7

HAMILTONIAN SYSTEMS VS. TWIST MAPS

In this chapter, we explore the relationship between Hamiltonian systems and symplectic twist maps. We will assume that the reader is familiar with the material reviewed in Appendix 1, which introduces Hamiltonian systems in cotangent bundles and some of their fundamental properties. 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. This method generalizes 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.

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. Almost no knowledge of Riemannian geometry is assumed here. 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 39. D 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 . In Section 40, 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 41 that in many cases, a symplectic twist map may be written as the time 1 of a (time dependent) Hamiltonian system. Most of this last section is courtesy of M. Bialy and L. Polterovitch, who graciously let us publish their proof of suspension of symplectic twist maps for the first time in this book.

38. Case Study: The Geodesic Flow

A. A Few Facts About Riemannian Geometry

Hamiltonian Approach to the Geodesic Flow. 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*:

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

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, thanks to the Cauchy-Schwartz inequality, action extremals are length extremals and vice versa (with the difference that action extremals come with a specified parameterization, see Milnor (1969)). 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 59 of Appendix 1, we use the Legendre transform to compute the corresponding Hamiltonian function. In local coordinates $\boldsymbol{q}$ in M, we can write

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

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

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

In particular, $\frac{\partial^2 L_0}{\partial \boldsymbol{v}^2}$ is nondegenerate. Hence the Legendre condition (see Appendix 1) 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 geodesics are parameterized at constant speed.

The *geodesic flow* is the Hamiltonian flow h_0^t 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 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 $\mathcal{L}^{-1}h_0^t\mathcal{L}$ 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 Section 59 in Appendix 1). In the present case, if $(\boldsymbol{q}(t), \boldsymbol{p}(t))$ is 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 $(\boldsymbol{q}_0, \boldsymbol{p}_0) = (\boldsymbol{q}(0), A^{-1}_{\boldsymbol{q}_0}\dot{\boldsymbol{q}}(0))$.

Exponential Map. We now want to establish a fundamental result of Riemannian 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 $\boldsymbol{q}(0) = \boldsymbol{q}_0$ and $\dot{\boldsymbol{q}}(0) = \boldsymbol{v}$. 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 *Let M be a compact Riemannian manifold. The map $Exp : TM \to M \times M$*

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

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:

$$\text{Dis}(\boldsymbol{q}, exp_{\boldsymbol{q}}(\boldsymbol{v})) = \|\boldsymbol{v}\| \,. \tag{38.2}$$

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 $\boldsymbol{Q}$. One way to paraphrase this theorem is by saying that, any two close by points are joined by a unique, short enough, geodesic segment.

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

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

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

We now show that Exp is an embedding when restricted to $V_\epsilon = \{(\boldsymbol{q}, \boldsymbol{v}) \in TM \mid \|\boldsymbol{v}\| \leq \epsilon\}$, where ϵ is as above. Since we proved that Exp is a local diffeomorphism on V_ϵ, it is enough to check the injectivity. Let two elements in V_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 V_ϵ. But, by our choice of V_ϵ this implies these elements are the same.

Finally, we show that $\text{Dis}(\boldsymbol{q}, exp_{\boldsymbol{q}}(\boldsymbol{v})) = \|\boldsymbol{v}\|$ whenever $\|\boldsymbol{v}\| \leq \epsilon$. As a length minimizer, the shortest path giving the distance between two points is also an action minimizer, and hence a geodesic. Since Exp is an embedding of V_ϵ in $M \times M$, exp is one to one on $V_\epsilon \cap T_{\boldsymbol{q}}M$ and the unique geodesic that joins $\boldsymbol{q}$ and $exp_{\boldsymbol{q}}(\boldsymbol{v})$ in $exp(V_\epsilon \cap T_{\boldsymbol{q}}M)$ 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 38.2c)). The only way Formula (38.2) may fail is if there were a shorter geodesic joining $\boldsymbol{q}$ and $exp_{\boldsymbol{q}}(\boldsymbol{v})$ not in $exp(V_\epsilon \cap T_{\boldsymbol{q}}M))$. 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 38.2 a) Check that, in local coordinates, Hamilton's equations for the geodesic flow write:

$$\begin{aligned}\dot{\boldsymbol{q}} &= A_{(\boldsymbol{q})}\boldsymbol{p} \\ \dot{\boldsymbol{p}} &= -\left\langle \frac{\partial A_{(\boldsymbol{q})}}{\partial \boldsymbol{q}}\boldsymbol{p}, \boldsymbol{p} \right\rangle\end{aligned} \tag{38.3}$$

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 $\boldsymbol{q}(t) = exp_{q_0}(t\boldsymbol{v})$, $\|\dot{\boldsymbol{q}}(t)\| = \|\boldsymbol{v}\|$ for all t.
d) Show that $\mathrm{Dis}(\boldsymbol{q}(0), \boldsymbol{q}(t)) = |t|\, \|\boldsymbol{p}(0)\|$

Exercise 38.3 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 38.3 is the key to the following:

Proposition 38.4 *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 $U_\epsilon = \{(\boldsymbol{q},\boldsymbol{p}) \in T^*M \mid \|\boldsymbol{p}\| \le \epsilon\}$, for ϵ small enough. More generally, given any $R > 0$, there is a $t_0 > 0$ (or given any t_0 there is an R) such that, for any $t \in [-t_0, t_0]$, h_0^t is a symplectic twist map on the set $U_R = \{(\boldsymbol{q},\boldsymbol{p}) \mid \mid \|\boldsymbol{p}\| \le R\}$. 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 (see Theorem 59.7 in Appendix 1). Define $Exp^\# = Exp \circ \mathcal{L}^{-1}$. By Theorem 38.3, $Exp^\#$ 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,

$$(38.4) \qquad (h_0^1)^* \boldsymbol{p}d\boldsymbol{q} - \boldsymbol{p}d\boldsymbol{q} = dS, \quad \text{with} \quad S(\boldsymbol{q},\boldsymbol{p}) = \int_\gamma \boldsymbol{p}d\boldsymbol{q} - H_0 dt$$

where γ is the curve $h_0^t(\boldsymbol{q},\boldsymbol{p}),\ t \in [0,1]$ (see Theorem 59.7 in Appendix 1). 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 38.2) and $H_0 = \frac{1}{2}\langle A_{(\boldsymbol{q})}\boldsymbol{p},\boldsymbol{p}\rangle = \frac{1}{2}\|\boldsymbol{p}\|^2$, the integral simplifies:

$$(38.5) \qquad \int_\gamma \boldsymbol{p}d\boldsymbol{q} - H_0 dt = \int_0^1 \frac{1}{2}\langle A_{(\boldsymbol{q})}\boldsymbol{p},\boldsymbol{p}\rangle - \frac{1}{2}\|\boldsymbol{p}(t)\|^2 dt = \int_0^1 \frac{1}{2}\|\boldsymbol{p}(t)\|^2 dt.$$

But the integrand is H_0, which is constant along γ. Hence, using Theorem 38.3, 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 $\boldsymbol{Q}$ in the region where $(\boldsymbol{q},\boldsymbol{p}) \mapsto (\boldsymbol{q},\boldsymbol{Q})$ is a diffeomorphism. $\square$

Remark 38.5 1) Note that the proof of Proposition 38.4 equates the action of a geodesic segment between two points to the generating function evaluated at this pair of points.
2) 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 $\frac{1}{2}\mathrm{Dis}^2(0,\theta)$ at $\theta = \pi$ is π, whereas the right derivative is $-\pi$: the function Dis^2 is not differentiable at this point.

The following will be instrumental in the proof of Theorem 43.1.

Corollary 38.6 *Let* $h_0^s(\boldsymbol{q},\boldsymbol{p}) = (\boldsymbol{Q}_s, \boldsymbol{P}_s)$ *be the time* s *of the geodesic flow, then:*

$$\partial_1 \mathrm{Dis}(\boldsymbol{q},\boldsymbol{Q}_s) = -sign(s).\frac{\boldsymbol{p}}{\|\boldsymbol{p}\|} \quad and \quad \partial_2 \mathrm{Dis}(\boldsymbol{q},\boldsymbol{Q}_s) = sign(s).\frac{\boldsymbol{P}_s}{\|\boldsymbol{P}_s\|} \tag{38.6}$$

Proof. From Proposition 38.4, 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 (see Milnor (1969)). As before, let h_0^1 be the time 1 map[14] of the geodesic flow with Hamiltonian H_0. Fix some neighborhood U of the zero section in T^*M. Proposition 38.4 implies 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, $i.e.$ 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 the set $X_N(U)$ of sequences in M such that $(\boldsymbol{q}_k,\boldsymbol{q}_{k+1}) \in \psi(U)$, where we write $\psi = \psi_{h_0^{\frac{1}{N}}}$. 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 $c_k\colon [\frac{k}{N}, \frac{k+1}{N}] \to T^*M$ of the geodesic flow that joins $(\boldsymbol{q}_k,\boldsymbol{p}_k)$ to $(\boldsymbol{q}_{k+1},\boldsymbol{P}_k)$.

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

The projection $\pi(c_k)$ on M is a geodesic, parameterized at constant speed equal to the norm of $\boldsymbol{p}_k$. As noted in Remark 38.5, $S(\boldsymbol{q}_k, \boldsymbol{q}_{k+1}) = \frac{1}{2}\mathrm{Dis}^2(\boldsymbol{q}_k, \boldsymbol{q}_{k+1})$ is also the action of c_k: $S(\boldsymbol{q}_k, \boldsymbol{q}_{k+1}) = \int_{c_k} \boldsymbol{p}d\boldsymbol{q} - Hdt$. Hence W is the sum of the actions of the c_k's, *i.e.* the action of the curve C obtained by the concatenation of the c_k's.

The curve C can be described as a *broken geodesic*: in general it 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, $\boldsymbol{P}_k = \boldsymbol{p}_{k+1}$ (see Remark 23.3 and Exercise 26.4), 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 the *finite dimensional* subspace of broken geodesics, which is parameterized by elements of $X_N(U)$, 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[15] 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 Section 16 in Milnor (1969)). 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.

D. The Standard Map on Cotangent Bundles of Hyperbolic Manifolds

In this subsection, we use our understanding of the relation between geodesic flow and symplectic twist maps to define the Standard Map on the cotangent bundle of any compact hyperbolic manifold. Recall that a *hyperbolic manifold* M of dimension n is a manifold that can be covered by the hyperbolic half space $\mathbb{H}^n = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_n > 0\}$ given the Riemannian metric $ds^2 = \frac{1}{x_n^2}\sum_1^n dx_k^2$, which has constant negative curvature. Geodesics on $\mathbb{H}^n$ are open semi circles or straight lines perpendicular to the boundary $\{x_n = 0\}$. Here, the relevant property of the geometry of $\mathbb{H}^n$, and hence of any hyperbolic manifold, is that the exponential map at each point is a *global* diffeomorphism between the fiber and $\mathbb{H}^n$, a corollary of the Hopf-Rinow Theorem (Gallot, Hulin and Lafontaine (1987),

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

p. 99). The generalization of the standard map that we present now is in fact valid on any Riemannian manifold with the property that $T^*\tilde{M} \cong \tilde{M} \times \tilde{M}$.

Proposition 38.7 *Let $S : \mathbb{H}^n \times \mathbb{H}^n \to \mathbb{R}$ be given by:*

$$S(\boldsymbol{q}, \boldsymbol{Q}) = \frac{1}{2}\mathrm{Dis}^2(\boldsymbol{q}, \boldsymbol{Q}) + V(\boldsymbol{q}),$$

*where $V : \mathbb{H}^n \to \mathbb{R}$ is some C^2 function, and Dis is the distance given by the hyperbolic metric. Then S is the generating function for a symplectic twist map that we call the **generalized standard map** on $\mathbb{H}^n$. Furthermore, if V is equivariant under a group of isometries Σ of $\mathbb{H}^n$ representing the fundamental group of the hyperbolic manifold $M = \mathbb{H}^n/\Sigma$, then S is the generating function for a lift of a symplectic twist map on T^*M.*

Proof. We show that S complies with the hypothesis of Proposition 26.2 where we take $M = \mathbb{H}^n, U = T^*\mathbb{H}^n \cong \mathbb{H}^n \times \mathbb{R}^n$. Let $h_0^1(\boldsymbol{q}, \boldsymbol{p}) = (\boldsymbol{Q}, \boldsymbol{P})$ be the time 1 map of the geodesic flow on $T^*\mathbb{H}^n$. The assumption that the exponential is a *global* diffeomorphism for this metric means that $\boldsymbol{p} \to \boldsymbol{Q}(\boldsymbol{q}_0, \boldsymbol{p})$ is a *global* diffeomorphism $\{\boldsymbol{q}_0\} \times \mathbb{R}^n \to \mathbb{H}^n$ for each fixed $\boldsymbol{q}_0$ and thus h_0^1 is a (global) symplectic twist map . Likewise $\boldsymbol{P} \to \boldsymbol{q}(\boldsymbol{Q}_0, \boldsymbol{P})$ is a diffeomorphism because h_0^{-1}, the inverse of a symplectic twist map is a symplectic twist map itself. Since, according to Proposition 38.4, $\frac{1}{2}\mathrm{Dis}^2$ is the generating function for h_0^1, we have established that the maps $\boldsymbol{Q} \mapsto \partial_1 \frac{1}{2}\mathrm{Dis}^2(\boldsymbol{q}_0, \boldsymbol{Q})$ and $\boldsymbol{q} \mapsto \partial_2 \frac{1}{2}\mathrm{Dis}^2(\boldsymbol{q}, \boldsymbol{Q}_0)$ are both diffeomorphisms for each *fixed* $\boldsymbol{q}_0, \boldsymbol{Q}_0$. Coming back to our full generating function, we have proven that:

$$\boldsymbol{q} \mapsto \partial_2 S(\boldsymbol{q}, \boldsymbol{Q}_0) = \partial_2 \frac{1}{2}\mathrm{Dis}^2(\boldsymbol{q}, \boldsymbol{Q}_0)$$

is a diffeomorphism.

$$\boldsymbol{Q} \mapsto \partial_1 S(\boldsymbol{q}_0, \boldsymbol{Q}) = \partial_1 \frac{1}{2}\mathrm{Dis}^2(\boldsymbol{q}_0, \boldsymbol{Q}) + dV(\boldsymbol{q}_0)$$

must also be a diffeomorphism $\mathbb{H}^n \to T_{\boldsymbol{q}_0}\mathbb{H}^n$ since we added a translation by the constant $dV(\boldsymbol{q}_0)$ to a diffeomorphism. Proposition 26.2 concludes the proof that S is the generating function for a twist map of $T^*\mathbb{H}^n$. The last statement of the proposition is an easy consequence of Exercise 26.5. □

39. Decomposition of Hamiltonian Maps into Twist Maps

In Subsection A, we generalize Theorem 38.4 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 (Subsection B), giving a generalization of the fundamental Theorem 38.1. In Subsection C, 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. Finally, in Subsection D we prove a theorem of decomposition of Hamiltonian maps into symplectic twist maps , whether or not they satisfy the Legendre condition.

A. Legendre Condition Vs. Twist Condition

Heuristics. Remember that Hamiltonian maps, which are time t maps of Hamiltonian systems, are exact symplectic (Theorem 59.7) 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 map 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 (1986a) 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.

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

$$\boldsymbol{p} \mapsto H_{\boldsymbol{p}}(\boldsymbol{q},\boldsymbol{p},t) \tag{39.1}$$

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). We let the reader check that, although we have written it in a chart of conjugate coordinates in T^*M, this condition is coordinate independent. We give examples of systems satisfying these conditions after the proof of the theorem.

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

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

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 this section, we fix a and abbreviate h_a^{a+t} by h^t.

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 $\pi \circ h^{[0,\epsilon_3]}(\pi^{-1}(B_i) \cap U(K)) \subset B_i'$. Such an ϵ_3 exists since there are finitely many B_i's and 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 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 $U(K)$: the inverse is then automatically continuous. 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 guarantees that R is smooth in all its variables. Indeed, its precise expression is (see *eg.* 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. □

Examples 39.2 We give two classes of examples. In the first class, the Hamiltonian is not assumed to be convex. We characterize the Hamiltonians in local charts of a cotangent bundles. Again, the following conditions are coordinate independent.

• Let $H(\boldsymbol{q},\boldsymbol{p},t) = \frac{1}{2}\langle A_{(\boldsymbol{q},t)}\boldsymbol{p},\boldsymbol{p}\rangle + V(\boldsymbol{q},t)$ 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 independence of $\boldsymbol{p}$). Hence this class contains, but is substantially larger than, the classical mechanical systems.

• 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 satisfies the global Legendre condition. This is a direct consequence of Lemma 25.4. If we remove the lower bound on the smallest eigenvalue, one can show (see Exercise 39.3) 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 39.1, are also satisfied if $H_{\boldsymbol{pp}}$ is positive on a compact set U invariant under the flow (see Exercise 39.4).

Exercise 39.3 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. Give an example where this embedding is not onto.

Exercise 39.4 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 39.1).

B. Lagrangian Formulation Of Theorem 39.1

The following proposition, which is a reformulation of Theorem 39.4 in Lagrangian terms, is a generalization of the fundamental Theorem 38.1. It guarantees the existence and *uniqueness* of Euler-Lagrange solutions between any two close by points. A time that the solution is traversed has to be specified within a compact interval. In Chapter 9, 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 39.5 *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 K on the velocity, 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 $\|\dot{\boldsymbol{q}}(a)\| \leq K$.*

Remark 39.6 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 differences, and sources of confusion, when trying to generalize 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 59 of Appendix 1). H satisfies the global Legendre condition and $\mathcal{L}^{-1}(\boldsymbol{q}, \boldsymbol{p}) = (\boldsymbol{q}, H_{\boldsymbol{p}})$ (see Remark 59.1). In particular Theorem 39.1 applies to the Hamiltonian H. Let

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

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

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

in the tangent bundle. Theorem 39.1 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 $\|\dot{\boldsymbol{q}}(a)\| = \|H_{\boldsymbol{p}}\| \leq K$. As in the proof of Theorem 39.1, 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 $\epsilon^2 R$, $\boldsymbol{q}$ still is 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 39.1: 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 $(\boldsymbol{q}, \boldsymbol{p}) \in V(K)$, $\|\dot{\boldsymbol{q}}(a)\| = \|H_p(\boldsymbol{q}, \boldsymbol{p}, a)\| \leq K$. Finally, this solution is unique. Otherwise, by the uniqueness of solutions of O.D.E.'s, there would be $\boldsymbol{p} \neq \boldsymbol{p}'$ such that $\pi \circ h^\epsilon(\boldsymbol{q}, \boldsymbol{p}) = \pi \circ h^\epsilon(\boldsymbol{q}, \boldsymbol{p}')$, a contradiction to the twist condition. $\square$

C. Global Twist: The Case Of The Torus

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

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

Here A denotes a constant matrix, $\boldsymbol{c}$ a constant in $\mathbb{R}^n$ and K_1 a positive real. We stress that, in Assumption 2, A (and hence $H_{\boldsymbol{pp}}$) is *not* necessarily positive definite.

Theorem 39.7 *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 39.8 Proposition 39.7 holds for $h_a^{a+\epsilon}$ whenever it does for h^ϵ: $h_a^{a+\epsilon}$ 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 Assumption 2. The strategy is to estimate $\left\|\frac{\partial \boldsymbol{q}(\epsilon)}{\partial \boldsymbol{p}}^{-1}\right\|$ and use Proposition 25.3 to turn the local Assumptions 1 and 2 into global twist condition. 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* [16]

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

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

Lemma 39.9 *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 :

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

[16] 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)).

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

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

and hence:

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

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

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

and we conclude by setting $\epsilon = |t - t_0|$. □

We now proceed with the proof of Theorem 39.7 . By Lemma 39.9 we can write: $U(s) = Id + O_1(s)$ where $\|O_1(s)\| < 2Ks$, for s small enough ($i.e.$ such that $e^{Ks} < 2$). Integrating Equation (39.2) on both sides then yields:

$$U(\epsilon) = Id + \int_0^{\epsilon} J\nabla^2 H(h^s(\boldsymbol{z})).(Id + O_1(s))ds \tag{39.3}$$

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 $U(\epsilon)$. From Equation (39.3), and Assumption 1 (1) we know it is of the form:

$$\boldsymbol{b}_\epsilon(\boldsymbol{z}) = \int_0^{\epsilon} H_{\boldsymbol{pp}}(h^s(\boldsymbol{z}))ds + \int_0^{\epsilon} O_2(s)ds \tag{39.4}$$

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

$$C\|\boldsymbol{v}\|^2 < \langle H_{\boldsymbol{pp}}(\boldsymbol{z})\boldsymbol{v}, \boldsymbol{v}\rangle < C^{-1}\|\boldsymbol{v}\|^2, \tag{39.5}$$

we deduce that:

$$(\epsilon C - K^2\epsilon^2)\|\boldsymbol{v}\|^2 < \langle \boldsymbol{b}_\epsilon(\boldsymbol{z})\boldsymbol{v}, \boldsymbol{v}\rangle < (\epsilon C^{-1} + K^2\epsilon^2)\|\boldsymbol{v}\|^2 \tag{39.6}$$

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

$$\sup_{\boldsymbol{z}\in\mathbb{R}^{2n}} \left\|\boldsymbol{b}_\epsilon^{-1}(\boldsymbol{z})\right\| < K', \tag{39.7}$$

for some positive K'. We can now apply Proposition 25.3 to show that h^ϵ is a symplectic twist map with a generating function S defined on all of $\mathbb{R}^{2n}$.

Remark 39.10 The above proof shows that h^ϵ satisfies a certain convexity condition :

$$\left\langle \boldsymbol{b}_\epsilon^{-1}\boldsymbol{v}, \boldsymbol{v} \right\rangle = \left\langle \left(\frac{\partial \boldsymbol{q}}{\partial \boldsymbol{p}}(\epsilon) \right)^{-1} \boldsymbol{v}, \boldsymbol{v} \right\rangle \geq a \left\| \boldsymbol{v} \right\|^2, \quad \forall \boldsymbol{v} \in \mathbb{R}^n. \tag{39.8}$$

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

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

and M the corresponding sup, (39.6) implies:

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

We now adapt the above proof to Assumption 2. Note that under this assumption, we can still derive (39.4) : the boundary condition (2) implies that $\nabla^2 H$ is bounded. Since H is C^2, and $H_{\boldsymbol{pp}} = A$ outside a compact set, $H_{\boldsymbol{pp}}(h^s\boldsymbol{z})$ is uniformly close to $H_{\boldsymbol{pp}}(\boldsymbol{z})$ for small s, and thus the first matrix integral in (39.4) is non singular for $\boldsymbol{z}$ and small s. Thus $\boldsymbol{b}_\epsilon(\boldsymbol{z})$ is also nonsingular for small ϵ. Since $\boldsymbol{b}_\epsilon(\boldsymbol{z}) = \epsilon A$ outside of the compact set $\|\boldsymbol{p}\| \leq K_1$, the set of matrices $\{\boldsymbol{b}_\epsilon(\boldsymbol{z}) \mid \boldsymbol{z} \in \mathbb{R}^n\}$ is compact and hence Inequality (39.7) holds and, again, we can apply Proposition 25.3 . $\square$

D. Decomposition Of Hamiltonian Maps Into Twist Maps

When the time ϵ maps of a Hamiltonian system are symplectic twist maps for $\epsilon < \epsilon^*$, one can readily decompose the time 1 map into such twist maps. Take a time independent Hamiltonian, for example. Its time 1 map h^1 can be written:

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

and, for $N > 1/\epsilon^*$, 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:

$$h^1 = h^1_{\frac{N-1}{N}} \circ (h^{\frac{N-1}{N}}_{\frac{N-2}{N}}) \circ \ldots h^{\frac{k+1}{N}}_{\frac{k}{N}} \circ \ldots h_0^{\frac{1}{N}} \tag{39.9}$$

and each $h_{\frac{k}{N}}^{\frac{k+1}{N}}$ is a symplectic twist map by assumption on our Hamiltonian. as the next Proposition shows. What may be more surprising, and gives a greater scope to the use of symplectic twist maps, 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 (1991) applied in his variational proof of the Poincaré-Birkhoff Theorem (see Chapter 1). This will work with either of the following 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 $\sup \left\|\nabla^2 H_t\right\| < K$. The domain $U = T^*\mathbb{T}^n$.

Proposition 39.11 (Decomposition) *Let $H(\boldsymbol{z},t)$ be a Hamiltonian function satisfying Assumptions 3 or 4, or the hypothesis of either Theorem 39.1 or Theorem 39.7. Then h^1, the time 1 map of its corresponding Hamiltonian system, can be decomposed into a finite product of symplectic twist maps (defined on the domain U corresponding to the various assumptions):*

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

Proof. We have given the trivial proof above for Hamiltonians that satisfies the hypothesis of Theorems 39.1 and 39.7. 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 a symplectic twist map on $U(K)$. That such an s exists is proven in Proposition 38.4. We can write:

(39.10)

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

One can check that, at each successive step of the composition, the points remain in $U(K)$. The map G is a symplectic twist map, by assumption, and $G^{-1} \circ h_{\frac{k}{N}}^{\frac{k+1}{N}}$ is also a symplectic twist map by openness of the set of twist maps on a compact neighborhood (see Exercise 26.6).

Suppose now that H satisfies Assumption 4. Let $G(\boldsymbol{q}, \boldsymbol{p}) = (\boldsymbol{q} + \boldsymbol{p}, \boldsymbol{p})$, our favorite symplectic twist map on $T^*\mathbb{T}^n$. Decompose h^1 as in Equation (39.10). We now show that $G^{-1} \circ h_{\frac{k}{N}}^{\frac{k+1}{N}}$ is also a symplectic twist map. Lemma 39.6 implies that $h_t^{t+\epsilon}$ satisfies $\left\|Dh_t^{t+\epsilon} - Id\right\| < \epsilon K e^{K\epsilon}$. Hence

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

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

40. Return Maps in Hamiltonian Systems

We show that return maps around a periodic orbit of a Hamiltonian system is exact symplectic. If the periodic orbit is elliptic, the return map has an elliptic fixed point, and thus, generically, it is a symplectic twist map around this point (see Section 91).

Consider a time independent Hamiltonian on $\mathbb{R}^{2n+2}$, with the 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 $\{H = H_0\}$ where $H_0 = H(\gamma(0))$. 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 $\gamma(0)$. 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. We will now work in a *simply connected* neighborhood $\mathcal{O}$ of z_0 in Σ, parameterized by $(q_1, \dots, q_n, p_1, \dots, p_n)$. Since $dq_0 = 0$ in $\mathcal{O}$, the restriction of Ω_0 is in fact

$$\omega \overset{def}{=} \Omega_0\big|_{\mathcal{O}} = \sum_{k=1}^{n} dq_k \wedge dp_k.$$

To prove that $\mathcal{R}$ is exact symplectic, use Formula 59.9 of Appendix 1which states that, for any closed curve in $\mathcal{O}$, or more generally for any closed 1–chain c in $\mathcal{O}$,

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

since c and $\mathcal{R}c$ are on the same trajectory tube. We now show that $\int_{\mathcal{R}c} Hdt = \int_c Hdt = 0$. This is due to the fact that $d(Hdt) = dH \wedge dt = 0 \wedge dt = 0$ since $H = H_0$ on $\mathcal{O}$. Since $\mathcal{O}$ is simply connected, Poincaré's Lemma shows that Hdt is an exact form and hence its integrals along the closed curves c and $\mathcal{R}c$ are null. Now we have $\int_{\mathcal{R}c} \boldsymbol{p}d\boldsymbol{q} = \int_c \boldsymbol{p}d\boldsymbol{q}$ for any closed curve c in $\mathcal{O}$ and Exercise 58.6 implies that $\mathcal{R}$ is exact symplectic.

41. Suspension of Symplectic Twist Maps by Hamiltonian Flows

Moser (1986a) showed how to suspend a twist map of the annulus into a time 1 map of a (time dependent) Hamiltonian system satisfying the fiber convexity $H_{pp} > 0$. In subsection A we present a suspension theorem for higher dimensional symplectic twist maps announced in Bialy & Polterovitch (1992b), 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 41.1 (Bialy and Polterovitch) *Let F be a symplectic twist map with generating function S satisfying:*

$$\partial_{12} S(\boldsymbol{q}, \boldsymbol{Q}) \quad \textit{is symmetric and negative nondegenerate.} \tag{41.1}$$

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:

(41.2) (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 $\boldsymbol{q} + t(\boldsymbol{Q} - \boldsymbol{q})$;

(41.2) (b) $S(\boldsymbol{q}, \boldsymbol{Q}) = \displaystyle\int_0^1 L(\boldsymbol{q} + t(\boldsymbol{Q} - \boldsymbol{q}), \boldsymbol{Q} - \boldsymbol{q}, t)dt$;

(41.2) (c) L is strictly convex with respect to $\boldsymbol{v}$: $\frac{\partial^2 L}{\partial \boldsymbol{v}^2}$ is positive definite.

(41.2) (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 41.1: (41.2) (a) and (b) imply that F is the time 1 map of the Hamiltonian H, (41.2) (c) implies that $H_{\boldsymbol{pp}}$ is convex (see Exercise 59.2) and (41.2) (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 (41.2) (c) is satisfied then (41.2) (a) is equivalent to the following equation:

(41.2) (a′) $$\frac{\partial^2 L}{\partial \boldsymbol{v} \partial \boldsymbol{q}} \boldsymbol{v} + \frac{\partial^2 L}{\partial \boldsymbol{v} \partial t} - \frac{\partial L}{\partial \boldsymbol{q}} = 0.$$

Lemma 41.2 *Set* $R_{ij}(\boldsymbol{q}, \boldsymbol{v}, t) = -\dfrac{\partial^2 S}{\partial q_i \partial Q_j}(\boldsymbol{q} - t\boldsymbol{v}, \boldsymbol{q} + (1 - t)\boldsymbol{v})$. *Then the following holds:*

(41.3) (a) $R_{ij} = R_{ji}$;

(41.3) (b) $\dfrac{\partial R_{ij}}{\partial v_k} = \dfrac{\partial R_{ik}}{\partial v_j}$;

(41.3) (c) $\dfrac{\partial R_{ij}}{\partial q_k} = \dfrac{\partial R_{ik}}{\partial q_j}$;

(41.3) (d) $\dfrac{\partial R_{ij}}{\partial t} + \sum_l \dfrac{\partial R_{lj}}{\partial q_i} v_l = 0$

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 41.3 *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:*

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

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

(41.4) (c) *L satisfies Equation (41.2) (a').*

Proof. Rewrite L as follows:

(41.5)

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

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 (41.5) 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 $\gamma(s) = (\boldsymbol{q}, s\boldsymbol{v}, t)$. Fixing $\boldsymbol{q}$ and t, Equation (41.3) (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 (41.4) (a) follows. The proof of (41.4) (b) is similar. We now prove (41.4) (c). In view of (41.4) (a), the left hand side I of (41.2) (a') can be written as follows:

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

where a_k is the k^{th} integral in the above expression. Rewrite a_3 using (41.3) (c) as follows:

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

The first term is equal to a_1. Therefore:

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

Equation (41.3) (d) 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 41.4 *Assume that L satisfies (41.2) (a'). Then the following holds:*

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

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

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

Proof. Equation (41.6) (c) is a consequence of (41.6) (a), which we now prove. The same argument also proves (41.6) (b). If L satisfies (41.2) (a') or equivalently (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,

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

$$\int_0^1 \left\{ -\frac{\partial L}{\partial v_i}(\boldsymbol{q} + t(\boldsymbol{Q} - \boldsymbol{q}), \boldsymbol{Q} - \boldsymbol{q}, t) + (1-t)\frac{d}{dt}\left(\frac{\partial L}{\partial v_i}(\boldsymbol{q} + t(\boldsymbol{Q} - \boldsymbol{q}), \boldsymbol{Q} - \boldsymbol{q}, t)\right)\right\} dt$$

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

□

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

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

Lemma 41.5

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

(41.7) (b) *If L satisfies* (41.2) (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 41.1. Let L be the function defined in Lemma 41.3. From (41.6) (c) and (41.4) (b), we get:

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

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 (41.2) (a)-(d). We prove these properties one by one.
1. We proved in (41.4) (c) that L satisfies (41.2) (a'), and hence (41.2) (a). Statement (41.7) (b) proves that L_f does as well.
2. From (41.7) (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 (41.2) (b).

3. $\dfrac{\partial^2 L_f}{\partial \boldsymbol{v}^2} = \dfrac{\partial^2 L}{\partial \boldsymbol{v}^2} = (R_{ij}) = -\dfrac{\partial^2 S}{\partial \boldsymbol{q}\partial \boldsymbol{Q}}(\boldsymbol{q} - t\boldsymbol{v}, \boldsymbol{q} + (1-t)\boldsymbol{v})$. Since this last matrix is positive definite by Hypothesis (41.1), 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 $\frac{\partial f}{\partial \boldsymbol{q}}$ are also periodic in $\boldsymbol{q}$. Using the definitions and (41.6) (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 $\frac{\partial f}{\partial \boldsymbol{q}}$ are periodic. This finishes the proof of our claim, and hence that of Theorem 41.1. □

B. Suspension Without Convexity

If we let go of the symmetry of $\frac{\partial^2 S}{\partial \boldsymbol{q}\partial \boldsymbol{Q}}$ (but keep some form of definiteness) in Theorem 41.1 , 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 41.1, first appeared in Golé (1994c). I am indebted to F. Tangermann for a useful discussion about this theorem.

Theorem 41.6 *Let $F(\boldsymbol{q}, \boldsymbol{p}) = (\boldsymbol{Q}, \boldsymbol{P})$ be a symplectic twist map of $T^*\mathbb{T}^n$ whose differential $\boldsymbol{b}(\boldsymbol{z}) = \frac{\partial \boldsymbol{Q}(\boldsymbol{z})}{\partial \boldsymbol{p}}$ satisfies:*

$$\text{(41.8)} \qquad \inf_{\boldsymbol{z}\in T^*\mathbb{T}^n} \langle \boldsymbol{b}^{-1}(\boldsymbol{z})\boldsymbol{v}, \boldsymbol{v}\rangle > a\,\|\boldsymbol{v}\|, \quad a > 0, \ \forall \boldsymbol{v} \neq 0 \in \mathbb{R}^n.$$

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

Remark 41.7 Condition (41.8) tells us that F does not twist infinitely much. Note that (41.8) holds when H satisfies Assumptions 1 and 2 (see Remark 39.10).

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 (41.8) translates to:

$$\inf_{(\boldsymbol{q},\boldsymbol{Q})\in\mathbb{R}^{2n}} \langle -\partial_{12} S(\boldsymbol{q}, \boldsymbol{Q})\boldsymbol{v}, \boldsymbol{v}\rangle > a\,\|\boldsymbol{v}\|\,, \quad a > 0, \forall \boldsymbol{v} \neq 0 \in \mathbb{R}^n. \tag{41.9}$$

The following lemma show that (41.9) implies the hypothesis of Proposition 25.2, which in turn shows that whenever we have a function on $\mathbb{R}^{2n}$ which is suitably periodic and satisfies (41.9) , it is the generating function for some symplectic twist map.

Lemma 41.7 *Let $\{A_x\}_{x\in\Lambda}$ be a family of $n \times n$ real matrices satisfying:*

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

Then :

$$\det A_x \neq 0 \quad \text{and} \quad \sup_{x\in\Lambda} \left\|A_x^{-1}\right\| < a^{-1}.$$

We postpone the proof of this lemma. 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

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

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:

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

Hence S_t generates a smooth family $F_t,\ t \in (0,1]$ of symplectic twist maps. In fact $F_t(\boldsymbol{q}, \boldsymbol{p}) = (\boldsymbol{q} - (af(t))^{-1}\boldsymbol{p}, \boldsymbol{p}), \quad t \leq 1/2$, so that $\lim_{t\to 0^+} F_t = Id$. Define:

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

where ψ_{F_t} is the change of coordinates given by the fact that F_t is twist. Since $\psi_{F_t}(\boldsymbol{q}, \boldsymbol{p}) = (\boldsymbol{q}, \boldsymbol{q} - (af(t))^{-1}\boldsymbol{p}), \quad t \leq 1/2,$

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

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 59.7, F_t is a Hamiltonian isotopy. □

Proof of Lemma 41.7.. That $\det A_x \neq 0$ is obvious from the assumption: A has no kernel. For all non zero $\boldsymbol{v} \in \mathbb{R}^n$, we have:

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

Also:

$$\inf_{\|\boldsymbol{v}\|=1} \|A_x \boldsymbol{v}\| \geq \inf_{\|\boldsymbol{v}\|=1} |\langle A_x \boldsymbol{v}, \boldsymbol{v}\rangle| = \inf_{\boldsymbol{v} \in \mathbb{R}^n - \{0\}} \frac{|\langle A_x \boldsymbol{v}, \boldsymbol{v}\rangle|}{\|\boldsymbol{v}\|^2}$$

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:

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

□