university of groningen

faculty of mathematics
and natural sciences

Calculus of variations and its applications

Bachelor Project Mathematics

November 2015

Student: M. H. Mudde

First supervisor: Dr. A. E. Sterk

Second supervisor: Prof. dr. A. J. van der Schaft

Abstract

In this thesis, the calculus of variations is studied. We look at how optimization problems are solved using the Euler-Lagrange equation. Functions that satisfy this equation and the prescribed boundary conditions, must also satisfy Legendre's condition and there must be no conjugate points in the interval in order to be the minimizer.
We generalize the Euler-Lagrange equation to higher dimensions and higher order derivatives to solve not only one-dimensional problems, but also multi-dimensional problems. At last we investigate the canonical form of the Euler-Lagrange equation.

Keywords: optimization, functional, Euler-Lagrange equation, canonical form, Hamiltonian.

Acknowledgements

I want to thank my first supervisor Dr. A. E. Sterk for being the best supervisor I could wish for. Even though I do not live in Groningen, he helped me in the best possible way. I could send him my drafts whenever I wanted and as many time I wanted, and every time he gave me adequate feedback.

I also want to thank my second supervisor Prof. dr. A. J. van der Schaft, for immediately willing to be my second supervisor.

Contents

Chapter 1

Introduction

This thesis is about the *calculus of variations.* The calculus of variations is a field of mathematics about solving optimization problems. This is done by minimizing and maximizing functionals. The methods of calculus of variations to solve optimization problems are very useful in mathematics, physics and engineering. Therefore, it is an important field in contemporary research. However, the calculus of variations has a very long history, which is interwoven with the history of mathematics.

In 1696, Johan Bernoulli came up with one of the most famous optimization problems: the brachistochrone problem. His brother Jakob Bernoulli and the Marquis de l'Hôpital immediately were interested in solving this problem, but the first major developments in the calculus of variations appeared in the work of Leonhard Euler. He started in 1733 with some important contributions in his Elementa Calculi Variationum. Then also Joseph-Louis Lagrange and Adrien-Marie Legendre came up with some important contributions.
These big names were not the only contributors to the calculus of variations. Isaac Newton, Gottfried Leibniz, Vincenzo Brunacci, Carl Friedrich Gauss, Siméon Poisson, Mikhail Ostrogradsky and Carl Jacobi also worked on the subject. Not forget to mention Karl Weierstrass: he was the first to place the subject on an unquestionable foundation.
In the 20th century, David Hilbert, Emmy Noether, Leonida Tonelli, Henri Lebesgue and Jacques Hadamard studied the subject.

The calculus of variations is thus a subject with a long history, a huge importance in classical and contemporary research and a subject where many big names in mathematics and physics have worked on. Therefore, it is a very interesting subject to study.

The goal of this thesis is to give an idea how optimization problems are solved using the method of calculus of variations. The basic tools and ideas that are needed will be explained. When the theory behind the calculus of variations is understood, some basic problems will be solved. This will give a good impression of the subject.

Since we need to optimize functionals, in chapter 2 we will first explain what functionals are. We do this by comparing it to functions.
In chapter 3 we will study the most important concept of this thesis, the Euler-Lagrange equation. Extremals need to satisfy this equation in order to be a candidate for the optimizer. We will derive this formula for the basic one-dimensional case. In chapter 3 we will also discuss the necessary and sufficient conditions under which the candidate is a minimum.
In chapter 4 we will illustrate how the found Euler-Langrange equation is used in solving some famous, basic minimization problems.
In chapter 5 we will generalize the Euler-Lagrange equation to higher dimensions and higher order derivatives. With these Euler-Lagrange equations, we will solve two multi-dimensional problems in chapter 6.
Finally, in chapter 7 we will study the canonical form of the Euler-Lagrange equation. This is used when an optimization problem is not easily solved using the Euler-Lagrange equation.

Chapter 2

Functions and functionals

Calculus of variations is a subject that deals with *functionals*. So in order to understand the method of calculus of variations, we first need to know what functionals are. In a very short way, a functional is a function of a function. To make it more clear what a functional is, we compare it to functions.

2.1 Functions

Consider the function $y = f(x)$. Here, x is the independent input variable. To each x belonging to the function domain, the function operator f associates a number y which belongs to the function range. We call $y = y(x)$ the dependent variable.
There is a difference between *single-valued* functions and *multivalued* functions. The first one means that to each input belongs exactly one output. The second case means that an input can produce multiple outputs.
This short review of functions enables us to move on to functionals.

2.2 Functionals

A functional is also an input-output machine, but it consists of more boxes. A simple functional consists of the following: we start again with an independent input variable x. The function f produces a dependent variable $y = y(x)$. This $y(x)$ is then the input for the functional operator J.
Different notations for the functional are used in the literature:
$J = J(x, y)$ and $J = F[y]$.
There exist also functionals that depend on function derivatives. We then have the following: we start again with the independent input variable x. The function f produces the primary dependent variable $y = y(x)$. We then take the derivative of the dependent variable so we obtain $y' = dy/dx$. We now have three inputs for the functional operator J. Hence the output is $J[y] = J(x, y, y')$.

In this example we have the first derivative as input, but functionals can also depend on higher order derivatives.
The differences between functions, functionals depending on a function and functionals depending on a function and its derivative are summarized in figure 2.1 [1].

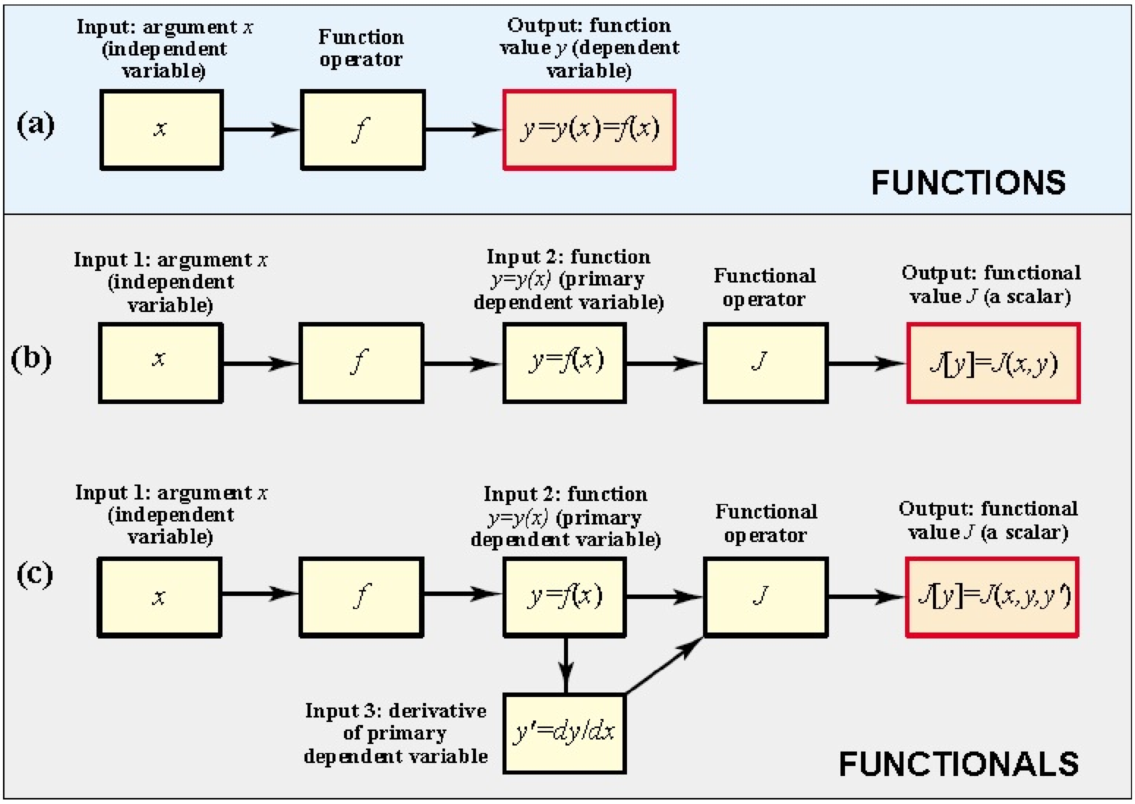

Figure 2.1: The input-output machines for the three different cases. (a) functions of the form $y = f(x)$, (b) functionals of the form $J[y] = J(x, y)$ and (c) functionals of the form $J[y] = J(x, y, y')$ [1].

2.3 Admissible functions

We now know the differences between functions and functionals, and we know that functionals need functions. Not all functions can be substituted into a functional. In this section we are going to look at the so called *admissible functions*: functions that can be substituted into a functional.
It turns out that admissible functions satisfy the following conditions [1]:

1. $\mathcal{C}^1$ continuity. This means that $y(x)$ needs to be continuously differentiable. In particular, $y'(x)$ is integrable.

2. Essential boundary conditions need to be satisfied. This means that the prescribed end values $y(a) = y_a$ and $y(b) = y_b$ need to be satisfied.
3. Single-valuedness (optional). To each x belongs exactly one y.

The set of all admissible functions is called the *admissible class*.

Figure 2.2 [1] shows admissible and inadmissible functions. For the admissible functions, all three conditions are satisfied.

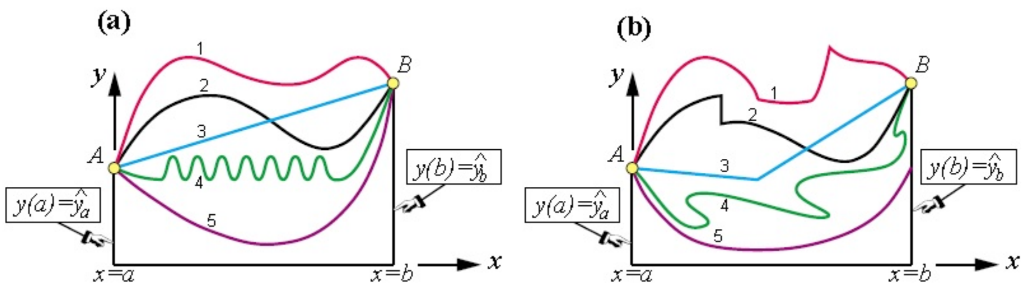

Figure 2.2: In (a), all functions 1 to 5 are admissible. In (b) all functions are inadmissible: 1 and 3 are not smooth, 2 is discontinuous, 4 is multivalued and 5 violates the essential boundary conditions [1].

Chapter 3

Euler-Lagrange Equation

The calculus of variations is thus all about solving optimization problems. The point is to find the function that optimizes the functional. To understand this method, we first look at the *basic one-dimensional functional*, also known as the *objective functional*:

$$J[y] = \int_a^b L(x, y, y')\,\mathrm{d}x \tag{3.1}$$

We use L instead of F because the integrand is called the *Lagrangian*, named after Joseph-Louis Lagrange, an Italian mathematician who was very important in the development of calculus of variations.

To obtain a unique minimizing function, we need to specify boundary conditions. Dirichlet boundary conditions are commonly used:

$$y(a) = \alpha, \qquad y(b) = \beta \tag{3.2}$$

3.1 The first variation

To determine the extrema of a function, we know from Calculus that an extremum is a point x at which $y(x)$ has a minimum or a maximum. At this point x, we know that $y'(x) = 0$ (i.e. the function's gradient vanishes).

We want to find out how this works for functionals. We want to find the functions that make the functional stationary. The functions that do this job are called *extremals*. To find these we need to look at the first variation (differential) of the functional. An extremal is either a maximum, minimum or an inflection. To figure this out, we need the second variation of the functional. Before looking at the second variation, we first discuss the first variation.

For the first variation test, we need to calculate the gradient of the functional. We follow the procedure of [2]. We first impose an inner product $\langle y; v\rangle$ on the underlying function space. The gradient $\nabla J[y]$ is then defined by the basic directional derivative formula:

$$\langle \nabla J[y]; v\rangle = \frac{d}{dt} J[y + tv]|_{t=0} \tag{3.3}$$

Where we take the standard L^2-inner product:

$$\langle f; g\rangle = \int_a^b f(x)g(x)\,\mathrm{d}x \tag{3.4}$$

In equation (3.3), v is called the variation in the function y. We sometimes write $v = \delta y$. We use the term *variational derivative* for the gradient. This is often written as δJ. The term variational is used a lot of times. This is where the name *calculus of variations* comes from.

What we have to do now is to compute the derivative of $J[y + tv]$ for each fixed y and v. Let $h(t) = J[y + tv]$:

$$h(t) = J[y + tv] = \int_a^b L(x, y + tv, y' + tv')\,\mathrm{d}x \tag{3.5}$$

Computing the derivative of $h(t)$ gives:

$$\begin{aligned} h'(t) &= \frac{d}{dt}\int_a^b L(x, y + tv, y' + tv')\,\mathrm{d}x \\ &= \int_a^b \frac{d}{dt} L(x, y + tv, y' + tv')\,\mathrm{d}x \\ &= \int_a^b \left[v\frac{\partial L}{\partial y}(x, y + tv, y' + tv') + v'\frac{\partial L}{\partial y'}(x, y + tv, y' + tv')\right]\mathrm{d}x \end{aligned}$$

Where we used the chain rule. Now evaluating the derivative at $t = 0$ gives:

$$\langle \nabla J[y]; v\rangle = \int_a^b \left[v\frac{\partial L}{\partial y}(x, y, y') + v'\frac{\partial L}{\partial y'}(x, y, y')\right]\mathrm{d}x \tag{3.6}$$

Equation (3.6) is known as the *first variation* of the functional $J[y]$. We call $\langle \nabla J[y]; v\rangle = 0$ the *weak form* of the variational principle.

What we want to do next is to obtain an explicit formula for $\nabla J[y]$.
The first step is to write equation (3.6) in a different way. Using the L^2-inner product (3.4) we write:

$$\langle \nabla J[y]; v\rangle = \int_a^b \nabla J[y] v\,\mathrm{d}x = \int_a^b h\, v\,\mathrm{d}x$$

Where we set $h(x) = \nabla J[y]$.

We now have the following:

$$\int_a^b h\, v \,\mathrm{d}x = \int_a^b \left[v \frac{\partial L}{\partial y}(x, y, y') + v' \frac{\partial L}{\partial y'}(x, y, y') \right] \mathrm{d}x$$

This is almost the same form, except for the v'. We fix this using integration by parts. Therefore, let:

$$r(x) \equiv \frac{\partial L}{\partial y'}(x, y, y')$$

The equation results in:

$$\int_a^b h\, v \,\mathrm{d}x = \int_a^b \left[v \frac{\partial L}{\partial y}(x, y, y') + v'\, r \right] \mathrm{d}x$$

Then for the second term in the equation we obtain:

$$\int_a^b r(x)\, v'(x) \,\mathrm{d}x = [r(b)v(b) - r(a)v(a)] - \int_a^b r'(x)\, v(x) \,\mathrm{d}x \tag{3.7}$$

So we have to compute $r'(x)$. This is done using the chain rule:

$$\begin{aligned} r'(x) &= \frac{d}{dx} \left(\frac{\partial L}{\partial y'}(x, y, y') \right) \\ &= \frac{\partial^2 L}{\partial x \partial y'}(x, y, y') + y' \frac{\partial^2 L}{\partial y \partial y'}(x, y, y') + y'' \frac{\partial^2 L}{\partial (y')^2}(x, y, y') \end{aligned} \tag{3.8}$$

So now we have found an expression for equation (3.7). The only thing we are left with are the boundary terms $r(b)v(b) - r(a)v(a)$. It follows that this is not problematic, if we look at the prescribed boundary conditions:

$$y(a) = \alpha, \qquad y(b) = \beta$$

The varied function $\hat{y}(x) = y(x) + tv(x)$ has to remain in the set of functions that satisfy the boundary conditions. So we obtain:

$$\hat{y}(a) = y(a) + tv(a) = \alpha, \qquad \hat{y}(b) = y(b) + tv(b) = \beta$$

Therefore, $v(x)$ must satisfy:

$$v(a) = 0, \qquad v(b) = 0 \tag{3.9}$$

Hence the boundary terms of equation (3.7) vanish.
Equation (3.7) becomes:

$$\int_a^b r(x)\, v'(x) \,\mathrm{d}x = - \int_a^b r'(x)\, v(x) \,\mathrm{d}x$$

Hence, equation (3.6) becomes:

$$\langle \nabla J[y]; v \rangle = \int_a^b \nabla J[y]\, v \, \mathrm{d}x = \int_a^b v \left[\frac{\partial L}{\partial y}(x, y, y') - \frac{d}{dx} \left(\frac{\partial L}{\partial y'}(x, y, y') \right) \right] \mathrm{d}x$$

This holds for all $v(x)$. We obtain our final explicit result:

$$\nabla J[y] = \frac{\partial L}{\partial y}(x, y, y') - \frac{d}{dx} \left(\frac{\partial L}{\partial y'}(x, y, y') \right) \tag{3.10}$$

We call this the variational derivative of the objective functional (3.1). So the gradient of a functional is a function.

To find the *critical functions* $y(x)$, we set $\nabla J[y] = 0$:

$$\nabla J[y] = \frac{\partial L}{\partial y}(x, y, y') - \frac{d}{dx} \frac{\partial L}{\partial y'}(x, y, y') = 0 \tag{3.11}$$

Which is a second order ordinary differential equation, called the *Euler-Lagrange equation*, named after Leonhard Euler and Joseph-Louis Lagrange, two of the most important contributors to the calculus of variations. Equation (3.11) can be rewritten as:

$$E = \frac{\partial L}{\partial y} - \frac{\partial^2 L}{\partial x \partial y'} - y' \frac{\partial^2 L}{\partial y \partial y'} - y'' \frac{\partial^2 L}{\partial (y')^2} = 0 \tag{3.12}$$

Any solution to the Euler-Lagrange equation that satisfies the boundary conditions, is a potential candidate for the minimizing function. This is all stated in the following theorem:

Theorem 1. *Suppose the Lagrangian function is at least twice continuously differentiable:* $L(x, y, y') \in C^2$. *Then any* C^2 *optimizer* $y(x)$ *to the corresponding functional* $J[y] = \int_a^b L(x, y, y')\, \mathrm{d}x$, *subject to the selected boundary conditions, must satisfy the Euler-Lagrange equation* (3.11). □

3.2 Legendre transform

We can rewrite the Euler-Lagrange equation in a different form, using the so called *Legendre transform* or *dual form* [1]. This may be useful if L is of a certain form, which will be explained in the next section.
To derive the Legendre transform, we first differentiate $L = L(x, y, y')$ with respect to x and use the chain rule:

$$\frac{dL}{dx} = \frac{\partial L}{\partial x} + \frac{\partial L}{\partial y} y' + \frac{\partial L}{\partial y'} y'' \tag{3.13}$$

We then expand a variation of the second term of the Euler-Lagrange equation (3.11):

$$\frac{d}{dx}\left(y'\,\frac{\partial L}{\partial y'}\right)=\frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right)y'+\frac{\partial L}{\partial y'}\,y'' \tag{3.14}$$

Subtracting (3.14) from (3.13) we obtain:

$$\frac{dL}{dx}-\frac{d}{dx}\left(y'\,\frac{\partial L}{\partial y'}\right)=\frac{\partial L}{\partial x}+\frac{\partial L}{\partial y}\,y'+\frac{\partial L}{\partial y'}\,y''-\frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right)y'-\frac{\partial L}{\partial y'}\,y''$$

Collecting terms we obtain:

$$\frac{d}{dx}\left(L-y'\,\frac{\partial L}{\partial y'}\right)-\frac{\partial L}{\partial x}=\left(\frac{\partial L}{\partial y}-\frac{d}{dx}\frac{\partial L}{\partial y'}\right)y'=E\,y'$$

Since E vanishes in (3.12), we can write:

$$\bar{E}=\frac{d}{dx}\left(L-y'\,\frac{\partial L}{\partial y'}\right)-\frac{\partial L}{\partial x}=\big(L-L_{y'}y'\big)'-L_x=0 \tag{3.15}$$

Where we call $\bar{E}$ the Legendre transform of the Euler-Lagrange equation, named after Adrien-Marie Legendre, a French mathematician.

3.3 Degenerate cases

Now we have two forms of the Euler-Lagrange equation, we can look at some degenerate cases of L. These cases will show why it is useful to have two forms [1].

1. L independent of x: Then $\partial L/\partial x=0$. Equation (3.15) becomes: $\big(L-L_{y'}y'\big)'=0$. Hence we have: $L-y'\frac{\partial L}{\partial y'}=C$, which is a first order ODE.

2. L independent of y: Then $\partial L/\partial y=0$. Equation (3.11) becomes: $\frac{d}{dx}(\frac{\partial L}{\partial y'})=0$. Hence we have: $\frac{\partial L}{\partial y'}=C$, which is a first order ODE.

3. L independent of x and y: Then $\partial L/\partial x=0$ and $\partial L/\partial y=0$. Equation (3.11) becomes: $y'=C_1$. Integrating again gives: $y=C_1x+C_2$.

4. L independent of y': Then $\partial L/\partial y'=0$. Equation (3.11) becomes: $\partial L/\partial y=0$. Hence we have: $L=L(x)$.

3.4 The second variation

As said in the beginning of this chapter, we need the second variation test to figure out whether a critical function is a minimizing function or not. In order to understand this, we first go back to the second derivative test for functions. We follow the explanation of [3].
For functions depending only on one variable, we have the following:

1. If $f''(x) < 0$, then f has a local maximum at x.
2. If $f''(x) > 0$, then f has a local minimum at x.
3. If $f''(x) = 0$, then the test is inconclusive.

For the multi-variable case, we need the Hessian matrix to determine if a critical point is a maximum, minimum or saddle point.
The Hessian matrix of a function f depending on two variables (x and y), is the following:

$$H(x, y) = \begin{pmatrix} f_{xx}(x, y) & f_{xy}(x, y) \\ f_{yx}(x, y) & f_{yy}(x, y) \end{pmatrix}$$

Let $D(x, y)$ be the determinant. Then:

$$D(x, y) = \det(H(x, y)) = f_{xx}(x, y)\, f_{yy}(x, y) - (f_{xy}(x, y))^2$$

Let (a, b) be a critical point of f (i.e. $f_x(a, b) = 0$ and $f_y(a, b) = 0$). Then the second partial derivative test states:

1. If $D(a, b) > 0$ and $f_{xx}(a, b) > 0$, then (a, b) is a local minimum of f.
2. If $D(a, b) > 0$ and $f_{xx}(a, b) < 0$, then (a, b) is a local maximum of f.
3. If $D(a, b) < 0$, then (a, b) is a saddle point of f.
4. If $D(a, b) = 0$, then the test is inconclusive.

For functions depending on more than two variables, we look at the eigenvalues of the Hessian matrix at the critical point. Let $(a, b, ...)$ be a critical point:

1. If the Hessian is positive definite (all eigenvalues at $(a, b, ...)$ are positive), then $(a, b, ...)$ is a local minimum of f.
2. If the Hessian is negative definite (all eigenvalues at $(a, b, ...)$ are negative), then $(a, b, ...)$ is a local maximum of f.
3. If the Hessian has both positive and negative eigenvalues at $(a, b, ...)$, then $(a, b, ...)$ is a saddle point of f.
4. For cases different than those three listed above, the test is inconclusive.

The justification for this is based on the second order Taylor expansion of the objective function at the critical point. For the second variation test for functionals, we need to expand the objective functional near the critical function. We follow the explanation of [2].
Using the objective functional with variation $v(x)$, we obtain:

$$h(t) = J[y + tv] = \int_a^b L(x, y + tv, y' + tv')\, \mathrm{d}x$$

The second derivative $h''(t)$ at $t = 0$ is:

$$h''(0) = \int_a^b \left[v^2 \frac{\partial^2 L}{\partial y^2}(x,y,y') + 2v\,v' \frac{\partial^2 L}{\partial y \partial y'}(x,y,y') + (v')^2 \frac{\partial^2 L}{\partial (y')^2}(x,y,y') \right] \mathrm{d}x$$

Where we write:

$$A(x) = \frac{\partial^2 L}{\partial y^2}(x,y,y'), \qquad B(x) = \frac{\partial^2 L}{\partial y \partial y'}(x,y,y'), \qquad C(x) = \frac{\partial^2 L}{\partial (y')^2}(x,y,y') \tag{3.16}$$

Hence we obtain:

$$h''(0) = \int_a^b \left[A\,v^2 + 2\,B\,v\,v' + C\,(v')^2 \right] \mathrm{d}x \tag{3.17}$$

The coefficients A, B, and C, are found by evaluating the second order derivatives of the Lagrangian.
Just like for functions, we need positive definiteness of the second variation in order to obtain a minimizer for the functional. Hence we want that $h''(0) > 0$ for all non-zero variations $v(x) \neq 0$.
So if the integrand is positive definite at each point, we obtain $A(x)\,v^2 + 2\,B(x)\,v\,v' + C(x)\,(v')^2 > 0$. Whenever $a < x < b$ and $v(x) \neq 0$, then $h''(0) > 0$ is also positive definite.

For the justification of this, we take the second order Taylor expansion of $h(t) = J[y + tv]$:

$$h(t) = J[y + tv] = J[y] + t\,K[y;v] + \frac{1}{2}t^2\,Q[y;v] + \dots$$

The Taylor expansion of h around 0 gives:

$$h(t) = h(0) + th'(0) + \frac{1}{2}t^2 h''(0) + \dots$$

Hence $h'(0) = K[y;v]$. From earlier calculations we know that $h'(0) = \langle \nabla J[y]; v \rangle$. So we obtain:

$$h'(0) = K[y;v] = \langle \nabla J[y]; v \rangle$$

If y is a critical function, the first order term vanish and hence we obtain:

$$K[y;v] = \langle \nabla J[y]; v \rangle = 0$$

For all functions $v(x)$.
Therefore, the second derivative terms $h''(0) = Q[y;v]$ determine whether the critical function $y(x)$ is a minimum, maximum or neither.
The results that we just found are stated in the following theorem:

Theorem 2. *A necessary and sufficient condition for the functional $J[y]$ to have a minimum for $y = y(x)$ is that the Euler-Lagrange equation vanishes and*

$$h''(0) = Q[y; v] > 0$$

For a maximum, we replace $>$ by $<$. □

In order to state another necessary condition, consider (3.17):

$$h''(0) = Q[y; v] = \int_a^b \left[A\, v^2 + 2\, B\, v\, v' + C\, (v')^2\right] \mathrm{d}x$$

Integrating by parts and using the boundary conditions (3.9) for $v(x)$, we find:

$$\int_a^b 2\frac{\partial^2 L}{\partial y \partial y'} v v' \,\mathrm{d}x = -\int_a^b \left(\frac{d}{dx}\frac{\partial^2 L}{\partial y \partial y'}\right) v^2 \,\mathrm{d}x$$

Therefore, we can write (3.17) as:

$$h''(0) = Q[y; v] = \int_a^b (P(v')^2 + Qv^2)\,\mathrm{d}x$$

Where:

$$P = P(x) = \frac{\partial^2 L}{\partial (y')^2}, \qquad Q = Q(x) = \frac{\partial^2 L}{\partial y^2} - \frac{d}{dx}\frac{\partial^2 L}{\partial y \partial y'} \tag{3.18}$$

We now state the following lemma:

Lemma 1. *A necessary condition for the quadratic functional*

$$h''(0) = Q[y; v] = \int_a^b (P(v')^2 + Qv^2)\,\mathrm{d}x \tag{3.19}$$

defined for all functions $v(x)$ such that $v(a) = v(b) = 0$, to be non-negative is that

$$P(x) \geq 0 \qquad (a \leq v \leq b)$$

Proof. We follow the proof of [4].
Suppose for a contradiction that $P(x) \geq 0$ does not hold, i.e. suppose that $P(x_0) = -2\beta, (\beta > 0)$ at some point x_0 in $[a, b]$. Since P is continuous, there exists an $\alpha > 0$ such that $(x_0 - \alpha, x_0 + \alpha) \subset [a, b]$, and:

$$P(x_0) < -\beta \qquad (x_0 - \alpha \leq x \leq x_0 + \alpha)$$

The next step is to construct a function $v(x)$ such that (3.19) is negative. Let:

$$v(x) = \begin{cases} \sin^2\left(\frac{\pi(x - x_0)}{\alpha}\right) & \text{for} \quad x_0 - \alpha \leq x \leq x_0 + \alpha \\ 0 & \text{otherwise} \end{cases} \tag{3.20}$$

Then:

$$\int_a^b (P(v')^2 + Qv^2)\,\mathrm{d}x = \int_{x_0-\alpha}^{x_0+\alpha} P\frac{\pi^2}{\alpha^2}\sin^2\left(\frac{2\pi(x-x_0)}{\alpha}\right)\mathrm{d}x + \int_{x_0-\alpha}^{x_0+\alpha} Q\sin^4\left(\frac{\pi(x-x_0)}{\alpha}\right)\mathrm{d}x < -\frac{2\beta\pi^2}{\alpha} + 2M\alpha \tag{3.21}$$

Where:

$$M = \max_{a\leq x\leq b} |Q(x)| \tag{3.22}$$

For sufficiently small α, the RHS of (3.21) becomes negative, and hence (3.19) is negative. Hence we have proven that if $h''(0) \geq 0$ then $P(x) \geq 0$. □

Using this lemma and the second variation, we can now state Legendre's theorem:

Theorem 3 (Legendre)**.** *A necessary condition for the functional*

$$J[y] = \int_a^b L(x, y, y')\,\mathrm{d}x, \qquad y(a) = \alpha, \qquad y(b) = \beta$$

to have a minimum for the curve $y = y(x)$ is that the inequality, which we call Legendre's condition

$$\frac{\partial^2 L}{\partial (y')^2} \geq 0$$

is satisfied at every point of the curve.

Proof. In order to have a minimum, we found that the second variation must be positive definite. Lemma 1 states that a necessary condition for the second variation to be positive definite, is that:

$$P(x) = \frac{\partial^2 L}{\partial (y')^2} \geq 0.$$

□

3.5 Necessary and sufficient conditions that determine the nature of an extremum

In the previous section we found necessary conditions for an optimizer to be a minimum or a maximum. In this section we want to find conditions which are both necessary and sufficient.

Consider the quadratic functional:

$$\int_a^b (P(v')^2 + Qv^2)\,\mathrm{d}x \tag{3.23}$$

Where:

$$P = \frac{\partial^2 L}{\partial (y')^2}, \qquad Q = \frac{\partial^2 L}{\partial y^2} - \frac{d}{dx}\frac{\partial^2 L}{\partial y \partial y'} \tag{3.24}$$

From now on, we do not use that (3.23) is a second variation, but we treat it as a new, independent problem.
In the previous section we found that the condition:

$$P(x) \geq 0 \qquad (a \leq x \leq b)$$

is a necessary condition for the functional (3.23) to be non-negative for all admissible $v(x)$ (lemma 1).
In this section we assume that the strict inequality holds:

$$P(x) > 0 \qquad (a \leq x \leq b)$$

Our goal is to find conditions which are both necessary and sufficient for the functional (3.23) to be positive definite for all admissible $v(x) \neq 0$.
The Euler-Lagrange equation of (3.23) is:

$$Qv - \frac{d}{dx}Pv' = 0 \tag{3.25}$$

This is a second order differential equation, which we call *Jacobi's equation*.
The function $v(x) = 0$ satisfies (3.25) and the boundary conditions:

$$v(a) = 0, \qquad v(c) = 0, \qquad (a < c \leq b)$$

It is possible that equation (3.25) has other, non-trivial solutions that satisfy the boundary conditions. Therefore, we introduce the following definition [4]:

Definition 1. *The point $\tilde{a}$ ($\neq a$) is said to be conjugate to the point a if the equation* (3.25) *has a solution which vanishes for $x = a$ and $x = \tilde{a}$ but is not identically zero.*

What we then have, if $v(x)$ is a solution of (3.25), $v(x) \neq 0$ and $v(a) = v(c) = 0$, then $Cv(x)$ is also a solution, where C is a constant and $C \neq 0$. Hence we can impose a normalization on $v(x)$. We shall assume that C must be chosen in such a way that $v'(a) = 1$ (if $v(x) \neq 0$ and $v(a) = 0$, then $v'(a)$ must be non-zero, because of the uniqueness theorem for (3.25)).

We then arrive at the following theorem:

Theorem 4. *If*

$$P(x) > 0 \qquad (a \leq x \leq b)$$

and if the interval $[a, b]$ *contains no points conjugate to* a*, then the functional*

$$\int_a^b (P(v')^2 + Qv^2)\,\mathrm{d}x \tag{3.26}$$

is positive definite for all $v(x)$ *such that* $v(a) = v(b) = 0$*.*

Proof. We follow the proof of [4].
To prove theorem 4, we want to reduce the functional (3.26) to:

$$\int_a^b P(x)\phi^2(x)\,\mathrm{d}x$$

Where ϕ^2 is an expression which cannot be identically zero, unless $v(x) = 0$. Then the functional (3.26) will be positive definite.
To do so, we add the quantity $(wv^2)'$ to the integrand of (3.26), where $w(x)$ is a differentiable function. We can do this, because the value of the integrand will not change, because $v(a) = v(b) = 0$ implies:

$$\int_a^b (wv^2)'\,\mathrm{d}x = 0$$

Our next step is to select a function $w(x)$ such that the expression:

$$P(v')^2 + Qv^2 + \frac{d}{dx}(wv^2) = P(v')^2 + 2wvv' + (Q + w')v^2 \tag{3.27}$$

is a perfect square. So $w(x)$ has to be chosen in such a way, that it is a solution of the equation (for details see [4]):

$$P(Q + w') = w^2 \tag{3.28}$$

This is called a *Ricatti equation.* If (3.28) holds, (3.27) can be rewritten as:

$$P\left(v' + \frac{w}{P}v\right)^2$$

Hence, if (3.28) has a solution which is defined on the whole interval $[a, b]$, then the functional (3.26) can be transformed into:

$$\int_a^b P\left(v' + \frac{w}{P}v\right)^2\,\mathrm{d}x \tag{3.29}$$

and hence is non-negative.
If (3.29) vanishes for some $v(x)$, then:

$$v' + \frac{w}{P}v = 0 \tag{3.30}$$

since $P(x) > 0$ for $a \leq x \leq b$. Therefore, $v(a) = 0$ implies $v(x) = 0$, because of the uniqueness theorem for (3.30). Hence (3.30) is positive definite.
Proving the theorem reduces to showing that the absence of conjugate points to a in $[a, b]$ guarantees that (3.28) has a solution defined on the whole interval $[a, b]$.
We can reduce (3.28) to a second order differential equation, using a change of variables. Let:

$$w = -\frac{u'}{u}P \tag{3.31}$$

Where u is a unknown function. We then obtain:

$$Qu - \frac{d}{dx}(Pu') = 0 \tag{3.32}$$

This is just the Euler-Lagrange equation (3.25) of the functional (3.26).
If there are no conjugate points to a in $[a, b]$, then (3.32) has a solution which does not vanish in $[a, b]$. Then there exists a solution of (3.28) given by (3.32), which is defined on the whole of $[a, b]$. □

We just showed that the absence of point conjugate to a in $[a, b]$ is sufficient for the functional (3.26) to be positive definite. Now we are going to show that this is also a necessary condition.

Lemma 2. *If the function $v = v(x)$ satisfies the equation*

$$Qv - \frac{d}{dx}(Pv') = 0$$

and the boundary conditions

$$v(a) = v(b) = 0 \tag{3.33}$$

then

$$\int_a^b (P(v')^2 + Qv^2)\,\mathrm{d}x = 0$$

Proof. Using integration by parts and (3.33), we immediately obtain:

$$0 = \int_a^b \left[-\frac{d}{dx}(Pv') + Qv\right] v\,\mathrm{d}x = \int_a^b (P(v')^2 + Qv^2)\,\mathrm{d}x$$

□

Theorem 5. *Assume that $P(x) > 0$ on $[a, b]$. If the functional*

$$\int_a^b (P(v')^2 + Qv^2)\,\mathrm{d}x \tag{3.34}$$

is positive definite for all $v(x)$ such that $v(a) = v(b) = 0$, then the interval $[a, b]$ contains no points conjugate to a.

Proof. We follow the proof of [4].
The first step is to construct a family of positive definite functionals, depending on a parameter t. For $t = 1$, we obtain the functional (3.34). For $t = 0$ we obtain:

$$\int_a^b (v')^2 \, \mathrm{d}x$$

For this functional, there are certainly no conjugate points to a in $[a, b]$.
We now want to prove that when the parameter t is varied continuously from 0 to 1, no conjugate points can appear in the interval $[a, b]$. Consider the following functional:

$$\int_a^b [t(P(v')^2 + Qv^2) + (1-t)(v')^2] \, \mathrm{d}x \tag{3.35}$$

This is positive definite for all t, where $0 \le t \le 1$, since (3.34) is positive definite by hypothesis. The corresponding Euler-Lagrange equation is:

$$\begin{aligned} 2tQv - \frac{d}{dx}\left(2tPv' + 2v' - 2tv'\right) = 0 \\ tQv - \frac{d}{dx}[tP + (1-t)]v' = 0 \end{aligned} \tag{3.36}$$

Let $v(x,t)$ be the solution of (3.36) such that $v(a,t) = 0, v_x(a,t) = 1, \forall t, 0 \le t \le 1$. For $t = 1$, the solution reduces to the solution $v(x)$ of (3.25) satisfying $v(a) = 0, v'(a) = 1$. For $t = 0$ the solution reduces to the solution of the equation $v''(x) = 0$ satisfying $v = x - a$.
Suppose that $[a, b]$ contains a point $\tilde{a}$ conjugate to a. This means, suppose that $v(x, 1)$ vanishes at some point $x = \tilde{a}$ in $[a, b]$. Then $\tilde{a} \ne b$, since otherwise:

$$\int_a^b (P(v')^2 + Qv^2) \, \mathrm{d}x = 0$$

for a function $v(x) \ne 0$ satisfying $v(a) = v(b) = 0$ (lemma 2), which would contradict our assumption that the functional (3.34) is positive definite. Therefore, the only thing we are left with, is showing that $[a, b]$ contains no interior point $\tilde{a}$ conjugate to a .
In order to prove this, consider the set of all points $(x, t), a \le x \le b$, satisfying $v(x,t) = 0$. If this set is non-empty, it represents a curve in the xt-plane. At each point where $v(x,t) = 0$, the derivative $v_x(x,t)$ is different from zero. According to the implicit function theorem, $v(x,t) = 0$ defines a continuous function $x = x(t)$ in the neighbourhood of each such point.
By hypothesis, $(\tilde{a}, 1)$ lies on this curve. We now have the following results, see figure 3.1 [4]. Starting from $(\tilde{a}, 1)$, the curve:

A. cannot terminate inside the rectangle $a \le x \le b, 0 \le t \le 1$. This would contradict the continuous dependence of $v(x,t)$ on t.

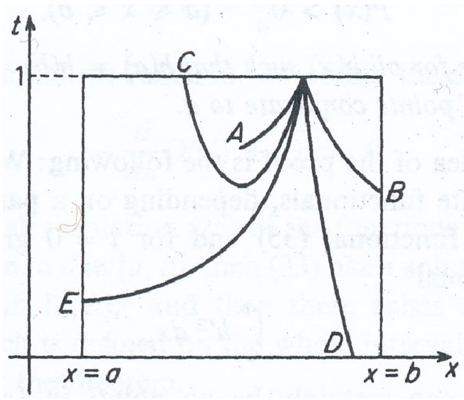

Figure 3.1: Curves in the xt-plane [4].

B. cannot intersect the line $x = b, 0 \leq t \leq 1$. This would contradict the assumption that the functional is positive definite for all t (same as in lemma 2).

C. cannot intersect the line $a \leq x \leq b, t = 1$. Then for some t we would have $v(x,t) = 0, v_x(x,t) = 0$ simultaneously.

D. cannot intersect the line $a \leq x \leq b, t = 0$. For $t = 0$, equation (3.36) reduces to $v''(x) = 0$, whose solution $v = x - a$ would only vanish for $x = a$.

E. cannot approach the line $x = a, 0 \leq t \leq 1$. Then for some t, we would have $v_x(a,t) = 0$, which contradicts our hypothesis.

Taking all these conditions together, we conclude that there is no curve that satisfies all these conditions. The proof is complete. □

Theorem 6. *Assume that $P(x) > 0$ on $[a,b]$. If the functional*

$$\int_a^b (P(v')^2 + Qv^2)\,\mathrm{d}x \tag{3.37}$$

is non-negative for all $v(x)$ such that $v(a) = v(b) = 0$, then the interval $[a,b]$ contains no interior points conjugate to a.

Proof. We follow the proof of [4].
If the functional (3.37) is non-negative, the functional (3.35) is positive definite for all t, except possibly at $t = 1$. Hence the proof of theorem 5 is still valid, except for the use of the lemma to prove that $\tilde{a} = b$ is impossible. Therefore, the possibility that $\tilde{a} = b$ is not excluded. □

We now arrive at the final theorem, where we combine theorem 4 and 5.

Theorem 7. *Assume that $P(x) > 0$ on $[a, b]$. The functional*

$$\int_a^b (P(v')^2 + Qv^2)\,\mathrm{d}x$$

is positive definite for all $v(x)$ such that $v(a) = v(b) = 0$ if and only if the interval $[a, b]$ contains no points conjugate to a. □

Summarizing our results so far, we have the following necessary conditions for an extremum:
If the functional $J[y] = \int_a^b L(x, y, y')\,\mathrm{d}x, y(a) = \alpha, y(b) = \beta$ has an extremum for $y = y(x)$, then:

1. The curve $y = y(x)$ satisfies the Euler-Lagrange equation (3.11).
2. The Legendre condition is satisfied ($\partial^2 L/\partial(y')^2 \geq 0$ for a minimum, $\partial^2 L/\partial(y')^2 \leq 0$ for a maximum).
3. The interval (a, b) contains no conjugate points to a.

We now focus on the sufficient conditions.

Theorem 8. *Suppose the functional $J[y] = \int_a^b L(x, y, y')\,\mathrm{d}x, y(a) = \alpha, y(b) = \beta$ satisfies the following conditions:*

1. *The curve $y = y(x)$ satisfies the Euler-Lagrange equation* (3.11).
2. *The strict Legendre condition is satisfied ($P(x) = \partial^2 L/\partial(y')^2 > 0$ for a minimum, $P(x) < 0$ for a maximum).*
3. *The interval $[a, b]$ contains no conjugate points to a.*

Then the functional $J[y]$ has an extremum for $y = y(x)$.

Proof. We follow the proof of [4].
Suppose the interval $[a, b]$ contains no conjugate points to a and suppose $P(x) > 0$. Then because of continuity of the solution of Jacobi's equation (3.25) and of $P(x)$, there exists a larger interval $[a, b+\epsilon]$ which still contains no conjugate points to a and such that $P(x) > 0$ in $[a, b+\epsilon]$. Consider:

$$\int_a^b (P(v')^2 + Qv^2)\,\mathrm{d}x - \alpha^2 \int_a^b (v')^2\,\mathrm{d}x \tag{3.38}$$

With Euler-Lagrange equation:

$$Qv - \frac{d}{dx}[(P - \alpha^2)v'] = 0 \tag{3.39}$$

We then have the following: Since $P(x)$ is positive in $[a, b+\epsilon]$ and thus has a positive greatest lower bound, and since the solution of (3.39) satisfying $v(a) = 0, v'(a) = 1$ depends continuously on α for all sufficiently small α, we have:

1. $P(x) - \alpha^2 > 0, a \leq x \leq b$

2. The solution of (3.39) satisfying $v(a) = 0, v'(a) = 1$ does not vanish for $a < x \leq b$. Using theorem 4, this implies that the functional (3.38) is positive definite for all sufficiently small α. This means there exists a $c > 0$ such that:

$$\int_a^b (P(v')^2 + Qv^2)\,\mathrm{d}x > c \int_a^b (v')^2\,\mathrm{d}x \tag{3.40}$$

As a consequence of (3.40), $J[y]$ has a minimum. If $y = y(x)$ is the extremal and $y = y(x) + h(x)$ is a sufficiently close neighbouring curve, then:

$$J[y+v] - J[y] = \int_a^b (P(v')^2 + Qv^2)\,\mathrm{d}x + \int_a^b (\xi v^2 + \eta(v')^2)\,\mathrm{d}x \tag{3.41}$$

(See [4] for this formula). In this equation, $\xi(x), \eta(x) \to 0$ uniformly for $a \leq x \leq b$ as $\|v\|_1 \to 0$. Using the Schwarz-inequality:

$$v^2(x) = \left(\int_a^x v'\,\mathrm{d}x\right)^2 \leq (x-a)\int_a^x (v')^2\,\mathrm{d}x \leq (x-a)\int_a^b (v')^2\,\mathrm{d}x$$

This means:

$$\int_a^b v^2\,\mathrm{d}x \leq \frac{(b-a)^2}{2}\int_a^b (v')^2\,\mathrm{d}x$$

This implies:

$$\left|\int_a^b (\xi v^2 + \eta(v')^2)\,\mathrm{d}x\right| \leq \epsilon\left(1 + \frac{(b-a)^2}{2}\right)\int_a^b (v')^2\,\mathrm{d}x \tag{3.42}$$

if $|\xi(x)| \leq \epsilon, |\eta(x)| \leq \epsilon$.
We can chose $\epsilon > 0$ arbitrarily small. Hence, using (3.40) and (3.42), we obtain:

$$J[y+v] - J[y] = \int_a^b (P(v')^2 + Qv^2)\,\mathrm{d}x + \int_a^b (\xi v^2 + \eta(v')^2)\,\mathrm{d}x > 0$$

for sufficiently small $\|v\|_1$.
We conclude that $y = y(x)$ is indeed a minimum of the functional $J[y] = \int_a^b L(x, y, y')\,\mathrm{d}x$ with boundary conditions $y(a) = \alpha, y(b) = \beta$.

□

Chapter 4

Applications of the Euler-Lagrange equation

In the previous chapter we found that the function $y = y(x)$ has to satisfy the Euler-Lagrange equation (3.11) in order to be a candidate for the minimizer. In this chapter we are going to look at the application of the Euler-Lagrange equation to some optimization problems.

4.1 Curves of shortest length

We start with the most elementary problem in the calculus of variations. In this problem, we want to find the curve of shortest length connecting two points $\mathbf{a} = (a, \alpha), \mathbf{b} = (b, \beta) \in \mathbb{R}^2$ in a plane. See figure 4.1 [2].

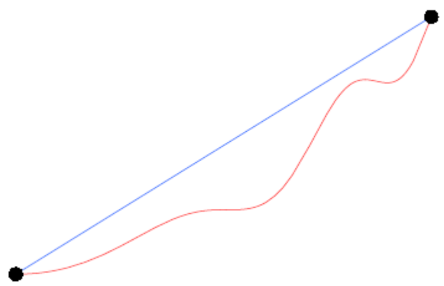

Figure 4.1: Paths between two points. [2]

We know from calculus that the formula for the arc length integral is [3]:

$$J[y] = \int_a^b \sqrt{1 + (y')^2}\,\mathrm{d}x$$

Hence the Lagrangian is:

$$L(x, y, y') = \sqrt{1 + (y')^2}$$

Our goal is to find the minimizer of the arc length integral that satisfies the boundary conditions:

$$y(a) = \alpha, \qquad y(b) = \beta$$

The minimizer has to satisfy the Euler-Lagrange equation (3.11), so we calculate the partial derivatives:

$$\frac{\partial L}{\partial y} = 0, \qquad \frac{\partial L}{\partial y'} = \frac{y'}{\sqrt{1+(y')^2}}$$

The Euler-Lagrange equation (3.11) becomes:

$$\begin{aligned} -\frac{d}{dx}\frac{y'}{\sqrt{1+(y')^2}} &= 0 \\ -\frac{y''}{(1+(y')^2)^{3/2}} &= 0 \end{aligned}$$

Hence we find the second order ODE:

$$y'' = 0 \tag{4.1}$$

The solution of (4.1) is:

$$y = c_1 x + c_2$$

This solution has to satisfy the prescribed boundary conditions.

$$y(a) = c_1 a + c_2 = \alpha, \qquad y(b) = c_1 b + c_2 = \beta$$

Subtracting $y(a)$ from $y(b)$ to find the constants c_1 and c_2:

$$\begin{aligned} c_1(b-a) &= \beta - \alpha \\ c_1 &= \frac{\beta-\alpha}{b-a} \\ c_2 &= \alpha - \frac{\beta-\alpha}{b-a}a \end{aligned}$$

Hence we find the minimizer:

$$\begin{aligned} y &= \frac{\beta-\alpha}{b-a}x + \alpha - \frac{\beta-\alpha}{b-a}a \\ &= \frac{\beta-\alpha}{b-a}(x-a) + \alpha \end{aligned} \tag{4.2}$$

We found that there is only one solution that satisfies the Euler-Lagrange equation and the boundary conditions. Since we know that we can minimize distance (there is always a shortest path between two points), our solution must be the minimizer. We conclude that the shortest path between two points is a straight line.

The next step is to show that the second variation test proves that our solution is indeed the minimizer.
First we calculate the second order partial derivatives:

$$\frac{\partial^2 L}{\partial y^2} = 0, \qquad \frac{\partial^2 L}{\partial y \partial y'} = 0, \qquad \frac{\partial^2 L}{\partial (y')^2} = \frac{1}{(1+(y')^2)^{3/2}}$$

Substituting (4.2) we find:

$$A(x) = 0, \qquad B(x) = 0, \qquad C(x) = \frac{(b-a)^3}{[(b-a)^2 + (\beta-\alpha)^2]^{3/2}} = k$$

Hence the second variation (3.17) becomes:

$$h''(0) = \int_a^b k(v')^2 \, \mathrm{d}x$$

Where $k > 0$ is a constant.
$h''(0) = 0$ if and only if $v(x)$ is a constant function. But we have that $v(x)$ needs to satisfy the boundary conditions $v(a) = v(b) = 0$. Hence $h''(0) > 0$ for all allowable non-zero variations. We conclude that the straight line is indeed a minimizer for the arc length functional.
We also see that the Legendre condition is satisfied:

$$\frac{\partial^2 L}{\partial (y')^2} = \frac{1}{(1+(y')^2)^{3/2}} > 0$$

and that there are no conjugate points.

In this problem we showed that the second variation test proves that our solution was indeed a minimizer, but in most of the problems we do not need the second variation, because if we know that a problem has a minimizer (we know that we can minimize length, time etc.) and we find only one solution that satisfies the Euler-Lagrange equation and the prescribed boundary conditions, then this solution must be the minimizer. The same reasoning holds for problems that are maximizable.

4.2 Minimal surface of revolution

In this section we study a simple version of the minimal surface problem, which is called the minimal surface of revolution problem. A surface of revolution is a surface created by rotating a curve around an axis. In this problem we take the x-axis. Our goal is to find the curve $y = y(x)$ joining two given points $\mathbf{a} = (a, \alpha), \mathbf{b} = (b, \beta) \in \mathbb{R}^2$ such that the surface of revolution created by rotating the curve around the x-axis has the least surface area.

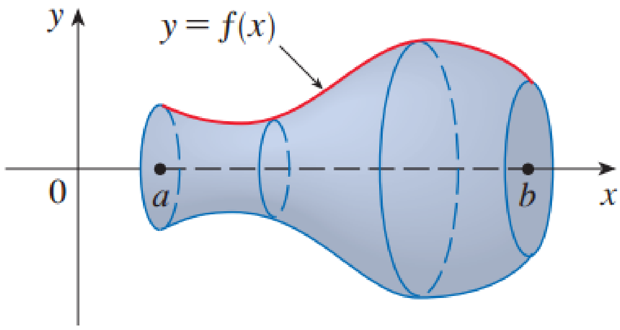

Figure 4.2: Surface of revolution obtained by rotating $y = f(x)$ around the x-axis. [3]

Each cross-section of this surface is a circle centered on the x-axis. See figure 4.2 [3].
The surface area of the surface obtained by rotating the curve $y = y(x), a \leq x \leq b$ about the x-axis is given by [3]:

$$J[y] = \int_a^b 2\pi\, |y|\, \sqrt{1 + (y')^2}\, \mathrm{d}x \tag{4.3}$$

Without loss of generality, we assume that $y = y(x) \geq 0\ \forall\, x$. Hence we can omit the absolute value bars. We also ommit the irrelevant factor 2π. Hence our goal is to minimize the following functional:

$$J[y] = \int_a^b y\, \sqrt{1 + (y')^2}\, \mathrm{d}x \tag{4.4}$$

The Lagrangian is:

$$L(x, y, y') = y\, \sqrt{1 + (y')^2}$$

Calculating the partial derivatives:

$$\frac{\partial L}{\partial y} = \sqrt{1 + (y')^2}, \qquad \frac{\partial L}{\partial y'} = \frac{y\, y'}{\sqrt{1 + (y')^2}}$$

The Euler-Lagrange equation (3.11) becomes:

$$\begin{aligned} \sqrt{1 + (y')^2} - \frac{d}{dx} \frac{y\, y'}{\sqrt{1 + (y')^2}} &= 0 \\ \frac{1 + (y')^2 - y\, y''}{(1 + (y')^2)^{3/2}} &= 0 \end{aligned} \tag{4.5}$$

The result is a non-linear second order ODE:

$$1 + (y')^2 - y\, y'' = 0 \tag{4.6}$$

This is difficult to solve. Therefore we use the following trick: we multiply (4.5) by y':

$$y' \left(\frac{1 + (y')^2 - y\, y''}{(1 + (y')^2)^{3/2}} \right) = \frac{d}{dx} \frac{y}{\sqrt{1 + (y')^2}} = 0$$

We obtain:

$$\frac{y}{\sqrt{1 + (y')^2}} = c \tag{4.7}$$

We see that the LHS is a constant. This means that the LHS is a *first integral* for the differential equation. Rewriting (4.7):

$$y' = \frac{\sqrt{y^2 - c^2}}{c}$$

This is a first order ODE. We use separation of variables:

$$\begin{aligned} \frac{dy}{dx} &= \frac{\sqrt{y^2 - c^2}}{c} \\ \int \frac{c}{\sqrt{y^2 - c^2}}\, \mathrm{d}y &= \int \mathrm{d}x \\ \int \frac{c}{\sqrt{y^2 - c^2}}\, \mathrm{d}y &= x + c_2 \end{aligned} \tag{4.8}$$

We use the trigonometric substitution $y = c \cosh t$:

$$\frac{dy}{dt} = c \sinh t \;\rightarrow\; \mathrm{d}y = c \sinh \mathrm{d}t$$

Substituting y and $\mathrm{d}y$ in (4.8):

$$\int \frac{c^2 \sinh t}{\sqrt{c^2(\cosh t - 1)}}\, \mathrm{d}t = x + c_2$$

Using the identity $\cosh^2 - \sinh^2 t = 1$:

$$\int c\, \mathrm{d}t = x + c_2$$

$$ct + c_3 = x + c_2$$

We take the constants c_2 and c_2 together in the constant c_2 and we substitute $t = \cosh^{-1} \frac{y}{c}$:

$$\begin{aligned} \cosh^{-1} \frac{y}{c} &= \frac{x + c_2}{c} \\ y &= c\, \cosh \left(\frac{x + c_2}{c} \right) \end{aligned} \tag{4.9}$$

Hence we found the solution of the Euler-Lagrange equation (4.5). The curve of (4.9) is known as a *catenary.* Catenaries are used a lot in engineering (bridges, roofs and arches).

Returning to the boundary conditions, we have three possibilities [2]:

1. There is only one value of the two integration constants c and c_2 such that (4.9) satisfies the boundary conditions. Then this catenary is the unique curve that minimizes the surface of revolution.

2. There are two different values of c and c_2 such that (4.9) satisfies the boundary conditions. Then one of these is the minimizer. The other one is a spurious solution: one that corresponds to a saddle point.

3. There are no values of c and c_2 such that (4.9) satisfies the boundary conditions. This happens when the points **a** and **b** are far apart. The film that spans the two circular wires breaks then apart into two circular disks. There is no surface of revolution that has a smaller surface area than the two disks. The function that minimizes this situation can be approximated by a sequence of functions that give progressively smaller values to the surface area functional (4.3), but the actual minimum is not attained among the class of smooth functions (for details, see [4]).

4.3 The brachistochrone problem

The brachistochrone problem is the most famous classical variational problem. The Greek word brachistochrone means minimal time. In this problem, we want to shape a wire between two points in such a way, that a bead slides from one end to the other in minimal time. This problem was posed by Johan Bernoulli, a Swiss mathematician, in 1696. At the time Bernoulli posed this famous problem, he was a professor of mathematics at the University of Groningen.

As starting point of the bead, we take the origin $\mathbf{a} = (0, 0)$. To avoid minus signs, we taken the y-axis downwards. Then the graph will be given by the function $y = y(x) \geq 0$. This graph will end in $\mathbf{b} = (b, \beta)$, where $b > 0$ and $\beta > 0$. See figure 4.3 [2].
We know that speed equals the time derivative of distance:

$$v = \frac{ds}{dt}$$

The travelled distance is equal to the arc length of the curve $y = y(x)$. We obtain using separation of variables:

$$v = \sqrt{1 + (y')^2}\,\frac{dx}{dt}$$

$$\int \mathrm{d}t = \int \frac{\sqrt{1 + (y')^2}}{v}\,\mathrm{d}x$$

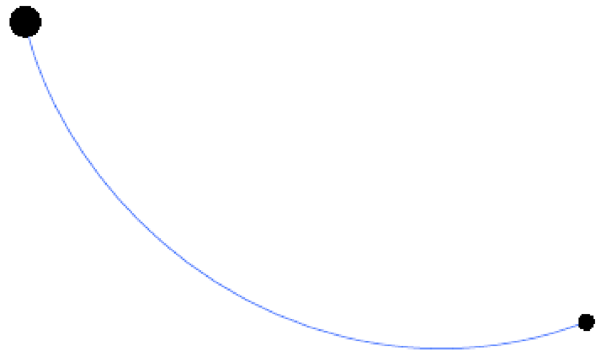

Figure 4.3: Brachistochrone problem: wire between two points. [2]

Then the total travel time is:

$$T[y] = \int_0^b \frac{\sqrt{1 + (y')^2}}{v} \, \mathrm{d}x \tag{4.10}$$

Where v is the speed of descent.

We use the conservation of energy law to determine the speed v. The kinetic energy of the bead is $\frac{1}{2}mv^2$, where m is the mass of the bead. The potential energy of the bead at height $y = y(x)$ is equal to $-mgy$, where g is the gravitational constant. At the initial height there is zero potential energy level. Initially, the bead is at rest. Hence there is no kinetic and potential energy. We assume that frictional forces are negligible. Then the conservation of energy implies tat the total energy is equal to 0. Hence:

$$\frac{1}{2}mv^2 - mgy = 0$$

Rewriting this, we find an expression for v:

$$v = \sqrt{2gy} \tag{4.11}$$

Substituting v in (4.10) we obtain:

$$T[y] = \int_0^b \sqrt{\frac{1 + (y')^2}{2gy}} \, \mathrm{d}x \tag{4.12}$$

Hence we need to minimize (4.12) to obtain the shape of $y = y(x)$. This function has to satisfy the boundary conditions:

$$y(0) = 0, \qquad y(b) = \beta \tag{4.13}$$

The Lagrangian of (4.12) is:

$$L(x, y, y') = \sqrt{\frac{1 + (y')^2}{y}}$$

We omit the irrelevant factor $\sqrt{2g}$. Computing the partial derivatives gives:

$$\frac{\partial L}{\partial y} = -\frac{\sqrt{1 + (y')^2}}{2y^{3/2}}, \qquad \frac{\partial L}{\partial y'} = \frac{y'}{\sqrt{y(1 + (y')^2)}}$$

Then the Euler-Lagrange equation (3.11) becomes:

$$\begin{aligned} -\frac{\sqrt{1 + (y')^2}}{2y^{3/2}} - \frac{d}{dx} \frac{y'}{\sqrt{y(1 + (y')^2)}} &= 0 \\ -\frac{2yy'' + (y')^2 + 1}{2\sqrt{y(1 + (y')^2)}} &= 0 \end{aligned} \tag{4.14}$$

Thus the minimizing functions need to solve the following non-linear second order ODE:

$$2yy'' + (y')^2 + 1 = 0 \tag{4.15}$$

Solving this equation is possible, but it is difficult. In chapter 7 we will study a new method to solve problems. At the end of that chapter we will actually solve the brachistochrone problem using this new technique.

Chapter 5

Multi-dimensional problems

Just as for one-dimensional optimization problems, we need the Euler-Lagrange equation to find candidates for the minimizer in multi-dimensional problems.
In physics, the second variation is not always needed, because the Euler-Lagrange boundary value problems suffice to single out the physically relevant solutions. In physical problems we know if a problem is minimizable/-maximizable. Hence if we find only one solution that satisfies the Euler-Lagrange equation and the boundary conditions, then this solution must be the minimizer/maximizer.
In order to find the Euler-Lagrange equations for the multi-dimensional variational problems, we are going to generalize the basic objective functional. First we look at a first order variational problem of a functional depending on two independent variables. After this, we use the same steps to generalize the Euler-Lagrange equation for several independent and dependent variables.
Then we are going to discuss the problem of a functional depending on the second order derivative. After this, we generalize the found Euler-Lagrange equation to higher order derivatives.
At the end of this chapter we provide an overview of the so found Euler-Lagrange equations.

5.1 Two independent variables

In this section we look at a functional with two independent variables x and y, dependent variable $u = u(x, y)$ and its partial derivatives $\partial u/\partial x = u_x$ and $\partial u/\partial y = u_y$. We follow again the procedure of [2].

Consider the following functional:

$$J[u] = \iint_\Omega L(x, y, u, u_x, u_y)\, \mathrm{d}x\, \mathrm{d}y \tag{5.1}$$

Which is a double integral over the domain $\Omega \subset \mathbb{R}^2$.
The optimization problem involves finding the function $u = f(x, y)$ that optimizes $J[u]$, and satisfies the prescribed boundary conditions. We take Dirichlet boundary conditions for simplicity.

$$u(x, y) = g(x, y) \qquad \text{for} \qquad (x, y) \in \partial\Omega \tag{5.2}$$

We follow the same procedure as for the one-dimensional case, so we first look at the first variation to find the optimizers:

$$h(t) \equiv J[u + tv] = \iint_\Omega L(x, y, u + tv, u_x + tv_x, u_y + tv_y)\,\mathrm{d}x\,\mathrm{d}y$$

The varied function $\hat{u}(x, y) = u(x, y) + tv(x, y)$ has to remain in the set of functions that satisfy the boundary conditions. So we obtain:

$$\hat{u}(x, y) = u(x, y) + tv(x, y) = g(x, y)$$

Therefore, $v(x, y)$ must satisfy:

$$v(x, y) = 0 \qquad \text{for} \qquad (x, y) \in \partial\Omega \tag{5.3}$$

Under the conditions of (5.3), the function $h(t)$ will have a minimum at $t = 0$. Then $h'(0) = 0$. Computing $h'(t)$ gives:

$$\begin{aligned}
h'(t) &= \frac{d}{dt}\iint_\Omega L(x, y, u + tv, u_x + tv_x, u_y + tv_y)\,\mathrm{d}x\,\mathrm{d}y \\
&= \iint_\Omega \frac{d}{dt} L(x, y, u + tv, u_x + tv_x, u_y + tv_y)\,\mathrm{d}x\,\mathrm{d}y \\
&= \iint_\Omega v\frac{\partial L}{\partial u}(x, y, u + tv, u_x + tv_x, u_y + tv_y) \\
&+ v_x\frac{\partial L}{\partial u_x}(x, y, u + tv, u_x + tv_x, u_y + tv_y) \\
&+ v_y\frac{\partial L}{\partial u_y}(x, y, u + tv, u_x + tv_x, u_y + tv_y)\,\mathrm{d}x\,\mathrm{d}y
\end{aligned}$$

Where we used the chain rule. Now evaluating the derivative at $t = 0$ gives:

$$h'(0) = \\ \iint_\Omega \left(v\frac{\partial L}{\partial u}(x, y, u, u_x, y_x) + v_x\frac{\partial L}{\partial u_x}(x, y, u, u_x, u_y) + v_y\frac{\partial L}{\partial u_y}(x, y, u, u_x, u_y)\right)\mathrm{d}x\,\mathrm{d}y \tag{5.4}$$

What we want to do next is to obtain an explicit formula for $\nabla J[u]$.
The first step is to write (5.4) in a different way, using an inner product:

$$h'(0) = \langle \nabla J[u]; v\rangle = \iint_\Omega h(x, y)\,v(x, y)\,\mathrm{d}x\mathrm{d}y$$

Where we set $h(x, y) = \nabla J[u]$.

We now have the following:

$$\iint_\Omega h\, v \,\mathrm{d}x\,\mathrm{d}y = \iint_\Omega \left(v\frac{\partial L}{\partial u} + v_x\frac{\partial L}{\partial u_x} + v_y\frac{\partial L}{\partial u_y} \right) \mathrm{d}x\,\mathrm{d}y$$

The terms that are problematic are v_x and v_y. Removing these terms turns out to be more difficult than for the one-dimensional case. Because we have a double integral, we need integration by parts that is based on Green's Theorem.

Green's Theorem states [3]:

$$\iint_\Omega \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \mathrm{d}A = \oint_{\partial\Omega} P\,\mathrm{d}x + Q\,\mathrm{d}y$$

For simplicity we write:

$$w_1 = \frac{\partial L}{\partial u_x} \qquad \text{and} \qquad w_2 = \frac{\partial L}{\partial u_y}$$

We obtain [2]:

$$\iint_\Omega \frac{\partial v}{\partial x} w_1 + \frac{\partial v}{\partial y} w_2 \,\mathrm{d}x\,\mathrm{d}y = \oint v(-w_2\,\mathrm{d}x + w_1\,\mathrm{d}y) - \iint_\Omega v\left(\frac{\partial w_1}{\partial x} + \frac{\partial w_2}{\partial y} \right) \mathrm{d}x\,\mathrm{d}y \tag{5.5}$$

We found in (5.3) that $v(x, y) = 0$, hence the boundary integral vanishes. Substituting $w_1 = \partial L/\partial u_x$ and $w_2 = \partial L/\partial u_y$ gives:

$$\iint_\Omega \left(v_x\frac{\partial L}{\partial u_x} + v_y\frac{\partial L}{\partial u_y} \right) \mathrm{d}x\,\mathrm{d}y = -\iint_\Omega v\left[\frac{\partial}{\partial x}\left(\frac{\partial L}{\partial u_x} \right) + \frac{\partial}{\partial y}\left(\frac{\partial L}{\partial u_y} \right) \right] \mathrm{d}x\,\mathrm{d}y$$

Substituting this result in (5.4):

$$h'(0) = \iint_\Omega v\left[\frac{\partial L}{\partial u} - \frac{\partial}{\partial x}\left(\frac{\partial L}{\partial u_x} \right) - \frac{\partial}{\partial y}\left(\frac{\partial L}{\partial u_y} \right) \right] \mathrm{d}x\,\mathrm{d}y = \langle \nabla J[u]; v \rangle \tag{5.6}$$

Hence we found the first variation of the functional:

$$\nabla J[u] = \frac{\partial L}{\partial u} - \frac{\partial}{\partial x}\left(\frac{\partial L}{\partial u_x} \right) - \frac{\partial}{\partial y}\left(\frac{\partial L}{\partial u_y} \right)$$

The gradient must vanish in order to find the critical functions. Hence the optimizer $u(x, y)$ must satisfy the following *Euler-Lagrange equation*:

$$E = \frac{\partial L}{\partial u} - \frac{\partial}{\partial x}\left(\frac{\partial L}{\partial u_x} \right) - \frac{\partial}{\partial y}\left(\frac{\partial L}{\partial u_y} \right) = 0 \tag{5.7}$$

and the prescribed boundary conditions.

5.2 Several independent variables

In this section we are going to look at functionals depending on several independent variables $x_1, x_2, \ldots, x_m$, with dependent variable $u(x_1, x_2, \ldots, x_m)$ and its derivatives $u_{x_1}, u_{x_2}, \ldots, u_{x_m}$. The derivation of the Euler-Lagrange equation for m independent variables is almost the same as for 2 independent variables. Therefore, we sometimes omit some steps.

Consider the following functional:

$$J[u] = \int \ldots \int_{\Omega} L(x_1, x_2, \ldots, x_m, u, u_{x_1}, u_{x_2}, \ldots, u_{x_m})\, \mathrm{d}x_1 \ldots \mathrm{d}x_m \tag{5.8}$$

Which is a multiple integral over the domain $\Omega \subset \mathbb{R}^m$.
In the rest of this section we will write $\int_\Omega$ instead of $\int \ldots \int_\Omega$ and $\mathrm{d}\Omega$ instead of $\mathrm{d}x_1 \ldots \mathrm{d}x_m$.

The optimization problem involves finding the function $u = f(x_1, \ldots, x_m)$ that optimizes $J[u]$, and satisfies the prescribed boundary conditions. We take Dirichlet boundary conditions for simplicity.

$$u(x_1, \ldots, x_m) = g(x_1, \ldots, x_m) \qquad \text{for} \qquad (x_1, \ldots, x_m) \in \partial\Omega \tag{5.9}$$

We first look at the first variation to find the candidates for the optimizer. Consider the function:

$$h(t) \equiv J[u + tv] =$$
$$\int_\Omega L(x_1, x_2, \ldots, x_m, u + tv, u_{x_1} + tv_{x_1}, u_{x_2} + tv_{x_2}, \ldots, u_{x_m} + tv_{x_m})\, \mathrm{d}\Omega$$

The varied function $\hat{u}(x_1, \ldots, x_m) = u(x_1, \ldots, x_m) + tv(x_1, \ldots, x_m)$ has to remain in the set of functions that satisfy the boundary conditions. So $v(x_1, \ldots, x_m)$ must satisfy:

$$v(x_1, \ldots, x_m) = 0 \qquad \text{for} \qquad (x_1, \ldots, x_m) \in \partial\Omega \tag{5.10}$$

Under the condition of (5.9), the function $h(t)$ will have a minimum at $t = 0$. Then $h'(0) = 0$. Computing $h'(t)$ at $t = 0$ gives:

$$\begin{aligned} h'(0) = \int_\Omega & v \frac{\partial L}{\partial u}(x_1, \ldots, x_m, u, u_{x_1}, \ldots, u_{x_m}) \\ & + \sum_{i=1}^{m} \left(v_{x_i} \frac{\partial L}{\partial u_{x_i}}(x_1, \ldots, x_m, u, u_{x_1}, \ldots, u_{x_m}) \right) \mathrm{d}\Omega \end{aligned} \tag{5.11}$$

To obtain an explicit formula for $\nabla J[u]$, we write (5.11) in a different way, using an inner product:

$$h'(0) = \langle \nabla J[u]; v \rangle = \int_\Omega h(x_1, \ldots, x_m)\, v(x_1, \ldots, x_m)\, \mathrm{d}\Omega$$

Where we set $h(x_1, \ldots, x_m) = \nabla J[u]$.

We now have the following:

$$\int_\Omega h\, v \, \mathrm{d}\Omega = \int_\Omega v \frac{\partial L}{\partial u}(x_1, \ldots, x_m, u, u_{x_1}, \ldots, u_{x_m}) + \sum_{i=1}^{m} \left(v_{x_i} \frac{\partial L}{\partial u_{x_i}}(x_1, \ldots, x_m, u, u_{x_1}, \ldots, u_{x_m}) \right) \mathrm{d}\Omega$$

The terms that are problematic are v_{x_i}. Removing these terms turns out to be even more difficult than in the previous section. Because we have a multiple integral, we need integration by parts that is based on the divergence theorem.
The Divergence Theorem states [3]:

$$\iint_S \mathbf{F} \bullet \mathrm{d}\mathbf{S} = \iiint_E \operatorname{div} \mathbf{F} \, \mathrm{d}V$$

For simplicity we write:

$$w_i = \frac{\partial L}{\partial u_{x_i}}$$

We obtain [1]:

$$\int_\Omega \frac{\partial}{\partial x_i}(v\, w_i) \, \mathrm{d}\Omega = \int_\Gamma w_i\, n_i\, v \, \mathrm{d}\Gamma \tag{5.12}$$

Where Γ is the boundary of the domain Ω and $\mathbf{n} = (n_1, n_2, \ldots, n_m)$ is the external normal vector.
Expanding the left hand side:

$$\int_\Omega \frac{\partial w_i}{\partial x_i} v \, \mathrm{d}\Omega + \int_\Omega w_i \frac{\partial v}{\partial x_i} \, \mathrm{d}\Omega = \int_\Gamma w_i\, n_i\, v \, \mathrm{d}\Gamma \tag{5.13}$$

We found in (5.10) that $v(x_1, \ldots, x_m) = 0$, hence the boundary integral vanishes. Substituting $w_i = \partial L / \partial u_{x_i}$ gives:

$$\int_\Omega \sum_{i=1}^{m} v_{x_i} \frac{\partial L}{\partial u_{x_i}} \, \mathrm{d}\Omega = - \int_\Omega v \sum_{i=1}^{m} \frac{\partial}{\partial x_i} \left(\frac{\partial L}{\partial u_{x_i}} \right) \mathrm{d}\Omega$$

Substituting this result in (5.11):

$$h'(0) = \int_\Omega v \left[\frac{\partial L}{\partial u} - \sum_{i=1}^{m} \frac{\partial}{\partial x_i} \left(\frac{\partial L}{\partial u_{x_i}} \right) \right] \mathrm{d}\Omega = \langle \nabla J[u]; v \rangle \tag{5.14}$$

Hence we found the first variation of the functional:

$$\nabla J[u] = \frac{\partial L}{\partial u} - \sum_{i=1}^{m} \frac{\partial}{\partial x_i} \left(\frac{\partial L}{\partial u_{x_i}} \right)$$

The gradient must vanish in order to find the critical functions. Hence the optimizer $u(x_1, \dots, x_m)$ must satisfy the following *Euler-Lagrange equation*:

$$E = \frac{\partial L}{\partial u} - \sum_{i=1}^{m} \frac{\partial}{\partial x_i}\left(\frac{\partial L}{\partial u_{x_i}}\right) = 0 \tag{5.15}$$

and the prescribed boundary conditions.

5.3 Several dependent variables but one independent variable

In this section we are going to study functionals depending on several dependent variables $y_1, y_2, ..., y_n$ which depend on the independent variable x: $y_1 = y_1(x), y_2 = y_2(x), ..., y_n = y_n(x)$, with the presence of their first derivatives $y_1', y_2', ..., y_n'$, where $y_i' = dy_i/dx$.

Consider the functional:

$$J[y_1, \dots, y_n] = \int_a^b L(x, y_1, y_2, ..., y_n, y_1', y_2', ..., y_n')\,\mathrm{d}x \tag{5.16}$$

It is convenient to use vector notation:

$$J[\mathbf{y}] = \int_a^b L(x, \mathbf{y}, \mathbf{y}'), \qquad \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \quad \mathbf{y}' = \begin{bmatrix} y_1' \\ \vdots \\ y_n' \end{bmatrix} \tag{5.17}$$

Where boldface letters indicate a vector.
The optimization problem involves finding the function $\mathbf{y} = \mathbf{y}(x)$ that optimizes $J[\mathbf{y}]$, and satisfies the prescribed boundary conditions:

$$\mathbf{y}(a) = \boldsymbol{\alpha}, \qquad \mathbf{y}(b) = \boldsymbol{\beta} \tag{5.18}$$

To find the optimizers, consider the function:

$$h(t) \equiv J[\mathbf{y} + t\mathbf{v}] = \int_a^b L(x, \mathbf{y} + t\mathbf{v}, \mathbf{y}' + t\mathbf{v})\,\mathrm{d}x$$

The varied function $\hat{\mathbf{y}}(x) = \mathbf{y}(x) + t\mathbf{v}(x)$ has to remain in the set of functions that satisfy the boundary conditions. So we obtain:

$$\hat{\mathbf{y}}(a) = \mathbf{y}(a) + t\mathbf{v}(a) = \boldsymbol{\alpha}, \qquad \hat{\mathbf{y}}(b) = \mathbf{y}(b) + t\mathbf{v}(b) = \boldsymbol{\beta}$$

Therefore, $\mathbf{v}(x)$ must satisfy:

$$\mathbf{v}(a) = 0, \qquad \mathbf{v}(b) = 0 \tag{5.19}$$

Under the conditions of (5.19), the function $h(t)$ will have a minimum at $t = 0$. Then $h'(0) = 0$. Computing $h'(t)$ gives:

$$\begin{aligned}
h'(t) &= \frac{d}{dt}\int_a^b L(x, \mathbf{y} + t\mathbf{v}, \mathbf{y}' + t\mathbf{v}')\,\mathrm{d}x \\
&= \int_a^b \frac{d}{dt} L(x, \mathbf{y} + t\mathbf{v}, \mathbf{y}' + t\mathbf{v}')\,\mathrm{d}x \\
&= \int_a^b v_1 \frac{\partial L}{\partial y_1}(x, \mathbf{y} + t\mathbf{v}, \mathbf{y}' + t\mathbf{v}') + \cdots + v_n \frac{\partial L}{\partial y_n}(x, \mathbf{y} + t\mathbf{v}, \mathbf{y}' + t\mathbf{v}') \\
&+ v_1' \frac{\partial L}{\partial y_1'}(x, \mathbf{y} + t\mathbf{v}, \mathbf{y}' + t\mathbf{v}') + \cdots + v_n' \frac{\partial L}{\partial y_n'}(x, \mathbf{y} + t\mathbf{v}, \mathbf{y}' + t\mathbf{v}')\,\mathrm{d}x \\
&= \int_a^b \sum_{i=1}^n \left(v_i \frac{\partial L}{\partial y_i}(x, \mathbf{y} + t\mathbf{v}, \mathbf{y}' + t\mathbf{v}') + v_i' \frac{\partial L}{\partial y_i'}(x, \mathbf{y} + t\mathbf{v}, \mathbf{y}' + t\mathbf{v}') \right)\,\mathrm{d}x
\end{aligned}$$

Where we used the chain rule. Now evaluating the derivative at $t = 0$ gives:

$$h'(0) = \int_a^b \sum_{i=1}^n \left(v_i \frac{\partial L}{\partial y_i}(x, \mathbf{y}, \mathbf{y}') + v_i' \frac{\partial L}{\partial y_i'}(x, \mathbf{y}, \mathbf{y}') \right)\,\mathrm{d}x \tag{5.20}$$

What we want to do next is to obtain an explicit formula for $\nabla J[y]$.
The first step is to write (5.20) in a different way, using an inner product:

$$h_i'(0) = \langle \nabla J[y_i]; v_i \rangle = \int_a^b h_i(x)\, v_i(x)\,\mathrm{d}x$$

Where we set $h_i(x) = \nabla J[y_i]$.

We now have the following:

$$\int_a^b h_i\, v_i\,\mathrm{d}x = \int_a^b v_i \frac{\partial L}{\partial y_i}(x, \mathbf{y}, \mathbf{y}') + v_i' \frac{\partial L}{\partial y_i'}(x, \mathbf{y}, \mathbf{y}')\,\mathrm{d}x$$

The v_i' term is problematic. We remove this using integration by parts.
Let:

$$r_i(x) \equiv \frac{\partial L}{\partial y_i'}(x, \mathbf{y}, \mathbf{y}')$$

The equation results in:

$$\int_a^b h_i\, v_i\,\mathrm{d}x = \int_a^b v_i \frac{\partial L}{\partial y_i}(x, \mathbf{y}, \mathbf{y}') + v_i'\, r_i\,\mathrm{d}x \tag{5.21}$$

Then for the second term in the equation we obtain:

$$\int_a^b r_i(x)\, v_i'(x)\,\mathrm{d}x = [r_i(b)v_i(b) - r_i(a)v_i(a)] - \int_a^b r_i'(x)\, v_i(x)\,\mathrm{d}x \tag{5.22}$$

Since $v_i(a) = 0$ and $v_i(b) = 0$ (5.19), the result is:

$$\int_a^b r_i(x)\, v_i'(x)\,\mathrm{d}x = -\int_a^b r_i'(x)\, v_i(x)\,\mathrm{d}x \tag{5.23}$$

Substituting equation (5.23) in equation (5.21):

$$\begin{aligned}\int_a^b h_i\, v_i\,\mathrm{d}x &= \int_a^b v_i \left[\frac{\partial L}{\partial y_i}(x, \mathbf{y}, \mathbf{y}') - r_i'\right]\,\mathrm{d}x \\ &= \int_a^b v_i \left[\frac{\partial L}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial L}{\partial y_i'}\right)\right]\,\mathrm{d}x\end{aligned} \tag{5.24}$$

So we have:

$$h_i'(0) = \int_a^b v_i \left[\frac{\partial L}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial L}{\partial y_i'}\right)\right]\,\mathrm{d}x = \langle \nabla J[y_i]; v_i \rangle \tag{5.25}$$

Hence we found the first variation of the functional:

$$\nabla J[y_i] = \frac{\partial L}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial L}{\partial y_i'}\right)$$

The gradient must vanish in order to find the critical functions. Hence the optimizer $\mathbf{y}(x)$ must satisfy the following system of *Euler-Lagrange equations*:

$$E_i = \frac{\partial L}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial L}{\partial y_i'}\right) = 0 \qquad i = 1, \ldots, n \tag{5.26}$$

and the prescribed boundary conditions.

5.4 Second order derivative

In this section we are going to look at functionals depending on one independent variable x, but with the presence of the first derivative $y' = dy/dx$ and the second derivative $y'' = d^2y/dx^2$.

Consider the functional:

$$J[y] = \int_a^b L(x, y, y', y'')\,\mathrm{d}x \tag{5.27}$$

The optimization problem involves finding the function $y = y(x)$ that optimizes $J[y]$, and satisfies the prescribed boundary conditions:

$$\begin{aligned} y(a) &= \alpha_0, \qquad y'(a) = \alpha_1 \\ y(b) &= \beta_0, \qquad y'(b) = \beta_1 \end{aligned} \tag{5.28}$$

To find the optimizers, consider:

$$h(t) \equiv J[y+tv] = \int_a^b L(x, y+tv, y'+tv'+y''+tv'')\,\mathrm{d}x$$

The varied function $\hat{y}(x) = y(x)+tv(x)$ has to remain in the set of functions that satisfy the boundary conditions. So we obtain:

$$\begin{aligned} \hat{y}(a) &= y(a)+tv(a) = \alpha_0, \qquad & \hat{y}'(a) &= y'(a)+tv'(a) = \alpha_1 \\ \hat{y}(b) &= y(b)+tv(b) = \beta_0, \qquad & \hat{y}'(b) &= y'(b)+tv'(b) = \beta_1 \end{aligned}$$

Therefore, $v(x)$ must satisfy:

$$v(a) = v'(a) = 0, \qquad v(b) = v'(b) = 0 \tag{5.29}$$

Under the conditions of (5.29), the function $h(t)$ will have a minimum at $t = 0$. Then $h'(0) = 0$. Computing $h'(t)$ gives:

$$\begin{aligned} h'(t) &= \frac{d}{dt}\int_a^b L(x, y+tv, y'+tv', y''+tv'')\,\mathrm{d}x \\ &= \int_a^b \frac{d}{dt} L(x, y+tv, y'+tv', y''+tv'')\,\mathrm{d}x \\ &= \int_a^b v\frac{\partial L}{\partial y}(x, y+tv, y'+tv', y''+tv'') \\ &+ v'\frac{\partial L}{\partial y'}(x, y+tv, y'+tv', y''+tv'') \\ &+ v''\frac{\partial L}{\partial y''}(x, y+tv, y'+tv', y''+tv'')\,\mathrm{d}x \end{aligned}$$

Where we used the chain rule. Now evaluating the derivative at $t = 0$ gives:

$$\begin{aligned} &h'(0) = \\ &\int_a^b \left(v\frac{\partial L}{\partial y}(x, y, y', y'') + v'\frac{\partial L}{\partial y'}(x, y, y', y'') + v''\frac{\partial L}{\partial y''}(x, y, y', y'') \right)\,\mathrm{d}x \end{aligned} \tag{5.30}$$

What we want to do next is to obtain an explicit formula for $\nabla J[y]$.
The first step is to write (5.30) in a different way, using an inner product:

$$h'(0) = \langle \nabla J[y]; v \rangle = \int_a^b h(x)\,v(x)\,\mathrm{d}x$$

Where we set $h(x) = \nabla J[y]$.

We now have the following:

$$\int_a^b h\,v\,\mathrm{d}x = \int_a^b \left(v\frac{\partial L}{\partial y} + v'\frac{\partial L}{\partial y'} + v''\frac{\partial L}{\partial y''} \right)\,\mathrm{d}x$$

The terms that are problematic are v' and v''. We remove these terms using integration by parts.
Let:

$$r(x) \equiv \frac{\partial L}{\partial y'}(x, y, y', y''), \qquad s(x) \equiv \frac{\partial L}{\partial y''}(x, y, y', y'')$$

The equation results in:

$$\int_a^b h\, v\, \mathrm{d}x = \int_a^b \left[v \frac{\partial L}{\partial y}(x, y, y', y'') + v'\, r + v''\, s \right] \mathrm{d}x \tag{5.31}$$

Then for the second term in the equation we obtain:

$$\int_a^b r(x)\, v'(x)\, \mathrm{d}x = [r(b)v(b) - r(a)v(a)] - \int_a^b r'(x)\, v(x)\, \mathrm{d}x \tag{5.32}$$

Since $v(a) = 0$ and $v(b) = 0$ (5.29), the result is:

$$\int_a^b r(x)\, v'(x)\, \mathrm{d}x = -\int_a^b r'(x)\, v(x)\, \mathrm{d}x \tag{5.33}$$

For the third term we obtain:

$$\begin{aligned} \int_a^b s(x)\, v''(x)\, \mathrm{d}x &= \big[s(b)v'(b) - s(a)v'(a)\big] - \int_a^b s'(x)\, v'(x)\, \mathrm{d}x \\ \int_a^b s'(x)\, v'(x)\, \mathrm{d}x &= s'(b)v(b) - s'(a)v(a) - \int_a^b s''(x)\, v(x)\, \mathrm{d}x \end{aligned} \tag{5.34}$$

Since $v(a) = v'(a) = 0$ and $v(b) = v'(b) = 0$ (5.29), the result is:

$$\int_a^b s(x)\, v''(x)\, \mathrm{d}x = \int_a^b s''(x)\, v(x)\, \mathrm{d}x \tag{5.35}$$

Substituting (5.33) and (5.34) in (5.31):

$$\begin{aligned} \int_a^b h\, v\, \mathrm{d}x &= \int_a^b v \left[\frac{\partial L}{\partial y}(x, y, y', y'') - r' + s'' \right] \mathrm{d}x \\ &= \int_a^b v \left[\frac{\partial L}{\partial y} - \frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right) + \frac{d^2}{dx^2}\left(\frac{\partial L}{\partial y''}\right) \right] \mathrm{d}x \end{aligned} \tag{5.36}$$

So we have:

$$h'(0) = \int_a^b v \left[\frac{\partial L}{\partial y} - \frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right) + \frac{d^2}{dx^2}\left(\frac{\partial L}{\partial y''}\right) \right] \mathrm{d}x = \langle \nabla J[y]; v \rangle \tag{5.37}$$

Hence we found the first variation of the functional:

$$\nabla J[y] = \frac{\partial L}{\partial y} - \frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right) + \frac{d^2}{dx^2}\left(\frac{\partial L}{\partial y''}\right)$$

The gradient must vanish in order to find the critical functions. Hence the optimizer $y(x)$ must satisfy the following *Euler-Lagrange equation*:

$$E = \frac{\partial L}{\partial y} - \frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right) + \frac{d^2}{dx^2}\left(\frac{\partial L}{\partial y''}\right) = 0 \tag{5.38}$$

and the prescribed boundary conditions.

5.5 Higher order derivatives

In the previous chapter we found the Euler-Lagrange equation for a functional depending only on the first derivative. In the previous section we found the Euler-Lagrange equation for a functional depending on the second derivative.

For higher order derivatives, the boundary terms that we obtain after each integration by parts vanish, since we have the following boundary conditions:

$$\begin{aligned} y(a) &= \alpha_0, & y'(a) &= \alpha_1, & \dots, && y^{n-1}(a) &= \alpha_{n-1} \\ y(b) &= \beta_0, & y'(b) &= \beta_1, & \dots, && y^{n-1}(b) &= \beta_{n-1} \end{aligned} \tag{5.39}$$

The varied function $\hat{y}(x) = y(x) + tv(x)$ has to remain in the set of functions that satisfy the boundary conditions. So we obtain:

$$\begin{aligned} \hat{y}^{(k)}(a) &= y^{(k)}(a) + tv^{(k)}(a) = \alpha_k \\ \hat{y}^{(k)}(b) &= y^{(k)}(b) + tv^{(k)}(b) = \beta_k \qquad\qquad k = 0, \dots, n-1 \end{aligned}$$

Therefore, $v(x)$ must satisfy:

$$\begin{aligned} v^{(k)}(a) &= 0 \\ v^{(k)}(b) &= 0 \qquad\qquad k = 0, \dots, n-1 \end{aligned} \tag{5.40}$$

It turns out that for a functional containing the third derivative, triple integration by parts adds an extra term to the Euler-Lagrange equation:

$$-\frac{d^3}{dx^3}\left(\frac{\partial L}{\partial y'''}\right)$$

In general, for the n-th order derivative, integration by parts is n-times used and hence:

$$(-1)^n \frac{d^n}{dx^n}\left(\frac{\partial L}{\partial y^{(n)}}\right)$$

is added.

We conclude that for a functional containing the n-th order derivative, $y(x)$ has to satisfy the following Euler-Lagrange equation and the prescribed boundary conditions (5.39) in order to be a candidate for the optimizer:

$$E = \frac{\partial L}{\partial y} - \frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right) + \frac{d^2}{dx^2}\left(\frac{\partial L}{\partial y''}\right) - \dots + (-1)^n \frac{d^n}{dx^n}\left(\frac{\partial L}{\partial y^{(n)}}\right) \tag{5.41}$$

5.6 Overview of the Euler-Lagrange equations

In this section we provide an overview of all of the Euler-Lagrange equations that we found in this chapter. This is done in table 5.1.

L depending on	Euler-Lagrange equation
(x, y, y') One independent variable One dependent variable	$$\frac{\partial L}{\partial y} - \frac{d}{dx}\frac{\partial L}{\partial y'}$$
(x, y, u, u_x, u_y) Two independent variables One dependent variable	$$\frac{\partial L}{\partial u} - \frac{\partial}{\partial x}\left(\frac{\partial L}{\partial u_x}\right) - \frac{\partial}{\partial y}\left(\frac{\partial L}{\partial u_y}\right)$$
$(x_1, \ldots, x_m, u, u_{x_1}, \ldots, u_{x_m})$ Several independent variables One dependent variable	$$\frac{\partial L}{\partial u} - \sum_{i=1}^{m} \frac{\partial}{\partial x_i}\left(\frac{\partial L}{\partial u_{x_i}}\right)$$
$(x, \mathbf{y}, \mathbf{y'})$ One independent variable Several dependent variables	$$\frac{\partial L}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial L}{\partial y_i'}\right)$$
(x, y, y', y'') Second order derivative One independent variable One dependent variable	$$\frac{\partial L}{\partial y} - \frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right) + \frac{d^2}{dx^2}\left(\frac{\partial L}{\partial y''}\right)$$
$(x, y, y', \ldots, y^{(n)})$ Higher order derivatives One independent variable One dependent variable	$$\frac{\partial L}{\partial y} - \frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right) + \frac{d^2}{dx^2}\left(\frac{\partial L}{\partial y''}\right) - \ldots + (-1)^n \frac{d^n}{dx^n}\left(\frac{\partial L}{\partial y^{(n)}}\right)$$

Table 5.1: Overview of the Euler-Lagrange equations for functionals depending on different variables and derivatives.

Chapter 6

Applications of the multi-dimensional Euler-Lagrange equations

In this section we are going to solve two multi-dimensional problems, using the Euler-Lagrange equations that we found in the previous chapter. The first problem is a problem involving two independent variables, the second problem involves the second order derivative.

6.1 Minimal surfaces

The minimal surface problem is about finding the surface of least total area, among all those whose boundary is the closed curve $C \subset \mathbb{R}^3$ [2]. Our goal is thus to minimize the surface area integral:

$$\operatorname{area} S = \iint_S \mathrm{d}S$$

over all possible surfaces $S \subset \mathbb{R}^3$ with boundary curve $\partial S = C$.
The minimal surface problem is also known as *Plateau's problem*, named after Joseph Plateau, a French physicist of the nineteenth century.
In the mid twentieth century, they finally found a solution to the simplest minimal surface problem. The minimal surface problem is still important in engineering, architecture and biology (foams, domes, cell membranes).
In section 4.2 we already solved a simple version: the minimal surface of revolution problem.

To solve the minimal surface problem, we assume that the bounding curve C projects down to a simple closed curve $\Gamma = \partial\Omega$ that bounds an open domain $\Omega \subset \mathbb{R}^2$. See figure 6.1 [2].

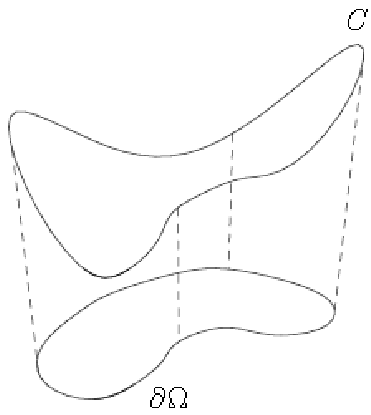

Figure 6.1: Projection of the bounding curve C to the simple closed curve $\Gamma = \partial\Omega$ for the minimal surface problem. [2]

The curve $C \subset \mathbb{R}^3$ is given by $z = g(x,y)$, for $(x,y) \in \Gamma = \partial\Omega$. The minimal surface S will be described as the graph of a function $z = u(x,y)$ parametrized by $(x,y) \in \Omega$.

We know that the surface area is then given by [3]:

$$J[u] = \iint_\Omega \sqrt{1 + \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2}\, \mathrm{d}x\, \mathrm{d}y \tag{6.1}$$

Our goal is thus to find the function $z = u(x,y)$ that minimizes (6.1) and satisfies the prescribed Dirichlet boundary conditions:

$$u(x,y) = g(x,y) \qquad \text{for} \qquad (x,y) \in \partial\Omega \tag{6.2}$$

The Lagrangian is given by:

$$L = \sqrt{1 + (u_x)^2 + (u_y)^2}$$

We see that we have two independent variables x, y and one dependent variable u. Using table 5.1, we find that the function $z = u(x,y)$ has to satisfy the following Euler-Lagrange equation:

$$\frac{\partial L}{\partial u} - \frac{\partial}{\partial x}\left(\frac{\partial L}{\partial u_x}\right) - \frac{\partial}{\partial y}\left(\frac{\partial L}{\partial u_y}\right) = 0 \tag{6.3}$$

Therefore, we calculate the partial derivatives:

$$\frac{\partial L}{\partial u} = 0, \qquad \frac{\partial L}{\partial u_x} = \frac{u_x}{\sqrt{1 + u_x^2 + u_y^2}}, \qquad \frac{\partial L}{\partial u_y} = \frac{u_y}{\sqrt{1 + u_x^2 + u_y^2}}$$

Then the Euler-Lagrange equation (6.3) becomes:

$$-\frac{\partial}{\partial x}\frac{u_x}{\sqrt{1+u_x^2+u_y^2}}-\frac{\partial}{\partial y}\frac{u_y}{\sqrt{1+u_x^2+u_y^2}}=0$$
$$\frac{-(1+u_y^2)u_{xx}+2u_xu_yu_{xy}-(1+u_x^2)u_{yy}}{(1+u_x^2+u_y^2)^{3/2}}=0$$

The result is a non-linear second order PDE:

$$(1+u_y^2)u_{xx}-2u_xu_yu_{xy}+(1+u_x^2)u_{yy}=0 \tag{6.4}$$

A function that satisfies the minimal surface equation (6.4) and satisfies the boundary conditions (6.2), minimizes the surface area.
The only problem is that this equation is difficult to solve. Solving this equation is beyond the scope of this thesis.
In 1930, Jesse Douglas and Tibor Radó both found a general solution to the problem. In [5] and [6] you can read their results.

6.2 Second order derivative problem

We do not have a concrete physical problem involving higher order derivatives, but consider the following functional [7]:

$$J[y]=\int_0^{\pi/2}(y'')^2-y^2+x^2\,\mathrm{d}x \tag{6.5}$$

With prescribed boundary conditions:

$$\begin{aligned} y(0)&=1, \qquad & y'(0)&=0,\\ y(\frac{\pi}{2})&=0, \qquad & y'(\frac{\pi}{2})&=-1 \end{aligned} \tag{6.6}$$

Our goal is to find the function $y=y(x)$ that extremizes the functional (6.5) and satisfies the boundary conditions (6.6).

The corresponding Lagrangian is:

$$L(x,y,y',y'')=(y'')^2-y^2+x^2$$

Since L depends on the independent variable x, dependent variable y and its second order derivative y'', we find using table 5.1 that the function $y=y(x)$ has to satisfy the following Euler-Lagrange equation:

$$E=\frac{\partial L}{\partial y}-\frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right)+\frac{d^2}{dx^2}\left(\frac{\partial L}{\partial y''}\right)=0 \tag{6.7}$$

We calculate the partial derivatives:

$$\frac{\partial L}{\partial y} = -2y, \qquad \frac{\partial L}{\partial y'} = 0, \qquad \frac{\partial L}{\partial y''} = 2y''$$

The Euler-Lagrange equation (6.7) becomes:

$$\begin{aligned} -2y + \frac{d^2}{dx^2} 2y'' &= 0 \\ -2y + 2y'''' &= 0 \\ y'''' - y &= 0 \end{aligned} \tag{6.8}$$

Which is a fourth order ODE.
In order to solve (6.8), let:

$$\begin{aligned} y &= e^{\lambda x}, \\ y'''' &= \lambda^4 e^{\lambda x} \end{aligned}$$

Substituting y and y'''' in (6.8) gives:

$$\begin{aligned} &\lambda^4 e^{\lambda x} - e^{\lambda x} = 0 \\ &\lambda^4 - 1 = 0 \\ &\lambda = 1 \quad \vee \quad \lambda = -1 \quad \vee \quad \lambda = i \quad \vee \quad \lambda = -i \\ y(x) &= c_1 e^x + c_2 e^{-x} + c_3 e^{ix} + c_4 e^{-ix} \\ &= c_1 e^x + c_2 e^{-x} + c_3 \cos(x) + c_4 \sin(x) \end{aligned}$$

Using the boundary conditions (6.6), we find:

$$c_1 = 0, \qquad c_2 = 0, \qquad c_3 = 1, \qquad c_4 = 0$$

Hence we find that the extremizing function is $y(x) = \cos(x)$. This is the only solution that satisfies the Euler-Lagrange equation (6.7) and satisfies the boundary conditions (6.6). Since this is not a clear problem about length, time or surface, we do not know if this problem has a minimizer or a maximizer. The only thing we can say right now is that this problem either has no minimizer, or the minimizer is $y(x) = \cos(x)$. If this problem has no minimizer, then our solution is the maximizer. Hence, we need to know the physical meaning of this problem to give a conclusion about the extremizer.

Chapter 7

The canonical form of the Euler-Lagrange equations and related topics

In the previous chapter and in chapter 4 we saw that the Euler-Lagrange equations are very useful in solving optimization problems, but in section 4.3 we found out that we could not easily solve the brachistochrone problem. In order to solve this problem, we need a new technique.
In this chapter we will show that the Euler-Lagrange equation can be rewritten in the so called *canonical form*. In the first section we will explain what the canonical form is. Then we will show that this form is useful in solving difficult optimization problems like the brachistochrone problem. We will use the results from this chapter to actually solve this problem.

7.1 The canonical form

Recall the functional with one independent variable, several dependent variables, and its first order derivatives (5.16):

$$J[y_1,\dots,y_n] = \int_a^b L(x, y_1, y_2, ..., y_n, y_1', y_2', ..., y_n')\,\mathrm{d}x$$

We found the corresponding Euler-Lagrange system of n second order differential equations (5.26):

$$E_i = \frac{\partial L}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial L}{\partial y_i'}\right) = 0 \qquad i = 1,\dots,n$$

We can rewrite this into a system of $2n$ first order differential equations. If we regard $y_1',\dots,y_n'$ as n new functions independent of $y_1,\dots,y_n$ we can write:

$$\frac{dy_i}{dx} = y_i', \qquad L_{y_i} - \frac{d}{dx}L_{y_i'} = 0 \tag{7.1}$$

We obtain a more convenient and symmetric system using canonical variables instead of the variables $x, y_1, \ldots, y_n, y_1', \ldots, y_n'$.
First, let:

$$p_i = L_{y_i'}, \qquad i = 1, \ldots, n \tag{7.2}$$

and suppose that the Jacobian [3]:

$$\frac{\partial(p_1, \ldots, p_n)}{\partial(y_1', \ldots, y_n')} = \det\|L_{y_i' y_k'}\|$$

is non-zero. Then we can solve the equations (7.2) for $y_1', \ldots, y_n'$ as functions of the variables $x, y_1, \ldots, y_n, p_1, \ldots, p_n$ [4].

Second, we express the function $L(x, y_1, \ldots, y_n, y_1', \ldots, y_n')$ in terms of a new function $H(x, y_1, \ldots, y_n, p_1, \ldots, p_n)$ related to L by the formula [4]:

$$H = -L + \sum_{i=1}^{n} y_i' L_{y_i'} \equiv -L + \sum_{i=1}^{n} y_i' p_i \tag{7.3}$$

Where H is called the *Hamiltonian function* corresponding to J.
We now have a transformation from $x, y_1, \ldots, y_n, y_1', \ldots, y_n', L$ to the canonical variables $x, y_1, \ldots, y_n, p_1, \ldots, p_n, H$.
Our goal is to transform the Euler-Lagrange equations in terms of these variables. Therefore, we need to express the partial derivatives of L with respect to y_i, evaluated for constant $x, y_1', \ldots, y_n'$ in terms of the partial derivatives of H with respect to y_i, evaluated for constant $x, p_1, \ldots, p_n$. To avoid lengthy calculations, we first write the expression for the differential of the function H. Then we use the fact that the first differential of a function does not depend on the choice of independent variables. Using this, we obtain the required formulas.

Using equation (7.3), we can write:

$$dH = -dL + \sum_{i=1}^{n} dy_i' p_i + \sum_{i=1}^{n} y_i' dp_i \tag{7.4}$$

so that:

$$dH = -\frac{\partial L}{\partial x} dx - \sum_{i=1}^{n} \frac{\partial L}{\partial y_i} dy_i - \sum_{i=1}^{n} \frac{\partial L}{\partial y_i'} dy_i' + \sum_{i=1}^{n} dy_i' p_i + \sum_{i=1}^{n} y_i' dp_i \tag{7.5}$$

Substituting (7.2): $p_i = L_{y_i'}$, the dy_i' terms cancel and we obtain:

$$dH = -\frac{\partial L}{\partial x} dx - \sum_{i=1}^{n} \frac{\partial L}{\partial y_i} dy_i + \sum_{i=1}^{n} y_i' dp_i \tag{7.6}$$

And this shows why it is useful to use canonical variables, because now we do not need to express dy_i' in terms of x, y_i and p_i.
The only thing we need to do is to write down the appropriate coefficients of the differentials in the RHS of (7.6), to obtain the partial derivatives of H:

$$\frac{\partial H}{\partial x} = -\frac{\partial L}{\partial x}, \qquad \frac{\partial H}{\partial y_i} = -\frac{\partial L}{\partial y_i}, \qquad \frac{\partial H}{\partial p_i} = y_i' \tag{7.7}$$

This shows that $\partial L/\partial y_i$ and y_i' are connected with the partial derivatives of H.
Then the Euler-Lagrange equations of (7.1) are:

$$\frac{dy_i}{dx} = \frac{\partial H}{\partial p_i}, \qquad \frac{dp_i}{dx} = -\frac{\partial H}{\partial y_i} \qquad i = 1, \dots, n \tag{7.8}$$

This system of $2n$ first order differential equations is called the *canonical system of Euler-Lagrange equations.*

7.2 First integrals of the Euler-Lagrange equations

We call a *first integral* of a system of differential equations a function which has a constant value along each integral curve of the system.
In this section we want to find the first integrals of the canonical system (7.8).
The most easy case is that of a functional L not depending on x explicitly: $L(y_1, \dots, y_n, y_1', \dots, y_n')$. Then:

$$H = -L + \sum_{i=1}^{n} y_i' p_i$$

does not depend on x explicitly. Using the chain rule:

$$\frac{dH}{dx} = \sum_{i=1}^{n} \left(\frac{\partial H}{\partial y_i} \frac{dy_i}{dx} + \frac{\partial H}{\partial p_i} \frac{dp_i}{dx} \right) \tag{7.9}$$

Using the canonical form of the Euler-Lagrange equations (7.8) we can write (7.9) as:

$$\frac{dH}{dx} = \sum_{i=1}^{n} \left(\frac{\partial H}{\partial y_i} \frac{\partial H}{\partial p_i} - \frac{\partial H}{\partial p_i} \frac{\partial H}{\partial y_i} \right) = 0 \tag{7.10}$$

along each extremal. Hence we can conclude that if L does not depend on x explicitly, then $H(y_1, \dots, y_n, p_1, \dots, p_n)$ is a first integral of the Euler-Lagrange equations.

Now we consider an arbitrary function, which may or may not, depend on x explicitly:

$$\Phi = \Phi(y_1, \dots, y_n, p_1, \dots, p_n)$$

Our goal is to figure out under which conditions Φ will be a first integral of the system (7.8).
Along each integral curve, we have:

$$\begin{aligned}\frac{d\Phi}{dx} &= \sum_{i=1}^{n}\left(\frac{\partial\Phi}{\partial y_i}\frac{dy_i}{dx} + \frac{\partial\Phi}{\partial p_i}\frac{dp_i}{dx}\right) \\ &= \sum_{i=1}^{n}\left(\frac{\partial\Phi}{\partial y_i}\frac{\partial H}{\partial p_i} - \frac{\partial\Phi}{\partial p_i}\frac{\partial H}{\partial y_i}\right) = [\Phi, H]\end{aligned} \tag{7.11}$$

Where we call $[\Phi, H]$ the *Poisson bracket* of Φ and H. Hence we found the formula:

$$\frac{d\Phi}{dx} = [\Phi, H] \tag{7.12}$$

So we can conclude that a necessary and sufficient condition for $\Phi = \Phi(y_1, \dots, y_n, p_1, \dots, p_n)$ to be a first integral of the Euler-Lagrange equations (7.8) is that the Poisson bracket $[\Phi, H]$ vanish.

7.3 Legendre transformation

There exists another way to reduce the Euler-Lagrange equations into canonical form, which will be called the *Legendre transformation.* This method consists of replacing the variational problem by an equivalent problem, in such a way that the Euler-Lagrange equations for the new problem are the same as for the canonical Euler-Lagrange equations.

In order to understand this method, we first go back to functions. We are going to look at extrema of functions of n variables.
Let $n = 1$ and suppose we want to find a minimum of the function $f(\xi)$, where $f(\xi)$ is strictly convex (i.e. $f''(\xi) > 0$).
Let:

$$p = f'(\xi) \tag{7.13}$$

Which is called the *tangential coordinate.* By assumption we have:

$$\frac{dp}{d\xi} = f''(\xi) > 0$$

Since $f(\xi)$ is strictly convex, we have that any point on the curve $\eta = f(\xi)$ is uniquely determined by the slope of its tangent.
Let:

$$H(p) = -f(\xi) + p\xi \tag{7.14}$$

The transformation from ξ and $f(\xi)$ to p and $H(p)$ is called the *Legendre transformation.* We obtain:

$$\frac{dH}{dp} = \xi \tag{7.15}$$

This gives:

$$\frac{d^2H}{dp^2} = \frac{d\xi}{dp} = \frac{1}{\frac{dp}{d\xi}} = \frac{1}{f''(\xi)} > 0 \tag{7.16}$$

Since $f''(\xi) > 0$. This shows that $H(p)$ is also strictly convex.
If we apply the Legendre transformation to p and $H(p)$, we obtain using (7.15):

$$-H(p) + pH'(p) = f(\xi) - pH'(p) + pH'(p) = f(\xi) \tag{7.17}$$

Hence we can conclude that the Legendre transform is an *involution*: a transformation which is its own inverse.

We are now going to prove the following theorem.

Theorem 9. *If:*

$$-H(p) + \xi p \tag{7.18}$$

is regarded as a function of two variables, then:

$$f(\xi) = \max_p [-H(p) + \xi p]. \tag{7.19}$$

Proof. Using (7.17), we know that (7.18) reduces to $f(\xi)$ when the following condition is satisfied:

$$\frac{\partial}{\partial p}[-H(p) + \xi p] = -H'(p) + \xi = 0$$
$$\xi = H'(p)$$

Thus $f(\xi)$ is an extremum of the function $-H(p) + \xi p$.
Now we show it is a maximum.

$$\frac{\partial^2}{\partial p^2}[-H(p) + \xi p] = -H''(p) < 0$$

Since H is strictly convex (7.16). □

Hence we have proven that the extremum of $f(\xi)$ is also an extremum of (7.18):

$$\min_\xi f(\xi) = \min_\xi \max_p [-H(p) + \xi p]$$

We can generalize this to functions of n independent variables. Therefore, let:

$$f(\xi_1, \ldots, \xi_n)$$

Such that:

$$\det \| f_{\xi_i} f_{\xi_k} \| \neq 0 \tag{7.20}$$

Let:

$$p_i = f_{\xi_i}, \qquad i = 1, \ldots, n \tag{7.21}$$

We can write H as:

$$H(p_1, \ldots, p_n) = -f + \sum_{i=1}^{n} \xi_i p_i$$

As before, we can write:

$$f(\xi_1, \ldots, \xi_n) = \underset{p_1,\ldots,p_n}{\text{ext}} \left[-H(p_1, \ldots, p_n) + \sum_{i=1}^{n} p_i \xi_i \right]$$

and

$$\underset{\xi_1,\ldots,\xi_n}{\text{ext}} f(\xi_1, \ldots, \xi_n) = \underset{\xi_1,\ldots,\xi_n,p_1,\ldots,p_n}{\text{ext}} \left[-H(p_1, \ldots, p_n) + \sum_{i=1}^{n} p_i \xi_i \right]$$

Hence the extremum of $f(\xi_1, \ldots, \xi_n)$ is also an extremum of:

$$-H(p_1, \ldots, p_n) + \sum_{i=1}^{n} p_i \xi_i$$

Our next goal is to apply our result to functionals. Consider the basic functional:

$$J[y] = \int_a^b L(x, y, y')\, \mathrm{d}x \tag{7.22}$$

Let:

$$p = L_{y'}(x, y, y') \tag{7.23}$$

and

$$H(x, y, p) = -L + py' \tag{7.24}$$

The new functional is:

$$J[y, p] = \int_a^b [-H(x, y, p) + py']\, \mathrm{d}x \tag{7.25}$$

Where y and p are two independent functions. (7.25) is the same functional as (7.22), if p is as in (7.23). The corresponding Euler-Lagrange equations are:

$$-\frac{\partial H}{\partial y} - \frac{dp}{dx} = 0, \qquad -\frac{\partial H}{\partial p} + \frac{dy}{dx} = 0 \tag{7.26}$$

which are just the canonical equations for (7.22).
We now need to prove that (7.22) and (7.25) have their extrema for the same curves, because then the following Euler-Lagrange equation:

$$-\frac{\partial L}{\partial y} - \frac{d}{dx}\frac{\partial L}{\partial y'} = 0 \tag{7.27}$$

and the canonical Euler-Lagrange equations (7.26) are equivalent.

It turns out that if we subject $H(x, y, p)$ to a Legendre transformation, we get back to $L(x, y, y')$:

$$dH = -\frac{\partial L}{\partial x}dx - \frac{\partial L}{\partial y}dy + y'dp$$

So:

$$\frac{\partial H}{\partial p} = y'$$

The Legendre transform gives:

$$-H + p\frac{\partial H}{\partial p} = L - py' + py' = L \tag{7.28}$$

The only thing we are left with is to show that $J[y]$ is an extremum of $J[y, p]$ when p is varied and y fixed:

$$J[y] = \operatorname*{ext}_{p} J[y, p] \tag{7.29}$$

Because then, when both p and y are varied, an extremum of $J[y, p]$ will be an extremum of $J[y]$.
$J[y, p]$ does not contain p'. Hence we have the situation as in case 4 of section 3.3. We have:

$$\frac{\partial}{\partial p}[-H + py'] = 0$$

This means:

$$y' = \frac{\partial H}{\partial p}$$

This implies (7.29), since:

$$-H + p\frac{\partial H}{\partial p} = L$$

Hence we have shown that the variational problems (7.22) and (7.25) are equivalent, and also that the Euler-Lagrange equations (7.26) and (7.27) are equivalent.
It turns out that this equivalence also holds for functionals depending on several functions.

Example Consider the following functional:

$$\int_a^b (Py'^2 + Qy^2)\,\mathrm{d}x \tag{7.30}$$

Where P and Q depend on x.

$$\begin{aligned}
L &= Py'^2 + Qy^2 \\
p &= L_{y'} = 2Py' \\
H &= -L + py' = -Py'^2 - Qy^2 + 2Py'^2 = Py'^2 - Qy^2 \\
y' &= \frac{p}{2P} \\
H &= \frac{p^2}{4P^2} - Qy^2
\end{aligned}$$

Hence the canonical Euler-Lagrange equations are (7.26):

$$\frac{dp}{dx} = 2Qy, \qquad \frac{dy}{dx} = \frac{p}{2P}$$

While the usual Euler-Lagrange equation (7.27) is:

$$2yQ - \frac{d}{dx}(2Py') = 0 \tag{7.31}$$

7.4 Canonical transformations

In this section, we ware going to look for transformations under which the Euler-Lagrange equations preserve their canonical form. In order to do so, we first prove the invariance under coordinate transformations of the Euler-Lagrange equation:

$$L_y - \frac{d}{dx}L_{y'} = 0.$$

To prove the invariance under coordinate transformations of the Euler-Lagrange equation, we are going to look at a change of variables. Suppose that instead of the rectangular plane coordinates x and y, we use curvilinear coordinates u and v, such that:

$$\begin{aligned} x &= x(u,v), \\ y &= y(u,v) \end{aligned} \qquad \begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix} \neq 0 \tag{7.32}$$

Then the curve $y = y(x)$ in the xy-plane corresponds to the curve $v = v(u)$ in the uv-plane. The functional:

$$J[y] = \int_a^b L(x, y, y')\,\mathrm{d}x$$

transforms into:

$$J_1[v] = \int_{a_1}^{b_1} L\left[x(u,v), y(u,v), \frac{y_u + y_v v'}{x_u + x_v v'}\right](x_u + x_v v')\,\mathrm{d}u$$

Where:

$$L_1(u,v,v') = L\left[x(u,v), y(u,v), \frac{y_u + y_v v'}{x_u + x_v v'}\right](x_u + x_v v')$$

Our goal is to show that if $y = y(x)$ satisfies the Euler-Lagrange equation:

$$\frac{\partial L}{\partial y} - \frac{d}{dx}\frac{\partial L}{\partial y'} = 0, \tag{7.33}$$

then $v = v(u)$ satisfies the Euler-Lagrange equation:

$$\frac{\partial L_1}{\partial v} - \frac{d}{du}\frac{\partial L_1}{\partial v'} = 0 \tag{7.34}$$

Therefore, let $\Delta\sigma$ denote the area bounded by the curves $y = y(x)$ and $y = y(x) + h(x)$. Let $\Delta\sigma_1$ denote the are bounded by the curves $v = v(u)$ and $v = v(u) + \eta(u)$.
By the standard formula for the transformation of areas [4], the limit as $\Delta\sigma, \Delta\sigma_1 \to 0$ of the ratio $\Delta\sigma/\Delta\sigma_1$ approaches the Jacobian:

$$\begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix}$$

Which is non-zero by hypothesis.
Hence if:

$$\lim_{\Delta\sigma\to 0} \frac{J[y+h] - J[y]}{\Delta\sigma} = 0,$$

then:

$$\lim_{\Delta\sigma_1\to 0} \frac{J_1[v+\eta] - J_1[v]}{\Delta\sigma_1} = 0.$$

Hence $v(u)$ satisfies (7.34) if $y(x)$ satisfies (7.33).
We conclude that the choice of coordinate system does not change the fact that a curve is an extremal or not.

We found that the Euler-Lagrange equation is invariant under coordinate transformations. This invariance property also holds for the canonical form. Because of the symmetry between y_i and p_i in the canonical form, even more general changes of variables are possible. We can transform the variables x, y_i, p_i into new variables:

$$\begin{aligned} Y_i &= Y_i(x, y_1 \ldots, y_n, p_1, \ldots, p_n) \\ P_i &= P_i(x, y_1 \ldots, y_n, p_1, \ldots, p_n) \end{aligned} \tag{7.35}$$

This means that p_i can transform independently of y_i.
It turns out that some transformations change the form of the canonical equations. Therefore it is important to study the conditions we need to

impose on the transformations (7.35) in order to continue them to be in canonical form when written in new variables.
We obtain the following new equations:

$$\frac{dY_i}{dx} = \frac{\partial H^*}{\partial P_i}, \qquad \frac{dP_i}{dx} = -\frac{\partial H^*}{\partial Y_i} \tag{7.36}$$

Where $H^* = H^*(x, Y_1, \ldots, Y_n, P_1, \ldots, P_n)$ for some new function.
Transformations of the from (7.35) that preserve the canonical form are called *canonical transformations.* To find these, remember that:

$$\frac{dy_i}{dx} = \frac{\partial H}{\partial p_i}, \qquad \frac{dp_i}{dx} = -\frac{\partial H}{\partial y_i} \tag{7.37}$$

are the canonical Euler-Lagrange equations for the functional:

$$J[y_1, \ldots, y_n, p_1, \ldots, p_n] = \int_a^b \left(\sum_{i=1}^{n} p_i y_i' - H \right) \mathrm{d}x \tag{7.38}$$

Where y_i and p_i are $2n$ independent functions. The new Y_i and P_i need to satisfy (7.36) for some function H^*. The functional that has the Euler-Lagrange equations of (7.36) is:

$$J^*[Y_1, \ldots, Y_n, P_1, \ldots, P_n] = \int_a^b \left(\sum_{i=1}^{n} P_i Y_i' - H^* \right) \mathrm{d}x \tag{7.39}$$

The functionals (7.38) and (7.39) involve the same variables y_i and p_i, but they represent two different problems. Transforming (7.37) into (7.38) using a coordinate transformation is thus the same as the requirement that the two variational problems are equivalent.
Two variational problems are equivalent if the integrands of the corresponding functions differ from each other by a total differential [4]. This means:

$$\sum_{i=1}^{n} p_i \,\mathrm{d}y_i - H \,\mathrm{d}x = \sum_{i=1}^{n} P_i \,\mathrm{d}Y_i - H^* \,\mathrm{d}x + \mathrm{d}\Phi(x, y_1, \ldots, y_n, p_1, \ldots, p_n) \tag{7.40}$$

for some function Φ.
We conclude the following: if a transformation (7.35) from x, y_i, p_i to x, Y_i, P_i is such that there exists a function Φ satisfying (7.40), then the transformation (7.35) is canonical.
The function Φ is called the *generating function.* If Φ is given, we can find the corresponding canonical transformation. Collecting terms in (7.40), we can write:

$$\mathrm{d}\Phi = \sum_{i=1}^{n} p_i \,\mathrm{d}y_i - \sum_{i=1}^{n} P_i \,\mathrm{d}Y_i + (H^* - H) \,\mathrm{d}x$$

Hence we find that:

$$p_i = \frac{\partial \Phi}{\partial y_i}, \qquad P_i = \frac{\partial \Phi}{\partial Y_i}, \qquad H^* = H + \frac{\partial \Phi}{\partial x} \tag{7.41}$$

So we have found the connection between the old variables and the new variables, and also an expression for H^*. If Φ does not depend on x explicitly, then $H^* = H$. Then to find H^*, we only need to replace y_i and p_i in H by their expressions in terms of Y_i and P_i.

We can also express the generating function in terms of y_i and P_i instead of y_i and Y_i. Therefore, we first rewrite (7.40):

$$\mathrm{d}\left(\Phi + \sum_{i=1}^{n} P_i\, Y_i\right) = \sum_{i=1}^{n} p_i\, \mathrm{d}y_i + \sum_{i=1}^{n} Y_i\, \mathrm{d}P_i + (H^* - H)\, \mathrm{d}x$$

We then obtain a new generating function:

$$\Psi(x, y_1, \ldots, y_n, P_1, \ldots, P_n) = \Phi + \sum_{i=1}^{n} P_i\, Y_i \tag{7.42}$$

We find the canonical transformation:

$$p_i = \frac{\partial \Psi}{\partial y_i}, \qquad Y_i = \frac{\partial \Psi}{\partial P_i}, \qquad H^* = H + \frac{\partial \Psi}{\partial x} \tag{7.43}$$

7.5 Noether's theorem

In section 6.2 we studied first integrals of the Euler-Lagrange equation. We found that the Euler-Lagrange equations for a functional where the Lagrangian does not depend on x explicitly: $L(y_1, \ldots, y_n, y_1', \ldots, y_n')$, has the first integral:

$$H = -L + \sum_{i=1}^{n} y_i' p_i$$

If L does not depend on x explicitly, we can replace x by a new variable:

$$x^* = x + \epsilon \tag{7.44}$$

and then L and the functional:

$$\int_a^b L(y_1, \ldots, y_n, y_1', \ldots, y_n')\, \mathrm{d}x \tag{7.45}$$

remains the same. Then H is a first integral of the Euler-Lagrange equations corresponding to (7.45) if and only if (7.45) is invariant under the transformation (7.44).

For the general case, there is also a connection between the existence of certain first integrals of the Euler-Lagrange equations and the invariance of the corresponding functional under certain transformations of the variables $x, y_1, \ldots, y_n$.
First we are going to explain what we mean by the invariance of a functional under transformations. Consider the following functional:

$$J[y_1, \ldots, y_n] = \int_{x_0}^{x_1} L(x, y_1, \ldots, y_n, '_1, \ldots, y_n')\,\mathrm{d}x$$

In short notation we can write this as:

$$J[\mathbf{y}] = \int_{x_0}^{x_1} L(x, \mathbf{y}, \mathbf{y}')\,\mathrm{d}x \tag{7.46}$$

Where $\mathbf{y}$ is the n-dimensional vector $(y_1, \ldots, y_n)$ and $\mathbf{y}'$ the n-dimensional vector $(y_1', \ldots, y_n')$.
Consider the following transformation:

$$\begin{aligned} x^* &= \Phi(x, y_1, \ldots, y_n, '_1, \ldots, y_n') = \Phi(x, \mathbf{y}, \mathbf{y}') \\ y_i^* &= \Psi_i(x, y_1, \ldots, y_n, '_1, \ldots, y_n') = \Psi_i(x, \mathbf{y}, \mathbf{y}'), \qquad i = 1, \ldots, n \end{aligned} \tag{7.47}$$

This transformation carries the curve γ with equation:

$$\mathbf{y} = \mathbf{y}(x) \qquad (x_0 \leq x \leq x_1)$$

into another curve γ^*. If we replace y and y' in (7.47) by $y(x)$ and $y'(x)$ and eliminate x from the resulting $n+1$ equations, we obtain:

$$\mathbf{y}^* = \mathbf{y}^*(x^*) \qquad (x_0^* \leq x^* \leq x_1^*)$$

for γ^*, where $\mathbf{y}^* = (y_1^*, \ldots, y_n^*)$.

Definition 2. *The functional* (7.46) *is said to be invariant under the transformation* (7.47) *if* $J[\gamma^*] = J[\gamma]$*, i.e. if:*

$$\int_{x_0}^{x_1} L\left(x^*, \boldsymbol{y}^*, \frac{\mathrm{d}\boldsymbol{y}^*}{\mathrm{d}x}\right)\mathrm{d}x^* = \int_{x_0}^{x_1} L\left(x, \boldsymbol{y}, \frac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}x}\right)\mathrm{d}x$$

We now provide two examples, one that is invariant under a certain transformation and the other is not invariant under the same transformation.

Example 1 Consider the following functional:

$$J[\mathbf{y}] = \int_{x_0}^{x_1} (\mathbf{y}')^2\,\mathrm{d}x$$

This functional is invariant under the transformation:

$$x^* = x + \epsilon, \qquad \mathbf{y}^* = \mathbf{y} \tag{7.48}$$

Given a curve γ with equation:

$$\mathbf{y} = \mathbf{y}(x) \qquad (x_0 \le x \le x_1),$$

the curve γ^* obtained by shifting γ a distance ϵ along the x-axis, has equation:

$$\mathbf{y}^* = \mathbf{y}(x^* - \epsilon) = \mathbf{y}^*(x^*) \qquad (x_0 + \epsilon \le x^* \le x_1 + \epsilon)$$

Then:

$$\begin{aligned} J[\gamma^*] &= \int_{x_0^*}^{x_1^*} \left[\frac{\mathrm{d}\mathbf{y}^*(x^*)}{\mathrm{d}x^*}\right]^2 \mathrm{d}x^* = \int_{x_0+\epsilon}^{x_1+\epsilon} \left[\frac{\mathrm{d}\mathbf{y}(x^*-\epsilon)}{\mathrm{d}x^*}\right]^2 \mathrm{d}x^* \\ &= \int_{x_0}^{x_1} \left[\frac{\mathrm{d}\mathbf{y}(x)}{\mathrm{d}x}\right]^2 \mathrm{d}x = J[\gamma] \end{aligned}$$

Example 2 Consider the integral:

$$J[\mathbf{y}] = \int_{x_0}^{x_1} x(\mathbf{y}')^2 \, \mathrm{d}x$$

This functional is not invariant under the transformation (7.48). Following the same procedure as in example 1, we obtain:

$$\begin{aligned} J[\gamma^*] &= \int_{x_0^*}^{x_1^*} x^* \left[\frac{\mathrm{d}\mathbf{y}^*(x^*)}{\mathrm{d}x^*}\right]^2 \mathrm{d}x^* = \int_{x_0+\epsilon}^{x_1+\epsilon} x^* \left[\frac{\mathrm{d}\mathbf{y}(x^*-\epsilon)}{\mathrm{d}x^*}\right]^2 \mathrm{d}x^* \\ &= \int_{x_0}^{x_1} (x+\epsilon) \left[\frac{\mathrm{d}\mathbf{y}(x)}{\mathrm{d}x}\right]^2 \mathrm{d}x = J[\gamma] + \epsilon \int_{x_0}^{x_1} \left[\frac{\mathrm{d}\mathbf{y}(x)}{\mathrm{d}x}\right]^2 \mathrm{d}x \ne J[\gamma] \end{aligned}$$

Suppose now that we have a family of transformations:

$$\begin{aligned} x^* &= \Phi(x, \mathbf{y}, \mathbf{y}'; \epsilon), \\ y_i^* &= \Psi(x, \mathbf{y}, \mathbf{y}'; \epsilon) \qquad i = 1, \ldots, n \end{aligned} \tag{7.49}$$

Where Φ and Ψ are differentiable with respect to ϵ and $\epsilon = 0$ correspond to the identity transformation:

$$\begin{aligned} \Phi(x, \mathbf{y}, \mathbf{y}'; 0) &= x, \\ \Psi(x, \mathbf{y}, \mathbf{y}'; 0) &= y_i \end{aligned} \tag{7.50}$$

Then we obtain the following theorem, which we call *Noether's Theorem*:

Theorem 10 (Noether)**.** *If the functional:*

$$J[\boldsymbol{y}] = \int_{x_0}^{x_1} L(x, \boldsymbol{y}, \boldsymbol{y}')\,\mathrm{d}x \tag{7.51}$$

is invariant under the family of transformations (7.50) *for arbitrary* x_0 *and* x_1*, then:*

$$\sum_{i=1}^{n} L_{y_i'}\psi_i + \left(L - \sum_{i=1}^{n} y_i L_{y_i'}\right)\phi = \text{constant} \tag{7.52}$$

along each extremal of $J[\boldsymbol{y}]$*, where:*

$$\begin{aligned} \phi(x, \boldsymbol{y}, \boldsymbol{y}') &= \left.\frac{\partial \Phi(x, \boldsymbol{y}, \boldsymbol{y}'; \epsilon)}{\partial \epsilon}\right|_{\epsilon=0} \\ \psi_i(x, \boldsymbol{y}, \boldsymbol{y}') &= \left.\frac{\partial \Psi_i(x, \boldsymbol{y}, \boldsymbol{y}'; \epsilon)}{\partial \epsilon}\right|_{\epsilon=0} \end{aligned} \tag{7.53}$$

i.e. every one-parameter family of transformations leaving $J[\boldsymbol{y}]$ *invariant leads to a first integral of its system of Euler-Lagrange equations.*

Proof. We follow the proof of [4].
Suppose ϵ is small. Then by Taylor's theorem:

$$\begin{aligned} x^* &= \Phi(x, \mathbf{y}, \mathbf{y}'; \epsilon) = \Phi(x, \mathbf{y}, \mathbf{y}'; 0) + \epsilon \left.\frac{\partial \Phi(x, \mathbf{y}, \mathbf{y}'; \epsilon)}{\partial \epsilon}\right|_{\epsilon=0} + o(\epsilon) \\ y_i^* &= \Psi_i(x, \mathbf{y}, \mathbf{y}'; \epsilon) = \Psi_i(x, \mathbf{y}, \mathbf{y}'; 0) + \epsilon \left.\frac{\partial \Psi_i(x, \mathbf{y}, \mathbf{y}'; \epsilon)}{\partial \epsilon}\right|_{\epsilon=0} + o(\epsilon) \end{aligned}$$

Using (7.50) and (7.53) we can write this as:

$$\begin{aligned} x^* &= x + \epsilon\phi(x, \mathbf{y}, \mathbf{y}') + o(\epsilon) \\ y_i^* &= y_i + \epsilon\psi_i(x, \mathbf{y}, \mathbf{y}') + o(\epsilon) \end{aligned} \tag{7.54}$$

In Gelfand and Fomin, [4] they state that the variation of J can be written as:

$$\delta J = \left[\sum_{i=1}^{n} L_{y_i'}\delta y_i + \left(L - \sum_{i=1}^{n} y_i' L_{y_i'}\right)\delta x\right]\Bigg|_{x=x_0}^{x=x_1} \tag{7.55}$$

Assuming that:

$$y_i = y_i(x)$$

and using (7.55), we can write an expression for the variation of $J[y]$ corresponding to the transformation (7.54).
Since:

$$\delta x = \epsilon\phi, \qquad \delta y_i = \epsilon\psi_i,$$

the result is:

$$\delta J = \epsilon \left[\sum_{i=1}^{n} L_{y'_i} \psi_i + \left(L - \sum_{i=1}^{n} y'_i L_{y'_i} \right) \phi \right] \Bigg|_{x=x_0}^{x=x_1}$$

By hypothesis we have that $J[\mathbf{y}]$ is invariant under (7.54). This means that δJ vanishes:

$$\begin{aligned} & \left[\sum_{i=1}^{n} L_{y'_i} \psi_i + \left(L - \sum_{i=1}^{n} y'_i L_{y'_i} \right) \phi \right] \Bigg|_{x=x_0} \\ = & \left[\sum_{i=1}^{n} L_{y'_i} \psi_i + \left(L - \sum_{i=1}^{n} y'_i L_{y'_i} \right) \phi \right] \Bigg|_{x=x_1} \end{aligned}$$

Since x_0 and x_1 are arbitrary, (7.52) holds along each extremal. □

We can write (7.52) in terms of the canonical variables p_1 and H:

$$\sum_{i=1}^{n} p_i \psi_i - H\phi = \text{constant} \tag{7.56}$$

Example 3 Consider the following functional:

$$J[\mathbf{y}] = \int_{x_0}^{x_1} L(\mathbf{y}, \mathbf{y}')\,\mathrm{d}x \tag{7.57}$$

Since this functional does not depend on x explicitly, it is invariant under the transformations:

$$x^* = x + \epsilon, \qquad y_i^* = y_i \tag{7.58}$$

Using (7.54), we find:

$$\phi = 1, \qquad \psi = 0$$

Then (7.56) reduces to:

$$H = \text{constant}$$

This means that the Hamiltonian H is constant along each extremal of $J[\mathbf{y}]$. We can conclude the following: For a functional of the form (7.57), which does not depend on x explicitly, the Hamiltonian H is a first integral of the system of Euler-Lagrange equations.

7.6 Back to the brachistochrone problem

We use the theory of this chapter to solve the brachistochrone problem.
In section 4.3 we found that the Lagrangian of the brachistochrone problem is:

$$L(x, y, y') = \sqrt{\frac{1 + (y')^2}{y}} \tag{7.59}$$

This Lagrangian does not depend on x explicitly. Hence, as stated before, then the Hamiltonian H is a first integral for the Euler-Lagrange equation. In the beginning of this chapter we found that the Hamiltonian H is defined as:

$$H = -L + y'\frac{\partial L}{\partial y'} \tag{7.60}$$

The Hamiltonian for (7.59) is thus:

$$H(x, y, y') = \frac{1}{\sqrt{y(1 + (y')^2)}}$$

This defines a first integral. We have $H(x, y, y') = \text{constant}$:

$$\frac{1}{\sqrt{y(1 + (y')^2)}} = k$$

This can be rewritten in terms of a new constant:

$$y(1 + (y')^2) = c \tag{7.61}$$

Where $c = 1/k^2$.
Solving for y' we find:

$$\begin{aligned} y' &= \sqrt{\frac{c-y}{y}} \\ \frac{dy}{dx} &= \sqrt{\frac{c-y}{y}} \end{aligned}$$

Separation of variables gives:

$$\begin{aligned} \int \sqrt{\frac{y}{c-y}}\,\mathrm{d}y &= \int \mathrm{d}x \\ \int \sqrt{\frac{y}{c-y}}\,\mathrm{d}y &= x + k \end{aligned}$$

To solve the LHS, we use the trigonometric substitution:

$$y = \frac{1}{2}c(1 - \cos\theta)$$

Substituting y and $\mathrm{d}y = \frac{1}{2}\,c\,\sin\theta\,\mathrm{d}\theta$:

$$x + k = \frac{1}{2}c\int \sqrt{\frac{1-\cos\theta}{1+\cos\theta}}\,\sin\theta\,\mathrm{d}\theta$$

Using $\sin^2\theta = 1 - \cos^2\theta$, we find:

$$\begin{aligned} x + k &= \frac{1}{2}c\int (1 - \cos\theta)\,\mathrm{d}\theta \\ &= \frac{1}{2}c(\theta - \sin\theta) \end{aligned}$$

Hence the general solution of the corresponding Euler-Lagrange equation consists of a family of cycloids:

$$x = r(\theta - \sin\theta) + d, \qquad y = r(1 - \cos\theta)$$

Where we use $r = \frac{1}{2}c$ and $d = -k$.
The curve must pass through the origin, so $d = 0$. We only need to determine r. In this thesis, we do not treat the theory to solve this. For the interested reader, we refer to [4]. In this book they treat the concept of variable end points. They state that the parameter r is uniquely prescribed by the right hand boundary condition. Then the following should hold:

$$\frac{\partial L}{\partial y'} = 0$$

And hence:

$$\frac{y'}{\sqrt{y(1 + (y')^2)}} = 0 \qquad \text{for} \quad x = b$$

We find that $y' = 0$ for $x = b$. This means that the tangent to the curve at its right end point must be horizontal. Hence $r = \frac{b}{\pi}$. We finally arrive at our solution: The required curve is given by:

$$x = \frac{b}{\pi}(\theta - \sin\theta), \qquad y = \frac{b}{\pi}(1 - \cos\theta)$$

We conclude that the minimizing curve of the brachistochrone problem is a cycloid.

Chapter 8

Conclusion

In this thesis we studied the calculus of variations. We found out that this method is used to solve optimization problems. Solving these problems consists of minimizing or maximizing functionals. Therefore, our first job was to explain what a functional is. We found out that the difference between functions and functionals is, that functionals not only have the independent variable as input but also the dependent variable and its derivative(s).
We then saw that functions that can be substituted into a functional must be C^1 continuous, they must satisfy the boundary conditions, and to each x must belong exactly one y.

After discussing functionals we looked at the basic one dimensional functional $J[y]$ with Dirichlet boundary conditions. We studied the first variation to figure out which functions are candidates for the optimizer. We found out that any optimizer $y(x)$ to the functional $J[y]$ must satisfy the Euler-Lagrange equation and the prescribed boundary conditions. To figure out whether this candidate is indeed the minimizer we looked at the second variation. We found that the necessary and sufficient conditions that we need to determine the nature of an optimizer must satisfy the following conditions:

1. The curve $y = y(x)$ satisfies the Euler-Lagrange equation.
2. The strict Legendre condition is satisfied ($P(x) = \partial^2 L/\partial(y')^2 > 0$ for a minimum, $P(x) < 0$ for a maximum).
3. The interval $[a, b]$ contains no conjugate points to a.

With these conditions, we could solve three basic one dimensional minimization problems. We found out that the shortest path between two points is a straight line and that the solution to the minimal surface of revolution problem is a catenary. Solving the brachistochrone problem turned out to be difficult, so we postponed it to the last chapter.

Since we could now only solve the most elementary problems, we wanted to find out how to solve multi-dimensional problems. We derived the Euler-Lagrange equations for two independent variables, several independent variables, several dependent variables, second order derivative and higher order derivatives using the same method as for the one dimensional case.

After this we actually solved two multi-dimensional problems. The first problem, which is called Plateaus problem, about minimal surfaces we could not solve completely, but we found the non-linear second order PDE that the function must satisfy.
Then we solved a problem involving the second order derivative.

Since we could not solve the brachistochrone problem, we studied in our final chapter a new way to solve optimization problems. We used canonical variables instead of the usual variables. Using the Hamiltonian function we found the canonical system of Euler-Lagrange equations. We also found out that if the Lagrangian does not depend on x, then the Hamiltonian is a first integral of the system of Euler-Lagrange equations. If the Lagrangian does depend on x, then a function is a first integral of the system of Euler-Lagrange equations if and only if the Poisson bracket $[\Phi, H]$ vanishes.
Our final results were that the Euler-Lagrange equation and the canonical Euler-Lagrange equations are equivalent and that the Euler-Lagrange equation is invariant under coordinate transformations. We ended this thesis with Noether's theorem: every one-parameter family of transformations leaving $J[y]$ invariant leads to a first integral of its system of Euler-Lagrange equations.
We used the results of this last chapter to solve the brachistochrone problem and found out that the minimizing curve was a cycloid.

This thesis was a good introduction for everyone who wants to get an idea of what calculus of variations is. But the calculus of variations is much broader than what we have discussed here. For example variable end point(s), the theory of fields and problems of optimal control are important in the calculus of variations. Therefore, we refer the interested reader to [4], [8] and [7] to learn more about this interesting field of mathematics.

Bibliography

[1] C. Fellipa. *Introduction to Variational Calculus.* University of Colorado, Colorado, 2013.

[2] P. J. Olver. *Introduction to the Calculus of Variations.* University of Minnesota, Minnesota, 2015.

[3] J. Stewart. *Calculus, Early Transcendentals.* Brooks Cole, Canada, 7th international metric edition edition, 2012.

[4] I.M. Gel'fand and S.V. Fomin. *Calculus of Variations.* Dover Publications, New York, 2000.

[5] J. Douglas. *Solution of the problem of Plateau.* Transactions of the American Mathematical Society 33 nr.1, 1931.

[6] R. Tibor. *On Plateau's problem.* Annals of Mathematics 31 nr.3, 1930.

[7] L. D. Elsgolc. *Calculus of Variations.* Dover Publications, New York, 2007.

[8] H. Sagan. *Introduction to the Calculus of Variations.* Dover Publications, New York, 1969.