1 Brief History

It is common to look for maximum or minimum values of functions using the critical points. Nevertheless it is widely known that not all critical points turn out to correspond to relative extrema.

Example 1. For \( f : \mathbb{R}^2 \rightarrow \mathbb{R} \), defined by

\[
 f(x, y) = x^2 - y^2,
\]

we have that \( \nabla f = (f_x, f_y) = (0, 0) \) at \( (0, 0) \) and yet \( f(0, 0) \) is neither a maximum nor a minimum of \( f \).

The typical sort of problem we will study is to find among all functions with prescribed boundary conditions, those which minimize a given functional. More precisely, we will try to find

\[
 \min \{ I(u) : u \in X \text{ and } u = u_0 \text{ on } \partial \Omega \}
\]

where

\[
 I(u) = \int_{\Omega} F(x, u(x), Du(x)) \, dx,
\]

and \( \Omega \subset \mathbb{R}^n \) is a bounded open set, \( u : \Omega \rightarrow \mathbb{R}^m \), \( Du \in \mathbb{R}^{mn} \), \( F \in C(\Omega \times \mathbb{R}^m \times \mathbb{R}^{mn} \rightarrow \mathbb{R} \), \( u_0 \) is a given function and \( X \) is a real Banach space.

The history of the calculus of variations is older than the greek civilization; it is widely known the interest of ancient civilizations in solving problems related to “optimal forms”. The development of calculus implied a development of the calculus of variations. The first textbook on the area is due to Euler: “A method for finding curves enjoying certain maximum or minimum properties”, 1744. Euler expresses his belief that every effect in nature follows a maximum or minimum rule.

The calculus of variations has generated a great deal of mathematical work since then. During the 19th century, new methods were developed and important contributions were made by, among many others, Lagrange, Riemann, Weierstrass, Jacobi, Hamilton. These are the
so called classical methods. During the 20th century, different techniques were introduced by Hilbert and Lebesgue in connection with the study of the integral. These methods are now known as the direct methods.

We will list some classical examples:

- **Dido’s Problem**: given a certain closed curve, choose its shape so that it encloses the largest possible area.

- **Brachistochrome Problem**: Given two points \((x_0, y_0), (x_1, y_1)\) in the plane, choose the curve that joins them such that a particle under gravity slides from one to the other in the least amount of time possible. This was solved in 1696 by Bernoulli. The solution is obtained by finding the minimizer of the functional
  \[
  I(u) = \int_{x_0}^{x_1} \frac{\sqrt{1 + (u'(x))^2}}{\sqrt{2gu(x)}} \, dx,
  \]
  with \(u(x_0) = y_0, \quad u(x_1) = y_1\). \(^1\)

- **Fermat’s Principle of Least Time**: Given two points \((x_0, y_0), (x_1, y_1)\) in the plane, a light ray from one to the other, chooses the path between the points which requires the least time to transverse. Once again the solution corresponds to the minimizer of
  \[
  I(u) = \int_{x_0}^{x_1} \frac{\sqrt{1 + (u'(x))^2}}{\gamma(t, u(x))} \, dx,
  \]
  where \(\gamma\) is the speed of light in the medium.

- **Shortest Distance**: Given two points \((x_0, y_0), (x_1, y_1)\) in the plane, we want to find the curve joining them with smallest length. We obtain it, by minimizing the functional
  \[
  I(u) = \int_{x_0}^{x_1} \sqrt{1 + (u'(x))^2} \, dx.
  \]

If we work in a more general setting, we are given a Riemannian metric, and
\[
L(u) = I(u) = \int_0^1 \left( \sum a_{ij}(u(t))u_iu_j \right)^{1/2} \, dt
\]

Usually we try to minimize the length of the curve, i.e., find \(p\) such that

\[
I_1(p) \leq I_1(u)
\]

for all admissible \(u\). \(p\) is called a **minimal geodesic**

\(^1\)Notice that \(\sqrt{1 + (u'(x))^2} \, dx = dS\) and \(\sqrt{2gu(x)} = v\). Hence the integrand function is \(\frac{dS}{v}\).
- Hamilton’s Principle: Consider the system formed with \( n \) particles of mass \( m_i, i = 1, \ldots, n \). We will denote by \( x_i(t) \) the position of the \( i \)th particle at time \( t \) and \( \dot{x}_i(t) \) its velocity. The kinetic energy associated to this system is

\[
K(t) = \sum_{i=1}^{n} \frac{m_i}{2} \dot{x}_i^2,
\]

and the potential energy is a function \( V = V(x_1, \ldots, x_n) \). Then Hamilton’s Principle states that the motion of the system, for \( t_0 \leq t \leq t_1 \) is such that

\[
I(x) = \int_{t_0}^{t_1} \left( \sum_{i=1}^{n} \frac{m_i}{2} \dot{x}_i^2 - V(x_1, \ldots, x_n) \right) dt
\]

is extremized.

- Dirichlet’s Principle: Let \( D \subset \mathbb{R}^n \) be any bounded domain with smooth boundary, and consider the boundary value problem

\[
\begin{cases}
\Delta u = 0 & \text{for } x \in D \\
u = \phi & \text{on } \partial D
\end{cases}
\]

where \( \phi \) is any prescribed continuous function defined on \( \partial D \). The solutions of this problem minimize the functional

\[
\int_D |\nabla u|^2 dx,
\]

over the class of functions satisfying the boundary condition \( u = \phi \).

- Plateau’s Problem: Given \( \gamma \) a smooth, Jordan curve in \( \mathbb{R}^3 \), find the surface of minimal area bounded by \( \gamma \), i.e., find \( u : B_1(0) \subset \mathbb{R}^2 \to \mathbb{R}^3 \) such that \( u|_{\partial B_1(0)} = \gamma \) and

\[
A(u) = \int_{B_1(0)} \det(|\nabla u|^T |\nabla u|)^{1/2} dx
\]

is minimal.

The critical points of the functional \( I \), which are smooth functions will satisfy a differential equation, the so called Euler-Lagrange equations, together with some auxiliary conditions.

The classical approach to minimization problems consists in constructing the associated Euler equation, solve it and go back to prove that the solution is in fact a minimum value of \( I \). On the other hand, direct methods work directly with the functionals.

In \( \mathbb{R}^n \), suppose that we have a function with 2 relative minimum, satisfying a coercive property \( (f(x) \to \infty \text{ when } |x| \to \infty) \). If \( n = 1 \) then \( f \) must also have a relative maximum. If \( n > 1 \), we can’t guarantee the existence of a maximum, but \( f \) must have at least one more critical point. Our goal is to extend these sort of considerations to functionals defined other than in \( \mathbb{R}^n \).
Another type of application is to prove the existence of special type of solutions of an ODE. Consider
\[ \dot{x} = f(x), \quad x \in [a, b] \quad \text{and} \quad f(a) = f(b) = 0. \]
Then \( x(t) \equiv a \) and \( x(t) \equiv b \) are equilibrium solutions of the equation, i.e. they are time independent solutions of the boundary value problem. One possible thing to do is to prove the existence of heteroclinic solutions, asymptotic to the equilibrium, that is we seek for solutions of the ordinary differential equation satisfying
\[ x(t) \to a \quad \text{as} \quad t \to \infty, \]
and
\[ x(t) \to b \quad \text{as} \quad t \to -\infty. \]
Another class of problems is to solve variational problems associated to functionals exhibiting some sort of symmetry proprieties. Consider \( f \in C^1(\mathbb{R}^n, \mathbb{R}) \) such that \( f(-x) = f(x) \) for all \( x \in \mathbb{R}^n \). Let’s look at \( f|_{S^{n-1}} \). The critical points occur at antipodal points, and if one corresponds to a maximum value of \( f \) its antipodal will correspond to a minimum value of \( f \). In this case it is easy to prove that \( f \) has at least \( n \) pairs of critical points.

**Example 2.** For the special case
\[ f(x) = \frac{1}{2}(Mx, x), \]
where \( M \) is a \( n \times n \) symmetric matrix, with the constraint \( |x| = 1 \). Using Lagrange multipliers, we can show that the critical points of \( f \) must satisfy the equation
\[ Mx = \lambda x \]
and we notice that this is just the eigenvalue equation. Since a symmetric matrix has \( n \) distinct eigenvalues, \( f \) will have at least \( n \) distinct critical points.

## 2 Outline of the Course

1. Classical Methods: Euler equations and applications;

2. Topics in the calculus in Banach Spaces: function spaces, properties of functionals, Fréchet differentiability.

3. Direct methods: Minima of functionals.

4. Applications:

   Solutions of BVP, Periodic solutions, heteroclinic and homoclinic solutions.

5. Geodesics: Riemannian metric, curves of shortest length;
3 Euler Equation and Applications

The classical approach of the calculus of variations consists of, as remarked earlier, to find the critical points of a given functional. As in the finite dimensional case, one way of studying the problem is by finding the zeros of the “derivative” of $I$, $I'(u) = 0$, which are known as the Euler-Lagrange equations. We will derive the Euler-Lagrange equations in some specific cases.

3.1 The One-dimensional Case

Consider

$$I(u) = \int_a^b F(x, u, u') \, dx,$$

where $F = F(x, u, p) \in C^2([a, b] \times \mathbb{R}^2, \mathbb{R})$. The natural set of admissible functions is

$$\Gamma = \{ u \in C^1([a, b], \mathbb{R}) : u(a) = \alpha, \, u(b) = \beta \},$$

for some given real numbers $\alpha, \beta$. Our goal is to find

$$\inf_{u \in \Gamma} I(u)$$

Assume that there exists $v \in \Gamma$ such that $I(v) = \inf_{u \in \Gamma} I(u)$. We will try to obtain some necessary conditions for this to hold.

Pick any function $\varphi \in C^2([a, b], \mathbb{R})$ such that $\varphi(a) = \varphi(b) = 0$, and let $\delta \in \mathbb{R}$. It is easy to check that the function $v + \delta \varphi$ is still in $C^1([a, b], \mathbb{R})$ and by construction verifies

$$(v + \delta \varphi)(a) = \alpha, \quad (v + \delta \varphi)(b) = \beta$$
Hence $v + \delta \varphi \in \Gamma$. Fixing $\varphi$, define

$$
\psi(\delta) = I(v + \delta \varphi) = \int_a^b F(x, v + \delta \varphi, v' + \delta \varphi') \, dx.
$$

Since $v$ is the minimizer of $I$

$$
\psi(0) \leq \psi(\delta), \quad \forall \delta \in \mathbb{R}.
$$

Then $\psi$ has a global minimum at $\delta = 0$. On the other hand, the regularity of $F$ and the differentiability of $v + \delta \varphi$ imply that $\psi \in C^1(\mathbb{R}, \mathbb{R})$. It is known that the minimum of a differentiable real function is attained at the zero of the derivative. Thus

$$
\psi'(0) = 0
$$

and we can derive the equation

$$
\int_a^b \left( F_u(x, v, v') \varphi + F_p(x, v, v') \varphi' \right) \, dx = 0,
$$

denoted as the weak form of the Euler equation. Notice that, although we have derived the equation for fixed $\varphi$, it will hold for any function $\varphi$ in the conditions we assumed.

If we further assume that $v \in C^2([a, b], \mathbb{R})$, we can integrate by parts the second term of the equation, and obtain

$$
\int_a^b F_u(x, v, v') \varphi \, dx + F_p(x, v, v') \varphi|_a^b - \int_a^b \frac{d}{dx} \left( F_p(x, v, v') \right) \varphi \, dx = 0
$$

As a consequence of the boundary conditions on $\varphi$

$$
F_p(x, v, v') \varphi|_a^b = 0,
$$

and therefore

$$
\int_a^b \left( F_u(x, v, v') - \frac{d}{dx} F_p(x, v, v') \right) \varphi \, dx = 0,
$$

for all $\varphi \in C^2([a, b], \mathbb{R}), \varphi(a) = \varphi(b) = 0$.

**Homework 1.** Prove that, if for $h \in C([a, b])$

$$
\int_a^b h(x) \varphi(x) \, dx = 0,
$$

for all $\varphi \in C([a, b]), \varphi(a) = \varphi(b) = 0$, then $h(x) \equiv 0$ on $[a, b]$.

Using the above result, we can finally derive the Euler equation

$$
F_u(x, v, v') - \frac{d}{dx} F_p(x, v, v') = 0,
$$
or computing the derivative, we obtain the expanded form of the Euler equation

\[ F_{pp}(x, v, v')v'' + F_{pu}(x, v, v')v' + F_{px}(x, v, v') - F_u(x, v, v') = 0, \]

with \( v(a) = \alpha, v(b) = \beta \). Minimizing \( I \) over \( \Gamma \) leads to this boundary value problem. It is a necessary condition for the existence of solution.

**Example 3.** Consider

\[ I(u) = \int_a^b \sqrt{1 + (u')^2} \, dx, \quad u(a) = \alpha, \quad u(b) = \beta. \]

As we remarked before, the minimizer of \( I \) is the curve of shortest length joining \((a, \alpha)\) to \((b, \beta)\). Then,

\[ F(x, u, u') = \sqrt{1 + (u')^2} \]

and the Euler equation is

\[ \frac{d}{dx} \frac{u'}{\sqrt{1 + (u')^2}} = 0, \]

which implies that

\[ u'(x) = c\sqrt{1 + (u')^2} \quad \text{for} \quad |c| < 1. \]

Solving for \( u' \) we obtain

\[ (u'(x))^2 = \frac{c^2}{1 - c^2}, \]

and therefore we can conclude that \( u' \) is constant. Hence the solution of the Euler equation is the linear function through the points \((a, \alpha)\) to \((b, \beta)\),

\[ u(x) = \frac{x - a b^{-a}}{b - a} \beta + \frac{b - x b^{-a}}{b - a} \alpha. \]

Nevertheless we still have to prove that \( u \) is exactly the solution of the minimization problem - it might exist a minimizer of \( I \) not in \( C^2 \).

**Example 4.** Consider

\[ I(u) = \int_a^b F(u, u') \, dx, \]

where \( F \) is as before, but independent of \( x \). Then

\[ \frac{d}{dx} \left( F(u, u') - u' F_p(u, u') \right) = F_u u' + F_p u'' - u'' F_p - u' \frac{d}{dx} F_p = u' \left( F_u - \frac{d}{dx} F_p \right) \]

If \( u \) is a solution of the Euler equation

\[ \frac{d}{dx} \left( F(u, u') - u' F_p(u, u') \right) = 0, \]
which implies that the solutions of the Euler equation satisfy the first order differential equation

\[ F(u, u') - u' F_u(u, u') = \text{const.} \]  

which is a simpler equation to solve. As an application of this, consider the brachistochrome problem,

\[ F(x, u, u') = \sqrt{1 + (u')^2} \]

which is independent of \( t \). Using equation (1) evaluated at a solution of the Euler equation, we obtain

\[ \frac{\sqrt{1 + (v')^2}}{\sqrt{2gv}} - \frac{(v')^2}{\sqrt{2gv(1 + (v')^2)}} = c_1 \]

which implies

\[ \frac{1}{\sqrt{2gv(1 + (v')^2)}} = c_1. \]

Take \( c = \sqrt{2gc_1} \), and solving for \( v' \)

\[ (v')^2 = \frac{1}{c^2v} - 1. \]

Assume that \( v' > 0 \), which is a reasonable assumption since we are trying to reach the point \((x_1, y_1)\) in the least possible time. Then

\[ \frac{dv}{dx} = \sqrt{\frac{1 - c^2v}{c^2v}} \iff \frac{dx}{dv} = \sqrt{\frac{c^2v}{1 - c^2v}}. \]

We will make some changes of variable, in order to obtain a parametrization of the solution. Take \( w = c^2v \). Then the equation becomes

\[ c^2 \frac{dx}{dw} = \sqrt{\frac{w}{1 - w}}. \]

Making the trigonometric substitution \( w = \sin^2 \theta \),

\[ \frac{dx}{dw} = \frac{dx}{d\theta} \frac{d\theta}{dw} = \frac{dx}{d\theta} \frac{1}{2\sin \theta \cos \theta} \]

Replacing in the differential equation

\[ \frac{dx}{d\theta} = \frac{1}{c^2(2\sin \theta \cos \theta)^{-1}} \frac{dx}{dw} = \frac{2}{c^2} \sin^2 \theta \]

Solving the differential equation

\[ x(\theta) = \alpha + \frac{1}{c^2} (\theta - \frac{1}{2} \sin(2\theta)) \]
For simplicity, assume that the initial point is at \((0,0)\), implying \(\alpha = 0\). For \(R = \frac{1}{2}\), the parametric solutions of the problem are

\[x(\theta) = R(2\theta - \sin(2\theta)) \text{ and } v(\theta) = \frac{\sin^2 \theta}{c^2} = R(1 - \cos(2\theta))\]

Recalling that the parametric equations for a cycloid are

\[X(\phi) = R(\phi - \sin \phi), \quad Y(\phi) = R(1 - \cos \phi)\]

the solution for our problem is a cycloid.

### 3.2 The Vectorial Case

We will now derive the Euler equation associated to the some minimization problem, for \(x \in \mathbb{R}\) and \(u(x) = (u_1(x), \ldots, u_n(x))\). We will take

\[\Gamma = \{ u \in C^1([a,b], \mathbb{R}^n) : u(a) = \alpha, u(b) = \beta, \alpha, \beta \in \mathbb{R}^n \},\]

as the class of admissible functions. Once again we suppose that there exists \(v \in \Gamma\) minimizing \(I\). Our goal is to derive a necessary condition satisfied by \(v\). Pick \(\phi \in C^1([a,b], \mathbb{R}^n)\) with \(\phi(a) = \phi(b) = 0\), and let \(\delta \in \mathbb{R}\) be arbitrary. Define

\[\psi(\delta) = I(v + \delta \phi)\]

and for fixed \(\phi\), \(\psi\) is a differentiable function in \(\delta\). Moreover, since \(v\) is the minimum of \(I\) over \(\Gamma\), and \(v + \delta \phi \in \Gamma\)

\[\psi(0) \leq \psi(\delta), \quad \forall \delta \in \mathbb{R},\]

which will imply, as in the previous case, that \(\psi'(0) = 0\). Hence we can write weak form of the Euler equation

\[
\int_a^b \left( F_{u_1}(x,v,v') \phi_1 + \ldots + F_{u_n}(x,v,v') \phi_n + F_{p_1}(x,v,v') \phi_1' + \ldots + F_{p_m}(x,v,v') \phi_m' \right) dx = 0.
\]

If, furthermore \(v \in C^2([a,b], \mathbb{R}^n)\), we can integrate by parts and conclude

\[
\int_a^b \left( F_u(x,v,v') \cdot \phi - \frac{d}{dx} F_p(x,v,v') \cdot \phi \right) dx = 0, \quad \forall \phi.
\]

Using a slight variation of the result of Homework 1

\[F_u(x,v,v') - \frac{d}{dx} F_p(x,v,v') = 0,\]

and this system of ordinary differential equations represents the Euler equation for this problem.

\[\text{Take } \phi = (\varphi_1, 0, \ldots, 0) \text{ and use Homework 1 to conclude that the first component of } h \equiv 0. \text{ Use this process inductively.}\]
Example 5. Consider the functional

$$I(p, q) = \int_{a}^{b} (p \cdot \dot{q} - H(p, q)) \, dt \quad , \quad p, q \in \mathbb{R}^n.$$  

The quantity $p \cdot \dot{q} - H(p, q)$ is usually denoted by Lagrangian, $\int_{a}^{b} p \cdot \dot{q} \, dt$ the action integral and $H(p, q)$ the Hamiltonian. By Hamilton’s Principle, the motion of a system of $n$ particles with no friction extremize $I(p, q)$. Let’s derive the Hamilton’s equation: the necessary condition that any minimizer of $I(p, q)$ must satisfy. If $(p, q)$ is a minimizer of $I$, for $\varphi, \psi \in C^1$ with $\varphi(a) = \varphi(b) = 0$ and $\psi(a) = \psi(b) = 0$, we define

$$\Psi(\delta) = \int_{a}^{b} \left( (p + \delta \varphi) \cdot (\dot{q} + \delta \dot{\psi}) - H(p + \delta \varphi, q + \delta \psi) \right) \, dt.$$  

As in the previous cases, $\Psi$ is a differentiable function of $\delta$ attaining its minimum value at $\delta = 0$, and therefore

$$0 = \Psi'(0) = \int_{a}^{b} \left( \varphi \cdot \dot{q} + p \cdot \dot{\psi} - H_p(p, q) \cdot \varphi - H_q(p, q) \cdot \psi \right) \, dt.$$  

Assuming that $p$ is twice differentiable, we can integrate by parts and as a consequence of the boundary conditions

$$0 = \int_{a}^{b} \left( (\dot{q} - H_p(p, q)) \cdot \varphi - (\dot{p} + H_q(p, q)) \cdot \psi \right) \, dt \quad , \quad \forall \varphi, \psi$$  

Then we can write Hamilton’s equation (which is a $2n$ dimensional system)

$$\begin{cases} \dot{q} = H_p(p, q) \\ \dot{p} = -H_q(p, q) \end{cases}$$  

For the simplest case

$$H(p, q) = \frac{1}{2} |p|^2 + V(q),$$  

that is, in the case where the Hamiltonian is the total energy of the system,

$$\begin{cases} \dot{q} = p \\ \dot{p} = -\nabla V(q) \end{cases}$$  

which can be written in the form

$$\ddot{q} + \nabla V(q) = 0$$  

corresponding to Newton’s second law of motion.
3.3 Case $F = F(x, u, u', ..., u^{(k)})$

Let us study next the case

$$I(u) = \int_{a}^{b} F(x, u, u', ..., u^{(k)}) dx.$$ 

The minimization of such functional is connected to problems of elasticity. We will minimize $I$ over the natural class of functions, $\Gamma$, of the $k$ differentiable functions $u$ verifying the boundary conditions

$$(u, u', ..., u^{(k-1)})(a) = (\alpha, \alpha_1, ..., \alpha_{k-1}) \text{ and } (u, u', ..., u^{(k-1)})(b) = (\beta, \beta_1, ..., \beta_{k-1}).$$

for some given $(\alpha, \alpha_1, ..., \alpha_{k-1})$, $(\beta, \beta_1, ..., \beta_{k-1}) \in \mathbb{R}^n$. Suppose that there is a minimizer of $I$, $v \in \Gamma$.

**Homework 2.** Derive the correspondent Euler equations (You might need to make some further assumptions).

3.4 The Scalar Case

Another problem is to minimize

$$I(u) = \int_{\Omega} F(x, u, Du) \, dx,$$

where $\Omega \subset \mathbb{R}^n$ is a bounded, open set with smooth boundary. The set of admissible functions is

$$\Gamma = \{ u \in C^1(\Omega, \mathbb{R}) : u|_{\partial \Omega} = f(x) \}$$

where $f$ is a given smooth function. Suppose it exists $v \in \Gamma$ such that $I(v) = \inf_{\Gamma} I$. To find the Euler equations we will proceed as in the previous examples. Take $\varphi \in C^1(\Omega, \mathbb{R})$ such that $\varphi|_{\partial \Omega} = 0$. Then $v + \delta \varphi \in \Gamma$ and we can define

$$\psi(\delta) = I(v + \delta \varphi) = \int_{\Omega} F(x, v + \delta \varphi, (v_{x_1} + \delta \varphi_{x_1}, ..., v_{x_n} + \delta \varphi_{x_n})) dx.$$ 

Although it is not as simple as the previous cases, we can prove that $\psi$ is a differentiable function of $\delta$, and

$$\psi(0) \leq \psi(\delta) \quad , \quad \forall \delta \in \mathbb{R}$$

Hence

$$0 = \psi'(\delta) = \int_{\Omega} \left( F_u(x, v, Du) \varphi + F_{p_1}(x, v, Du) \varphi_{x_1} + \ldots + F_{p_n}(x, v, Du) \varphi_{x_n} \right) dx$$

$$= \int_{\Omega} \left( F_u(x, v, Du) \varphi + F_p(x, v, Du) \cdot \nabla \varphi \right) dx.$$
Assuming that $v \in C^2$, we can use the Divergence Theorem. We recall that

$$\int_{\Omega} \nabla \cdot v \, dx = \int_{\partial \Omega} v \cdot \nu \, dS,$$

where $\nu(x)$ is the outward pointing unit normal to $\partial \Omega$. Look at

$$\nabla \cdot (\varphi F_p) = \sum_i \frac{\partial}{\partial x_i} (\varphi F_{p_i}) = \sum_i \frac{\partial \varphi}{\partial x_i} F_{p_i} + \varphi \sum_i \frac{\partial F_{p_i}}{\partial x_i} = \nabla \varphi \cdot F_p + \varphi \nabla \cdot F_p$$

which implies

$$\nabla \varphi \cdot F_p = \nabla \cdot (\varphi F_p) - \varphi \nabla \cdot F_p$$

Using this equality, we can write the weak Euler equation as

$$0 = \int_{\Omega} \left( F_u(x, v, Dv) - \nabla \cdot F_p \right) \varphi + \int_{\Omega} \nabla \cdot (\varphi F_p)$$

and applying the Divergence Theorem

$$0 = \int_{\Omega} \left( F_u(x, v, Dv) - \nabla \cdot F_p \right) \varphi + \int_{\partial \Omega} F_p \varphi \cdot \nu dS$$

Since $\varphi = 0$ on $\partial \Omega$

$$\int_{\Omega} \left( F_u(x, v, Dv) - \nabla \cdot F_p \right) \varphi \, dx = 0,$$

for all such $\varphi$. Using the usual arguments, one can show that this implies

$$F_u(x, v, Dv) - \nabla \cdot F_p = 0$$

or

$$F_u(x, v, Dv) - \sum_i \frac{\partial}{\partial x_i} F_{p_i}(x, v, Dv) = 0 \quad (2)$$

**Example 6.** Take $F(x, u, Du) = \frac{1}{2} |\nabla u|^2 - G(x, u)$. Using (2), the correspondent Euler equation is

$$-G_u(x, u) - \sum_i \frac{\partial}{\partial x_i} u_{x_i} = 0$$

which can be written as

$$\begin{cases} G_u(x, u) + \Delta u = 0 & , \ x \in \Omega \\ u = f & , \ \text{on} \ \partial \Omega \end{cases}$$
3.5 Free Boundary Problem

We will consider next the case, where no boundary conditions are assumed. We will minimize

$$I(u) = \int_{a}^{b} F(x, u, u') \, dx$$

over the class of admissible functions

$$\Gamma = C^1([a,b], \mathbb{R}).$$

This problem is usually denoted by free boundary problem. Proceeding as in the previous cases, for arbitrary $\varphi \in \Gamma$ and $v$ the minimizer of $I$ over $\Gamma$, the weak Euler equation is verified, i.e.,

$$\int_{a}^{b} \left( F_u(x, v, v') \varphi + F_p(x, v, v') \varphi' \right) \, dx = 0 \quad \forall \varphi \in \Gamma$$

Further assuming that $v \in C^2$, we can integrate by parts and obtain

$$\int_{a}^{b} \left( F_u(x, v, v') \varphi - \frac{d}{dx} F_p(x, v, v') \varphi \right) \, dx + F_p(x, v, v') \varphi |_{a}^{b} = 0 \quad (3)$$

Choose $\varphi \in \Gamma$ such that $\varphi(a) = \varphi(b) = 0$. For this particular case equation (3) becomes

$$\int_{a}^{b} \left( F_u(x, v, v') - \frac{d}{dx} F_p(x, v, v') \right) \varphi \, dx = 0$$

and, as before, this implies that $v$ is a solution of the Euler equation

$$F_u(x, v, v') - \frac{d}{dx} F_p(x, v, v') = 0 \quad (4)$$

For an arbitrary $\varphi \in \Gamma$, (3) and (4) yields

$$0 = F_p(b, v(b), v'(b)) \varphi(b) - F_p(a, v(a), v'(a)) \varphi(a).$$

Choose $\varphi$ such that $\varphi(a) = 0$ and $\varphi(b) = 1$. Then

$$F_p(b, v(b), v'(b)) = 0. \quad (5)$$

Reversing the roles of $a$ and $b$, we also obtain

$$F_p(a, v(a), v'(a)) = 0. \quad (6)$$

We conclude that although our problem has no boundary conditions, the minimizer must verify some sort of boundary conditions. (5) and (6) are denoted as natural boundary conditions.
3.6 Constrained Variational Problem

We want to minimize

\[ I(u) = \int_a^b F(x, u, u') \, dx \]

subject to the integral constraint

\[ J(u) = \int_a^b G(x, u, u') \, dx = c \]

for some given constant \( c \) and \( G \in C^2([a, b] \times \mathbb{R}^2, \mathbb{R}) \). Define the class

\[ \Gamma = \{ u \in C^1 : J(u) = c \}. \]

If \( v \) is a minimizer of \( I \) over the \( \Gamma \), consider \( \varphi \in C^1([a, b], \mathbb{R}) \) and \( \delta \in \mathbb{R} \) such that \( v + \delta \varphi \in \Gamma \). Define \( \psi(\delta) = I(v + \delta \varphi) \) subject to \( \chi(\delta) \equiv J(v + \delta \varphi) = c \). By construction, the minimum value of \( \psi \) verifying \( \chi(\delta) = c \) is achieved at \( \delta = 0 \). Then using Lagrange multipliers

\[
\frac{d}{d\delta} \left( \psi(\delta) + \lambda \chi(\delta) \right) = 0 \quad \text{at} \quad \delta = 0. \tag{7}
\]

**Example 7.** Minimize \( \int_0^\pi (u'(t))^2 \, dt \) subject to \( u(0) = u(\pi) = 0 \) and \( \int_0^\pi u^2(t) \, dt = 1 \). Notice that the constraint excludes the obvious candidate \( u \equiv 0 \). By (7) we may write the equation

\[ \int_0^\pi (u' \varphi' + \lambda u \varphi) \, dt = 0 \]

If \( u \in C^2([a, b], \mathbb{R}) \)

\[ \frac{d}{dx} \left( F_p(x, u, u') + \lambda G_p(x, u, u') \right) - (F_u(x, u, u') + \lambda G_u(x, u, u')) = 0, \]

which is equivalent to

\[ \begin{cases} u'' - \lambda u = 0 \\ u(0) = u(\pi) = 0 \end{cases} \]

This is an eigenvalue problem for ordinary differential equations: we are looking for the nontrivial solutions of the boundary value problem. It is known that the eigenfunctions of this problem are \( u_k(t) = \sin(kt) \), corresponding to the eigenvalues \( \lambda_k = -k^2 \).

**Homework 3.** Find the curve with endpoints \( u(a) = \alpha, u(b) = \beta \) of shortest length which has area \( A \) below it.

To derive the Euler equation, we need to assume that the minimizer is a \( C^2 \) function. Solving the Euler equation yields a \( C^2 \) candidate for the minimum of \( I \). But since the class of admissible functions is, roughly speaking, \( C^1 \), the minimum value of \( I \) may be achieved in \( C^1 \setminus C^2 \). We would like to derive a procedure, which allows us to show when the solution of the Euler equation is in fact the minimizer of the functional.

\[^3\text{Other type of constraints may be } |u(x)| = 1, \text{ or more generally that } u(x) \text{ lies on a given manifold.}\]
Example 8. In the case of the curve with shortest length, we saw that the solution of the Euler equation is the straight line through \((a, \alpha)\) and \((b, \beta)\). We will prove that in fact is the minimizer of \(I\). Writing the solution explicitly

\[
v(x) = \frac{x-a}{b-a} \beta + \frac{b-x}{b-a} \alpha
\]

we can compute

\[
I(v) = \int_a^b \sqrt{1 + (v')^2} \, dx = \int_a^b \sqrt{1 + \left(\frac{\beta - \alpha}{b-a}\right)^2} \, dx = \sqrt{(b-a)^2 + (\beta - \alpha)^2}
\]

which is the formula for the distance between two points.

Notice that minimizing \(I\), depends on our choices of the endpoints \(A(a, \alpha), B(b, \beta)\). Fix \(A\) at \((0,0)\) and let \(B = (x, y)\) change. For each \((x, y)\) we obtain a function \(v(t)\) and

\[
I(v) = (x^2 + y^2)^{1/2} \equiv S(x, y).
\]

Define

\[
\varphi(x, y, p) = S_x(x, y) + S_y(x, y)p,
\]

For arbitrary \(\gamma \in \Gamma = \{u \in C^1([0, x]) : u(0) = 0 \text{ and } u(x) = y\}\), we will have

\[
\varphi(t, \gamma(t), \gamma'(t)) = S_x(t, \gamma(t)) + S_y(t, \gamma(t))\gamma'(t) = \frac{d}{dt} S(t, \gamma(t))
\]

Then

\[
\int_0^x \varphi(t, \gamma, \gamma') \, dt = \int_0^x \frac{d}{dt} S(t, \gamma) \, dt = S(x, y) = I(v) \tag{8}
\]

On the other hand, if

\[
\varphi(x, y, p) \leq \sqrt{1 + p^2}, \quad \forall x, y, p \tag{9}
\]

then

\[
\int_0^x \varphi(t, \gamma, \gamma') \, dt \leq \int_0^x \sqrt{1 + (\gamma')^2} \, dt = I(\gamma) \tag{10}
\]

It is easy to see that (8) and (10) imply

\[
I(v) \leq I(\gamma), \quad \forall \gamma \in \Gamma
\]

and therefore \(v\) is the minimizer of \(I\) over \(\Gamma\). All we have to do is prove that (9) holds. Notice that

\[
\varphi(x, y, p) = \frac{x}{\sqrt{x^2 + y^2}} + \frac{yp}{\sqrt{x^2 + y^2}} = \frac{(x, y) \cdot (1, p)}{\sqrt{x^2 + y^2}}
\]

Using Schwarz inequality

\[
(x, y) \cdot (1, p) \leq \sqrt{x^2 + y^2} \sqrt{1 + p^2}
\]

and (9) is verified.
Let’s try to generalize this procedure, in order to prove that the solution of the Euler equation is in fact the minimizer of the functional
\[
I(u) = \int_a^b F(x, u, u') \, dt
\]
over the class
\[
\Gamma = \{ u \in C^1([a, b]) : u(a) = \alpha \text{ and } u(b) = \beta \}.
\]
Fixing \( A \) at \((0, 0)\) and letting \( B(x, y) \) change, for \( v \) the minimizer of \( I \) over \( \Gamma \)
\[
I(v) = S(x, y).
\]
Defining
\[
\varphi(x, y, p) = S_x(x, y) + S_y(x, y)p.
\]
for any
\[
\gamma \in \Gamma = \{ u \in C^1([0, x]) : u(0) = 0 \text{ and } u(x) = y \},
\]
as above
\[
\varphi(t, \gamma(t), \gamma'(t)) = \frac{d}{dt} S(t, \gamma(t))
\]
and therefore
\[
I(v) = \int_0^x \varphi(t, \gamma(t), \gamma'(t)) \, dt.
\]
If we can prove
\[
\varphi(x, \gamma(t), \gamma'(t)) \leq \frac{d}{dt} S(t, \gamma(t))
\]
doing the same analysis we can prove that the solution of the Euler equation is in fact the minimizer of \( I \) over \( \Gamma \).

To summarize, we consider the functional
\[
I(u) = \int_a^b F(x, u, u') \, dx,
\]
where \( F \) is a smooth function in its arguments. The procedure we have been following is

1. Choose a class of admissible functions, \( \Gamma \).

2. **Existence of Minimizer.** Assume that there exists \( v \in \Gamma \) such that \( I(v) = \inf \Gamma \). This implies that \( v \) is a solution of the weak Euler equation
\[
\int_a^b \left( F_u(x, v, v') \varphi + F_p(x, v, v') \varphi' \right) \, dx = 0
\]
for all \( \varphi \in C^1([a, b]) \) such that \( \varphi(a) = \alpha \) and \( \varphi(b) = \beta \).

3. **Regularity.** Assume that \( v \in C^2([a, b]) \) and the Euler equation is satisfied.

So far we have been working with the class
\[
\Gamma = \{ u \in C^1([a, b], \mathbb{R}) : u(a) = \alpha, u(b) = \beta \}
\]
but notice that we can choose to work with other classes of functions. The bigger the class \( \Gamma \) is, the smaller the infimum might be. For example if we take the set of piecewise \( C^1 \) functions, instead of \( C^1 \),
\[
\inf_{PC^1} I \leq \inf_{C^1} I.
\]
**Example 9.** Consider \( I(u) = \int_{-1}^{1} (1 - (u')^2)^2 dx \) with \( u(-1) = u(1) = 0 \). It is obvious that \( I(u) \geq 0 \), which implies that \( \inf I \geq 0 \). But notice that \( I(v) = 0 \) for \( v(x) = |x| \). Hence \( \inf I = 0 \) and therefore \( |x| \) is one minimizer of \( I \) which is not \( C^1([-1,1]) \).

**Homework 4.** Find all the minimizers of \( I \) as in Example 9, that lie in \( PC^1([-1,1]) \). Notice that none of it is in \( C^1([-1,1]) \).

**Example 10.** Consider \( I(u) = \int_{-1}^{1} |1 - \dot{u}|^2 u^2 dx \) with \( u(-1) = 0 \) and \( u(1) = 1 \). Once again \( \inf I \geq 0 \) and for \( v(x) = \max\{0, x\} \), \( I(v) = 0 \). Hence \( \inf I = 0 \). Check that \( v \) is the only minimizer of \( I \).

**Example 11.** Let \( I(u) = \int_{-1}^{1} (2x - u')^2 u^2 dx \) with \( u(-1) = 0 \) and \( u(1) = 1 \). Then \( I(v) \geq 0 \) and for
\[
v(x) = \begin{cases} 
0 & \text{if } x \leq 0 \\
x^2 & \text{if } x > 0 
\end{cases}
\]
\( I(v) = 0 \). Note that \( v \in C^1 \) but still not in \( C^2 \) (meaning that it is not solution of the Euler equation).

## 4 Topics in Calculus in Banach Spaces

In order to solve our minimization problems in a more general setting, we need some notions in an abstract set. Consist of

- working on appropriate spaces,
- continuity properties,
- notions of differentiability,

Let \( E \) be a real Banach space.

- \( C([a, b], \mathbb{R}) \) with norm \( \|u\| = \max_{x \in [a, b]} |u(x)| \);
- \( C([a, b], \mathbb{R}^n) \) with norm \( \|u\| = \max_{x \in [a, b]} |u(x)| \);
- \( C(\Omega, \mathbb{R}), \Omega \subset \mathbb{R}^n \) compact, with norm \( \|u\| = \max_{x \in \Omega} |u(x)| \);
- \( C^1([a, b]) \) with norm \( \|u\|_{C^1([a, b])} = \|u\| + \|u'\| \);
- \( L^p([a, b]) \) with norm \( \|u\|_p = \left( \int_a^b |u(x)|^p dx \right)^{1/p} \).
The particular case $p = 2$, $L^2([a, b])$ is a Hilbert space: the norm comes from the inner product

$$(u, v)_{L^2([a, b])} = \int_a^b u(x)v(x) \, dx$$

In an integral setting, $L^2$ is the analogue of continuous functions. The analogues of the spaces $C^p$ are the Sobolev spaces.

$W^{1,2}([a, b])$ is the completion of $C^1([a, b])$ under the norm

$$\|u\|_{W^{1,2}([a, b])} = \left( \int_a^b \left( u^2(x) + (u'(x))^2 \right) \, dx \right)^{1/2}.$$  

In the case $[a, b] = [0, 2\pi]$, a function in $L^2([0, 2\pi])$ may be seen as a Fourier series

$$u(x) = \sum_{k \in \mathbb{Z}} a_k e^{ikx},$$

and by Parseval’s identity

$$\int_0^{2\pi} |u(x)|^2 \, dx = \sum_{k \in \mathbb{Z}} |a_k|^2.$$  

If $u \in W^{1,2}([0, 2\pi])$, the same idea can be applied to $u'$ and therefore

$$\|u\|_{W^{1,2}([0, 2\pi])} = \sum_{k \in \mathbb{Z}} (1 + k^2) a_k^2.$$  

$W^{1,2}_T(\mathbb{R})$ is the set of $T$-periodic functions in $W^{1,2}$. Is the closure of $C^1_T(\mathbb{R})$ under the norm of $W^{1,2}([0, T])$.

$W^{1,2}_0([a, b])$ is the set of $W^{1,2}$ functions vanishing at $a$ and $b$, is the closure of $C^1_0([a, b])$ under the $W^{1,2}([a, b])$ norm.

**Lemma 1. (Sobolev Embedding Theorem)** If $u \in W^{1,2}([a, b])$ then $u \in C([a, b])$, and

$$\|u\|_{C([a, b])} \equiv \|u\|_{L^\infty([a, b])} \leq \alpha \|u\|_{W^{1,2}([a, b])},$$

for some positive constant $\alpha$.

**Proof:**

Suppose that $u \in C^1([a, b])$ and consider $a \leq y < x \leq b$. Then

$$u(x) - u(y) = \int_y^x u'(t) \, dt,$$
which implies that

\[ |u(x) - u(y)| \leq \int_y^x |u'(t)| \, dt \leq |b - a|^{1/2} \left( \int_a^b |u'(t)|^2 \, dt \right)^{1/2} \]

On the other hand, since \(|u(x) - u(y)| \geq ||u(x)| - |u(y)||

\[ |u(x)| \leq |u(y)| + |b - a|^{1/2} \left( \int_a^b |u'(t)|^2 \, dt \right)^{1/2} \]

Integrating in \(y\) over \([a, b]\),

\[ |u(x)| \leq \frac{1}{|b - a|} \int_a^b |u(y)| \, dy + \frac{1}{|b - a|^{1/2}} \left( \int_a^b |u'(t)|^2 \, dt \right)^{1/2} \leq \frac{1}{|b - a|} (b - a)^{1/2} \left( \int_a^b |u(t)|^2 \, dt \right)^{1/2} + \frac{1}{|b - a|^{1/2}} \left( \int_a^b |u'(t)|^2 \, dt \right)^{1/2} \]

Taking the supremum over \(x \in [a, b]\) the inequality is proved. Consider now the case \(u \in W^{1,2}([a, b])\). By definition, there exist a sequence \((u_m) \subset C^1([a, b])\) such that

\[ \|u_m - u\|_{W^{1,2}([a, b])} \to 0 \text{ as } m \to \infty \]

Since \(u_m\) is a convergent sequence in \(W^{1,2}([a, b])\), it will be a Cauchy sequence in \(W^{1,2}([a, b])\). By definition \(u_m \in C^1([a, b])\) and as a consequence , for \(m, j \in \mathbb{N}\)

\[ \|u_m - u_j\|_{C([a, b])} \leq \alpha \|u_m - u_j\|_{W^{1,2}([a, b])} \]

Hence \(u_m\) is a Cauchy sequence in \(C([a, b])\). Then \(u_m\) converges uniformly in \(C([a, b])\) to some function \(v \in C([a, b])\). But notice that if \(u_m \to v\) in \(C([a, b])\)

\[ \int_a^b |u_m - v|^2 \, dx \leq \max_{x \in [a, b]} |u_m(x) - v(x)||b - a| \to 0 \text{ as } m \to \infty \]

implying \(u_m \to v\) in \(L^2([a, b])\). By uniqueness of limit \(u \equiv v\). In particular this implies that \(u \in C([a, b])\).

\[ \square \]

**Definition 1.** A function is Holder continuous of order \(\alpha\), if

\[ |u(x) - u(y)| \leq K|x - y|^\alpha, \quad 0 < \alpha < 1. \]

**Corollary 1.** If \(u \in W^{1,2}([a, b])\) then \(u\) is Holder continuous of order \(1/2\).

**Homework 5.** Prove the result.
As a reference we state the general Sobolev Theorem, which is of extreme importance in the applications to PDE’s.

**Theorem 1. (Sobolev Imbedding Theorem)** Let \( \Omega \subset \mathbb{R}^n \) be bounded with \( \partial \Omega \) smooth. Then

\[
 u \in W^{1,2}_0(\Omega) \implies u \in L^{\frac{2n}{n-2}}(\Omega)
\]

for \( n \geq 3 \), and there exists \( c = c(\Omega, n) \) such that

\[
 \|u\|_{L^s(\Omega)} \leq c\|u\|_{W^{1,2}_0(\Omega)}, \quad \forall s \in \left[1, \frac{2n}{n-2}\right]
\]

This inequality implies that

\[
 W^{1,2}_0(\Omega) \subset L^s(\Omega), \quad \forall s \in \left[1, \frac{2n}{n-2}\right]
\]

and this imbedding is compact if \( s < \frac{2n}{n-2} \).

For the proof see

Adams - Sobolev Spaces;

Friedman - PDE.

### 4.1 Continuity Properties of Functionals

Let \( E \) be a real Banach space, and consider \( I : E \to \mathbb{R} \).

**Definition 2.** \( I \) is continuous at \( u \in E \) if

\[
 \forall \epsilon > 0 \exists \delta = \delta(\epsilon, u) : \|v - u\|_E \leq \delta \implies |I(v) - I(u)| \leq \epsilon.
\]

\( I \) is continuous on \( E \), \( I \in C(E, \mathbb{R}) \), if it is continuous on every point on \( E \).

**Definition 3.** \( I \) is sequentially continuous at \( u \in E \) if

\[
 \forall u_m \to u \text{ in } E \quad I(u_m) \to I(u).
\]

**Homework 6.** Prove that \( I \) is continuous at \( u \in E \) iff \( I \) is sequentially continuous at \( u \).

**Example 12.** Consider

\[
 I(u) = \int_a^b F(x, u, u') \, dx
\]

where \( F \) is continuous in its arguments (\( F \in C([a, b] \times \mathbb{R}^2, \mathbb{R}) \)), and \( E = C^1([a, b]) \). For any \( u \in E \), let \( (u_m) \subset E \) such that \( \|u_m - u\|_E \to 0 \) as \( m \to \infty \). Since \( F(x, u_m, u_m') \to F(x, u, u') \) uniformly for \( x \in [a, b] \). Hence \( I \) is sequentially continuous and as a consequence continuous.
Example 13. Consider

\[ I(u) = \int_{a}^{b} F(x, u, u') \, dx \]

defined in \( E = W^{1,2}([a, b]) \). Notice that if \( F \) is just continuous, \( I \) may not be well defined in \( E \). As an example of this, consider \( F(x, u, u') = (u'(x))^4 \), \([a, b] = [0, 1] \) and let \( u(x) = x^{3/4} \). Then \( u \in W^{1,2}([0,1]) \) and yet \( I(u) = \infty \). Therefore we need some further conditions on \( F \). We will suppose that \( F \) verifies the Growth Condition

\[ |F(x, y, z)| \leq \gamma(x, y)z^2 + \beta(x, y) \]

where \( \gamma, \beta \) are continuous in its arguments. This condition means that \( F \) can't grow more than quadratically in \( z \) (since \( u' \in L^2 \) this should be enough for \( I \) to be well defined). Let us start to prove that \( I \) is well defined. For \( u \in W^{1,2}([a, b]) \) let \( \|u\|_{W^{1,2}([a, b])} = M \). By Sobolev Embedding Theorem \( \|u\|_{C([a, b])} \leq \alpha M \), and

\[
|F(x, u, u')| \leq \gamma(x, u(x))|u'(x)|^2 + \beta(x, u(x)) \\
\leq \max_{x \in [a, b], |u| \leq \alpha M} |\gamma(x, u(x))||u'(x)|^2 + \max_{x \in [a, b], |u| \leq \alpha M} |\beta(x, u(x))| \\
\leq M_1|u'(x)|^2 + M_2
\]

for some positive constants \( M_1, M_2 \). Hence

\[ I(u) \leq M_1\|u\|_{W^{1,2}([a, b])} + (b-a)M_2 \]

and thus \( I \) is well defined in \( W^{1,2}([a, b]) \). To prove that \( I \) is continuous we need some more assumptions. The ideal situation is that \( F \in C^2 \) in its arguments, but we don't need so much. We will assume that \( F \in C^1([a, b] \times \mathbb{R}^2, \mathbb{R}) \) and that

\[ |F_y(x, y, z)| \leq \gamma_1(x, y)|z|^2 + \beta_1(x, y) \]

and

\[ |F_z(x, y, z)| \leq \gamma_2(x, y)|z| + \beta_2(x, y) \]

for \( \gamma_i, \beta_i, i = 1,2 \), continuous in its arguments. Notice that the second condition makes sense, since we assumed that \( F \) grows at most quadratically in \( z \), \( F_z \) grows at most linearly in \( z \). Then

\[
I(v) - I(u) = \int_{a}^{b} (F(x, u, u') - F(x, v, v')) \, dx \\
= \int_{a}^{b} \int_{0}^{1} \frac{d}{ds} F(x, u + s(v - u), u' + s(v' - u')) \, ds \, dx \\
= \int_{a}^{b} \int_{0}^{1} \left( F_u(x, u + s(v - u), u' + s(v' - u'))(v-u) + F_z(x, u + s(v - u), u' + s(v' - u'))(v-u') \right) \, ds \, dx
\]

21
Using the growth conditions on the derivatives

\[ |I(v) - I(u)| \leq \int_a^b \int_0^1 \left( \gamma_1(x, u + s(v - u))|u'| + s(v' - u')|^2|v - u| + \beta_1(x, u + s(v - u))|v - u| + \gamma_2(x, u + s(v - u))|u'| + s(v' - u')||v' - u'||v - u| + \beta_1(x, u + s(v - u))|v' - u'| \right) dsdx \]

Let \( M = \max\{\|u\|_{W^{1,2}([a, b])}, \|v\|_{W^{1,2}([a, b])}\} \). Then both \( \|u\|_{W^{1,2}([a, b])} \) and \( \|v\|_{W^{1,2}([a, b])} \) are bounded by \( M \), and by Sobolev inequality, both \( \|u\|_{C([a, b])} \) and \( \|v\|_{C([a, b])} \) are bounded by \( \alpha M \). Notice that we can make a rough estimate

\[ |u + s(v - u)| \leq |u| + s(|u| + |v|) \leq 3\alpha M \]

Thus

\[ \int_a^b \int_0^1 \gamma_1(x, u + s(v - u))|u'| + s(v' - u')|^2|v - u| dsdx \leq \max_{x \in [a, b], |u + s(v - u)| \leq 3\alpha M} \Gamma_1(M) \max_{x \in [a, b]} \Gamma_1(M) \max_{x \in [a, b]} \|v(x) - u(x)\| \int_a^b \int_0^1 |u'| + s(v' - u')|^2 dsdx \]

and similarly

\[ \int_a^b \int_0^1 \beta_1(x, u + s(v - u))|v - u| dsdx \leq B_1(M) \|u - v\|_{C([a, b])} \]

On the other hand using Schwarz inequality

\[ \int_a^b \int_0^1 \gamma_2(x, u + s(v - u))|u'| + s(v' - u')||v' - u'|| |v - u| dsdx \leq \Gamma_2(M) \|u' - v'\|_{L^2([a, b])} \]

and

\[ \int_a^b \int_0^1 \beta_1(x, u + s(v - u))|v' - u'| dsdx \leq B_2(M) \|u' - v'\|_{L^2([a, b])} \]

Finally, it is easy to check that

\[ |I(v) - I(u)| \leq K(M) \|u - v\|_{W^{1,2}([a, b])} \]

and this proves that \( I \) is continuous on \( E \).

**Homework 7.** Prove that \( I(u) = \int_a^b G(x, u(x))dx \) defined on \( E = W^{1,2}([a, b]) \), is continuous, where \( G \in C([a, b] \times \mathbb{R}, \mathbb{R}) \).
We will now review some of the most important results in Functional Analysis. The next proposition states that the distance to a closed, convex subset of a Hilbert space is achieved.

**Proposition 1.** Let \( K \) be a closed, convex set contained in \( E \). Then for all points \( u \in E \), there is a unique \( v \in K \) such that
\[
\| u - v \| = \inf_{w \in K} \| u - w \|.
\]

**Proof:**

Let \( d = \inf_{w \in K} \| u - w \| \). By definition of infimum there exists \((v_m) \subset K\) such that \( \| u - v_m \| \downarrow d \) as \( m \to \infty \). By the parallelogram law
\[
2(\| u - v_m \|^2 + \| u - v_n \|^2) = \| v_m - v_n \|^2 + \| 2u - (v_n + v_m) \|^2 = \| v_m - v_n \|^2 + 4\| u - \frac{v_n + v_m}{2} \|^2
\]
\( K \) is convex and therefore \( \frac{v_n + v_m}{2} \in K \), since is the midpoint of the segment line joining \( v_m \) and \( v_n \). Moreover \( d \) is the infimum, hence
\[
\| u - \frac{v_n + v_m}{2} \|^2 \geq d^2
\]
Thus
\[
4d^2 \geq \| v_m - v_n \|^2 + 4d^2
\]
which implies that \( \| v_m - v_n \| \leq 0 \). Then \( (v_m) \) is a Cauchy sequence, and since \( K \) is closed \( v_m \to v \) as \( m \to \infty \) for some \( v \in K \). To prove the uniqueness, let’s suppose that there are two such elements \( v, w \). Take the sequence \( v, w, v, w, \ldots \). By definition this is a minimizing sequence. But we just proved that any minimizing sequence is a Cauchy sequence. Henceforth it must have a unique limit, which is a contradiction unless \( v = w \).

\[ \diamond \]

**Theorem 2.** (Projection Theorem) Let \( F \) be a closed linear subspace of \( E \). Then, for all \( u \in E \) there exists unique \( v \in F \) and \( w \in F^\perp \) such that \( u = v + w \).

**Proof:**

Since a closed subspace is a closed convex set, by the previous proposition, there exists \( v \in F \) such that
\[
\| u - v \| = \inf_{w \in F} \| u - w \|
\]
Choose any \( f \in F \). By definition of \( v \) and the fact that \( v + \alpha f \in F \) for all \( \alpha \in \mathbb{R} \)
\[
\| u - v \|^2 \leq \| u - v - \alpha f \|^2 = \| u - v \|^2 - 2\alpha(u - v, f) + \alpha^2\| f \|^2.
\]
If \( \alpha \in \mathbb{R}^+ \),
\[
(u - v, f) \leq \frac{\alpha}{2}\| f \|^2, \quad \forall f \in F, \; \alpha \in \mathbb{R}^+
\]

23
and letting $\alpha \to 0$

$$(u - v, f) \leq 0, \quad \forall f \in F$$

On the other hand, for $\alpha \in \mathbb{R}^-$ converging to 0, we will obtain

$$(u - v, f) \geq 0, \quad \forall f \in F$$

Then $(u - v, f) = 0$ for all $f \in F$, and therefore $u - v \in F^\perp$. Since $u = v + (u - v)$, the result is proved. To show uniqueness, suppose that $u = v_1 + w_1 = v_2 + w_2$ with $v_1, v_2 \in F$ and $w_1, w_2 \in F^\perp$. Then $v_1 - v_2 = w_1 - w_2$ and since $F \cap F^\perp = \{0\}$ we must necessarily have $v_1 = v_2$ and $w_1 = w_2$.

◊

**Corollary 2.** If $F$ is a closed subspace of $E$ and $F \neq E$, there exists $w \in F^\perp \setminus \{0\}$.

**Proof:**
By hypothesis $F \neq E$, and therefore $\exists u \in E \setminus F$. By the Projection Theorem

$$u = v + w, \quad v \in F, \quad w \in F^\perp$$

and $w \neq 0$ otherwise $u = v \in F$.

◊

**Definition 4.** Let $E$ be a real Banach space. $l : E \to \mathbb{R}$ is a linear functional on $E$, if

$$\forall \alpha, \beta \in \mathbb{R} \quad \forall u, v \in E \quad l(\alpha u + \beta v) = \alpha l(u) + \beta l(v)$$

$l$ is bounded if

$$\exists M \in \mathbb{R}^+ : \quad |l(\varphi)| \leq M\|\varphi\|_E \quad \forall \varphi \in E$$

The class

$$E' = \{\text{bounded linear functional over } E\}$$

is known as the dual space of $E$. We define

$$\|l\| = \inf_{E} \frac{|l(\varphi)|}{\|\varphi\|} = \sup_{\|\varphi\|=1} |l(\varphi)|$$

which is a norm in $E'$. Moreover, $E'$ is a Banach space relatively to $\|\cdot\|_{E'}$.

Notice that as a consequence of the definition, if a linear functional $l$ is bounded, then

$$|l(\varphi)| \leq \|l\|_{E'} \|\varphi\|_E, \quad \forall \varphi \in E$$

**Homework 8.** If $l$ is a bounded linear functional on $E$, prove that $l$ is bounded iff $l$ is continuous.
Consider the set of all bounded linear functionals on $E$. In the case of $E$ being a Hilbert space its dual is known.

**Theorem 3. (Riesz Representation Theorem)** If $E$ is a real Hilbert space and $l \in E'$, there exists a unique $u \in E$ such that

$$l(e) = (u, e)_E, \quad \forall e \in E.$$  

Moreover

$$\|l\|_{E'} = \|u\|_E.$$

**Proof:**

Let $N$ be the null space of $l$, i.e.

$$N = \{ n \in E : l(n) = 0 \}.$$  

It is known that $N$ is a closed linear subspace of $E$. If $E = N$ then $u = 0$ and we are done. Otherwise, we are in the conditions of Corollary 2, and there exists $v \in N^\perp \setminus \{0\}$, and in particular $l(v)$ cannot be 0. Choose any $e \in E$. By linearity

$$l(e - l(e) \frac{v}{l(v)} e) = l(e) - l(e) = 0$$

which implies that $e - l(e) \frac{v}{l(v)} e \in N$. Then

$$0 = (v, e - l(e) \frac{v}{l(v)} e) = (v, e) - l(e) \frac{v}{l(v)} \|v\|^2$$

implying that

$$l(e) = \frac{(e, v)}{\|v\|^2} l(v) = \frac{l(v) v}{\|v\|^2}, \quad \forall e \in E.$$  

Since $\frac{l(v) v}{\|v\|^2} \in E$ the existence result is proved. To prove the uniqueness, suppose that for $u \neq w$

$$l(e) = (u, e) = (w, e), \quad \forall e \in E.$$  

Then

$$\|u - w\|^2 = (u - w, u - w) = (u, u - w) - (w, u - w) = l(u - w) - l(u - w) = 0$$

which implies that $u = w$. To prove the norm equality, notice that

$$\|u\|_E^2 = |(u, u)_E| = |l(u)| \leq \|l\|_{E'} \|u\|_E,$$

and therefore

$$\|u\|_E \leq \|l\|_{E'}. \quad (11)$$

25
On the other hand, by definition
\[ \|l\|_{E'} = \sup_{\|\varphi\|_E = 1} |l(\varphi)| \]
which means that
\[ \forall \epsilon > 0 \exists e \in E : \|e\| = 1 \text{ and } |l(e)| \geq \|l\|_{E'} - \epsilon. \]
But also
\[ |l(e)| = (u, e) \leq \|u\|_E \|e\|_E = \|u\|_E \]
Then
\[ \|u\|_E \geq |l(e)| \geq \|l\|_{E'} - \epsilon. \]
Letting \( \epsilon \to 0 \)
\[ \|u\|_E \geq \|l\|_{E'}. \] (12)
(11) and (12) imply \( \|u\|_E = \|l\|_{E'} \).

\[ \boxed{} \]

4.2 Weak Convergence of Sequences

Definition 5. If \( E \) is a real Banach space, \( u_m \) converges weakly to \( u \), \( u_m \rightharpoonup u \) in \( E \), if
\[ l(u_m) \to l(u) \text{ for all continuous linear functional } l \text{ on } E. \]

In the case of a Hilbert space, by Riesz Representation Theorem
\[ u_m \rightharpoonup u \text{ if } \forall \varphi \in E : (\varphi, u_m) \to (\varphi, u) \text{ as } m \to \infty. \]

In the particular case \( E = L^2([a, b]) \)
\[ u_m \rightharpoonup u \text{ if } \forall \varphi \in L^2([a, b]) : \int_a^b \varphi(x)u_m(x) \, dx \to \int_a^b \varphi(x)u(x) \, dx \text{ as } m \to \infty. \]

Example 14. In \( E = L^2([0, 2\pi]) \) take \( u_m(x) = \cos(mx) \). We will determine the weak limit of \( u_m \) in \( E \). Since any \( \varphi \in E \) can be written as a Fourier series,
\[ \varphi(x) = \sum_{i \in \mathbb{N}} a_i \cos(ix) + b_i \sin(ix) \]
then
\[ < \varphi, u_m > = \int_0^{2\pi} \varphi(x)u_m(x) \, dx = \int_0^{2\pi} a_m(\cos(mx))^2 \, dx = \pi a_m \]
By Parseval’s identity
\[ \infty > \|\varphi\|^2_E = \sum_{i \in \mathbb{N}} (a_i^2 + b_i^2) \]
implying that both \( a_i, b_i \to 0 \) as \( i \to \infty \). In particular, this implies that \( < \varphi, u_m > \to 0 \) as \( m \to \infty \), and therefore \( u_m \rightharpoonup 0 \) as \( m \to \infty \). We remark that \( u_m \not\to 0 \) as \( m \to \infty \).
**Theorem 4.** Let $E$ be a Hilbert space, and $(u_m) \subset E$ be bounded. Then $(u_m)$ has a weakly convergent subsequence, i.e., there exists $u \in E$ such that $(u_m, \varphi) \to (u, \varphi)$ for all $\varphi \in E$.

**Proof:**

Suppose we can find a subsequence $(u_{m_i})$ of $(u_m)$ such that $(u_{m_i}, \varphi)$ is a Cauchy sequence for all $\varphi \in E$. Then, we can define

$$l(\varphi) = \lim_{i \to \infty} (u_{m_i}, \varphi).$$

Since $(u_m)$ is bounded in $E$, there exists $M > 0$ such that $\|u_m\|_E \leq M$ for all $m \in M$, and as a consequence

$$|(u_m, \varphi)| \leq \|u_m\| \|\varphi\| \leq M \|\varphi\|$$

Then

$$|l(\varphi)| \leq M \|\varphi\|,$$

and therefore $l \in E'$. By the Riesz Representation Theorem, there exists $u \in E$ such that

$$l(\varphi) = (u, \varphi), \quad \forall \varphi \in E$$

By construction

$$\lim_{i \to \infty} (u_{m_i}, \varphi) = l(\varphi) = (u, \varphi), \quad \forall \varphi \in E$$

as wanted. It remains to show that there exists such subsequence $(u_{m_i})$. We can simplify the problem a little further. If

$$F = \overline{\text{span}}(u_m),$$

we know that $F$ is a closed linear subspace of $E$. By the Projection Theorem

$$\forall \varphi \in E, \varphi = f + g \text{ for } f \in F \text{ and } g \in F^\perp.$$ 

Therefore

$$(u_m, \varphi) = (u_m, f) + (u_m, g) = (u_m, f)$$

since $(u_m) \subset F$ and $g \in F^\perp$. So it’s enough to prove that $(u_m, \varphi)$ is a Cauchy sequence for all $\varphi \in F$. Consider $(u_m, u_1)$. By assumption

$$|(u_m, u_1)| \leq \|u_m\| \|u_1\| \leq M^2.$$ 

Hence $|(u_m, u_1)|$ is a bounded set of $\mathbb{R}$ which implies that we can extract a convergent subsequence, which we will denote as

$$(u_{1m}, u_1)$$

Next, consider $(u_{1m}, u_2)$. Once again this is a bounded sequence in $\mathbb{R}$ and therefore admits a convergent subsequence

$$(u_{2m}, u_2)$$
Proceeding inductively we construct the convergent sequence \((u_{km}, u_k)\). With a diagonalization process we can take \(u_{mj} = u_{jj}\). To show that \((u_{jj}, \varphi)\) is a Cauchy sequence, for all \(\varphi \in F\), notice that the result is obvious for any \(\varphi\) which is a linear combination of \(u_j\)'s, since by construction \((u_{jj}, \varphi)\) is convergent for any such \(\varphi\). It remains to show that the same thing holds for an element on the closure. Take any \(\psi \in F\) and any \(\epsilon > 0\), we can find a finite linear combination of elements of \((u_m)\), \(\epsilon\) close to \(\psi\), i.e., we can find

\[
f = \sum_{i=1}^{k} \alpha_i u_i : \|f - \psi\| \leq \epsilon.
\]

Then

\[
|(u_{ii} - u_{jj}, \psi)| \leq (u_{ii} - u_{jj}, \psi - f) + (u_{ii} - u_{jj}, f).
\]

On one hand

\[
|(u_{ii} - u_{jj}, \psi - f)| \leq \|u_{ii} - u_{jj}\|\|\psi - f\| \leq 2M\epsilon,
\]

on the other hand, since \(f \in F\) as we remarked earlier, by construction

\[(u_{ii} - u_{jj}, f) \to 0 \quad \text{as} \quad i, j \to \infty.
\]

Hence, for \(i, j\) sufficiently large \(|(u_{ii} - u_{jj}, \psi)|\) is as small as we want.

\[\diamondsuit\]

### 4.3 Lower Semicontinuity of Functionals

**Definition 6.** \(I : E \to \mathbb{R}\) is lower semicontinuous at \(x\) if for all sequences \(x_m \to x\) on \(E\)

\[I(x) \leq \lim \inf_{m \to \infty} I(x_m).\]

**Example 15.** Let \(I : \mathbb{R} \to \mathbb{R}\) defined by

\[I(x) = \begin{cases} 
|x| + 1 & \text{if} \quad x \neq 0 \\
0 & \text{if} \quad x = 0
\end{cases}
\]

The definition of lower semicontinuity allows the function to jump down: note that \(I(0) = 0 \leq \lim_{x \to 0} I(x) = 1\).

**Definition 7.** If \(I : E \to \mathbb{R}\), with \(E\) a Banach space, then \(I\) is weakly lower semi-continuous at \(u \in E\), if for all sequences \(u_m \to u\)

\[I(u) \leq \lim \inf_{m \to \infty} I(u_m).
\]
Example 16. Let \( I(u) = \|u\|^2 \) defined on a Hilbert space \( E \). Recall that in Example 14, we proved that a sequence of unit vectors, converges weakly to 0, so \( I \) can’t be continuous in a reasonable way. For \( u_m \rightharpoonup u \)

\[
0 \leq \|u_m - u\|^2 = \|u_m\|^2 - 2(u_m, u) + \|u\|^2,
\]

which implies that

\[
2(u_m, u) - \|u\|^2 \leq \|u_m\|^2. \tag{13}
\]

Choose a subsequence \((u_{m_i})\)

verifying

\[
\lim_{i \to \infty} \|u_{m_i}\| = \liminf_{m \to \infty} \|u_m\|.
\]

By (13) applied to \((u_{m_i})\)

\[
2(u_{m_i}, u) - \|u\|^2 \leq \|u_{m_i}\|^2.
\]

Letting \( i \to \infty \)

\[
\|u\|^2 \leq \liminf_{m \to \infty} \|u_m\|^2,
\]

and therefore \( \| \cdot \|^2 \) is weakly lower semicontinuous.

Definition 8. If \( E \) is a Banach space, \( I: E \to \mathbb{R} \) is said to be weakly sequentially continuous at \( u \in E \), if whenever \( u_m \rightharpoonup u \), \( I(u_m) \to I(u) \).

Notice that functionals which are weakly sequentially continuous have some kind of compactness properties.

Remarks:

1 - If \( I \) is weakly sequentially continuous then \( I \) is weakly lower semicontinuous.

2 - If \( I_1 \) is weakly lower semicontinuous and \( I_2 \) is weakly sequentially continuous, then \( I_1 + I_2 \) is weakly lower semicontinuous.

Example 17. Let \( E = W^{1,2}([a, b]) \) and

\[
I_2(u) = \int_a^b g(x, u) \, dx
\]

where \( g \in C([a, b] \times \mathbb{R}, \mathbb{R}) \). First notice that \( I_2 \) is well defined in \( E \) (if \( u \in E \), then \( u \) is continuous and therefore \( g(\cdot, u) \) is continuous). We claim that \( I_2 \) is weakly sequentially continuous. Let \( u_m \rightharpoonup u \) on \( E \). By definition

\[
(u_m, \varphi) \to (u, \varphi), \quad \forall \varphi \in W^{1,2}([a, b])
\]

Then \((u_m, \varphi)\) is a convergent sequence in \( \mathbb{R} \) and therefore is bounded.
Theorem 5. Banach-Steinhaus Theorem Suppose that $E$ is a Hilbert space, and consider $S \subset E$. If
\[ \sup_{x \in S} |(x, \varphi)| \leq A \varphi, \quad \forall \varphi \in E \]
then $S$ is a bounded subset of $E$.

By direct application of the Theorem, we can conclude that $(u_m)$ is bounded in $E$. Hence, there exists $M \geq 0$ such that
\[ \|u_m\|_E \leq M, \quad \forall m \in \mathbb{N} \]
By the Sobolev inequality
\[ \|u_m\|_{L^\infty} \leq \alpha M \]
and therefore $(u_m)$ is bounded in $L^\infty([a, b])$. As we remarked before, if $u \in W^{1,2}([a, b])$ then $u$ is Holder continuous of order $1/2$. In particular, for all $m \in \mathbb{N}$ and $x, y \in [a, b]$
\[ \frac{|u_m(x) - u_m(y)|}{|x - y|^{1/2}} \leq \|u_m'\|_{L^2([a, b])}^2 \leq M^2 \]
and therefore $(u_m)$ is an equicontinuous family of functions.

Theorem 6. (Ascoli-Árzella Theorem) If $F \subset C(K, \mathbb{R})$, where $K \subset \mathbb{R}^n$ is compact, and $F$ is a uniformly bounded and equicontinuous set, then $F$ is pre-compact (any bounded sequence has a convergent subsequence).

Then, along some subsequence, there exists $v \in C([a, b])$ such that $u_m, \rightarrow v$ uniformly, i.e.,
\[ \max_{x \in [a, b]} |u_m(x) - v(x)| \rightarrow 0 \quad \text{as} \quad i \rightarrow \infty \]
This implies
\[ \int_a^b u_m(x)\varphi(x)dx \rightarrow \int_a^b v(x)\varphi(x)dx, \quad \forall \varphi \in E \]
On the other hand
\[ |\int_a^b u_m(x)\varphi(x)dx| \leq \|u_m\|_{L^2} \|\varphi\|_{L^2} \leq \|u_m\|_E \|\varphi\|_E \leq M \|\varphi\|_E \]
and therefore we can conclude that $\int_a^b u_m(x)\varphi(x)dx$ is a bounded linear functional on $E$. Since $u_m \rightarrow u$,
\[ \int_a^b u_m(x)\varphi(x)dx \rightarrow \int_a^b u(x)\varphi(x)dx, \quad \forall \varphi \in E \]
By uniqueness of limit
\[ \int_a^b u(x)\varphi(x)dx \rightarrow \int_a^b v(x)\varphi(x)dx, \quad \forall \varphi \in E \]
Since $C^1([a,b])$ is dense in $E$, we will also have
\[ \int_a^b (u - v)(x)\varphi(x)dx = 0, \quad \forall \varphi \in C^1([a,b]) \]
which as we have seen before, this implies that $u = v$. Hence $u_m \to u$ in $L^\infty([a,b])$, as $i \to \infty$. The next thing to prove is that all the sequence converges. If not, we could find another subsequence $(\tilde{u}_m)$ such that $\|\tilde{u}_m - u\| \geq \epsilon$ for some positive $\epsilon$. But as above $(\tilde{u}_m)$ is a uniformly bounded sequence and equicontinuous. Ascoli-Arzella Theorem implies the existence of $\tilde{v} \in C([a,b])$ such that $\tilde{u}_m \to \tilde{v}$ in $L^\infty([a,b])$, as $i \to \infty$. Using the same arguments we can prove that $u = v$. Then
\[ u_m \to u \quad \text{in} \quad E \quad \Rightarrow \quad u_m \to u \quad \text{in} \quad L^\infty([a,b]) \]
which implies that $I_2(u_m) \to I_2(u)$ as we wanted.

**Example 18.** Let $E = W^{1,2}([a,b], \mathbb{R})$ and
\[ I(u) = \int_a^b \frac{1}{2} |u'|^2 dx \]
We want to prove that $I$ is weakly lower semicontinuous on $E$. First notice that we can write $I$ in the form
\[ I(u) = \frac{1}{2} \int_a^b |u|^2 + |u'|^2 \, dx - \frac{1}{2} \int_a^b |u'| \, dx \equiv I_1(u) + I_2(u) \]
In Example 16 we proved that $I_1$ is weakly lower semicontinuous, and in Example 17 we proved that $I_2$ is weakly sequentially continuous. Hence $I$ is weakly lower semicontinuous.

### 4.4 Minima of Functionals

**Theorem 7.** Let $E$ be a Hilbert space, $K \subset E$ bounded and closed under weak convergent sequences. Let $I : E \to \mathbb{R}$ be a weakly lower semicontinuous functional. Then there exists $v \in K$ such that
\[ I(v) = \inf_K I > -\infty \]

**Proof:**

Choose $(u_m) \subset K$ a minimizing sequence, i.e., $I(u_m) \searrow \inf I$. Since $K$ is a bounded set, $(u_m)$ is bounded in $E$. Then there exists a subsequence $u_{m_i} \subset K$ such that $u_{m_i} \to v$, and $v \in K$ (since $K$ is closed). By the weak lower semicontinuity of $I$
\[ I(v) \leq \liminf_{i \to \infty} I(u_{m_i}) = \inf_K I \]
Then $I(v) = \inf_K I > -\infty$.  

\[ \diamond \]
Corollary 3. If $E$ is a Hilbert space, $I : E \rightarrow \mathbb{R}$ weakly lower semicontinuous and $I(u) \rightarrow \infty$ as $\|u\| \rightarrow \infty$, then $I$ attains its infimum over $E$, i.e.,

$$\exists v \in E : I(v) = \inf_E I$$

Proof:

Choose $z \in E$ such that $I(z) < \infty$. By the coercivity hypothesis, for $\|u\| > R$, $I(u) > I(z)$ for $R$ sufficiently large. Therefore

$$\inf_E I = \inf_{\|u\| \leq R} I(u)$$

But then we are in the conditions of Theorem 7 and the result follows.

\[\Box\]

Corollary 4. If $I$ is weakly sequentially continuous on $K$, where $K$ is as in the Theorem 7, then $I$ achieves its infimum and supremum on $K$.

Proof:

Since $I$ is weakly sequentially continuous, both $I$ and $-I$ are weakly lower semicontinuous. By Theorem 7, $I$ and $-I$ achieve the infimum on $K$. Moreover $-\inf_k (-I) = \sup_K I$.

\[\Box\]

Proposition 2. Let $I : K \rightarrow \mathbb{R}$ with $I$ weakly sequentially continuous, $K \subset E$ bounded, and $E$ is an infinite dimensional Banach space. Then

$$\overline{I(\partial K)} \supset I(\text{int } K)$$

Proof:

Since $E$ is an infinite dimensional Banach space, we can consider a sequence of unit vectors $(e_n)_{n \in \mathbb{N}}$ verifying $\|e_m\| = 1$ and $e_m \rightharpoonup 0$ as $m \rightarrow \infty$. Let $u \in \text{int } K$ (then $I(u) \subset I(\text{int } K)$), and choose $t_m$ such that

$$u_m = u + t_m e_m \in \partial K \quad (14)$$

Notice that $|t_m| = \|u_m - u\|$, and since $K$ is bounded, $t_m$ is a bounded sequence of real numbers. Letting $m \rightarrow \infty$ in (14), by construction $e_m \rightharpoonup 0$ and $t_m$ is bounded. Then the right hand side

$$u + t_m e_m \rightharpoonup u$$

and as a consequence $u_m \rightharpoonup u$. Since $I$ is weakly sequentially continuous, $I(u_m) \rightarrow I(u)$. But notice that $I(u_m) \in I(\partial K)$. Thus

$$I(u) \in \overline{I(\partial K)}$$

as wanted.

\[\Box\]

As an immediate consequence we can state the following result.
Corollary 5. (Maximum/Minimum Modulus Theorem) In the same hypothesis of Proposition 2

\[
\sup_K I = \sup_{\partial K} I, \quad \inf_K I = \inf_{\partial K} I
\]

Example 19. Let \( V \in C^2([a, b] \times \mathbb{R}^n, \mathbb{R}) \), and consider the corresponding Hamiltonian system 

\( (HS) \)

\[
\ddot{q} + V_q(t, q) = 0
\]

As we have seen before, \((HS)\) is the Euler equation for the minimization of the functional

\[
I(q) = \int_a^b L(q) \, dt
\]

where \( L(q) = \frac{1}{2} |\dot{q}|^2 - V(t, q) \) is the associated Lagrangian.

As a simple Special Case, consider \( n = 1 \) and \( V(q) = 1 - \cos q \). Then the Euler equation become

\[
\ddot{q} + \sin q = 0, \quad q(a) = \alpha, \quad q(b) = \beta
\]

which is the equation of motion of a simple pendulum.

To prove that \( I \) has a minimum value, we can’t work in such a generality. So we further assume that

\[
V(t, x) \leq 0, \quad \forall t \in [a, b] \text{ and } x \in \mathbb{R}^n
\]

\( (V1) \) implies \( L(q) \geq 0 \) (and therefore \( I \) is bounded by below), and \( I(q) \) is well defined for \( q \in E = W^{1,2}([a, b], \mathbb{R}^n) \). By the Sobolev Theorem we can assume that \( q \) is continuous and thus it makes sense to consider the boundary conditions in \( E \).

Consider the problem of minimizing \( I \) over

\[
\Gamma = \{ q \in W^{1,2}([a, b], \mathbb{R}^n) : q \text{ satisfies BC} \}
\]

and let

\[
c = \inf_{\Gamma} I(q)
\]

As an immediate consequence of \( (V1) \), \( c \geq 0 \). Let \( (q_m) \subset \Gamma \) be a minimizing sequence, i.e., \( I(q_m) \searrow c \) as \( m \to \infty \). Then \( I(q_m) \) is a bounded sequence in \( \mathbb{R} \) and therefore we can assume that \( \exists M \geq 0 \) such that \( I(q_m) \leq M \). Using \( (V1) \)

\[
M \geq I(q_m) \geq \int_a^b \frac{1}{2} |\dot{q}_m|^2 dt = \frac{1}{2} \|q_m\|_{L^2([a, b])}^2
\]

and therefore \( (\dot{q}_m) \) is a bounded sequence in \( L^2([a, b]) \). On the other hand

\[
|q_m(t)| \leq |q_m(a)| + \int_a^b |\dot{q}_m| dt
\]

\[
\leq \alpha + |b - a|^{1/2} \|\dot{q}_m\|_{L^2([a, b])}
\]

\[
\leq \alpha + |b - a|^{1/2} \sqrt{2M} \equiv M_1
\]

33
Taking the supremum over \( t \in [a, b] \), we obtain \( \|q_m\|_{L^\infty} \leq M_1 \) and as a consequence

\[
\|q_m\|_{L^2([a, b])}^2 = \int_a^b |q_m|^2 dt \leq \|q_m\|_{L^\infty}^2 |b - a| = M_1^2 |b - a| \tag{16}
\]

(15) and (16) imply that \((q_m)\) is bounded in \( W^{1,2}([a, b]) \). Then, there exists \( v \in E \) such that \( q_m \rightharpoonup v \) in \( E \) (possibly along a subsequence that we will denote again by \( q_m \)). Also, notice that

\[
I(q) = \frac{1}{2} \int_a^b (|\dot{q}|^2 + |q|^2) dt - \int_a^b \left( \frac{1}{2} |\dot{q}|^2 + V(t, q) \right) dt = \frac{1}{2} \|q\|_{W^{1,2}}^2 - I_2(q)
\]

and as we have seen before \( \| \cdot \| \) is weakly lower semicontinuous and \( I_2 \) is weakly continuous (since \( \frac{1}{2}|\dot{q}|^2 + V(t, q) \) is continuous). Hence \( I \) is weakly lower semicontinuous, which implies that

\[
I(v) \leq \liminf I(q_m) = c \tag{17}
\]

Next we will prove that \( v \in \Gamma \). By its definition \( v \in E \), so all we have to show is that it satisfies the boundary conditions. But note that since \( q_m \) converges weakly to \( v \) in \( E \), by Sobolev Theorem \( q_m \rightarrow v \) uniformly and therefore \( v(a) = \lim q_m(a) = \alpha \) and \( v(b) = \lim q_m(b) = \beta \). Thus \( v \in \Gamma \) and as a consequence

\[
I(v) \geq c \tag{18}
\]

(17) and (18) imply \( I(v) = c \). We still have to show that \( v \in C^2([a, b]) \) and that it satisfies (HS). We will do this later.

**Example 20.** In Example 19, we used the condition \( V(t, x) \leq 0 \) to obtain estimates on a minimizing sequence in \( E \). In this example we will assume a growth condition instead

\[
|V(t, x)| \leq A + B|x|^{\gamma}, \quad \text{with } 0 < \gamma < 2 \tag{V1}
\]

where \( A \) and \( B \) are positive real numbers. Condition \((V1)\) forces \( V \) not to grow fast, in fact it forces \( V \) to grow less than quadratically. Let \((q_m)\) be a minimizing sequence, our goal is to obtain estimates for \((q_m)\) in \( E \). Once we do that, the rest follows as in Example 19. Since \( I(q_m) \) converges, there exists \( M \geq 0 \) such that \( I(q_m) \leq M \) for all \( m \in \mathbb{N} \). Then

\[
\int_a^b \frac{1}{2} |\dot{q_m}|^2 dt \leq M + \int_a^b |V(t, q_m)| \, dt
\]

and by \((V1)\)

\[
\frac{1}{2} \|q_m\|_{L^2([a, b])}^2 \leq M + (b - a)A + B \int_a^b |q_m|^\gamma dt
\]

As noticed, our goal is to get an estimate on \( \int_a^b |q_m|^\gamma dt \) in terms of \( \|q_m\|_{L^2} \). Since \( q_m \) is Holder continuous of order \( 1/2 \), for \( t \in [a, b] \)

\[
|q(t)| \leq |q(a)| + (t - a)^{1/2} \left( \int_a^t |\dot{q}|^2 dt \right)^{1/2}
\]
and therefore
\[ \frac{1}{2} \| \dot{q}_m \|_{L^2([a,b])}^2 \leq M_1 + B \int_a^b \left( \alpha + (t - a)^{1/2} \left( \int_a^t |q_m|^2 \, dt \right)^{1/2} \right)^{\gamma/2} \, dt \] \hspace{1cm} (19)

Notice that if 
\[ z > w \text{ then } (z + w)^\gamma \leq (2z)^\gamma \]
and if
\[ z < w \text{ then } (z + w)^\gamma \leq (2w)^\gamma \]
In either case
\[ (z + w)^\gamma \leq (2z)^\gamma + (2w)^\gamma \]
Using this fact in (19)
\[ \frac{1}{2} \| \dot{q}_m \|_{L^2([a,b])}^2 \leq M_1 + 2 \gamma B \int_a^b \left( \alpha^{\gamma/2} + (t - a)^{\gamma/2} \left( \int_a^b |\dot{q}_m|^2 \, dt \right)^{\gamma/2} \right) \, dt \]
\[ \equiv M_2 + 2 \gamma B \int_a^b (t - a)^{\gamma/2} \left( \int_a^b |\dot{q}_m|^2 \, dt \right)^{\gamma/2} \, dt \]
\[ \leq M_2 + M_3 \left( \int_a^b |\dot{q}_m|^2 \, dt \right)^{\gamma/2} \]

Lemma 2. (Young’s Inequality) If \( x, y \in \mathbb{R}^+ \) and \( \sigma, \tau \geq 1 \) satisfying \( \frac{1}{\sigma} + \frac{1}{\tau} = 1 \), then
\[ xy \leq \frac{x^\sigma}{\sigma} + \frac{y^\tau}{\tau} \]

Homework 9. Prove the result. Hint: Use the sign of the second derivative of the logarithm function.

Choosing \( x = \frac{M_3}{\epsilon} \) and \( y = \epsilon \left( \int_a^b |\dot{q}_m|^2 \, dt \right)^{\gamma/2} \), the inequality implies
\[ M_3 \left( \int_a^b |\dot{q}_m|^2 \, dt \right)^{\gamma/2} \leq \frac{M_3^\gamma}{\sigma} + \epsilon \frac{\gamma}{2} \left( \int_a^b |\dot{q}_m|^2 \, dt \right)^{\gamma/2} \]
Take \( \tau \frac{\gamma}{2} = 1 \) (notice that since \( \gamma < 2 \), \( \tau > 1 \)), and \( \sigma \) such that the inequality holds. Then
\[ M_3 \left( \int_a^b |\dot{q}_m|^2 \, dt \right)^{\gamma/2} \leq \frac{M_3^\gamma}{\sigma} + \frac{\gamma}{2} \epsilon^{2/\gamma} \| \dot{q}_m \|_{L^2([a,b])}^2 \]
Finally
\[ \frac{1}{2} \| \dot{q}_m \|_{L^2([a,b])}^2 \leq M_2 + \frac{M_3^\gamma}{\sigma} + \frac{\gamma}{2} \epsilon^{2/\gamma} \| \dot{q}_m \|_{L^2([a,b])}^2 \]
Choose \( \epsilon \) so small so that \( \epsilon^\gamma \frac{\gamma}{2} = \frac{1}{4} \), and conclude
\[ \frac{1}{2} \| \dot{q}_m \|_{L^2([a,b])}^2 \leq M_4 \]
and from here we proceed as in Example 19.
4.5 Differentiability in Banach Spaces

Let $E$, $F$ be real Banach spaces and $L : E \to F$ a linear map.

**Example 21.** 1. Let $E = F = \mathbb{R}^n$ and $Lu = Au$ where $A$ is a $(n \times n)$ matrix.

2. Let $E = C^1([a,b])$, $F = C([a,b])$ and $Lu = u'$.

3. Let $E = C^2([a,b])$, $F = C([a,b])$ and $Lu = u'' + a(x)u$ where $a$ is a given continuous function.

**Definition 9.** $L$ is a bounded map if there exists $M > 0$ such that

\[
\sup_{x \neq 0, x \in E} \frac{\|Lx\|_F}{\|x\|_E} \leq M \quad \text{or} \quad \sup_{\|x\|=1} \|Lx\|_F \leq M
\]

We can easily prove, like in the case of linear functionals, that a linear map is bounded iff it is continuous.

Define

\[ \mathcal{L}(E,F) = \{ \text{Continuous Linear maps} \} \]

which is a Banach space under the norm

\[ \|L\|_{\mathcal{L}(E,F)} = \sup_{\|x\|=1} \|Lx\|_F \]

**Definition 10.** Let $\mathcal{O} \subset E$ be open and suppose $T : \mathcal{O} \to F$. Then $T$ is Fréchet Differentiable at $u \in \mathcal{O}$ if there exists $L = L(u) \in \mathcal{L}(E,F)$ such that

\[ \forall \epsilon > 0 \ \exists \delta = \delta(\epsilon, u) : \|v\|_E \leq \delta \Rightarrow \|T(u + v) - T(u) - L(v)\|_F \leq \epsilon \|v\|_E \]

The linear map $L = L(u)$, is called the Fréchet derivative of $T$ at $u$, denoted by $T'(u)$.

**Example 22.** 1. Let $E = \mathbb{R}^n$, $F = \mathbb{R}^m$ and $T : \mathbb{R}^n \to \mathbb{R}^m$. Then the Fréchet derivative of $T$ at $u \in E$ is $T'(u) = (\frac{\partial T}{\partial u})$, which is the Jacobian matrix of $T$.

2. $T \in \mathcal{L}(E,F)$. Then, by linearity

\[ \|T(u + v) - T(u) - T'(u)v\| = \|T(v) - T'(u)v\| \]

and therefore choosing $T'(u)v = T(v)$ the definition is verified independently of $u$.

3. For $E = C([0,1])$, consider $T : E \to E$ defined by $T(u) = u^3$. Then

\[ T(u + v) - T(u) = (u + v)^3 - u^3 = 3uv^2 + 3u^2v + v^3 \]

Looking at the linear part in $v$, we guess that $T'(u) = 3u^2$. In fact

\[ \|T(u + v) - T(u) - T'(u)v\| = \|3uv^2 + v^3\| = o(\|v\|) \]

as wanted.
Remarks:

1. The Fréchet derivative $T'(u)$ is unique. If there were two linear maps $S'(u)$ and $T'(u)$, verifying the definition,

$$\|T(u + v) - T(u) - T'(u)v\| = o(\|v\|)$$

and

$$\|T(u + v) - T(u) - S'(u)v\| = o(\|v\|)$$

for $\|v\| \leq \delta$. Then

$$\|T'(u) - S'(u)\|_{L(E,F)} = \sup_{\|v\| = \delta} \|(T'(u) - S'(u))v\| \leq \|T'(u)v - T(u + v) + T(u)\| + \|S'(u)v - T(u + v) + T(u)\| \leq o(\|v\|) = \epsilon\|v\|$$

for all $\epsilon > 0$, implying $S'(u) = T'(u)$.

2. If $T$ is Fréchet differentiable at $u$, then $T$ is continuous at $u$, since

$$\|T(u + v) - T(u)\| \leq \|T'(u)v\| + o(\|v\|) \leq M\|v\| + o(\|v\|)$$

which implies the continuity at $u$ since $M\|v\| + o(\|v\|)$ is small whenever $\|v\|$ is small.

3. If $S$ and $T$ are Fréchet differentiable at $u$ then $S + T$ is Fréchet differentiable at $u$, and

$$(S + T)'(u) = S'(u) + T'(u)$$

Example 23. Let $E = C^1([0, 1])$, $F = \mathbb{R}$ and consider

$$I(u) = \int_0^1 G(x, u, u') \, dx$$

where $G = G(x, u, p) \in C^1([0, 1] \times \mathbb{R}^2, \mathbb{R})$. Let us study the Fréchet differentiability of $I$.

$$I(u + v) - I(u) = \int_0^1 \left(G(x, u + v, u' + v') - G(x, u, u')\right) \, dx = \int_0^1 \int_0^1 \frac{d}{ds} G(x, u + sv, u' + sv') \, ds \, dx$$

Looking at the “linear part” in $v$, we guess

$$I'(u)v = \int_0^1 \left(G_u(x, u, u')v + G_p(x, u, u')v'\right) \, dx$$
To show the claim we need to prove that

\[ |I(u + v) - I(u) - \int_0^1 (G_u(x, u, u')v + G_p(x, u, u')v') \, dx| = o(\|v\|)_{C^1([0,1])} \]

By the above computations

\[
I(u + v) - I(u) - \int_0^1 (G_u(x, u, u')v + G_p(x, u, u')v') \, dx \\
= \int_0^1 \int_0^1 \left[ (G_u(x, u + sv, u' + sv')v + G_p(x, u + sv, u' + sv')v') ds \\
- (G_u(x, u, u')v - G_p(x, u, u')v') \right] \, dx \\
= \int_0^1 \int_0^1 \left[ (G_u(x, u + sv, u' + sv') - G_u(x, u, u'))v \\
+ (G_p(x, u + sv, u' + sv') - G_p(x, u, u'))v' \right] \, ds \, dx
\]

Since \( G \in C^1([0,1] \times \mathbb{R}^2) \), both \( G_u \) and \( G_u' \) are continuous. Hence, for \( \|v\|_E \leq \delta \)

\[ |G_u(x, u + sv, u' + sv') - G_u(x, u, u')| \leq \max_{x \in [0,1], \|v\|_E \leq \delta} |G_u| \epsilon \equiv M_1 \epsilon \]

and similarly

\[ |G_p(x, u + sv, u' + sv') - G_p(x, u, u')| \leq \max_{x \in [0,1], \|v\|_E \leq \delta} |G_u'| \epsilon \equiv M_2 \epsilon \]

Then

\[ |I(u + v) - I(u) - I'(u)v| \leq \int_0^1 (M_1|v| + M_2|v'|) \, dx \leq M_3 \max_{x \in [0,1]}(|v(x)| + |v'(x)|) = M_3\|v\|_E \]

**Definition 11.** If \( T \) is Fréchet differentiable for all \( u \in O \) and \( T'(u) \) is continuous (in the topology of \( L(E, F) \)), we write that \( T \in C^1(O, F) \).

We have seen in the examples, that somehow we have to guess the Fréchet derivative. This may not always be as simple as it has been so far. We will introduce a weaker notion of derivative, the Gateaux derivative, that in particular will help us to find the Fréchet derivative.

**Definition 12.** Let \( T : O \to F \), where \( O \subset E \) is open. Then \( T \) is Gateaux differentiable at \( u \in O \) if

\[ T'_G(u)v = \lim_{h \to 0} \frac{T(u + hv) - T(u)}{h} \]

exists for all \( v \in E \).
The difference between the two notions of derivative is that the Gateaux derivative involves only $\| \cdot \|_F$, while the Fréchet derivative involves both $\| \cdot \|_E$ and $\| \cdot \|_F$. In the Euclidean space, the Gateaux derivative corresponds to the directional derivative.

If $T$ is Fréchet differentiable then is Gateaux differentiable and $T'(u) = T'_G(u)$. The converse may not be true.

We are interested in studying maps $I \in C^1(E, \mathbb{R})$, where $E$ is either a Banach or a Hilbert space. Then $I'(u) \in \mathcal{L}(E, \mathbb{R})$, and via the Riesz Representation Theorem $I'(u)$ will be associated to some $w \in E$.

Consider $I \equiv I_1 + I_2$, where

$$I_1(q) = \int_a^b \frac{1}{2} |\dot{q}|^2 dt$$

and for $V \in C^1([a, b] \times \mathbb{R}, \mathbb{R}^n)$

$$I_2(q) = \int_a^b -V(t, q) dt$$

defined in $E = W^{1,2}([a, b], \mathbb{R}^n)$. We will prove that $I \in C^1(E, \mathbb{R})$.

**Lemma 3.** $I_1 \in C^1(E, \mathbb{R})$.

**Proof:**

By its definition

$$I_1(q + v) = \int_a^b \frac{1}{2} |\dot{q} + \dot{v}|^2 dt = \int_a^b \frac{1}{2} |\dot{q}|^2 + (\dot{q}, \dot{v}) + \frac{1}{2} |\dot{v}|^2 dt$$

$$= I_1(q) + I_1'(q)v + o(\|v\|)$$

Hence, it is obvious that

$$I_1'(q)v = \int_a^b (\dot{q}, \dot{v}) dt$$

We must prove that $I_1'$ is continuous in the appropriate topology. We will show that if $q_m \to q$ in $E$ then $I_1'(q_m) \to I_1'(q)$ in $\mathcal{L}(E, F)$.

$$|I_1'(q_m)\varphi - I_1'(q)\varphi| = \left| \int_a^b (q_m - q) \cdot \dot{\varphi} dt \right| \leq \|q_m - q\|_{L^2([a,b])} \|\dot{\varphi}\|_{L^2([a,b])}$$

$$\leq \|q_m - q\|_E \|\varphi\|_E$$

Then

$$\sup \frac{|I_1'(q_m)\varphi - I_1'(q)\varphi|}{\|\varphi\|_E} \leq \|q_m - q\|_E$$

and therefore

$$\|I_1'(q_m) - I_1'(q)\|_{E'} \leq \|q_m - q\|_E$$

So $I_1 \in C^1(E, \mathbb{R})$. $\Diamond$
Lemma 4.

\[ I_2(q) = \int_a^b V(t, q) \, dt \in C^1(E, \mathbb{R}) \]

for all \( V \in C^1([a, b] \times \mathbb{R}^n, \mathbb{R}) \).

Proof:

To figure out what the Fréchet derivative might be, we will start by computing the Gateaux derivative. For any \( \varphi \in E \)

\[
\frac{I_2(q + h\varphi) - I_2(q)}{h} = \frac{1}{h} \int_a^b \left( V(t, q + h\varphi) - V(t, q) \right) \, dt
\]

\[
= \frac{1}{h} \int_a^b \int_0^1 \frac{d}{ds} V(t, q + hs\varphi) \, ds \, dt
\]

\[
= \frac{1}{h} \int_a^b \int_0^1 V_q(t, q + hs\varphi) \, h\varphi \, dt
\]

which implies

\[
(I'_2(q))_{G\varphi} = \lim_{h \to 0} \frac{I_2(q + h\varphi) - I_2(q)}{h} = \int_a^b V_q(t, q)\varphi \, dt
\]

We now have to verify that this is the Fréchet derivative, i.e., if

\[
|I_2(q + \varphi) - I_2(q) - \int_a^b V_q(t, q)\varphi \, dt| = \int_a^b \left| V(t, q + \varphi) - V(t, q) - V_q(t, q)\varphi \right| \, dt = o(\|\varphi\|_E)
\]

Pick any \( \epsilon > 0 \). Since \( V \in C^1([a, b], \mathbb{R}^n) \), for all \( (t, z) \in [a, b] \times \mathbb{R}^n \) there exists \( \delta = \delta(\epsilon, t, z) \) such that

\[
|w| < \delta \quad \Rightarrow \quad |V(t, z + w) - V(t, z) - V_q(t, z)w| \leq \epsilon |w|
\]

(20)

since \( t \in [a, b] \) and this is a compact interval, \( \delta = \delta(\epsilon, z). \) If furthermore we restrict to \( |z| \leq M \), then \( \delta = \delta(\epsilon, M). \) Using Sobolev inequality, if \( q \in E \), then \( \|q\|_{L^\infty([a,b])} \leq \alpha\|q\|_E. \) Choose \( M = \alpha\|q\|_E + 1 \) and \( \varphi \in E \) such that \( \|\varphi\|_E \leq \delta. \) Without loss of generality we can assume that \( \alpha\delta \leq 1. \) Then

\[
\|q + \varphi\|_{L^\infty([a,b])} \leq \alpha(\|q\|_E + \|\varphi\|_E) \leq \alpha\|q\|_E + 1 = M
\]

Hence we are in the conditions of (20) and therefore

\[
|V(t, q + \varphi) - V(t, q) - V_q(t, q)\varphi| \leq \epsilon |\varphi(t)|
\]

which implies

\[
|I_2(q + \varphi) - I_2(q) - \int_a^b V_q(t, q) \cdot \varphi \, dt| \leq \int_a^b \epsilon |\varphi(t)| \, dt \leq \epsilon (b - a)^{1/2}\|\varphi\|_E
\]

40
Thus $I_2$ is Fréchet differentiable. We still have to prove that $I_2'(q)$ is continuous in $E'$. Let $q_m \to q$ in $E$.

$$\|I_2'(q_m) - I_2'(q)\|_{\mathcal{L}(E,\mathbb{R})} = \sup_{\|\varphi\|_E = 1} |I_2'(q_m)\varphi - I_2'(q)\varphi|$$

$$= \sup_{\|\varphi\|_E = 1} \left| \int_a^b (V_q(t, q_m) - V_q(t, q)) \cdot \varphi \, dt \right|$$

$$\leq \sup_{\|\varphi\|_E = 1} \|V_q(t, q_m) - V_q(t, q)\|_{L^2([a,b])} \|\varphi\|_{L^2([a,b])}$$

$$\leq \|V_q(t, q_m) - V_q(t, q)\|_{L^2([a,b])}$$

Since, by Sobolev Theorem, $q_m$ is converging uniformly to $q$, and $V \in C^1([a,b] \times \mathbb{R}^n)$

$$\|V_q(t, q_m) - V_q(t, q)\|_{L^2([a,b])} \to 0$$
as $m \to \infty$ and therefore $I_2'(q)$ is continuous.

$\diamond$

By Lemma 3 and Lemma 4, we can conclude that $I = I_1 - I_2 \in C^1(E, \mathbb{R})$.

**Definition 13.** If $I \in C^1(E, \mathbb{R})$, $E$ is a real Banach space, $u$ is a critical point of $I$ if $I'(u) \equiv 0$, that is $I'(u)v = 0$ for all $v \in E$.

**Lemma 5.** If $I \in C^1(E, \mathbb{R})$, a local maximum or minimum of $I$ is a critical point of $I$.

**Proof:**

Let $v$ be a local minimum of $I$. As a consequence, for any $\varphi \in E$ and $\delta \in \mathbb{R}^+$, $I(v + \delta \varphi) \geq I(v)$, implying

$$\frac{I(v + \delta \varphi) - I(v)}{\delta} \geq 0$$

Letting $\delta \to 0$ we obtain $I'_G(v)\varphi \geq 0$, but replacing $\varphi$ by $-\varphi$, it is obvious that $I'_G(v)\varphi \leq 0$. Hence $I'_G(v)\varphi = 0$ for all $\varphi \in E$.

$\diamond$

Summarizing a little bit, we are proving the existence of a solution of (HS), by minimizing the functional

$$I(q) = \int_a^b \left( \frac{1}{2} |\dot{q}|^2 - V(t, q) \right) \, dt$$

over the class

$$\Gamma = \{q \in E : q(a) = \alpha, q(b) = \beta\}$$

So far, we proved the existence of $v \in \Gamma$ such that $\inf_I I = I(v)$, and that $I \in C^1(E, \mathbb{R})$. To finish our problem we still have to prove a couple of things: that $v$ is a critical point of $I$, and that $v \in C^2([a,b], \mathbb{R}^n)$ and that way, as we have seen, it will satisfy the Euler equation.
Lemma 6. If \( v \) is the minimizer of \( I \) over \( \Gamma \), then \( v \) is a critical point of \( I \).

Proof:
Choose \( \varphi \in E \) and arbitrary \( \delta \in \mathbb{R} \). If \( v \) is the minimizer of \( I \) over \( \Gamma \), in order to have \( v + \delta \varphi \in \Gamma \), we must restrict to the case when \( \varphi(a) = \varphi(b) = 0 \). Then, as in the proof of Lemma 5,

\[
I'(v)\varphi = 0 \quad \forall \varphi \in E
\]

with \( \varphi(a) = \varphi(b) = 0 \). This means that we still have some work to do to show that \( v \) is a critical point of \( I \) in \( \Gamma \). If we consider the special case, \( \alpha = \beta = 0 \), we can replace \( E = W^{1,2}([a,b], \mathbb{R}^n) \) by \( E_0 = W^{1,2}_0([a,b], \mathbb{R}^n) \) which is a proper subspace of \( E \). We showed that \( I \in C^1(E, \mathbb{R}) \), and it is easy to understand that with similar arguments one can show that \( I \in C^1(E_0, \mathbb{R}) \). If this is the case, \( v \) is a critical point since \( I'(v)\varphi = 0 \) for all \( \varphi \in E_0 \).

To prove the general case, we have to use a trick to reduce the problem to \( E_0 \). Choose any function \( p \in E \) such that \( p(a) = \alpha \) and \( p(b) = \beta \). In particular \( p \) can be the linear function

\[
p(t) = \frac{t - a}{b - a} \beta + \frac{b - t}{b - a} \alpha
\]

Define

\[
J(q) = I(q + p)
\]

where \( q \in E_0 \). Notice that we proved that if \( Q \) is a minimizer of \( J \) over \( E_0 \), then \( J'(Q) \equiv 0 \). But, if \( v \) is a minimizer of \( I \) over \( E \), by its definition, \( v - p \in E_0 \) and is a minimizer of \( J \) over \( E_0 \). Then

\[
I'(v) = J'(v - p) = 0
\]

meaning that \( v \) is a critical point of \( I \).

Lemma 7. If \( v \in \Gamma \) is such that \( \inf \Gamma I = I(v) \), then \( v \in C^2([a,b], \mathbb{R}^n) \).

Proof:
Consider the linear boundary value problem

\[
\ddot{q} + V_q(t, v) = 0
\]

\[
q(a) = \alpha
\]

\[
q(b) = \beta
\]

Homework 10. Solve this Boundary Value Problem (write the solution explicitly).

There is \( w \in C^2([a,b], \mathbb{R}^n) \) solution to this problem. Then

\[
\ddot{w} + V_q(t, v) = 0 \quad , \quad w(a) = \alpha \ , \ w(b) = \beta
\]
Take any \( \varphi \in E_0 \), multiply both sides of the equation and integrate over \([a,b]\), yields

\[
\int_a^b (\ddot{w} + V_q(t,v)) \cdot \varphi dt = 0, \quad \forall \varphi \in E_0
\]

Integrate by parts the first term of the integral

\[
\dot{w}\varphi|_a^b - \int_a^b \dot{w} \cdot \varphi dt + \int_a^b V_q(t,v) \cdot \varphi dt = 0, \quad \forall \varphi \in E_0
\]

and since \( \varphi(a) = \varphi(b) = 0 \)

\[
\int_a^b \left( \dot{\varphi} - V_q(t,v) \right) \cdot \varphi dt = 0, \quad \forall \varphi \in E_0
\]

(21)

On the other hand, by Lemma 6

\[
\int_a^b \left( \dot{v} \cdot \varphi - V_q(t,v) \right) \cdot \varphi dt = 0, \quad \forall \varphi \in E_0
\]

(22)

Subtracting equations (21) and (22)

\[
\int_a^b (\ddot{w} - \dot{v}) \cdot \varphi dt = 0
\]

for all \( \varphi \in E_0 \). In particular, since \( w - v \in E_0 \), we can take \( \varphi = w - v \), and

\[
\int_a^b |\ddot{w} - \dot{v}|^2 dt = 0
\]

which implies \( |\ddot{w} - \dot{v}| = 0 \) a.e. in \([a,b]\).

**Lemma 8.** If \( \varphi \in W^{1,2}([a, b], \mathbb{R}^n) \) and \( \varphi' = 0 \) a.e., then \( \varphi = \text{constant} \).

**Proof:**

Suppose that \( \varphi \in C^1([a, b]) \). Then we can write

\[
\varphi(x) = \varphi(a) + \int_a^x \varphi'(t) dt
\]

and since \( \varphi' = 0 \), we conclude \( \varphi(x) = \varphi(a) \) for all \( x \in [a, b] \). By definition, every function of \( E \) is the limit (in the \( E \) norm) of \( C^1([a, b]) \) functions. Then the result holds for an arbitrary \( \varphi \in E \).

\[\diamondsuit\]

If \( w - v \) is constant, using the boundary conditions we conclude that \( w - v = 0 \) which implies in particular that \( v \in C^2([a, b], \mathbb{R}^n) \).

\[\diamondsuit\]
We have shown that any minimizer of $I$ is in $C^2([a,b])$ and therefore is a solution of the Euler equation. Let’s look now at a more general setting. Consider

$$I(u) = \int_a^b F(x, u, u')dx$$

and suppose that there exists a critical point $v$. We want to indicate a method in how to check that the critical point has more regularity than the one it seems. All we have to start with is the weak form of the Euler equation

$$I'(v)\varphi = 0, \text{ for all } \varphi \text{ in an appropriate class of functions},$$

and suppose for the moment that $v \in W^{3,2}$ (which corresponds to $v \in C^3$), and take $\varphi = v''$. Then (23) is equivalent to

$$\int_a^b (F_u(x, v, v')v'' + F_p(x, v, v')v''')dx = 0$$

Integrate by parts

$$\int_a^b F_p(x, v, v')v'''dx = F_p(x, v, v')v''|_a^b - \int_a^b (F_{px}(x, v, v') + F_{pu}(x, v, v')v' + F_{pp}(x, v, v')v'')v''dx$$

and assume that $\varphi \equiv 0$ at the boundary points. Since $v'' = \varphi'$, that will imply $v''(a) = v''(b) = 0$ and therefore

$$\int_a^b F_p(x, v, v')v'''dx = -\int_a^b (F_{px}(x, v, v') + F_{pu}(x, v, v')v' + F_{pp}(x, v, v')v'')v''dx$$

Hence

$$\int_a^b F_{pp}(x, v, v')(v'')^2dx = -\int_a^b F_p(x, v, v')v'''dx - \int_a^b (F_{px}(x, v, v') + F_{pu}(x, v, v')v')v''dx$$

Notice that the left hand side is quadratic in $v''$, while the right hand side is linear in $v''$. Further assume that

$$F_{pp}(x, y, z) \geq \gamma > 0$$

Then

$$\gamma \|v''\|^2_{L^2} \leq \int_a^b F_{pp}(x, v, v')(v'')^2dx$$

and we can use Schwarz inequality to obtain estimates on $\|v''\|_{L^2}$.

How can we get around the assumptions we made? In the general case we have to localize. Consider

$$\varphi(t) = (\chi(t)v^h)^{-h}$$
where $\chi$ is a cut off function and

$$v^h(x) = \frac{v(x + h) - v(x)}{h}$$

$\varphi$ is an approximation of the second derivative. Then instead of integrating by parts, we can difference by parts and conclude that $v \in W^{2,2}$ ($v' \in C^1$) which is enough to write the Euler equation (a.e. sense). Then using the same arguments we can show that $v''$ exists.

5 Applications to Differential Equations

5.1 Existence of Periodic Solutions of (HS)

After having proved the existence of solution for the boundary value problem, we are interested in finding some special solutions of (HS): we will start by proving the existence of periodic solutions. Suppose we are in a simplified setting, i.e., we are solving

$$\ddot{q} + V'(q) = 0$$

where $V \in C^1(\mathbb{R}, \mathbb{R})$. If $q$ is a solution, multiplying both sides of the equation by $\dot{q}$

$$\dot{q} \cdot \ddot{q} + V'(q) \cdot \dot{q} = 0$$

which is equivalent to

$$\frac{d}{dt} \left( \frac{1}{2} |\dot{q}|^2 + V(q) \right) = 0$$

Then

$$\frac{1}{2} |\dot{q}|^2 + V(q) = \text{constant}$$

meaning that the total energy of the system is conserved. Since we are working in $\mathbb{R}$, we can make the analysis in the phase plane: the solutions of (HS) will be the level curves of the energy function, i.e., will correspond to the curves

$$\frac{1}{2} \dot{p}^2 + V(q) = c$$

Suppose that we can prove the following

if $V(q) = H$ and $V'(q) = 0$ then $x(t) = q$ is an equilibrium solution of (HS);

Then, in any other type of solution, $V'(q) \neq 0$, and we can interpret the condition

$$\frac{1}{2} \dot{p}^2 + V(q) = c$$

as a $C^1$ manifold (the gradient of the expression is everywhere nonzero - the curve has no edges). Consider any initial point on such manifold and turn on time:
since energy is constant it will remain on the curve

\[ V'(q) \neq 0 \] means that it can’t stop

If the curve is closed, the two conditions correspond to a periodic solution (this is the qualitative way to obtain a periodic solution).

If \( n \) is arbitrary, things are not as intuitive. We will use the calculus of variations.

Let \( T > 0 \) arbitrary, and consider

\[
\ddot{q} + V_q(t, q) = 0 \quad \text{(HS)}
\]

where \( V \in C^2(\mathbb{R} \times \mathbb{R}^n, \mathbb{R}) \) is \( T \)-periodic in \( t \).

**Question:** Is there a periodic solution to this problem?

Finding periodic solutions is equivalent to find solutions of (HS) with boundary conditions \( q(0) = q(T) \) and \( q'(0) = q'(T) \).

**Example 24.**

1. \( \ddot{q} + a(t) \sin q = 0 \), where \( a \) is continuous and periodic. If \( a(t) = 1 \), we have the model of the simple pendulum.

2. \( \ddot{q} + \frac{q}{R(q)} = 0 \) where \( R \) is a periodic function, is the equation for the interaction between earth/sun.

3. **n-body problem:** Let \( q_1, \ldots, q_n \) be the position function of \( n \) bodies. By the second law of \( \text{Newton} \)

\[
m_i \ddot{q}_i = \frac{\partial V}{\partial q_i}
\]

where \( V \) is the \( n \) body potential

\[
V(q) = \sum_{i \neq j} \frac{m_i m_j}{|q_i - q_j|}
\]

which is a singular potential.

We will seek the periodic solutions of (HS), by minimizing the corresponding functional

\[
I(q) = \int_0^T \left( \frac{1}{2} |\dot{q}|^2 - V(t, q) \right) dt
\]

over \( E_T = W^{1,2}_T(\mathbb{R}, \mathbb{R}^n) \). The following example show us that it is not enough to consider

\[
V(t, x) \leq 0 \quad , \quad t \in [0, T] , \quad x \in \mathbb{R}^n
\]

If \( (q_m) \) is a minimizing sequence, \( I(q_m) \leq M \) implies \( (\dot{q}_m) \) bounded in \( L^2([0, T]) \), but it is not enough to prove that \( (q_m) \) is bounded in \( E_T \).
Example 25. Consider
\[ I(q) = \int_0^T \left( \frac{1}{2} |q|^2 + \frac{1}{1 + |q|^2} \right) dt \]

Take \( q_m = R_m \to \infty \) as \( m \to \infty \), where \( R_m \) is a constant sequence (which means in particular that is periodic). Then

\[ I(q_m) = \int_0^T \frac{1}{1 + |R_m|^2} dt \to 0 \quad \text{as} \quad m \to \infty \]

and therefore \( \inf_{E_T} I(q) = 0 \). But it is easy to see that there is no \( T \)-periodic function for which 0 is achieved.

Theorem 8. If \( V(t, x) \in C^2(\mathbb{R} \times \mathbb{R}^n, \mathbb{R}) \) is \( T \)-periodic, for some \( T > 0 \), and verifies
\[(V1) \quad -V(t, x) \geq \alpha |x|^\beta - \gamma , \ t \in [0, T] \ x \in \mathbb{R}^n\]
and \( \alpha, \beta, \gamma \) fixed positive constants, then (HS) admists at least one \( T \)-periodic solution.

Proof:
As mentioned, we will prove the Theorem using a variational argument. For \((q_m) \subset E_T\), a minimizing sequence, there exists \( M > 0 \) such that

\[ M \geq \int_0^T \left( \frac{1}{2} |\dot{q}_m|^2 - V(t, q_m) \right) dt \]

Using condition (V1)
\[ \int_0^T \left( \frac{1}{2} |\dot{q}_m|^2 + \alpha |q_m|^\beta \right) dt \leq M + \gamma T \]

This implies directly that

\[ \|\dot{q}_m\|_{L^2([0,T])} \leq 2(M + \gamma T) \]

and

\[ \int_0^T |q_m|^{\beta} dt \leq \frac{1}{\alpha}(M + \gamma t) \quad (24) \]

Then \((\dot{q}_m)\) is bounded in \( L^2([0,T]) \). To prove that \((q_m)\) is also bounded \( L^2([0,T]) \), we have to consider two distinct cases:

Case 1. \( \beta > 2 \)

Using Holder’s inequality
\[ \|q_m\|_{L^2([0,T])} = \int_0^T |q_m|^2 dt \leq \left( \int_0^T 1^p dt \right)^{1/p} \left( \int_0^T (|q_m|^2)^q dt \right)^{1/q} \]

47
Choose \( q \) such that \( 2q = \beta \), and \( p \) such that \( \frac{1}{p} + \frac{1}{q} = 1 \) (notice that \( \beta > 2 \) implies \( q > 1 \) and therefore \( p \) is well defined). Then

\[
\|q_m\|_{L^2([0,T])}^2 \leq T^{1/p} \left( \int_0^T |q_m|^{\beta} \, dt \right)^{2/\beta}
\]

As a consequence of (24)

\[
\|q_m\|_{L^2([0,T])}^2 \leq T^{1/p} \left( \frac{1}{\alpha} (M + \gamma T) \right)^{2/\beta}
\]

and we can conclude that \( (q_m) \) is bounded in \( E_T \).

**Case 2.** \( 0 < \beta \leq 2 \)

As noticed, in this case we can’t use Holder’s inequality. Any \( q \in E_T \) can be written as a Fourier series

\[
q(t) = \sum_{j \in \mathbb{Z}} a_j e^{\frac{i2\pi j}{T} t} = a_0 + \text{remaining terms}
\]

The above decomposition \( q(t) = a_0 + R(t) \), corresponds to the decomposition of our space in 2 orthogonal subspaces, i.e., \( a_0 \in \mathbb{R}^n \) and \( R(t) \in (\mathbb{R}^n)^\perp \). By definition

\[
a_0 = \frac{1}{T} \int_0^T q(t) \, dt \equiv [q]
\]

which is the mean value or average of \( q \). Define

\[
q = [q] + \hat{q}
\]

Computing the derivative

\[
\dot{q}(t) = \sum_{j \neq 0} a_j \frac{i2\pi j}{T} e^{\frac{i2\pi j}{T} t}
\]

and by Parseval’s identity, we can easily verify that

\[
\|\hat{q}\|_{L^2([0,T])}^2 = \|q - [q]\|_{L^2([0,T])}^2 = \sum_{j \neq 0} a_j^2
\]

On the other hand

\[
\|\hat{\dot{q}}\|_{L^2([0,T])}^2 = \|\dot{q}\|_{L^2([0,T])}^2 = \sum_{j \neq 0} \left( a_j \frac{2\pi j}{T} \right)^2 \geq \left( \frac{2\pi}{T} \right)^2 \sum_{j \neq 0} a_j^2 = \left( \frac{2\pi}{T} \right)^2 \|\hat{q}\|_{L^2([0,T])}^2
\]

i.e.

\[
\|\dot{q}\|_{L^2([0,T])}^2 \leq c(T) \|\hat{q}\|_{L^2([0,T])}^2
\]
We will now proceed in finding bounds for \((q_m)\) in \(E_T\). Since
\[
\|\dot{q}_m\|_{L^2([0,T])}^2 \leq 2M_1
\]
we can write
\[
\|\dot{\hat{q}}_m\|_{L^2([0,T])}^2 \leq C(T)\|\dot{q}_m\|_{L^2([0,T])}^2 = C_1(T)\|\dot{q}_m\|_{L^2([0,T])}^2
\]
and therefore \(\dot{\hat{q}}_m\) is bounded in \(L^2([0,T])\). It remains to show that \([q_m]\) is also bounded, and since it is a constant we have to find a bound for \(||[q_m]|\). Start noticing that
\[
|[q_m]| = |q_m - \hat{q}_m| \leq |q_m| + |\hat{q}_m|
\]
Raising to \(\beta\)
\[
||q_m||^\beta \leq \left(|q_m| + |\hat{q}_m|\right)^\beta \leq 2^\beta \left(|q_m|^\beta + |\hat{q}_m|^\beta\right)
\]
as used before. Integrating \(t\) over \([0,T]\)
\[
T||q_m||^\beta \leq 2^\beta \int_0^T \left(|q_m|^\beta + |\hat{q}_m|^\beta\right) dt
\]
Once again, by (24)
\[
\int_0^T |q_m|^\beta dt \leq \frac{M_1}{\alpha}
\]
and therefore
\[
||q_m||^\beta \leq \frac{2^\beta}{T} \left( \frac{M_1}{\alpha} + T||\hat{q}_m||_{L^\infty([0,T])}^\beta \right)
\]
Using Sobolev Inequality
\[
||q_m||^\beta \leq \frac{2^\beta}{T} \left( \frac{M_1}{\alpha} + cT||\hat{q}_m||_{E_T}^\beta \right)
\]
Since we proved that \((\dot{q}_m)\) is bounded in \(E_T\), we can finally conclude that \((q_m)\) is bounded in \(E_T\) and therefore \((q_m)\) admits a weakly convergent subsequence (which we will denote again by \((q_m)\)), i.e., there exists \(v \in E_T\) such that \(q_m \rightharpoonup v \) in \(E_T\) and \(q_m \rightarrow v\) in \(L^\infty([0,T])\). As before I is weakly lower semicontinuous, implying
\[
I(v) \leq \lim inf_{m \rightarrow \infty} I(q_m)
\]
Since \(v \in E_T\),
\[
I(v) = \inf_{q \in E_T} I(q)
\]
We still have to show that \(v \in C^2([0,T])\) (the variational characterization implies that it is a \(T\)-periodic solution of (HS)). Let us start proving that \(v\) is a critical point of \(I\), i.e., that \(I'(v)\varphi = 0\), for all \(\varphi \in E_T\). But in this case this is straightforward, since for any \(\varphi \in E_T\) and \(\delta \in \mathbb{R}\), \(v + \delta\varphi \in E_T\) (so this case is easier than the BVP). Thus
\[
0 = I'(v)\varphi = \int_0^T (\dot{v} \cdot \dot{\varphi} - V_0(t,v)\varphi) \ dt, \ \forall \varphi \in E_T
\]
We want to show next that this implies \( v \in C^2([0, T], \mathbb{R}^n) \). We will use a variation of the regularity argument used for the BVP. Given that \( v \) is the minimizer of \( I \) over \( E_T \), we want to solve the linear boundary value problem

\[
\begin{align*}
  w'' & = -V_q(t, v) \\
  w(0) & = w(T) \\
  w'(0) & = w'(T)
\end{align*}
\]

Nevertheless, we must point out that the solution is not unique: if \( w \) is such a solution, then \( w + kT \) is also a solution. Hence we must impose an additional condition \([w] = [v]\)

**Homework 11.** Show that there exists a unique solution to this problem. (Hint: Notice that as a consequence of the periodicity, \([V_q] = 0\) and then use calculus or Fourier series).

Let \( w \in C^2([0, T]) \) be the solution of the considered problem. For an arbitrary \( \varphi \in E_T \), multiply both sides of the equation and integrate over \( t \in [0, T] \). Then

\[
\int_0^T \varphi \cdot w'' \, dt = - \int_0^T \varphi \cdot V_q(t, v) \, dt
\]

Integrating by parts the left hand side

\[
w' \cdot \varphi|_0^T - \int_0^T \dot{\varphi} \cdot \dot{w} \, dt = - \int_0^T \varphi \cdot V_q(t, v) \, dt
\]

Since \( \varphi \) and \( \dot{w} \) are \( T \)-periodic,

\[
\int_0^T \dot{\varphi} \cdot \dot{w} \, dt = \int_0^T \varphi \cdot V_q(t, v) \, dt
\]

Subtracting the weak Euler equation

\[
\int_0^T \varphi \cdot (\dot{v} - \dot{w}) \, dt = 0
\]

Take \( \varphi = w - v \), we can conclude that \( \dot{v} - \dot{w} = 0 \) and henceforth \( w = v \) =constant. But by our normalization \([w] = [v]\) which implies that \( w = v \), in particular \( v \in C^2([0, T]) \).

\( \Diamond \)

When seeking for periodic solutions of (HS), by minimizing \( I \) over \( E_T \), some problems may arise:
1. If $V$ is independent of $t$ then is obviously $T$-periodic, for all $T$, which implies by our results, that there exists a $T$ periodic solution for all $T > 0$. But notice, that since we assumed

$$-V(q) \geq \alpha |q|^{\beta} - \gamma, \quad \alpha, \beta, \gamma > 0$$

$-V$ is very large for large values of $|q|$ and therefore it must have a global minimum $\xi$. Consider $q(t) \equiv \xi$. By construction $V'(\xi) = 0$ and hence $\xi$ is a solution of (HS) (it is an equilibrium solution). If this is the case there may not exist any other periodic solution.

2. If $V$ is $T$ periodic in $t$, then $V$ is $kT$ periodic in $t$, for all $k > 0$. By our results, there exists a solution of (HS) which is $kT$ periodic, but it may be $k$ copies of the $T$ periodic solution. To show the existence of subharmonic solutions we need more work.

**Homework 12.** *In the same setting as in Theorem 8, for $V = V(q)$, what (“a priori” conditions must we impose, so that, for $v_1, v_2$ verifying

$$I(v_1) = \inf_{E_T} I \equiv c_T$$

and

$$I(v_2) = \inf_{E_{2T}} I \equiv c_{2T}$$

exist and are geometrically distinct.*

**Example 26. (Multiple Pendulum Type of Solutions)** *We consider the system

$$\ddot{q} + V_q(t, q) = 0 \quad (HS)$$

where $V = V(t, q) \in C^2(\mathbb{R} \times \mathbb{R}^n, \mathbb{R})$, $T$ periodic in $t$ and $T_i$ periodic in $q_i$. The simplest example is the simple pendulum

$$\ddot{q} + \alpha(t) \sin q = 0$$

where $\alpha$ is a periodic function. In the multiple pendulum, the variables $q_i$ correspond to the several angles. We will consider the case

$$V(t, q) = W(t, q) - f(t) \cdot q$$

where $f$ is continuous, $T$ periodic and satisfies $[f] = 0$ (to ensure the existence of periodic solutions) and $W$ satisfies the same conditions as $V$. For this $V$, the equation takes the form

$(HS)$

$$\ddot{q} + W_q(t, q) = f(t)$$

**Theorem 9.** *Under the stated conditions, there exists a $T$-periodic solution of (HS).*

**Proof:**

We will prove the Theorem, by minimizing

$$I(q) = \int_0^T \left( \frac{1}{2} |\dot{q}|^2 - V(t, q) \right) dt$$

51
over $E_T = W_T^{1,2}(\mathbb{R}, \mathbb{R}^n)$. Let 
\[ c_T = \inf_{E_T} I \]
and $(q_m) \subset E_T$ a minimizing sequence. Using the periodicity of $W$ and the condition $[f] = 0$, it is obvious that 
\[ I(q + (k_1T_1, ..., k_nT_n)) = I(q) \]
for all $k_1, ..., k_n \in \mathbb{Z}$. This invariance property is important, since with it we will control the norm of a minimizing sequence. As before, we can decompose 
\[ q_m = [q_m] + \hat{q}_m \]
and by the invariance property we can replace $(q_m)$ by any $\tilde{q}_m = q_m + (k_1T_1, ..., k_nT_n)$, with $(k_1, ..., k : n) \in \mathbb{Z}^n$. We will normalize by choosing $(k_1, ..., k_n)$ so that 
\[ 0 \leq [(\tilde{q}_m)_i] < T_i \]
Without loss of generality, assume that $(q_m)$ verifies 
\[ 0 \leq [(q_m)_i] < T_i \]
which is a very strong control on the mean value of the minimizing sequence.

Let’s prove that the minimizing sequence is bounded in $E_T$. We can start to assume that $I(q_m) \leq M$. We will consider two cases $f = 0$ and $f \neq 0$. If $f = 0$, 
\[ \frac{1}{2} \int_0^T |\dot{q}_m|^2 dt \leq M + \int_0^T V(t, q_m) dt \]
Since $V$ is periodic and $C^2([0, T])$, it is uniformly bounded in $C([0, T])$, and therefore 
\[ \frac{1}{2} \int_0^T |\dot{q}_m|^2 dt \leq M + (\max_{t \in [0, T]} V(t, q_m(t)))T \leq M_1 \]
uniformly in $m$. From here we proceed as in the previous examples. If $f \neq 0$, 
\[ I(q_m) = \int_0^T \left( \frac{1}{2} |\dot{q}_m|^2 - W(t, q_m) + f(t) \cdot q_m \right) dt \leq M \]
then 
\[ \int_0^T \frac{1}{2} |\dot{q}_m|^2 dt = \int_0^T \frac{1}{2} |\dot{\tilde{q}}_m|^2 dt \leq M + \int_0^T W(t, q_m) dt + \int_0^T |f \cdot q_m| dt \]
Since $W$ is $C^2([0, T])$ and periodic, is uniformly bounded, and therefore 
\[ \int_0^T W(t, q_m) dt \leq M_1 \quad \forall m \]
On the other hand
\[ \int_0^T f \cdot q_m \, dt = \int_0^T f \cdot ( [q_m] + \dot{q}_m ) \, dt = \int_0^T f \cdot [q_m] \, dt + \int_0^T f \cdot \dot{q}_m \, dt \]

Since by hypothesis \([f] = 0\),
\[ \int_0^T f \cdot q_m \, dt = \int_0^T f \cdot \dot{q}_m \, dt \leq \| f \|_{L^2([0,T])} \| \dot{q}_m \|_{L^2([0,T])} \]

On one hand, as we have seen previously
\[ \| \dot{q}_m \|_{L^2([0,T])} \leq C(T) \| \dot{q}_m \|_{L^2([0,T])} \]
on the other hand
\[ \| f \|_{L^2([0,T])} \leq T^{1/2} \| f \|_{L^\infty([0,T])} \]

As a consequence
\[ \int_0^T f \cdot q_m \, dt \leq C(T) \| f \|_{L^\infty([0,T])} \| \dot{q}_m \|_{L^2([0,T])} \]

Then
\[ \frac{1}{2} \| \dot{q}_m \|_{L^2}^2 = \int_0^T \frac{1}{2} |\dot{q}_m|^2 \, dt \leq M_1 + M_2 \| \dot{q}_m \|_{L^2([0,T])} \]

Notice that the left hand side is quadratic in \( \| \dot{q}_m \|_{L^2} \) while the right hand side is linear. Solving the quadratic inequality, we conclude that \( \| \dot{q}_m \|_{L^2([0,T])} \) is bounded uniformly in \( m \).

\[ \diamond \]

Remarks:

1. As a consequence of the invariance property, there is no uniqueness of minimizer. Instead there is a lattice of minimizers. This suggests the existence of more critical points (we will do this later, using the Mountain Pass Theorem).

2. This was a simplified version of the multiple pendulum. In the more general case, the kinetic energy
\[ K(t) = \int_0^T \sum_{i,j} a_{i,j}(q) \dot{q}_i \cdot \dot{q}_j \, dt \]

where \((a_{i,j}(q))\) is a symmetric, positive definite matrix, \(T_i\) periodic in \(q_i\). The harder part is to prove the weak lower semicontinuity.

Example 27. (Singular Systems) Our next example deals with potentials, having a singularity. For example
\[ V(q) = -\frac{1}{|q|^\beta}, \quad \beta > 0 \]
These are potentials arising from the 2-body problem. Our goal is to find a periodic solution winding around the singularity. We will start by formulating the problem. Consider

\[(HS) \quad \ddot{q} + V'(q) = 0\]

where \(V\) satisfies

\[(V1) \quad V \in C^2(\mathbb{R}^2 \setminus \{(0,0)\});\]

\[(V2) \quad V(x) \leq 0 \text{ for all } x \in \mathbb{R}^2 \setminus \{(0,0)\};\]

\[(V3) \quad V(x) \to -\infty \text{ as } |x| \to 0.\]

To prove the existence of such solution we will minimize the functional

\[I(q) = \int_0^T \left( \frac{1}{2} |\dot{q}|^2 - V(q) \right) dt\]

over the class

\[\Gamma_T = \{ q \in W^{1,2}_T(\mathbb{R}, \mathbb{R}^2) : W N_0(q) \neq 0 \}\]

We have to explicit what we mean by \(W N_0(q) \neq 0\). For \(q \in \Gamma_T\) we can consider its characterization in polar coordinates with respect to 0:

\[q(t) = R(t)e^{i\theta(t)}\]

where \(R(0)\) is in the positive \(x\)-axis and \(\theta(0) = 0\). Since \(q(t)\) is continuous, \(\theta\) is continuous. We will define

\[W N_0|_T = \Theta(T) - \Theta(0)\]

As a consequence of the periodicity of the functions in \(E_T\), we will have

\[W N_0|_T = 2k\pi\]

for some \(k \in \mathbb{Z}\). \(k\) will denote the number of times \(q\) winds around 0 in the interval \([0,T]\).

Before we proceed, we can make a couple of simplifications. Since \(V\) does not depend explicitly in \(t\), for \(q \in \Gamma_T\), \(q(-t)\) is the same curve running in the opposite direction, and

\[I(q(-t)) = I(q(t))\]

Hence if \(v \in \Gamma_T\) is a minimizer of \(I\), so it will \(v(-t)\). Consider

\[\Gamma_T = \Gamma_T^+ \cup \Gamma_T^-\]

and by our considerations, it is enough to minimize over \(\Gamma_T^+\). Define

\[c_T = \inf_{\Gamma_T^+} I\]

54
and consider \((q_m) \subset \Gamma_T^+\) a minimizing sequence. Then, there exists \(M\) such that \(I(q_m) \leq M\) and together with condition (V2) implies \(\|q_m\|_{L^2([0,T])}\) is uniformly bounded. As we have seen previously, to find bounds on \(\|q_m\|_{L^2([0,T])}\) is equivalent to find bounds for \([q_m]_m\) where \(q_m = [q_m] + \hat{q}_m\). As before, we can prove that \(\|\hat{q}_m\|_{L^2([0,T])}\) is bounded uniformly in \(m \in \mathbb{N}\). If the mean value piece, \([q_m]_m\), is unbounded, then along a subsequence, \([q_m]_m \to \infty\) as \(i \to \infty\). But, if this is the case, there is no winding, and therefore for \(i\) sufficiently large, \(q_m \notin \Gamma_T^+\), which is a contradiction. Thus we can conclude that \([q_m]_m\) is uniformly bounded, and as a consequence \((q_m)\) is bounded in \(W^{1,2}_T(\mathbb{R}, \mathbb{R}^2)\). Then, there exists \(v \in W^{1,2}_T(\mathbb{R}, \mathbb{R}^2)\), such that, possibly along a subsequence
\[
q_m \rightharpoonup v \quad \text{in} \quad W^{1,2}_T(\mathbb{R}, \mathbb{R}^2)
\]
and
\[
q_m \to v \quad \text{in} \quad L^\infty_T(\mathbb{R}, \mathbb{R}^2)
\]
The only thing left to prove is that \(v \in \Gamma_T^+\), and since \(v \in W^{1,2}_T(\mathbb{R}, \mathbb{R}^2)\) by construction, we have to show that \(WN_0v|_T > 0\). But this doesn’t happen necessarily: the limit may pass through the origin, in which case the winding number does not make sense anymore.

**Example 28.** Let \(V(q) \approx -\frac{1}{|q|}\) (the gravitational potential), and \(q(t) \approx t^\alpha\) near the origin. Then we need two things to happen:
\[
t^\alpha \in W^{1,2}([0,T]) \quad \text{and} \quad I(t^\alpha) < \infty
\]
For the first condition to occur
\[
q \in W^{1,2}([0,T]) \quad \Leftrightarrow \quad \alpha > \frac{1}{2}
\]
since if \(q(t) \approx t^\alpha\) then \(\dot{q}(t) \approx t^{\alpha-1}\), and
\[
\int t^{2\alpha-2} dt < \infty \quad \Leftrightarrow \quad 2\alpha - 2 > -1
\]
On the other hand
\[
-\int V(t^\alpha) dt = -\int \frac{1}{t^\alpha} dt < \infty \quad \Leftrightarrow \quad \alpha < 1
\]
Then for \(\frac{1}{2} < \alpha < 1\) everything is finite. So it is possible that the minimizer approaches the origin.

Such a minimizer is called a collision orbit. To avoid the existence of such orbits we have to assume an extra condition on the potential, which as we will see, keeps candidates to minimizers away from the origin.

(V4) **Strong Force Condition:** There exists a neighborhood of 0, \(N\), and a function \(U \in C^1(\mathbb{R}^2)\) such that
\[
|U'(q)|^2 \leq -V(q) \quad \text{for all} \quad q \in \mathbb{N} \setminus \{0\}
\]
$|U(x)| \rightarrow \infty$ as $x \rightarrow 0$.

**Homework 13.** Show that there exists $V$ satisfying such condition. Consider $V(q) = -\frac{1}{|q|^\alpha}$ for an appropriate $\alpha$.

**Lemma 9.** Let $V$ verify (V1)-(V4). If for $q \in E_T$, $I(q) \leq M$ for some $M > 0$, there exists $\beta = \beta(M)$ such that

$$|q(t)| > \beta$$

for all $t \in [0,T]

**Proof:**

Let’s suppose that $q(t) \in N \setminus \{0\}$, for $t \in [\sigma,s]$, where $N$ is given by (V4). Then

$$U(q(s)) - U(q(\sigma)) = \int_{s}^{\sigma} \frac{d}{dt}U(q(t)) \, dt = \int_{s}^{\sigma} U'(q(t)) \cdot \dot{q}(t) \, dt$$

$$\leq \left( \int_{s}^{\sigma} |U'(q(t))|^2 \, dt \right)^{1/2} \left( \int_{s}^{\sigma} |\dot{q}|^2 \, dt \right)^{1/2}$$

$$\leq \left( \int_{s}^{\sigma} -V(q) \, dt \right)^{1/2} \left( \int_{s}^{\sigma} |\dot{q}|^2 \, dt \right)^{1/2}$$

$$\leq \sqrt{2}I(q)$$

Since $I(q) \leq M$

$$U(q(s)) - U(q(\sigma)) \leq \sqrt{2}M$$

Further suppose that $q(s) \in \partial N$. Based on what we have shown

$$|U(q(\sigma))| \leq C$$

for some constant $C$. But since $|U|$ blows up as we approach the origin, $q(\sigma)$ can’t get too close to it and therefore

$$|q(\sigma)| \geq \beta(M)$$

for all $\sigma$.

$\diamond$

**Homework 14.** Using Lemma , finish the proof of the existence of a periodic solution for the singular hamiltonian system.

### 5.2 Existence of Heteroclinic Solutions of (HS)

Consider

$$\ddot{q} + V_q(t,q) = 0 \quad \text{(HS)}$$

where $V$ satisfies the following conditions
(V1) \( V = V(t, q) \in C^2(\mathbb{R} \times \mathbb{R}^n, \mathbb{R}) \) is 1-periodic in \( t \) and \( q_i, 1 \leq i \leq n \),

(V2) \( V(t, 0) = 0 > V(t, x) \) for all \( x \in \mathbb{R}^n \setminus \mathbb{Z}^n \) and all \( t \in \mathbb{R} \).

Condition (V2) implies that the potential \( V \) has a global maximum at \( q = 0 \). The periodicity of \( V \) implies that the global maximum, 0, is achieved at all integer points (lattice of maximum). Then,

\[
V_q(t, 0) = V_q(t, x) = 0, \quad \forall x \in \mathbb{Z}^n
\]

and as a consequence all points in \( \mathbb{Z}^n \) are equilibrium solutions of (HS).

**Question:** Are there solutions of (HS) connecting 0 with other equilibrium solutions? That is, we want to find solutions of (HS) having the following asymptotic behavior

\[
Q(-\infty) = \lim_{t \to -\infty} Q(t) = 0
\]

and

\[
Q(\infty) = \lim_{t \to \infty} Q(t) = \eta \in \mathbb{Z}^n \setminus \{0\}
\]

These type of solutions (connecting equilibrium solutions) are known as **heteroclinic solutions**.

**Example 29. (Simple Pendulum)**

Consider

\[
\ddot{q} = \sin(2\pi q)
\]

which as mentioned earlier is the equation of motion of the simple pendulum, and corresponds to the potential

\[
V(q) = \frac{1}{2\pi} (\cos(2\pi q) - 1), \quad q \in \mathbb{R}
\]

It is easy to see, that for \( n \in \mathbb{Z}, q(t) \equiv n \) is an equilibrium solution of the differential equation. Furthermore, it is known that near the minimum values of \( V \), the curves

\[
\frac{1}{2} \dot{q}^2 + V(q) = C
\]

represent the periodic solutions of (HS).

**Homework 15.** Sketch the phase plane for this equation.

**Theorem 10.** If \( V \) satisfies (V1)-(V2), for each \( \xi \in \mathbb{Z}^n \) there is \( \eta \in \mathbb{Z}^n \setminus \{\xi\} \) and a solution of (HS) heteroclinic from \( \xi \) to \( \eta \).

Without loss of generality we will assume \( \xi = 0 \). Notice that we have to change our previous arguments, since in this case we are working in an unbounded interval. Our goal is to find a minimizer for

\[
I(q) = \int_{\mathbb{R}} \left( \frac{1}{\tau} |\dot{q}|^2 - V(t, q) \right) dt
\]

over an appropriate class of functions: the elements of this class must verify the following:
- have the right asymptotic behavior;
- give sense to $I(q)$.

We start by remarking that we can’t expect the functions of our class to be square integrable: in particular we want \[ \lim_{t \to \infty} (q(t) = \eta \neq 0. \] This means $W^{1,2}(\mathbb{R}, \mathbb{R}^n)$ is not a reasonable choice. We will choose 
\[ E = W^{1,2}_{\text{loc}}(\mathbb{R}, \mathbb{R}^n) = \{q \in W^{1,2}([a, b], \mathbb{R}^n) : \forall a < b \in \mathbb{R}\} \]
which is a Hilbert space when endowed with the norm 
\[ \|q\|_E^2 = \int_\mathbb{R} |\dot{q}|^2 dt + |q(0)|^2 \]
and define the class
\[ \Gamma = \{q \in E : q(-\infty) = 0 \text{ and } q(\infty) \in \mathbb{Z}^n \setminus \{0\}\} \]
For 
\[ c = \inf_{q \in E} I(q), \]
as an immediate consequence of condition (V2), $c \geq 0$. The Theorem follows from the following propositions:

**Proposition 3.** There exists $Q \in \Gamma$ such that $I(Q) = c$.

**Proposition 4.** $Q$ is a solution of (HS) satisfying $|\dot{Q}(t)| \to 0$ as $|t| \to \infty$.

The definition of $\Gamma$ and the asymptotic behavior of $\dot{Q}$, will imply that $Q$ is an heteroclinic solution of (HS) as wanted. Let’s assume Proposition 3 for the moment, and prove Proposition 4.

**Proof:** (of Proposition 4)

Choose a pair of real numbers $r < s$ and a function $\varphi \in C^2(\mathbb{R}, \mathbb{R})$ such that $\varphi(t) = 0$ for $t \notin (r, s)$. If $\delta \in \mathbb{R}$ is arbitrary, the function $(Q + \delta \varphi)(t)$ behaves like $Q(t)$ for large values of $t$, since $\varphi$ has compact support. This implies, in particular, that $Q + \delta \varphi \in \Gamma$, and by definition of $Q$
\[ \psi(\delta) = I(Q) - I(Q + \delta \varphi) \leq 0 = \psi(0) \]
Note that $\psi$ is a continuously differentiable function of $\delta$, implying
\[ \psi'(0) = \int_r^s \left( \dot{Q} \cdot \varphi - V_q(t, Q) \cdot \varphi \right) dt = 0 \tag{25} \]
Let $x \in C^2([r, s])$ denote the unique solution of the linear problem
\[
\begin{align*}
\ddot{x} &= -V_q(t, Q) \\
x(s) &= Q(s) \\
x(r) &= Q(r)
\end{align*}
\]
and consider \( \varphi \) as above. Multiplying both sides of the equation by \( \varphi \), and integrate in \( t \in [r, s] \)
\[
\int_r^s \left( \dot{x} \cdot \varphi - V_q(t, Q) \cdot \varphi \right) dt = 0 \tag{26}
\]
Subtracting (25) of (26)
\[
\int_r^s (x - Q) \cdot \dot{\varphi} dt = 0
\]
In the previous examples we made the choice \( \varphi = x - Q \). We can’t do it in this case, since we don’t know if \( Q \in C^1([r, s]) \). Consider \( \varphi \) such that \( \ddot{\varphi} = x - Q \) in \( (r, s) \) and \( \varphi(r) = \varphi(s) = 0 \).
Integrating by parts
\[
0 = \int_r^s (x - Q) \cdot \ddot{\varphi} dt = -\int_r^s (x - Q) \cdot \dot{\varphi} dt = -\int_r^s |x - Q|^2 dt
\]
which implies that \( x = Q \) and therefore \( Q \in C^2((r, s)) \). Since \( r \) and \( s \) were chosen arbitrarily, \( Q \in C^2(\mathbb{R}, \mathbb{R}^n) \), and as a consequence \( Q \) is a classical solution of (HS). It remains to show that
\[
|\dot{Q}(t)| \to 0 \quad \text{as} \quad |t| \to \infty
\]
By Proposition 3,
\[
Q(-\infty) = 0 \quad \text{and} \quad Q(\infty) = \eta \in \mathbb{Z}^n \setminus \{0\}
\]
and by the first part of the proof, \( Q \) verifies (HS). Hence
\[
\lim_{|t| \to \infty} (\dot{Q} + V_q(t, Q)) = 0
\]
Since \( V_q(t, Q) \to 0 \) as \( |t| \to \infty \), we conclude
\[
\dot{Q}(\pm \infty) = 0
\]
To conclude the desired asymptotic behavior on \( \dot{Q} \) we need an interpolation result.

**Lemma 10.** Let \( \varphi \in C^2([a, b], \mathbb{R}^n) \). Then, for all \( \epsilon > 0 \), there exists \( k = k(\epsilon, b - a) \) such that
\[
\|\dot{\varphi}\|_{L^\infty([a, b])} \leq \epsilon \|\ddot{\varphi}\|_{L^\infty([a, b])} + k \|\varphi\|_{L^\infty([a, b])}
\]
If we assume the Lemma,
\[
\|\dot{Q}\|_{L^\infty([-1, 1])} \leq \epsilon \|\ddot{Q}\|_{L^\infty([-1, 1])} + k \|Q\|_{L^\infty([-1, 1])}
\]
Letting \( t \to -\infty \)
\[
\dot{Q}(t) \to 0 \quad \text{as} \quad t \to -\infty
\]
For the other side, apply Lemma 10 to \( \dot{Q}(t) = Q(t) - \eta \).

\[\diamond\]
Proof: (of Lemma 10)

Let $\epsilon > 0$ be arbitrary. Then there exists $n \in \mathbb{N}$ such that

$$\frac{\epsilon}{n} < \frac{|b-a|}{2}$$

Without loss of generality, take $n = 1$. For $x \in [a,b]$ choose $z \in [a,b]$ such that $|x-z| = \epsilon$. By the Mean value Theorem, there exists $y \in (x,z)$ (or $y \in (z,x)$) such that

$$\frac{\varphi(x) - \varphi(z)}{x-z} = \varphi'(y)$$

Then

$$\varphi'(x) = \varphi'(y) + \int_y^x \varphi''(t) dt = \frac{\varphi(x) - \varphi(z)}{x-z} + \int_y^x \varphi''(t) dt$$

Hence, for all $x$

$$|\varphi'(x)| \leq \frac{2}{\epsilon}\|\varphi''\|_{L^\infty[a,b]} + \epsilon\|\varphi''\|_{L^\infty[a,b]}$$

Taking the supremum over $x \in [a,b]$ we obtain the result.

Homework 16. Prove the analogous result for $L^2([a,b])$.

Before we prove Proposition 3, let us make some remarks. For $k \in \mathbb{Z}$ denote $\tau_k(q) = q(t-k)$. Then, as a consequence of the periodicity of $V$

$$I(\tau_k(q)) = I(q)$$

i.e., $I$ is invariant under time phase shifts. This has consequences for our problem.

Suppose that $Q$ is a critical point of $I$. Then $q_m = \tau_m(Q)$ is a minimizing sequence, since $I(q_m) = I(Q) = c$ for all $m \in \mathbb{Z}$. But notice that the sequence $\tau_m(Q) = Q(t-m)$ converges weakly to 0 as $m \to \infty$ and $0 \not\in \Gamma$, so we need to be very careful - we need to control the sequence so that this doesn’t happen.

With the invariance property we can normalize a minimizing sequence in an “good” form. If $Q \in \Gamma$, then $Q(-\infty) = 0$ and $Q(\infty) = \eta \in \mathbb{Z}^n \setminus \{0\}$. For $0 < \rho < \frac{1}{3}$, $Q$ will intersect $\partial B_\rho(0)$ at some $\hat{t}(Q)$. We will use this fact to normalize a minimizing sequence, which will imply, in particular, that the limit will not be the trivial solution.

Proof: (Proposition 3)

Let $(q_m) \subset \Gamma$ be a minimizing sequence. For $0 < \rho < \frac{1}{3}$, denote by $\hat{t}(q_m)$ the first time $q_m$ hits the boundary of $B_\rho(0)$. Choose $k_m \in \mathbb{Z}$, so that

$$\hat{t}(\tau_{k_m}(q_m)) \in [0,1) \quad \forall m$$

60
Without loss of generality, assume $k_m = 0$, for all $m$, i.e., assume that for all $m$, $q_m$ verifies
\[ |q_m(t)| < \rho \quad \forall t \in (-\infty, \hat{t}(q_m)) \]
and
\[ |q_m(\hat{t}(q_m))| = \rho \quad \text{for} \quad \hat{t}(q_m) \in [0, 1) \]
As usual, $I(q_m) \searrow c$ implies the existence of $M > 0$ for which $I(q_m) \leq M$. Thus
\[ M \geq I(q_m) \geq \int_{\mathbb{R}} \left( \frac{1}{2} |\dot{q}_m|^2 - V(t, q_m) \right) dt \]
and as a consequence of (V2)
\[ \|\dot{q}_m\|_{L^2}^2 \leq 2M \]
Then $(\dot{q}_m)$ is bounded in $L^2(\mathbb{R}, \mathbb{R}^n)$. On the other hand
\[ |q_m(0)| \leq |q_m(\hat{t}(q_m))| + |\int_{\hat{t}(q_m)}^0 \dot{q}_m(t) \, dt| \leq \rho + |\dot{t}(q_m)|^{1/2} \|\dot{q}_m\|_{L^2} \]
Since, by our normalization, $|\dot{t}(q_m)| < 1$, we conclude
\[ |q_m(0)| \leq \rho + \sqrt{2M} \]
Hence
\[ \|q_m\|_E^2 \leq 2\sqrt{2M} + \rho \]
as wanted. Then, there exists $Q \in E$ such that, possibly along a subsequence (again denoted by $q_m$), $q_m$ converges weakly to $Q$ in $E$. For $L > 0$, by definition, $(q_m)$ is bounded in $W^{1,2}([-L, L], \mathbb{R}^n)$. Then, as a consequence of the normalization
\[ q_m(t) - q_m(0) \leq \int_0^t |\dot{q}_m| \, dt \leq \rho + L^{1/2} \sqrt{2M} \]
This proves that $(q_m)$ is an equicontinuous family, which implies that $q_m \to Q$ in $L^\infty([-L, L], \mathbb{R}^n)$. Since $L$ is arbitrary, we can conclude that $q_m \to Q$ in $L^\infty_{\text{loc}}(\mathbb{R}, \mathbb{R}^n)$. The strong convergence implies that there exists $\hat{t} = \hat{t}(Q) \in [0, 1]$ such that $Q(\hat{t}) \in \partial B_\rho(0)$. Hence $Q(t)$ can not be the trivial solution. We will start by proving that $I(Q) < \infty$.

For $L \in \mathbb{R}^+$, since $q_m \to Q$ in $W^{1,2}([-L, L])$, and
\[ I(q)|_{-L}^L = \int_{-L}^L \mathcal{L}(q) \, dt \]
is weakly lower semicontinuous,
\[ \int_{-L}^L \mathcal{L}(Q) \, dt \leq \liminf_{m \to \infty} \int_{-L}^L \mathcal{L}(q_m) \, dt \leq \liminf_{m \to \infty} \int_{\mathbb{R}} \mathcal{L}(q_m) \, dt \leq M \]
Taking the limit as $L \to \infty$, we conclude $I(Q) \leq M < \infty$ as wanted. Next we want to show that $Q \in \Gamma$, i.e., that

$$Q(-\infty) = 0 \quad \text{and} \quad Q(\infty) = \eta \in \mathbb{Z}^n \setminus \{0\}$$

For $\rho$ as in the normalization, choose $0 < \epsilon < \rho$, and set

$$\beta(\epsilon) = \inf_{x \notin B_\epsilon(\mathbb{Z}^n), t \in \mathbb{R}} (-V(t, x))$$

As an immediate consequence of condition (V2), $\beta(\epsilon)$ is strictly positive.

**Lemma 11.** If $w \in E$, consider real numbers $\sigma < s$ such that $w(t) \notin B_\epsilon(\mathbb{Z}^n)$ for $t \in [\sigma, s]$. Then

$$\int_\sigma^s \mathcal{L}(w) \ dt \geq \sqrt{2\beta(\epsilon)} |w(\sigma) - w(s)|$$

**Proof:**

Let $l \equiv |w(\sigma) - w(s)|$. Then

$$l = |\int_\sigma^s \dot{w}(t) \ dt| \leq |s - \sigma|^{1/2} \left( \int_\sigma^s |\dot{w}(t)|^2 \ dt \right)^{1/2}$$

implying

$$\int_\sigma^s \mathcal{L}(w) \ dt = \int_\sigma^s \left( \frac{1}{2} |\dot{w}(t)|^2 - V(t, w) \right) \ dt \geq \frac{l^2}{2|s - \sigma|} + \beta(\epsilon) |s - \sigma|$$

For $z = |\sigma - s|$, we have

$$\int_\sigma^s \mathcal{L}(w) \ dt \geq \frac{l^2}{2z} + \beta(\epsilon) z \equiv \varphi(z)$$

It is easy to prove that $\varphi$ has a minimum value achieved at $z = \frac{l}{\sqrt{2\beta(\epsilon)}}$. Then

$$\int_\sigma^s \mathcal{L}(w) \ dt \geq \varphi(\frac{l}{\sqrt{2\beta(\epsilon)}}) = \sqrt{2\beta(\epsilon)} l$$

as wanted.

\[ \diamond \]

**Corollary 6.** If $I(w) < \infty$, then $w \in L^\infty(\mathbb{R}, \mathbb{R}^n)$.

**Homework 17.** Prove Corollary 6.

**Lemma 12.** If $w \in E$ and $I(w) < \infty$, then $w(\pm \infty) \in \mathbb{Z}^n$.

**Proof:**

We will prove $w(-\infty) \in \mathbb{Z}^n$, that $w(\infty) \in \mathbb{Z}^n$ follows in a similar form. By Corollary 6, $w \in L^\infty(\mathbb{R}, \mathbb{R}^n)$. Then both $\limsup_{t \to -\infty} w(t)$ and $\liminf_{t \to -\infty} w(t)$ exist, and as a consequence, there
exists a sequence $t_m \to -\infty$ and $\xi \in \mathbb{R}^n$ such that $w(t_m) \to \xi$ as $m \to \infty$. Suppose for the moment that we can prove

$$t \to -\infty \implies w(t) \to \xi$$

(27)

Then $\xi$ must necessarily belong to $\mathbb{Z}^n$. If not, for large values of $t$, $V(t, w)$ is strictly positive, i.e., there are strictly positive constants $L$ and $\gamma$ such that

$$-V(t, w) > \gamma, \quad \forall t \leq -L$$

As a consequence

$$I(w) > \int_{\mathbb{R}} -V(t, w) \, dt > \int_{-\infty}^{-L} \gamma \, dt = \infty$$

contradicting the hypothesis. All we have to prove is that (27) holds. Once again arguing by contradiction, if not, there exists $\delta > 0$ and sequences $(t_m), (s_m) \subset \mathbb{R}$, converging to $-\infty$ as $m \to \infty$, and such that $w(t_m) \in \partial B_\delta(\xi)$ and $w(s_m) \in \partial B_{2\delta}(\xi)$.

Without loss of generality, we may assume that $0 < \delta < \frac{\rho}{3}$, so that

$$B_{2\delta}(\xi) \setminus B_\delta(\xi) \cap \mathbb{Z}^n = \emptyset$$

This is obviously true for $\xi \in \mathbb{Z}^n$, and for $\xi \notin \mathbb{Z}^n$ we can take $\delta$ as small as necessary. Furthermore, we may also assume that $s_m \in (t_m, t_{m+1})$, and for $t \in (t_m, s_m)$ require that

$$w(t) \in B_{2\delta}(\xi) \setminus B_\delta(\xi)$$

(refine the sequence if necessary), i.e., for $t \in (t_m, s_m)$, $w(t)$ does not belong to $B_\delta(\xi)$. Thus, we are in the conditions of Lemma 12, and therefore

$$\infty > I(w) \geq \sum_m \int_{t_m}^{s_m} \mathcal{L}(w) \, dt \geq \sum_m \sqrt{2\beta(\delta)}\delta = \sqrt{2\beta(\delta)}\delta(\#\text{terms}) = \infty$$

which is a contradiction.

$$\diamondsuit$$

We will proceed the proof of the Theorem. By the previous Lemma, $Q(-\infty) \in \mathbb{Z}^n$. On the other hand by the normalization

$$Q(t) \in \overline{B_\rho(0)}, \quad \forall t < 0$$

Then

$$Q(-\infty) \in \mathbb{Z}^n \cap \overline{B_\rho(0)} = \{0\}$$

and by our choice of $\rho$, we must necessarily have $Q(-\infty) = 0$. Also as a consequence of Lemma 12, $Q(\infty) \in \mathbb{Z}^n$. But we have to be careful, since $Q(\infty) = 0$ (which may happen if the
minimizing sequence breaks down in the limit). Consider \( 0 < \delta \ll \rho \) and suppose \( Q(\infty) = 0 \). Then, there exists \( T = T(\delta) > 0 \) such that \( Q(t) \in B_\delta(0) \) for all \( t \geq T \). Since \( q_m \to Q \) in \( L^\infty_{\text{loc}}(\mathbb{R}, \mathbb{R}^n) \), for \( m \) sufficiently large \( q_m(t) \in B_{2\delta}(0) \) for \( t \in [T, T + 1] \). We will prove that this yields a contradiction (the intuitive idea is that we can cut the extra part \( q_m|_{-\infty}^T \) in \( (q_m) \) and construct a new minimizing sequence \( v_m \) with \( I(v_m) < I(q_m) \).) Define

\[
v_m(t) = \begin{cases} 
0 & \text{if } t \leq T \\
(t - T)q_m(T + 1) & \text{if } t \in [T, T + 1] \\
q_m(t) & \text{if } t \geq T + 1 
\end{cases}
\]

It is easy to check that \( v_m \in \Gamma \) and

\[
I(v_m) - I(q_m) = \int_T^{T+1} \mathcal{L}(v_m) \, dt - \int_{-\infty}^{T+1} \mathcal{L}(q_m) \, dt
\]

\[
= \int_T^{T+1} \left( \frac{1}{2} |q_m(T + 1)|^2 - V(t, v_m) \right) \, dt - \int_{-\infty}^{T+1} \mathcal{L}(q_m) \, dt
\]

By our assumptions

\[
\int_T^{T+1} \frac{1}{2} |q_m(T + 1)|^2 \, dt = 2\delta^2
\]

and for \( t \in [T, T + 1] \)

\[
|v_m(t)| \leq \delta
\]

which implies, by the properties of \( V \),

\[
\int_T^{T+1} -V(t, v_m) \, dt \leq \epsilon(\delta)
\]

where \( \epsilon(\delta) \to 0 \) as \( \delta \to 0 \). On the other hand, for \( t \in (-\infty, T + 1] \), \( q_m(t) \) starts at 0 for \( t = -\infty \), hits the boundary of \( B_\rho(0) \) for some \( t \in [0, 1] \), gets back to \( B_{2\delta}(0) \) for \( t \in [T, T + 1] \) and goes back to another point in the lattice. So it must cross at least twice, the boundary of \( B_\rho(0) \setminus B_{\rho/2}(0) \) (remember that \( \delta \ll \rho \) so we may assume that \( 2\delta < \rho/2 \)). By Lemma 11

\[
\int_{-\infty}^{0} \mathcal{L}(q_m) \, dt \geq 2\sqrt{2\beta}\rho/2 = \rho\sqrt{2\beta}(\rho/2)
\]

Finally

\[
I(v_m) - I(q_m) \leq 2\delta^2 + \epsilon(\delta) - \rho\sqrt{2\beta}(\rho/2)
\]

and we can choose \( \delta \) sufficiently small, so that

\[
2\delta^2 + \epsilon(\delta) < \frac{\rho}{2}\sqrt{2\beta}(\rho/2)
\]

and therefore

\[
I(v_m) - I(q_m) < 0
\]

64
contradicting the fact that $q_m$ is a minimizing sequence of $c = \inf_{\Gamma} I$. Therefore $Q(\infty) \neq 0$, and as a consequence there exists $\eta \in \mathbb{Z}^n \setminus \{0\}$ such that $Q(\infty) = 0$. Then $Q \in \Gamma$. If this is the case

$$I(Q) \geq c$$  \hspace{1cm} (28)

For $\epsilon > 0$ and $m$ large, using the definition of infimum

$$I(q_m) \leq c + \epsilon$$

Then

$$c + \epsilon \geq I(q_m) \geq \int_{-L}^{L} \mathcal{L}(q_m) \, dt$$

for all $L > 0$. Letting $m \to \infty$, and as a consequence of the weakly lower semicontinuity

$$c + \epsilon \geq \int_{-L}^{L} \mathcal{L}(Q) \, dt$$

for all $L \in \mathbb{R}^+$ and $\epsilon > 0$. Finally, letting $L \to \infty$ and $\epsilon \to 0$

$$I(Q) \leq c$$  \hspace{1cm} (29)

Conditions (28) and (29) imply $I(Q) = c$.

Using the same type of arguments, we can prove that this is not the only heteroclinic solution of (HS) leaving 0.

**Theorem 11.** If $V$ satisfies (V1)-(V2), for any $\xi \in \mathbb{Z}^n$, there are distinct $\eta_0, \eta_1, \ldots, \eta_n \in \mathbb{Z}^n \setminus \{0\}$, and at least $Q_0, \ldots, Q_n$ solutions of (HS), where $Q_i$ is heteroclinic from $\xi$ to $\eta_i$, $0 \leq i \leq n$.

In other words, the Theorem states that there are at least $n+1$ distinct heteroclinics leaving each point in $\mathbb{Z}^n$.

**Proof:**

Once again we will assume that $\xi = 0$. By Theorem 10, there exists $Q_0$ and $\eta_0$ as wanted. Define the new class

$$\Gamma_1 = \{ q \in \Gamma : q(\infty) \not\in \text{span}_\mathbb{Z}\eta_0 \}$$

i.e., the class of curves in $W^{1,2}_{\text{loc}}(\mathbb{R}, \mathbb{R}^n)$, with $q(-\infty) = 0$ and such that $q(\infty)$ is not any multiple of $\eta_0$. Define

$$c_1 = \inf_{\Gamma_1} I(q)$$

We claim that there is a solution of (HS) which minimizes $I$ over $\Gamma_1$. Arguing as before, if $(q_m)$ is a minimizing sequence, $(q_m)$ is bounded in $E$ (after normalization). Then there exists $Q_1 \in E$ such that $q_m \to Q_1$ weakly in $W^{1,2}_{\text{loc}}(\mathbb{R}, \mathbb{R}^n)$ and strongly in $L^\infty_{\text{loc}}(\mathbb{R}, \mathbb{R}^n)$. To show that in fact $Q_1 \in \Gamma_1$, as last time it is easy to show that $Q_1(-\infty) = 0$. The possible dangers are
- \( Q_1(\infty) = 0 \) which will be excluded as in last time;

- \( Q_1(\infty) \in \text{span}_N\eta_0 \). For simplicity assume that \( Q_1(\infty) = \eta_0 \). For \( 0 < \delta << \rho \), there is \( T = T(\delta) \) for which \( |Q_1(t) - \eta_0| \leq \delta \) for all \( t \geq T(\delta) \). As a consequence of the strong convergence, \( |q_m(t) - \eta_0| \leq 2\delta \) for \( t \in [T, T+1] \) and \( m \) sufficiently large. If this is the case, define

\[
    p_m(t) = \begin{cases} 
    0 & \text{if } t \leq T \\
    (T - t)(q_m(T + 1) - \eta_0) & \text{if } T \leq t \leq T + 1 \\
    q_m(t) - \eta_0 & \text{if } t \geq T + 1 
    \end{cases}
\]

(i.e., we are considering the piece of \( q_m \), from \( \eta \) to \( q_m(\infty) = \eta_1 \) translated by \( \eta_1 \)). Then \( p_m : 0 \rightarrow \eta_1 - \eta_0 \not\in \text{span}_N\eta_0 \), which implies that \( p_m \in \Gamma_2 \), and as before we can take \( \delta \) so small so that \( I(p_m) < I(q_m) \) contradicting the fact that \( q_m \) is a minimizing sequence.

So we produced another heteroclinic solution from \( 0 \) to \( \eta_1 \not\in \text{span}_N\eta_0 \). Why did we cut out all \( \text{span}_N\eta_0 \) and not only \( \eta_0 \)? If we just cut out \( \eta_0 \) the limit of the minimizing sequence may be just a heteroclinic chain.

We proceed inductively. Define

\[
    \Gamma_2 = \{ q \in \Gamma_1 : q(\infty) \not\in \text{span}_N\eta_1 \}
\]

and

\[
    c_2 = \inf_{\Gamma_2} I(q)
\]

and we can continue this process until we span \( \mathbb{Z}^n \), so we obtain \( n + 1 \) distinct heteroclinics.

**Remark:** Choose a pair of points \( 0 \) and \( \xi \). Can we find an heteroclinic solution such that \( Q(-\infty) = 0 \) and \( Q(\infty) = \xi \)? We can’t prove the existence of such heteroclinic, but we can prove that there is at least an heteroclinic chain - the infimum may not be achieved.

### 5.3 Existence of Homoclinic Solutions for a Singular Hamiltonian System

Consider

\[
    \ddot{q} + V_q(t, q) = 0 \quad (HS)
\]

where \( V \) satisfies the following

(V1) There exists \( \xi \in \mathbb{R}^2 \setminus \{0\} \) such that \( V = V(t, q) \in C^2(\mathbb{R} \times \mathbb{R}^2 \setminus \{\xi\}, \mathbb{R}) \), and \( V \) is 1-periodic in \( t \);

(V2) \( \lim_{x \to \xi} V(t, x) = -\infty \) uniformly in \( t \in [0, 1] \);

(V3) **Strong Force Potential:** There exists a neighborhood of \( \xi, \mathcal{N} \), and \( U \in C^1(\mathcal{N} \setminus \{\xi\}, \mathbb{R}) \) such that
\[ |U(x)| \to \infty \text{ as } x \to \xi \]
\[ |U'(x)|^2 \leq -V(t,x) \text{ uniformly in } t \in [0,1) \text{ and } x \in N \setminus \{\xi\}. \]

(V4) \( V(t,0) = 0 > V(t,x) \) for all \( x \neq 0 \)

(V5) There exists \( V_0 < 0 \) such that \( \limsup_{|x| \to \infty} V(t,x) \leq V_0. \)

Condition (V5) implies that the potential is strictly negative at infinity, and therefore 0 is a global maximum of \( V. \)

**Theorem 12.** If \( V \) satisfies (V1)-(V5), there exists a pair of solutions of (HS), \( q^+ \) and \( q^- \), which are homoclinic to 0 and wind, respectively positively and negatively around \( \xi. \)

We remark that, if \( V \) is independent of \( t \), \( q^+ \) and \( q^- \) represent the same curve runned in opposite directions.

We will prove the Theorem by minimizing

\[ I(q) = \int_\mathbb{R} \left( \frac{1}{2} |\dot{q}|^2 - V(t,q) \right) dt \geq 0 \]

in an appropriate subclass of \( E = W^{1,2}_{\text{loc}}(\mathbb{R}, \mathbb{R}^2). \) Consider

\[ \Lambda = \{ q \in E : q(\pm \infty) = 0 \text{ and } q(t) \neq \xi \forall t \in \mathbb{R} \} \]

**Proposition 5.** If \( I(q) < \infty \), there exists \( \gamma > 0 \) such that

\[ |q(t) - \xi| > \gamma, \quad \forall t \in \mathbb{R} \]

Notice that \( \gamma \) depends only on the bounds for \( I(q). \)

**Proposition 6.** If \( I(q) < M \) then \( q \in L^\infty(\mathbb{R}, \mathbb{R}^2) \) and \( q(t) \to 0 \) as \( |t| \to \infty. \)

**Homework 18.** Prove Proposition 6.

Choose \( q \in \Lambda \), such that \( I(q) < \infty. \) By Proposition 6 the Range of \( q \) is a closed curve, and by Proposition 5 it avoids \( \xi. \) Then it makes sense in considering the winding number of \( q \) with respect to \( \xi. \) By definition

\[ \WN_{\xi}(q) = \frac{1}{2\pi} (\Theta(\infty) - \Theta(-\infty)) \]

where we are considering

\[ q(t) = R(t)e^{i\Theta(t)} \]

a representation of \( q \) in polar coordinates with respect to \( \xi. \) Let

\[ \Gamma^\pm = \{ q \in \Lambda : \pm \WN_{\xi}(q) > 0 \} \]

and

\[ c^\pm = \inf_{q \in \Gamma^\pm} I(q) \]

The proof of the Theorem follows from the next two Propositions.

67
Proposition 7. $c^+ > 0$ and there exists $Q^\pm \in \Gamma^\pm$ such that $I(Q^\pm) = c^\pm$.

Proposition 8. $Q^\pm$ is a solution of (HS), verifying $|\dot{Q}^\pm(\pm\infty)| = 0$.

The proof of Proposition 8 follows like the analogous results in our previous examples, so we will omit it. The variational characterization of $Q$ and Proposition 8 prove Theorem 12.

Proof:(Proposition 7)

We will prove the existence of $Q^+ \in \Gamma^+$ with $I(Q^+) = c^+$. The proof for the existence of $Q^-$ is similar. Let $(q_m)$ be a minimizing sequence for $\Gamma^+$. Then there exists $M > 0$ for which $I(q_m) \leq M$, which give us bounds for $\|q_m\|_{L^2(\mathbb{R}, \mathbb{R}^2)}$ uniformly for $m \in \mathbb{N}$. As before we can normalize the minimizing sequence, since we are working with an unbounded domain. Let

$$S = \{\theta \xi : \theta \geq 1\}$$

i.e., $S$ is the ray emanating from $\xi$ and running to $\infty$. If $q \in \Gamma^+$, it must intersect $S$, at least once. Replacing $q_m$ by $\tau_{k_m}(q_m) = q_m(t - k_m)$, $k_m \in \mathbb{Z}$, we can choose

$$\begin{cases} q_m((-\infty, 0) \cap S = \emptyset \\ q_m([0, 1)) \cap S \neq \emptyset \end{cases}$$

that is, we can choose the minimizing sequence, so that, for every $m \in \mathbb{N}$, the first time $q_m$ intersects $S$ is in the unit interval. Moreover, Proposition 6 implies that $\|q_m\|_{L^\infty(\mathbb{R}, \mathbb{R}^2)} \leq k(M)$. Then

$$\|q_m\|^2_E = \|\dot{q}_m\|^2_{L^2(\mathbb{R}, \mathbb{R}^2)} + |q_m(0)|^2 \leq \|\dot{q}_m\|^2_{L^2(\mathbb{R}, \mathbb{R}^2)} + \|q_m\|^2_{L^\infty(\mathbb{R}, \mathbb{R}^2)} \leq M_1$$

and as a consequence, there exists $Q$ in $E$ such that $q_m \rightharpoonup Q$ weakly in $W^{1,2}_{\text{loc}}(\mathbb{R}, \mathbb{R}^2)$ and strongly in $L^\infty_{\text{loc}}(\mathbb{R}, \mathbb{R}^2)$. The strong convergence of $q_m$ to $Q$ in bounded intervals, imply

$$Q([0, 1]) \cap S \neq \emptyset$$

Using the weakly lower semicontinuity, it is easy to show that, for all $L \in \mathbb{R}^+$,

$$\int_{-L}^L \mathcal{L}(Q) \ dt \leq \liminf_{m \to \infty} I(q_m) \leq M$$

Letting $L \to \infty$, we conclude that $I(Q)$ is finite. Proposition 5 implies that $Q(t) \neq \xi$, for all $t \in \mathbb{R}$, and by Proposition 6, $Q(\pm \infty) = 0$. Then $Q \in \Lambda$. To show that $Q \in \Gamma^+$, it remains to prove that $\text{WN}_\xi(Q) > 0$. Once this is proved, it is easy to show that $I(Q) = c^+$, which will finish the proof. Arguing indirectly, suppose that $\text{WN}_\xi(Q) \leq 0$. Since $Q(\pm \infty) = 0$, for $\delta > 0$, there exists $T = T(\delta) > 0$, such that

$$|Q(t)| \leq \frac{\delta}{2}, \quad \forall |t| \geq T - 1$$

Then, for $\epsilon = \epsilon(\delta) \to 0$ as $\delta \to 0$,

$$-\frac{\epsilon}{8} \leq \text{WN}_\xi(Q)|_{-\infty}^{T} \leq \frac{\epsilon}{8}$$

68
and
\[-\frac{\epsilon}{8} \leq \mathrm{WN}_\xi(Q)^\infty_T \leq \frac{\epsilon}{8}\]

As a consequence
\[
\mathrm{WN}_\xi(Q) = \mathrm{WN}_\xi(Q)^{-T} + \mathrm{WN}_\xi(Q)^{T} + \mathrm{WN}_\xi(Q)^{-\infty_T} \leq 0
\]

implies
\[
\mathrm{WN}_\xi(Q)^{T} \leq \frac{\epsilon}{4}
\]

Since \(q_m \to Q\) in \(L^\infty_{\text{loc}}(\mathbb{R}, \mathbb{R}^2)\), \(q_m\) converges uniformly to \(Q\) in \([-T, T]\), and therefore, for \(m\) sufficiently large
\[
\mathrm{WN}_\xi(q_m)^{-T} \leq \frac{\epsilon}{2}
\]

On the other hand, since by the normalization, \(q_m((\infty, -T)) \cap S = \emptyset\), we may also assume (since in this interval there is not much change in the angle)
\[
\mathrm{WN}_\xi(q_m)^{-\infty_T} < \frac{\epsilon}{8}.
\]

Finally, by construction \(q_m \in \Gamma^+\), for all \(m \in \mathbb{N}\), and therefore
\[
\mathrm{WN}_\xi(q_m) \geq 1
\]

for all \(m \in \mathbb{N}\). Putting all this information together, we can conclude
\[
\mathrm{WN}_\xi(q_m)^{\infty_T} \geq 1 - \frac{\epsilon}{2} - \frac{\epsilon}{8} \geq 1 - \frac{5\epsilon}{8} \approx 1
\]

We will prove that this yields a contradiction (intuitively, \((q_m)^{\infty_T}\) does the job of a minimizing sequence by itself, so we can cut the extra part \(q_m^{-T}\) and obtain \((p_m)\) with \(I(p_m) < I(q_m)\)). Consider the comparison function
\[
p_m(t) = \begin{cases} 
0 & \text{if } t \leq T - 1 \\
(t - (T - 1))q_m(T) & \text{if } T - 1 \leq t \leq T \\
q_m(t) & \text{if } t \geq T 
\end{cases}
\]

By construction, \(\mathrm{WN}_\xi(q_m)^{\infty_T} \geq 1\), which implies that \(p_m \in \Gamma^+\), for all \(m \in \mathbb{N}\). On the other hand
\[
I(p_m) - I(q_m) = \int_{T-1}^{T} \left(\frac{1}{2}[q_m(T)]^2 - V(t, p_m)\right)dt - \int_{\infty}^{T} \mathcal{L}(q_m) dt
\]

For \(t \in (\infty, T)\), \(q_m(t)\) starts at 0 for \(t = -\infty\), touches \(S\) for some \(t \in [0, 1]\) and returns to \(B_\delta(0)\) at time \(T\). Thus
\[
q_m(t) \in B_{\xi\delta}(0) \setminus B_{\xi\delta/2}(0)
\]

for \(t \in [\sigma, \tau] \subset (\infty, T)\). This implies
\[
\int_{-\infty}^{T} \mathcal{L}(q_m) dt > \int_{\sigma}^{T} \mathcal{L}(q_m) dt \geq \sqrt{2\beta(|\xi|/2)|\xi|/2} \equiv \gamma > 0
\]

69
On the other hand
\[ \int_{T-1}^{T} \frac{1}{2} |q_m(T)|^2 dt = \frac{\delta^2}{2} \]
and there exists \( \dot{\epsilon} = \dot{\epsilon}(\delta) \to 0 \) as \( \delta \to 0 \), such that
\[ \int_{T-1}^{T} -V(t', p_m) dt' \leq \dot{\epsilon} \]
Finally
\[ I(p_m) - I(q_m) \leq \frac{\delta^2}{2} + \dot{\epsilon} - \gamma \]
and we can take \( \delta \) sufficiently small and \( m \) large, so that
\[ \frac{\delta^2}{2} + \dot{\epsilon} < \frac{\gamma}{2} \]
Then
\[ I(p_m) - I(q_m) \leq -\frac{\gamma}{2} < 0 \]
contradicting the definition of \( q_m \). Hence \( WN_\xi(Q) > 0 \) as wanted, and we finished the proof.

5.4 Existence of Heteroclinic Solutions to Periodics

Consider
\[ \ddot{q} + V_q(t, q) = 0 \quad (HS) \]
where

(V1) \( V = V(t, q) \in C^2(\mathbb{R} \times \mathbb{R}^n, \mathbb{R}) \) is 1-periodic in \( t \) and in \( q_1, ..., q_n \).

We proved that, for such a potential, there exists at least a solution of (HS) which is 1-periodic. We obtained it, by minimizing
\[ I_1(q) = \int_{0}^{1} \mathcal{L}(q) \, dt \]
over \( E_1 = W^{1,2}_1(\mathbb{R}, \mathbb{R}^n) \). Let
\[ c_1 = \inf_{E_1} I_1 \]
and define
\[ K_1 = \{ q \in E_1 : I_1(q) = c_1 \} \]
As mentioned, \( K_1 \) is nonempty, and as a consequence of (V1), if \( v \in K_1 \) then \( v + j \in K_1 \) for all \( j \in \mathbb{Z}^n \), i.e. \( K_1 \) contains a lattice of 1-periodic solutions. We are interested in finding solutions of (HS) which are heteroclinic to these 1-periodic solutions. We need some extra assumptions to prove their existence.
(V2) \( V \) is time reversible, i.e.,
\[
V(-t, x) = V(t, x) \quad \forall t \in \mathbb{R}, \; x \in \mathbb{R}^n
\]

We will start by showing some technical results, which are consequence of condition (V2).

**Lemma 13.** If (V1) and (V2) are satisfied then
\[
c_1 = 2 \inf_{q \in W^{1,2}([0,1],\mathbb{R}^n)} \int_0^{1/2} \mathcal{L}(q) \, dt
\]

Moreover, if \( v \in \mathcal{K}_1 \) then
\[
v(-t) = v(t) \quad \forall t \in \mathbb{R}
\]

**Proof:**
Choose \( q \in E^{1/2} \equiv W^{1,2}([0,1],\mathbb{R}^n) \), and define \( w \) by
\[
w(t) = \begin{cases} 
q(t) & t \in [0, \frac{1}{2}] \\
q(1 - t) & t \in [\frac{1}{2}, 1]
\end{cases}
\]
Extending \( w \) periodically to \( \mathbb{R} \), and denoting the extension again by \( w \), it is obvious that \( w \in E_1 \). Furthermore
\[
\int_{1/2}^{1} \mathcal{L}(w) \, dt = \int_{1/2}^{1} \frac{1}{2} |q(1 - t)|^2 - V(t, q(1 - t)) \, dt
\]
\[
= \int_0^{1/2} \frac{1}{2} |q(s)|^2 - V(1 - s, q) \, ds
\]

As a consequence of (V1) and (V2), we can conclude
\[
\int_{1/2}^{1} \mathcal{L}(w) \, dt = \int_0^{1/2} \mathcal{L}(q) \, dt
\]
and therefore
\[
I_1(w) = 2 \int_0^{1/2} \mathcal{L}(q) \, dt
\]
Thus, for all \( q \in E^{1/2} \) there exists \( w \in E_1 \) for which
\[
I_1(w) = 2 \int_0^{1/2} \mathcal{L}(q) \, dt
\]
Taking the infimum over the appropriate sets
\[
2 \inf_{q \in E^{1/2}} \int_0^{1/2} \mathcal{L}(q) \, dt = \int_0^{1} \mathcal{L}(w) \, dt \geq \inf_{E_1} \int_0^{1} \mathcal{L}(w) \, dt = c_1
\]
On the other hand, for any $v \in K_1$

$$c_1 = I_1(v) = \int_0^{1/2} \mathcal{L}(v) \, dt + \int_{1/2}^1 \mathcal{L}(v) \, dt \equiv a + b$$

If $a \leq b$ (implying $\frac{a}{2} \leq a$), we can define

$$w(t) = \begin{cases} v(t) & \text{if } t \in [0, \frac{1}{2}] \\ v(1 - t) & \text{if } t \in [\frac{1}{2}, 1] \end{cases}$$

i.e., we are taking the restriction of $v$ to the interval where $\min\{a, b\}$ is achieved and extend it to the other interval as an even function. As a consequence

$$2I_1(w|_{[0, \frac{1}{2}]} \leq c_1$$

Taking the infimum over $E^{1/2}$, equality is achieved in (30) as wanted. It remains to prove that any element of $K_1$ is an even function. For $v, w$ as above, since they both minimize $I_1$, they are solutions of (HS). Moreover, by construction, they agree in $[0, \frac{1}{2}]$ or in $[\frac{1}{2}, 1]$. In either case, by the uniqueness of solution for the associated IVP, they must equal in $[0, 1]$. Since $w$ is an even function about $\frac{1}{2}$, also

$$v(1 - t) = v(t)$$

But $v$ is 1-periodic, which implies

$$v(-t) = v(t), \quad \forall t \in \mathbb{R}$$

as wanted.

\[\diamondsuit\]

**Corollary 7.** If $(V1)$ and $(V2)$ are satisfied

$$\overline{\tau}_1 \equiv \inf_E I_1 = c_1$$

**Proof:**

Since $E$ is a bigger space, we must necessarily have $\overline{\tau}_1 \leq c_1$. We want to show that equality holds, so let us assume that $\overline{\tau} < c_1$ and reach a contradiction. If that is the case, there exists $q \in E$ such that $I_1(q) < c_1$. We can write $I_1(q)$ as

$$I_1(q) = \int_0^{1/2} \mathcal{L}(q) \, dt + \int_{1/2}^1 \mathcal{L}(q) \, dt \equiv a + b$$

Defining $w$ as the restriction to the interval where $\min\{a, b\}$ is achieved, extend it to the other interval as an even function about $1/2$, and further extend it periodically to $\mathbb{R}$, we construct $w \in E_1$. But

$$I_1(w) \leq I_1(q) < c_1 \equiv \inf_{E_1} I_1$$

72
which contradicts the definition of infimum.

We will now make another assumption, in order to restrict the set $K_1$—it may be a very large set, containing a continuous of solutions.

(*) $K_1$ consists of isolated points.

How reasonable is this condition? Suppose $v \in K_1$ and consider (HS), where $V$ is replaced by $W(t, x) = V(t, x) - \epsilon|x - v(t)|^2$. Then

$$I_1(q) = \int_0^1 \left( \frac{1}{2} |q(t)|^2 - V(t, q) + \epsilon|q - v(t)|^2 \right) dt$$

Notice that the extra term $\epsilon|q - v|^2$ will make the distinction between the elements of $K_1$, but we lost periodicity.

**Homework 19.** Use this idea to show that there is a function $V$ verifying (V1), (V2) and (*).

We can normalize any element $v \in K_1$, in a way that

$$0 \leq [v_i] < 1 \quad 1 \leq i \leq n$$

**Lemma 14.** If (V1), (V2) and (*) are verified, there is only a finite number of normalized functions in $K_1$.

**Proof:**

If not, there would be infinitely many, and therefore we could find a sequence of distinct elements $(w_m) \subset K_1$, verifying

$$0 \leq [(w_m)_i] < 1 \quad 1 \leq i \leq n$$

Since $I_1(w_m) = c_1$, there exists $M = M(c_1, \max V)$, such that

$$\|\dot{w}_m\|_{L^2([0,1],\mathbb{R}^n)} \leq M$$

Then, we can conclude that $(w_m)$ is bounded in $W^{1,2}([0,1],\mathbb{R}^n)$, and by Sobolev inequality, there exists $M_1 = M_1(M)$ such that

$$\|w_m\|_{L^\infty_{[0,1]}} \leq M_1$$

On the other hand, for all $m \in \mathbb{N}, w_m$ is a solution of (HS), and therefore

$$\dot{w}_m = -V_q(t, w_m)$$
By (V1), $V_q$ is continuous, and since $(w_m)$ is uniformly bounded, there exists $M_2$ such that 
\[ \|w_m\|_{L^\infty([0,1],\mathbb{R}^n)} \leq M_2 \]
Finally, by interpolation, there exists $M_3$ such that 
\[ \|w_m\|_{L^\infty([0,1],\mathbb{R}^n)} \leq M_3 \]
Then $(w_m)$ is bounded uniformly in $C^2([0,1],\mathbb{R}^n)$. By Ascoli-Arzela Theorem, it must converge in $C^1([0,1],\mathbb{R}^n)$, possibly along a subsequence, to another point in $\mathcal{K}_1$ contradicting (*).

As an immediate consequence we can state the following result.

**Lemma 15.** If (V1)-(V2) and (*) holds, there exists $b > 0$ such that 
\[ |v(0) - w(0)| > b \]
for normalized $v, w \in \mathcal{K}_1$ with $v \neq w$

**Proof:**

Notice that we can have $v(0) - w(0) = 0$. If that was the case, since they are both even functions, their derivatives will be odd functions, and as a consequence $\dot{v}(0) = \dot{w}(0)$. But then, both $v$ and $w$ are solutions of the IVP 
\[
\begin{cases}
\ddot{q} + V_q(t, q) = 0 \\
q(0) = 0 = \dot{q}(0)
\end{cases}
\]
and by uniqueness of solution, $v \equiv w$ which contradicts the fact that $v \neq w$. So, the only possible situation where we can’t find such $b$ is if we have a sequence $(v_m) \subset \mathcal{K}_1$, such that 
\[ |v_m(0) - v_n(0)| \to 0 \quad \text{as} \quad m, n \to \infty \]
Using familiar arguments, we can prove that $v_m$ converges for some $v \in \mathcal{K}_1$, violating condition (*).

**Corollary 8.** If 
\[ \gamma = \inf \{ \|v - w\|_{L^\infty}, \ v \neq w \in \mathcal{K}_1 \} \]
then $\gamma \geq b$.

This is obvious, since 
\[
\gamma = \inf \{ \|v - w\|_{L^\infty}, \ v \neq w \in \mathcal{K}_1 \} = \inf_{v,w} \sup_t \ |v(t) - w(t)| \geq \inf_{v,w} |v(0) - w(0)| \geq b
\]
For $v \in \mathcal{K}_1$ and $r > 0$, denote 
\[ B_r(v) = \{ q \in E : \|q - v\|_{E_1} < r \} \]
Proposition 9. If (V1)-(V2) and (*) hold, for any \(0 < \rho < \frac{\gamma}{4}\), there exists \(\alpha = \alpha(\rho) > 0\), such that

(1) \(B_\rho(v) \cap B_\rho(w) = \emptyset\) if \(v \neq w\) in \(\mathcal{K}_1\).

(2) \(I_1(u) > c_1\) for \(u \in B_\rho(\mathcal{K}_1) \setminus \mathcal{K}_1\).

(3) \(I_1(u) \geq c_1 + \alpha\) if \(u \in E \setminus B_\rho(\mathcal{K}_1)\).

Proof:

(1) Is just a consequence of Corollary 8.

(2) By definition of \(\mathcal{K}_1\) and Corollary 8.

(3) Suppose this is not the case. Then, for all \(\alpha \geq 0\),

\[ I_1(u) \leq c_1 + \alpha \]

and as a consequence, there exists \((u_m) \subset E \setminus B_\rho(\mathcal{K}_1)\) such that

\[ I_1(u_m) \to c_1 \quad m \to \infty \]

As before, we can assume that, for all \(m \in \mathbb{N}\)

\[ 0 \leq |u_{m_i}| < 1 \quad 1 \leq i \leq n \]

and, we can conclude that \((u_m)\) is a bounded sequence in \(E\), and therefore, possibly along a subsequence, it converges weakly in \(E\) and strongly in \(L^\infty_{\text{loc}}(\mathbb{R}, \mathbb{R}^n)\) to some \(u \in E\). Since \(I_1\) is weakly lower semicontinuous

\[ I_1(u) \leq \liminf_{m \to \infty} I_1(u_m) = c_1 \]

By definition of \(c_1\), \(I_1(u) = c_1\), and thus \(u \in \mathcal{K}_1\). For \(w_m = u_m - u\),

\[
I_1(u_m) = I_1(u + w_m) = \int_0^1 \left( \frac{1}{2} |\dot{u} + \dot{w}_m|^2 - V(t, u + w_m) \right) dt
\]

\[
= \int_0^1 \left( \frac{1}{2} |\dot{u}|^2 + \frac{1}{2} |\dot{w}_m|^2 + \dot{u} \cdot \dot{w}_m - V(t, u + w_m) \right) dt
\]

\[
= \frac{1}{2} \|w_m\|_{E_1}^2 + I_1(u) + \int_0^1 \left( \dot{u} \cdot \dot{w}_m - \frac{1}{2} |w_m|^2 - V(t, u + w_m) + V(t, u) \right) dt
\]

Notice that,

. \(w_m \in E \setminus B_\rho(\mathcal{K}_1)\), as a consequence of condition (*)

. \(\|w_m\|_E \geq \rho\) since we considered \((u_m) \in E \setminus B_\rho(\mathcal{K}_1)\) and \(v \in \mathcal{K}_1\)
. since \( u_m \to u \) in \( E_1 \),
\[
\int_0^1 (\dot{u} \cdot \dot{w}_m) dt \to 0 \quad m \to \infty
\]
. since \( u_m \to u \) in \( L^\infty([0,1], \mathbb{R}^n) \)
\[
\int_0^1 |w_m|^2 dt \to 0 \quad m \to \infty
\]
and by the continuity of \( V \), also
\[
V(t, u + w_m) - V(t, u) \to 0 \quad m \to \infty
\]
Then, taking the limit as \( m \to \infty \),
\[
c_1 = \lim_{m \to \infty} I_1(u_m) \geq c_1 + \frac{1}{2} \rho^2
\]
which is a contradiction.

\begin{align*}
\text{Theorem 13.} & \quad \text{If (V1)-(V2) and (*) hold, for any } v \in K_1, \text{ there exists } w \in K_1 \setminus \{v\} \text{ and a solution } Q \text{ of (HS) such that } Q(t) - v(t) \to 0 \text{ as } t \to -\infty \text{ and } Q(t) - w(t) \to 0 \text{ as } t \to \infty. \text{ By its definition, } Q \text{ is a heteroclinic solution from } v \text{ to } w. \\
& \text{Unfortunately, we can’t prove the Theorem by minimization of } \\
I(q) &= \int_{\mathbb{R}} \mathcal{L}(q) dt \\
\text{Notice that, since we want } \quad & Q(t) - v(t) \to 0 \quad \text{as } t \to -\infty \\
\text{there exists } T \in \mathbb{N} \text{ such that } \quad & Q(t) \approx v(t) \quad \forall t \leq -T \\
\text{As a consequence } \quad & \int_{-\infty}^{-T} \mathcal{L}(Q) dt \approx \int_{-\infty}^{-T} \mathcal{L}(v) dt = \sum_{k=-\infty}^{-T} \int_{k-1}^{k} \mathcal{L}(v) dt = \sum_{k=-\infty}^{-T} c_1 = \infty
\]
Therefore \( I \) is not well defined for the functions we want to find, and we have to formulate the variational problem in a different way.

For \( q \in W^{1,2}_{\text{loc}}(\mathbb{R}, \mathbb{R}^n) \) and \( p \in \mathbb{Z} \), define
\[
a_p(q) = \int_{p}^{p+1} \mathcal{L}(q) dt - c_1
\]
Notice that \( W^{1,2}([p, p+1], \mathbb{R}^n) \) can be identified with \( W^{1,2}([0,1], \mathbb{R}^n) \) by a simple change of coordinates.
Homework 20. Show that \( a_p(q) \geq 0 \), for all \( p \in \mathbb{Z} \).

We define the renormalized functional

\[
J(q) = \sum_{p \in \mathbb{Z}} a_p(q)
\]

We will minimize \( J \) over the set of admissible curves

\[
\Gamma = \{ q \in W^{1,2}_{\text{loc}}(\mathbb{R}, \mathbb{R}^n) : q(-\infty) - v(-\infty) = 0 \text{ and } \exists w \in \mathcal{K}_1 \setminus \{ v \} : q(\infty) - w(\infty) = 0 \}
\]

It is easy to check that \( \Gamma \) is nonempty: define the function that coincides with \( v \) in \((-\infty, 1)\) runs to \( w \) for \( t \in [0, 1]\) and coincides with \( w \) for \((1, \infty)\). Further define

\[
c = \inf_{\Gamma} J(q)
\]

The proof of the Theorem follows from two propositions.

**Proposition 10.** There exists \( Q \in \Gamma \) such that \( J(Q) = c \).

**Proposition 11.** \( Q \) is a solution of (HS) and \( \dot{Q}(t) - \dot{v}(t) \to 0 \) as \( t \to -\infty \) and \( \dot{Q}(t) - \dot{w}(t) \to 0 \) as \( t \to \infty \).

The variational characterization of \( Q \), and the asymptotic behavior of \( \dot{Q} \), imply that \( Q \) is in fact a heteroclinic solution between two periodic solutions of (HS). The proof of Proposition 11 follows the same ideas as previous proofs, so we will omit it.

**Proof:** (Proposition 10)

Let \((q_m) \subset \Gamma \) be a minimizing sequence for \( J \) over \( \Gamma \). The periodicity of \( V \) implies that

\[
J(\tau_k(q_m)) = J(q_m(t - k)) = J(q_m)
\]

for all \( k \in \mathbb{Z} \). Thus, we can replace \( q_m \) by an appropriate choice of \( \tau_k q_m \) verifying

\[
|\tau_k q_m(t) - v(t)| < \rho \text{ for all } t < 0 \text{ and } \\
|\tau_k q_m(t) - v(t)| = \rho \text{ for some } t \in [0, 1).
\]

Without loss of generality we can assume that \( k_m = 0 \) for all \( m \). Our goal is to obtain bounds for \( q_m \) in \( W^{1,2}_{\text{loc}}(\mathbb{R}, \mathbb{R}^n) \). Since \( J(q_m) \to c \), as \( m \to \infty \), there exists \( M \geq 0 \) such that \( J(q_m) \leq M \) for all \( m \in \mathbb{N} \). For fixed \( L \in \mathbb{R}^+ \), we can find \( k \in \mathbb{N} \) such that \( k - 1 \leq L < k \). Then

\[
M \geq J(q_m) = \sum_{p \in \mathbb{Z}} a_p(q_m) \geq \sum_{p=-k}^{k-1} a_p(q_m) \\
= \sum_{p=-k}^{k-1} \left( \int_{p}^{p+1} \mathcal{L}(q_m) \, dt - c_1 \right) \\
= \int_{-L}^{L} \mathcal{L}(q_m) \, dt - 2kc_1 \geq \int_{-L}^{L} \mathcal{L}(q_m) \, dt - 2Lc_1
\]

77
This implies
\[
\int_{-L}^{L} \frac{1}{2} |\dot{q}_m|^2 \, dt \leq M + 2Lc_1 + \int_{-L}^{L} V(t, q_m) \, dt
\]
Since \( V \) is continuous and periodic, \( \int_{-L}^{L} V(t, q_m) \, dt \) is bounded independently of \( m \), and therefore
\[
\int_{-L}^{L} \frac{1}{2} |\dot{q}_m|^2 \, dt \leq K(L)
\]
for all \( m \in \mathbb{N} \) and all \( L \in \mathbb{R}^+ \). On the other hand, as a consequence of the normalization
\[
|q_m(0)| \leq |v(0)| + \rho
\]
Then \( q_m \) is bounded in \( W^{1,2}([-L, L], \mathbb{R}^n) \) for every \( L > 0 \). Since \( W^{1,2}([-L, L], \mathbb{R}^n) \) is a Hilbert space, for each \( L \in \mathbb{R}^+ \), \( q_m \) admits a convergent subsequence. By a diagonalization process we can find \( Q \in W^{1,2}(\mathbb{R}, \mathbb{R}^n) \) such that \( q_m \to Q \) weakly in \( W^{1,2}_{loc}(\mathbb{R}, \mathbb{R}^n) \). We have to show that \( Q \in \Gamma \) and that \( J(Q) = c \). Let’s start by proving that \( J(Q) < \infty \). Consider \( L \in \mathbb{N} \) arbitrary, then
\[
\sum_{p=-L}^{L-1} a_p(Q) = \sum_{p=-L}^{L-1} \left( \int_{-L}^{L} \mathcal{L}(Q) \, dt - c_1 \right) = \int_{-L}^{L} \mathcal{L}(Q) \, dt - 2Lc_1
\]
Since, \( \int_{a}^{b} \mathcal{L}(q) \, dt \) is weakly lower semicontinuous
\[
\sum_{p=-L}^{L-1} a_p(Q) \leq \liminf_{m \to \infty} \int_{-L}^{L} \mathcal{L}(q_m) \, dt - 2Lc_1 = \liminf_{m \to \infty} \sum_{p=-L}^{L-1} a_p(q_m) \leq \liminf_{m \to \infty} J(q_m) \leq M
\]
for all \( L \in \mathbb{N} \). Letting \( L \to \infty \) we conclude \( J(Q) \leq M < \infty \), as wanted. As an immediate consequence
\[
a_p(Q) \to 0 \quad \text{as} \quad |p| \to \infty
\]
For \( Q_p = Q|_{p+1}^{p+1} \) (which we can view as an element of \( W^{1,2}([0,1], \mathbb{R}^n) \)), we can choose \( p \in \mathbb{Z} \) so that
\[
a_p(Q) \leq \frac{1}{2} \alpha(\rho)
\]
where \( \alpha \) is as defined in Proposition 9. Then
\[
I_1(Q_p) \leq \frac{1}{2} \alpha(\rho) + c_1
\]
and we conclude that
\[
Q_p \in B_\rho(K_1) \setminus K_1
\]
If not we would have
\[
I_1(Q_p) \geq c_1 + \alpha(\rho)
\]
as a consequence of Proposition 9, contradicting (31). Then, there exists \( v_p \in K_1 \) such that

\[
\|Q_p - v_p\|_{W^{1,2}([p,p+1],\mathbb{R}^n)} \leq \rho
\]

and by Sobolev inequality

\[
\|Q_p - v_p\|_{L^\infty([p,p+1],\mathbb{R}^n)} \leq 2\rho
\]

On the other hand, we normalized \((q_m)\) so that

\[
|q_m(t) - v(t)| \leq \rho, \quad \forall t \leq 0
\]

For fixed \( t \), letting \( m \to \infty \) we will also obtain

\[
|Q(t) - v(t)| \leq \rho, \quad \forall t \leq 0
\]

Then, for large negative \( p \)

\[
\|v_p - v\|_{L^\infty([p,p+1],\mathbb{R}^n)} \leq \|v_p - Q_p\|_{L^\infty([p,p+1],\mathbb{R}^n)} + \|Q_p - v\|_{L^\infty([p,p+1],\mathbb{R}^n)} \leq 3\rho
\]

Recall that we made the choice of

\[
\rho < \frac{\gamma}{4} = \frac{1}{4} \inf_{z \neq w} \|z - w\|_{L^\infty([0,1],\mathbb{R}^n)}
\]

and therefore

\[
\|v_p - v\|_{L^\infty([p,p+1],\mathbb{R}^n)} \leq \rho < \frac{3}{4} \gamma < \gamma
\]

which implies \( v = v_p \) in \([p, p + 1]\) for all large negative \( p \). Thus

\[
\|Q_p - v\|_{W^{1,2}([p,p+1],\mathbb{R}^n)} \leq 2\rho
\]

for all large negative \( p \), and as a consequence \((Q_p)\) is a bounded sequence in \( W^{1,2}([0,1]) \). Thus, there exists \( Q^* \in W^{1,2}([0,1]) \) such that \( Q_p \to Q^* \) weakly in \( W^{1,2}([0,1],\mathbb{R}^n) \) and strongly in \( L^\infty([0,1],\mathbb{R}^n) \). By the weak lower semicontinuity

\[
I_1(Q^*) \leq \liminf_{p \to -\infty} I_1(Q_p)
\]

But, since \( a_p(Q) = a_p(Q_p) \to 0 \) as \( p \to -\infty \) we will have

\[
\lim_{p \to -\infty} I_1(Q_p) = c_1
\]

and therefore \( I_1(Q^*) \leq c_1 \). This together with the fact \( Q^* \in W^{1,2}([0,1]) \), implies \( I_1(Q^*) = c_1 \), i.e., \( Q^* \in K_1 \). On the other hand

\[
|Q_p(t) - v(t)| \leq \rho
\]

letting \( p \to -\infty \),

\[
|Q^*(t) - v(t)| \leq \rho
\]
and by choice of $\rho$, we must necessarily have $Q^*(t) = v(t)$. We showed that a subsequence of $Q_p$ converges uniformly to $v$. The uniqueness of limit implies that all the sequence converges to $v$. This gives the right asymptotic behavior as $t \to -\infty$.

We will now prove that
$$
\lim_{t \to \infty} |Q(t) - w(t)| = 0
$$
for some $w \in K_1 \setminus \{v\}$. Our previous reasoning is valid, up to the step
$$
\|Q_p - v_p\|_{L^{\infty}((0,1], \mathbb{R}^n)} < 2\rho
$$
(32)
In the following arguments we can’t use the normalization, so we will have to work harder. (32) implies, in particular, that
$$
|Q_p(1) - v_p(1)| \leq 2\rho
$$
and
$$
|Q_{p+1}(0) - v_{p+1}(0)| \leq 2\rho
$$
But since $v_p \in K_1$, we will also have
$$
|Q_{p+1}(0) - v(0)| \leq 2\rho
$$
Then
$$
|v_p(0) - v_{p+1}(0)| \leq 4\rho
$$
and since $\rho < \frac{\gamma}{4}$,
$$
|v_p(0) - v_{p+1}(0)| \leq \gamma
$$
By Corollary 8, we can conclude that $v_p = v_{p+1}$ for $p$ large. If we denote the common function by $w$, it is easy to prove that, (using the arguments we used to prove $Q(-\infty) - v(-\infty) = 0$), as $t \to \infty$, $|Q(t) - w(t)| \to 0$. Nevertheless, we still have to check that $w(t) \neq v(t)$. For that purpose, we will use a comparison argument. Let’s suppose that $w = v$. For $0 < \delta << \alpha(\frac{\rho}{2})$, (for $\alpha(\rho/2)$ as in Proposition 9), since $Q(t) \to v(t)$ as $t \to \infty$, there exists $T = T(\delta) > 1$ such that
$$
|Q(t) - v(t)| \leq \delta, \quad \forall t \geq T
$$
Since $q_m \to Q$ in $L^{\infty}_{\text{loc}}(\mathbb{R}, \mathbb{R}^n)$, for $m$ sufficiently large
$$
|q_m(t) - v(t)| \leq 2\delta
$$
for $t \in [T-1, T]$. Define the comparison function
$$
v_m(t) = \begin{cases} 
    v(t) & \text{if } t \leq T - \delta \\
    \frac{(t-(T-\delta))}{\delta}q_m(T) + \frac{T-\delta}{\delta}v(T-\delta) & \text{if } T - \delta \leq t \leq T \\
    q_m(t) & \text{if } t \geq T
\end{cases}
$$
It is easy to check that $v_m \in \Gamma$. Moreover
$$
J(v_m) - J(q_m) = \sum_{p \in \mathbb{Z}} (a_p(v_m) - a_p(q_m))
$$
On one hand, \( v_m(t) = v(t) \) for all \( t \leq T - \delta \), and therefore
\[
\sum_{-\infty}^{T-1} a_i(v_m) = 0
\]
On the other hand, \( v_m(t) = q_m(t) \) for \( t \geq T \), implying
\[
\sum_{T}^{\infty} a_i(q_m) = \sum_{T}^{\infty} a_i(v_m)
\]
Hence
\[
J(v_m) - J(q_m) = a_{T-1}(v_m) - \sum_{p=-\infty}^{T-1} a_p(q_m)
\]
In the interval \([0,1)\) \( q_m \) hits the boundary of \( B_p(v) \) and returns to \( B_\delta(v) \) for \( t \in [T-1,T] \). In particular, we can find a negative integer \( p_0 \leq T \) such that \( \|q_m - v\|_{W^{1,2}(\{p_0, p_0+1\}, \mathbb{R}^n)} \geq \frac{\rho}{2} \)
By Proposition 9, c)
\[
\int_{p_0}^{p_0+1} L(q_m) \, dt \geq c_1 + \alpha(\rho/2)
\]
which is equivalent to \( a_{p_0}(q_m) \geq \alpha(\rho/2) \)
Then
\[
J(v_m) - J(q_m) \leq a_{T-1}(v_m) - a_{p_0}(q_m) \leq a_{T-1}(v_m) - \alpha(\rho/2)
\]
Finally, let us now analyse \( a_{T-1}(v_m) \).
\[
a_{T-1}(v_m) = \int_{T-1}^{T} L(v_m) \, dt - c_1 = \int_{T-1}^{T} (L(v_m) - L(v)) \, dt
\]
By construction, \( v_m = v \) for \( t \in [T-\delta, T] \), and as a consequence
\[
a_{T-1}(v_m) = \int_{T-\delta}^{T} (L(v_m) - L(v)) \, dt
\]
We proved before that \( v \in C^2(\mathbb{R}, \mathbb{R}^n) \), therefore
\[
\left| \int_{T-\delta}^{T} L(v) \, dt \right| \leq \max_{t \in [T-\delta, T]} \left( \frac{1}{2} |\dot{v}|^2 + V(t, v) \right) \leq \delta M_1
\]
for some positive constant \( M_1 \). On the other hand, for \( t \in [T-\delta, T] \), also by construction, \( |v_m(t)|! \leq \epsilon(\delta) \to 0 \) as \( \delta \to 0 \), and thus
\[
\left| \int_{T-\delta}^{T} V(t, v_m) \, dt \right| \leq M_2 \delta
\]
for some positive constant $M_2$. Lastly, by its definition

$$\left| \int_{T-\delta}^{T} \frac{1}{2} |\dot{v}_m|^2 dt \right| = \frac{1}{2} \int_{T-\delta}^{T} \left| \frac{q_m(T) - v(T-\delta)}{\delta} \right|^2 dt$$

and

$$|q_m(T) - v(T-\delta)| \leq |q_m(T) - v(T)| + |v(T) - v(T-\delta)| \leq 2\delta + \|\dot{v}\|_{L^\infty[T-\delta,T]} \delta$$

Using, once again, the fact that $v \in C^2(\mathbb{R}, \mathbb{R}^n)$, we can conclude that there exists $M_3 \geq 0$ such that

$$|q_m(T) - v(T-\delta)| \leq M_3 \delta$$

Hence

$$\left| \int_{T-\delta}^{T} \frac{1}{2} |\dot{v}_m|^2 dt \right| \leq M_3 \delta$$

Finally, for some positive $M$

$$J(v_m) - J(q_m) \leq M\delta - \alpha\left(\frac{\rho}{2}\right)$$

and we can choose $\delta$ so small so that $M\delta < \alpha(\frac{\rho}{2})$, implying

$$J(v_m) < J(q_m)$$

which is a contradiction to the fact that $(q_m)$ is a minimizing sequence for $J$ over $\Gamma$. To finish the proof it remains to show that $J(Q) = c$. Since $J(Q) < \infty$

$$\sum_{-L}^{L} a_p(Q) \geq J(Q) - \epsilon$$

Hence

$$c = \lim_{m \to \infty} J(q_m) \geq \lim_{m \to \infty} \left( \int_{-L}^{L} \mathcal{L}(q_m) \ dt - c_1 \right)$$

$$\geq \left( \int_{-L}^{L} \mathcal{L}(Q) \ dt - c_1 \right) = \sum_{-L}^{L} a_p(Q)$$

Letting $L \to \infty$ we obtain $J(Q) \leq c$. Since $Q \in \Gamma$, by definition of $c$, $J(Q) = c$. 

\[\Diamond\]
6 Geodesics

Let $\mathcal{M}$ be a metric space, i.e., a set equipped with a map $d : \mathcal{M} \times \mathcal{M} \to \mathbb{R}$, verifying

(M1) $d(x, y) = 0$ iff $x = y$;
(M2) $d(x, y) = d(y, x)$;
(M3) $d(x, z) \leq d(x, y) + d(y, z)$.

Intuitively, we can interpret $d(x, y)$ as the distance from $x$ to $y$.

Example 30. 1. $\mathcal{M} = \mathbb{R}^n$ with $d(x, y) = |x - y|$.

2. $\mathcal{M} = C([0, 1], \mathbb{R})$ with

$$d(x, y) = \int_0^1 |x(t) - y(t)| \, dt$$

A function $f$ defined in $I \subset \mathbb{R}$, a closed interval, verifying $f \in C(I, \mathcal{M})$ is a curve or path in $\mathcal{M}$. Consider $\mathcal{P}$ as the set of all partitions of the interval $I = [a, b]$, i.e.

$$\mathcal{P} = \{ P = (t_0, ..., t_k) : a = t_0 < t_1 < ... < t_k = b \}$$

For any $P \in \mathcal{P}$, we can define the length of $f$ relatively to $P$ as

$$L(Pf) = \sum_{i=1}^{k} d(f(t_{i+1}), f(t_i))$$

Definition 14. We define Lenght of $f$ as

$$L(f) = \sup_{P \in \mathcal{P}} L(Pf)$$

Moreover, $f$ is rectifiable if $L(f) < \infty$.

Homework 21. Let $f \in C(I, \mathcal{M})$ be rectifiable. Show that:

(a) For all $\tau \leq s \leq t$ in $I = [a, b]$

$$L(f|_{[\tau, t]}) = L(f|_{[\tau, s]}) + L(f|_{[s, t]})$$

(b) For any $\psi \in C([c, d], [a, b])$ strictly monotone, with $\psi(c) = a$ and $\psi(d) = b$, and $h = f \circ \psi \in C([c, d], \mathcal{M})$, then

$$L(h) = L(f)$$

that is the length is independent of the parametrization.
Example 31. Let $\mathcal{M} = \mathbb{R}^n$, with $d(x, y) = |x - y|$ and $f \in C^1([a, b], \mathbb{R}^n)$. Then, using the Mean Value Theorem, for any $P \in \mathcal{P}$

$$L(Pf) = \sum_{i=1}^{k} |f(t_{i+1}) - f(t_i)| = \sum_{i=1}^{k} |f'(\theta_i)||t_{i+1} - t_i|$$

for some $\theta_i \in [t_i, t_{i+1}]$. Taking the supremum over all partitions, we conclude

$$L(f) = \sup_{P \in \mathcal{P}} L(Pf) = \int_a^b |f'(t)| \, dt$$

Define\n
$$\mathcal{R} = \{ f \in C(I, \mathcal{M}) : L(f) < \infty \}$$

due to $\mathcal{M}$ the set of rectifiable curves on $\mathcal{M}$. Then $\mathcal{R}$ can be made a metric space, by introducing the metric

$$\rho(f, g) = \max_{t \in I} d(f(t), g(t)) ; \quad f, g \in \mathcal{R}$$

Example 32. If $\mathcal{M} = \mathbb{R}^n$, then $\rho$ corresponds to the maximum norm:

$$\rho(f, g) = \max_{t \in [0,1]} |f(t) - g(t)|$$

The next example shows that the length map, $L : \mathcal{R} \rightarrow \mathbb{R}$, is not continuous, in the topology defined by $\rho$.

Example 33. Let $f : [0, 1], \mathbb{R})$ be defined by $f(t) = 1 - |1 - 2t|$, and consider $F : \mathbb{R} \rightarrow \mathbb{R}$ the periodic extension of $f$ to $\mathbb{R}$. Define the sequence $f_m \in \mathcal{R}$ by $f_m(t) = \frac{F(mt)}{m}$. It is easy to check that, for all $m \in \mathbb{N}$, $L(f_m) = \sqrt{5}$. But $f_m \rightarrow 0$ as $m \rightarrow \infty$ and $L(\text{limit curve}) = 1$. Hence $L(f_m) \not\rightarrow L(\text{limit curve})$ as $m \rightarrow \infty$.

Theorem 14. $L$ is lower semicontinuous on $\mathcal{R}$.

Proof:

We want to prove that

$$L(f) \leq \liminf_{m \to \infty} L(f_m)$$

for any $(f_m) \subset \mathcal{R}$, $f \in \mathcal{R}$ such that $\rho(f_m, f) \to 0$ as $m \to \infty$. Choose $s \leq L(f) = \sup_{P \in \mathcal{P}} L(Pf)$. By definition of supremum, we can find a $P_0 \in \mathcal{P}$, such that

$$s < L(P_0 f) \leq L(f)$$

and

$$L(P_0 f) = \sum_{i=1}^{k} d(f(t_{i-1}), f(t_i))$$

84
Pick $\epsilon > 0$ and $g \in \mathcal{R}$, close to $f$ in the sense
\[ \rho(f, g) < \frac{\epsilon}{k} \]

Then, for each $i \in \{1, \ldots, k\}$
\[ d(f(t_{i-1}), f(t_i)) \leq d(f(t_{i-1}), g(t_{i-1})) + d(g(t_{i-1}), g(t_i)) + d(g(t_{i-1}), f(t_i)) \]
\[ \leq 2\frac{\epsilon}{k} + d(g(t_{i-1}), g(t_i)) \]

Thus
\[ L(P_0 f) \leq 2\epsilon + L(P_0 g) \]
which is equivalent to
\[ L(P_0 g) \geq L(P_0 f) - 2\epsilon > s - 2\epsilon \]

Taking the supremum over all $P \in \mathcal{P}$, we conclude
\[ L(g) \geq s - 2\epsilon \]

Since $\rho(f_m, f) \to 0$, as $m \to \infty$, for $m$ large, so that $\rho(f_m, f) < \frac{\epsilon}{k}$, also
\[ L(f_m) \geq s - 2\epsilon \]

Taking the limit as $m \to \infty$
\[ \liminf_{m \to \infty} L(f_m) \geq s - 2\epsilon \]
for all $\epsilon > 0$, which implies
\[ \liminf_{m \to \infty} L(f_m) \geq s \quad \forall s < L(f) \]
and therefore
\[ \liminf_{m \to \infty} L(f_m) \geq L(f) \]

We want to prove the existence of a curve of shortest length joining two arbitrary points of a compact metric space. Before, we will prove a technic result which states that we can parametrize any curve in $\mathcal{R}$ in a way that a strong continuity condition is verified.

**Proposition 12.** For $f \in \mathcal{R}$, there exists $F \in \mathcal{R}$ such that

1. $F(0) = f(0)$ and $F(1) = f(1)$;
2. $L(F) = L(f)$;
3. $\forall s, t \in [0, 1] \ d(F(s), F(t)) \leq L(f)|s - t|$. 

85
Proof:
For $t \in [0, 1]$, define $\gamma(t) = L(f|_0^t)$. Then, it is obvious that $\gamma(0) = L(f|_0^0) = 0$, $\gamma(1) = L(f|_0^1) = L(f)$ and $\gamma$ is monotone non-decreasing. We are going to prove that $\gamma$ is continuous. For any $t \in [0, 1]$ and $\epsilon > 0$, since $f$ is continuous on a compact set, there exists $\delta = \delta(\epsilon) > 0$, such that
\[
|t - \tau| \leq \delta \Rightarrow d(f(t), f(\tau)) \leq \frac{\epsilon}{2}
\]
Choose a partition $P = (t_0, ..., t_k)$ of $[0, 1]$, verifying
\[
L(Pf) > L(f) - \epsilon
\]
We will assume that $t_i - t_{i-1} \leq \delta$, $1 \leq i \leq k$ and choose $\tau$ so that $|t - \tau| < \min(t_i - t_{i-1})$
By our choices, either $t, \tau$ are in the same partition interval, or they belong to adjacent intervals. We claim that, if this is the case, then
\[
|\gamma(\tau) - \gamma(t)| \leq 3\epsilon \tag{33}
\]
which will prove that $\gamma$ is continuous in $t$, for all $t \in [0, 1]$. Suppose (33) does not hold, i.e.,
\[
|\gamma(\tau) - \gamma(t)| > 3\epsilon
\]
Using the additivity property of the length, also
\[
3\epsilon < |\gamma(t) - \gamma(\tau)| = L(f|_t^\tau)
\]
We will prove that this yields a contradiction if the worst case $t_{i-1} < t < t_i < \tau < t_{i+1}$ occurs. On one hand
\[
L(Pf) + \epsilon > L(f(t_{i-1})^1) + L(f(t_{i-1})^1) + L(f(t_{i+1})^1)
\]
On the other hand
\[
L(Pf) = \sum_{j \neq i, i+1} d(f(t_{j-1}), f(t_j)) + d(f(t_{i-1}), f(t_i)) + d(f(t_i), f(t_{i+1})) + \epsilon
\]
and, by definition
\[
L(f(t_{i-1})^1) + L(f(t_{i+1})^1) > \sum_{j \neq i, i+1} d(f(t_{j-1}), f(t_j))
\]
Also, by our assumption
\[
L(f(t_{i-1})^1) \geq L(f|_t^\tau) > 3\epsilon
\]
86
Then
\[ d(f(t_{i-1}), f(t_i)) + d(f(t_i), f(t_{i+1})) = L(Pf) - \sum_{j \neq i, i+1} d(f(t_{j-1}), f(t_j)) \geq L(Pf) - L(f(t_{i-1}^1)) - L(f(t_{i+1}^1)) = L(Pf) - L(f) + L(f(t_{i-1}^1)) < -\epsilon + 3\epsilon = 2\epsilon \]

But, the continuity of \( f \) implies
\[ d(f(t_{i-1}), f(t_i)) + d(f(t_i), f(t_{i+1})) \leq \epsilon \]
which is a contradiction, and we conclude that \( \gamma \) is continuous.

Our goal is to define the inverse function of \( \gamma \), but we have a problem since \( \gamma \) may not be strictly monotone. Define \( \varphi : [0, L(f)] \rightarrow [0, 1] \), by
\[ \varphi(r) = \inf \{ t \in [0, 1] : \gamma(t) = r \} \]
and, our first approximation to \( F, f^* : [0, L(f)] \rightarrow \mathcal{M} \), defined by
\[ f^*(r) = f(\varphi(r)) \]
Notice that, \( f^* \) is well defined, since \( \varphi \) is well defined. Moreover, if \( \gamma(t) = r \) for \( t \) minimal, then \( f^*(r) = t \). Also,
\[ f^*(0) = f(\varphi(0)) = \inf \{ t \in [0, 1] : L(f(t)_0) = 0 \} = f(0) ; \]
\[ f^*(L(f)) = f(\inf \{ t \in [0, 1] : L(f(t)_0) = L(f) \}) \equiv f(t^*) \).
If \( t^* = 1 \) then \( f^*(L(f)) = f(1) \) and we are done. Otherwise \( t^* < 1 \) and therefore, by construction, \( f(t) \equiv \) constant for \( t \geq t^* \), implying that \( f(t) = f(1) \) for all \( t \geq t^* \). In both cases \( f^*(L(f)) = f(1) \).
Moreover
\[ d(f^*(s), f^*(t)) = d(f(s), f(t)) \quad (34) \]
which can be proved, using the same minimality argument we used above.

We will prove next, some properties of \( f^* \).

(1) \( f^* \) is continuous: by definition, \( f^* \) is the composition of a continuous function with \( \varphi \) which can be discontinuous at a jump point. But, notice that, if we approach by below there is no problem, since \( \varphi \) is continuous. If we approach by above, there is also no problem, since \( f \equiv \) constant in the “jump”.

87
(2) For all $t, r$ such that $\gamma(t) = r$

$$L(f^t_0) = l(f^r_0)$$

Consider $P_r = (r_0, ..., r_k)$ a partition of the interval $[0, r]$. Then, using (34)

$$L(P_r f^*) = \sum_{i=1}^{k} d(f^*(r_{i-1}), f^*(r_i)) = \sum_{i=1}^{k} d(f(r_{i-1}), f(r_i)) \leq L(f^t_0)$$

As a consequence

$$L(P_r f^*) \leq L(f^t_0)$$

and taking the supremum over all partitions $P_r$,

$$L(f^*_0) \leq L(f^t_0)$$

Play the other way around, we will obtain

$$L(f^*_0) \geq L(f^t_0)$$

which implies the equality.

As a consequence of property (2)

$$r = \gamma(t) = L(f^t_0) = L(f^r_0) \quad \forall r \in [0, L(f)]$$

which will also be true for $r + s$, i.e.

$$r + s = L(f^*_0^{r+s}) = L(f^*_0) + L(f^*_r^{r+s})$$

Subtracting the above equalities

$$s = L(f^*_r^{r+s})$$

which is equivalent to (making $r + s = \sigma$)

$$L(f^*_\sigma) = \sigma - s$$

We denote $f^*$ by the \textit{reparametrization of $f$ by arclength}. In particular, by definition of $L$

$$d(f^*(r)f^*(s)) \leq L(f^*_r) \leq |s - r| \quad (35)$$

Finally, for $t \in [0, 1]$, define

$$F(t) = f^*(L(f)t)$$

Then

$$F(0) = f^*(0) = f(0) \text{ and } F(1) = f^*(L(f)) = f(1).$$

$$L(F) = L(f^*) = L(f), \text{ since } F \text{ is just a reparametrization of } f^*.$$
As a consequence of (35)
\[
d(F(t), F(\tau)) = d(f^*(L(f)t), f^*(L(f)\tau)) \leq |L(f)t - L(f)\tau| = L(f)|t - \tau|
\]

\[\Box\]

**Theorem 15.** Let \(M\) be a compact metric space and \(x, y \in M\). Define
\[
S = \{ f \in \mathcal{R} : f(0) = x \text{ and } f(1) = y \}
\]
If \(S \neq \emptyset\), there exists \(g \in S\) such that \(L(g) \leq L(f)\) for all \(f \in S\).

**Proof:**
Take a minimizing sequence in \(S\), \((f_m)\) such that
\[
M \geq L(f_m) \ \sup_n \inf_S L
\]
By Proposition 12, we can replace \((f_m)\) by \((F_m)\) where \((F_m)\) has the additional property (3). Then
\[
d(F_m(s), F_m(t)) \leq L(f_m)|s - t| \leq M|s - t|
\]
meaning that \((F_m)\) is an equicontinuous sequence in \(C([0,1], M)\). Moreover for \(x = F_m(0)\)
\[
d(x, F_m(t)) \leq M|t| \ \forall t \in [0,1]
\]
implying that it is uniformly bounded for \(t \in [0,1]\). By the Ascoli-Arzela Theorem some subsequence of \(F_m\) will converge uniformly to some function \(g\), and \(g(0) = x, g(1) = y\). Also
\[
L(g) \leq \lim \inf_{f_m} L(F_m) = \inf_S L \leq M
\]
and therefore \(g\) is rectifiable \((g \in \mathcal{R})\). Hence \(g \in S\). This fact and (36) imply
\[
L(g) = \inf_S L
\]
\[\Box\]

### 6.1 Curves defined on a Manifold

**Definition 15.** \(f\) is absolutely continuous on \(I = [a,b] \subset \mathbb{R}\) if for all \(\epsilon > 0\) there exists \(\delta = \delta(\epsilon) > 0\), such that, whenever \([a_1, b_1], ..., [a_k, b_k]\), closed, disjoint subintervals of \([a,b]\), then
\[
\sum_{i=1}^{k} |f(b_i) - f(a_i)| < \epsilon
\]

Note that, if \(f\) is absolutely continuous then \(f\) is uniformly continuous (just take \(k = 1\)).
Definition 16. \( f \) defined on \( [a,b] \) is of Bounded Variation (\( f \in BV([a,b]) \)), if for all \( P = (x_0, ..., x_1) \in \mathcal{P} \) partition of the interval \( [a,b] \),

\[
\sup_{P \in \mathcal{P}} \sum |f(x_{i-1}) - f(x_i)| < \infty
\]

Notice that in \( \mathbb{R} \) the notion of bounded variation is equivalent to the notion of rectifiable. Also, it is easy to check that if \( f \in AC([a,b]) \), then \( f \in BV([a,b]) \). But, the same thing may not be true, if we consider \( f \in AC(\mathbb{R}) \).

Example 34. Consider \( f(x) = x \) and \( f(x) = \sin x \).

Theorem 16. \( f \in AC([a,b]) \) iff \( \int_a^T f'(t) \, dt = f(T) - f(a) \)

You can find the proof in Ch 8 of Rudin's Real and Complex Analysis.

We will now generalize the notion of length to a curve defined on a manifold.

Let \( M \subset \mathbb{R}^n \) be a \( k \) dimensional manifold, and consider \( c \in AC([0,T],M) \), i.e., \( c(t) \in M \) for all \( t \in [0,T] \).

Definition 17. We define the length of \( c \), \( L(c) \), by

\[
L(c) = \int_0^T |\dot{c}(t)| \, dt
\]

Since \( c \in AC([0,T]) \), the definition makes sense, as a consequence of Theorem 16. As we have seen in a previous example, if \( c \) is a smooth curve, \( L(c) \) coincides with our previous definition.

Definition 18. We define the energy of \( c \), \( E(c) \), by

\[
E(c) = \frac{1}{2} \int_0^T |\dot{c}(t)|^2 \, dt
\]

We observe, that \( E(c) \) may not be well defined, if \( \dot{c} \notin L^2([0,T]) \). So we have to make the extra assumption

\[
c \in W^{1,2}([0,T],M)
\]

Lemma 16. (Invariance Properties)

1. For \( i : \mathbb{R}^n \to \mathbb{R}^n \), defined by \( i(y) = Ay + b \) where \( A \) is an orthogonal matrix (\( i \) is the euclidean isometry), then

\[
L(ic) = L(c) \quad \text{and} \quad E(ic) = E(c)
\]

2. If \( \varphi \in C^1([0,s],[0,T]) \) is a diffeomorphism, then

\[
L(c \circ \varphi) = L(c)
\]
Proof: In order to prove (2), notice that, by definition

\[ L(c \circ \varphi) = \int_0^s \left| \frac{d}{ds} c(\varphi(s)) \right| ds = \int_0^s |\dot{c}(\varphi)| |\dot{\varphi}(s)| ds \]

\[ = \int_0^T |\dot{c}(t)| dt = L(c) \]

\[ \diamond \]

Unfortunately, \( E \) is not invariant under this type of transformations. But we can observe the following relationship. As a consequence of Schwarz inequality.

\[ L(c) = \int_0^T |\dot{c}(t)| dt \leq T^{1/2} \left( \int_0^T |\dot{c}(t)|^2 dt \right)^{1/2} = \sqrt{2TE(c)} \]

Recall that

\[ \int fg < \left( \int f^2 \right)^{1/2} \left( \int g^2 \right)^{1/2} \]

unless \( f = \alpha g \) and equality holds. Then, if \(|\dot{c}| = \) is constant, equality holds, and therefore

\[ L(c) = \sqrt{2TE(c)} \] (37)

If this is the case, \( c \) is said to be parametrized proportionally to the arclength. If \(|\dot{c}| = 1\), then \( c \) is said to be parametrized by the arclength.

Lemma 17. Let \( c : [0, L(c)] \to \mathbb{R}^n \) be a curve parametrized on \([0, L(c)]\). Then the parametrization of \( c \) by arclength (or the parametrization of \( E \) proportional to arclength), minimizes \( E \) over the class of reparametrizations \( \varphi : [0, L(c)] \to [0, L(c)] \).

Proof:

Denote by \( c^* \) the parametrization of \( c \) by arclength. Then \( L(c) = L(c^*) \) and by (37)

\[ L(c) = L(c^*) = \sqrt{2L(c)E(c^*)} \]

Squaring both sides, \( L(c) = 2E(c^*) \). As remarked above, if \( \tilde{c} \) is any other parametrization, then

\[ L(c) = L(\tilde{c}) < \sqrt{2L(c)E(\tilde{c})} \]

Squaring both sides, and comparing

\[ E(c^*) < E(\tilde{c}) \]

which concludes the proof.

\[ \diamond \]
Definition 19. $M \subseteq \mathbb{R}^n$ is a $k$-dimensional manifold, if for all $p \in M$ there are neighborhoods $W = W(p) \subseteq \mathbb{R}^n$, $U \subseteq \mathbb{R}^k$ and a diffeomorphism $f : U \to W$, such that $f(U) = M \cap W$. The triple $(f, U, W)$ is denoted by local chart for $M$.

Suppose that $c([0, T]) \subset f(U) \subset M$ (i.e., the image of the curve lies in a single local chart). By definition, we can find a curve $\gamma(t) \in U$ such that

$$c(t) = f(\gamma(t))$$

and we can express $L(c)$ in terms of $f$ and $\gamma$. We will get more complicated expressions but we no longer have to worry about the constraint (since $U$ is an open set).

Homework 22. If $f$ is smooth, $c$ is absolutely continuous in $[0, T]$ and $\gamma(t) \in U$, prove that $\gamma \in AC([0, T], U)$.

We want to express $L$ and $E$ in terms of $f$ and $\gamma$. For $c = (c_1, ..., c_n)$, by definition

$$L(c) = \int_0^T |\dot{c}(t)|^2 dt = \int_0^T \left( \sum_{m=1}^n (\dot{c}_m(t))^2 \right)^{1/2} dt$$

Since

$$c(t) = f(\gamma(t)) = (f_1(\gamma_1, ..., \gamma_k), ..., f_n(\gamma_1, ..., \gamma_k))$$

using the chain rule

$$\dot{c}_m = \sum_{i=1}^k \frac{\partial f_m}{\partial x_i}(\gamma) \cdot \dot{\gamma}_i$$

and, as a consequence

$$L(c) = \int_0^T \left( \sum_{i,j=1}^k \sum_{m=1}^n \left( \frac{\partial f_m}{\partial x_i}(\gamma) \cdot \dot{\gamma}_i \right) \left( \frac{\partial f_m}{\partial x_j}(\gamma) \cdot \dot{\gamma}_j \right) \right)^{1/2} dt$$

In a similar way

$$E(c) = \int_0^T \frac{1}{2} |\dot{c}|^2 dt = \frac{1}{2} \int_0^T \sum_{i,j=1}^k \sum_{m=1}^n \left( \frac{\partial f_m}{\partial x_i}(\gamma) \cdot \dot{\gamma}_i \right) \left( \frac{\partial f_m}{\partial x_j}(\gamma) \cdot \dot{\gamma}_j \right) dt$$

If we set

$$g_{ij}(x) = \sum_{m=1}^n \frac{\partial f_m}{\partial x_i}(x) \frac{\partial f_m}{\partial x_j}(x)$$

$[g_{ij}(x)]_{i,j=1}^n$ defines a matrix $G(x)$, which is denoted the metric tensor with respect to the chart $f$. Then

$$L(c) = \int_0^T \left( \sum_{i,j=1}^k g_{ij}(\gamma) \dot{\gamma}_i \dot{\gamma}_j \right)^{1/2} dt$$
and
\[ E(c) = \int_0^T \frac{1}{2} \sum_{i,j=1}^k g_{ij}(\gamma) \dot{\gamma}_i \dot{\gamma}_j \, dt \]

**Homework 23.** Show that \( G \) is a symmetric and positive definite matrix, i.e,

a) \( g_{ij} = g_{ji} \)

b) \( Gy \cdot y \geq 0 \) and \( Gy \cdot y = 0 \) iff \( y = 0 \).

The expression
\[ \sum_{i,j=1}^k g_{ij} \dot{\gamma}_i \dot{\gamma}_j \]
is called the [riemannian metric](https://en.wikipedia.org/wiki/Riemannian_metric). Our initial interest is to find curves defined in \( M \) of smallest length or smallest energy.

If we look at the two expressions we derived for \( L(c) \):
\[ L(c) = \int_0^T |\dot{c}(t)| \, dt = \int_0^T \left( \sum_{i,j=1}^k g_{ij} \dot{\gamma}_i \dot{\gamma}_j \right)^{1/2} \]
we observe that, while the first one depends on the manifold, the second one depends on \( g \). The equality suggests that the representation does not depend on the charts.

**Lemma 18.** The representation
\[ L(c) = \int_0^T \left( \sum_{i,j=1}^k g_{ij} \dot{\gamma}_i \dot{\gamma}_j \right)^{1/2} \, dt \]
where
\[ g_{ij}(x) = \sum_{m=1}^n \frac{\partial f_m}{\partial x_i} \frac{\partial f_m}{\partial x_j} \]
does not depend of the choice of \( f, U, W \).

**Proof:**
Consider two different charts, \( f : U \to W \) and \( \tilde{f} : \tilde{U} \to \tilde{W} \) such that \( c([0, T]) \subset f(U) \) and \( c([0, T]) \subset \tilde{f}(\tilde{U}) \). By definition, we can find two continuous maps, \( \gamma, \tilde{\gamma} \), such that
\[ c(t) = f(\gamma(t)) = \tilde{f}(\tilde{\gamma}(t)) \quad , \forall t \in [0, T] \]
Define \( \phi : f^{-1}(f(U) \cap \tilde{f}(\tilde{U})) \to \tilde{f}^{-1}(f(U) \cap \tilde{f}(\tilde{U})) \) by
\[ \phi = \tilde{f}^{-1} \circ f \]
We want to show that
\[ \sum_{i,j=1}^{k} g_{ij}(x) \dot{\gamma}^i \dot{\gamma}^j = \sum_{i,j=1}^{k} \tilde{g}_{ij}(x) \dot{\gamma}^i \dot{\gamma}^j \]

By definition, for \( x \in \tilde{U} \)
\[ \tilde{g}_{ij}(x) = \sum_{m=1}^{n} \frac{\partial \tilde{f}_m}{\partial x_i} \frac{\partial \tilde{f}_m}{\partial x_j} \]

Since \( f(\gamma(t)) = c(t) = \tilde{f}(\tilde{\gamma}(t)) \)

then
\[ \tilde{\gamma}(t) = \varphi(\gamma(t)) \]

Therefore, we can compute the derivatives, using the chain rule
\[ \dot{\gamma}_i = \frac{d}{dt}(\varphi(\gamma(t)))_i = \sum_{l=1}^{k} \frac{\partial \varphi_i}{\partial x_l}(\gamma) \gamma_l (38) \]

On the other hand
\[ g_{ij}(x) = \sum_{m=1}^{n} \frac{\partial f_m}{\partial x_i} \frac{\partial f_m}{\partial x_j} = \sum_{m=1}^{n} \frac{\partial(\tilde{f}(\varphi))_m}{\partial x_i} \frac{\partial(\tilde{f}(\varphi))_m}{\partial x_j} \]

since \( f = \tilde{f} \circ \varphi \). Carrying out the derivatives
\[ \frac{\partial f_m}{\partial x_i} = \sum_{l=1}^{k} \frac{\partial \tilde{f}_m}{\partial x_l} \frac{\partial \varphi_l}{\partial x_i} \]

and as a consequence, we obtain a relation between the two metric tensors
\[ g_{ij}(x) = \sum_{m=1}^{n} \left( \sum_{l=1}^{k} \frac{\partial \tilde{f}_m}{\partial x_l} \frac{\partial \varphi_l}{\partial x_i} \right) \left( \sum_{l=1}^{k} \frac{\partial \tilde{f}_m}{\partial x_l} \frac{\partial \varphi_l}{\partial x_j} \right) \]

Finally, using (38)
\[ \sum_{i,j=1}^{k} g_{ij}(x) \dot{\gamma}^i \dot{\gamma}^j = \sum_{i,j=1}^{k} \sum_{l,t=1}^{k} \tilde{g}_{lt}(x) \frac{\partial \varphi_l}{\partial x_i} \frac{\partial \varphi_t}{\partial x_j} \dot{\gamma}^i \dot{\gamma}^j \]

\[ = \sum_{l,t=1}^{k} \tilde{g}_{lt}(x) \dot{\gamma}^l \dot{\gamma}^t \]

Hence we conclude that the expressions for \( L \) and \( E \) are independent of the chart.

\[ \diamond \]

94
Next, we will extend the notions for the case when the curve doesn’t lie in a single element of the chart. Let \( c : [0, 1] \to M \), and take any partition \( P = (t_0, ..., t_k) \) of the interval \([0, 1]\). Choose a parametrization of the curve, such that
\[
c([t_i, t_{i+1}]) \subset f_i(U_i)
\]
For each \( f_i : U_i \to W_i, 1 \leq i \leq n \), is a single chart and therefore has an associated metric tensor \((g'_{ij}(x))\). By the additivity property of the length \( L(c) = \sum_{1}^{m} L(c|_{t_{i-1}^{t_{i}}}) = \sum_{t_{i-1}}^{t_{i}} \left( \sum_{i,j=1}^{k} g'_{ij}(\gamma_{i}(t)) \dot{\gamma}_{i} \dot{\gamma}_{j} \right)^{1/2} \)
where \( c(t) = f_{\nu}(\gamma_{\nu}(t)) \) for \( t \in [t_{\nu-1}, t_{\nu}] \). Nevertheless, to avoid extra difficulties in the notation, we will abbreviate
\[
L(c) = \int_{0}^{1} \left( \sum_{i,j=1}^{k} g_{ij}(\gamma(t)) \dot{\gamma}_{i} \dot{\gamma}_{j} \right)^{1/2} dt
\]
and likewise for \( E \).

We want to find the curve of shortest length joining two arbitrary points of a \( k \)-dimensional differentiable manifold \( M \) of \( \mathbb{R}^{n} \). The expressions we derived for \( L \) and \( E \), using the local coordinates of \( M \), are undoubtedly messier, but we no longer have to deal with the constraint \( c(t) \in M \) for all \( t \in [0, T] \): from now on, it is enough to consider curves, \( \gamma \), defined on an open set of \( \mathbb{R}^{k} \). The purpose of this is, if we want to derive the differential equation satisfied by a minimizer, \( v \), of \( L \), to argue as before, we need to be sure that \( v + \delta \varphi \in M \) for some appropriate \( \varphi \). This may not be an easy task to do.

To solve the problem, for \( x, y \in M \), consider
\[
\Gamma_{xy} = \{ q \in W^{1,2}([0, T], M) : q(0) = x, q(T) = y \}
\]
As remarked, we want to find the curve of \( \Gamma_{xy} \) of shortest length. Lemma 17 suggests that it is better to minimize the energy, \( E(c) \), instead, since its minimizers may be parametrized with respect to arclength. Nevertheless, we need to be sure that the curves that minimize \( E \), also minimize \( L \).

**Proposition 13.** Let \( \tilde{q} \in \Gamma_{xy} \) be a minimizer of \( E \) over \( \Gamma_{xy} \). Then

1. \( \tilde{q} \) is parametrized proportionally to arclength.
2. \( \tilde{q} \) minimizes \( L(q) \) over the class \( \Gamma_{xy} \).

**Proof:**

1. If this is was not the case, then
\[
L(\tilde{q}) = \int_{0}^{T} |\dot{\tilde{q}}| dt \leq \left( \int_{0}^{T} |\dot{\tilde{q}}|^2 dt \right)^{1/2} = \sqrt{2TE(\tilde{q})}
\]

95
If \( q^* \) is the parametrization of \( \tilde{q} \) over \([0, 1]\) proportional to arclength, as a consequence of Lemma 17

\[
L(\tilde{q}) = L(q^*) = \sqrt{2TE(q^*)} < \sqrt{2TE(\tilde{q})}
\]

implying \( E(q^*) < E(\tilde{q}) \) which is a contradiction.

2. If not, there exists \( \tilde{q} \in \Gamma \) with \( L(\tilde{q}) < L(q) \). Since parametrization doesn’t change length, we can assume that \( \tilde{q} \) is parametrized proportional to arclength. But then, Lemma 17 implies

\[
\sqrt{2TE(\tilde{q})} < \sqrt{2TE(q)}
\]

which is a contradiction.

Assume for the moment, that the minimizers of \( E \) are sufficiently regular. Since

\[
E(c) = \int_0^T \sum_{i,j=1}^k g_{ij}(\gamma) \dot{\gamma}_i \dot{\gamma}_j dt
\]

we know how to derive the associated Euler-Lagrange equations

\[
\frac{d}{dt} \frac{\partial}{\partial \dot{\gamma}_l} \left( \sum_{i,j=1}^k g_{ij}(\gamma) \dot{\gamma}_i \dot{\gamma}_j \right) - \frac{\partial}{\partial \gamma_l} \left( \sum_{i,j=1}^k g_{ij}(\gamma) \dot{\gamma}_i \dot{\gamma}_j \right) = 0
\]

for \( 1 \leq l \leq k \). Carrying out the partial derivatives

\[
2 \frac{d}{dt} \sum_{j=1}^k g_{lj}(\gamma) \ddot{\gamma}_j - \frac{\partial}{\partial \gamma_l} \left( \sum_{i,j=1}^k g_{ij}(\gamma) \dot{\gamma}_i \dot{\gamma}_j \right) = 0
\]

which is equivalent to

\[
\sum_{j=1}^k g_{lj} \ddot{\gamma}_j + \sum_{i,j=1}^k \frac{\partial g_{lj}}{\partial \gamma_i} \dot{\gamma}_i \dot{\gamma}_j - \frac{1}{2} \sum_{i,j=1}^k \frac{\partial g_{ij}}{\partial \gamma_l} (\dot{\gamma}_i \dot{\gamma}_j) = 0 \tag{39}
\]

Since \( G \) is a positive definite, symmetric matrix, it is nonsingular, and therefore it admits an inverse matrix, \( G^{-1} \). Let \( G^{-1} = (g^{ij})_{i,j=1}^k \), where

\[
\sum_{j=1}^k g^{ij} g_{jl} = \delta_{il} = \begin{cases} 1 & \text{if } i = l \\ 0 & \text{if } i \neq l \end{cases}
\]

Multiplying all the terms in (39) by the inverse matrix, we obtain

\[
\ddot{\gamma}_l + \frac{1}{2} \sum_{i,j,p=1}^k g^{pl} \left( \frac{\partial g_{lj}}{\partial \gamma_i} \dot{\gamma}_i \dot{\gamma}_j - \frac{\partial g_{ij}}{\partial \gamma_l} \dot{\gamma}_i \dot{\gamma}_j \right) = 0
\]

Recall that the metric tensor depends of the partial derivatives of \( f \), which is at least \( C^3(\mathbb{R}^k, \mathbb{R}^n) \), and as a consequence, the coefficients of the differential equation are at least \( C^1 \) functions. Thus, we can conclude that the associated initial value problem, has a unique local solution.
Definition 20. The solutions of the above system of ODE’s, are called geodesics.

The usual form of the Geodesic equation is

$$\ddot{\gamma} + \sum_{i,j=1}^{k} \Gamma_{ij}^l(\gamma(t)) \dot{\gamma}_i \dot{\gamma}_j = 0$$

for $1 \leq l \leq k$, where

$$\Gamma_{ij}^l = \frac{1}{2} \sum_{p=1}^{k} g^{kp}(g_{pj,i} + g_{ip,j} - g_{ij,p})$$

and

$$g_{ij,l} = \frac{\partial g_{ij}}{\partial x_l}$$

The terms $\Gamma_{kj}^p$ are denoted the Christoffel Symbols.

Example 35. Compute the geodesics in sphere, i.e., for

$$M = S^2 = \{ x \in \mathbb{R}^3 : |x| = 1 \}$$

The solutions of the Euler equation are denoted Great Circles. (It takes 2 coordinate patches: $\Omega_1 = S^2 \setminus \{N\}$ and $\Omega_2 = S^2 \setminus \{S\}$.)

Let’s now summarize a little bit:

(1) If $M$ is a metric space, for $c \in C([0,T], M)$, we defined

$$L(c) = \sup_{P \in \mathcal{P}} L(Pf)$$

and showed that, if $M$ is complete, connected and compact, there is a curve of shortest length, joining two arbitrary points of $M$.

(2) If $M$ is a differentiable, $k$-dimensional manifold of $\mathbb{R}^n$, for $c \in AC([0,T], M)$, we defined

$$E(c) = \frac{1}{2} \int_0^T |\dot{c}(t)|^2 dt$$

and

$$L(c) = \int_0^T |\dot{c}(t)| dt$$

\[ = \sum_{\nu=1}^{p} \int_{t_{\nu-1}}^{t_{\nu}} \left( \sum_{i,j=1}^{k} g^{\nu}_{ij}(\gamma) \dot{\gamma}_i^{\nu} \dot{\gamma}_j^{\nu} \right)^{1/2} dt \]

where the intervals $(t_{\nu-1}, t_{\nu})$ were chosen so that the image of $c|_{t_{\nu-1}}^{t_{\nu}}$ lies entirely in one coordinate chart, i.e., $c(t) = f_{\nu}(\gamma_{\nu}(t))$ for $t \in [t_{\nu-1}, t_{\nu}]$, where $f_{\nu} : U_{\nu} \rightarrow M$ and $U_{\nu} \subset \mathbb{R}^k$. We showed that the minimizers of $E$ are minimizers of $L$ which are parametrized proportionally to arclength. Moreover, the $C^2$ minimizers of $E$ are solutions of the Euler equations, which are denoted by geodesics.
Next, we will bring these two notions together. By introducing a metric in $M$, our manifold becomes a metric space, and by our earlier theory, there is a curve of shortest length. We will show that the curve is a solution of the geodesic equation.

**Definition 21.** For $x, y \in M$, consider

$$\Gamma_{xy} = \{ q \in AC([0, T], M) : q(0) = x, q(T) = y \}$$

and define the distance

$$d(x, y) = \inf \{ L(c) : c \in \Gamma_{xy} \}$$

**Proposition 14.** As defined, $d(\cdot, \cdot)$ is a metric on $M$, i.e., satisfies

(i) For all $x, y \in M$, $d(x, y) \geq 0$ and $d(x, y) = 0$ iff $x = y$;

(ii) $d(x, y) = d(y, x)$ for all $x, y \in M$;

(iii) $d(x, z) \leq d(x, y) + d(y, z)$, for all $x, y, z \in M$.

**Proof:**

(i) Since $L(c) \geq 0$, for all $c \in \Gamma_{xy}$, it is obvious that $d(x, y) \geq 0$, and also, if $x = y$, $d(x, y) = 0$. Suppose now that $d(x, y) = 0$. Then there exists a sequence $(c_m) \subset \Gamma_{xy}$ such that

$$L(c_m) = \int_0^T |\dot{c}_m(t)| \, dt \to 0 \quad \text{as} \quad m \to \infty$$

Since, for all $m \in \mathbb{N}$, $c_m \in AC([0, t], M)$,

$$c_m(t) = c_m(0) + \int_0^t |\dot{c}_m(s)| \, ds, \quad \forall t \in [0, T]$$

In particular, for $t = T$

$$c_m(T) = c_m(0) + \int_0^T |\dot{c}_m(s)| \, ds$$

Also, by definition of $\Gamma_{xy}$, for all $m \in \mathbb{N}$, $c_m(T) = y$ and $c_m(0) = x$. Then

$$y = x + \int_0^T |\dot{c}_m(s)| \, ds$$

Letting $m \to \infty$, we obtain $x = y$.

(ii) This is obvious by reparametrization of the curves.

(iii) By definition

$$d(x, z) = \inf \{ L(c) : c \in \Gamma_{xz} \}$$

Consider, for $y \in M$

$$\tilde{\Gamma}_{xz} = \{ \gamma \in \Gamma_{xz} : \exists t_0 \in [0, T] \text{ such that } \gamma(t_0) = y \}$$
i.e., $\tilde{\Gamma}_{xz}$ is the subclass of $\Gamma_{xz}$ of the curves passing through $y$. It is obvious that

$$\inf \{ L(c) : c \in \tilde{\Gamma}_{xz} \} = \inf \{ L(c_1) : c_1 \in \Gamma_{xy} \} + \inf \{ L(c_2) : c_2 \in \Gamma_{yz} \} = d(x, y) + d(y, z)$$

But, $\tilde{\Gamma}_{xz} \subset \Gamma_{xz}$, and therefore

$$\inf \{ L(c) : c \in \tilde{\Gamma}_{xz} \} \geq \inf \{ L(c) : c \in \Gamma_{xz} \}$$

Hence, we conclude that the triangular inequality is verified by $d$.

If $M$ is a compact manifold, $M \subset \mathbb{R}^n$, gets the induced topology from $\mathbb{R}^n$ (the open sets in $M$ are $U \cap M$ where $U$ is open in $\mathbb{R}^n$). On the other hand once we introduce a metric in $M$, we define another topology in $M$. It can be proved that these two topologies are equivalent.

For $x \in M$ let $d_x = d(x, \partial f(U))$ and define $r_x = \frac{1}{3} d_x$. We can cover $M$ by $B_{r_i}(x)$, and since $M$ is compact we can find $x_1, ..., x_j \in M$ such that

$$M \subset B_{r_1}(x_1) \cup ... B_{r_j}(x_j)$$

Define $r_0 = \min_i r_i$.

**Theorem 17.** Let $M$ be a compact, connected, differentiable $k$-dimensional manifold of $\mathbb{R}^n$. If, for $x, y \in M$, $d(x, y) \leq r_0$ then there exists a shortest geodesic joining $x$ and $y$.

**Proof:**

For $x \in M$, there exists $x_i \in M$ such that $x \in B_{r_i}(x_i)$. Then

$$d(y, x_i) \leq d(y, x) + d(x, x_i) \leq r_0 + r_i \leq 2 \frac{d_x}{3}$$

and therefore $x, y$ lie in the same coordinate patch - the one associated to $x_i$.

Let us minimize $E(c)$ over $\Gamma_{xy}$, where

$$E(c) = \frac{1}{2} \int_0^T |\dot{c}(t)|^2 dt = \frac{1}{2} \int_0^T k \sum_{i,j=1}^k g_{ij}(\gamma) \dot{\gamma}_i \dot{\gamma}_j dt$$

and $\gamma : [0, T] \rightarrow U_{x_i}$ is defined by $c(t) = f(\gamma(t))$, with $(f, U)$ the local coordinates associated to $x_i$. Thus, we are minimizing $E$ over the class

$$\tilde{\Gamma}_{xy} = \{ \gamma \in W^{1,2}([0, T], \mathbb{R}^k) : \gamma(0) = f^{-1}(x), \gamma(T) = f^{-1}(y) \}$$

Pick a minimizing sequence $(\gamma_m) \subset \tilde{\Gamma}_{xy}$. Then

$$E(\gamma_m) = \frac{1}{2} \int_0^T \sum_{i,j=1}^k g_{ij}(\gamma_m) \dot{\gamma}_m^i \dot{\gamma}_m^j dt$$
Since the matrix $G$ is positive definite, there is $\beta > 0$, such that

$$E(\gamma_m) \geq \frac{1}{2} \beta \int_0^T |\dot{\gamma}_m|^2 dt$$

and therefore $(\dot{\gamma}_m)$ is a bounded sequence in $L^2([0, T], \mathbb{R}^k)$. On the other hand, $\gamma_m(0) = f^{-1}(x)$, for all $m \in \mathbb{N}$, and therefore, for $t \in [0, T]$

$$\gamma_m(t) = \gamma_m(0) + \int_0^t |\dot{\gamma}_m| dt$$

implying

$$|\gamma_m(t)| \leq |f^{-1}(x)| + T^{1/2} \|\dot{\gamma}_m\|_{L^2([0, T], \mathbb{R}^k)}$$

Then, $(\gamma_m)$ is bounded in $L^\infty([0, T], \mathbb{R}^k)$, and as a consequence $(\gamma_m)$ is bounded in $L^2([0, T], \mathbb{R}^k)$. Finally, $(\gamma_m)$ is bounded in $W^{1,2}([0, T], \mathbb{R}^k)$, and there is $\gamma \in W^{1,2}([0, T], \mathbb{R}^k)$ such that $\gamma_m \to \gamma$ weakly in $W^{1,2}([0, T], \mathbb{R}^k)$ and strongly in $L^\infty([0, T], \mathbb{R}^k)$. The strong convergence implies that $\gamma(t) \subset \overline{U}$ and $\gamma(0) = f^{-1}(x)$, $\gamma(T) = f^{-1}(y)$. Let us prove that $E$ is weakly lower semicontinuous. Suppose $\gamma_m \to \gamma$ in $W^{1,2}([0, T], \mathbb{R}^k)$ (which as seen before, implies $\gamma_m \to \gamma$ in $L^\infty([0, T], \mathbb{R}^k)$). Then

$$E(\gamma_m) = \frac{1}{2} \int_0^T \sum_{i,j=1}^k g_{ij}(\gamma_m) \dot{\gamma}_m^i \dot{\gamma}_m^j dt$$

$$= \frac{1}{2} \int_0^T \left( \sum_{i,j=1}^k g_{ij}(\gamma) \dot{\gamma}_m^i \dot{\gamma}_m^j + \frac{1}{2} \sum_{i,j=1}^k (g_{ij}(\gamma_m) - g_{ij}(\gamma)) \dot{\gamma}_m^i \dot{\gamma}_m^j \right) dt$$

The strong convergence implies

$$g_{ij}(\gamma_m) - g_{ij}(\gamma) \to 0 ; \quad m \to \infty$$

Thus

$$\int_0^T \sum_{i,j=1}^k (g_{ij}(\gamma_m) - g_{ij}(\gamma)) \dot{\gamma}_m^i \dot{\gamma}_m^j dt \to 0 \quad as \quad m \to \infty$$

since $\int_0^T \sum_{i,j=1}^k \dot{\gamma}_m^i \dot{\gamma}_m^j dt$ is uniformly bounded. Finally

$$\liminf E(\gamma_m) = \liminf \frac{1}{2} \int_0^T \sum_{i,j=1}^k g_{ij}(\gamma) \dot{\gamma}_m^i \dot{\gamma}_m^j dt$$

$$= E(\gamma) + \frac{1}{2} \liminf \int_0^T \sum_{i,j=1}^k g_{ij}(\gamma) (\dot{\gamma}_m^i \dot{\gamma}_m^j - \dot{\gamma}_m^i \dot{\gamma}_m^j) dt$$

$$\geq E(\gamma) + \beta \liminf \int_0^T |\dot{\gamma}_m - \dot{\gamma}|^2 dt$$

100
as a consequence of the positive definiteness of $G$. Finally, the weak convergence implies the result. If this is the case, 

$$E(\gamma) = \inf_{f} E$$

We still have to show that the minimizer of the energy (and of the length) satisfies the following:

$$\gamma \subset U_{x_i} \Rightarrow c \subset f(U_{x_i})$$

If that is not the case, there exists $t_0$ such that $c(t_0) \in \partial(f(U_{x_i}))$. On one hand

$$r_0 \geq d(x, y) = L(c) = L(c|_{t_0}^{0}) + L(c|_{t_0}^{1})$$

Notice, that $c|_{t_0}^{0}$ and $c|_{t_0}^{1}$ are also minimizers of the length, over the class of curves joining $x$ to $c(t_0)$ and $c(t_0)$ to $y$ respectively. Then

$$r_0 \geq d(x, c(t_0)) + d(c(t_0), y) \geq d(x, c(t_0)) \geq d_x$$

On the other hand, as a consequence of $c(t_0) \in \partial(f(U_{x_i}))$, we can choose $z \in \partial f(U_{x_i})$, such that $d(x, z) = d_x$ and $d(x_i, z) \geq d_{x_i}$. Then

$$d(x_i, z) \leq d(x, x_i) + d(x, z) = d(x, x_i) + d_x$$

implying

$$d_x \geq d(x_i, z) - d(x, x_i) \geq d_{x_i} - r_{x_i} = 2r_{x_i} \geq 2r_0$$

Then $r_0 \geq d_x$ and $d_x \geq 2r_0$ which is obviously contradictory.

\[\diamond\]

We still need to prove the regularity of the minimizer: once we do that it will be a solution of the geodesic equation, i.e., $c$ is really a shortest geodesic. Let us assume that for the moment. The previous Theorem was preparatory to the next one.

**Theorem 18.** Let $M$ be as in Theorem 17. Then, for all $x, y \in M$, there exist a shortest geodesic joining $x$ to $y$.

**Proof:**

As remarked before, we can assume that all the curves are parametrized proportionally to arclength. Consider $(c_{m}) \subset \Gamma_{xy}$ a minimizing sequence for $L$, i.e., $L(c_{m}) \subseteq d(x, y)$ as $m \to \infty$. We can assume that, for all $m \in \mathbb{N}$, $L(c_{m}) \leq d(x, y) + 1 \equiv K$. For $x_0, ... x_j \in M$ such that $\bigcup_i B_{r_i}(x_i) \supset M$, for fixed $m$, consider $P^m = (t_{0}^{m}, ..., t_{j}^{m})$ a partition of the interval $[0, T]$, such that

$$c_{m|t_{i-1}^{m}}^{t_{i}^{m}} \in f_{i}(u_{i})$$

where $(f_{i}, U_{i})$ are the local coordinates associated to $x_{i}$, $1 \leq i \leq j$. Moreover, we can assume

$$t_{i}^{m} - t_{i-1}^{m} \leq \frac{r_{0}}{K}$$

101
Since $c_m$ is parametrized proportional to arclength,
\[ d(c_m(t_{i-1}^m), c_m(t_i^m)) \leq K \frac{r_0}{K} = r_0 \]
We are in the hypothesis of Theorem 17, and conclude that $c_m(t_{i-1}^m)$ and $c_m(t_i^m)$ can be joined by a shortest geodesic. Define $\bar{c}_m : [0, T] \to M$ by the curve obtained by gluing all the shortest geodesics, i.e., for $t \in [t_{i-1}^m, t_i^m]$, $\bar{c}_m(t)$ is the shortest geodesic joining $c_m(t_{i-1}^m)$ to $c_m(t_i^m)$. By construction
\[ L(\bar{c}_m) \leq L(c_m) \]
and thus $(\bar{c}_m) \subset \Gamma_{xy}$ is a new minimizing sequence, which we can assume to be reparametrized proportional to arclength. Notice that for each $1 \leq i \leq j$, $(\bar{c}_m(t_i^m))$ is a bounded sequence, and therefore, as $m \to \infty$, $t_i^m \to t_i \in [0, T]$, possibly along a subsequence. On the other hand, for $s, t \in [0, T]$,
\[ d(\bar{c}_m(t), \bar{c}_m(t)) \leq L(\bar{c}_m)|s - t| \leq K|s - t| \]
implying that $(\bar{c}_m)$ is an equicontinuous sequence. On the other hand, for $t \in [0, T]$
\[ d(x, \bar{c}_m(t)) \leq k|t| \leq KT \]
and as a consequence $(\bar{c}_m)$ is uniformly bounded. By Ascoli-Arzela Theorem, $\bar{c}_m \to \bar{c}$ in $C([0, T], M)$, as $m \to \infty$. The continuity implies
\[ \bar{c}_m(t_i^m) \to \bar{c}(t_i) , \quad m \to \infty \]
We claim that $\bar{c}$ is the shortest geodesic joining $x$ to $y$. By construction, $\bar{c}$ is the shortest geodesic connecting $\bar{c}(t_{i-1})$ to $\bar{c}(t_i)$, for all $1 \leq i \leq j$. But we could find $s \in (t_{i-1}, t_i)$, $\sigma \in (t_i, t_{i+1})$ and a curve $\hat{c}|_{t_{i-1}}^{t_{i+1}}$ through $\bar{c}(s)$ and $\bar{c}(\sigma)$, such that
\[ L(\hat{c}|_{t_{i-1}}^{t_{i+1}}) < L(\bar{c}|_{t_{i-1}}^{t_i}) + L(\bar{c}|_{t_i}^{t_{i+1}}) \]
If this was the case, we could define a new partition $(t_0, ..., t_{i-1}, s, \sigma, t_{i+1}, ..., t_j)$ and a new minimizing sequence $(\hat{c}_m) \subset \Gamma_{xy}$, with
\[ L(\hat{c}_m) < L(c_m) \]
which is a contradiction to the definition of $c_m$. 

\[ \diamond \]
To finish these proofs, we have to show the regularity of the minimizers. We will do it by a series of Lemmas.

**Lemma 19. (DuBois-Raymond)** Let $f \in L^1([a, b], \mathbb{R})$ such that
\[ \int_a^b f(x) \varphi'(x) \, dx = 0 \]
for all $\varphi \in C_0([a, b], \mathbb{R})$, such that $\varphi(x)$ exists a.e. for $x \in [a, b]$. Then, there exist $c \in \mathbb{R}$, such that $f(x) = c$ almost everywhere for $x \in [a, b]$. 

102
Proof:
Recall, \( x \) is a Lebesgue point for \( f \), if for all family \( \Omega_x \) of measurable sets containing \( x \)
\[
\frac{1}{|\Omega_x|} \int_{\Omega_x} f(y) \, dy \to f(x) \quad \text{as} \quad |\Omega_x| \to 0
\]
Moreover, if \( f \in L^1([a,b], \mathbb{R}) \), almost all points in \([a,b]\) are Lebesgue points for \( f \). Consider \( y < z \) Lebesgue points in \((a,b)\), and choose
\[
\varphi(x) = \begin{cases} 
1 & \text{if } x \in (y, z) \\
0 & \text{if } x \in (a, y - \epsilon) \cup (z + \epsilon, b) \\
\frac{1}{\epsilon}(x - y + \epsilon) & \text{if } x \in (y - \epsilon, y) \\
\frac{1}{\epsilon}(z - x + \epsilon) & \text{if } x \in (z, z + \epsilon)
\end{cases}
\]
Then, \( \varphi \) is continuous in \([a,b]\), \( \varphi(a) = \varphi(b) = 0 \) and \( \varphi \in C([a,b] \setminus \{y,z\}, \mathbb{R}) \). By hypothesis
\[
0 = \int_a^b f(x) \varphi'(x) \, dx = \frac{1}{\epsilon} \int_y^{y-\epsilon} f(x) \, dx - \frac{1}{\epsilon} \int_z^{z+\epsilon} f(x) \, dx
\]
Letting \( \epsilon \to 0 \) and since \( y, z \) are Lebesgue points we obtain \( f(y) = f(z) \). If we let \( z \) change in \((y,b]\), we proved that \( f(x) = c \) for all Lebesgue points, and since they are dense in \([a,b]\), \( f(x) = c \) a.e. on \([a,b]\).

\[\square\]

Lemma 20. Suppose that \( v \in C^1([a,b], \mathbb{R}^n) \), \( F = F(t, u, p) \in C^1([a,b] \times \mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}) \) and
\[
\int_a^b \left( F_u(t, v, v') \cdot \varphi + F_p(t, v, v') \cdot \varphi' \right) \, dt = 0
\]
for all \( \varphi \in C^1_0([a,b], \mathbb{R}^n) \) (notice that we are requiring that the Fréchet derivative of \( I \) is 0). Then, there exists \( c \in \mathbb{R}^n \) such that
\[
F_p(t, v, v') = c + \int_a^t F_u(s, v(s), v'(s)) \, ds
\]
for all \( t \in [a,b] \)

Let us study first the meaning of the Lemma. Since \( F_u(s, v(s), v'(s)) \) is continuous, the integral \( \int_a^t F_u(s, v(s), v'(s)) \, ds \) is a differentiable function and as a consequence of the equality \( F_p(t, v, v') \) is a differentiable function (as a composition). Therefore we can differentiate both sides of the equation and obtain another version of the Euler equation
\[
\frac{d}{dt} F = F_u
\]
Nevertheless we can’t use the chain rule in the left hand side, since we know that the composite is in \( C^1 \), but we know nothing about the pieces.
Proof:
Integrating by parts
\[
\int_a^b F_u(t, v, v') \cdot \varphi(t) \, dt = \left( \int_a^t F_u(s, v, v') \, ds \right) \cdot \varphi(t)|_a^b - \int_a^b \int_a^t \left( F_u(t, v, v') \, ds \right) \cdot \varphi'(t) \, dt
\]
Since, by hypothesis, \( \varphi(a) = \varphi(b) = 0 \)
\[
\int_a^b F_u(t, v, v') \cdot \varphi(t) \, dt = - \int_a^b \left( \int_a^t F_u(s, v, v') \, ds \right) \cdot \varphi'(t) \, dt
\]
Then
\[
\int_a^b \left( F_p(t, v, v') - \int_a^t F_u(s, v, v') \, ds \right) \cdot \varphi(t) \, dt = 0
\]
for all \( \varphi \in C^1_0([a, b], \mathbb{R}^n) \). By Lemma 19, there exists \( c \in \mathbb{R}^n \) such that
\[
F_p(t, v, v') - \int_a^t F_u(s, v, v') \, ds = c
\]
\( \Box \)

Homework 24. Prove that Lemma 20 is verified a.e. in \( t \), if we replace \( v \in C^1([a, b], \mathbb{R}^n) \) by \( v \) lipschitz continuous form \([a, b] \) to \( \mathbb{R}^n \). (note that \( |v'| \leq k \)).

Consider \( F = F(t, u, p) \in C^2([a, b] \times \mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}) \) verifying the conditions

(F1) there are positive constants, \( \alpha \) and \( \beta \), such that
\[
\alpha|p|^2 \leq F(t, u, p) \leq \beta|p|^2
\]
for all \( (t, u, p) \in \Omega \equiv [a, b] \times \mathbb{R}^n \times \mathbb{R}^n \);

(F2) there exists \( M(R) > 0 \), such that for \( (t, u, p) \in \Omega \) and \( |u| < R \)
\[
|F_u(t, u, p)| + |F_p(t, u, p)| \leq M(R)(1 + |p|^2)
\]
i.e., we assume a growth condition on the derivatives;

(F3) for all \( (t, u, p) \in \Omega \) and \( \xi \in \mathbb{R}^n \setminus \{0\} \)
\[
\sum F_{p_i p_j}(t, u, p)\xi_i \xi_j > 0
\]
which is a convexity assumption.
You can easily check, that

\[ F(t,u,p) = \sum_{i,j=1}^{k} g_{ij} \dot{p}_i \dot{p}_j \]

verifies conditions (F1)-(F3). Suppose further, that \( v \) is a local minimum of

\[ I(u) = \int_a^b F(t,u,u') \, dt \]

in \( W^{1,2}([a,b], \mathbb{R}^n) \) satisfying some prescribed boundary conditions.

**Proposition 15.** If \( F \) verifies (F1)-(F3) and \( v \in C^1([a,b], \mathbb{R}^n) \), then \( v \in C^2([a,b], \mathbb{R}^n) \).

**Proof:**

Since \( v \) is a local minimum of \( I \), \( I'(v) = 0 \), and Lemma 20 implies

\[ F_p(t,v(t),v'(t)) = c + \int_a^t F_u(s,v(s),v'(s)) \, ds \equiv f(t) \quad (40) \]

By hypothesis, \( F_u(s,v(s),v'(s)) \) is continuous, and the Fundamental Theorem of calculus allow us to conclude that \( F_p(t,v(t),v'(t)) \) is continuously differentiable for \( t \in (a,b) \). Define

\[ A(t,p) = F_p(t,v,p) - f(t) \]

As remarked, \( A \in C^1([a,b] \times \mathbb{R}^n, \mathbb{R}^n) \) and by (40)

\[ A(t,v') = 0 \quad \forall t \in [a,b] \]

On the other hand

\[ A_p(t,v') = F_{pp}(t,v,v') \]

and by hypothesis (F3), the matrix \( F_{pp}(t,v,v') \) is nonsingular. The Implicit Function Theorem implies that for each \( t_0 \in [a,b] \), there exists \( \xi_0 \in C_1(N_0, \mathbb{R}) \) such that, the solutions of the equation \( A(t,p) = 0 \) near \( (t_0,v'(t_0)) \) are given by \( (t,\xi_0(t)) \) for \( t \) near \( t_0 \). Since the zeros of \( A \) are precisely \( (t,v'(t)) \), we conclude \( \xi_0(t) = v'(t) \) for \( t \) near \( t_0 \). The differentiability of \( \xi_0 \) implies \( v' \in C^2([a,b], \mathbb{R}^n) \) as wanted.

\[ \diamond \]

**Theorem 19.** If \( F \in C^2([a,b] \times \mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}) \) verifies (F1)-(F3), and \( v \) is a local minimum of \( I \), then \( v \in C^2([a,b], \mathbb{R}^n) \) and satisfies the associated Euler equation. Moreover \( F \in C^k \) implies \( v \in C^k \).

**Proof:**

If \( v \) is a local minimum of \( I \), there is a real constant \( M \) such that \( I(v) \leq M \). Using (F1)

\[ M \geq \alpha \int_a^b |v'(t)|^2 \, dt \]

105
implying that \( v' \in L^2([a, b], \mathbb{R}^n) \). As we have seen before in previous examples, if \( v \) verifies some boundary conditions and \( v' \in L^2([a, b], \mathbb{R}^n) \), then \( v \in W^{1,2}([a, b], \mathbb{R}^n) \). By the Sobolev Theorem, we conclude \( v \in L^\infty([a, b], \mathbb{R}^n) \). Using hypothesis (F2) and the Dominated Convergence Theorem, we conclude that both \( F_u(t, v, v') \) and \( F_p(t, v, v') \) are bounded in \( L^1([a, b], \mathbb{R}^n) \). Since \( I'(v) = 0 \), \( v \) is a solution of the weak Euler equation

\[
\int_a^b (F_u(t, v, v') \varphi + F_p(t, v, v') \varphi') \, dt = 0 \quad \forall \varphi \in C_0^1([a, b], \mathbb{R}^n)
\]

and by Lemma 20, there exists a real number, \( c \), for which

\[
F_p(t, v, v') = c + \int_a^t F_u(s, v, v') \, ds
\]  

We claim that the map \( F_p(\cdot, \cdot, p) : \mathbb{R}^n \rightarrow \mathbb{R}^n \), is a global diffeomorphism. For fixed \( p_0 \), (F3) implies that \( F_{pp}(\cdot, \cdot, p_0) \) is a nonsingular matrix, and as a consequence \( F_p(\cdot, \cdot, p) \) is a local diffeomorphism. All we have to show is that it is globally 1-1 and onto. If it is not globally 1-1, there are \( p, p_1 \in \mathbb{R}^n \) such that

\[
0 = (F_p(\cdot, \cdot, p) - F_p(\cdot, \cdot, p_1)) \cdot (p - p_1)
\]

Using hypothesis (F3), we obtain the absurd \( 0 > 0 \). Thus \( F_p(\cdot, \cdot, p) \) is globally 1-1 (notice that this is a direct consequence of the convexity assumption). To prove that is onto, for \( w \in \mathbb{R}^n \), we want to show that there exists \( z \in \mathbb{R}^n \) (which will be unique as a consequence of the injectivity), such that

\[
F_p(\cdot, \cdot, z) = w
\]

We will prove this fact, minimizing \( F(\cdot, \cdot, x) - w \cdot x \) over \( \mathbb{R}^n \). Hypothesis (F1) implies that \( F(t, u, p) \) grows at least quadratically in \( p \), and therefore \( F(\cdot, \cdot, x) - wx \) is bounded by below. As a consequence, it admits a minimizer which we will denote by \( z \). On the other hand \( F(\cdot, \cdot, x) - wx \) is a continuously differentiable function, implying

\[
(F(\cdot, \cdot, x) - wx)'|_{x=z} = 0
\]

Then

\[
F_p(\cdot, \cdot, z) = w
\]

as wanted, and we proved that \( F_p(\cdot, \cdot, p) \) is a global diffeomorphism. If this is the case, equation (41) becomes

\[
w(t) = c + \int_a^t F_u(t, v, v') \, dt
\]
for a unique $z$ such that

$$F_p(t, v(t), z(t)) = w(t)$$

As we defined it, $w \in C([a, b], \mathbb{R}^n)$. We will show that also $z \in C([a, b], \mathbb{R}^n)$. Let $(t_m) \subset [a, b]$ such that $t_m \to t$, as $m \to \infty$. By (F1)

$$\alpha|z(t_m)|^2 - wz(t_m) \leq F(t_m, v(t_m), z(t_m)) - wz(t_m)$$

and since $z(t_m)$ is a minimizer of $F(\cdot, \cdot, p) - wp$

$$F(t_m, v(t_m), z(t_m)) - wz(t_m) \leq F(t_m, v(t_m), 0) \leq c$$

Then

$$\alpha|z(t_m)|^2 - wz(t_m) \leq c$$

implying that $(z(t_m))$ is bounded in $\mathbb{R}^n$, and therefore it will converge (possibly along a subsequence), for some $\zeta \in \mathbb{R}^n$. On the other hand,

$$F_p(t_m, v(t_m), z(t_m)) = w(t_m)$$

and letting $m \to \infty$

$$F_p(t, v(t), \zeta) = w(t)$$

and, by the definition of $w$, $\zeta$ is in fact the minimizer of $F(t, v(t), x) - w(t)x$, implying that $\zeta = z(t)$, and therefore $z$ is continuous. Moreover, $v'(t) = z(t)$ a.e. for $t \in [a, b]$. Finally, since $v \in W^{1,2}([a, b], \mathbb{R}^n)$, for all $t \in [a, b]$

$$v(t) = v(a) + \int_a^t |v'(s)| ds = v(a) + \int_a^t z(s) \, ds$$

The Fundamental theorem of calculus and the continuity of $z$, imply $v \in C^1([a, b], \mathbb{R}^n)$, and Proposition 15, implies $v \in C^2([a, b], \mathbb{R}^n)$.}

\[ \diamond \]

### 6.2 Geodesics in the Torus

Consider the $n$ torus $\Pi^n = \mathbb{R}^n \setminus \mathbb{Z}^n$. We are interested in finding closed geodesics in the torus, i.e., curves $\gamma \in C([a, b], \Pi^n)$, such that, for some $a, b \in \mathbb{R}^n$, $\gamma(a)$ is identified with $\gamma(b)$. In $\mathbb{R}^n$, they correspond to curves $\tilde{\gamma} \in C(\mathbb{R}, \mathbb{R}^n)$, verifying $\tilde{\gamma}_j = \tilde{\gamma} + j$ for some $j \in \mathbb{Z}^n$. For each $k \in \mathbb{Z}^n \setminus \{0\}$ define a homotopy class of the curves winding $k_i$ times in the $x_i$ direction, and suppose that the metric tensor $g_{ij}(x)$ is smooth, and satisfies $g_{ij}(x + e_l) = g_{ij}(x)$ for all $l \in \mathbb{Z}^n$ (is periodic in the coordinate directions). We can use $g$ to define a metric tensor

$$L(\gamma) = \int_a^b \sqrt{\sum g_{ij}(\gamma) \dot{\gamma}_i \dot{\gamma}_j} dt$$
for $\gamma \in AC([a, b], \Pi^n)$ or $\gamma \in W^{1,2}([a, b], \Pi^n)$. As a consequence

$$d(x, y) = \inf\{L(\gamma) : \gamma \in W^{1,2}([a, b], \mathbb{R}^n), \gamma(a) = x, \gamma(b) = y\}$$

induces a metric in $\Pi^n$ or in $\mathbb{R}^n$, and we can extend this notion to the length of a continuous function. If $\gamma \in C([0, 1], \mathbb{R}^n)$ then

$$L(\gamma) = \sup_P \sum d(\gamma(t_{i-1}), \gamma(t_i))$$

for any partition $P$ of the interval $[0, 1]$. The next Theorem states the existence of a curve of homotopy type $k$ of shortest length.

**Theorem 20.** For each $k \in \mathbb{Z}^n \setminus \{0\}$, define $\Gamma_k$ as the set of absolutely continuous curves of homotopy type $k$, i.e.,

$$\Gamma_k = \{\gamma \in W^{1,2}(\mathbb{R}, \mathbb{R}^n) : \exists T = T(\gamma) \text{ for which } \gamma(t + T) = \gamma(t) + k\}$$

(after time $T$ they differ by $k$). Then there exists $\gamma^*_k \in \Gamma_k$ such that

$$c_k = \inf_{\gamma \in \Gamma_k} L(\gamma) = L(\gamma^*_k)$$

**Proof:**

We start by noticing that if $\gamma \in \Gamma_k$, then $\tau_\theta \gamma(t) = \gamma(t - \theta) \in \Gamma_k$ for all $\theta \in \mathbb{Z}^n$. As a consequence, we have the freedom to choose $\gamma(0)$. We will assume $\gamma(0) \in [0, 1]^n$, and consider that all the curves are parametrized with respect to arclength. Let $(\gamma_m) \subset \Gamma_k$ be a minimizing sequence. Then, for some positive $M$, $L(\gamma_m) \leq M$, for all $m \in \mathbb{N}$, and by Lemma 12, for all $s, t \in \mathbb{R}$

$$d(\gamma_m(t), \gamma_m(s)) \leq L(\gamma_m)|s - t| \leq M|t - s|$$

We can conclude that $(\gamma_m)$ is an equicontinuous sequence with $L(\gamma_m)$ uniformly bounded. By Ascoli-Arzela Theorem, there exists $\gamma^*_k \in W^{1,2}(\mathbb{R}, \mathbb{R}^n)$ such that $\gamma_m \to \gamma^*_k$ uniformly, possibly along a subsequence. Since $L$ is lower semicontinuous

$$L(\gamma^*_k) \leq \liminf L(\gamma_m) = c_k$$

(42) implies $L(\gamma^*_k) = c_k$.

Since we proved that any minimizer of $L$ is $C^2([a, b], \mathbb{R}^n)$, for all $a, b \in \mathbb{R}$, $\gamma^*_k \in C^2(\mathbb{R}, \mathbb{R}^n)$, and consequently is a solution of the geodesic equation.

108
Definition 22. Two geodesics on $\Pi^n$, parametrized by arclength, will be called geometrically distinct if they do not differ by a phase shift.

If $\gamma^*_k \in \Gamma_k$ is a minimizer of $L$,

$$c_k = L(\gamma^*_k) = L(\gamma^*_k + j)$$

for any $j \in \mathbb{Z}^n$, which means, that in $\mathbb{R}^n$, there is an infinite number of representatives, geometrically distinct, for the minimizer of $L$ over $\Gamma_k$. Consider the case $n = 2$, and define

$$M_k = \{ \gamma \in \Gamma_k : L(\gamma) = c_k \}$$

Definition 23. A geodesic, $\gamma$, is said to be minimal if for all $a < b \in \mathbb{R}$, $\gamma|^{b}_a$ minimizes $L(\cdot)$ over the class of curves in $\mathbb{R}^n$ joining $a$ to $b$, that is

$$L(\gamma|^{b}_a) = d(\gamma(a), \gamma(b))$$

Theorem 21. (Morse’s Theorem)

If $\gamma^*_k \in M_k$, then $\gamma^*_k$ is a minimal geodesic.

To prove Theorem 21, we need a preliminary result.

Proposition 16. For $l \in \mathbb{N}$

$$c_{lk} = lc_k$$

Proof:

Choose some homotopy type $k = (k_1, k_2)$, and $\gamma^*_k$ a geodesic for $k$. It is obvious, that $\gamma^*_k$ runned $l$ times, is an element of $\Gamma_{lk}$, and as a consequence

$$c_{lk} = \inf_{\gamma \in \Gamma_{lk}} L(\gamma) \leq lc_k = L(\gamma^*_k \text{ runned } l \text{ times}) \tag{43}$$

We will prove, by induction on $l$, that $c_{lk} = lc_k$. The case $l = 1$ is trivially true. For $l = 2$, (43) implies

$$2c_k = \int_0^{2T(\gamma^*_k)} \sqrt{g(\gamma^*_k, \dot{\gamma}^*_k)} dt \geq c_{2k}$$

We will prove, that also

$$c_{2k} \geq 2c_k$$

Let $\gamma^*_{2k}$ be the minimizer of $L$ over $\Gamma_{2k}$, and assume that we can find a pair of numbers $\tau < s$, $\tau, s \in [0, T(\gamma^*_{2k})]$, such that

$$\gamma^*_{2k}(\tau) + k = \gamma^*_{2k}(s) \tag{44}$$

By definition, $\gamma^*_{2k}|^s_\tau \in \Gamma_k$, and as a consequence

$$L(\gamma^*_{2k}|^s_\tau) \geq c_k$$
If we consider the curve \( \tilde{\gamma} \) obtained by gluing \( \gamma_{2k}^\ast \vert_\tau \) to \( \gamma_{2k}^\ast \vert_\tau^{T(\gamma_{2k}^\ast + \sigma)} \), it is easy to verify that
\[
\tilde{\gamma}(T(\gamma_{2k}^\ast) + \tau) = \tilde{\gamma}(\tau) + 2k = \tilde{\gamma}(s) + k
\]
and therefore
\[
\gamma_{2k}^\ast \vert_{T(\gamma_{2k}^\ast) + \sigma} \in \Gamma_k
\]
implies
\[
L(\gamma_{2k}^\ast \vert_{T(\gamma_{2k}^\ast) + \sigma}) \geq c_k
\]
Then
\[
c_{2k} = L(\gamma_{2k} = L(\gamma_{2k}^\ast \vert_\tau) + L(\gamma_{2k}^\ast \vert_{T(\gamma_{2k}^\ast) + \sigma}) \geq 2c_k
\]
as wanted. This proves the case \( l = 2 \), assuming we can find \( \tau, s \) verifying property 44. Denote \( \gamma \) by \( \gamma_{2k}^\ast \) for simplicity in notation. \( \gamma \) lies between two lines of slope \( \frac{\sigma}{k} \), that we will denote by \( L_1 \) and \( L_2 \), respectively. Shift the lines until tangent to \( \gamma \). Without loss of generality assume \( \gamma(0) \in L_1 \). By construction there exists
\[
\sigma = \min\{t \in (0, T(\gamma)) : \gamma(t) \in L_2\}
\]
If \( \gamma(0) + k \in \gamma \) then we are done. Otherwise \( \gamma(0) + k \not\in \gamma \), but notice that \( \gamma(t) + k \) lies in the region bounded by \( \gamma(t) \) and \( L_1 \) for \( t \) near 0, and is in the region bounded by \( \gamma(t) \) and \( L_2 \) for \( t \) near \( \tau \). By continuity it must intersect \( \gamma \) for some \( \theta \in [0.\tau] \), which proves the result for \( l = 2 \).

For the general case, if the result is true for \( l \in \{1, \ldots, m - 1\} \) we want to show that
\[
c_{mk} = mc_k
\]
Once again it is obvious that \( mc_k \geq c_{mk} \), since \( \gamma_{k}^\ast \) runned \( m \) times belongs to \( \Gamma_{mk} \). To obtain the reverse inequality, it suffices to show the existence of \( 0 < \tau < s \leq T(\gamma_{mk}^\ast) \) such that
\[
\gamma_{mk}^\ast(\tau) + k = \gamma_{mk}^\ast(s)
\]
If this is the case, \( \gamma_{mk}^\ast \vert_{\tau} \in \Gamma_k \) (since the end points are \( k \)), and also
\[
\gamma_{mk}^\ast \vert_{T(\gamma_{mk}^\ast) + \tau} \in \Gamma_{(m-1)k}
\]
By induction
\[
c_{mk} = L(\gamma_{mk}^\ast) = L(\gamma_{mk}^\ast \vert_{\tau}) + L(\gamma_{mk}^\ast \vert_{T(\gamma_{mk}^\ast) + \tau}) \geq c_k + (m - 1)c_k = mc_k
\]
To find \( \tau \) and \( s \) we argue as for the case \( l = 2 \).

\[\diamondsuit\]

We can now prove Morse’s Theorem (Theorem 16)

**Proof:**
Take any pair of points \( a, b \in \mathbb{R}, a < b \). We want to show that
\[
d(\gamma^*_k(a), \gamma^*_k(b)) = L(\gamma|_a^b_k)
\]
where \( \gamma^*_k \) minimizes \( L \) over \( \Gamma_k \). If not, by definition, there exists \( \gamma : [a, b] \to \mathbb{R}^2 \), such that
\[
d(\gamma^*_k(a), \gamma^*_k(b)) = L(\gamma|_a^b_k) < L(\gamma|_a^b_k)
\]
and \( \gamma(a) = \gamma^*_k(a), \gamma(b) = \gamma^*_k(b) \). We will consider two distinct cases.

(i) Assume that, for some \( l \in \mathbb{N} \)
\[
\gamma^*_k(b) = \gamma^*_k(a) + lk
\]
Then
\[
L(\gamma|_a^b_k) = lc_k
\]
and by Proposition 16,
\[
L(\gamma|_a^b_k) = cl_k
\]
which implies that \( \gamma|_a^b_k \) minimizes \( L \) over \( \Gamma_{lk} \). Since by construction \( \gamma(a) = \gamma^*_k(a) \) and \( \gamma(b) = \gamma^*_k(b) \), also \( \gamma \in \Gamma_{lk} \). Thus
\[
c_{lk} \leq L(\gamma|_a^b_k) < L(\gamma|_a^b_k) = c_{lk}
\]
is an absurd.

(ii) Assume that, for all \( l \in \mathbb{N} \)
\[
\gamma^*_k(b) \neq \gamma^*_k(a) + lk
\]
If this is the case, extend \( \gamma^*_k \) to the interval \([b, t]\), such that
\[
\gamma^*_k(t) = \gamma^*_k(a) + lk
\]
for some \( l \in \mathbb{N} \). Consider the curve \( \gamma : [a, t] \to \mathbb{R}^2 \), defined by gluing \( \gamma|_a^b \) to \( \gamma|_a^b \). By construction \( \gamma \in \Gamma_{lk} \), and
\[
L(\gamma) = L(\gamma|_a^b) + L(\gamma|_a^b) \geq cl_k
\]
On the other hand
\[
L(\gamma) < L(\gamma|_a^b) = lc_k = c_{lk}
\]
These two inequalities imply that \( c_{lk} < c_{lk} \), which is obviously a contradiction.

\[\Box\]

**Definition 24.** A geodesic \( \gamma : \mathbb{R} \to \mathbb{R}^n \), which is asymptotic to geodesics \( \gamma^\pm \) as \( t \to \pm \infty \) is called a heteroclinic geodesic from \( \gamma^- \) to \( \gamma^+ \).

**Theorem 22.** (Morse-Hedlund)

For any \( k \in \mathbb{Z}^2 \setminus \{0\} \), let \( v^+, v^- \) be adjacent minimizers of \( L \) in \( M_k \). Then there exists a minimal geodesic heteroclinic from \( v^+ \) to \( v^- \) and a minimal geodesic heteroclinic from \( v^- \) to \( v^+ \).
We will prove Theorem 22, using a variational argument. Without loss of generality, assume \( T(v^\pm) = 1 \). For \( k \) fixed, pick another homotopy type \( l \neq k \). Then a curve of homotopy type \( l \), which we will denote by \( z_0 \) (and whose existence is guaranteed by Theorem 20, will be transversal to \( v^\pm \). We will assume, \( z_0(0) \in v^+(\mathbb{R}) \) and \( z_0(T(z_0)) = v^-(0) \). Moreover, for \( i \in \mathbb{Z} \), consider

\[
z_i = z_0 + ik
\]
i.e., the translates of \( z_0 \). Define \( \mathcal{R}_i \) as the region of \( \mathbb{R}^2 \), with boundaries \( v^+ \), \( z_{i+1} \), \( v^- \) and \( z_i \) and consider

\[
\Pi_i = \{ \gamma_i \in W^{1,2}([0,1],\mathbb{R}^2) : \gamma_i(0) \in z_i \ , \gamma_i(1) \in z_{i+1} \text{ and } \gamma_i([0,1]) \subset \mathcal{R}_i \}
\]
Furthermore, assume that, for all \( i \in \mathbb{Z} \), \( \gamma_i \) is parametrized proportionally to arclength. Let

\[
\Pi = \{ \gamma = (\gamma_i)_{i \in \mathbb{Z}} : \gamma \text{ verifies } (\Pi^1) - (\Pi^3) \}
\]
where

\[
(\Pi^1) \gamma_i \in \Pi_i \text{ for all } i \in \mathbb{Z};
\]

\[
(\Pi^2) \gamma_{i+1}(0) = \gamma_i(1) \text{ for all } i \in \mathbb{Z};
\]

\[
(\Pi^3) \text{ For } z_i(t_i(\gamma_i)) = \gamma_i(0) \text{ and } z_{i+1}(t_{i+1}(\gamma_i)) = \gamma_i(1) \text{ we require that }
\]

\[
t_{i+1}(\gamma) \leq t_i(\gamma) \quad \forall i \in \mathbb{Z}
\]
Notice that \( \Pi_i \) is endowed with the maximum topology and \( \Pi \) is endowed with the product topology. Define the length functional over this class (which as before is a renormalized functional), as

\[
J(\gamma) = \sum_{i \in \mathbb{Z}} (L(\gamma_i) - c_k)
\]
where \( c_k = \inf_{\Gamma_k} L = L(v^\pm) \).

**Lemma 21.**

1) There exists \( \gamma \in \Pi \) such that \( J(\gamma) < \infty \).

2) \( J \) is bounded by below on \( \Pi \).

3) If \( \gamma \in \Pi \) is such that \( J(\gamma) \leq M \) then

\[
\sum_{i \in \mathbb{Z}} |L(\gamma_i) - c_k| \leq M + 2c_l
\]
where \( c_l = \inf_{\Gamma_l} L = L(z_0) \).

4) \( J \) is weakly lower semicontinuous.
Proof:

1) Consider $\gamma$ as the curve obtained by gluing $v^-|_{-\infty}^0$ to $z_0|^T_0$ to $v^+|^{\infty}_1$. It is obvious that $\gamma \in \Pi$, and

$$J(\gamma) = \int_0^T (L(z_0) - c_k) \, dt = c_l - c_k$$

2) Let $\alpha_i = \gamma_i(0)$, $\alpha_{i+1}=\gamma_{i+1}(1)$ and consider $\beta_{i+1} = \alpha_1 + k$ (the translate by $k$ of $\alpha_i$). Then $\gamma_i + z_{i+1}|_{\alpha_{i+1}}^{\beta_{i+1}}$ is an element of $\Gamma_k$, and as a consequence

$$L(\gamma_i) + L(z_{i+1}|_{\alpha_{i+1}}^{\beta_{i+1}}) \geq c_k$$

implying

$$L(\gamma_i) - c_k \geq -L(z_{i+1}|_{\alpha_{i+1}}^{\beta_{i+1}})$$

Summing on $i \in \mathbb{Z}$, we obtain

$$J(\gamma) \geq -\sum_{i} L(z_{i+1}|_{\alpha_{i+1}}^{\beta_{i+1}}) \geq -L(z_0) = -c_l \quad \text{(45)}$$

3) Consider

$$N^- = \{i \in \mathbb{Z} : L(\gamma_i) - c_k \leq 0\}$$

For $i \in N^-$, using (45)

$$-\sum_{i \in N^-} (L(\gamma_i) - c_k) \leq c_l$$

For $n \in \mathbb{N}$ fixed

$$\sum_{i=-n}^{n} (L(\gamma_i) - c_k) = \sum_{i \in N^-, |i| \leq n} (L(\gamma_i) - c_k) + \sum_{i \in N^+, |i| \leq n} (L(\gamma_i) - c_k)$$

where $N^+ = \mathbb{Z} \setminus N^-$. Hence

$$\sum_{i \in N^+, |i| \leq n} (L(\gamma_i) - c_k) \leq J(\gamma) - \sum_{|i| > n} (L(\gamma_i) - c_k) + c_l$$

Since, by hypothesis, $J(\gamma) < M$, for $\epsilon > 0$ arbitrary, we can choose $n$ sufficiently large, so that

$$-\sum_{|i| > n} (L(\gamma_i) - c_k) \leq \epsilon$$

and as a consequence

$$\sum_{i \in N^+, |i| \leq n} (L(\gamma_i) - c_k) \leq M + \epsilon + c_l$$

113
Finally
\[ \sum_{i=-n}^{n} |L(\gamma_i) - c_k| = \sum_{i \in N^+, |i| \leq n} (L(\gamma_i) - c_k) - \sum_{i \in N^-, |i| \leq n} (L(\gamma_i) - c_k) \leq M + \epsilon + 2c_l \]
and letting \( \epsilon \to 0 \)
\[ \sum_{i=-n}^{n} |L(\gamma_i) - c_k| \leq M + 2c_l \]
We conclude
\[ \sum_{i \in \mathbb{Z}} |L(\gamma_i) - c_k| = \lim_{n \to \infty} \sum_{i=-n}^{n} |L(\gamma_i) - c_k| \leq M + 2c_l \]
as wanted.

4) Let \( \gamma^m \to \gamma \) in \( \Pi \), as \( m \to \infty \). For \( \epsilon > 0 \) arbitrary, there exists \( j_0 = j_0(\epsilon) \in \mathbb{N} \), such that
\[ \sum_{i=|j|}^{j} (L(\gamma_i) - c_k) \geq J(\gamma) - \epsilon \]
for all \( j \geq j_0 \). By the product topology, for all \( i \in \mathbb{Z} \), \( \gamma^m_i \to \gamma_i \) in \( L^\infty([0,1],\mathbb{R}^2) \), as \( m \to \infty \). Then
\[ \liminf_{m \to \infty} \sum_{i=|j|}^{j} (L(\gamma^m_i) - c_k) \geq \sum_{i=|j|}^{j} (L(\gamma_i) - c_k) \geq J(\gamma) - \epsilon \]
since \( L \) is weakly lower semicontinuous. Let \( j \to \infty \), we obtain
\[ \liminf_{m \to \infty} J(\gamma^m) \geq J(\gamma) - \epsilon , \quad \forall \epsilon > 0 \]
implying
\[ \liminf_{m \to \infty} J(\gamma^m) \geq J(\gamma) \]
as wanted.
\[ \Box \]
Note that, by property (3), if \( J(\gamma) \) is convergent, then the series is absolutely convergent.

**Proposition 17.** If \( \gamma \in \Pi \) and \( J(\gamma) < \infty \), then
\[ \|\gamma_j - v^+\|_{L^\infty([j,j+1])} \to 0 \text{ as } j \to \infty; \]
\[ \|\gamma_j - v^-\|_{L^\infty([j,j+1])} \to 0 \text{ as } j \to -\infty. \]

**Proof:**

114
For \( j > 0 \), we can consider \( \gamma_j \in C([0,1]) \) (just translate them). Since \( J(\gamma) < \infty \), \( L(\gamma_j) \to c \) as \( j \to \infty \). By construction, for all \( j \in \mathbb{Z} \), \( \gamma_j \in \mathcal{R}_j \), all the functions are uniformly bounded, and they are all parametrized proportionally to arclength. Hence

\[
L(\gamma_j|_s^t) \leq L(\gamma_j)|t-s|
\]

and as a consequence \((\gamma_j)\) are uniformly bounded and equicontinuous. Ascoli-Arzella Theorem imply the existence of \( \tilde{\gamma} \) such that \( \gamma_j \to \tilde{\gamma} \) as \( j \to \infty \). Since \( L \) is weakly lower semicontinuous

\[
L(\tilde{\gamma}) \leq \lim \inf L(\gamma_j) = c
\]

Moreover \( \tilde{\gamma} \in M_k \) (since also \( \tilde{\gamma}(1) - \tilde{\gamma}(0) = k \)). Then

\[
L(\tilde{\gamma}) = c
\]

which implies that \( \tilde{\gamma} \) is a minimizer of \( L \) in \( \mathcal{R}_0 \). By construction of \( \mathcal{R}_0 \), \( \tilde{\gamma} \) can only be one of \( v^+ \) or \( v^- \) (they are adjacent curves in \( M_k \)). By construction of the sequence \( s_j \) it must necessarily be \( v^+ \). By the uniqueness of limit all the sequence must converge.

\[\Box\]

**Proof:**

Let \((\gamma^m)\) be a minimizing sequence. Then there exist \( M \) such that

\[
m \geq J(\gamma^m) \setminus c^+
\]

By (3) of Lemma 0

\[
\sum |L(\gamma^m_j) - c| \leq M + 2L(z)
\]

implying that

\[
|L(\gamma^m_j) - c| \leq M + 2L(z)
\]

for all \( m,j \). For fixed \( j \)

\[
L(\gamma^m_j) \leq c + M + 2L(z) \quad \forall m
\]

This is a uniform bound for \( L(\gamma^m_j) \) in \( m \), which gives us an upper bound in the Lipschitz constant. Furthermore the fact that, by construction \( \gamma^m_j \in \mathcal{R}_j \) gives us an \( L^\infty \) bound. Then we can use Ascoli-Arzella, for each \( j \) there exists \( \gamma^*_j \in AC([0,1]) \) such that \( \gamma^m_j \to \gamma^*_j \). We want to show that \( \gamma^*_j \in \Pi_j \). By continuity, it is obvious that \( \gamma^*_j(0) \in Z_j \) and \( \gamma^*_j(1) \in z_{j+1} \). If also \( \gamma^*_j \) is parametrized proportional to arclength then \( \gamma^*_j \in \Pi_j \). For that purpose, it suffices to show that

\[
L((\gamma^*_j|_0^t) = tL(\gamma^*_j)
\]

for all \( t \in [0,1] \). Since \( \gamma^*_j \to \gamma^* \) in \( L^\infty \), we will have \( L(\gamma^*_j) \leq \lim \inf L(\gamma^m_j) \). We claim that equality is verified. If not, then

\[
L(\gamma^*_j) < \lim \inf L(\gamma^m_j)
\]

115
Since we already know that
\[ \gamma_j^*(0) = \lim \gamma_j^m(0) \quad \text{and} \quad \gamma_j^*(1) = \lim \gamma_j^m(1) \]
we can choose \( m \) large enough so that
\[ L(\gamma_j^*) + \epsilon < \lim \inf L(\gamma_j^m) \]
Then we can replace \( \gamma_j^m \) by \( \gamma_j^*(0) \) glued to \( \gamma_j^*(0) \) to \( \gamma_j^*(1) \). This is a new minimizing sequence, denoted by \( \tilde{\gamma}_j^m \), and by construction
\[ L(\tilde{\gamma}_j^m) < L(\gamma_j^m) \]
which is a contradiction. The same argument holds for \([0,1] \) replaced by \([0,t] \). Hence
\[ L(\gamma_j^*) = \lim \inf L(\gamma_j^m) \]
But
\[ L(\gamma_j^m|_0^t) = tL(\gamma_j^*) \]
Letting \( m \to \infty \), we obtain \( L(\gamma_j^*|_0^t) = tL(\gamma_j^*) \) which implies, as remarked earlier that \( \gamma_j^* \in \Pi_j \).

The monotonocity properties are consequence of the \( L^\infty \) convergence. Then, by definition of the product topology
\[ \gamma^m \to \gamma^* \]
in \( \Pi \). Hence \( \gamma^* \in \Pi \) and
\[ J(\gamma^*) \leq \lim \inf J(\gamma^m) = c^* \]
As a consequence \( J(\gamma^*) = c^* \). Next we have to show that \( \gamma^* \) is a solution of (HS). We will start by showing that \( \gamma_j \) is a solution of the geodesic equation, for all \( j \).

**Remark:**

Recall the geodesic equation
\[ \ddot{\gamma}_i + \sum_i (...) \text{quadratic term in } \dot{\gamma}_i = 0 \]
If \( \gamma(t) \) is a solution of the equation, by the form of the equation \( \gamma(\alpha t) \) is also a solution for all non zero \( \alpha \). This is the reason that we can always parametrize proportional to arclength.

To proceed the curve, look at all curves joining the endpoints. we proved earlier that there exists a minimizer for this problem, which is \( C^2 \) and must lie in \( R_i \). We claim that \( \gamma_i^* \) is such a minimizer. It is obvious that it must be, otherwise we could replace \( \gamma^* \) by a curve with smaller length which contradicts our previous arguments. Then \( \gamma_i^* \) is \( C^2 \) and therefore a solution of the geodesic equation. It remains to prove that there is no problem in the boundary of the \( R_i \). Let's reparametrize \( \gamma^* \) by arclength, taking \( \gamma^*(0) \in z_0 \) (we do change of variables so that \([0,1] \to [0,L(\gamma^*)), [1,2] \to [L(\gamma^*), 2L(\gamma^*)]\) and so on). We will show that \( \gamma^* \) is \( C^2 \) through 0.
Use the same arguments as before, but we have to worry that $\gamma^*(0)$ intersects $z_0$ in the right way. If not then we can find points $k$ apart and use this to shorten the minimizer.

Our next claim is that $\gamma^*$ is a minimal geodesic.

Consider

$$b = L(z) = \inf_{M_i} L$$

and choose a point $p$ lying between $v^+$ and $v^-$ such that

$$\tilde{c} = \inf_{M_k, p \in \gamma} L(\gamma) > c + 2b$$  \tag{46}

When does this condition hold?

picture

Take $p_0 = p$ and consider the translates $p_i$ by $k$. We claim

$$p_0 \notin z_0 \cup z_1$$

Suppose this was not the case. Then $p_0 \in z_0$ implying that $p_1 \in z_1$. Then, by definition $d(p_0, p_1) = \tilde{c}$. Consider the curve $z_0|_{p_0}$ glued to $v^+|_{0}$ glued to $z_1|_{p_1}$. This also a curve joining $p_0$ to $p_1$ and therefore

$$\tilde{c} \leq 2\alpha + c$$  \tag{47}

On the other hand, consider the curve $z_0|_{p_0}$ glued to $v^-|_{0}$ glued to $z_1|_{p_1}$ which is such curve, which yields

$$\tilde{c} \leq 2\beta + c$$  \tag{48}

Summing (47) and (48), we obtain

$$\tilde{c} < b + c$$

contradicting assumption (46) In a similar way, we can prove the existence of functions $h^+, h^-$ homoclinic to $v^+$ and $v^-$ respectively.

Picture

To prove the existence of $h^-$, for example, we will minimize the renormalized functional $J$ over an appropriate class of functions. Define

$$\Pi_i^- = \{ \gamma_i \in AC([0, 1]) : \gamma_i(0) \in z_i, \gamma_i(1) \in z_{i+1} \}$$

where we consider $\gamma_i$ parametrized with respect to arclength. Then

$$\Pi^- = \{ (\gamma_i)_{i \in \mathbb{Z}} : (\Pi_1) - (\Pi_4) \text{ are verified} \}$$

for

$$\Pi_1 \gamma_i \in \Pi_i^- \text{ for all } i \in \mathbb{Z};$$

$$\Pi_2 \gamma_{i+1}(0) = \gamma_i(1);$$

117
\( \Pi_3 \) \( \gamma_i \) lies below \( P_i \) for all \( i \neq 0 \) and \( \gamma_0 \) lies above \( P_0 \);
\( \Pi_4 \) \( \gamma_i \) gets closer to \( v^- \) as \( i \to \pm \infty \).

**Picture**

Except from the region \( P_{-1} \) to \( P_1 \) the \( \gamma_i \)'s lie below \( u_i \) and in this region lies above \( u_{-1}, u_0 \).

We consider the renormalized functional
\[
J(\gamma) = \sum_{i \in \mathbb{Z}} (L(\gamma_i) - c)
\]
As in our previous example, we can obtain a lower bound for \( J \) in \( \Pi^- \), i.e., for all \( \gamma \in \Pi^- \)
\[
J(\gamma) \geq -2L(z)
\]
Since the curve obtained by gluing \( v^- \) up to \( z_0 \), to \( z_0 \) up to \( v^+ \), to \( v^+ \) up to \( z_1 \) to \( v^+ \) is an element of \( \Pi^- \), and the value of \( J \) at this curve is \( 2b \), we can consider
\[
\inf_{\gamma \in \Pi^-} J(\gamma) \leq 2b
\]
Similarly to Proposition —, we can prove the result which will give us the right asymptotic behavior of the curves in \( \Pi^- \).

**Proposition 18.** If \( \gamma \in \Pi^- \) is such that \( J(\gamma) < \infty \), then
\[
\|\gamma_i - v^-\|_{L^\infty([i,i+1])} \to 0 \text{ as } i \to \pm \infty.
\]
Then we can prove the result stating the existence of such curves, as being the minimizers of \( J \) over \( \Pi^- \).

**Theorem 23.** There exists \( h^-_{P_i} \) such that \( J(h^-) = \inf_{\Pi^-} J \), and by construction \( h^- \) is a homoclinic geodesic.

The proof of this theorem follows closely the proof of Theorem —. What can go wrong?

**Picture**

It may happen that the limit of a minimizing sequence hits \( P_0 \). We will prove that is this is the case we have a contradiction.

Suppose that this is the case. Since \( J(h^-) < \infty \), Proposition 18 gives the right asymptotic behavior. Consider \( \epsilon_n \) as the distance between the intersections of \( h^- \) and \( v^- \).

As remarked above \( \epsilon_n \to 0 \) as \( n \to \pm \infty \). Look at the length of the curve \( h^-|_{t_{-n}} \), where \( t_{\pm n} \) denote the time of intersection with \( z_{\pm n} \) respectively. we can find two points in the curve which are \( k \) apart. Cut out that piece, and shift by \( k \). Then
\[
L(h^-|_{t_{-n}}) \geq L(\text{new curve}) + c
\]
Repeat this procedure $2n$ times,

$$L(h^{-|_{t_{-n}}}) \geq L(\text{last curve}) + 2nc \geq \tilde{c} + 2nc$$

Then

$$J(h^{-|_{t_{-n}}}) \geq \tilde{c} - c$$

and therefore

$$J(h^{-}) \geq \tilde{c} - c + \text{small error}$$

By assumption ...

$$J(h^{-}) \geq 2b + \text{small error}$$

But we proved that

$$\inf_{\Pi^-} J \leq 2b$$

and therefore we got the contradiction, i.e., $h^{-}$ can’t run through $P_0$.

We can generalize this type of solutions. We will associate to each type, a symbolic sequence which gives us the homotopy type of the geodesic.

Let

$$\Sigma = \{\sigma = (\sigma_i)_{i\in \mathbb{Z}} : \sigma_i \in \{+, -\}\}$$

Corresponding to each $\sigma \in \Sigma$ there is a function $h^{\sigma_i} + ik$ (translates of $h^\pm$). From the strip, we will get the following information

- to $a+$ corresponds a piece of the curve lying between $h^+$ and $v^-$,
- to $a-$ corresponds a piece of the curve lying between $h^-$ and $v^-$.

Finally we define $D_\sigma$ as the remaining region (note that there are no $P - i$’s in this region).

**Theorem 24.** If $\tilde{c} > c + 2b$, then, for all $\sigma \in \Sigma$ there exists a geodesic of homotopy type $\sigma$ on $D_\sigma$.

**Proof:**

Let $\rho, r \in \{+, -\}$, and define the subclass of $\Sigma$

$$\Sigma^{\rho, r} = \{\sigma \in \Sigma : \sigma_i = \rho \forall i \leq -L \text{ and } \sigma_i = r \forall i \geq L\}$$

where $L$ is a large positive integer. Then to each $\sigma \in \Sigma^{\rho, r}$ it corresponds a geodesic in $D_\sigma$, which is asymptotic to $v_\rho$ as $t \to -\infty$ and is asymptotic to $v_r$ as $t \to \infty$. We will start to sketch the proof for this special case and mimics the previous proof. We define

$$\Pi^\sigma_i = \{\gamma_i \in AC([0, 1]) : \gamma_i(0) \in z_i, \gamma_i(1) \in z_{i+1} \text{ and } \gamma_i \text{ lies in } D_\sigma \cap R_i\}$$
and $\Pi^\sigma$ as the class of functions obtained by gluing together the $\gamma_i \in \Pi^\sigma_i$ so that the curves match at the end points, the monotone intersections (in the regions where $\rho$, $r$ start to repeat). We minimize

$$J(\gamma) = \sum_{i \in \mathbb{Z}} (L(\gamma_i) - c)$$

over $\Pi^\sigma$ and obtain the minimizers using the same arguments. Notice, however that the proof is a little simpler: there is no danger in running through $p_i$ since the minimizer can’t intersect the boundary of $D_\sigma$. It remains to prove the case $\sigma \in \Sigma \setminus \Sigma^{\rho, r}$. We can’t use the previous arguments since $J(\gamma)$ may not be bounded. We will use an approximation argument instead. For any $\sigma \in \Sigma \setminus \Sigma^{\rho, r}$ and for $j \in \mathbb{N}$ consider $\sigma^j = (\sigma^j_i)_{i \in \mathbb{Z}}$, defined by

$$\sigma^j_i = \begin{cases} 
\sigma_i & \text{if } |i| < j \\
\sigma_j & \text{if } i \geq j \\
\sigma_{-j} & \text{if } i \leq -j
\end{cases}$$

It is easy to check that $\sigma^j \in \Sigma^{\rho, r}$ and therefore, for each $j \in \mathbb{N}$, there is a minimal geodesic $(G_j)_j \in \mathbb{N}$ of homotopy type $\sigma^j$. We claim that we can find bounds for $G_j$ in $C^2_{\text{loc}}$. If so, by Ascoli-Arzela we can find a convergent subsequence in $C^1_{\text{loc}}$, which will be a solution of the geodesic equation. This will be a minimal geodesic of homotopy type $\sigma$. It remains to prove the claim:

- $G_j$ is bounded in $L^\infty_{\text{loc}}$ for all $j$, since by construction they all are in the region $D_\sigma$;
- suppose we can find bound for $\dot{G}_j$ in $L^\infty_{\text{loc}}$,
- By the geodesic equation

$$\ddot{G}_j = \text{stuff depending on } G_j + \text{quadratic on } \dot{G}_j$$

Then we will get $L^\infty_{\text{loc}}$ on $\dot{G}_j$, which proves that $(G_j)$ is bounded in $C^2_{\text{loc}}$. It remains to show that $\dot{G}_j$ is bounded in $L^\infty_{\text{loc}}$. Let $x^j_{-n}$ and $x^j_n$ the intersections of $G_j$ with $z_{-n}$ and $z_n$ respectively. Then

$$L(G_j|_{x^j_{-n}}^{x^j_n}) = \int \sum a_{ik}(G_j) \dot{G}_{ij} \dot{G}_{jk} \, dt$$

and since the matrix $(a_{ij})$ is positive definite

$$L(G_j|_{x^j_{-n}}^{x^j_n}) \geq \beta \int |\dot{G}_{ij}|^2 \, dt$$

To get an upper bound for $L(G_j|_{x^j_{-n}}^{x^j_n})$, compute the length of some curve with the same end points, which will be greater since $G_j$ is a geodesic. Then we have upper bounds for $\|\dot{G}_j\|_{L^2}$. The $L^\infty$ bounds on $G_j$ plus the $L^2$ bounds on $\dot{G}_j$ gives us (via the differential equation) $L^1$ bounds on $\dot{G}_j$. Considering $\tau_j$, $\theta_j$ defined by $G_j(\tau_j) = x^j_{-n}$ and $G_j(\theta_j) = x^j_n$,

$$\dot{G}_j(\theta_j) - \dot{G}_j(\tau_j) = \int_{\tau}^{\theta} |\dot{G}_j| \, dt$$

120
implying
\[ |\dot{G}_j(\theta_j)| \leq |\dot{G}_j(\tau_j)| + \int_{\tau}^{\theta} |\ddot{G}_j| \, dt \]

Integrate with respect to \( \tau_j \) and divide by the length interval
\[ |\dot{G}_j(\theta_j)| \leq \frac{1}{\theta_j - \tau_j} \|\dot{G}_j\|_{L^1} + \int_{\tau}^{\theta} |\ddot{G}_j| \, dt \]

and via the Schwarz inequality we obtain the desired bound, if the difference in the times is sufficiently large. To finish up, let’s obtain the lower bound on the time difference.
\[ |\gamma(t_n - n) - \gamma(t_n)| \geq (2n + 1) \min \text{dist(neighborhood( } z_i ) \]

since in this time interval it winds \((2n + 1)\) times around the torus. On the other hand
\[ \int |\dot{\gamma}| \, dt \leq |t_n - t_{-n}|^{1/2} (\int |\dot{\gamma}|^2 \, dt)^{1/2} \]

Putting these inequalities together
\[ \frac{1}{|t_n - t_{-n}|} \leq k \int |\dot{\gamma}|^2 \, dt \leq c(n) \]

We proved the existence of some sort of chaotic dynamics - highly sensitivity to initial conditions.

Finally we remark that in a similar form we can prove the existence of periodic solutions, by requiring that \( \sigma_i + l = \sigma_i \) for all \( i \) and some \( l \).