Calculus of variations in the convex case :
an introduction to Fathi's weak KAM theorem
and Mather's theory of minimal invariant measures

Alain Chenciner, Barcelona july 2004

1st lecture. Calculus of variations in the convex case (local structures).

From Euler-Lagrange equations to the Poincaré-Cartan integral invariant, the Legendre transform and Hamilton's equations.

Exercices. Flows, differential forms, symplectic structures

2nd lecture. The Hamilton-Jacobi equation.

The solutions of Hamilton's equations as characteristics. Lagrangian submanifolds and geometric solutions of the Hamilton-Jacobi equation. Caustics as an obstruction to the existence of global solutions to the Cauchy problem.

Exercices. The geodesic flow on a torus of revolution as an example of a completely integrable system

3rd lecture. Minimizers.

Weierstrass theory of minimizers. Minimizing KAM tori, Existence of minimizers (Tonelli's theorem) and the Lax-Oleinik semi-group.

Exercices. Examples around the pendulum

4th lecture. Global solutions of the Hamilton-Jacobi equation

Weak KAM solutions as fixed points of the Lax-Oleinik semi-group; convergence of the semi-group in the autonomous case. Conjugate weak KAM solutions.

Exercices. Burger's equation and viscosity solutions.

5th lecture. Mather's theory. Class A geodesics and minimizing measures. The α and β functions as a kind of integrable skeleton
Exercices. The time-periodic case as a generalization of Aubry-Mather theory, Birkhoff billiards, Hedlund's example in higher dimension.

1st lecture. Calculus of variations in the convex case (local structures).

General convexity hypotheses. $M = \boldsymbol{T}^n = I\!R^n/Z\!\!Z^n$ is the n-dimensional torus (the theory works with an arbitrary compact manifold but the torus will allow us to work with global coordinates). The C^∞ (C^3 would be enough) Lagrangian $L(q, \dot{q}, t)$

$$L : T\boldsymbol{T}^n \times I\!R = \boldsymbol{T}^n \times I\!R^n \times I\!R \to I\!R$$

will be assumed to satisfy the "Mather" hypotheses (the third one will be explained later : it is only in case L depends effectively on the time variable t that it is not automatically satisfied) :

1) L is strictly convex in $\dot{q}$, that is (in the sense of quadratic forms) :

$$\forall q, \dot{q}, t, \ \frac{\partial^2 L}{\partial \dot{q}^2}(q, \dot{q}, t) > 0;$$

2) L is superlinear in $\dot{q}$:

$$\forall C \in I\!R, \exists D \in I\!R, \ \forall q, \dot{q}, t, \ L(q, \dot{q}, t) \geq C||\dot{q}|| - D,$$

that is $\lim_{||\dot{q}|| \to \infty} \frac{L(q,\dot{q},t)}{||\dot{q}||} = +\infty$ *uniformly in* (q, t).

3) the Euler-Lagrange flow associated to L is complete.

Path. A C^0 and piecewise C^1 mapping $\gamma : [a, b] \to \boldsymbol{T}^n$. When only minima of the action are concerned, it is more natural to work with *absolutely continuous paths.*

Action. To a path γ, one associates its action

$$\mathcal{A}_L(\gamma) = \int_a^b L(\gamma(t), \dot{\gamma}(t), t)\, dt.$$

Variation. A variation of γ is a mapping from $]-\epsilon, \epsilon[\times[a, b]$ to $\boldsymbol{T}^n$,

$$(u, t) \mapsto \Gamma(u, t) = \gamma_u(t), \quad \text{such that}$$

1) $\gamma_0 = \gamma$;
2) $\forall u$, γ_u is a path;
3) $\partial\Gamma/\partial u$ is continuous;
4) there exists a subdivision of $[a, b]$ into subintervals $[\tau_i, \tau_{i+1}]$ such that $\partial^2\Gamma/\partial u\partial t$ and $\partial^2\Gamma/\partial t\partial u$ are continuous (and hence equal) on the rectangles $]-\epsilon, \epsilon[\times[\tau_i, \tau_{i+1}]$.

Infinitesimal variation. It is the vector-field on $\boldsymbol{T}^n$ along γ defined by

$$X(t) = \frac{\partial\Gamma}{\partial u}(0, t).$$

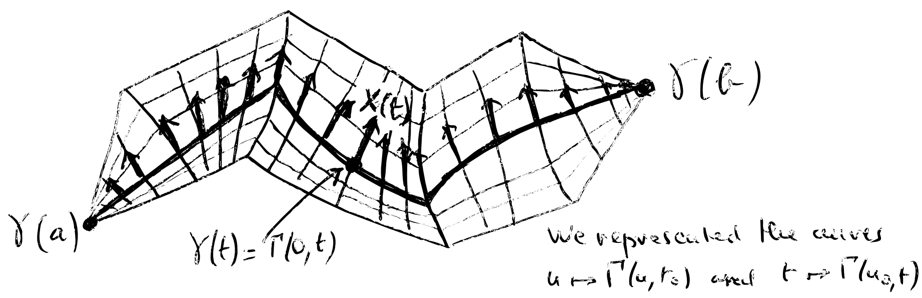

Figure 1

It is C^0 and piecewise C^1, and vanishes at a and b. The set of all these infinitesimal variations plays the rôle of the tangent space to the "manifold of paths".
Computing the derivative of the function $u \mapsto \mathcal{A}_L(\gamma_u)$ via an integration by parts, one gets

$$d\mathcal{A}_L(\gamma)X = \int_a^b \left[\frac{\partial L}{\partial \dot{q}}\big(\gamma(t), \dot{\gamma}(t), t\big) - \int_a^t \frac{\partial L}{\partial q}\big(\gamma(s), \dot{\gamma}(s), s\big) ds \right] \cdot \dot{X}(t)\, dt.$$

The following lemma is classical :

Lemma (Erdmann). *Let $\varphi : [a,b] \to R$ be continuous except possibly at a finite set of points. If $\int_a^b \varphi(t)\dot{\psi}(t)dt = 0$ for every C^0 and piecewise C^1 function $\psi : [a,b] \to \mathbb{R}$, which vanishes at a and b, the function φ coincides with the constant $\frac{1}{b-a}\int_a^b \varphi(t)dt$ at each point of continuity.*

Extremals. The paths γ such that $d\mathcal{A}_L(\gamma)X = 0$ for any infinitesimal variation X.

Euler-Lagrange equations (integral form). One deduces from the Erdmann lemma that a path γ is an extremal iff there exist constants C_i , $i = 1, 2 \cdots n$, such that, for $i = 1, 2 \cdots n$,

$$\frac{\partial L}{\partial \dot{q}_i}\big(\gamma(t), \dot{\gamma}(t), t\big) = \int_a^t \frac{\partial L}{\partial q_i}\big(\gamma(s), \dot{\gamma}(s), s\big) ds + C_i\ . \qquad (E')$$

Legendre mapping. The "general hypotheses" we made on L imply that the *Legendre mapping*

$$\Lambda : T\boldsymbol{T}^n \times \mathbb{R} = \boldsymbol{T}^n \times \mathbb{R}^n \times \mathbb{R} \to (\mathbb{R}^n)^* \times \boldsymbol{T}^n \times \mathbb{R} = T^*\boldsymbol{T}^n \times \mathbb{R}$$

defined by

$$\Lambda(q, \dot{q}, t) = (p, q, t), \quad p = \frac{\partial L}{\partial \dot{q}}(q, \dot{q}, t),$$

is a global diffeomorphism (strict convexity for all p of $\dot{q} \mapsto L(q, \dot{q}, t) - p \cdot \dot{q}$ implies the injectivity of Λ and surlinearity implies that it is proper, hence surjective). One says that L is *globally regular*. Using equations (E'), this implies immediately the

Regularity lemma. *Any extremal is as regular as L,*

and the following form of equations (E') :

Euler-Lagrange equations (differential form).

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot q_i}\big(\gamma(t),\dot\gamma(t),t\big)\right)=\frac{\partial L}{\partial q_i}\big(\gamma(t),\dot\gamma(t),t\big)\,,\quad i=1,\cdots,n. \qquad (E)$$

This amounts to computing $d\mathcal{A}_L(\gamma)\cdot X$ by the "other" integration by parts, which is permitted because γ is regular.

Intrinsic character of equations (E) : the Euler-Lagrange flow. It follows from the fact that Λ is a diffeomorphism that these equations define (time-dependant if L is) vector-fields X_L in $T\boldsymbol{T}^n$ and X_H^* in $T^*\boldsymbol{T}^n$ (the notation X_H^* will be explained below). These vector-fields are intrinsically defined (i.e. they do not depend on the choice of local or global coordinates on $\boldsymbol{T}^n$). Their flows will both be called the *Euler-Lagrange flow*.

Indeed, their variational origin implies that the Euler-Lagrange equations (E) take exactly the same form in any local or global coordinate system. In other words, the mapping $[L]_\gamma:[a,b]\to T^*\boldsymbol{T}^n$ defined by

$$[L]_\gamma(t)=\frac{\partial L}{\partial q}\big(\gamma(t),\dot\gamma(t),t\big)-\frac{d}{dt}\left(\frac{\partial L}{\partial \dot q}\big(\gamma(t),\dot\gamma(t),t\big)\right)\in T^*_{\gamma(t)}\boldsymbol{T}^n$$

is an intrinsically defined field of covectors tangent to $\boldsymbol{T}^n$ "along γ" and the derivative of the action can be written

$$d\mathcal{A}_L(\gamma)\cdot X=\int_a^b[L]_\gamma(t)\cdot X(t)\,dt.$$

Unconstrained variations and the Poincaré-Cartan integral invariant. The main structures of classical mechanics can be deduced from a single computation : the variations of the action when no constraints are imposed on the extremities of the paths γ_u or on their intervals of definition $[a(u),b(u)]$. We note $X_u(t)=\frac{\partial q}{\partial u}=\frac{\partial\Gamma}{\partial u}(u,t)$ the infinitesimal variations.

$$\begin{aligned}
&\frac{d}{du}\Big(\mathcal{A}_L\big(\gamma_u\big)\Big)=\frac{d}{du}\int_{a(u)}^{b(u)}L\left(\gamma_u(t),\dot\gamma_u(t),t\right)\,dt\\
&=\int_{a(u)}^{b(u)}\left[\left(\frac{\partial L}{\partial q}-\frac{d}{dt}\frac{\partial L}{\partial \dot q}\right)\big(\gamma_u(t),\dot\gamma_u(t),t\big)\right]\cdot X_u(t)\,dt\\
&+\frac{\partial L}{\partial \dot q}\big(\gamma_u(t),\dot\gamma_u(t),t\big)\cdot X_u(t)\,\Big|_{t=b(u)}-\frac{\partial L}{\partial \dot q}\big(\gamma_u(t),\dot\gamma_u(t),t\big)\cdot X_u(t)\,\Big|_{t=a(u)}\\
&+L\big(\gamma_u(t),\dot\gamma_u(t),t\big)\frac{db}{du}(u)\,\Big|_{t=b(u)}-L\big(\gamma_u(t),\dot\gamma_u(t),t\big)\frac{da}{du}(u)\,\Big|_{t=a(u)}\,,
\end{aligned}$$

a formula that we shall abreviate in

$$\frac{d\mathcal{A}_L}{du}=\int_a^b\left(\frac{\partial L}{\partial q}-\frac{d}{dt}\frac{\partial L}{\partial \dot q}\right)\cdot\frac{\partial q}{\partial u}\,dt\;+\;\left[\frac{\partial L}{\partial \dot q}\cdot\frac{\partial q}{\partial u}+L\frac{dt}{du}\right]_a^b.$$

If, in particular, γ_u is a family of extremals of $\int L\,dt$, we get

$$\frac{d\mathcal{A}_L}{du} = \left[\frac{\partial L}{\partial \dot q}\cdot\frac{\partial q}{\partial u} + L\,\frac{dt}{du}\right]_a^b .$$

We replace now the partial derivative $\frac{\partial q}{\partial u}$ (that is $\frac{\partial \Gamma}{\partial u}$), deprived of geometric meaning, by the "effective variation"

$$\frac{d}{du}\Big(\Gamma\big(u,t(u)\big)\Big) = \frac{dq}{du} = \frac{\partial q}{\partial u} + \frac{\partial q}{\partial t}\frac{dt}{du} = \frac{\partial q}{\partial u} + \dot q\,\frac{dt}{du}, \quad t(u) = a(u) \text{ or } b(u),$$

of the extremities of the path γ_u as a fonction of u. (figure 2.1). This transforms the expression of $\frac{d}{du}\left(\mathcal{A}_L(\gamma_u)\right)$ for a family of extremals into an identity between *differential 1-forms* on the interval $\mathcal{U}$ of definition of the parameter u :

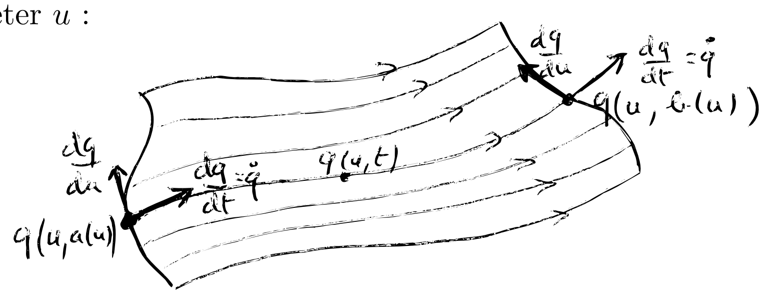

Figure 2

$$d\mathcal{A}_L = \delta_b^*\varpi_L - \delta_a^*\varpi_L,$$

where $\delta_a, \delta_b : \mathcal{U} \to T^*\boldsymbol{T}^n \times I\!R$ denote the mappings

$$\delta_t(u) = \left(\Gamma\big(u,t(u)\big), \frac{\partial \Gamma}{\partial t}\big(u,t(u)\big), t(u)\right), \quad t(u) = a(u) \text{ or } b(u),$$

and ϖ_L is the differential 1-form on $T\boldsymbol{T}^n \times I\!R$ defined by

$$\varpi_L = \frac{\partial L}{\partial \dot q}(q,\dot q,t)\cdot dq - \left(\frac{\partial L}{\partial \dot q}(q,\dot q,t)\cdot \dot q - L(q,\dot q,t)\right)dt.$$

Finally, we can simplify the formulas by transporting everything on the cotangent side with the Legendre diffeomorphism Λ. The function on $T^*\boldsymbol{T}^n \times I\!R$ defined by

$$H(p,q,t) = p\cdot \dot q - L(q,\dot q,t),$$

where $\dot q$ is expressed in terms of p,q,t via Λ is called the *Legendre transform* of L, or the *Hamiltonian* associated to the Lagrangian L. If ϖ_H denotes the 1-form on $T^*\boldsymbol{T}^n \times I\!R$ defined by

$$\varpi_H = p\cdot dq - H(p,q,t)dt,$$

the formula for the unconstrained variations of extremals becomes

$$d\mathcal{A}_L = (\Lambda\circ\delta_b)^*\varpi_H - (\Lambda\circ\delta_a)^*\varpi_H.$$

The 1-form ϖ_H is the *Poincaré-Cartan integral invariant* (*tenseur impulsion-énergie* in Cartan's terminology).

Rewriting the action. The action istself can now be written as the integral of $\varpi_H = p \cdot dq - Hdt$ on the *lift* $\Gamma^*(t) = \left(\frac{\partial L}{\partial \dot{q}}(\gamma(t), \dot{\gamma}(t), t), \gamma(t), t\right)$ to $T^*\boldsymbol{T}^n \times I\!R$ of the path $\gamma(t)$ in $\boldsymbol{T}^n$:

$$\mathcal{A}_L(\gamma) = \int_{\Gamma^*} \varpi_H.$$

This expression is the basis of *Hamilton's least action principle.*

From the integral invariant to the symplectic structure.

A paraphrase of equations (E) is that a path $t \mapsto \gamma(t)$ in $\boldsymbol{T}^n$ is an extremal if and only if the parametrized curve in $T^*\boldsymbol{T}^n \times I\!R$

$$t \mapsto \left(\frac{\partial L}{\partial \dot{q}}\big(\gamma(t), \dot{\gamma}(t), t\big), \gamma(t), t\right) = \Lambda\big(\gamma(t), \dot{\gamma}(t), t\big)$$

is an integral curve of the (time-dependant) vector-field

$$\Xi^*_H = (X^*_H, 1) = \Lambda_*(X_L, 1)$$

on $T^*\boldsymbol{T}^n \times I\!R$. The last formula of the preceding section then implies that, if C_a and C_b are two oriented *loops* in $T^*\boldsymbol{T}^n \times I\!R$, such that $C_b - C_a$ is the oriented boundary of a cylinder C generated by pieces of of integral curves of $\Xi^*_H = (X^*_H, 1)$, one has

$$\int_{C_a} p \cdot dq - H(p,q)dt = \int_{C_b} p \cdot dq - H(p,q)dt\,.$$

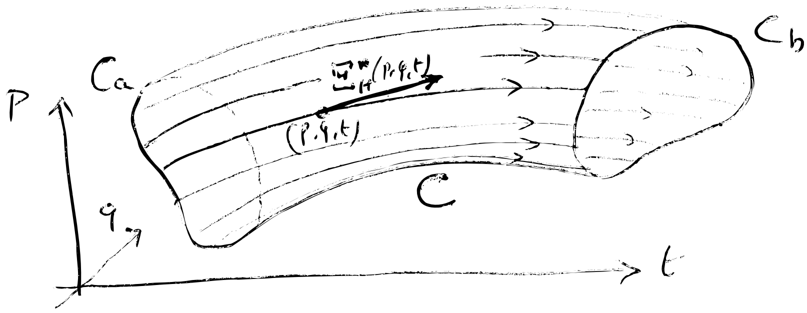

Figure 3

In Cartan's terminology, $\varpi_H = p \cdot dq - Hdt$ is a *relative* and *complete* integral invariant : relative because its invariance holds only if the integral is taken on loops C_i, complete because C_a and C_b are not supposed to be contained in slices where t is constant (i.e. C_b is not supposed to be the image of C_a under the element φ^b_a of the flow of Ξ^*_H.

Applying Stokes formula to small disks D_a et D_b contained respectively in the time slices $T^*\boldsymbol{T}^n \times \{a\}$ and $T^*\boldsymbol{T}^n \times \{b\}$ and such that $D_b = \varphi^b_a(D_a)$ is the image of D_a under the flow of Ξ^*_H, one gets the

Theorem. *The time-dependant vector-field X_H^* defined on $T^*\boldsymbol{T}^n$, preserves the* standard symplectic 2-form $\omega = \sum_{i=1}^n dp_i \wedge dq_i$.

A corollary of the preservation of the symplectic structure is

Liouville theorem. *The flow of the time-dependant vector-field X_H^* preserves the 2n-form ω^n, hence the Lebesgue measure (volume).*

The Hamilton equations. We deduce now the structure of the vector-field X_H^* (i.e. the structure of the Euler-Lagrange equations (E) seen from the cotangent side) from the following characterization of integral invariants :

(K) The 1-form ϖ_H *is an integral invariant of the vector-field Ξ_H^* if and only if, at each point $(p,q,t) \in T^*\boldsymbol{T}^n \times I\!R$, the vector $\Xi_H^*(p,q,t)$ belongs to the kernel of the bilinear form $d\varpi_H(p,q,t)$, i.e. if $i_{\Xi_H^*} d\varpi_H = 0$.*

The proof is a consequence of Stokes formula applied to oriented cylinders. This determines the direction of Ξ_H^*, hence X_H^*, because the kernel of

$$d(p\cdot dq - Hdt) = \sum_{i=1}^n \left[\left(dp_i + \frac{\partial H}{\partial q_i}dt\right) \wedge \left(dq_i - \frac{\partial H}{\partial p_i}dt\right)\right],$$

is easily seen to be 1-dimensional and generated at each point (p,q,t) by the vector $\left(-\frac{\partial H}{\partial q}(p,q,t), \frac{\partial H}{\partial p}(p,q,t), 1\right)$.

Finally, we get

$$X_H^* = \left(-\frac{\partial H}{\partial q_1}, \cdots - \frac{\partial H}{\partial q_n}, \frac{\partial H}{\partial p_1}, \cdots \frac{\partial H}{\partial p_n}\right).$$

Hence, when transported in $T^*\boldsymbol{T}^n$ by the Legendre diffeomorphism, the Euler-Lagrange equations (E) take the particularly symmetric form of the *Hamilton's equations* (or *canonical equations*) :

$$\begin{cases} \frac{dp_i}{dt} & = -\frac{\partial H}{\partial q_i}, \quad i = 1\cdots n, \\ \frac{dq_i}{dt} & = \frac{\partial H}{\partial p_i}, \quad i = 1\cdots n\,. \end{cases}$$

As the equations depend only on H, this justifies the notation X_H^*.

Symplectic changes of coordinates. If $\Phi(p,q) = (a,b)$ is symplectic, that is if $dp \wedge dq = da \wedge db$ (or more correctly $\Phi^*(\omega) = \omega$), the direct image of the Hamiltonian vector-field X_H^* is the Hamiltonian vector-field $X_{H\circ\Phi^{-1}}^*$.

The Legendre transform in the convex case. It follows from Hamilton's equations that the Legendre transform $L \mapsto H$ is involutive :

$$\begin{aligned} H(p,q,t) &= p\cdot \dot q - L(q,\dot q,t), \quad p = \frac{\partial L}{\partial \dot q}(q,\dot q,t), \\ L(q,\dot q,t) &= p\cdot \dot q - H(p,q,t), \quad \dot q = \frac{\partial H}{\partial p}(p,q,t). \end{aligned}$$

This symmetry makes it natural to write it in the following form, where the variables (q, t) play the role of mere parameters :

$$p \cdot \dot{q} = L(q, \dot{q}, t) + H(p, q, t).$$

The convexity of $\dot{q} \mapsto L(q, \dot{q}, t)$ is equivalent to that of $p \mapsto H(p, q, t)$ and if a function satisfies the general convexity hypotheses, so does its transform.

Young-Fenchel inequality. *For all $q, t, \dot{q}, p$, the following holds :*

$$p \cdot \dot{q} \leq L(q, \dot{q}, t) + H(p, q, t),$$

with equality if and only if $p = \dfrac{\partial L}{\partial \dot{q}}(q, \dot{q}, t)$.

Figure 4 illustrates in dimension 1 this variational definition of the Legendre transform. One also reads on this figure the interpretation of the transform as the passage from a punctual to a tangential equation.

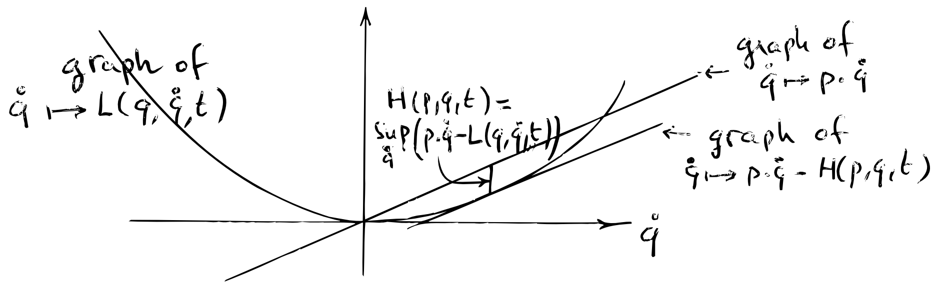

Figure 4

Autonomous Lagrangians. These are the Lagrangians $L(q, \dot{q})$ which do not depend explicitely on the time. It follows from Hamilton's equations that the (autonomous) vector-field X_H^* preserves the Hamiltonian H. In mechanics, this is the preservation of the *total energy.*

An elegant way of proving this property is to notice that the property (K), that is $i_{\Xi_H^*} d(p \cdot dq - Hdt) = 0$, is equivalent to

$$i_{X_H^*}\omega \ = \ -dH,$$

where ω is the symplectic form (in the time-dependant case, one must replace dH by $\partial H = dH - \frac{\partial H}{\partial t}$). Hence X_H^* is characterized by the property that, for any vector field Y on $T^*\boldsymbol{T}^n$, one has

$$\omega(X_H^*, Y) = -dH \cdot Y.$$

By analogy with the gradient of a Riemannian metric, one calls X_H^* the *symplectic gradient* of H. The conservation of energy amounts now to the identity

$$L_{X_H^*}H = dH(X_H^*) = -\omega(X_H^*, X_H^*) = 0.$$

An important feature of autonomous Hamiltonian systems is that *up to the parametrization, integral curves of the flow of X_H^* are completely determined by the sole geometry of the level hypersurfaces of H* : this is clear on figure 5 : the direction of $\mathrm{grad}_\omega H$ depends only on the direction of $\mathrm{grad}\, H$ and not on its length or orientation.

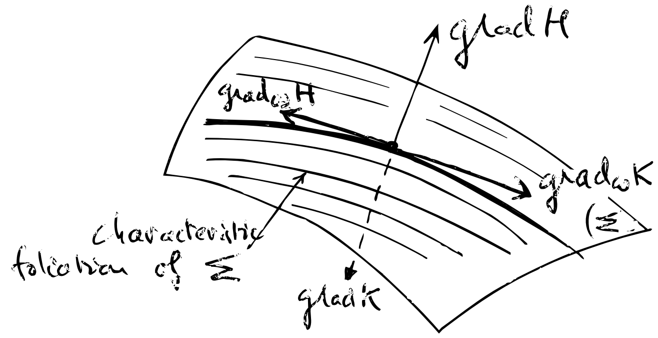

Figure 5 (H and K are regular equations of $H^{-1}(h) = K^{-1}(k)$ at x)

From time-dependant to time-independant. A time-dependant system can always be embedded into a time-independant one at the expense of adding dimensions and loosing track of time origin : indeed, the vector-field X_K^* on $T^*(\boldsymbol{T}^n \times I\!\!R)$ corresponding to the *extended* Hamiltonian

$$K(p, E, q, \tau) = E + H(p, q, \tau)$$

restricts to $\Xi_H^* = (X_H^*, 1)$ when one identifies the energy hypersurface $K^{-1}(0)$ with $T^*\boldsymbol{T}^n \times I\!\!R$. This extension may be useful even if H does not depend on time : because of the last component equal to 1, the geometry of the energy hypersurface $K \equiv E + H(p, q, t) = 0$ determines completely the vector-field Ξ_H^*, hence X_H^*.

The example of classical mechanics. The Lagrangian is the difference

$$L(q, \dot q) = \frac{1}{2}\, \dot q \cdot G(q)\dot q - V(q) = \frac{1}{2} g(q)(\dot q, \dot q) - V(q)$$

between *kinetic* and *potential* energy. The kinetic energy is defined by a Riemannian metric g on $M = \boldsymbol{T}^n$, that is for each q a positive definite quadratic form $g(q)$, represented by a symmetric matrix $G(q)$. When there is no potential V, the extremals are the *geodesics* of the metric. The Legendre transform $p = G(q)\dot q$ defines the conjugate momenta (the *impulsions*) p_i of the configuration variables q_i, the $\dfrac{\partial L}{\partial q_i}$ are the *forces* and the Hamiltonian is total energy, i.e. the sum of kinetic and potential energy

$$H(p, q) = \frac{1}{2}\, \dot q \cdot G(q)\dot q + V(q) = \frac{1}{2} p \cdot G(q)^{-1} p + V(q).$$

2nd lecture. The Hamilton-Jacobi equation.

A very simple completely integrable system : the geodesic flow of a flat torus. The Lagangian $L : T^*\boldsymbol{T}^2 = \mathbb{R}^2/(2\pi\mathbb{Z})^2 \times \mathbb{R}^2 \to \mathbb{R}$ is $L(q,\dot{q}) = \frac{1}{2}||\dot{q}||^2$. We shall write the coordinates $q = (\varphi, \psi)$ and $\dot{q} = (\dot{\varphi}, \dot{\psi})$ (figure 6).

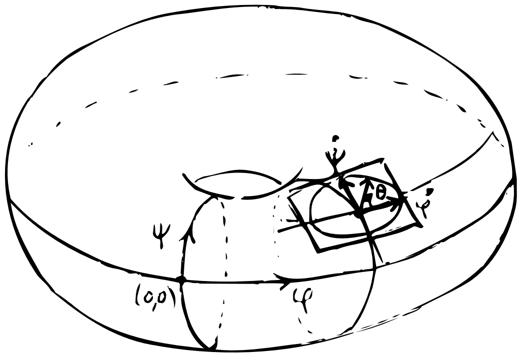

Figure 6

The Euler-Lagrange equation (E) is $\ddot{q} = 0$ and the extremals, the geodesics of $\boldsymbol{T}^2$, are the images by the canonical projection of the straight lines of $\mathbb{R}^2$ with an affine parametrization. The Legendre diffeomorphism is defined by $p = q$ and fixing the energy $H(p,q) = \frac{1}{2}||p||^2$ amounts to fixing the norm of the velocity. If the energy is different from 0, the energy hypersurface is diffeomorphic to $\boldsymbol{T}^3 = (\mathbb{R}/2\pi\mathbb{Z})^3$. The flow is depicted on figure 7.

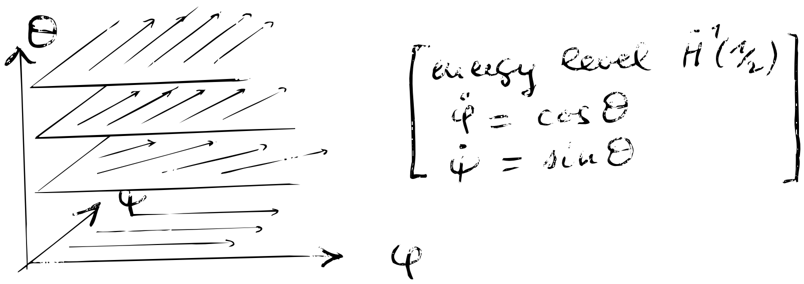

Figure 7

The whole *phase space* $T\boldsymbol{T}^2$ (or $T^*\boldsymbol{T}^2$) is foliated by the 2-dimensional tori $\dot{q} =$ constant (or $p =$ constant) which are invariant under the flow of X_L (or X_H^*). On these tori, the vector-field is constant (the flow is a flow of tranlations) and, depending on the rationality or irrationality of $\dot{\psi}/\dot{\varphi}$, the integral curves on the torus are all periodic or all dense.

Notice that the tori on which the integral curves are dense have a dynamical definition, as the closure of any of the integral curves they contain. This is not the case of the "periodic" tori which are a mere union of closed integral curves.

Opening of a resonance : the geodesic flow of a torus of revolution. We embed the 2-torus $\boldsymbol{T}^2 = \mathbb{R}^2/(2\pi\mathbb{Z})^2$ in $\mathbb{R}^3$ by the mapping ($r < 1$)

$$(\varphi, \psi) \mapsto \big((1 + r\cos\psi)\cos\varphi, (1 + r\cos\psi)\sin\varphi, r\sin\psi\big).$$

The image is invariant under rotation around the z axis. The extremals of the Lagrangian

$$L(\varphi, \psi, \dot\varphi, \dot\psi) = \frac{1}{2}\left((1 + r\cos\psi)^2\dot\varphi^2 + r^2\dot\psi^2\right)$$

are the geodesics of the induced metric, parametrized proportionnally to arc length.
The Euler-Lagrange equations are

$$\begin{cases} \frac{d}{dt}\big(1 + r\cos\psi)^2\dot\varphi\big) = 0\,, \\ \frac{d}{dt}\big(r^2\dot\psi\big) = -r\sin\psi(1 + r\cos\psi)\dot\varphi^2. \end{cases}$$

The first expresses the invariance under rotation around Oz and can be interpreted as the *conservation of angular momentum around Oz*. It is the analogue of the conservation of the angle θ in the flat case. Fixing the energy is fixing the velocity and the non-zero energy levels are diffeomorphic to the unit tangent bundle $T^1\boldsymbol{T}^2 \equiv \boldsymbol{T}^3$ with global angular coordinates (φ, ψ, θ) defined by choosing as third coordinate the *riemannian angle θ* :

$$\begin{cases} \dot\varphi &= \dfrac{\cos\theta}{1 + r\cos\psi}\,, \\ \dot\psi &= \dfrac{\sin\theta}{r}\,. \end{cases}$$

The first Euler-Lagrange equation becomes the constancy of the *Clairaut integral* :

$$(1 + r\cos\psi)\cos\theta = \text{ constant.}$$

Figures 9 represents the level curves of this function in the plane (ψ, θ). Figure 8 represents the level curves of the function θ, which plays for the flat torus the role of the Clairaut integral.

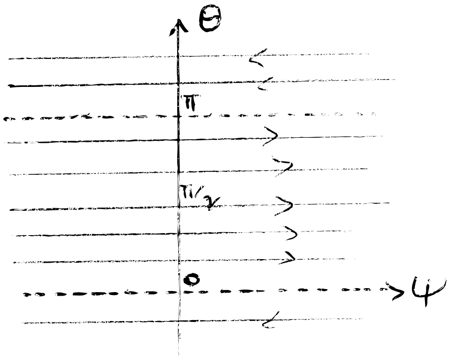

Figure 8

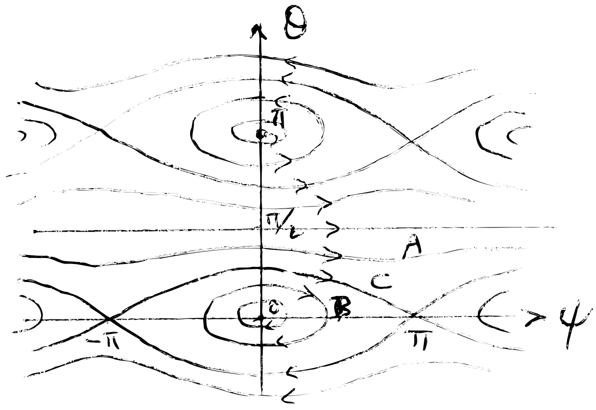

figure9

In the coordinates (φ, ψ, θ), the equations become

$$\begin{cases} \dfrac{d\varphi}{dt} & = \dfrac{\cos\theta}{1 + r\cos\psi}\,, \\ \dfrac{d\psi}{dt} & = \dfrac{\sin\theta}{r}\,, \\ \dfrac{d\theta}{dt} & = \dfrac{-\cos\theta\sin\psi}{1 + r\cos\psi}\,. \end{cases}$$

Because of the invariance under rotation, they are independant of φ, hence they admit a direct image in the torus (ψ, θ) which consists in forgetting the first equation. The same is of course true for the flat metric. The integral curves of this direct image are contained in the level curves of the Clairaut integral, which explains the arrows of figures 8 and 9.

In each open band $\theta \in \,]-\frac{\pi}{2} + k\pi, \frac{\pi}{2} + k\pi[\,$, $k \in \mathbb{Z}$, the flow looks qualitatively like the flow of a conservative pendulum. The rotations of the pendulum correspond to integral curves belonging to invariant tori which, as in the flat case, project biunivocally onto the configuration torus (φ, θ) (figure 10),

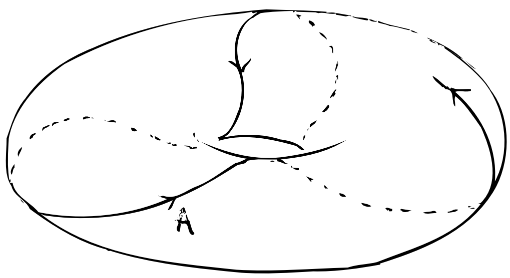

Figure 10

while oscillations correspond to integral curves belonging to invariant tori which project neither injectively nor surjectively but on an annulus whose boundary is a *caustic* (figure 11). These tori fill the *resonance zone.*

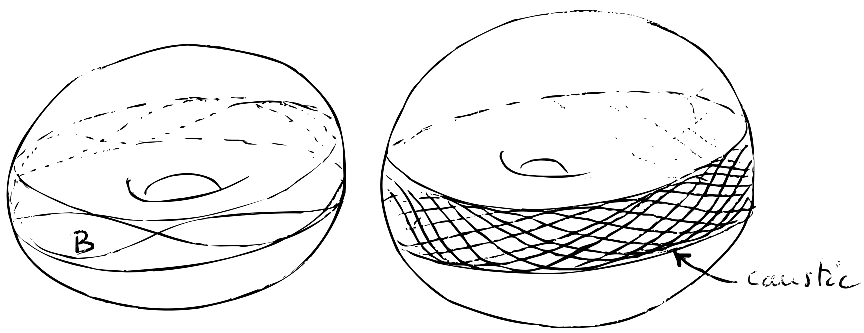

Figure 11

As in the flat case, in each of these invariant tori, integral curves are either all periodic or all dense.

A new feature is the existence in each non-zero energy level of 4 isolated periodic solutions, which correspond to the 2 geodesics defined by the intersection of the torus with the plane $z = 0$, each one with two possible directions of velocity. The inner one ($\psi = \pi, \theta = 0$ or $\psi = \pi, \theta = \pi$) is *hyperbolic* hence *unstable*. The set of integral curves with the same energy which are positively (negatively) asymptotic to it define the *stable* (*unstable*) manifold of this periodic orbit. These sets happen to coincide here. A corresponding geodesic is represented on figure 12. Their union is a surface which makes the transition between the two kinds of invariant tori oustside and inside the resonance zone.

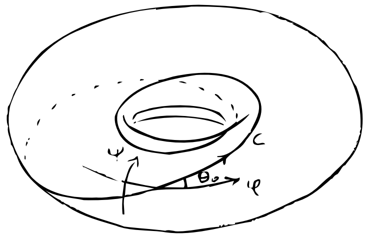

Figure 12

The outer one ($\psi = 0, \theta = 0$ or $\psi = 0, \theta = \pi$) is *elliptic*, hence stable. In its energy level, it is "surrounded" by invariant tori.

We have now two kinds of invariant sets dynamically defined : the invariant tori with dense integral curves and the stable = unstable manifolds of the hyperbolic periodic solutions.

Lagrangian submanifolds of $T^*\boldsymbol{T}^n$. In both exemples above, most of the *phase space* $T^*\boldsymbol{T}^2$ is foliated by invariant tori on which the flow of X_H^* can be shown to be a flow of translations in well chosen coordinates. This is obvious for the flat torus and a consequence of the invariance under rotation for the torus of revolution. The existence of such a foliation is a characteristic feature of the so-called *completely integrable* autonomous Hamiltonian systems. The following lemma shows that these tori are very special :

Lemma. *If the restriction of the flow of X_H^* to an invariant torus $\mathcal{T}$ is a flow of translations with dense orbits, $\mathcal{T}$ is isotropic, i.e. if j is its canonical injection in $T^*\boldsymbol{T}^n$, the pull-back $j^*\omega$ of the canonical symplectic form is identically zero.*

The proof is a consequence of the fact that $\omega = d\lambda$, where $\lambda = \sum_{i=1}^n p_i dq_i$ is the *Liouville form* on $T^*\boldsymbol{T}^n$. If $j^*\omega = \sum_{i<j} a_{ij}(u_1, \cdots, u_k) du_i \wedge du_j$ in coordinates $u_1, \cdots, u_k$ on $\mathcal{T}$ such that the flow of X_H^* becomes a flow of translations $\Phi_t(u) = u + tv$, the fact that $\Phi^*\omega = \omega$ implies that the functions a_{ij} are constant along the integral curves contained in $\mathcal{T}$ *(this would not be the case if $d\Phi_t(u)$ was not the Identity).* As these integral curves are dense, the a_{ij} are constant, hence equal to 0 because $j^*\omega = d(j^*\lambda)$ is a coboundary.

Notice that in the completely integrable cases that we studied above, an easy argument of continuity implies that all invariant tori (and not only the ones with dense integral curves), and also the stable = unstable invariant manifolds of the hyperbolic periodic solutions, share the property $j^*\omega = 0$. This property will play a fundamental role in the sequel :

Definition-Proposition. *A submanifold V of $T^*\boldsymbol{T}^n$ is called isotropic if $j^*\omega = 0$, where $j : V \to T^*\boldsymbol{T}^n$ is the canonical inclusion. The dimension of an isotropic submanifold is at most n. If it is equal to n, the submanifold is called Lagrangian.*

The bound on the dimension is an exercise in symplectic algebra : at each point (p,q), the bilinear form $\omega(p,q)$ is non degenerate, hence an isotropic subspace (i.e. a linear subspace contained in its orthogonal) is at most of dimension $2n/2 = n$.

Each invariant Lagrangian submanifold that we found in the integrable examples is contained in a single energy level. This is a consequence of the conservation of energy when the submanifold is the closure of a single solution and the others follow by continuity. This property has a very important converse :

Proposition. *let $H : T^*\boldsymbol{T}^n$ be an autonomous Hamiltonian. Every Lagrangian submanifold V of $T^*\boldsymbol{T}^n$ contained in a regular energy level $H^{-1}(h)$ is invariant under the flow of X_H^*.*

The proof is again an exercise in symplectic algebra : because of the maximality of the dimension of V among isotropic submanifolds, it is enough to show that at each point $m \in H^{-1}(h)$, the vector $X_H^*(m)$ belongs to (in fact generates) the kernel of $i_h^*\omega(m) = d(i_h^*\lambda)(m)$, where i_h is the canonical injection of $H^{-1}(h)$ in $T^*\boldsymbol{T}^n$ and $\lambda = p\cdot dq$ is the Liouville form. Indeed, if $X_H^*(m)$ was not contained in T_mV, the linear subspace genrated by $X_H^*(m)$ and T_mV would be isotropic of dimension $n+1$, a contradiction.
We have already proved this when $X_H^*|_{H^{-1}(h)}$ is replaced by Ξ_H^*, $H^{-1}(h)$ is replaced by $K^{-1}(0) \equiv T^*\boldsymbol{T}^n \times I\!R \subset T^*(\boldsymbol{T}^n \times I\!R)$, and the 1-form $i_h^*\lambda$ is replaced by the Poincaré-Cartan integral invariant $p\cdot dq - Hdt$. As this is the only case that we need, we leave the general assertion as an exercise.

Remark. A Hamiltonian flow is a very particular one as it preserves the symplectic 2-form ω, hence in particular the volume. *Its restriction to a Lagrangian submanifold V, on the contrary, does not satisfy any a priori constraint : every vector-field X on V is the restriction of a Hamiltonian flow defined on a neighborhood of V.* The simplest example is obtained when $V \equiv \boldsymbol{T}^n$ is the zero-section $p = 0$ of $T^*\boldsymbol{T}^n$: if $X(q)$ is vector-field on V, the Hamiltonian $H(p,q) = p\cdot X(q)$ is such that the restriction of X_H^* to V coincides with X (but it is not convex in p !).

Lagrangian graphs and the Hamilton-Jacobi equations. All invariant tori of the geodesic flow of a flat torus are graphs of a mapping $q \mapsto p(q)$, that is sections of the projection $(p, q,) \mapsto q$. For the torus of revolution, only those not contained in the resonance zone are graphs in the same way. The invariant manifolds of the hyperbolic periodic solutions are the union of two pieces, each of which is a graph.

Lemma. *If the Lagrangian submanifold V of $T^*\boldsymbol{T}^n = (\mathbb{R}^n)^* \times \boldsymbol{T}^n$ is a graph, it is the graph of a mapping of the form $p = a + ds(q)$, where $a = (a_1, \cdots, a_n) \in (\mathbb{R}^n)^*$ and $s : \boldsymbol{T}^n \to \mathbb{R}$.*

The proof is an easy calculation : the graph V of the mapping $q \mapsto p(q)$ is Lagrangian if and only if the 2-form $\sum_{i=1}^n dp(q) \wedge dq = \sum_{i,j} \frac{\partial p_i}{\partial q_j}(q) dq_j \wedge dq_i$ on $\boldsymbol{T}^n$ is identically 0. But this means that $\frac{\partial p_i}{\partial q_j}(q) = \frac{\partial p_j}{\partial q_i}(q)$ for all i, j. This implies that there exists a function $\sigma : \mathbb{R}^n \to \mathbb{R}$ such that for all i, $p_i(q) = \frac{\partial \sigma}{\partial q_i}(q)$. Hence there exist constants a_i (the *periods* of σ) and a function $s : \boldsymbol{T}^n \to \mathbb{R}$ such that for all i, $p_i(q) = a_i + \frac{\partial s}{\partial q_i}(q)$.

Corollary. *A Lagrangian graph V contained in the energy level $H^{-1}(h)$ of an autonomous Hamiltonian is of the form $\{(p, q), p = a + ds\}$, where s is a solution of the partial differential equation $H(a + ds(q), q) = h$.*

Definition. *The time-independant Hamilton-Jacobi equations associated to the Hamiltonian $H(p, q)$ are the equations of the form $H(ds(q), q) = h$. The time-dependant Hamilton-Jacobi equation associated to the Hamiltonian $H(p, q, t)$ is the equation $\frac{\partial S}{\partial t}(q, t) + H(\frac{\partial S}{\partial q}(q, t), q, t) = 0$. After identification of $K^{-1}(0)$ with $T^*\boldsymbol{T}^n \times \mathbb{R}$, it is nothing but the time-independant Hamilton-Jacobi equation $K(dS(q, t), q, t) = 0$, where K is defined by $K(p, E, q, \tau) = E + H(p, q, \tau)$.*

The modified Hamiltonian and Lagrangian. According to the above Corollary, each Lagrangian graph contained in an energy level of an autonomous Hamiltonian is the graph of the derivative of a solution of the Hamilton-Jacobi equation associated to a Hamiltonian

$$H_a(p, q) = H(a + p, q)$$

where $a \in (\mathbb{R}^n)^*$ should indeed be thought of as a cohomology class in $\mathcal{H}^1(\boldsymbol{T}^n, \mathbb{R})$. Such a Hamiltonian is easily seen to be the Legendre transform of the Lagrangian

$$L_a(q, \dot{q}) = L(q, \dot{q}) - a \cdot \dot{q} = L(q, \dot{q}) - \sum_{i=1}^n a_i \dot{q}_i,$$

which satisfies the same hypotheses as the original one.

Solving geometrically the Cauchy problem for the Hamilton-Jacobi equation. We shall be mostly interested in the time-dependant equation. The solution is contained in the figure 13 which explains how singularities (*caustics*) do occur which prevent the existence of a global solution as a function but allow for the existence of a "multiform" solution whose graph is a Lagrangian submanifold of $K^{-1}(0) \equiv T^*T^n \times I\!\!R \subset T^*(T^n \times I\!\!R)$. The *initial condition* at time t_0 being a function $u : \boldsymbol{T}^n \to I\!\!R$, the graph of this multiform solution is the union of the images of the graph of $du(q)$ under the flow of Ξ_H^*.

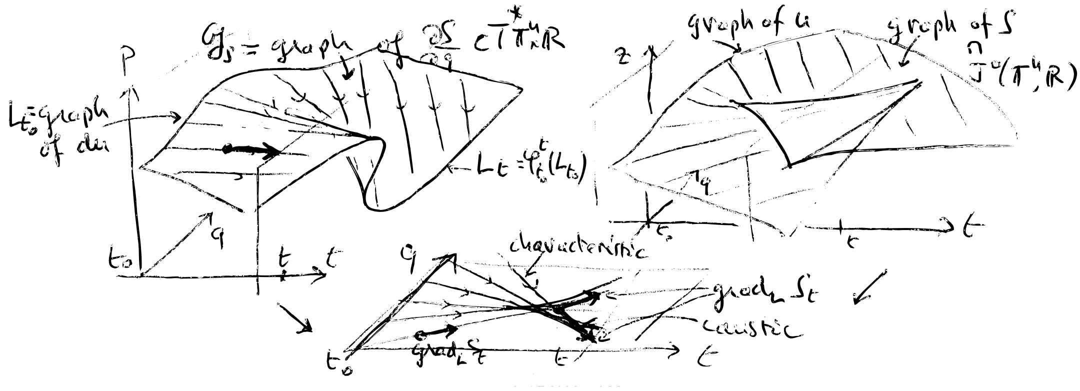

Figure 13

Figure 14 illustrates the resolution of the Cauchy problem for the time-independant Hamilton-Jacobi equation associated to the geodesic flow of the standard flat metric on $I\!\!R^2$ (or $\boldsymbol{T}^2$). The Cauchy data is chosen to be constant on a hypersurface $\mathcal{F}$. The *rays*, projections of the integral curves contained in the graph of the multiform solution, are the straight lines (geodesics) orthogonal to $\mathcal{F}$. They form what is classically called a *field of extremals*. The level hypersurfaces of the solution are the *wave fronts*. They are equidistant and everywhere orthogonal to the rays. This is the classical duality between waves and rays. In this case, the caustic is the envelope of the rays, that is the envelope of the normals to $\mathcal{F}$ (*evolute*).

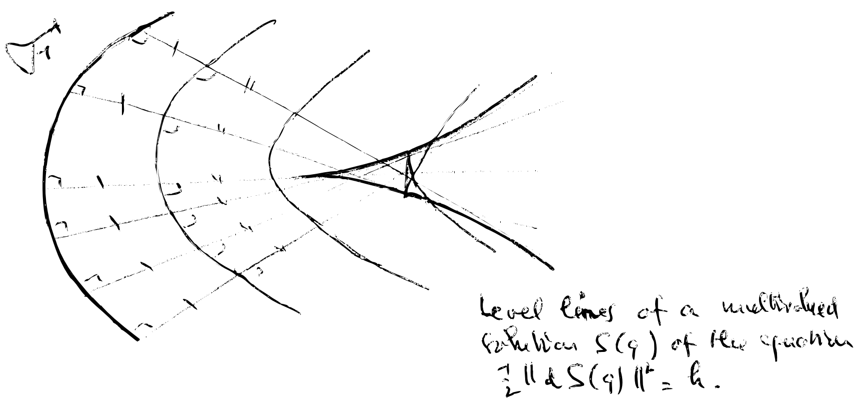

Figure 14

KAM and weak KAM. *Finally, finding invariant tori under X_H^* which are Lagrangian graphs in the energy level $H^{-1}(h)$ of an autonomous Hamiltonian is the same as finding GLOBAL solutions of the Hamilton-Jacobi equations associated to the Hamiltonians H_a.*

Such solutions do not exist in general but the K.A.M. theory asserts that a Cantor set of global solutions exists when H is a small C^k-perturbation of a completely integrable Hamiltonian (k not too small). The weak K.A.M. theorem of Fathi asserts that such solutions exist in a weak sense (a priori the funtion u is only Lipshitz and it is a *viscosity* solution of the equation in the sense of Lions, Papanicolaou, Varadhan, who had already proved the theorem in the case of the torus) as long as the convexity hypotheses on H (or L) are satisfied. These weak solutions define semi-invariant sets which contain as invariant subsets the so-called *Mather sets*, which generalize to any dimension the *Hedlund-Aubry-Mather* theory.

Characteristics. Coming back to the time-dependant Hamilton-Jacobi equation, i.e. to figure 13, we shall call *characteristics* associated to a geometrical (i.e. a priori *multivalued*) solution $S(q,t)$, the projections on space-time $\boldsymbol{T}^n \times I\!\!R$ of the integral curves of Ξ_H^* contained in the "graph" $\mathcal{G}_S \subset T^*\boldsymbol{T}^n \times I\!\!R$ of the space derivative $(q,t) \mapsto \frac{\partial S}{\partial q}(q,t)$. If we identify $T^*\boldsymbol{T}^n \times I\!\!R$ with $K^{-1}(0)$, where K is the extended Hamiltonian on $T^*(\boldsymbol{T}^n \times I\!\!R)$, $\mathcal{G}_S$ becomes a Lagrangian submanifold of $T^*(\boldsymbol{T}^n \times I\!\!R)$. The characteristics are the graphs of the solutions $t \mapsto q(t)$ of the *multivalued* differential equation

$$\frac{dq}{dt} = \frac{\partial H}{\partial p}\left(\frac{\partial S}{\partial q}(q,t), q, t\right) = \mathrm{grad}_L S_t(q). \tag{$\mathcal{C}$}$$

To explain the notation $\mathrm{grad}_L S_t$, recall that, by the Legendre diffeomorphism, the above equation is equivalent to $\frac{\partial S}{\partial q}(q,t) = \frac{\partial L}{\partial \dot q}(q, \frac{dq}{dt})$. Defining S_t by the formula $S_t(q) = S(q,t)$, we write it

$$dS_t(q) = \frac{\partial L}{\partial \dot q}(q, \frac{dq}{dt}). \tag{$\mathcal{C}'$}$$

Recalling that, for each (q,t), the mapping $\dot q \mapsto \frac{\partial L}{\partial \dot q}(q, \dot q, t)$ is a linear isomorphism from $I\!\!R^n$ to $(I\!\!R^n)^*$, it becomes natural to denote by $\mathrm{grad}_L S_t$ the right-hand side of the (multivalued) differential equation $(\mathcal{C})$: this notation is reminiscent of the transformation of the derivative of a function into its gradient via the linear isomorphism between tangent and cotangent vectors given by a Riemannian metric.

The multivaluedness of the time-dependant vector field $\mathrm{grad}_L S_t$ is a reflection of the fact that beyond the caustic, several characteristics pass through a given point (q,t) of space-time.

3rd lecture. Minimizers.

Weierstrass theory. We consider now a (*true, univalued*) solution $S(q,t)$, defined on a certain interval of time $[t_0,t_1]$, of the time-dependant Hamilton-Jacobi equation $\frac{\partial S}{\partial t}+H\left(\frac{\partial S}{\partial q},q,t\right)=0$. The vector-field $\mathrm{grad}_L S_t$ that we just defined is now univalued and the graphs of its integral curves, i.e. the characteristics associated to S, form a *field of extremals.*

Let $c:[t_0,t_1]\to \boldsymbol{T}^n$ be a segment of extremal whose lift to $T^*\boldsymbol{T}^n\times I\!R$, $C^*(t)=\left(\frac{\partial L}{\partial \dot q}\big(c(t),\dot c(t),t\big),c(t),t\right)$, is contained in the Lagrangian graph $\mathcal{G}_S$ defined by the equation $p=\frac{\partial S}{\partial q}(q,t)$

Proposition. *The segment of extremal c does minimize the action among absolutely continuous paths $\gamma:[t_0,t_1]\to\boldsymbol{T}^n$ with the same extremities, whose graph remains in the domain of definition of S.*

The idea of the proof is to lift to $\Gamma^*(t)=\left(\frac{\partial S}{\partial q}(\gamma(t),t),\gamma(t),t\right)\subset\mathcal{G}_S$ any curve $\gamma(t)$ that we want to compare to $c(t)$ (figure 15) and to use the fact that $\mathcal{G}_S$ is exact Lagrangian, to get

$$\mathcal{A}_L(c)=\int_{\Gamma^*}(p\cdot dq-H(p,q,t)dt).$$

Indeed, on $\mathcal{G}_S$, we have $p\cdot dq-H(p,q,t)dt=dS(q,t)$ because $p=\frac{\partial S}{\partial q}(q,t)$ and $H(p,q,t)=-\frac{\partial S}{\partial t}(q,t)$. This implies that

$$\mathcal{A}_L(c)=\int_{C^*}(p\cdot dq-H(p,q,t)dt)=\int_{t_0}^{t_1}d\big(S(q(t),t)\big)=\big[S\big(q(t),t\big)\big]_{t_0}^{t_1}$$

does not depend on the path on which one integrates as long as this path is contained in $\mathcal{G}_S$.

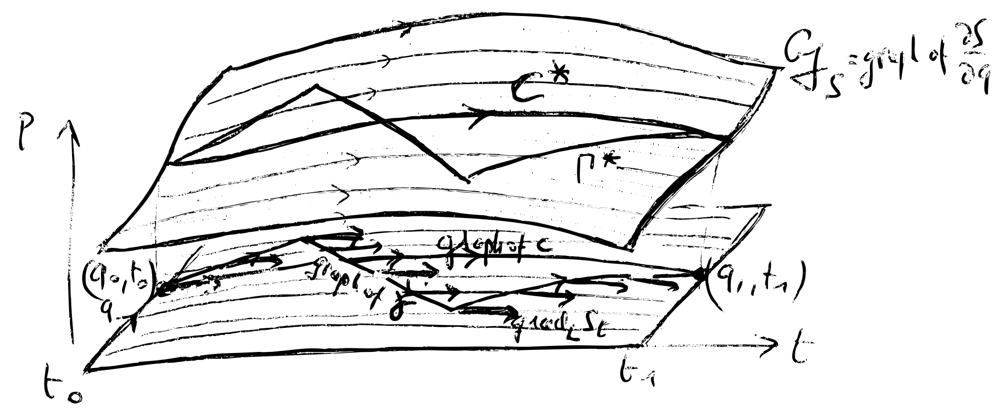

Figure 15 (in $T^*\boldsymbol{T}^n\times I\!R\equiv K^{-1}(0)$)

The Young-Fenchel inequality then implies that the difference of the actions

$$\mathcal{A}_L(c) - \mathcal{A}_L(\gamma) = \int_{t_0}^{t_1} \left[\pi(t)\cdot\dot\gamma(t) - H\big(\pi(t),\gamma(t),t\big) - L\big(\gamma(t),\dot\gamma(t),t\big)\right] dt,$$

where $\pi(t) = \frac{\partial S}{\partial q}\big(\gamma(t),t\big)$, is the integral of an everywhere ≤ 0 function.

Corollary. *The solution $S(q,t)$ of the time-dependant Hamilton-Jacobi equation defined on the (small enough) interval $[t_0,t]$ with initial condition $S(q,t_0) = u(q)$ is given by :*

$$S(q,t) = \min_{\gamma,\gamma(t)=q} \left[u(\gamma(t_0)) + \int_{t_0}^{t} L\big(\gamma(s),\dot\gamma(s),s\big)ds\right],$$

where the min *is taken over all absolutely continuous paths $\gamma : [t_0,t] \to \boldsymbol{T}^n$ such that $\gamma(t) = q$.*

The proof is the same as for the Proposition : c is replaced by the unique extremal whose graph is the characteristic associated to S such that $c(t) = q$ and γ by a path defined on the interval $[t_0,t]$ and such that $\gamma(t) = q$. Then, as above,

$$\mathcal{A}_L(c) = \int_{C^*} (p\cdot dq - Hdt) = \int_{\Delta^*} (p\cdot dq - Hdt),$$

where Δ^* is composed of the lift to $\mathcal{G}_S$ of a path in $\boldsymbol{T}^n \times \{t_0\}$ joining $c(t_0)$ to $\gamma(t_0)$, followed by the lift Γ^* of γ (figure 16). One concludes because the part of $\mathcal{G}_S$ above $t = t_0$ coincides with the graph of du.

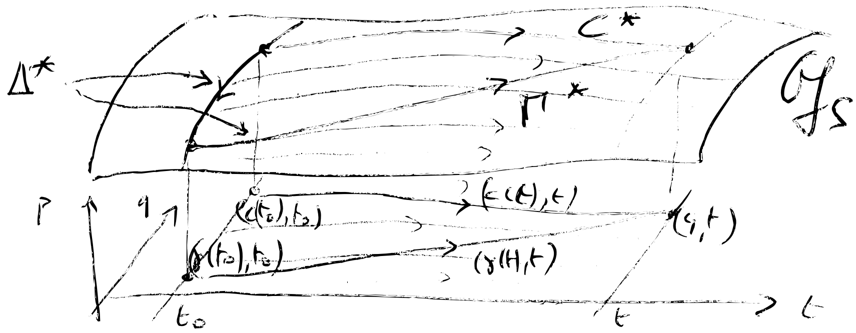

Figure 16

The Lax-Oleinik semi-group (autonomous case). The solution of the Cauchy problem defined above is in general only *local in time* : caustics appear as soon as the extremals in the corresponding field start intersecting each other. We now globalize it at the expense of regularity by taking in the global situation the *same* formula as in the local one : this amounts to cutting the *swallowtails* of the graph of the multiform function S. By keeping only for each (q,t) the lowest of the values of $S(q,t)$, one obtains the (discontinuous) graph of a Lipcshitz solution of the Hamilton-Jacobi equation (figure 17) .

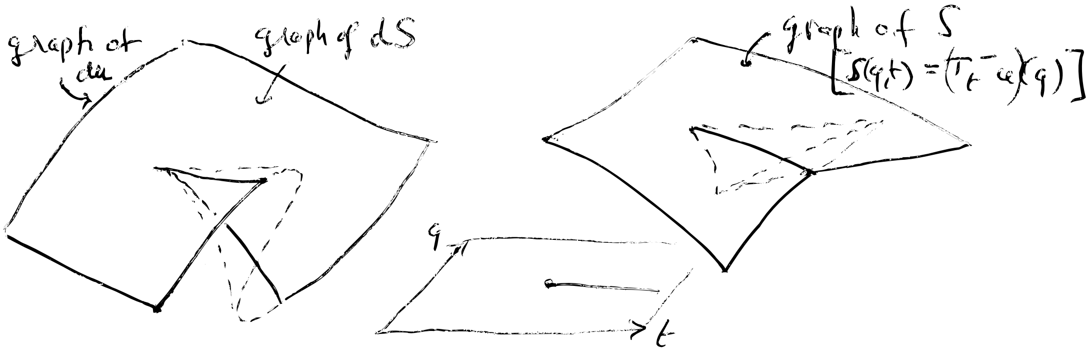

Figure 17

An astonishing feature of the result is that we get a global (weak) solution of the Cauchy problem *even when the initial condition $u(q)$ is only continuous.* If we approach u by C^1 functions u_n, the behaviour of the derivatives du_n may become wild as n tends to infinity but still the truncated global solutions corresponding to initial conditions u_n have a nice limit. The complete statement is the following :

Theorem (Existence of the Lax-Oleinik semi-group in the autonomous case). *1) The formula (for $t \geq 0$)*

$$\left(T_t^- u\right)(q) = \inf_\gamma \left[u\big(\gamma(0)\big) + \int_0^t L\big(\gamma(s), \dot\gamma(s)\big) ds \right],$$

where the inf *is taken over all absolutely continuous paths $\gamma : [0,t] \to \boldsymbol{T}^n$ such that $\gamma(t) = q$, defines a semi-group $\{T_t^-\}_{t\geq 0}$ of mappings from the space of continuous functions $C^0(\boldsymbol{T}^n, I\!R)$ to itself;*
2) For all q, t, there exists a minimizing extremal $\gamma : [0,t] \to \boldsymbol{T}^n$ such that $\gamma(t) = q$ and

$$(T_t^- u)(q) = u(\gamma(0)) + \int_0^t L\big(\gamma(s), \dot\gamma(s)\big) ds; \qquad (*)_{u,q,t}$$

3)$||T_t^- u - T_t^- v||_0 \leq ||u - v||_0$;
4) $T_t^-(u + c) = (T_t^- u) + c$;
5) At each point where $S(q,t) = (T_t^- u)(q)$ has a derivative, it is a true solution of the time-dependant Hamilton-Jacobi equation;
6) The same is true for the semi-group $\{T_t^+\}_{t\geq 0}$ defined by

$$(T_t^+ u)(q) = \sup_\gamma \left[u\big(\gamma(0)\big) - \int_{-t}^0 L\big(\gamma(s), \dot\gamma(s)\big) ds \right],$$

where the sup is taken over absolutely continuous paths $[-t, 0] \to \boldsymbol{T}^n$ such that $\gamma(-t) = q$.

The main tools in the proof are
1) the existence of minimizing extremals (*Tonelli's theorem*);
2) the regularity of minimizing extremals (*Weierstrass local theory*);
3) an easy but fundamental *lemma of a priori compactness*

Tonelli existence theorem. *Let M be any compact manifold and $L : TM \times I\!R \to I\!R$ a Lagrangian which is strictly convex and superlinear. For any two points $q', q'' \in M$ and any interval $[t', t'']$, there exists an absolutely continuous path $\gamma : M \to I\!R$ such that $\gamma(t') = q', \gamma(t'') = q''$ and for any absolutely continuous path $\gamma_1 : [t', t''] \to M$ with the same extremities, one has $\int_{t'}^{t''} L\big(\gamma_1(s), \dot\gamma_1(s), s\big) \geq \int_{t'}^{t''} L\big(\gamma(s), \dot\gamma(s), s\big)$.*

This classical is result is based on the following key fact : the topologies on the space of absolutely continuous paths which insure the continuity of the action functional $\mathcal{A}_L$, for instance the norm topology of the Sobolev space H^1, are too strong to allow for enough compact sets to handle minimizing sequences. For weaker topologies which have plenty of compact sets, as the C^0 one, $\mathcal{A}_L$ is only *lower semi-continuous* (it is well known that the length of a uniformly convergent sequence of curves can drop down at the limit) but this is enough to deal with the minimization problem. In fact, one can prove that the set of absolutely continuous paths $\gamma : [a, b] \to M$ with a bounded action is compact in the C^0 topology !

Weierstrass regularity theorem. *If moreover the Euler-Lagrange flow is complete, i.e. if L satisfies the general convexity hypotheses, the minimizing curves are minimizing extremals and hence as regular as the Lagrangian.*

The proof is done in three steps :
1) That small enough extremals are minimizing among absolutely continuous curves with the same end-points results from a local application of the Weierstrass theory (Proposition at the beginning of the section) : indeed, any sufficiently small piece of extremal is contained in a piece of exact Lagrangian graph.
2) One then shows that between points sufficiently close together, there exists a small extremal. More precisely, given any positive constant C, there exists a positive ϵ such that if $(q_t, \dot q_t) = \varphi_t(q, \dot q)$ is the image of $(q, \dot q)$ under the Euler-Lagrange flow in $T\boldsymbol{T}^n$ and $|t| < \epsilon$, the mapping $\dot q \mapsto q_t$ (initial velocity to final position) is a diffeomorphism of the ball $||\dot q|| \leq C$ onto a subset of $\boldsymbol{T}^n$ containing the ball of center q and radius $C|t|/2$ (the problem is local and we can use any metric in $I\!R^n$).
3) This last property implies immediately that a minimizing curve $\gamma(t)$ coincides with an extremal in the neighborhood of any t_0 at which it is differentiable. As an absolutely continuous function is differentiable almost everywhere, this proves *Tonelli's partial regularity theorem.* To conclude, one uses the completeness of the Euler-Lagrange flow. In the autonomous case, this completeness is a consequence of the compactness of the energy levels, itself a consequence of the superlinearity of the Hamiltonian.

Lemma of a priori compactness. *Let M be a compact manifold. If $t > 0$ is given, there exists a compact subset K_t of TM such that, for any mimiz-*

ing extremal $\gamma : [a,b] \to M$ with $b-a \geq t$ and any $s \in [a,b]$, $\big(\gamma(s), \dot\gamma(s)\big)$ belongs to K_t. In other words, there is an a priori bound on the velocities of a minimizing extremal, which depends only on the length of its interval of definition.

The proof of this lemma is an elementary estimation of the action of a small geodesic (in any Riemannian metric) joining two nearby points at constant velocity in time t.

Sketch of proof of the theorem. The theorems of Tonelli and Weierstrass imply the existence, for each $t > 0$, and each $q_0, q \in \boldsymbol{T}^n$, of a minimizing extremal $\gamma_{q_0} : [0,t] \to \boldsymbol{T}^n$ such that $\gamma_{q_0}(0) = q_0$, $\gamma_{q_0}(t) = q$. Then, for each $u \in C^0(\boldsymbol{T}^n, I\!R)$,

$$(T_t^- u)(q) = \inf_{q_0 \in T^n} \left\{ u(q_0) + \int_0^t L\big(\gamma_{q_0}(s), \dot\gamma_{q_0}(s)\big) ds \right\}.$$

But $(\gamma_{q_0}(s), \dot\gamma_{q_0}(s)) = \varphi_s(q_0, \dot\gamma_{q_0}(0))$ is the image of $(q_0, \dot\gamma_{q_0}(0))$ under the Euler-Lagrange flow φ_s and the lemma of a priori compactness implies that $(q_0, \dot\gamma_{q_0}(0)) \in K_t$. Hence we can find a sequence of initial points $q_j \in \boldsymbol{T}^n$ such that $(q_j, \dot\gamma_{q_j}(0))$ converges, which implies the convergence of the γ_{q_j} to a minimizing extremal γ verifying the equation $(*)_{u,q,t}$ of part 2) of the Theorem. Note that, by Weierstrass' theorem, all minimizing extremals are regular curves. Assertions 1), 3), 4) follow from the existence of γ.

An important (and classical) property of T_t^- is decribed in the following lemma, whose proof is typical of the techniques involved :

Lemma 1. *For all $t > 0$ and all $u \in C^0(\boldsymbol{T}^n, I\!R)$, the function $T_t^- u$ is Lipschitzian with Lipschitz constant independant of u (but dependant of t).*

The proof can be read on figure 18 :

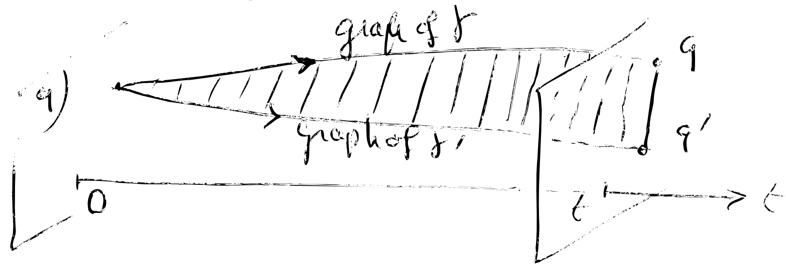

Figure 18

To compare $(T_t^- u)(q)$ and $(T_t^- u)(q')$, one uses a minimizing extremal γ ending in q at time t and verifying the equality in part 2) of the Theorem; one builds from it a curve $\gamma'(s) = \gamma(s) + \frac{s}{t}(q' - q)$ joining $\gamma(0)$ to q' (as everything is local, one can work in $I\!R^n$). It remains to use the minimizing property of T_t^- and the fact that $\gamma'(0) = \gamma(0)$ to get

$$(T_t^- u)(q') - (T_t^- u)(q) \leq \int_0^t L\big(\gamma'(s), \dot\gamma'(s)\big) ds - \int_0^t L\big(\gamma(s), \dot\gamma(s)\big) ds.$$

Using the explicit definition of γ' and the fact that the derivative of L is bounded on the compact K_t, one bounds the right-hand side by $k(t)||q-q'||$. To prove the lemma, it only remains to exchange the roles of q and q'.

The proof of 5) is based on a refinement (to next order) of the above lemma, whose proof is similar :

Lemma 2. *If $t_0 > 0$, there is a constant $K = K(t_0) > 0$ such that for any $u \in C^0(\boldsymbol{T}^n, I\!R)$ and any minimizing extremal γ satisfying $(*)_{u,q,t}$, one has*

$$(T_t^- u)(q') - (T_t^- u)(q) \leq \frac{\partial L}{\partial \dot{q}}\big(q, \dot{\gamma}(t)\big)(q' - q) + K||q' - q||^2.$$

Coroallary. *If $T_t^- u$ is differentiable at q and if γ satisfies $(*)_{u,q,t}$, one has*

$$d(T_t^- u)(q) = \frac{\partial L}{\partial \dot{q}}\big(q, \dot{\gamma}(t)\big). \qquad (\mathcal{C}')$$

In particular, γ is unique and its graph is a characteristic associated to the function $S(q,t) = (T_t^- u)(q)$.

Now 5) is a consequence of the following (intuitive) fact :

Differentiability along the characteristics. *By restricting to a smaller interval $[0, t']$, $t' < t$, a minimizing extremal $\gamma : [0, t] \to \boldsymbol{T}^n$ with $\gamma(t) = q$ which satisfies $(*)_{u,q,t}$, one gets a curve with the same property, i.e. satisfying $(*)_{u,\gamma(t'),t'}$. In other words, the restriction of a characteristic is still a characteristic !*

As a consequence, the function $s \mapsto (T_s^- u)(\gamma(s))$ is given by the formula $(*)_{u,\gamma(s),s}$ with the same γ for all s, hence it is differentiable on the whole interval $[0, t]$ and its derivative is $L\big(\gamma(s), \dot{\gamma}(s)\big)$.

Let us set $S(q,t) = (T_t^- u)(q)$. At a point (q,t) of differentiability of S, the above property becomes

$$\frac{d}{ds}\left(S(q,t)\right) = \frac{\partial S}{\partial q}(q,t) \cdot \dot{\gamma}(t) + \frac{\partial S}{\partial t}(q,t) = L\big(q, \dot{\gamma}(t)\big).$$

It implies that, at a point (q,t) of differentiability, S satisfies the equation $\frac{\partial S}{\partial t}(q,t) + H\left(\frac{\partial S}{\partial q}(q,t), q\right) = 0$ if and only if

$$-\frac{\partial S}{\partial q}(q,t) \cdot \dot{\gamma}(t) + L\big(q, \dot{\gamma}(t)\big) + H\left(\frac{\partial S}{\partial q}(q,t), q\right) = 0,$$

which is equivalent to $\frac{\partial S}{\partial q}(q,t) = \frac{\partial L}{\partial \dot{q}}(q, \dot{\gamma}(t))$ (equality in Young-Fenchel inequality). But this is the equation of characteristics $(\mathcal{C}')$ that we already know to be satisfied.

Finally, we quote without proof the following converse of the unicity of γ in the corollary of Lemma 2 :

Differentiability of $T_t^- u$ and the unicity of characteristics. *The point (q, t) is a point of differentiability of $S(q, t) = (T_t^- u)(q)$ if and only if it is the extremity of a unique characteristic.*

Figure 19 illustrates this :

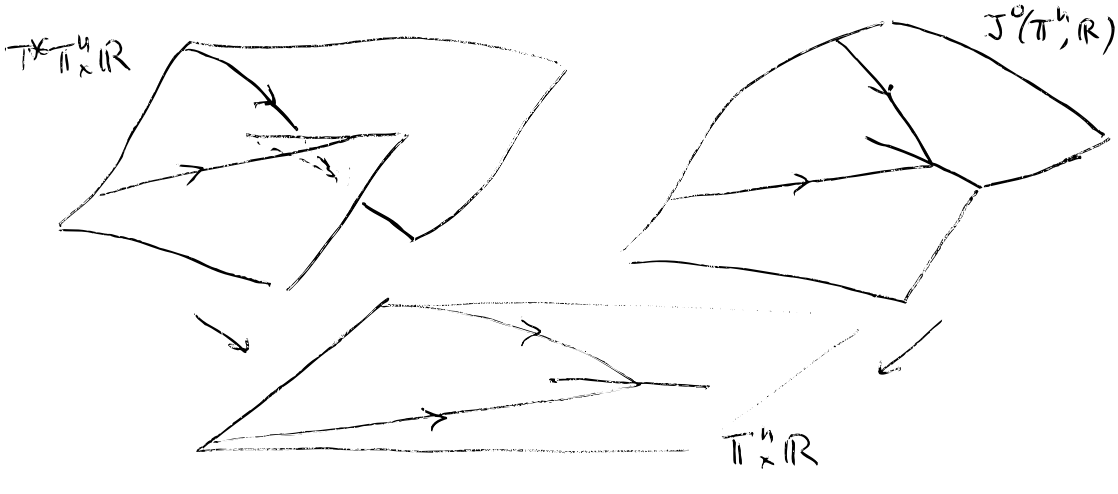

Figure 19

4th lecture. Fathi's weak KAM theorem (autonomous case). Following Fathi, we deduce from the existence of the Lax-Oleinik semi-group the existence of weak KAM solutions (in fact viscosity solutions) of the time-independant Hamilton-Jacobi equation $H(du(q),q) = c$ for a well-chosen c (equal to $\alpha(0)$ in Mather's notation).

First one notices that the following properties are equivalent to one another:

(1) $\exists c,\ H(du(q),q) = c$;

(2) $\exists c,\ \dfrac{\partial S}{\partial t} + H\big(\dfrac{\partial S}{\partial q}, q\big) = 0,$ where $S(q,t) = u(q) - ct$;

(3) u is a fixed point of the semi-group $u \mapsto T_t^- u + ct$;

(4) u represents a fixed point of $T_t^- u$ in $C^0(\boldsymbol{T}^n, I\!R)/I\!R$.

One then proves the existence of a fixed point by a Leray-Shauder type fixed point argument. To conveniently state the theorem, we introduce the following definitions :

Domination. *Given a real number c, we say that the function $u : \boldsymbol{T}^n \to I\!R$ is dominated by $L+c$ (and we write $u \prec L+c$) if for any interval $[a,b]$, $a \le b$ and any absolutely continuous curve $\gamma : [a,b] \to \boldsymbol{T}^n$ we have*

$$u\big(\gamma(b)\big) - u\big(\gamma(a)\big) \le \int_a^b L\big(\gamma(s), \dot\gamma(s)\big)ds + c(b-a).$$

Proposition. *$u \prec L + c$ if and only if u is locally Lipschitzian and $H(du(q),q) \le c$ at each point q where u has a derivative.*

Calibration. We suppose that $u \prec L + c$. The curve $\gamma : [a,b] \to \boldsymbol{T}^n$ is said to be (u, L, c)-calibrated if

$$u\big(\gamma(b)\big) - u\big(\gamma(a)\big) = \int_a^b L\big(\gamma(s), \dot\gamma(s)\big)ds + c(b-a).$$

Weak KAM theorem. *Let $L(q,\dot q)$ be a time-independant Lagrangian on $T\boldsymbol{T}^n$ of class at least C^3, which satisfies the general convexity hypotheses. There exist Lipschitz functions $u_-, u_+ : \boldsymbol{T}^n \to I\!R$ and a constant $c \in I\!R$ (the Mañé's critical energy) such that*
1) $u_-, u_+ \prec L + c$,
2) $\forall q \in \boldsymbol{T}^n$, there exists $\gamma_-^q :]-\infty, 0] \to \boldsymbol{T}^n$ and $\gamma_+^q :]0, +\infty] \to \boldsymbol{T}^n$ with $\gamma_-^q(0) = \gamma_+^q(0) = q$, such that, for all $t \ge 0$,

$$u_-(q) - u_-(\gamma_-^q(-t)) = ct + \int_{-t}^0 L\big(\gamma(s), \dot\gamma(s)\big)ds,$$

$$u_+(\gamma_+^q(t)) - u_+(q) = ct + \int_0^t L\big(\gamma(s), \dot\gamma(s)\big)ds,$$

3) $u_\pm$ satisfies $H(du_\pm(q), q) = c$ at each point q where it has a derivative.

Properties of weak KAM solutions. $u_\pm$ are not necessarily unique (doubly periodic pendulum).

The following property is the analogue of the one that we discussed for $T_t^- u$: the discontinuities of the derivative of a weak KAM solution are the intersection points of at least two rays (characteristics).

Proposition (differentiability and unicity of calibration). *A weak KAM solution u_- has a derivative at q if and only if there exists a UNIQUE (L, c, u_-)-calibrated path $\gamma_-^q :]-\infty, 0] \to \boldsymbol{T}^n$ such that $\gamma_-^q(0) = q$ (in other words, there is a unique characteristic which arrives at q).*

A consequence is that u_- is differentiable at every point of a characteristic except possibly at its extremity. This comes from the fact that, because they are necessarily regular, *two characteristics cannot meet except at a common end-point.*

Invariant measures and unicity of Mañé's critical energy. Let M_L be the set of all Borel probability measures μ on TM which are invariant under the flow of X_L. If $(q_s, \dot{q}_s) = \varphi_s(q_0, \dot{q}_0) \in T\boldsymbol{T}^n$ is an extremal, i.e. an integral curve of X_L, integrating the domination inequality against an invariant measure $\mu \in M_L$, we get

$$\int_{TT^n} \big(u_-(q_t) - u_-(q_0)\big) d\mu \leq \int_{TT^n} d\mu \left[\int_0^t L(\varphi_s(q_0, \dot{q}_0)) ds + ct \right].$$

Dividing both sides by t and using the invariance of μ we get, when $t \to \infty$,

$$c = - \inf_{\mu \in M_L} \int L d\mu.$$

Minimizing measures can indeed be constructed which are supported in the α-limit sets of the minimizing extremals γ_-^q (figure 20) or the ω-limit sets of the γ_+^q : let t_n be a sequence of times tending to $+\infty$, μ_n be defined by

$$\mu_n(f) = \frac{1}{t_n} \int_{-t_n}^0 f\big(\gamma_-^q(s), \dot{\gamma}_-^q(s)\big) ds$$

and μ be the weak limit of a subsequence of the μ_n's. One easily checks that $\int L d\mu = -c$.

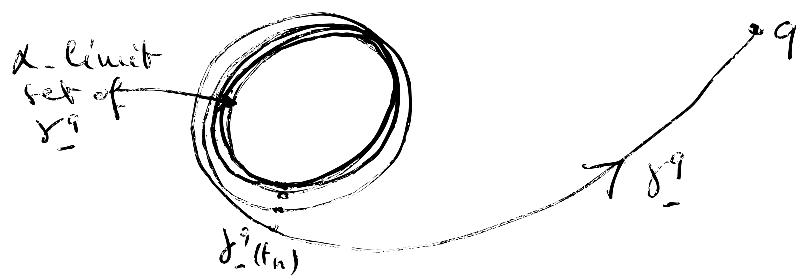

Figure 20

Mañé's critical energy and Hill's region for classical systems. When $L(q,\dot{q}) = \frac{1}{2}||\dot{q}||^2 - V(q)$, one checks that $c = \max_q V(q)$, i.e. the value of the energy under which the *Hill's region* (projection on $\mathbf{T}^n$ of an energy level) is not the whole configuration space $\mathbf{T}^n$.

The fundamental Lipschitz estimates. We now address the regularity problem of weak KAM solutions. Let $\mathcal{S}_-$ (resp. $\mathcal{S}_+$) denote the set of all weak KAM solutions u_- (resp. u_+).

Proposition. *The following assertions are equivalent :*
1) $u : \mathbf{T}^n \to I\!\!R$ is of class C^1 and belongs to $\mathcal{S}_-$;
2) $u : \mathbf{T}^n \to I\!\!R$ is of class C^1 and belongs to $\mathcal{S}_+$;
3) $u : \mathbf{T}^n \to I\!\!R$ belongs to $\mathcal{S}_- \cap \mathcal{S}_+$;
4) $u : \mathbf{T}^n \to I\!\!R$ is of class C^1 and $\exists c \in I\!\!R$ such that $H(du(q),q) = c$.
IN ALL CASES, $du(q)$ IS LOCALLY LIPSCHITZ (i.e., u is not only C^1 but $C^{1,1}$).

The main tool for the proof of the Lipschitz estimates is Lemma 2 of the third lecture. As $u_- = T_t^- u_- + ct$ and $u_+ = T_t^+ u_+ - ct$, it implies that, if u belongs to $\mathcal{S}_- \cap \mathcal{S}_+$, it satisfies

$$\left| u(q') - u(q) - \frac{\partial L}{\partial \dot{q}}\big(q, \dot{\gamma}^q(0)\big)(q'-q) \right| \leq k||q'-q||^2.$$

Here γ^q is the concatenation of γ_-^q and γ_+^q, which is easily proved to be a minimizing extremal, hence regular. One concludes with the following lemma, whose proof is elementary but tricky :

Characterization of $C^{1,1}$ functions. *The following assertions are equivalent :*
1) $\exists k > 0$ such that, $\forall q, \exists l_q$, a linear form, with

$$|u(q+h) - u(q) - l_q(h)| \leq k||h||^2;$$

2) The function u is of class $C^{1,1}$, $du(q) = l_q$ and $\exists C > 0$ such that

$$|du(q')(h) - du(q)(h)| \leq C||q-q'|| \times ||h||.$$

These Lipschitz estimates play a fundamental role in Mather's theory. As was shown by Herman, they are intimately related to the positive definiteness assumption

5th lecture. Mather's theory.

The Mather set. The graph of the derivative of a weak KAM solution is semi-invariant under the flow of X_H. Mather's theory shows that it is, in a generalized sense, made of pieces of invariant manifolds of a "weakly hyperbolic" fully invariant set, the (image under the Legendre diffeomorphism of the) *Mather set* :

Definition. *The Mather set $\tilde{\mathcal{M}}_0 \subset T\boldsymbol{T}^n$ is the closure of the union of the supports of all invariant Borel probability measures which minimize $\int L d\mu$, that is such that $\int L d\mu = -c$, where c is the Mañé energy. $\tilde{\mathcal{M}}_0$ is invariant under the flow φ_t of X_L.*

Integrating the inequality $u \prec L + c$ against an invariant measure, one gets

Proposition (universal calibration). *Let $(q, \dot{q}) \in \tilde{\mathcal{M}}_0$ and let $\gamma(t)$ be the extremal with initial conditions $(q, \dot{q})$, that is $\varphi_t(q, \dot{q}) = (\gamma(t), \dot{\gamma}(t)) \in \tilde{\mathcal{M}}_0$. Then, for any $u \prec L + c$ and any $t \leq t'$, $\gamma|_{[t,t']}$ is (L, c, u)-calibrated.*

As such a u exists (for instance a weak KAM solution), this proposition implies that the extremals contained in the projection $\mathcal{M}_0$ of $\tilde{\mathcal{M}}_0$ on $\boldsymbol{T}^n$ are minimizing.

The structure theorem. *The Mather set $\tilde{\mathcal{M}}_0 \in T\boldsymbol{T}^n$ is a Lipschitz graph over its projection $\mathcal{M}_0$ in $\boldsymbol{T}^n$. Its image under the Legendre diffeomorphism $\Lambda : T\boldsymbol{T}^n \to T^*\boldsymbol{T}^n$ (which is well defined in the autonomous case) is contained in the critical energy level $H^{-1}(c)$.*

The proof, which uses Lemma 2 of section 3 in a crucial way, is completely analogous to that of the fundamental Lipschitz estimates at the end of section 4. True, we are not dealing any more with weak KAM solutions, but the situation is quite as good thanks to the above proposition which can be restated as :

On $\mathcal{M}_0$, a function $u \prec L + c$ is as good as a weak KAM solution .

One gets that any u such that $u \prec L + c$, is differentiable at $q \in \mathcal{M}_0$ and that, for each $(q, \dot{q}) \in \tilde{\mathcal{M}}_0$, one has $du(q) = \frac{\partial L}{\partial \dot{q}}(q, \dot{q})$. This implies that $\dot{q}$ is uniquely determined by q, i.e. that $\tilde{\mathcal{M}}_0$ is a graph over $\mathcal{M}_0$, precisely the image under the inverse Legendre diffeomorphism of the graph of du. Finally, we get as in section 4 that the map $q \mapsto \big(du(q), q\big)$ from $\mathcal{M}_0$ to $T^*\boldsymbol{T}^n$ is Lipschitzian with a Lipschitz constant independant of u. In other words, the restriction of u to $\mathcal{M}_0$ is $C^{1,1}$ when a priori u is only Lipschitzian. These estimates are a fundamental feature of Hamilton-Jacobi theory. They have numerous avatars, like the a priori Lipschitz estimates on invariant curves of monotone twist mappings (a direct application of the extension of the theory to time periodic Lagrangians) or the Lipschitz property of Buseman functions, which correspond to the Lagrangian of a geodesic flow : if a C^1 function satisfies $||du|| \equiv 1$, it is automatically $C^{1,1}$.

The following diagram summarizes the situation :

$$\begin{tikzcd}
\tilde{\mathcal{M}}_0 \subset T\mathbb{T}^n \arrow[rr, "\Lambda", "\cong"'] \arrow[ddr] & & T^*\mathbb{T}^n \supset \Lambda(\tilde{\mathcal{M}}_0) \arrow[ddl] \\
(q,\dot q) & & \left(q, du(q)=\frac{\partial L}{\partial \dot q}(q,\dot q)\right) \\
& \mathbb{T}^n \supset \mathcal{M}_0 \ni q \arrow[ul] \arrow[ur] &
\end{tikzcd}$$

Conjugate weak KAM solutions. Weak KAM solutions are determined by their restriction to the Mather set $\mathcal{M}_0$. This allows to pair weak KAM solutions u_+, u_- so that the graphs of their (almost everywhere defined) derivatives play the role of (generalized) stable and unstable manifolds to $\tilde{\mathcal{M}}_0$.

Convergence of the Lax-Oleinik semi-group in the autonomous case. This is a kind of a generalized λ-*lemma*, which states that (*in the autonomous case only*), for any $u \in C^0(M, I\!\!R)$, the limits when $t \to \infty$ of $T_t^- + ct$ and $T_t^+ - ct$ exist and are weak KAM solutions u_- or u_+.

Mather's alpha function as an averaged Hamiltonian. Using a control, that is replacing the Lagrangian $L(q,\dot q)$ by $L_a(q,\dot q) = L(q,\dot q) - a\cdot\dot q$, one defines a critical energy c_a and a Mather set $\mathcal{M}_a$. Note that, if replacing L by L_a does not change the Euler-Lagrange equations, it DOES CHANGE THE MINIMIZERS of the action integral. The simplest example is the geodesic flow of the flat torus : adding a mass to better distinguish between the tangent and cotangent sides, let us take $L(q,\dot q) = \frac{m}{2}||\dot q||^2$. The Lagrangian L_a can be written

$$L_a(q,\dot q) = \frac{m}{2}||\dot q||^2 - a\cdot\dot q = \frac{m}{2}||\dot q - \frac{1}{m}a||^2 - \frac{||a||^2}{2m},$$

and the minimizers are immediately seen to be such that $\dot q = \frac{1}{m}a$, that is $p = a$. We have "controlled" (the word is from Kaloshin) the velocity (or momentum) of the minimizers. Recall that the Legendre transform of $L_a(q,\dot q)$ is $H_a(p,q) = H(p+a,q)$.

Definition-Proposition. *Mather's alpha function is the function*

$$\alpha : \mathcal{H}^1(\boldsymbol{T}^n, I\!\!R) \equiv (I\!\!R^n)^* \text{ defined by } \alpha(a) = c_a.$$

It is convex and superlinear.

It can be checked that for any compact manifold M, α is naturally defined on the first cohomology group of M, that is : the Mañé energy $c(L-\varpi)$ depends only on the cohomology class of the closed 1-form ϖ.

In the case of the torus, we can interpret α as an averaged Hamiltonian in the following sense : let us pretend that to each $a \in (\mathbb{R}^n)^*$ we can associate in a differentiable way (even continuity is false because of non-unicity !) a weak KAM solution u_a. At each point q where u_a is differentiable, we have

$$H(a + du_a(q), q) = \alpha(a).$$

Setting $S(a,q) = a \cdot q + u_a(q)$, this is equivalent to

$$H\left(\frac{\partial S}{\partial q}(a,q), q\right) = \alpha(a).$$

The function S is of course not $\mathbb{Z}^n$-periodic in q, that is not defined on $\boldsymbol{T}^n$, but its derivative is. Hence S can be used as the *generating function* of the *symplectic transformation*

$$\Phi : \mathbb{R}^n \times \boldsymbol{T}^n \to \mathbb{R}^n \times \boldsymbol{T}^n, \ \ \Phi(p,q) = (a,b),$$

defined by

$$\begin{aligned} p &= \frac{\partial S}{\partial q}(a,q) = a + \frac{\partial u_a}{\partial q}(q), \\ b &= \frac{\partial S}{\partial a}(a,q) = q + \frac{\partial u_a}{\partial a}(q). \end{aligned}$$

We have $dp \wedge dq + db \wedge da = d^2 S = 0$, hence $da \wedge db = dp \wedge dq$, which is the preservation of the canonical symplectic form. This implies that in the new coordinates (a,b), the Hamiltonian vector-field X_H^* becomes $X_{H\circ\Phi^{-1}}^*$, that is X_α^*. As α does not depend on the variables b, Hamilton's equations take the particularly simple *completely integrable* form

$$\frac{da_i}{dt} = 0, \ \ \frac{db_i}{dt} = \frac{d\alpha}{da_i}(a),$$

which is similar to the one defining the geodesic flow of the flat torus.

This is not astonishing. If for each a there exists a unique weak KAM solution u_a which is differentiable, the collection of the graphs of these functions defines a foliation of the phase space by Lagrangian tori. The existence of such a foliation would imply in turn the existence of action-angle coordinates in which the flows on the invariant tori are linear.

In general, this foliation is neither uniquely nor eveywhere defined but nevertheless, it can be thought of as a kind of (non uniquely defined) *integrable skeleton* made of KAM tori which are graphs (if they exist) and (non uniquely defined) pieces of stable and unstable manifolds of "weakly hyperbolic" Mather sets. In the case of the geodesic flow of a torus of revolution, this amounts to forgetting the whole (open) resonance zone. In the case of a monotone twist map, which according to Moser is always the Poincaré return map of the Euler-Lagrange flow of a periodic Lagrangian whch satisfies the general convexity hypotheses, we get the union of the invariant curves and (non uniquely defined) pieces of stable and unstable manifolds of the Aubry-Mather sets.

Mather's beta functions as an averaged Lagrangian. Legendre-Fenchel duality converts Mather convex and superlinear α function into the convex and superlinear function

$$\beta : \mathbb{R}^n \equiv \mathcal{H}_1(\boldsymbol{T}^n), \mathbb{R}) \to \mathbb{R}$$

defined by

$$\beta(\rho) = \max_{a\in(\mathbb{R}^n)^*} \left(\langle a, \rho\rangle - \alpha(a)\right), \text{or equivalently}\, \alpha(a) = \max_{\rho\in\mathbb{R}^n} \left(\langle a, \rho\rangle - \beta(\rho)\right).$$

The rotation number of an invariant probability measure. Let μ be a Borel probability measure on $T\boldsymbol{T}^n$, *invariant under the flow of* X_L (we have denoted by M_L the set of these measures).

Proposition-Definition. *The rotation number $\rho(\mu) \in \mathbb{R}^n \equiv \mathcal{H}_1(\boldsymbol{T}^n, \mathbb{R})$ of μ is uniquely defined by the identity*

$$\int_{TT^n} \varpi d\mu = \langle [\varpi], \rho(\mu)\rangle\,,$$

valid for any closed 1-form ϖ on $\boldsymbol{T}^n$ ($[\varpi]$ is the cohomology class of ϖ).

The existence of $\rho(\mu)$ is proved by showing that, for any coboundary $\varpi = d\theta$ and any invariant (this hypothesis is fundamental) measure μ, the integral $\int \varpi d\mu$ is equal to 0.

The interpretation of the beta function is given by the following

Proposition. *The beta function is given by the formula*

$$\beta(\rho) = \min_{\mu\in M_L, \rho(\mu)=\rho} \int_{TT^n} L d\mu.$$

At first sight, the proof looks like a formal rewriting of the definition of α :

$$\begin{aligned}\alpha(a) &= -\min_{\mu\in M_L} \int (L_a) d\mu = \max_{\mu\in M_L} \left(\int a d\mu - \int L d\mu\right)\\ &= \max_{\rho\in R^n} \left(\max_{\mu\in M_L, \rho(\mu)=\rho} \left(\langle a, \rho\rangle - \int L d\mu\right)\right) = \langle a, \rho\rangle - \min_{\mu\in M_L, \rho(\mu)=\rho} \int L d\mu,\end{aligned}$$

which one compares to the definition of α in terms of β.

The justification comes from the

Fundamental proposition. *For any $\rho \in I\!R^n \equiv \mathcal{H}_1(\boldsymbol{T}^n, I\!R)$, there exists an invariant measure $\mu \in M_L$ with finite action $\int L d\mu$ and whose rotation vector $\rho(\mu)$ is equal to ρ.*

WARNING : THIS PROPOSITION BECOMES FALSE IF INVARIANT MEASURES ARE REPLACED BY INTEGRAL CURVES (Hedlund examples on $\boldsymbol{T}^3$, see below). This is the main reason why introducing invariant measures, whose supports are collections of integral curves, is unavoidable.

Mather's theory of minimal measures as a generalization of Hedlund's theory. The following theorem of Mather relates the Mather sets and the $\tilde{\boldsymbol{T}}^n$-minimizers. These are the extremals which, when lifted to the universal covering $I\!R^n$ of $\boldsymbol{T}^n$, minimise the action integral between any two of their points. They are the natural generalization to higher dimensions of Hedlund's *class A geodesics* on the 2-torus.

Theorem. *For any $a \in (I\!R^n)^*$, an extremal which is contained in $\mathcal{M}_a$ is a $\tilde{\boldsymbol{T}}^n$-minimizer.*

This does not mean that any vector ρ can be the rotation vector of a $\tilde{\boldsymbol{T}}^n$-minimizer. In fact, if Hedlund had shown that for any Riemannian metric on the 2-torus and any real number ρ, there exist class A geodesics which, in the universal covering $I\!R^2$, stay at a bounded distance of a straight line of slope ρ and hence have the rotation number ρ, he had also given an example of a Riemannian metric on $\boldsymbol{T}^3$ for which class A geodesics exist only for three rotation vectors. Indeed, to achieve a rotation vector, one needs in general to take averages on a set of extremals, not on a single one.

The last figure indicates why Mather sets are more likely in general to have gaps than to cover the whole $\boldsymbol{T}^2$: if we slightly deform the flat metric by making a small localized bump, $\boldsymbol{T}^2$-minimizers (i.e. class A geodesics) will avoid the bump to stay minimizing and this creates a gap.

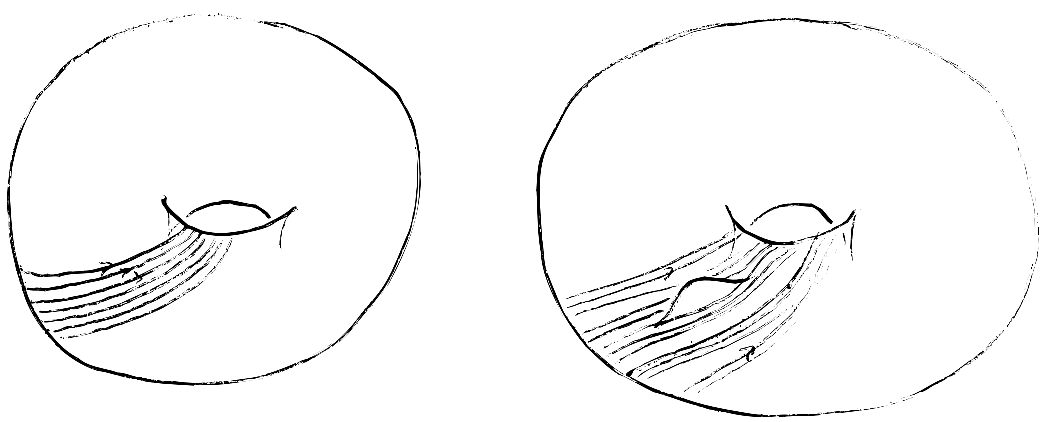

Figure 21