\boldsymbol{v}\rangle|}{\|\boldsymbol{v}\|^2} = \inf_{\|\boldsymbol{v}\|=1}|\langle A_x\boldsymbol{v},\boldsymbol{v}\rangle| \leq \inf_{\|\boldsymbol{v}\|=1}\|A_x\boldsymbol{v}\|$$

so that $\inf_{x\in\Lambda}\inf_{\|\boldsymbol{v}\|=1}\|A_x\boldsymbol{v}\| > a$. But:

$$\inf_{\|\boldsymbol{v}\|=1}\|A_x\boldsymbol{v}\| = \inf_{\boldsymbol{v}\in\mathbb{R}^n-\{0\}}\frac{\|A_x\boldsymbol{v}\|}{\|\boldsymbol{v}\|} = inf_{\boldsymbol{v}\in\mathbb{R}^n-\{0\}}\frac{\|\boldsymbol{v}\|}{\left\|A_x^{-1}\boldsymbol{v}\right\|}$$

so that, finally:

$$\sup_{x\in\Lambda}\left\|A_x^{-1}\right\| = \left(\inf_{x\in\Lambda}\inf_{\boldsymbol{v}\in\mathbb{R}^n-\{0\}}\frac{\|\boldsymbol{v}\|}{\|A_x\boldsymbol{v}\|}\right)^{-1} < a^{-1}.$$

□

41.1 Return Maps In Hamiltonian Systems

A. RETURN MAPS OF HAMILTONIAN SYSTEMS ARE SYMPLECTIC

Consider a time independent Hamiltonian on $\mathbb{R}^{2n+2}$, with its standard symplectic structure $\Omega_0 = \sum_{k=0}^n dq_k \wedge dp_k$. Assume that we have a periodic trajectory γ for the Hamiltonian flow. It must then lie in the energy level $H = H(\gamma(0))$, since H is time independent. Take any $2n+1$ dimensional open disk $\tilde{\Sigma}$ which is transverse to γ at $\gamma(0)$, and such that $\tilde{\Sigma}$ intersects γ only at $\gamma(0)$.

Fig. 41. 2.

Such a disk clearly always exists, if γ is not a fixed point. In fact, one can assume that, in a local Darboux chart, $\tilde{\Sigma}$ is the hyperplane with equation $q_0 = 0$: this is because in the construction of Darboux coordinates, one can start by choosing an arbitrary nonsingular differentiable function as one of the coordinate function (see Arnold (1978), section 43, or Weinstein (1979) , Extension Theorem, lecture 5.)

Define $\Sigma = \tilde{\Sigma} \cap \{H = H_0\}$. It is a standard fact (true for periodic orbits of *any* C^1 flow) that the Hamiltonian flow h^t admits a Poincaré return map $\mathcal{R}$, defined on Σ around $\boldsymbol{z}_0$, by $\mathcal{R}(\boldsymbol{z}) = h^{t(\boldsymbol{z})}(\boldsymbol{z})$, where $t(\boldsymbol{z})$ is the first return time of $\boldsymbol{z}$ to Σ under the flow (see Hirsh & Smale (1974), Chapter 13).

We claim that $\mathcal{R}$ is symplectic, with the symplectic structure induced by Ω_0 on Σ.

Since $\tilde{\Sigma}$ is transverse to γ, we may assume that:

$$\dot{q}_0 = \frac{\partial H}{\partial p_0} \neq 0$$

on $\tilde{\Sigma}$. Hence, by the Implicit Function Theorem, the equation

$$H(0, q_1 \dots, q_n, p_0, \dots, p_n) = H_0$$

implies that p_0 is a function of $(q_1, \dots, q_n, p_1, \dots, p_n)$. This makes the latter variables a system of local coordinates for Σ, and since $dq_0 = 0$ on Σ, the restriction of Ω_0 is in fact

$$\omega = \Omega_0\big|_{\Sigma} = \sum_{k=1}^{n} dq_k \wedge dp_k.$$

To prove that $\mathcal{R}$ is symplectic, remember that, by (41.-1) , for any closed curve in Σ, or more generally for any closed 1–chain c in Σ,

$$\int_{\mathcal{R}c} \boldsymbol{p}d\boldsymbol{q} - Hdt = \int_{c} \boldsymbol{p}d\boldsymbol{q} - Hdt$$

since c and $\mathcal{R}c$ are on the same trajectory tube. Here $\mathcal{R}c$ represent the chain in $\mathrm{I\!R}^{2n+2} \times \mathrm{I\!R}$ given by $(\mathcal{R}(c(s)), t^{c(s)})$.

This equality implies that the function $S(\boldsymbol{z}) = \int_{\boldsymbol{z}_0}^{\boldsymbol{z}} \mathcal{R}^*(\boldsymbol{p}d\boldsymbol{q} - Hdt) - (\boldsymbol{p}d\boldsymbol{q} - Hdt)$ is well defined. But, on Σ, the differential of the form inside this integral is $\mathcal{R}^*\omega - \omega$, since both dq_0 and dH are zero there. Hence $\mathcal{R}^*\omega - \omega = d^2 S = 0$, i.e., $\mathcal{R}$ is symplectic. □

B. TWISTING AROUND ELLIPTIC FIXED POINTS

We now follow Moser (1977). If 0 is an elliptic fixed point, that is $DR(0)$ has all its eigenvalues on the unit circle, a normal form theorem ???(find ref.) says that (generically?) the map R is, around 0 given by:

$$\begin{aligned} Q_k &= q_k cos\Phi_k(\boldsymbol{q},\boldsymbol{p}) - p_k sin\Phi_k(\boldsymbol{q},\boldsymbol{p}) + f_k(\boldsymbol{q},\boldsymbol{p}) \\ P_k &= q_k sin\Phi_k(\boldsymbol{q},\boldsymbol{p}) + p_k cos\Phi_k(\boldsymbol{q},\boldsymbol{p}) + g_k(\boldsymbol{q},\boldsymbol{p}) \\ \Phi_k(\boldsymbol{q},\boldsymbol{p}) &= \alpha_k + \sum_{l=1}^{n} \beta_{kl}(q_l^2 + p_l^2). \end{aligned}$$

where the error term f_k, g_k are C^1 and have vanishing derivatives up to order 3 at the origin. We now show how this map is, in "polar coordinates" a symplectic twist map of $T^*\mathbb{T}^n$, *whenever the matrix* $\{\beta_{kl}\}$ *is non singular*. Let V be a punctured neighborhood of 0 such that: $0 < \sum_k (q_k^2 + p_k^2) < \epsilon$. We introduce on V new coordinates (r_k, θ_k) by:

$$\boldsymbol{q}_k = \sqrt{2r_k\epsilon}\cos 2\pi\theta_k' \quad p_k = \sqrt{2r_k\epsilon}\sin 2\pi\theta_k$$

where θ_k is determined modulo 1. One can check that V is transformed into the "annular" set:

$$U = \{(r_k, \theta_k) \in \mathbb{T}^n \times \mathrm{I\!R}^n \mid \sum_k \left(2r_k - \frac{1}{2n}\right)^2 < \frac{1}{4n^2}\}$$

Since the symplectic form $d\boldsymbol{q} \wedge d\boldsymbol{p}$ is transformed into $\epsilon d\boldsymbol{r} \wedge d\boldsymbol{\theta}$, R remains symplectic in these new coordinates, with the symplectic form $d\boldsymbol{r} \wedge d\boldsymbol{\theta}$. In fact, it is exact symplectic in U. Remember that to check this, it is enough to show that, for any closed curve γ:

$$\int_{R\gamma} \boldsymbol{r}d\boldsymbol{\theta} = \int_{\gamma} \boldsymbol{r}d\boldsymbol{\theta}.$$

It is easy to see that $2\epsilon r_k d\theta_k = p_k dq_k - q_k dp_k$, so by Stokes' theorem:

$$2\epsilon \int_{\gamma} \boldsymbol{r}d\boldsymbol{\theta} = \int_{\partial D} \boldsymbol{p}d\boldsymbol{q} - \boldsymbol{q}d\boldsymbol{p} = -2\int_D \omega$$

where D is a 2 manifold in V with boundary $\partial D = \gamma$. Since R preserves ω in V, it must preserve the last integral, and hence the first. To see that R satisfies the two other conditions for being a symplectic twist map, we just write it in the new coordinates:

$$\begin{aligned}\Theta_k &= \theta_k + \psi_k(\boldsymbol{r}) + o_1(\epsilon) \\ R_k &= r_k + o_1(\epsilon) \\ \psi_k &= \alpha_k + \epsilon \sum_{l+1}^{n} 2\beta_{kl} r_l .\end{aligned}$$

where $\epsilon^{-1} o_1(\epsilon, \boldsymbol{\theta}, \boldsymbol{r})$ and its first derivatives in $\boldsymbol{r}, \boldsymbol{\theta}$ tend to 0 uniformally as $\epsilon \to 0$. We can rewrite this as:

$$\mathcal{R}(\boldsymbol{\theta}, \boldsymbol{r}) = (\boldsymbol{\theta} + \epsilon \boldsymbol{B}\boldsymbol{r} + \boldsymbol{\alpha} + o_1(\epsilon), \boldsymbol{r} + o_1(\epsilon)) .$$

So for small ϵ, the condition $\det\, \partial\boldsymbol{\Theta}/\partial\boldsymbol{r} \neq 0$ is given by the nondegeneracy of $\boldsymbol{B} = \{\beta_{kl}\}$, one uses the fact that $\mathcal{R}$ is C^1 close to a completely integrable symplectic twist map to show that $\mathcal{R}$ is twist in U (the twist condition is open.) The fact that it is homotopic to Id derives from Exercise 23.2.

Note that the set V and therefore U are not invariant under $\mathcal{R}$. However, it is still possible to show the existence of infinitely many periodic points for $\mathcal{R}$: this is the content of the Birkhoff– Lewis theorem (see Moser (1977)) .

Remarks HAMrem and HAMgrad are 39.8 , Corollary HAMpartial is 38.4 , Proposition HAMdecompone and HAMdecomptwo are (39.8) and 39.9 , Theorem HAMexp is 38.1 , HAMhamstm is 39.2 . Section HAMsecgeom is 38.0, Theorem HAMthmbp is 40.1