CHAPTER 1

First Order Equations

1. The constant velocity transport equation — characteristics

We consider the equation

$$\text{(CVT)} \qquad \frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = 0.$$

If $u(x,t)$ is a continuously differentiable solution, then by the several variable chain rule

$$\frac{du(x_0+ct,t)}{dt} = \frac{\partial u}{\partial x}\frac{d(x_0+ct)}{dt} + \frac{\partial u}{\partial t} = cu_x + u_t = 0$$

for any $x_0, t \in \mathbb{R}$. This means that any solution is constant along the lines $x(t) = x_0 + ct$ $(t \in \mathbb{R})$. These lines are called the characteristics of the equation.

2. The initial value problem

Suppose we are given the values $u(x,0) = F(x)$ of the solution at time $t = 0$ for all $x \in \mathbb{R}$. Then for any $(x,t) \in \mathbb{R}^2$ one has

$$u(x,t) = u(x-ct,0) = F(x-ct).$$

In other words, if there is a solution with the prescribed initial values then it must be $u(x,t) = F(x-ct)$.

On the other hand, if $F : \mathbb{R} \to \mathbb{R}$ is C^1 then you can verify by substituting that $u(x,t) = F(x-ct)$ satisfies the transport equation (CVT), i.e. $u_t + cu_x = 0$.

3. Solutions without derivatives

Much of the modern theory of partial differential equations deals with "generalized solution" of one kind or another. Here is an example that begins to show why one would want to do that.

Consider the solutions with initial values $u_n(x,0) = \arctan(nx)$. The solutions are

$$u_n(x,t) = \arctan(n(x-ct)).$$

What happens if we let $n \to \infty$? We get

$$u_\infty(x,0) = \lim_{n\to\infty} u_n(x,0) = \lim_{n\to\infty} \arctan nx = \begin{cases} +\pi/2 & (x>0), \\ -\pi/2 & (x<0), \end{cases}$$

and

$$u_\infty(x,t) = \lim_{n\to\infty} u_n(x,t) = \lim_{n\to\infty} \arctan n(x-ct) = \begin{cases} +\pi/2 & (x>ct), \\ -\pi/2 & (x<ct), \end{cases}$$

The function $u_\infty(x,t)$ is the limit of solutions, but it is not continuous and therefore not differentiable because it has a jump discontinuity at $x = ct$. It can't be a solution of the PDE $u_t + cu_x = 0$ because the derivatives in the equation aren't defined at all (x,t).

Nevertheless, u_∞ is a function of the form $F(x-ct)$ and it arises as the limit of actual solutions, so we would like to call it a solution. The discontinuous solution u_∞ is itself not a solution of the PDE, but it does give a good description of all smooth solutions that are close to u_∞.

This leads to the question: how do we change the definition of solution to allow for discontinuous solutions?

4. Method of characteristics

There is a general method, called the method of characteristics for finding the solutions to a first order equation of the form

$$\frac{\partial u}{\partial t} + a(x,t)\frac{\partial u}{\partial x} = f(x,t,u) \tag{1}$$

in which $a : \mathbb{R}^2 \to \mathbb{R}$ and $f : \mathbb{R}^3 \to \mathbb{R}$ are continuously differentiable functions.

4.1. Definition. A characteristic for the equation (1) is a function $x_c : [t_0, t_1] \to \mathbb{R}$ that is defined on some interval $[t_0, t_1] \subset \mathbb{R}$ and that satisfies the characteristic equation

$$\frac{dx_c(t)}{dt} = a\big(x_c(t), t\big) \qquad \text{for all } t \in [t_0, t_1] \tag{2}$$

Often the graph of x_c, i.e. the curve $\{(t, x_c(t)) \in \mathbb{R}^2 \mid t_0 \leqslant t \leqslant t_1\}$ is also called a characteristic of the equation (1).

4.2. Theorem. If $x_c : [t_0, t_1] \to \mathbb{R}$ is a characteristic of equation (1) then the function $u_c(t) = u(x(t), t)$ satisfies

$$u_c'(t) = f\big(x_c(t), t, u_c(t)\big). \tag{3}$$

In the special case where $f(x,t,u) = 0$ for all (x,t,u) it follows that $u_c(t) = u_c(t_0)$ for all $t \in [t_0, t_1]$, i.e. if the right hand side f in the equation (1) is just 0, then any C^1 solution u is constant along a characteristic.

Note that $u_c(t)$ is a function of one variable and (3) is an ordinary differential equation.

Proof. To derive (3) let $x_c : \mathbb{R} \to \mathbb{R}$ be a characteristic of the PDE (1), so that $x_c : \mathbb{R} \to \mathbb{R}$ satisfies (2). Then the several variable chain rule implies

$$\frac{du(x_c(t),t)}{dt} = u_x\big(x_c(t),t\big)\cdot x_c'(t) + u_t\big(x_c(t),t\big).$$

Since u satisfies (1), we can write this as

$$\frac{du(x_c(t),t)}{dt} = u_x \cdot x_c'(t) - a(x_c(t),t)\cdot u_x + f(x_c(t), t, u(x_c(t),t)).$$

The characteristic equation (2) implies $x_c'(t) = a(x_c(t),t)$ so the first two terms on the left cancel. This proves (3).

We next show that if $f(x,t,u) = 0$ for all x, t, u, then solutions are constant along characteristics. If $f = 0$ and if $x_c : [t_0, t_1] \to \mathbb{R}$ is a characteristic, then (3) implies that $u_c(t) = u(x_c(t), t)$ is a function of one variable that is differentiable for all $t \in [t_0, t_1]$, and that satisfies $u_c'(t) = 0$ for all $t \in [t_0, t_1]$. This implies that $u_c(t)$ does not depend on t, i.e. $u_c(t) = u_c(t_0)$ for all $t \in [t_0, t_1]$. *Q.E.D.*

4.3. The recipe for solving initial value problems for 1st order PDE. We can use the above Theorem to solve an initial value problem for the partial differential equation (1), i.e.

$$\frac{\partial u}{\partial t} + a(x,t)\frac{\partial u}{\partial x} = f(x,t,u) \tag{1}$$

Special case where the right hand side $f(x,t,u) = 0$. In this case (1) is the equation $u_t + a(x,t)u_x = 0$, and we can proceed as follows. If we are given the initial values $u_0(x) = u(x,0)$ of a solution for all $x \in \mathbb{R}$, and we want to find value of the solution u at some point (x_1, t_1) in space and time, then

- find the characteristic through (x_1, t_1), i.e. find the solution $x_c : \mathbb{R} \to \mathbb{R}$ of
$$x_c'(t) = a(x_c(t), t), \qquad x_c(t_1) = x_1$$
- Check that the characteristic is defined for all $t \in [0, t_1]$
- It follows from equation (3) that $\frac{du(x_c(t),t)}{dt} = 0$ for all $t \in [0, t_1]$. Therefore $u(x_c(t), t)$ does not depend on t and we have
$$u(x_1, t_1) = u\big(x_c(t_1), t_1\big) = u\big(x_c(0), 0\big) = u_0(x_c(0)).$$

General case with arbitrary right hand side $f(x,t,u)$.

- find the characteristic through (x_1, t_1), i.e. find the solution $x_c : \mathbb{R} \to \mathbb{R}$ of
$$x_c'(t) = a(x_c(t), t), \qquad x_c(t_1) = x_1$$
- Check that the characteristic is defined for all $t \in [0, t_1]$ * It follows from equation (3) that $\dfrac{du(x_c(t),t)}{dt} = f(x_c(t), t, u(x_c(t), t))$ for all $t \in [0, t_1]$. To find $u(x_c(t), t)$ define $u_c(t) = u(x_c(t), t)$ and solve the differential equation
$$u_c'(t) = f(x_c(t), t, u_c(t)), \qquad u_c(0) = u_0(x_c(0)).$$
- If the solution $u_c(t)$ is defined for all $t \in [0, t_1]$ then we have
$$u(x_1, t_1) = u\big(x_c(t_1), t_1\big) = u_c(t_1).$$

5. Equations for the derivatives of a solution

If u is a solution of
$$u_t + \sin(x)u_x = u^2$$
then the derivative $v = u_x$ also satisfies a partial differential equation. We get this equation by differentiating the equation for u with respect to x on both sides:
$$\frac{\partial\big(u_t + \sin(x)u_x\big)}{\partial x} = \frac{\partial u^2}{\partial x} \implies \frac{\partial u_t}{\partial x} + \cos(x)\frac{\partial u}{\partial x} + \sin(x)\frac{\partial u_x}{\partial x} = 2u\frac{\partial u}{\partial x}$$
Rewrite the second order derivatives that appeared as derivatives of v using
$$\frac{\partial u_x}{\partial x} = \frac{\partial v}{\partial x} \qquad \text{and} \qquad \frac{\partial u_t}{\partial x} = \frac{\partial^2 u}{\partial x \partial t} = \frac{\partial^2 u}{\partial t \partial x} = \frac{\partial u_x}{\partial t} = v_t.$$
We get
$$v_t + \cos(x)\frac{\partial u}{\partial x} + \sin(x)\frac{\partial v}{\partial x} = 2u\frac{\partial u}{\partial x}$$

Finally, replace every $\frac{\partial u}{\partial x}$ by v:

$$v_t + \sin(x)v_x = 2uv - \cos(x)v$$

6. A nonlinear equation

Consider the so-called inviscid Burger's equation

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0. \tag{4}$$

If $u(x,t)$ is a C^1 solution then any level set of u on which $u_x \neq 0$ is a straight line.

Suppose a level set is a graph $x = x(t)$, i.e. suppose that for some function $x = x(t)$ one has $u(x(t),t) = c$ for all t. Then

$$\left.\begin{aligned} 0 = \frac{dc}{dt} &= \frac{du(x(t),t)}{dt} \\ &= u_x(x(t),t)x'(t) + u_t(x(t),t) \\ &= u_x(x(t),t)x'(t) - u(x(t),t)u_x(x(t),t) \\ &= u_x(x(t),t)x'(t) - cu_x(x(t),t) \\ &= \big(x'(t) - c\big)u_x(x(t),t) \end{aligned}\right\} \implies x'(t) = c.$$

7. Problems

7.1. Suppose $u : [0,\infty) \times [0,\infty) \to \mathbb{R}$ is a C^1 solution of

$$u_t + xu_x = 0.$$

1. Find the characteristics of the equation.
2. If you are given the initial condition $u(x,0) = u_0(x)$ for all $x \in [0,\infty)$, can you compute $u(x,t)$ for all $x,t \geq 0$?
3. Show that the function $v(x,t) = u_x(x,t)$ satisfies the partial differential equation $v_t + xv_x = -v$.
4. Show that along a characteristic $x(t)$ of the equation for v one has $u_x(x(t),t) = Ae^{Bt}$ and find A, B.

Solution.

1. The characteristic equation is $\frac{dx}{dt} = x$, so the characteristics are $x_c(t) = x_c(0)e^t$.

2. To compute $u(x,t)$ consider the characteristic for which $x_c(t) = x$. For this characteristic we have $x = x_c(t) = x_c(0)e^t$ and therefore $x_c(0) = xe^{-t}$. Along the characteristic we have $\frac{d}{dt}u(x_c(t),t) = 0$, so $u(x_c(t),t) = u(x_c(0),0)$. Substitute $x_c(t) = x$ and $x_c(0) = xe^{-t}$ and we get $u(x,t) = u(xe^{-t},0) = u_0(xe^{-t})$.

3. Differentiate the equation $u_t + xu_x = 0$ with respect to x:

$$\begin{aligned} \frac{\partial(u_t + xu_x)}{\partial x} = 0 &\implies u_{tx} + \frac{\partial x}{\partial x}u_x + xu_{xx} = 0 && \text{(use } u_{tx} = u_{xt}\text{)} \\ &\implies (u_x)_t + v + xv_x = 0 && \text{(use } u_x = v, u_{xx} = v_x\text{)} \\ &\implies v_t + xv_x = -v. \end{aligned}$$

4. The characteristics for the v equation are determined by $\frac{dx}{dt} = x$, so they are the same as the characteristics for the equation for u. They are given by $x_c(t) = x_c(0)e^t$. Along a characteristic we have $\frac{dv(x_c(t),t)}{dt} = -v(x_c(t),t)$. Solving this ordinary differential equation we get $v(x_c(t),t) = v(x_c(0),0)e^{-t}$, i.e.

$$u_x(x_c(t),t) = u_x(x_c(0),0)e^{-t}.$$

7.2. Suppose $u : [0,\infty) \times [0,\infty) \to \mathbb{R}$ is a C^1 solution of

$$u_t = x^2 u_x.$$

1. Find the characteristics of the equation for u.
2. If you are given the initial condition $u(x,0) = u_0(x)$ for all $x \in [0,\infty)$, can you compute $u(x,t)$ for all $x,t \geq 0$?

Solution. The Partial Differential Equation for u is $u_t - x^2 u_x = 0$, so the differential equation for the characteristics is $\frac{dx}{dt} = -x^2$ ......note the minus sign! Solving this equation we get:

$$-\frac{1}{x^2}\frac{dx}{dt} = 1 \implies \frac{d\,\frac{1}{x}}{dt} = t \implies \frac{1}{x} = t + C \implies x_c(t) = \frac{1}{t+C}.$$

To find the value of $u(x,t)$ for given x,t we need to know the characteristic for which $x_c(t) = x$. The constant C and the initial value $x_c(0)$ for this characteristic follow from

$$x = \frac{1}{t+C} \implies C = \frac{1}{x} - t \implies x_c(0) = \frac{1}{0+C} = \frac{1}{1/x - t} = \frac{x}{1-xt}.$$

Along the characteristic we have

$$(\dagger) \qquad \frac{du(x_c(t'),t')}{dt'} = 0$$

and we would like to conclude that $u(x_c(t),t) = u(x_c(0),0)$. But we can only conclude this if we know that $(\dagger)$ holds for all $t' \in [0,t]$. The equation $(\dagger)$ holds at all t' at which the characteristic $x_c(t')$ is defined. The characteristic x_c is given by

$$x_c(t') = \frac{1}{t'+C}$$

so it is defined for all $t' \in [0,t]$ if $t' + C \neq 0$ for all $t' \in [0,t]$. This holds if either $C > 0$ or $C < -t$. Using $C = \frac{1}{x} - t$ we now check for which x,t this is true.

It is given that $x > 0$ so $C = \frac{1}{x} - t > -t$ is true. We still have to see when $C > 0$. This holds if $\frac{1}{x} > t$, i.e. if $xt < 1$.

We can therefore only determine the value of $u(x,t)$ if $xt < 1$ (assuming $x,t \geq 0$).

7.3. About the inviscid Burger's equation. Let $u : \mathbb{R} \times [0,T) \to \mathbb{R}$ be a C^2 solution to the inviscid Burger's equation (4). Let $a \in \mathbb{R}$ be given, and define $c = u(a,0)$.

1. What can you say about the level set $\{(x,t) \mid u(x,t) = c\}$? 2. Show that the function $v(x,t) = u_x(x,t)$ satisfies the partial differential equation

$$v_t + uv_x + v^2 = 0.$$

3. Show that $S(t) = u_x(a+ct,t)$ satisfies $S'(t) = -S(t)^2$. 4. Assume that $u(x,0) = \frac{2}{1+e^x}$ and compute $u_x(t,t)$ for all $t > 0$; show that there is a $T > 0$ such that

$$\lim_{t \nearrow T} u_x(t,t) = -\infty.$$

Solution.

1. On the level set $\{(x,t) \mid u(x,t) = c\}$ we have $u_t = -cu_x$. If $u_x \neq 0$ along the level set, then the Implicit Function Theorem says that near each point (x_0, t_0) with $u(x_0, t_0) = c$, the level set is the graph of a function $x = x_c(t)$. The derivative of this function follows from

$$0 = \frac{dc}{dt} = \frac{du(x_c(t),t)}{dt} = \frac{\partial u}{\partial x} x_c'(t) + \frac{\partial u}{\partial t} \implies x_c'(t) = -\frac{u_t}{u_x} = -\frac{-cu_x}{u_x} = c.$$

Thus the level set is a straight line with slope $x_c'(t) = c$. Since $u(a,0) = c$ we also know that $x_c(0) = a$, so $x_c(t) = a + ct$.

2. Differentiate the equation $u_t + uu_x = 0$ with respect to x to get

$$u_{tx} + u_x u_x + uu_{xx} = 0 \implies u_{xt} + u_x u_x + uu_{xx} = 0 \implies v_t + v^2 + uv_x = 0.$$

3. We have

$$S'(t) = \frac{dv(a+ct,t)}{dt} = cv_x(a+ct,t) + v_t(a+ct,t)$$

Furthermore $u(a+ct,t) = c$, so

$$S'(t) = u(a+ct,t)v_x(a+ct,t) + v_t(a+ct,t) = -v(a+ct,t)^2 = -S(t)^2.$$

4. The initial function is $u(x,0) = \dfrac{2}{1+e^x}$, which satisfies $u(0,0) = 1$. The level set through $(0,0)$ is therefore a line with slope 1, and is thus given by $x(t) = t$: the level set through $(0,0)$ is $\{(t,t) \mid t \geq 0\}$.

By the previous problem $S(t) = u_x(t,t)$ satisfies $S'(t) = -S(t)^2$.

At $t = 0$ we have

$$S(0) = u_x(0,0) = \left(\frac{\partial}{\partial x}\frac{2}{1+e^x}\right)_{x=0} = -\frac{1}{2}.$$

Solving $S'(t) = -S(t)^2$, $S(0) = -\frac{1}{2}$ we find $S(t) = \dfrac{1}{t-2}$. Hence $\lim_{t\nearrow 2} S(t) = -\infty$.

CHAPTER 2

The 1D Wave Equation

1. Derivation, vibrating string, vibrating membranes

The wave equation with several space variables is

$$\frac{\partial^2 u}{\partial t^2} - \Delta u = 0, \qquad \Delta u \overset{\text{def}}{=} \frac{\partial^2 u}{\partial x_1^2} + \cdots + \frac{\partial^2 u}{\partial x_n^2}$$

or more generally

$$\frac{\partial^2 u}{\partial t^2} - c^2 \Delta u = 0,$$

where $c > 0$ is a positive real number called the wave speed.

The unknown is a function $u : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$.

The equation appears in physics in the following contexts

- vibrating string ($n = 1$, slides for the derivation)
- vibrating membrane ($n = 2$)
- propagation of light waves, or of sound waves ($n = 3$)
- more …

2. D'Alembert's solution ($n = 1$)

2.1. Changing coordinates. Suppose that $u(x, t)$ satisfies

(1) $$u_{tt} - u_{xx} = 0 \text{ for all } (x, t).$$

We introduce new coordinates (p, q) related to our original space-time coordinates (x, t) by

$$p = x + t, \qquad q = x - t$$

If we know the (p, q) coordinates of a point then we can recover the (x, t) coordinates of the point from

$$x = \frac{p + q}{2}, \qquad t = \frac{p - q}{2}.$$

Consider the quantity u expressed as a function of the (p, q) coordinates, i.e. we consider the new function

$$v(p, q) \overset{\text{def}}{=} u\Big(\frac{p + q}{2}, \frac{p - q}{2}\Big).$$

and ask what equation it satsifies. The computation is easiest if we first write $u(x, t)$ in terms of v, i.e.

$$u(x, t) = v(x + t, x - t).$$

We can substitute this in the wave equation and apply the chain rule several time. We compute the derivatives one by one. The first derivative is

$$\begin{aligned} u_t &= \frac{\partial v(\overbrace{x+t}^{=p}, \overbrace{x-t}^{=q})}{\partial t} \\ &= v_p(x+t, x-t)\frac{\partial(x+t)}{\partial t} + v_q(x+t, x-t)\frac{\partial(x-t)}{\partial t} \\ &= v_p(x+t, x-t) - v_q(x+t, x-t) \end{aligned}$$

Differentiate again:

$$\begin{aligned} u_{tt} &= \frac{\partial\{v_p(x+t, x-t) - v_q(x+t, x-t)\}}{\frac{\partial V}{\partial x} ialt} \\ &= v_{pp} - v_{pq} - v_{qp} + v_{qq} \quad \text{at } (x+t, x-t) \\ &= v_{pp} - 2v_{pq} + v_{qq} \quad \text{at } (x+t, x-t) \end{aligned}$$

Similarly, the second x derivative is

$$u_{xx} = v_{pp} + 2v_{pq} + v_{qq} \quad \text{at } (x+t, x-t)$$

Therefore

$$u_{tt} - u_{xx} = v_{pp} - 2v_{pq} + v_{qq} - \left(v_{pp} + 2v_{pq} + v_{qq}\right) = -4v_{pq}$$

We see that if u is a C^2 function, then u satisfies the wave equation if and only if v satisfies the equation

(2) $$\frac{\partial^2 v}{\partial p \partial q} = 0 \text{ for all } (p, q).$$

The point of choosing the coordinates p, q is that the new equation (2) is easier to solve than the original wave equation (1).

2.2. Theorem — C^2 solutions to the Wave Equation are sums of traveling waves.

(a) If $u : \mathbb{R}^2 \to \mathbb{R}$ is a C^2 solution of the wave equation $u_{tt} = u_{xx}$ then there exist C^2 functions $F, G : \mathbb{R} \to \mathbb{R}$ such that $u(x, t) = F(x+t) + G(x-t)$ for all $(x, t) \in \mathbb{R}^2$.

(b) If $F, G : \mathbb{R} \to \mathbb{R}$ are C^2 functions then $u(x, t) = F(x+t) + G(x-t)$ satisfies the wave equation.

Proof. We begin with the proof of (a).

If u satisfies the wave equation then we have shown that $v(p, q)$ defined above satisfies

$$\frac{\partial^2 v}{\partial q \partial p} = \frac{\partial v_p}{\partial q} = 0.$$

Hence v_p does not depend on q, i.e. $v_p(p, q)$ is a function of p only, i.e. there is a function $f : \mathbb{R} \to \mathbb{R}$ such that

$$\frac{\partial v}{\partial p} = f(p)$$

We can now integrate this. Let $F : \mathbb{R} \to \mathbb{R}$ be an antiderivative of f. Then there is a function $G : \mathbb{R} \to \mathbb{R}$ such that

$$v(p, q) = \int f(p)\, dp + G(q) = F(p) + G(q)$$

Substitute $p = x + t$, $q = x - t$ and we get

$$u(x, t) = F(x - t) + G(x + t)$$

If we are given the function $u : \mathbb{R}^2 \to \mathbb{R}$ then we can find the functions F and G up to a constant from the formulas

$$u(s/2, s/2) = F(s) + G(0), \qquad u(s/2, -s/2) = F(0) + G(s).$$

We find these formulas by setting $x = t = s/2$ (for F) or $x = s/2, t = -s/2$ (for G) in $u(x, t) = F(x + t) + G(x - t)$.

If u is a C^2 function then it follows from $F(s) = u(s/2, s/2) - G(0)$ that

$$F'(s) = \tfrac{1}{2} u_x(\tfrac{s}{2}, \tfrac{s}{2}) + \tfrac{1}{2} u_t(\tfrac{s}{2}, \tfrac{s}{2}), \qquad F''(s) = \tfrac{1}{4} u_{xx}(\tfrac{s}{2}, \tfrac{s}{2}) + \tfrac{1}{2} u_{xt}(\tfrac{s}{2}, \tfrac{s}{2}) + \tfrac{1}{4} u_{tt}(\tfrac{s}{2}, \tfrac{s}{2})$$

Therefore F also has two derivatives, and they are continuous.

Now we turn to the proof of (b). If F, G are given functions, and if $u(x, t) = F(x - t) + G(x + t)$, then

$$u_t(x, t) = F'(x + t) - G'(x - t), \qquad u_x(x, t) = F'(x + t) + G'(x - t)$$

and hence

$$u_{tt}(x, t) = F''(x + t) + G''(x - t), \qquad u_{xx}(x, t) = F''(x + t) + G''(x - t)$$

This implies that $u_{tt} = u_{xx}$, so u satisfies the wave equation. *Q.E.D.*

2.3. Theorem — solution in terms of the initial values and velocities. If u is a C^2 solution of the wave equation, then

$$u(x, t) = \frac{u(x - t, 0) + u(x + t, 0)}{2} + \frac{1}{2} \int_{x-t}^{x+t} u_t(\xi, 0) \mathrm{d}\xi$$

Conversely, if $U, V : \mathbb{R} \to \mathbb{R}$ are two functions of which U is C^2 and V is C^1, then

$$u(x, t) = \frac{U(x - t) + U(x + t)}{2} + \frac{1}{2} \int_{x-t}^{x+t} V(\xi) \mathrm{d}\xi$$

is a solution of the wave equation which satisfies

$$u(x, 0) = U(x), \qquad u_t(x, 0) = V(x) \qquad \text{for all } x \in \mathbb{R}.$$

Proof. If u is a solution then it is of the form $u(x, t) = F(x + t) + G(x - t)$ for certain functions F and G. To find these functions we note that

$$u_t(x, t) = F'(x + t) - G'(x - t).$$

Set $t = 0$:

$$u(x, 0) = F(x) + G(x), \quad u_t(x, 0) = F'(x) - G'(x)$$

Differentiate the first equation with respect to x,

$$u_x(x, 0) = F'(x) + G'(x), \quad u_t(x, 0) = F'(x) - G'(x)$$

which we solve for F' and G':

$$F'(x) = \frac{u_x(x,0)+u_t(x,0)}{2}, \qquad G'(x) = \frac{u_x(x,0)-u_t(x,0)}{2}$$

The fundamental theorem of calculus implies

$$\begin{aligned} F(x) &= F(0) + \frac{1}{2}\int_0^x \big\{u_x(\xi,0)+u_t(\xi,0)\big\}\,d\xi \\ &= F(0) + \frac{1}{2}u(x,0) - \frac{1}{2}u(0,0) + \frac{1}{2}\int_0^x u_t(\xi,0)\,d\xi \end{aligned}$$

Similarly,

$$G(x) = G(0) + \frac{1}{2}u(x,0) - \frac{1}{2}u(0,0) - \frac{1}{2}\int_0^x u_t(\xi,0)\,d\xi$$

It follows that the solution $u(x,t)$ is given by

$$\begin{aligned} u(x,t) &= F(x+t) + G(x-t) \\ &= \frac{u(x+t,0)+u(x-t,0)}{2} + \frac{1}{2}\int_0^{x+t} u_t(\xi,0)\,d\xi - \frac{1}{2}\int_0^{x-t} u_t(\xi,0)\,d\xi \\ &\qquad + F(0) + G(0) - u(0,0). \end{aligned}$$

We can combine the two integrals:

$$\int_0^{x+t} u_t(\xi,0)\,d\xi - \int_0^{x-t} u_t(\xi,0)\,d\xi = \int_0^{x+t} u_t(\xi,0)\,d\xi + \int_{x-t}^0 u_t(\xi,0)\,d\xi = \int_{x-t}^{x+t} u_t(\xi,0)\,d\xi.$$

The terms $F(0)+G(0)-u(0,0)$ cancel because

$$u(x,t) = F(x+t) + G(x-t) \implies u(0,0) = F(0) + G(0).$$

Therefore we end up with

$$u(x,t) = \frac{u(x-t,0)+u(x+t,0)}{2} + \frac{1}{2}\int_{x-t}^{x+t} u_t(\xi,0)\,d\xi$$

as claimed.

3. Weak and classical solutions

For many partial differential equations the naive notion of a "solution" is not satisfactory, as there may be functions that do not have enough derivatives to allow verification of the equation, but which one really would like to call solutions, for other reasons. This has led to the introduction of many theories of "generalized solution." Here we will see one version of such a theory.

3.1. Definition. A classical solution of the wave equation is a C^2 function $u : \mathbb{R}^2 \to \mathbb{R}$ that satisfies the equation, i.e. for which $u_{tt} = u_{xx}$ holds at all $(x,t) \in \mathbb{R}^2$.

A weak solution of the wave equation is a continuous function that satisfies

$$(\star) \qquad u(x+h,t) + u(x-h,t) = u(x,t+h) + u(x,t-h)$$

for all $x, t \in \mathbb{R}$ and $h > 0$.

The new notion of solution has the following features:

- checking if some function u satisfies $(\star)$ does not require one to differentiate u

- Classical solutions, which were our original idea of what a "solution" ought to be, should still be solutions
- Weak solutions should have some of the same properties as classical solutions

3.2. Theorem. If $F, G : \mathbb{R} \to \mathbb{R}$ are continuous functions, then $u(x,t) = F(x+t) + G(x-t)$ is a weak solution to the Wave Equation.

Proof. For any $x, t \in \mathbb{R}$ and $h > 0$ we have

$$u(x+h,t) + u(x-h,t) = F(x+h+t) + G(x+h-t) + F(x-h+t) + G(x-h-t)$$
$$u(x,t+h) + u(x,t-h) = F(x+t+h) + G(x-t-h) + F(x+t-h) + G(x-t+h).$$

Therefore $(\star)$ holds. *Q.E.D.*

3.3. Theorem. Any classical solution of the wave equation is also a weak solution.

Proof. Since u is a C^2 solution of the wave equation there exist functions F, G such that $u(x,t) = F(x+t) + G(x-t)$ for all x, t. The previous theorem then implies that u is a weak solution *Q.E.D.*

3.4. Theorem. If u is a weak solution, and if u is C^2, then u is also a classical solution of the wave equation.

Proof. Suppose u is a C^2 function satisfying $(\star)$. Then we differentiate $(\star)$ on both sides twice with respect to h and set $h = 0$

$$\begin{aligned} & u(x+h,t) + u(x-h,t) = u(x,t+h) + u(x,t-h) \\ \Longrightarrow \quad & u_x(x+h,t) - u_x(x-h,t) = u_t(x,t+h) - u_t(x,t-h) \\ \Longrightarrow \quad & u_{xx}(x+h,t) + u_{xx}(x-h,t) = u_{tt}(x,t+h) + u_{tt}(x,t-h) \\ \Longrightarrow \quad & 2u_{\times}(x,t) = 2u_{tt}(x,t) \end{aligned}$$

Therefore u satisfies the wave equation. *Q.E.D.*

3.5. Theorem. If $u : \mathbb{R}^2 \to \mathbb{R}$ is a weak solution of the Wave Equation, then there exist continuous function $F, G : \mathbb{R} \to \mathbb{R}$ such that

$$u(x,t) = F(x+t) + G(x-t)$$

for all x, t.

Proof. Define

$$F(x) = u\left(\frac{x}{2}, \frac{x}{2}\right), \qquad G(x) = u\left(\frac{x}{2}, -\frac{x}{2}\right) - u(0,0).$$

and consider the function

$$v(x,t) \stackrel{\text{def}}{=} F(x+t) + G(x-t).$$

We will show that $u(x,t) = v(x,t)$ for all x, t. To do this we begin by verifying

- $u(t,t) = v(t,t)$ for all $t \in \mathbb{R}$
- v is a continuous function
- v also satisfies the condition $(\star)$

Let $h > 0$ be given. Then by repeatedly using the condition $(\star)$ we conclude that

$$u((m-n)h, (m+n)h) = v((m-n)h, (m+n)h) \text{ for all } m, n \in \mathbb{Z}$$

If $(x, t) \in \mathbb{R}^2$ is a given point, then we can find sequences $m_k, n_k \in \mathbb{Z}$ such that $(m_k - n_k)2^{-k} \to x$ and $(m_k + n_k)2^{-k} \to t$ as $k \to \infty$. Since u and v are continuous functions we then have

$$u(x,t) = \lim_{k\to\infty} u\left(\frac{m_k - n_k}{2^k}, \frac{m_k + n_k}{2^k}\right) = \lim_{k\to\infty} v\left(\frac{m_k - n_k}{2^k}, \frac{m_k + n_k}{2^k}\right) = v(x,t).$$

Q.E.D.

4. Problems

4.1. Suppose that $u(x, t)$ is a solution of

$$u_{tt} - 2u_{xt} - 15u_{xx} = 0 \tag{†}$$

For certain values of $c \in \mathbb{R}$ the function $u(x, t) = F(x - ct)$ is a solution of (†) for any $F : \mathbb{R} \to \mathbb{R}$ that is C^2. Find all $c \in \mathbb{R}$ with this property.

Solution. Substitute $u(x, t) = F(x - ct)$ in the equation:

$$u_t = -cF'(x - ct), \quad u_x = F'(x - ct),$$
$$u_{tt} = c^2F''(x - ct), \quad u_{xt} = u_{tx} = -cF''(x - ct), \quad u_{xx} = F''(x - ct)$$

implies

$$u_{tt} - 2u_{xt} - 15u_{xx} = \left(c^2 + 2c - 15\right)F''(x - ct).$$

Since F is allowed to be any C^2 function, we cannot assume that $F''(x - ct) = 0$ for all x, t. Therefore u is a solution for every choice of F if and only if

$$c^2 + 2c - 15 = 0, \text{ i.e. iff } c = +3 \text{ or } c = -5.$$

4.2. Consider the PDE

$$u_{tt} - u_{xt} - 2u_{xx} = 0 \tag{‡}$$

and consider the coordinate transformation

$$\begin{matrix} r = x - t \\ s = x + 2t \end{matrix} \iff \begin{matrix} x = \frac{1}{3}(2r + s) \\ t = \frac{1}{3}(-r + s) \end{matrix}$$

1. Which differential equation does the function $v(r, s) = u\left(\frac{2r+s}{3}, \frac{-r+s}{3}\right)$ satisfy? (The computation is easier if you begin with $u(x, t) = v(x - t, x + 2t)$ and substitute that in the equation for u).
2. Find the general solution for (‡).

Solution.

1. Substitute $u(x, t) = v(x - t, x + 2t) = v(r, s)$ in the equation:

$$\begin{aligned} u_t &= -v_r + 2v_s \\ u_x &= v_r + v_s \end{aligned} \implies \begin{aligned} u_{tt} &= v_{rr} - 4v_{rs} + 4v_{ss} \\ u_{xt} &= -v_{rr} + v_{rs} + 2v_{ss} \\ u_{xx} &= v_{rr} + 2v_{rs} + v_{ss} \\ \hline u_{tt} - u_{xt} - 2u_{xx} &= 9v_{rs} \end{aligned}$$

2. Hence u is a solution of the equation if and only if $v_{rs} = 0$,i.e. if $v(r, s) = F(r) + G(s)$ for certain functions $F, G : \mathbb{R} \to \mathbb{R}$. The general solution to the equation is therefore

$$u(x, t) = F(x - t) + G(x + 2t).$$

4.3. Problem: Differentiability of d'Alembert's solution. Suppose $f : \mathbb{R} \to \mathbb{R}$ is C^2 and $g : \mathbb{R} \to \mathbb{R}$ is C^1, and consider

$$u(x, t) = \frac{f(x + t) + f(x - t)}{2} + \frac{1}{2} \int_{x-t}^{x+t} g(\xi)\, d\xi.$$

1. Find C^2 functions $F.G : \mathbb{R} \to \mathbb{R}$ such that $u(x, t) = F(x + t) + G(x - t)$.
2. Show that u is a C^2 function. (This function appears in our formulation of d'Alembert's solution in Theorem 2.3; we never checked that u is twice differentiable, and this problem asks you to do that.)

Solution. 1. By the fundamental theorem of calculus we can write

$$\int_{x-t}^{x+t} g(\xi)\, d\xi = \int_0^{x+t} g(\xi)\, d\xi - \int_0^{x-t} g(\xi)\, d\xi.$$

Therefore the function u can be written as

$$u(x, t) = \frac{1}{2} f(x + t) + \frac{1}{2} \int_0^{x+t} g(\xi)\, d\xi + \frac{1}{2} f(x - t) - \frac{1}{2} \int_0^{x-t} g(\xi)\, d\xi.$$

If we now define

$$F(x) = \frac{1}{2} f(x) + \frac{1}{2} \int_0^x g(\xi)\, d\xi \quad \text{and} \quad G(x) = \frac{1}{2} f(x) - \frac{1}{2} \int_0^x g(\xi)\, d\xi$$

then we have $u(x, t) = F(x + t) + G(x - t)$.

2. Since g is C^1 the integral $\int_0^x g(\xi) d\xi$ is C^2. We are also given that f is C^2 so it follows that the two functions F and G are C^2. Therefore u is also C^2.

4.4. Problem: the initial value problem for weak solutions. Suppose $f : \mathbb{R} \to \mathbb{R}$ is C^1 and $g : \mathbb{R} \to \mathbb{R}$ is continuous, and again consider

$$u(x, t) = \frac{f(x + t) + f(x - t)}{2} + \frac{1}{2} \int_{x-t}^{x+t} g(\xi)\, d\xi.$$

1. Find C^1 functions $F.G : \mathbb{R} \to \mathbb{R}$ such that $u(x, t) = F(x + t) + G(x - t)$.

2. Show that $u(x, 0) = f(x)$ and $u_t(x, 0) = g(x)$.
3. Show that for every C^1 function $f : \mathbb{R} \to \mathbb{R}$ and every continuous $g : \mathbb{R} \to \mathbb{R}$ there is a weak solution to the wave equation with $u(x, 0) = f(x)$ and $u_t(x, 0) = g(x)$ for all $x \in \mathbb{R}$.

Solution. 1. Define F and G to be the same functions as in the previous problem.

2. Substitute $t = 0$ is the definition of u:

$$u(x, 0) = \frac{f(x + 0) + f(x - 0)}{2} + \frac{1}{2} \int_{x-0}^{x+0} g(\xi)\, d\xi = f(x).$$

To check the time derivative we differentiate $u(x,t)$ with respect to t and then set $t=0$:

$$\begin{aligned} u_t &= \frac{f'(x+t)-f'(x-t)}{2} + \frac{\partial}{\partial t}\left\{\frac{1}{2}\int_{x-t}^{x+t} g(\xi)\,d\xi\right\} \\ &= \frac{f'(x+t)-f'(x-t)}{2} + \frac{g(x+t)+g(x-t)}{2} \\ \implies u_t(x,0) &= \frac{f'(x+0)-f'(x-0)}{2} + \frac{g(x+0)+g(x-0)}{2} = g(x). \end{aligned}$$

3. If f is C^1 and g is C^0, then F and G are C^1 functions. In particular, they are continuous, so that $u(x,t) = F(x+t) + G(x-t)$ is a weak solution of the Wave Equation. The calculation in part 2 of this problem shows that $u(x, 0 = f(x)$ and $u_t(x,0) = g(x)$ for all $x \in \mathbb{R}$.

CHAPTER 3

Fourier Series and the Wave Equation

1. The complex exponential

Many of the computations involving Fourier series are simpler when we use the complex exponential instead of sine and cosine, so in this section we will consider complex valued solutions to the wave equation.

1.1. Euler's Formulas. By definition

$$e^{i\theta} = \cos\theta + i\sin\theta \quad \text{for all } \theta \in \mathbb{R}.$$

This implies

$$e^{-i\theta} = \cos\theta - i\sin\theta \quad \text{for all } \theta \in \mathbb{R}$$

and hence

$$\cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2}, \qquad \sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}$$

1.2. Derivative of the complex exponential. Euler's definition of $e^{i\theta}$ implies

$$\frac{de^{i\theta}}{d\theta} = ie^{i\theta}.$$

In the language of Linear Algebra, $e^{i\theta}$ is an eigenvector of the linear transformation $f \mapsto \frac{df}{d\theta}$ with eigenvalue i.

1.3. Trigonometric polynomials. A function of the form

$$f(x) = A_0 + \sum_{n=1}^{N} \big(A_n \cos nx + B_n \sin nx\big)$$

where $A_n, B_n \in \mathbb{C}$ are constants is called a trigonometric polynomial. Using Euler's formulas we can rewrite a trigonometric polynomial in terms of complex exponentials:

$$\begin{aligned} f(x) &= A_0 + \sum_{n=1}^{N} \big(A_n \cos nx + B_n \sin nx\big) \\ &= A_0 + \sum_{n=1}^{N} \frac{A_n - iB_n}{2} e^{inx} + \frac{A_n + iB_n}{2} e^{-inx} \\ &= \sum_{n=-N}^{N} \hat{f}_n e^{inx}, \end{aligned}$$

provided we define

$$\hat{f}_n = \begin{cases} (A_n - iB_n)/2 & n < 0 \\ (A_n + iB_n)/2 & n > 0 \\ A_0 & n = 0 \end{cases}$$

2. Finite string, Fourier solution

2.1. The wave equation is linear. If $u, v : \mathbb{R}^2 \to \mathbb{R}$ are solutions to the wave equation, then $w(x,t) = au(x,t) + bv(x,t)$ is also a solution for any choice of $a, b \in \mathbb{R}$. This is true both for classical solutions and for weak solutions; i.e. if u, v are classical solutions, then $w = au + bv$ is a classical solution, and if u, v are weak solutions, then $w = au + bv$ is a weak solution.

2.2. Solutions that are periodic in space. Instead of considering arbitrary solutions to the wave equation, we look for solutions for which

$$u(x + 2\pi, t) = u(x,t)$$

holds for all x, t. By direct substitution in the equation we can verify that each of the functions

$$e^{int}e^{inx}, \qquad e^{-int}e^{inx}$$

are classical solutions of the wave equation, and that they are 2π-periodic in the x variable. It follows that for any choice of $\hat{a}_n, \hat{b}_n \in \mathbb{C}$ the linear combination

$$u(x,t) = \sum_{n=-N}^{N} \hat{a}_n e^{int}e^{inx} + \hat{b}_n e^{-int}e^{inx} \tag{5}$$

is again a solution of the wave equation. If we have infinitely many coefficients $\hat{a}_n, \hat{b}_n$ then we could try to show that the series converges and that its limit is a weak or classical solution of the wave equation.

Sometimes it is convenient to rewrite (5) by applying Euler's formula $e^{\pm int} = \cos nt \pm i \sin nt$, with result

$$u(x,t) = \sum_{n=-N}^{N} \hat{u}_n \cos(nt)e^{inx} + \hat{v}_n \sin(nt)e^{inx} \tag{6}$$

where $\hat{u}_n = \hat{a}_n + \hat{b}_n$ and $\hat{v}_n = i(\hat{a}_n - \hat{b}_n)$.

2.3. The initial value problem. Suppose we want to find a solution to

$$\begin{aligned} &u_{tt} = u_{xx}, \text{ for all } x, t \\ &u(x + 2\pi, t) = u(x,t), \text{ for all } x, t \\ &u(x,0) = f(x), \quad u_t(x,0) = g(x) \text{ for all } x \end{aligned}$$

by looking for a function $u(x,t)$ that is given by (6). The function already satisfies the PDE, and it is 2π-periodic in the x variable, so we only have to choose the coefficients $\hat{u}_n, \hat{v}_n$ so that u satisfies the initial conditions. It follows from (6) that

$$f(x) = u(x,0) = \sum_n \hat{u}_n e^{inx}, \qquad g(x) = u_t(x,0) = \sum_n -n\hat{v}_n e^{inx}.$$

If f and g are trigonometric polynomials given by

$$f(x) = \sum_{n=-N}^{N} \hat{f}_n e^{inx}, \qquad g(x) = \sum_{n=-N}^{N} \hat{g}_n e^{inx},$$

then we should choose $\hat{u}_n, \hat{v}_n$ so that

$$\hat{u}_n = \hat{f}_n, \qquad \hat{v}_n = \frac{\hat{g}_n}{n}.$$

Thus the solution is

$$u(x,t) = \sum_{n=-N}^{N} \left\{ \hat{f}_n \cos(nt) + \frac{\hat{g}_n}{n} \sin(nt) \right\} e^{inx}. \tag{7}$$

2.4. How to find the coefficients $\hat{f}_n, \hat{g}_n$ if we know f and g. If $f(x) = \sum_{-N}^{N} \hat{f}_n e^{inx}$ then Fourier multiplied the equation with e^{-ikx} and integrated from 0 to 2π:

$$\int_0^{2\pi} f(x)e^{-ikx} dx = \int_0^{2\pi} \sum_{n=-N}^{N} \hat{f}_n e^{i(n-k)x} dx = \sum_{n=-N}^{N} \hat{f}_n \int_0^{2\pi} e^{i(n-k)x} dx = 2\pi \hat{f}_k,$$

because when $k \neq n$

$$\int_0^{2\pi} e^{i(n-k)x} dx = \left[\frac{e^{i(n-k)x}}{i(n-k)} \right]_{x=0}^{\pi} = 0$$

while when $k = n$ we get

$$\int_0^{2\pi} e^{i(n-k)x} dx = \int_0^{2\pi} 1 \, dx = 2\pi.$$

Thus, if we know the function f then its coefficients are given by

$$\hat{f}_k = \frac{1}{2\pi} \int_0^{2\pi} f(x)e^{-ikx} dx \tag{8}$$

2.5. Including infinitely many terms. If we completely ignore questions about convergence, then we could let $N \to \infty$ and claim that if the initial functions f and g are given by

$$f(x) = \sum_{-\infty}^{\infty} \hat{f}_n e^{inx}, \qquad g(x) = \sum_{-\infty}^{\infty} \hat{g}_n e^{inx}, \tag{9}$$

then the 2π-periodic solution to the wave equation with initial values $u(x,0) = f(x)$ and $u_t(x,0) = g(x)$ is given by

$$u(x,t) = \sum_{-\infty}^{\infty} \{ \hat{f}_n \cos(nt) + \frac{\hat{g}_n}{n} \sin(nt) \} e^{inx} \tag{10}$$

So if you can write the initial function and time derivative as a Fourier series (9) then we have a solution given in (10). This leaves us with a few questions

- for which 2π-periodic functions f and g can we find a Fourier expansion (9)?
- in what sense do the series (9) and (10) converge?

Fourier's remarkable answer to the first question was: every 2π-periodic function f has a Fourier expansion and the coefficients can be found by computing the integrals in (8).

For example, the "sawtooth function" is given by

$$f(x) = \frac{\pi - x}{2} \text{ for } 0 < x < 2\pi$$

Fourier computed the coefficients in its expansion and claimed

$$f(x) = \sum_{n=1}^{\infty} \frac{\sin nx}{n} = \sin x + \tfrac{1}{2}\sin 2x + \tfrac{1}{3}\sin 3x + \cdots$$

Wikipedia[1] offers a one paragraph history of Fourier's claims.

3. Digression: convergence of Fourier series

There is not enough space in this aside on Fourier series to do the topic justice. If you want to read more, the book FOURIER ANALYSIS by Stein and Shakarchi (Princeton lectures in analysis) offers a good introduction. Alternatively, T.W.Körner's book FOURIER ANALYSIS (yes, same title) from Cambridge University press is very readable.

3.1. Questions: was Fourier right? Let f be a 2π-periodic function. Does the Fourier series for f converge, and does it converge to f? To be more precise, define the partial sums of the series

$$S_N f(x) = \sum_{-N}^{N} \hat{f}_n e^{inx}, \qquad \hat{f}_n = \frac{1}{2\pi}\int_0^{2\pi} e^{-inx} f(x)\, dx.$$

Under what conditions can we guarantee that $\lim_{N\to\infty} S_N f(x) = f(x)$? Is the converge uniform? What other notions of convergence might apply here?

3.2. The Dirichlet kernel. For any 2π-periodic function f the partial sums of the Fourier series of f are given by

$$S_N f(x) = \int_0^{2\pi} D_N(x-\xi) f(\xi)\, d\xi, \quad \text{where} \quad D_N(x) = \frac{1}{2\pi}\frac{\sin(N+\frac{1}{2})x}{\sin\frac{x}{2}}$$

The function D_N is called the Dirichlet kernel. It satisfies

$$\int_0^{2\pi} D_N(x-\xi)\, d\xi = 1 \text{ for all } x \in \mathbb{R}.$$

PROOF.

$$S_N f(x) = \frac{1}{2\pi}\sum_{n=-N}^{N}\int_0^{2\pi} e^{in(x-\xi)} f(\xi)\, d\xi \tag{11}$$

$$= \frac{1}{2\pi}\int_0^{2\pi}\sum_{n=-N}^{N} e^{in(x-\xi)} f(\xi)\, d\xi = \int_0^{2\pi} D_N(x-\xi) f(\xi)\, d\xi$$

[1] https://en.wikipedia.org/wiki/JosephFourier#TheAnalyticTheoryofHeat

where

$$D_N(x) = \frac{1}{2\pi} \sum_{-N}^{N} e^{inx}$$

We can compute by using the formula for geometric sums:

$$\begin{aligned}
\sum_{-N}^{N} e^{inx} &= e^{-iNx} + e^{-i(N-1)x} + \cdots + e^{i(N-1)x} + e^{iNx} \\
&= e^{-iNx}\left\{1 + e^{ix} + e^{2ix} + \cdots + e^{2iNx}\right\} \\
&= e^{-iNx} \frac{e^{(2N+1)ix} - 1}{e^{ix} - 1} \\
&= \frac{e^{i(N+\frac{1}{2})x} - e^{-i(N+\frac{1}{2})x}}{e^{ix/2} - e^{-ix/2}} \\
&= \frac{\sin(N + \frac{1}{2})x}{\sin \frac{x}{2}}
\end{aligned}$$

Q.E.D.

3.3. Riemann-Lebesgue Lemma. If $f : [a, b] \to \mathbb{R}$ is Riemann integrable, then

$$\lim_{\lambda\to\infty} \int_a^b \sin(\lambda x) f(x)\, dx = \lim_{\lambda\to\infty} \int_a^b \cos(\lambda x) f(x)\, dx = 0.$$

PROOF. We show that $\int \sin(\lambda x) f(x) dx \to 0$, the proof that $\int \cos(\lambda x) f(x) dx \to 0$ being nearly identical.

Let $\epsilon > 0$ be given. Then, by definition of Riemann integrability of the function f, there exists a partition $a = x_0 < x_1 < x_2 < \cdots < x_{m-1} < x_m = b$, and numbers $m_i < M_i$ such that * $m_k \leq f(x) \leq M_k$ for all $x \in [x_{k-1}, x_k]$, and all $k = 1, \ldots, m$ * $\sum_{k=1}^{m}(M_k - m_k)(x_k - x_{k-1}) < \epsilon/\big(2(b - a)\big)$

Define the step function

$$s(x) = m_k \text{ for all } x \in (x_{k-1}, x_k] \text{ and for all } k = 1, 2, \ldots, m$$

Then $s(x) \leq f(x)$ and $0 \leq \int_a^b (f(x) - s(x))\, dx < \epsilon/\big(2(b - a)\big)$.

To show that $\int_a^b \sin(\lambda x) f(x)\, dx \to 0$ we rewrite the integral as follows

$$\int_a^b \sin(\lambda x) f(x)\, dx = \int_a^b \sin(\lambda x) s(x)\, dx + \int_a^b \sin(\lambda x)\big(f(x) - s(x)\big)\, dx = A + B.$$

We can compute the first term A explicitly:

$$\begin{aligned}
\text{(12)} \quad A = \int_a^b \sin(\lambda x) s(x)\, dx &= \sum_{k=1}^{m} \int_{x_{k-1}}^{x_k} \sin(\lambda x) m_k dx \\
&= \sum_{k=1}^{m} m_k \left[\frac{-\cos(\lambda x)}{\lambda}\right]_{x_{k-1}}^{x_k} = \frac{1}{\lambda} \sum_{k=1}^{m} m_k \left[\cos \lambda x_{k-1} - \cos \lambda x_k\right]
\end{aligned}$$

which implies

$$|A| \leq \frac{2}{\lambda} \sum_{k=1}^{m} |m_k|.$$

Therefore, if we define $\lambda_\epsilon = \dfrac{\epsilon}{4\sum m_k}$ then $|A| < \epsilon/2$ holds for all $\lambda > \lambda_\epsilon$.

Next we show that B is also small. No matter what λ is, we always have

$$(13) \quad |B| = \left| \int_a^b \sin(\lambda x)\big(f(x) - s(x)\big)\, dx \right|$$

$$\le \int_a^b |\sin(\lambda x)|\, \big|f(x) - s(x)\big|\, dx \le (b-a)\cdot \frac{\epsilon}{2(b-a)} = \frac{\epsilon}{2}.$$

Therefore we get $|A + B| \le |A| + |B| < \epsilon/2 + \epsilon/2 = \epsilon$ for all $\lambda > \lambda_\epsilon$. *Q.E.D.*

3.4. Definition. A function $f : \mathbb{R} \to \mathbb{R}$ is called Lipschitz continuous at $x \in \mathbb{R}$ if there exists a $C > 0$ such that

$$|f(\xi) - f(x)| \le C|x - \xi| \text{ for all } \xi \in \mathbb{R}.$$

3.5. Convergence of a Fourier series at a point of Lipschitz continuity. Assume that $f : \mathbb{R} \to \mathbb{R}$ is 2π-periodic, Riemann integrable, and that f is Lipschitz continuous at a. Then

$$f(a) = \sum_{k\in\mathbb{Z}} \hat{f}_k e^{ika}, \text{ i.e. } \lim_{N\to\infty} \sum_{k=-N}^{N} \hat{f}_k e^{ika} = f(a)$$

Proof. It follows from $\int_0^{2\pi} D_N(\xi - a)d\xi = 1$ that $f(a) = \int_0^{2\pi} D_N(\xi - a)f(a)d\xi$. Therefore

$$S_N f(a) - f(a) = \int_0^{2\pi} D_N(\xi - a)\big(f(\xi) - f(a)\big)\, d\xi.$$

Substitute $\xi = a + s$, and use the definition of the Dirichlet kernel:

$$\begin{aligned} S_N f(a) - f(a) &= \int_{-\pi}^{\pi} D_N(s)\big(f(a+s) - f(a)\big)\, ds \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \sin\big((N+\tfrac12)s\big)\, \frac{f(a+s) - f(a)}{\sin s/2}\, ds \\ &= \int_{-\pi}^{\pi} \sin\big((N+\tfrac12)s\big)\, g(s)\, ds \end{aligned}$$

where

$$g(s) \stackrel{\text{def}}{=} \frac{f(a+s) - f(a)}{2\pi \sin s/2}.$$

This function is bounded because

$$|g(s)| \le \frac{C|s|}{2\pi|\sin(s/2)|} \le \frac{C}{2} \text{ for all } s \in (-\pi, \pi), s \ne 0.$$

Here we have used that $\sin x \ge \frac{\pi}{2}x$ for all $x \in [0, \frac{\pi}{2}]$. using the boundedness of g and the fact that f is Riemann integrable, it is now a somewhat lengthy exercise in real analysis to show that g is also Riemann integrable on $[-\pi, \pi]$. We can therefore invoke the Riemann-Lebesgue lemma and conclude that there is a λ_ϵ such that for all $\lambda > \lambda_\epsilon$ one has

$$\left| \int_{-\pi}^{\pi} \sin(\lambda s) g(s) ds \right| < \epsilon.$$

and therefore $|S_N f(a) - f(a)| < \epsilon$ for all $\lambda > \lambda_\epsilon$, i.e. for all $N > \lambda_\epsilon - \frac12$. *Q.E.D.*

3.6. Theorem. If f is a 2π-periodic C^2 function then its Fourier coefficients $\hat{f}_n$ satisfy

$$|\hat{f}_n| \leq \frac{\|f''\|_\infty}{n^2} \qquad \text{for all } n \neq 0.$$

The Fourier series $f(x) = \sum_{-\infty}^{\infty} \hat{f}_n e^{inx}$ converges uniformly, and absolutely.

Proof. Integration by parts in the definition of $\hat{f}_n$ gives us

$$\begin{aligned}(14)\quad 2\pi \hat{f}_n &= \int_0^{2\pi} e^{-inx} f(x)\, dx \\ &= \left[\frac{e^{-inx}}{-in} f(x)\right]_0^{2\pi} - \frac{1}{-in}\int_0^{2\pi} e^{-inx} f'(x)\, dx = \frac{1}{in}\int_0^{2\pi} e^{-inx} f'(x)\, dx.\end{aligned}$$

Integrating by parts again we get

$$2\pi \hat{f}_n = \frac{-1}{n^2}\int_0^{2\pi} e^{-inx} f''(x)\, dx.$$

This implies

$$\begin{aligned}(15)\quad 2\pi|\hat{f}_n| &= \left|\frac{-1}{n^2}\int_0^{2\pi} e^{-inx} f''(x)\, dx\right| \leq \frac{1}{n^2}\int_0^{2\pi} \left|e^{-inx} f''(x)\right|\, dx \\ &\leq \frac{1}{n^2}\int_0^{2\pi} \|f''\|_\infty dx = 2\pi \frac{\|f''\|_\infty}{n^2}.\end{aligned}$$

The terms in the Fourier series for f are bounded by

$$\left|\hat{f}_n e^{inx}\right| = |\hat{f}_n| \leq \frac{\|f''\|_\infty}{n^2} \qquad (n \neq 0).$$

Therefore, by the Weierstrass M-test (see Section 4.2 in the analysis appendix), the series $\sum_{-\infty}^{\infty} \hat{f}_n e^{inx}$ converges uniformly, and absolutely.

Since f is C^1 it is also Hölder continuous at all $x \in \mathbb{R}$. Therefore Theorem 3.5 implies that $\sum_{-\infty}^{\infty} \hat{f}_n e^{inx} = f(x)$ for all $x \in \mathbb{R}$. *Q.E.D.*

4. Fourier series and the L^2 Inner Product

4.1. An inner product. Let $\mathcal{R}_{\text{per}}$ be the set of Riemann integrable 2π-periodic functions, and let $\mathcal{C}_{\text{per}}$ be the subset of continuous 2π-periodic functions. For any two functions $f, g \in \mathcal{R}_{\text{per}}$ we define

$$\langle f, g\rangle = \int_0^{2\pi} f(x)\overline{g(x)}\, dx\,.$$

Then $\langle \dots \rangle$ has the following properties:

- Symmetry: $\langle f, g\rangle = \overline{\langle g, f\rangle}$
- Bilinear: $\langle f + g, h\rangle = \langle f, h\rangle + \langle g, h\rangle$ for all $f, g, h \in \mathcal{R}_{\text{per}}$
- Bilinear: $\langle cf, g\rangle = c\langle f, g\rangle$ and $\langle f, cg\rangle = \bar{c}\langle f, g\rangle$ for all $f, g \in \mathcal{R}_{\text{per}}$ and $c \in \mathbb{C}$.
- Nonnegative: $\langle f, f\rangle \geq 0$ for all $f \in \mathcal{R}_{\text{per}}$

4.2. Theorem. If $f \in \mathcal{C}_{\text{per}}$ and $\langle f, f\rangle = 0$ then $f = 0$. On the other hand there exist functions $f \in \mathcal{R}_{\text{per}}$ with $f \neq 0$ and $\langle f, f\rangle = 0$.

Proof. Arguing by contradiction we suppose $\langle f, f\rangle = 0$ and $f(x_0) \neq 0$ for some $x_0 \in [0, 2\pi)$. Since f is continuous, there is an $\delta > 0$ such that $|f(x)| \geq \frac{1}{2}|f(x_0)|$ for all $x \in (x_0 - \delta, x_0 + \delta)$. This implies

$$0 = \langle f, f\rangle = \int_0^{2\pi} |f(x)|^2 dx \geq \int_{x_0-\delta}^{x_0+\delta} |f(x)|^2 dx \geq \left(\tfrac{1}{2}|f(x_0)|\right)^2 \cdot (2\delta) > 0.$$

This is a contradiction, so we conclude that $f(x_0) = 0$ for all x_0.

To prove the second part of the theorem, consider the function f for which $f(x) = 0$ at all $x \in \mathbb{R}$ except $x = n\pi$ ($n \in \mathbb{Z}$) where $f(n\pi) = 1$. Then f is Riemann integrable and 2π-periodic, and $\langle f, f\rangle = 0$ even though $f \neq 0$. *Q.E.D.*

4.3. Theorem. $\mathcal{C}_{\text{per}}$ with $\langle f, g\rangle$ defined as above is a complex vector space with an inner product.

See [the analysis appendix](AppendixAnalysis.md) for a review of complex inner products.

We define the L^2 norm of $f \in \mathcal{R}_{\text{per}}$ to be $\|f\|_2$ where

$$\|f\|_2^2 = \langle f, f\rangle = \int_0^{2\pi} |f(x)|^2 \, dx.$$

4.4. The L^2 norm and the supremum norm. The L^2 norm and the supremum norm $\|f\|_\infty = \sup_x |f(x)|$ are generally different. They always satisfy

$$\|f\|_2 \leq \sqrt{2\pi}\|f\|_\infty \text{ for any } f \in \mathcal{R}_{\text{per}}.$$

This follows from the following short computation

$$\|f\|_2^2 = \int_0^{2\pi} |f(x)|^2 dx \leq \int_0^{2\pi} \|f\|_\infty^2 dx = 2\pi\|f\|_\infty^2.$$

4.5. Fourier coefficients as inner products. Let $e_k(x) = e^{ikx}$. Then for any $k \in \mathbb{Z}$ and any $l \neq k$ one has

$$\|e_k\|^2 = 2\pi, \text{ and } \langle e_k, e_l\rangle = 0, \text{ i.e. } e_k \perp e_l.$$

If $f = \hat{f}_{-N}e_{-N} + \cdots + \hat{f}_N e_N$ then we have

$$\langle f, e_k\rangle = \hat{f}_k\|e_k\|^2 = 2\pi\hat{f}_k.$$

The k^{th} Fourier coefficient of $f \in \mathcal{R}_{\text{per}}$ is given by

$$\hat{f}_k = \frac{1}{2\pi}\langle f, e_k\rangle$$

The N^{th} partial sum of the Fourier series of $f \in \mathcal{R}_{\text{per}}$ is given by

$$S_N f = \frac{1}{2\pi}\sum_{k=-N}^{N} \langle f, e_k\rangle e_k$$

4.6. Bessel's inequality. For any $f \in \mathcal{R}_{\text{per}}$ we have

$$\|S_N f\|_2^2 + \|f - S_N f\|_2^2 = \|f\|_2^2.$$

In particular $\|S_N f\|_2 \leq \|f\|_2$ always holds.

Proof. We begin by showing that $S_N f \perp (f - S_N f)$. For any integer k with $|k| \leq N$ we have

$$\begin{aligned}
\langle e_k, f - S_N f\rangle &= \langle e_k, f\rangle - \frac{1}{2\pi}\sum_{l=-N}^{n}\langle e_k, \langle f, e_l\rangle e_l\rangle \\
&= \langle e_k, f\rangle - \frac{1}{2\pi}\sum_{l=-N}^{n}\overline{\langle f, e_l\rangle}\langle e_k, e_l\rangle \\
&= \langle e_k, f\rangle - \frac{1}{2\pi}\sum_{l=-N}^{n}\langle e_l, f\rangle\langle e_k, e_l\rangle && \langle e_k, e_l\rangle = 0 \text{ if } k \neq l \\
&= \langle e_k, f\rangle - \frac{1}{2\pi}\langle e_k, f\rangle\langle e_k, e_k\rangle && \langle e_k, e_k\rangle = 2\pi \\
&= 0.
\end{aligned}$$

Thus $e_k \perp f - S_N f$ for all $k = -N, \ldots, N$. Since $S_N f$ is a linear combination of $e_{-N}, \ldots, e_N$ it follows that $S_N f \perp f - S_N f$.

Bessel's inequality now follows from Pythagoras: since $S_N f \perp f - S_N f$ we have

$$\|f\|_2^2 = \|f - S_N f + S_N f\|_2^2 = \|f - S_N f\|_2^2 + \|S_N f\|_2^2.$$

Q.E.D.

4.7. Partial sums as best approximations. If $g = \sum_{-N}^{N} \hat{g}_k e_k$ is a trigonometric polynomial of degree at most N then

$$\|f - S_N f\|_2 \leq \|f - g\|_2,$$

with strict inequality if $g \neq S_N f$.

Proof. First note that $S_N f - g$ is a linear combination of $e_{-N}, \ldots, e_N$, and therefore that $S_N f - g \perp S_N - f$. Then apply Pythagoras to $f - g = f - S_N f + S_N f - g$, to get

$$\|f - g\|_2^2 = \|f - S_N f\|_2^2 + \|S_N f - g\|_2^2.$$

This implies $\|f - S_N f\|_2 \leq \|f - g\|_2$, and if $\|f - S_N f\|_2 = \|f - g\|_2$ then $\|S_N f - g\|_2 = 0$. Since $S_N f - g$ is a trigonometric polynomial it is a continuous function. Therefore $\|S_N f - g\|_2 = 0$ implies that $S_f - g = 0$, i.e. $S_N f = g$. *Q.E.D.*

4.8. Convergence in L^2 of Fourier series. Let $f \in \mathcal{R}_{\text{per}}$. Then

$$\lim_{N\to\infty}\|S_N f - f\|_2 = 0$$

and

$$\|f\|_2^2 = \frac{1}{2\pi}\sum_{n\in\mathbb{Z}}|\hat{f}_n|^2.$$

This statement is known as the Plancherel or Parseval identity.

Proof. Let $\epsilon > 0$ be given. Since f is Riemann integrable, a C^2 function g exists that is 2π-periodic, and that satisfies

$$\int_0^{2\pi} |f(x) - g(x)|^2 dx \le \frac{\epsilon^2}{4}, \text{ i.e. } \|f - g\|_2 < \frac{\epsilon}{2}.$$

(details in lecture). Since g is C^2, we know that the Fourier series of g converges uniformly to g, i.e.

$$\|S_N g - g\|_\infty \to 0.$$

For any given $\epsilon_2 > 0$ we can therefore find an N_ϵ such that $||S_N g - g||_\infty < \frac{\epsilon}{2\sqrt{2\pi}}$ for all $N \ge N_{\epsilon_2}$. This implies

$$\|S_N g - g\|_2 \le \sqrt{2\pi}\, \|S_N g - g\|_\infty < \frac{\epsilon}{2}.$$

We now have for all $N \ge N_\epsilon$

$$\begin{aligned} \|S_N f - f\|_2 &= \|S_N f - S_N g + S_N g - g\|_2 \\ &\le \|S_N f - S_N g\|_2 + \|S_N g - g\|_2 \\ &= \|S_N (f - g)\|_2 + \|S_N g - g\|_2 \\ &\le \|f - g\|_2 + \|S_N g - g\|_2 \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

The Parseval–Plancherel identity now follows by letting $N \to \infty$ in $\|f\|_2^2 = \|S_N f\|_2^2 + \|f - S_N f\|_2^2$. *Q.E.D.*

5. Problems

5.1. Consider the solutions $u : \mathbb{R}^2 \to \mathbb{R}$ to the wave equation described in (5). Find functions $F, G : \mathbb{R} \to \mathbb{R}$ such that $u(x, t) = F(x + t) + G(x - t)$ for all x, t. (Hint: don't use the derivation of d'Alembert's solution but instead use $e^{i\theta} e^{i\phi} = e^{i(\theta+\phi)}$ and take a good look at (5)).

Solution. We have

$$\begin{aligned} u(x,t) &= \sum_{n=-N}^{N} \hat{a}_n e^{int} e^{inx} + \hat{b}_n e^{-int} e^{inx} \\ &= \sum_{n=-N}^{N} \hat{a}_n e^{in(x+t)} + \sum_{n=-N}^{N} \hat{b}_n e^{in(x-t)} \\ &= F(x+t) + G(x-t) \end{aligned}$$

provided we define

$$F(x) = \sum_{n=-N}^{N} \hat{a}_n e^{inx}, \qquad G(x) = \sum_{n=-N}^{N} \hat{b}_n e^{inx}.$$

5.2. Various questions about the Dirichlet kernel and Fourier coefficients.

(a) Show that $\int_0^{2\pi} D_N(s)\,ds = 1$.
(b) The k^{th} Fourier coefficient $\hat{f}_k$ of a 2π-periodic function $f : \mathbb{R} \to \mathbb{C}$ is defined as an integral from $x = 0$ to $x = 2\pi$. Show that $\hat{f}_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-ikx} f(x)\,dx$.

Solution. For (a) we compute:

$$\int_0^{2\pi} D_N(s)ds = \int_0^{2\pi} \frac{1}{2\pi}\sum_{-N}^{N} e^{ins}ds = \sum_{-N}^{N} \frac{1}{2\pi}\int_0^{2\pi} e^{ins}ds.$$

If $n \neq 0$ then $\int_0^{2\pi} e^{inx}dx = 0$, so the only non zero term in this sum is the one with $n = 0$:

$$\int_0^{2\pi} D_N(s)ds = \frac{1}{2\pi}\int_0^{2\pi} 1ds = 1.$$

(b) follows from the fact that $f(x)e^{-inx}$ is periodic with period 2π so that $\int_a^b f(x)e^{-inx}dx = \int_{2\pi+a}^{2\pi+b} f(x)e^{-inx}dx$ for any $a, b \in \mathbb{R}$ and in particular,

$$\int_{-\pi}^{0} f(x)e^{-inx}dx = \int_{\pi}^{2\pi} f(x)e^{-inx}dx.$$

Hence

$$\begin{aligned}\int_{-\pi}^{\pi} f(x)e^{-inx}dx &= \int_{-\pi}^{0} f(x)e^{-inx}dx + \int_0^{\pi} f(x)e^{-inx}dx \\ &= \int_{\pi}^{2\pi} f(x)e^{-inx}dx + \int_0^{\pi} f(x)e^{-inx}dx = \int_0^{2\pi} f(x)e^{-inx}dx.\end{aligned}$$

5.3. Let $f(x) = |\sin\frac{x}{2}|$.

(a) Compute the Fourier series of f (the integrals simplify if you write $\sin\frac{x}{2}$ in terms of complex exponentials.)
(b) For which $x \in \mathbb{R}$ does the Fourier series converge to f?
(c) What do you get if you set $x = 0$ in the Fourier expansion of f?

Solution. (a) The Fourier coefficients are given by $\hat{f}_k = \frac{1}{2\pi}\int_0^{2\pi}|\sin\frac{x}{2}|dx$. If $0 \le x \le 2\pi$ then $0 \le \frac{x}{2} \le \pi$ and thus $\sin\frac{x}{2} \ge 0$. It follows that in the integral we have to compute, we have $|\sin\frac{x}{2}| = \sin\frac{x}{2}$. We now compute

$$\begin{aligned}2\pi\hat{f}_k = \int_0^{2\pi}\sin\frac{x}{2}\; e^{-ikx}dx &= \frac{1}{2i}\int_0^{2\pi}\left(e^{ix/2} - e^{-ix/2}\right)e^{-ikx}dx \\ &= \frac{1}{2i}\int_0^{2\pi}\left(e^{-i(k-\frac{1}{2})x} - e^{-i(k+\frac{1}{2})x}\right)dx \\ &= \frac{1}{2i}\left[\frac{e^{-i(k-\frac{1}{2})x}}{-i(k-\frac{1}{2})} - \frac{e^{-i(k+\frac{1}{2})x}}{-i(k+\frac{1}{2})}\right]_0^{2\pi} \\ &= \frac{1}{2}\left[\frac{e^{-(2k-1)i\pi}-1}{-k+\frac{1}{2}} - \frac{e^{-(2k+1)i\pi}-1}{k+\frac{1}{2}}\right] \quad \text{(use } i^2 = -1\text{)}.\end{aligned}$$

By Euler's definition of the complex exponential we have $e^{(2k\pm1)i\pi} = e^{2k\pi i}e^{\pm\pi i} = e^{\pm\pi i} = -1$, so we get

$$\hat{f}_k = \frac{1}{2\pi}\frac{1}{2}\left[-\frac{2}{k-\frac{1}{2}} + \frac{2}{k+\frac{1}{2}}\right] = \frac{1}{2\pi}\frac{-1}{(k-\frac{1}{2})(k+\frac{1}{2})} = -\frac{1}{2\pi}\frac{1}{k^2-\frac{1}{4}}.$$

The Fourier series for $f(x) = |\sin\frac{x}{2}|$ is thus

$$-\sum_{k=-\infty}^{\infty}\frac{1}{2\pi}\frac{e^{ikx}}{k^2-\frac{1}{4}}.$$

(b) The function $g(x) = \sin\frac{x}{2}$ satisfies the Lipschitz condition because, because of the Mean Value Theorem. Indeed, for any $a, x \in \mathbb{R}$ with $a \neq x$ there is a $c \in (a, x)$ (or $c \in (x, a)$ if $x < a$) such that

$$\frac{g(x)-g(a)}{x-a} = g'(c) = \frac{1}{2}\cos\frac{c}{2}.$$

This implies

$$\left|\frac{g(x)-g(a)}{x-a}\right| = \left|\frac{1}{2}\cos\frac{c}{2}\right| \leq \frac{1}{2},$$

and hence

$$|\sin\frac{x}{2} - \sin\frac{a}{2}| \leq \frac{1}{2}|x-a|.$$

To show that $f(x) = |g(x)|$ also satisfies the Lipschitz condition we use the triangle inequality $\big||a|-|b|\big| \leq |a-b|$:

$$|f(x)-f(a)| = ||g(x)|-|g(a)|| \leq |g(x)-g(a)| \leq \frac{1}{2}|x-a|.$$

This is true for any $a \in \mathbb{R}$ and therefore the Fourier series $S_N f(a)$ converges to $f(a)$:

$$\forall a \in \mathbb{R}: \quad \lim_{N\to\infty} S_N f(a) = f(a).$$

(c) We may choose $a = 0$ and conclude that since $f(0) = |\sin\frac{0}{2}| = 0$

$$\lim_{N\to\infty}\frac{1}{2\pi}\sum_{-N}^{N}\frac{e^{ik\cdot 0}}{k^2-\frac{1}{4}} = 0 \implies \lim_{N\to\infty}\sum_{-N}^{N}\frac{1}{k^2-\frac{1}{4}} = 0.$$

We can leave the result in this form, but we can also rewrite it by using $\frac{1}{k^2-\frac{1}{4}} = \frac{1}{(-k)^2-\frac{1}{4}}$, which implies

$$\sum_{-N}^{N}\frac{1}{k^2-\frac{1}{4}} = \sum_{-N}^{-1}\frac{1}{k^2-\frac{1}{4}} + \sum_{1}^{N}\frac{1}{k^2-\frac{1}{4}} + \frac{1}{-\frac{1}{4}} = 2\sum_{1}^{N}\frac{1}{k^2-\frac{1}{4}} - 4$$

Hence, in the limit $N \to \infty$ we find

$$\sum_{1}^{\infty}\frac{1}{k^2-\frac{1}{4}} = 2, \text{ i.e. } \frac{1}{1-\frac{1}{4}} + \frac{1}{4-\frac{1}{4}} + \frac{1}{9-\frac{1}{4}} + \cdots = 2.$$

5.4. The "plucked string" function. Let $\wp$ be the 2π-periodic function that satisfies $\wp(x) = x$ for $x \in [0, \pi]$, and $\wp(x) = 2\pi - x$ for $x \in [\pi, 2\pi]$

(a) Compute the Fourier series of $\wp$
(b) For which $x \in \mathbb{R}$ does the Fourier series converge to $\wp$?
(c) What do you get if you set $x = 0$ in the Fourier expansion of $\wp$?

5.5. Solving ordinary differential equations with Fourier series.

(a) Let $f : \mathbb{R} \to \mathbb{C}$ be a 2π-periodic function that is C^1. Show that if $g = f'$ then $\hat{g}_k = ik\hat{f}_k$.
(b) Let $f : \mathbb{R} \to \mathbb{C}$ be a 2π-periodic function that is C^2. Show that if $g = f''$ then $\hat{g}_k = -k^2\hat{f}_k$.
(c) Show that there exist numbers $m_k \in \mathbb{C}$ such that the following holds for any C^2 2π-periodic function $f : \mathbb{R} \to \mathbb{C}$: if $f'' - 4f = g$ then $\hat{f}_k = m_k\hat{g}_k$.
(d) Find the Fourier series of the 2π-periodic solution $f : \mathbb{R} \to \mathbb{C}$ of $f''(x) - 4f(x) = \sin(x)$ $(0 \le x \le 2\pi)$.
(e) Find the Fourier series of the 2π-periodic solution $f : \mathbb{R} \to \mathbb{C}$ of $f''(x) - 4f(x) = \wp(x)$ $(0 \le x \le 2\pi)$ where $\wp$ is the "plucked string" function from problem 5.4.

5.6. Let $f : [0, 2\pi] \to \mathbb{R}$ be Riemann integrable, and let $\hat{f}_k$ be the k^{th} Fourier coefficient of f.

(a) Show that $|\hat{f}_k| \le \|f\|_\infty$ for all $k \in \mathbb{Z}$.
(b) Show that $\lim_{k\to\pm\infty} \hat{f}_k = 0$.
(c) Suppose that f is m times continuously differentiable. Show that $|\hat{f}_k| \le \frac{\|f^{(m)}\|_\infty}{|k|^m}$ for all $k \neq 0$.

Solution.

(a) $|\hat{f}_k| = \left|\frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-ikx}dx\right| \le \frac{1}{2\pi}\int_0^{2\pi} |f(x)|dx \le \|f\|_\infty$

(b) $\hat{f}_k = \frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-ikx}dx = \frac{1}{2\pi}\int_0^{2\pi} f(x)\cos(kx)dx - \frac{i}{2\pi}\int_0^{2\pi} f(x)\sin(kx)dx$ by Euler's formula. The Riemann–Lebesgue lemma implies that both integrals tend to zero as $k \to \infty$.

5.7. Let $z : \mathbb{R} \to \mathbb{R}$ be the 2π-periodic "saw tooth function" given by $z(x) = \frac{\pi - x}{2}$ for $0 < x < 2\pi$. Compute $\|z\|_2$ and the Fourier coefficients of z. Which identity do you find by applying the Plancherel–Parseval identity?

CHAPTER 4

The Laplacian and where you will find it

1. Calculus of Variations in 1D

1.1. Example: Fermat's principle. Suppose a region of space with $a < x < b$ is laminated vertically, and that the speed of light at each point is given by $v(x) > 0$ for some positive C^∞ function v. If the path of a light ray entering at $x = a$ and exiting at $x = b$ is the graph of a function $f : [a, b] \to \mathbb{R}$, then the time it takes the light to traverse this path is

$$T[f] = \int_a^b \frac{ds}{v(x)} = \int_a^b \frac{\sqrt{1 + f'(x)^2}}{v(x)}\, dx$$

FERMAT's principle says that if a light ray travels from $A = (a, y_a)$ to $B = (b, y_b)$ along the graph of $y = f(x)$ then f minimizes the travel time $T[f]$: this means that for any other $h : [a, b] \to \mathbb{R}$ with $h(a) = y_a$ and $h(b) = y_b$ one has $T[f] \leqslant T[h]$.

Theorem. If $f : [a, b] \to \mathbb{R}$ minimizes the travel time $T[f]$ amongst all $f : [a, b] \to \mathbb{R}$ with $f(a) = y_a$, $f(b) = y_b$, and if f is C^2 then

$$\frac{d}{dx}\left\{\frac{f'(x)}{v(x)\sqrt{1 + f'(x)^2}}\right\} = 0 \text{ for all } x \in [a, b]. \tag{16}$$

This equation is called the Euler–Lagrange equation for the minimizer of the travel time $T[f]$.

PROOF. Suppose that a function $f : [a, b] \to \mathbb{R}$ represents the fastest path from $(a, f(a))$ to $(b, f(b))$. Then, for any function $g : [a, b] \to \mathbb{R}$ and all $\epsilon \in \mathbb{R}$ we have

$$T[f + \epsilon g] \geq T[f],$$

and hence

$$\frac{d}{d\epsilon} T[f + \epsilon g]\Big|_{\epsilon=0} = 0.$$

We compute this derivative as follows:

$$\begin{aligned}
\frac{d}{d\epsilon} T[f + \epsilon g] &= \frac{d}{d\epsilon} \int_a^b \sqrt{1 + (f_x + \epsilon g_x)^2}\, \frac{dx}{v(x)} \\
&= \int_a^b \frac{\partial}{\partial \epsilon}\left\{\frac{\sqrt{1 + (f_x + \epsilon g_x)^2}}{v(x)}\right\} dx \\
&= \int_a^b \frac{g_x(f_x + \epsilon g_x)}{\sqrt{1 + (f_x + \epsilon g_x)^2}} \frac{dx}{v(x)}.
\end{aligned}$$

Set $\epsilon = 0$ and you find

$$\frac{d}{d\epsilon} T[f + \epsilon g]\Big|_{\epsilon=0} = \int_a^b g_x \frac{f_x}{\sqrt{1 + f_x^2}} \frac{dx}{v(x)}.$$

Finally, we integrate by parts,

$$\frac{d}{d\epsilon} T[f+\epsilon g]\Big|_{\epsilon=0} = \left[g(x) \frac{f'(x)}{v(x)\sqrt{1 + f'(x)^2}} \right]_{x=a}^{b} - \int_a^b g(x) \frac{d}{dx} \left\{ \frac{f'(x)}{v(x)\sqrt{1 + f'(x)^2}} \right\} dx$$

and use $g(a) = g(b) = 0$ to get

$$\frac{d}{d\epsilon} T[f + \epsilon g]\Big|_{\epsilon=0} = - \int_a^b g(x) \frac{d}{dx} \left\{ \frac{f'(x)}{v(x)\sqrt{1 + f'(x)^2}} \right\} dx \tag{17}$$

This equation is called the first variation of the travel time function T.

Fermat's principle says that if the graph of f is a possible light ray, then $\frac{d}{d\epsilon} T[f + \epsilon g] = 0$ at $\epsilon = 0$, i.e.

$$- \int_a^b g(x) \frac{d}{dx} \left\{ \frac{f'(x)}{v(x)\sqrt{1 + f'(x)^2}} \right\} dx \text{ for all } g \text{ with } g(a) = g(b) = 0. \tag{18}$$

Using Lemma 1.2 below, we conclude

$$\frac{d}{dx} \left\{ \frac{f'(x)}{v(x)\sqrt{1 + f'(x)^2}} \right\} = 0.$$

Q.E.D.

1.2. Lemma. If $h : [a, b] \to \mathbb{R}$ is a continuous function with the property that

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

holds for all C^∞ functions $g : [a, b] \to \mathbb{R}$ with $g(a) = g(b) = 0$, then $h(x) = 0$ for all $x \in [a, b]$.

2. The Euler–Lagrange equations

Let $L : [a, b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be a C^2 function and consider the integral

$$J[u] \stackrel{\text{def}}{=} \int_a^b L\big(x, f(x), f'(x)\big)\, dx.$$

The function L is called the Lagrangian of the integral $J[f]$.

We will think of $L = L(x, u, v)$ as a function of the three variables (x, u, v). For example, if

$$L(x, u, v) = \sqrt{1 + v^2}$$

then

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

is the length of the graph of f.

Similarly, the Lagrangian for which $J[f]$ becomes the time-traveled in Fermat's principle is

$$L_{\text{Fermat}}(x,u,v) = \frac{1}{v(x)}\sqrt{1+v^2}. \tag{19}$$

2.1. Theorem. Let $L : [a,b]\times\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ be a C^2 function, and define the integral

$$J[u] = \int_a^b L(x,u(x),u'(x))dx.$$

If $u : [a,b]\to\mathbb{R}$ is a C^2 function that minimizes $J[u]$ amongst all C^2 functions $u : [a,b]\to\mathbb{R}$ with $u(a)=y_a$, $u(b)=y_b$, then u satisfies

$$\frac{d}{dx}\left(\frac{\partial L}{\partial v}(x,u(x),u'(x))\right) - \frac{\partial L}{\partial u}(x,u(x),u'(x)) = 0. \tag{20}$$

This equation is called the Euler–Lagrange equation for the problem of minimizing $J[u]$ with prescribed boundary values $u(a), u(b)$.

In general the Euler–Lagrange equation is a second order differential equation for the unknown function $u : [a,b]\to\mathbb{R}$.

Proof. Let $g : [a,b]\to\mathbb{R}$ be a C^2 function with $g(a)=g(b)=0$. Then, because u minimizes the integral $J[u]$ we have $J[u+\epsilon g]\geq J[u]$ for all $\epsilon\in\mathbb{R}$. It follows that

$$\left.\frac{dJ[u+\epsilon g]}{d\epsilon}\right|_{\epsilon=0} = 0.$$

We now compute this derivative:

$$\begin{aligned}\frac{dJ[u+\epsilon g]}{d\epsilon} &= \frac{d}{d\epsilon}\int_a^b L(x,u(x)+\epsilon g(x),u'(x)+\epsilon g'(x))\,dx\\ &= \int_a^b \frac{\partial L(x,u(x)+\epsilon g(x),u'(x)+\epsilon g'(x))}{\partial\epsilon}\,dx\\ &= \int_a^b \left\{\frac{\partial L}{\partial u}(\cdots)g(x) + \frac{\partial L}{\partial v}(\cdots)g'(x)\right\}dx\end{aligned}$$

where $(\cdots) = (x,u(x)+\epsilon g(x),u'(x)+\epsilon g'(x))$. If we set $\epsilon=0$ then we get

$$\left.\frac{dJ[u+\epsilon g]}{d\epsilon}\right|_{\epsilon=0} = \int_a^b\left\{\frac{\partial L}{\partial u}(x,u(x),u'(x))\cdot g(x) + \frac{\partial L}{\partial v}(x,u(x),u'(x))\cdot g'(x)\right\}dx.$$

Integrate by parts in the second term:

$$\begin{aligned}\int_a^b \frac{\partial L}{\partial v}(x,u(x),u'(x))g(x)dx &= \left[\frac{\partial L}{\partial v}(x,u(x),u'(x))g'(x)\right]_{x=a}^b\\ &\quad - \int_a^b \frac{d}{dx}\left(\frac{\partial L}{\partial v}(x,u(x),u'(x))\right)\cdot g(x)\,dx\end{aligned}$$

The boundary terms vanish because $g(a)=g(b)=0$. Therefore we have

$$\int_a^b \frac{\partial L}{\partial v}(x,u(x),u'(x))g(x)dx = -\int_a^b \frac{d}{dx}\left(\frac{\partial L}{\partial v}(x,u(x),u'(x))\right)\cdot g(x)\,dx$$

and hence

$$\frac{dJ[u+\epsilon g]}{d\epsilon}\Big|_{\epsilon=0} = \int_a^b \Big\{\frac{\partial L}{\partial u}(x,u(x),u'(x)) - \frac{d}{dx}\frac{\partial L}{\partial v}(x,u(x),u'(x))\Big\}g(x)dx.$$

Our assumption that u minimizes $J[u]$ implies that $\frac{dJ[u+\epsilon g]}{d\epsilon}\Big|_{\epsilon=0}$ vanishes for any C^2 function $g:[a,b]\to\mathbb{R}$ with $g(a)=g(b)=0$. Thus

$$\int_a^b \Big\{\frac{\partial L}{\partial u}(x,u(x),u'(x)) - \frac{d}{dx}\frac{\partial L}{\partial v}(x,u(x),u'(x))\Big\}g(x)dx = 0$$

for all g with $g(a)=g(b)=0$.

We have also assumed that $u:[a,b]\to\mathbb{R}$ and the Lagrangian $L:[a,b]\times\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ are both twice differentiable functions. This implies that

$$\frac{\partial L}{\partial u}(x,u(x),u'(x)) - \frac{d}{dx}\frac{\partial L}{\partial v}(x,u(x),u'(x))$$

is a continuous function of $x\in[a,b]$. Using Lemma 1.2 again we conclude that

$$\frac{\partial L}{\partial u}(x,u(x),u'(x)) - \frac{d}{dx}\frac{\partial L}{\partial v}(x,u(x),u'(x)) = 0$$

for all $x\in[a,b]$. *Q.E.D.*

2.2. What did we just prove? The theorem in the previous section tells that

(a) **if** there is a minimizer u for $J[u]$, and
(b) **if** the minimizer is a C^2 function

then the minimizer u satisfies the Euler–Lagrange equation.

The same theorem does not claim any of the following statements:

(a) if $u:[a,b]\to\mathbb{R}$ is C^2 and satisfies the Euler–Lagrange equation, then u is a minimizer for $J[u]$
(b) every minimizer is C^2
(c) there always is a minimizer for $J[u]=\int_a^b L(x,u,u')dx$, no matter what L is.

2.3. Solving the Euler–Lagrange equation. In general the Euler–Lagrange equation is a second order differential equation and finding solutions using calculus methods may not be possible, but there are some special cases where the equation simplifies. One of those is where $L(x,u,v)$ does not depend on u (as in the Fermat principle case, see (19) and (16)). In that case the Euler-Lagrange equation

$$\frac{d}{dx}\left(\frac{\partial L}{\partial v}(x,u'(x))\right) = 0$$

from which we immediately see that

$$\frac{\partial L}{\partial v}(x,u'(x)) = C_1$$

for some constant C_1. Instead of a second order equation we now have a first order equation which we may be able to solve.

3. Problems

3.1. The Euler–Lagrange equation. Find the Lagrangian, determine the Euler–Lagrange equation (but don't solve it) for each of the following minimization problems. Be on the lookout for strange answers when you see "☡".

(a) $J[u] = \int_a^b \frac{1}{2} u'(x)^2 \, dx$ **Solution.**

$$
\begin{aligned}
L(x,u,v) &= \frac{1}{2} v^2 \\
L_u(x,u,v) &= 0 \implies L_u(x,u(x),u'(x)) = 0 \\
L_v(x,u,v) &= v \implies L_v(x,u(x),u'(x)) = u'(x)
\end{aligned}
$$

$$\text{EL-eqn:} \quad \frac{d}{dx}(u'(x)) - 0 = 0, \text{ i.e. } u''(x) = 0$$

(b) $J[u] = \int_a^b \left\{ \frac{1}{2} u'(x)^2 + u(x) \right\} dx$ **Solution.**

$$
\begin{aligned}
L(x,u,v) &= \frac{1}{2} v^2 + u \\
L_u(x,u,v) &= 1 \implies L_u(x,u(x),u'(x)) = 1 \\
L_v(x,u,v) &= v \implies L_v(x,u(x),u'(x)) = u'(x)
\end{aligned}
$$

$$\text{EL-eqn:} \quad \frac{d}{dx}(u'(x)) - 1 = 0, \text{ i.e. } u''(x) = 1$$

(c) $J[u] = \int_0^1 \left\{ \sqrt{1 + u'(x)^2} + u(x) \right\} dx$ **Solution.**

$$
\begin{aligned}
L(x,u,v) &= \frac{1}{2} \sqrt{1+v^2} + u \\
L_u(x,u,v) &= 1 \implies L_u(x,u(x),u'(x)) = 1 \\
L_v(x,u,v) &= \frac{v}{\sqrt{1+v^2}} \implies L_v(x,u(x),u'(x)) = \frac{u'(x)}{\sqrt{1+u'(x)^2}}
\end{aligned}
$$

$$\text{EL-eqn:} \quad \frac{d}{dx}\left(\frac{u'(x)}{\sqrt{1+u'(x)^2}} \right) - 1 = 0, \text{ i.e. } \frac{d}{dx}\left(\frac{u'(x)}{\sqrt{1+u'(x)^2}} \right) = 1$$

(d) $J[u] = \int_a^b \sin\big(u'(x)\big) \, dx$ **Solution.**

$$
\begin{aligned}
L(x,u,v) &= \sin(v) \\
L_u(x,u,v) &= 0 \implies L_u(x,u(x),u'(x)) = 0 \\
L_v(x,u,v) &= \cos(v) \implies L_v(x,u(x),u'(x)) = \cos\big(u'(x)\big)
\end{aligned}
$$

$$\text{EL-eqn:} \quad \frac{d}{dx}\big(\cos u'(x)\big) - 0 = 0, \text{ i.e. } \frac{d}{dx}\big(\cos u'(x)\big) = 0$$

(e) $J[u] = \int_a^b u'(x)dx$ **Solution.**

$$\begin{aligned}
L(x,u,v) &= v \\
L_u(x,u,v) &= 0 \implies L_u(x,u(x),u'(x)) = 0 \\
L_v(x,u,v) &= 1 \implies L_v(x,u(x),u'(x)) = 1 \\
\text{EL–eqn:} \quad \frac{d}{dx}(1) - 0 &= 0, \text{ i.e. } 0 = 0
\end{aligned}$$

The Euler–Lagrange equation is trivial in this case, and is satisfied by every function u.

(f) $J[u] = \int_a^b u(x)dx$ **Solution.**

$$\begin{aligned}
L(x,u,v) &= v \\
L_u(x,u,v) &= 1 \implies L_u(x,u(x),u'(x)) = 1 \\
L_v(x,u,v) &= 0 \implies L_v(x,u(x),u'(x)) = 0 \\
\text{EL–eqn:} \quad \frac{d}{dx}(0) - 1 &= 0, \text{ i.e. } 0 = 1
\end{aligned}$$

The Euler–Lagrange equation has no solutions.

3.2. Solve the Euler–Lagrange equations for each of the following minimization problems:

(a) $J[u] = \int_a^b \frac{1}{2}u'(x)^2\,dx$, $u(a) = A$, $u(b) = B$

(b) $J[u] = \int_0^\ell \left\{\frac{1}{2}u'(x)^2 + u(x)\right\}\,dx$, $u(0) = u(\ell) = 0$ ($\ell > 0$ is a constant)

Solution. (a) From the previous problem we know that the Euler–Lagrange equation is $u''(x) = 0$. The solutions to this equation are $u(x) = C_1x + C_2$ where C_1, C_2 are constants. The solution that satisfies $u(a) = A$, $u(b) = B$ is found by solving

$$\left.\begin{aligned} C_1a + C_2 = A \\ C_1b + C_2 = B \end{aligned}\right\} \iff C_1 = \frac{B-A}{b-a}, \quad C_2 = \frac{Ab - aB}{b-a},$$

so the unique solution to the Euler–Lagrange equations that satisfies the boundary conditions is

$$u(x) = \frac{B-A}{b-a}x + \frac{Ab-aB}{b-a}.$$

(b) We found in the problem 3.1 that the Euler–Lagrange equation is $u''(x) = 1$. The solutions to this equation area

$$u(x) = \tfrac{1}{2}x^2 + C_1x + C_2$$

We want to find the solution that satisfies $u(0) = 0$, $u(\ell) = 0$. The appropriate values of C_1, C_2 are $C_1 = 0, C_2 = -\frac{1}{2}\ell$, so that the only solution of the Euler–Lagrange equations that satisfy the given boundary conditions is $u(x) = \frac{1}{2}x^2 - \frac{1}{2}\ell x = -\frac{1}{2}x(\ell - x)$.

3.3. A strange minimization puzzle. Consider the problem of finding a function that minimizes $J[u] = \int_a^b u'(x)dx$ among all functions u with $u(a) = A$, $u(b) = B$. The Euler–Lagrange equation does not tell you very much (what does it say?).

What is the minimal value of $J[u]$?

Solution. We found in Problem 3.1 that the Euler–Lagrange equation is $0 = 0$, which is satisfied by every C^2 function u.

How can this be? The Fundamental Theorem of Calculus implies that for any C^1 function $u : [a, b] \to \mathbb{R}$ one has

$$J[u] = \int_a^b u'(x)dx = u(b) - u(a).$$

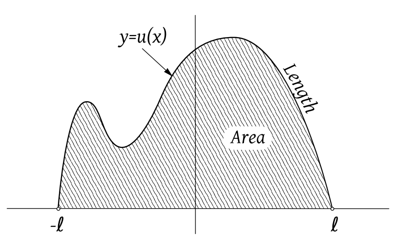
If the boundary values are prescribed by $u(a) = A, u(b) = B$, then the value $J[u] = B - A$ does not depend on u, and therefore every u with $u(a) = A, u(b) = B$ is a minimizer. The minimal value of $J[u]$ over all C^2 functions $u : [a, b] \to \mathbb{R}$ with $u(a) = A, u(b) = B$ is $B - A$ and it is attained by all such u.

3.4. Minimize length−area under the graph. Let $\ell > 0$ be some positive number. For any curve from $(-\ell, 0)$ to $(\ell, 0)$ that is the graph of a function $u : [-\ell, \ell] \to \mathbb{R}$ we consider

$$J[u] = \int_{-\ell}^{\ell} \left\{ \sqrt{1 + u'(x)^2} - u(x) \right\} dx,$$

i.e. $J[u]$ is the difference of the length of the graph of u and the area beneath the graph.

Note: this problem is a variation on Dido's problem which appears in the story of the founding of the city of Carthage in ∼800BC[1]. Imagine that the x-axis is the lake shore, and you have just been told that you can sell all the land you can enclose with a fence starting and ending at the points $(\pm\ell, 0)$ on the lake shore. The King will pay $1 per square foot of area, but you have to spend $1 per foot of fence you make. Your expense is the length of the fence minus the enclosed area. Which shape should the fence have to minimize your loss (or maximize your gain)?



(a) In an attempt to find a function u with the lowest possible value of $J[u]$, you could compute the Euler–Lagrange equation, and then solve the resulting differential equation. What do you find?
(b) Suppose $\ell = 3$ and consider for any $M > 0$ the function $u_M(x) = M(\ell^2 - x^2)$. Show that $\lim_{M\to\infty} J[u_M] = -\infty$. (The inequality $\sqrt{a^2 + b^2} \le |a| + |b|$ may be useful.)
(c) Let $m \in \mathbb{N}$ be a positive integer and consider for any $M > 0$ the function $u_M(x) = M(\ell^{2m} - x^{2m})$. For which $\ell > 1$ is it true that $\lim_{M\to\infty} J[u_M] = -\infty$? (The inequality $\sqrt{a^2 + b^2} \le |a| + |b|$ may again be useful.)

[1] See https://www.ams.org/journals/notices/201709/rnoti-p980.pdf or google "Dido's problem"

Solution.

(a) The Lagrangian is $L(x,u,v) = \sqrt{1+v^2} - u$, so that the Euler–Lagrange equations are

$$\frac{d}{dx}\left(\frac{u'(x)}{\sqrt{1+u'(x)^2}}\right) + 1 = 0$$

Integrate once. Any solution to the Euler–Lagrange equations satisfies

$$\frac{u'(x)}{\sqrt{1+u'(x)^2}} = -x + C_1$$

for some constant C_1, and thus

$$u'(x) = \frac{-x + C_1}{\sqrt{1-(x-C_1)^2}}.$$

Integrate again to get

$$u(x) = \sqrt{1-(x-C_1)^2} + C_2.$$

Rewriting this as

$$(x-C_1)^2 + (u(x)-C_2)^2 = 1$$

we see that the graph of u is a circle with radius 1 and center (C_1, C_2).

We are looking for a minimizer that satisfies $u(\pm\ell) = 0$, i.e. one for which the graph of u contains the points $(\pm\ell, 0)$. Since the graph of u is a part of a circle with radius 1, the distance between the points $(\pm\ell, 0)$ cannot be more than 2.

It follows that when $\ell > 1$ none of the solutions of the Euler–Lagrange equation satisfy the boundary condition $u(\pm\ell) = 0$. In this case there is no minimizer that is also C^2.

(b) Using the triangle inequality $\sqrt{a^2+b^2} \le |a| + |b|$ we find for any function u_M

$$J[u_M] = \int_{-\ell}^{\ell} \left\{\sqrt{1+u_M'(x)^2} - u_M(x)\right\} dx \le \int_{-\ell}^{\ell} \left\{1 + |u_M'(x)| - u_M(x)\right\} dx$$

For $u_M(x) = M(\ell^2 - x^2)$ we get

$$J[u_M] \le 2\ell + \int_{-\ell}^{\ell} \left\{2M|x| - M(\ell^2 - x^2)\right\} dx = 2\ell + 2M\ell^2 - \tfrac{4}{3}M\ell^3 = 2\ell - \tfrac{4}{3}M\ell^2\left(\ell - \tfrac{3}{2}\right)$$

In our problem $\ell = 3$ so $J[u_M] \le 6 - 18M$, which implies $\lim_{M\to\infty} J[u_M] = -\infty$.

(c) We now have $u_M(x) = M(\ell^{2m} - x^{2m})$ so that

$$\begin{aligned}
J[u_M] &\le \int_{-\ell}^{\ell} \left\{1 + |u_M'(x)| - u_M(x)\right\} dx \\
&= 2\ell + M \int_{-\ell}^{\ell} \left\{|u_M'(x)| - u_M(x)\right\} dx \\
&= 2\ell + M \int_{-\ell}^{\ell} \left\{2m|x|^{2m-1} - \ell^{2m} + x^{2m}\right\} dx \\
&= 2\ell + M\left\{2\ell^{2m} - 2\ell^{2m+1} + \tfrac{2}{2m+1}\ell^{2m+1}\right\} \\
&= 2\ell + 2M\ell^{2m}\left\{1 - \ell + \tfrac{1}{2m+1}\ell\right\} \\
&= 2\ell + 2M\ell^{2m}\left\{1 - \tfrac{2m}{2m+1}\ell\right\}.
\end{aligned}$$

If $\ell > 1$ then we can choose m so large that $\frac{2m}{2m+1}\ell > 1$. In that case we again get $\lim_{M\to\infty} J[u_M] = -\infty$. In particular, there is no minimizer.

3.5. Does a minimizer exist? Let $a > 0$ be a real constant and consider the problem of minimizing

$$J[u] = \int_0^1 x^a u'(x)^2 \, dx, \qquad u(0) = u(1) = 0.$$

(a) Find the Euler–Lagrange equation for this problem.
(b) Suppose $a = \frac{1}{2}$, and find the solution to the Euler–Lagrange equation with $u(0) = 0$, $u(1) = 1$. Is the solution you find C^2 on the whole interval $[0, 1]$?
(c) Suppose $a = 2$. Show that there is no solution to the Euler–Lagrange equation with $u(0) = 0$, $u(1) = 1$.

Solution. (a) The Lagrangian is $L(x, u, v) = x^a v^2$ so the Euler–Lagrange equation is

$$\frac{d}{dx}\big(2x^a u'(x)\big) = 0.$$

The problem doesn't ask you to solve the equation, but here is the solution anyway: Integrate to find that any solution of the EL equation satisfies $x^a u'(x) = C_1$ for some $C_1 \in \mathbb{R}$. Divide by x^a and integrate again and we see that if $a \neq 1$ then any solution of the EL equation satisfies

$$u(x) = \frac{C_1}{1-a} x^{1-a} + C_2$$

for certain constants C_1, C_2. If $a = 1$ then any solution of the EL-equation is

$$u(x) = C_1 \ln x + C_2.$$

(b) If $a = \frac{1}{2}$ then the solutions to the EL equation are given by $u(x) = 2C_1\sqrt{x} + C_2$. If $u(0) = 0$ then $C_2 = 0$ so $u(x) = 2C_1\sqrt{x}$. If in addition $u(1) = 1$, then $C_1 = \frac{1}{2}$, so the solution of the EL equation that satisfies the boundary conditions is $u(x) = \sqrt{x}$. This function is not differentiable at $x = 0$, so it is certainly not C^2 on the whole interval $[0, 1]$.

(c) If $a = 2$ then any solution to the EL equation is given by $u(x) = -C_1 x^{-1} + C_2$. This function is not defined at $x = 0$ unless $C_1 = 0$, so the only solutions that satisfy $\lim_{x\to 0} u(x) = 0$ are the ones with $C_1 = 0$, i.e. $u(x) = C_2$. The condition $u(0) = 0$ then also requires $C_2 = 0$, so we conclude that the only solution of the EL equation that satisfies $\lim_{x\to 0} u(x) = 0$ is the zero solution, $u(x) = 0$. This solution does not satisfy $u(1) = 1$, so there is no solution of the EL equation that satisfies the boundary conditions.

3.6. Noether's conserved quantity. Let $u : [a, b] \to \mathbb{R}$ be a C^2 solution to the Euler–Lagrange equation for the Lagrangian $L(x, u, v)$. Assume that the Lagrangian does not depend on x, i.e. $L = L(u, v)$. Show that

$$\frac{d}{dx}\left\{u'(x)\frac{\partial L}{\partial v}(u(x), u'(x)) - L(u(x), u'(x))\right\} = 0.$$

This is a special case of a much more general result due to Emmy Noether relating symmetries of a variational problem with conserved quantities for their Euler–Lagrange equations.

4. Calculus of Variations in nD

4.1. Example — the area of a surface. Let C be a curve in $\mathbb{R}^3$. We consider the problem of finding a two-dimensional surface $S \subset \mathbb{R}^3$ whose boundary is the curve C and which has the smallest possible area of all such surfaces.

Instead of considering the most general case, we will assume that the surface S is the graph of a function $z = u(x, y)$ of two variables. The domain of this function is some region $\mathcal{R} \subset \mathbb{R}^2$ in the plane. We will assume that

- $\mathcal{R}$ is an open subset of the plane,
- the boundary of the region is a "nice curve," i.e. a circle, ellipse, rectangle, a polygon, or a piecewise C^1 curve.

We write $\partial\mathcal{R}$ for the boundary curve.

In multivariable calculus one defines the area of the surface S to be

$$A[u] = \iint_{\mathcal{R}} \sqrt{1 + |\nabla u|^2}\, dA = \iint_{\mathcal{R}} \sqrt{1 + \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2}\, dx\, dy$$

If we use polar coordinates on the domain R then the surface area is given by

$$A[u] = \iint_{\mathcal{R}} \sqrt{1 + u_r^2 + \frac{1}{r^2} u_\theta^2}\;\; r\, d\theta\, dr.$$

Both forms of $A[u]$ are of the form

$$I[u] = \iint_{\mathcal{R}} L(x, y, u, v_x, v_y)\, dx\, dy$$

where $L : \mathcal{R} \times \mathbb{R}^3$ is a function of five variables. For the area integral A in Cartesian coordinates we should choose

$$L(x, y, u, v_x, v_y) = \sqrt{1 + v_x^2 + v_y^2}\;,$$

while for the area integral in polar coordinates we should choose

$$L(r, \theta, u, v_r, v_\theta) = r\sqrt{1 + v_r^2 + \frac{v_\theta^2}{r^2}}\;.$$

4.2. The Euler–Lagrange equation for multiple integrals. Let $\mathcal{R} \subset \mathbb{R}^n$ be an open domain, and let $L : \mathcal{R} \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ be a C^1 function. For any C^1 function $u : \mathcal{R} \to \mathbb{R}$ we define

$$I[u] = \iint_{\mathcal{R}} L\left(x, u(x), \tfrac{\partial u}{\partial x_1}(x), \ldots, \tfrac{\partial u}{\partial x_n}(x)\right)\, dx_1 \ldots dx_n$$

A shorter way of writing this integral is to abbreviate

$$\nabla u(x) = \big(u_{x_1}(x), \ldots, u_{x_n}(x)\big).$$

and

$$dx = dx_1 \cdots dx_n.$$

With this notation we can write

$$I[u] = \iint_{\mathcal{R}} L\big(x, u(x), \nabla u(x)\big)\, dx.$$

4.3. Definition. A C^1 function $u : \mathcal{R} \to \mathbb{R}$ with $u = g$ on $\partial\mathcal{R}$ minimizes $I[u]$ if for every C^1 function $\bar{u} : \mathcal{R} \to \mathbb{R}$ with $\bar{u} = u$ on $\partial\mathcal{R}$ one has $I[\bar{u}] \geq I[u]$.

4.4. Theorem. If L is C^2, and if $u : \mathcal{R} \to \mathbb{R}$ is C^2 minimizes $I[u]$ among all C^2 functions with $u = g$ on $\partial\mathcal{R}$, then u satisfies the Euler–Lagrange equation

$$\frac{\partial}{\partial x_1}\left(\frac{\partial L}{\partial v_1}\right) + \cdots + \frac{\partial}{\partial x_n}\left(\frac{\partial L}{\partial v_n}\right) = \frac{\partial L}{\partial u}. \tag{21}$$

Here the partial derivatives $\frac{\partial L}{\partial v_i}$ and $\frac{\partial L}{\partial u}$ are to be evaluated at $(x, u(x), u_{x_1}(x), \ldots, u_{x_n}(x))$.

4.5. Example: the Laplace equation and the Dirichlet integral. The Laplace equation is the partial differential equation

$$\Delta u = 0, \text{ where } \Delta u \stackrel{\text{def}}{=} \frac{\partial^2 u}{\partial x_1^2} + \cdots + \frac{\partial^2 u}{\partial x_n^2}. \tag{22}$$

The expression Δu is called the Laplacian of the function $u : \mathcal{R} \to \mathbb{R}$. Solutions of the Laplace equation are called harmonic functions.

Dirichlet introduced the Dirichlet integral

$$\mathcal{D}[u] = \frac{1}{2}\int_{\mathcal{R}} |\nabla u(x)|^2 \, dx_1 \cdots dx_n$$

in which

$$\nabla u = \begin{pmatrix} \partial u/\partial x_1 \\ \vdots \\ \partial u/\partial x_n \end{pmatrix} \text{ and } \|\nabla u\|^2 = \left(\frac{\partial u}{\partial x_1}\right)^2 + \cdots + \left(\frac{\partial u}{\partial x_n}\right)^2$$

are the gradient of the function $u : \mathcal{R} \to \mathbb{R}$ and its squared length.

Dirichlet observed that Laplace's equation is the Euler–Lagrange equation for the Dirichlet integral. To verify this we observe that $\mathcal{D}[u] = \iint_{\mathcal{R}} L(x, u, u_{x_1}, \ldots, u_{x_n}) \, dx_1 \cdots dx_n$ for the Lagrangian

$$L(x, u, v_1, \ldots, v_n) = \frac{1}{2}\{v_1^2 + \cdots + v_n^2\}.$$

We compute the associated Euler–Lagrange equations:

$$\begin{aligned} \frac{\partial L}{\partial u}(x, u, v_1, \ldots, v_n) = 0 \text{ and } \quad & \frac{\partial L}{\partial v_k}(x, u, v_1, \ldots, v_n) = v_k \\ \implies \text{Euler–Lagrange:} \quad & \frac{\partial u_{x_1}}{\partial x_1} + \cdots + \frac{\partial u_{x_n}}{\partial x_n} = 0 \\ \iff & \frac{\partial^2 u}{\partial x_1^2} + \cdots + \frac{\partial^2 u}{\partial x_n^2} = 0. \end{aligned}$$

4.6. Example: the Minimal Surface Equation. Suppose $n = 2$. If the graph of u minimizes the area $A[u]$, then the Lagrangian is

$$L(x, u, v_x, v_y) = \sqrt{1 + v_x^2 + v_y^2}$$

so that

$$\frac{\partial L}{\partial u} = 0, \qquad \frac{\partial L}{\partial v_x} = \frac{v_x}{\sqrt{1 + v_x^2 + v_y^2}}, \qquad \frac{\partial L}{\partial v_y} = \frac{v_y}{\sqrt{1 + v_x^2 + v_y^2}}.$$

The Euler–Lagrange equation for minimizers of the area integral $A[u]$ is therefore

$$\frac{\partial}{\partial x}\left(\frac{u_x}{\sqrt{1 + u_x^2 + u_y^2}}\right) + \frac{\partial}{\partial y}\left(\frac{u_y}{\sqrt{1 + u_x^2 + u_y^2}}\right) = 0. \tag{23}$$

This equation is related to the Laplace equation if one considers functions u whose partial derivatives are much smaller than 1. Namely, if we assume that $|u_x|, |u_y| \ll 1$, then $1 + u_x^2 + u_y^2 \approx 1$ and thus one would expect solutions to (23) to satisfy

$$\frac{\partial}{\partial x}\left(\frac{u_x}{\sqrt{1}}\right) + \frac{\partial}{\partial y}\left(\frac{u_y}{\sqrt{1}}\right) \approx 0,$$

i.e.

$$u_{xx} + u_{yy} \approx 0.$$

4.7. Proof of Theorem 4.4. For simplicity we assume $n = 2$. The general case can be proved in much the same way.

Let $h : \mathcal{R} \to \mathbb{R}$ be a C^2 function with $h = 0$ on $\partial\mathcal{R}$. Then $I[u + \varepsilon h] \geq I[u]$ for all $\varepsilon \in \mathbb{R}$. It follows that

$$\frac{dI[u + \varepsilon h]}{d\varepsilon}\Big|_{\varepsilon=0} = 0.$$

To compute the derivative of $I[u+\varepsilon h]$ with respect to ε we use $\nabla(u+\varepsilon h) = \nabla u + \varepsilon \nabla h$:

$$\begin{aligned}
&\frac{dI[u + \varepsilon h]}{d\varepsilon} = \\
&= \frac{d}{d\varepsilon} \iint_{\mathcal{R}} L\big(x, u(x) + \varepsilon h(x), \nabla u(x) + \varepsilon \nabla h(x)\big)\, dx \\
&= \iint_{\mathcal{R}} \frac{\partial}{\partial \varepsilon} L\big(x, u(x) + \varepsilon h(x), \nabla u(x) + \varepsilon \nabla h(x)\big)\, dx \\
&= \iint_{\mathcal{R}} \left\{ \frac{\partial L}{\partial u} \cdot h(x) + \frac{\partial L}{\partial v_{x_1}} \cdot \frac{\partial h}{\partial x_1}(x) + \cdots + \frac{\partial L}{\partial v_{x_n}} \cdot \frac{\partial h}{\partial x_n}(x) \right\} dx
\end{aligned}$$

where $\frac{\partial L}{\partial u}$, $\frac{\partial L}{\partial v_x}$, and $\frac{\partial L}{\partial v_y}$ are evaluated at $\big(x, u(x) + \varepsilon h(x), \nabla u(x) + \varepsilon \nabla h(x)\big)$. We now integrate by parts using the product rule

$$\frac{\partial L}{\partial v_{x_j}} \cdot \frac{\partial h}{\partial x_j} = \frac{\partial}{\partial x_j}\left(\frac{\partial L}{\partial v_{x_j}} \cdot h\right) - \frac{\partial}{\partial x_j}\left(\frac{\partial L}{\partial v_{x_j}}\right) \cdot h,$$

as well as Green's theorem

$$\begin{aligned}
\frac{dI[u + \varepsilon h]}{d\varepsilon} = &\iint_{\mathcal{R}} \left\{ \frac{\partial L}{\partial u} - \frac{\partial}{\partial x_1}\Big(\frac{\partial L}{\partial v_{x_1}}\Big) - \cdots - \frac{\partial}{\partial x_n}\Big(\frac{\partial L}{\partial v_{x_n}}\Big) \right\} h(x) dx. \\
&+ \oint_{\partial\mathcal{R}} \left\{ \boldsymbol{n}_{x_1} \frac{\partial L}{\partial v_{x_1}} h + \cdots + \boldsymbol{n}_{x_n} \frac{\partial L}{\partial v_{x_n}} h \right\} dS.
\end{aligned}$$

in which $\boldsymbol{n} = \boldsymbol{n}_{x_1}, \ldots, \boldsymbol{n}_{x_n})$ is the outward unit normal to $\partial\mathcal{R}$. Since $h = 0$ on $\partial\mathcal{R}$ the last integral vanishes and we find

$$\frac{dI[u+\varepsilon h]}{d\varepsilon} = \iint_{\mathcal{R}} \left\{ \frac{\partial L}{\partial u} - \frac{\partial}{\partial x_1}\Big(\frac{\partial L}{\partial v_{x_1}}\Big) - \cdots - \frac{\partial}{\partial x_n}\Big(\frac{\partial L}{\partial v_{x_n}}\Big) \right\} h(x) dx.$$

Keeping in mind that $\frac{\partial L}{\partial u}$, etc., are all evaluated at $(x, y, u(x) + \varepsilon h(x), \nabla u(x) + \varepsilon \nabla h(x))$, we set $\varepsilon = 0$ and find

(24) $$\frac{dI[u+\varepsilon h]}{d\varepsilon}\Big|_{\varepsilon=0} = \iint_{\mathcal{R}} \left\{ \frac{\partial L}{\partial u} - \frac{\partial}{\partial x_1}\Big(\frac{\partial L}{\partial v_{x_1}}\Big) - \cdots - \frac{\partial}{\partial x_n}\Big(\frac{\partial L}{\partial v_{x_n}}\Big) \right\} h(x) dx.$$

where $\frac{\partial L}{\partial u}$, etc., are now evaluated at $\big(x, u(x), \nabla u(x)\big)$.

Since u minimizes $I[u]$ we have shown that

(25) $$\iint_{\mathcal{R}} \left\{ \frac{\partial L}{\partial u} - \frac{\partial}{\partial x_1}\Big(\frac{\partial L}{\partial v_{x_1}}\Big) - \cdots - \frac{\partial}{\partial x_n}\Big(\frac{\partial L}{\partial v_{x_n}}\Big) \right\} h(x) dx = 0$$

for all C^2 functions $h : \mathcal{R} \to \mathbb{R}$ with $h = 0$ on $\partial\mathcal{R}$.

Finally, the fact that $L : \mathcal{R} \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ is C^2 and $u : \mathcal{R} \to \mathbb{R}$ also is C^2 implies

$$\frac{\partial L}{\partial u} - \frac{\partial}{\partial x_1}\Big(\frac{\partial L}{\partial v_{x_1}}\Big) - \cdots - \frac{\partial}{\partial x_n}\Big(\frac{\partial L}{\partial v_{x_n}}\Big)$$

is a continuous function on $\mathcal{R}$. It then follows from condition (25) that u satisfies equation (21). *Q.E.D.*

5. Convexity

5.1. Definitions. A subset $C \subset V$ of a real vector space V is called convex if the line segment connecting any two points in C is contained in C, i.e. if for all $x, y \in C$ and all $\theta \in (0, 1)$ one has $(1-\theta)x + \theta y \in C$.

If $C \subset V$ is a convex set then a function $f : C \to V$ is convex if for all $x, y \in C$ and all $\theta \in (0, 1)$ one has

$$f\big((1-\theta)x + \theta y\big) \le (1-\theta)f(x) + \theta f(y).$$

The function f is said to be strictly convex if

$$f\big((1-\theta)x + \theta y\big) < (1-\theta)f(x) + \theta f(y).$$

holds for all $x, y \in C$ and $\theta \in (0, 1)$.

5.2. Calculus Theorems on Convexity.

(a) If $f : (a, b) \to \mathbb{R}$ is twice differentiable and and convex then $f''(x) \ge 0$ for all $x \in (a, b)$.
(b) If $f : (a, b) \to \mathbb{R}$ is twice differentiable and if $f''(x) \ge 0$ for all $x \in (a, b)$, then f is convex.
(c) If $f : (a, b) \to \mathbb{R}$ is twice differentiable and if $f''(x) > 0$ for all $x \in (a, b)$, then f is strictly convex.
(d) Let $\mathcal{R} \subset \mathbb{R}^n$ be a convex open set. If $f : \mathcal{R} \to \mathbb{R}$ is twice differentiable and if $D^2 f(x) \ge 0$, i.e. if $D^2 f(x)$ is positive semi-definite for all $x \in \mathcal{R}$, then f is convex.
(e) If $f : \mathcal{R} \to \mathbb{R}$ is twice differentiable and if $D^2 f(x) > 0$, i.e. if $D^2 f(x)$ is positive definite for all $x \in \mathcal{R}$, then f is strictly convex.

In the last two statements $D^2 f(x)$ is the Hessian matrix of f, which, by definition, consists of all second order partial derivatives of f:

$$D^2 f(x) = \begin{pmatrix} f_{x_1x_1}(x) & \cdots & f_{x_1x_n}(x) \\ \vdots & & \vdots \\ f_{x_nx_1}(x) & \cdots & f_{x_nx_n}(x) \end{pmatrix}$$

The Hessian is a symmetric matrix. By definition the matrix $D^2 f(x)$ is positive semi-definite if for every vector $h \in \mathbb{R}^n$ one has

$$\langle h, D^2 f(x) \cdot h \rangle \geq 0.$$

It is positive definite if

$$\langle h, D^2 f(x) h \rangle > 0 \text{ for all } h \in \mathbb{R}^n, h \neq 0.$$

In practice one can see if $D^2 f(x)$ satisfies this condition by verifying

$$\frac{d^2 f(x + \theta h)}{d\theta^2}\bigg|_{\theta=0} \geq 0$$

i.e. regard $g(\theta) = f(x + \theta h)$ as a function of θ and compute its second derivative.

5.3. Example. The function $f(x) = |x|$ is convex, due to the triangle inequality. Namely, for all $x \in \mathbb{R}$ and $\theta \in (0, 1)$:

$$|(1-\theta)x + \theta y| \leq |(1-\theta)x| + |\theta y| = (1-\theta)|x| + \theta|y|.$$

5.4. Theorem about Convex Lagrangians. Let $L : \mathcal{R} \times \mathbb{R}^n \to \mathbb{R}$ be a C^2 function, and assume that for each $x \in \mathcal{R}$ the function $v \mapsto L(x, v)$ is a strictly convex function from $\mathbb{R}^n$ to $\mathbb{R}$, i.e. assume that for all $v, \bar{v} \in \mathbb{R}^n$ with $v \neq \bar{v}$, and for all $\theta \in (0, 1)$ one has

$$L(x, (1-\theta)v + \theta\bar{v}) < (1-\theta)L(x, v) + \theta L(x, \bar{v}).$$

Let $g : \partial\mathcal{R} \to \mathbb{R}$ be a given function. Then the function $I : C^1(\bar{\mathcal{R}}) \to \mathbb{R}$ defined by

$$I[u] = \iint_{\mathcal{R}} L(x, \nabla u(x))\, dx$$

has at most one C^1 minimizer $u : \mathcal{R} \to \mathbb{R}$ with $u = g$ on $\partial\mathcal{R}$.

(We again abbreviate)

Proof. Suppose $u, \bar{u} : \mathcal{R} \to \mathbb{R}$ both are minimizers with $u = \bar{u} = g$ on $\partial\mathcal{R}$. This means that for every C^1 function $v : \mathcal{R} \to \mathbb{R}$ with $v = g$ on $\partial\mathcal{R}$ one has $I[v] \geq I[u]$ and $I[v] \geq I[\bar{u}]$. In particular, we have $I[u] \geq I[\bar{u}]$ and $I[\bar{u}] \geq I[u]$, i.e. we have $I[u] = I[\bar{u}]$.

Choose a number $\theta \in (0, 1)$ (e.g. $\theta = \frac{1}{2}$), and consider the function $w(x) = (1-\theta)u(x) + \theta\bar{u}(x)$. Since u is a minimizer for I we have

$$\begin{aligned} I[u] \leq I[w] &= \iint_{\mathcal{R}} L(x, \nabla w(x))\, dx \\ &= \iint_{\mathcal{R}} L\big(x, (1-\theta)\nabla u(x) + \theta\nabla\bar{u}(x)\big)\, dx \end{aligned}$$

Convexity of L implies

$$L\big(x, (1-\theta)\nabla u(x) + \theta\nabla\bar{u}(x)\big) \leq (1-\theta)L\big(x, \nabla u(x)\big) + \theta L\big(x, \nabla\bar{u}(x)\big) \tag{26}$$

Thus we have

$$
\begin{aligned}
I[u] &\le \iint_{\mathcal{R}} \Big[(1-\theta)L\big(x, \nabla u(x)\big) + \theta L\big(x, \nabla \bar{u}(x)\big)\Big]\,dx \\
&= (1-\theta)\iint_{\mathcal{R}} L\big(x, \nabla u(x)\big)\,dx + \theta \iint_{\mathcal{R}} L\big(x, \nabla \bar{u}(x)\big)\,dx \\
&= (1-\theta)I[u] + \theta I[\bar{u}] \qquad (\text{recall } I[u] = I[\bar{u}]) \\
&= I[u].
\end{aligned}
$$

This shows that the inequalities above are actually equalities, and, in particular, the equality in (26) is an equality for all $x \in \mathcal{R}$. Since $L(x, v)$ is a strictly convex function of $v \in \mathbb{R}^n$, we conclude that $\nabla u(x) = \nabla \bar{u}(x)$ holds for all $x \in \mathcal{R}$, i.e. the difference function $u(x) - \bar{u}(x)$ is constant. Since $u - \bar{u} = 0$ on $\partial\mathcal{R}$, we have proved that $u(x) - \bar{u}(x) = 0$ for all $x \in \mathcal{R}$, i.e. u and $\bar{u}$ are the same function. *Q.E.D.*

5.5. A sufficient condition for a function to be a minimizer. If $L : \mathcal{R} \times \mathbb{R}^n$ is a convex function of v and if $u : \mathcal{R} \to \mathbb{R}$ is a C^2 function that satisfies the Euler–Lagrange equation, then for any other C^2 function $\bar{u} : \mathcal{R} \to \mathbb{R}$ with $\bar{u} = u$ on $\partial\mathcal{R}$ one has $I[\bar{u}] \ge I[u]$.

Proof. Let $h = \bar{u} - u$ and consider the function

$$g(t) = I[u + th].$$

We first show that g is a convex function, and then we show $g'(0) = 0$.

The function g is convex: let $t_0, t_1 \in [0, 1]$ and $\theta \in (0, 1)$ be given. Then

$$
\begin{aligned}
g((1-\theta)t_0 + \theta t_1) &= I[u + \big((1-\theta)t_0 + \theta t_1\big)h] \\
&= I[(1-\theta)(u + t_0 h) + \theta(u + t_1 h)] \\
&= \iint_{\mathcal{R}} L\big(x, (1-\theta)\nabla(u + t_0 h) + \theta\nabla(u + t_1 h)\big)\,dx \\
&\le (1-\theta)\iint_{\mathcal{R}} L\big(x, \nabla(u + t_0 h)\big)\,dx + \theta\iint_{\mathcal{R}} L\big(x, \nabla(u + t_1 h)\big)\,dx \\
&= (1-\theta)I[u + t_0 h] + \theta I[u + t_1 h] \\
&= (1-\theta)g(t_0) + \theta g(t_1).
\end{aligned}
$$

The derivative at $t = 0$: In the derivation of the Euler–Lagrange equation we found

$$g'(0) = \frac{dI[u + th]}{dt}\Big|_{t=0} = -\int_{\mathcal{R}} \left[\frac{\partial}{\partial x_1}\frac{\partial L}{\partial v_1} + \cdots + \frac{\partial}{\partial x_n}\frac{\partial L}{\partial v_n}\right] dx$$

The Euler–Lagrange equation states that the integrand in this integral vanishes, which implies $g'(0) = 0$.

We conclude by observing that a convex function $g : [0, 1] \to \mathbb{R}$ with $g'(0) = 0$ satisfies $g(t) \ge g(0)$ for all $t \in [0, 1]$ and hence $I[\bar{u}] = g(1) \ge g(0) = I[u]$. *Q.E.D.*

6. Problems

6.1. In this problem we let $f, g : C \to \mathbb{R}$ be convex functions defined on a convex subset $C \subset V$ of some real vector space V. Prove or give a counter example :

(a) $f + g$ is convex
(b) fg is convex
(c) $f - g$ is convex
(d) $h(x) = \phi\big(f(x)\big)$ is convex, if $\phi : \mathbb{R} \to \mathbb{R}$ is a convex function
(e) $h(x) = \max\{f(x), g(x)\}$ is a convex function
(f) $h(x) = \min\{f(x), g(x)\}$ is a convex function

Solution. (a) For any $x, y \in C$ and $\theta \in (0,1)$ it follows from convexity of f and g that

$$f\big((1-\theta)x + \theta y\big) \le (1-\theta)f(x) + \theta f(y), \text{ and } g\big((1-\theta)x + \theta y\big) \le (1-\theta)g(x) + \theta g(y).$$

Hence

$$\begin{aligned}(f+g)\big((1-\theta)x + \theta y\big) &= f\big((1-\theta)x + \theta y\big) + g\big((1-\theta)x + \theta y\big)\\ &\le (1-\theta)f(x) + \theta f(y) + (1-\theta)g(x) + \theta g(y)\\ &= (1-\theta)(f+g)(x) + \theta(f+g)(y)\end{aligned}$$

Therefore $f + g$ is convex.

(b) If $C = V = \mathbb{R}$ and $f(x) = -x$, $g(x) = x$, then f and g are convex, but $(fg)(x) = -x^2$ is not convex.

(c) Again let $C = V = \mathbb{R}$. If $f(x) = 0$ and $g(x) = x^2$ then f and g are convex, but $(f - g)(x) = -x^2$ is not convex.

(d) This is not true in general. For example, if $C = V = \mathbb{R}$ and if $\phi(x) = -x$, $f(x) = x^2$, then $h(x) = \phi(f(x)) = -x^2$ is not convex.

However, if ϕ is convex and also non decreasing ($\forall x < y : \phi(x) \le \phi(y)$), then $h(x) = \phi(f(x))$ is convex. Namely, for all $x, y \in C$ and $\theta \in \mathbb{R}$ one has

$$\begin{aligned} & f((1-\theta)x + \theta y) \le (1-\theta)f(x) + \theta f(y) \\ \implies\ & \phi\big(f((1-\theta)x + \theta y)\big) \le \phi\big((1-\theta)f(x) + \theta f(y)\big) && \text{because } \phi \text{ is nondecreasing}\\ & \qquad \le (1-\theta)\phi\big(f(x)\big) + \theta\phi\big(f(y)\big) && \text{because } \phi \text{ is convex}\end{aligned}$$

(e) $\max\{f, g\}$ is convex

(f) $\min\{f, g\}$ does not have to be convex. For example, if $C = V = \mathbb{R}$ and $f(x) = x$, $g(x) = -x$, then $h(x) = \min\{f(x), g(x)\} = -|x|$ is not convex.

6.2. In this problem let V be a real vector space and let $x \mapsto \|x\|$ be a norm on V. See §3.4 for the properties of a norm.

(a) Use the triangle inequality to show that $f(x) = |x|$ is a convex function.
(b) Show that $f : V \to \mathbb{R}$ defined by $f(x) = \|x\|$ is a convex function.
(c) Show that $f : V \to \mathbb{R}$ defined by $f(x) = \|x\|^m$ is a convex function for any $m \ge 1$. (Hint: consider part (d) of Problem 6.1.)
(d) Show that $h(x) = \sqrt{1 + \|x\|^2}$ is a convex function.

Solution. (a) for all $x, y \in \mathbb{R}$ and $\theta \in (0,1)$ we have $|(1-\theta)x + \theta y| \leq |(1-\theta)x| + |\theta y| = (1-\theta)|x| + \theta|y|$ because $\theta, 1-\theta > 0$.

(b) for all $x, y \in V$ and $\theta \in (0,1)$ we have $\|(1-\theta)x + \theta y\| \leq \|(1-\theta)x\| + \|\theta y\| = (1-\theta)\|x\| + \theta\|y\|$ because $\theta, 1-\theta > 0$.

(c) By computing its first and second derivatives we can verify that the function $\phi : [0,\infty) \to \mathbb{R}$ given by $\phi(x) = x^m$ is increasing and convex if $m \geq 1$. Since the norm $x \in V \mapsto \|x\|$ is a convex function, it follows that $f(x) = \|x\|^m$ also is convex.

(d) By computing the second derivative of $\phi(x) = \sqrt{1+x^2}$ we see that $\phi : [0,\infty) \to \mathbb{R}$ is increasing and convex. Since the norm $x \in V \mapsto \|x\|$ is convex it follows that $f(x) = \sqrt{1+\|x\|^2}$ also is convex.

6.3. Let $V = C^1(\bar{\mathcal{R}})$ be the real vector space of all C^1 functions defined on $\bar{\mathcal{R}}$. For any given function $g : \partial\mathcal{R} \to \mathbb{R}$ on the boundary of $\mathcal{R}$ we consider the set

$$V_g = \{u \in V : u = g \text{ on } \partial\mathcal{R}\},$$

i.e. the set of all $u \in V$ that have g as boundary value. Show that V_g is a convex subset of V.

6.4. Minimal surface in polar coordinates. If a surface is the graph of a function given in polar coordinates by $z = v(r,\theta)$, then the area of its graph is given by

$$A[v] = \iint_{\mathcal{R}} r\sqrt{1 + v_r^2 + r^{-2}v_\theta^2}\, dr\, d\theta$$

which one can also write as

$$A[v] = \iint_{\mathcal{R}} \sqrt{r^2 + r^2 v_r^2 + v_\theta^2}\, dr\, d\theta.$$

(a) Find the Euler–Lagrange equation for minmizers of the integral $A[u]$
(b) Suppose the function v does not depend on r, i.e. suppose $v(r,\theta) = F(\theta)$ for some C^2 function $F : \mathbb{R} \to \mathbb{R}$. What is the function F?

Solution. (a) The EL equation is

$$\frac{\partial}{\partial r}\left(\frac{r^2 v_r}{\sqrt{r^2 + r^2 v_r^2 + v_\theta^2}}\right) + \frac{\partial}{\partial \theta}\left(\frac{v_\theta}{\sqrt{r^2 + r^2 v_r^2 + v_\theta^2}}\right) = 0.$$

(b) If $v(r,\theta) = F(\theta)$ is a solution of the EL equation, then $v_r = 0$, $v_\theta = F'(\theta)$, and thus F is a solution of

$$\frac{\partial}{\partial \theta}\left(\frac{F'(\theta)}{\sqrt{r^2 + F'(\theta)^2}}\right) = 0.$$

This implies that there is a function $G(r)$ such that

$$\frac{F'(\theta)}{\sqrt{r^2 + F'(\theta)^2}} = G(r)$$

for all r, θ. Solving for $F'(\theta)$ we get

$$F'(\theta) = \frac{rG(r)}{\sqrt{1 - G(r)^2}}.$$

Here the right hand side does not depend on r, so there is a constant $C \in \mathbb{R}$ such that for all r

$$\frac{rG(r)}{\sqrt{1 - G(r)^2}} = C$$

which implies

$$v(r,\theta) = F(\theta) = C\theta.$$

We have shown that if there is a solution of the form $v(r,\theta) = F(\theta)$, then is must be given by $v = C\theta$ for some constant C. By direct substitution in the EL-equation we verify that $v = C\theta$ is indeed always a solution to the EL-equation.

The resulting surface (graph of $v(r,\theta) = C\theta$ in cylindrical coordinates) is known as the helicoid. (Google "helicoid" for images.)

6.5. A nonlinear differential equation.

(a) Find the Euler–Lagrange equation for the minimizers of

$$I[u] = \int_0^1 \left\{\frac{1}{2}u'(x)^2 + xe^{u(x)}\right\} dx$$

(b) Show that for all $x > 0$ $L(x,u,v) = \frac{1}{2}v^2 + xe^u$ is a strictly convex function of $(u,v) \in \mathbb{R}^2$.

(c) Show there is at most one C^2 function $u : [0,1] \to \mathbb{R}$ that satisfies the ordinary differential equation $u''(x) = xe^{u(x)}$ (for $0 < x < 1$), and for which $u(0) = u(1) = 0$.

Solution. (a) The EL-equation is an ordinary differential equation:

$$u''(x) + xe^{u(x)} = 0.$$

(b) We compute the Hessian matrix of the function $(u,v) \mapsto L(x,u,v)$, treating x as a constant: since $L_{uu} = xe^u$, $L_{uv} = L_{vu} = 0$, and $L_{vv} = 1$ we have

$$D^2L(x,u,v) = \begin{pmatrix} xe^u & 0 \\ 0 & 1 \end{pmatrix}$$

The eigenvalues of this matrix are xe^u and 1, both of which are positive for all $x > 0$, so the matrix is positive definite. It follows that $L(x,u,v)$ is a strictly convex function of (u,v).

Alternative argument: For any $h,k \in \mathbb{R}$ we consider the function $\phi(t) = L(x, u+th, v+tk)$ and show that it is strictly convex by computing its second derivative with respect to t at $t = 0$. We find

$$\begin{aligned}\phi'(t) &= \frac{d}{dt}\left(\tfrac{1}{2}(v+tk)^2 + xe^{u+th}\right) \\ &= (v+tk)k + xe^{u+th}h\end{aligned}$$

and thus

$$\phi''(t) = k^2 + xe^{u+th}h^2 \implies \phi''(0) = k^2 + xe^u h^2.$$

If $x > 0$ and if $(h,k) \neq (0,0)$ then we see that $\phi''(0) > 0$, which implies that $L(x,u,v)$ is a strictly convex function of (u,v).

(c) Any C^2 solution of the EL-equation is a minimizer of $I[u]$ (this was shown in section 5.5). If there are two C^2 solutions $u, \bar{u}$ then they are both minimizers, but since $L(x,u,v)$ is a strictly convex function of (u,v) it follows from section 5.4 that $u = \bar{u}$. Thus there cannot be more than one solution to the EL-equation that satisfies the boundary conditions $u(0) = u(1) = 0$.

CHAPTER 5

Analyzing Laplace's equation

1. Laplace's equation on a disc

If the domain $\mathcal{R}$ is the unit disc, i.e.

$$\mathcal{R} = \{(x, y) \mid x^2 + y^2 < 1\}$$

then there is an explicit representation of the solution of

$$\Delta u = 0 \text{ on } \mathcal{R}, \quad u = g \text{ on } \partial\mathcal{R}$$

where $g : \partial\mathcal{R} \to \mathbb{R}$ is a prescribed continuous function on the boundary. This explicit solution is given in polar coordinates by

$$u(r, \theta) = \int_{-\pi}^{\pi} P(r, \theta - \phi) g(\phi)\, d\phi, \tag{27}$$

where

$$P(r, \theta) = \frac{1}{2\pi} \frac{1 - r^2}{1 - 2r\cos\theta + r^2}. \tag{28}$$

In this section we prove that u is a harmonic function that equals g on the boundary, assuming only that g is continuous.

1.1. Laplace's equation in Polar Coordinates. To any point $(x, y) \in \mathcal{R}$ we can associate polar coordinates r, θ that are related to x, y by

$$x = r\cos\theta, \quad y = r\sin\theta.$$

If we have a function $f : \mathcal{R} \to \mathbb{C}$ then its representation in polar coordinate is the function

$$g(r, \theta) = f(r\cos\theta, r\sin\theta). \tag{29}$$

The function g is defined for $0 \leq r < 1$ and all $\theta \in \mathbb{R}$. It is 2π periodic in θ.

Conversely, if we have a function $g : [0, 1) \times \mathbb{R} \to \mathbb{C}$ with $g(r, \theta + 2\pi) = g(r, \theta)$ for all r, θ then there is a function $f : \mathcal{R} \to \mathbb{C}$ with $f(r\cos\theta, r\sin\theta) = g(r, \theta)$.

If $g(r, \theta)$ is the representation of $f : \mathcal{R} \to \mathbb{C}$ in polar coordinates, so that they are related by (29), then

$$f_{xx} + f_{yy} = g_{rr} + \frac{1}{r} g_r + \frac{1}{r^2} g_{\theta\theta}. \tag{30}$$

To verify this one differentiates the relation (29) twice with respect to r and twice with respect to θ, and substitutes the resulting expressions for $g_{rr}, g_r, g_{\theta\theta}$ in (30).

1.2. Lemma. For all $r \in [0,1)$ the Poisson kernel satisfies

$$P(r,\theta) = \frac{1}{2\pi}\sum_{-\infty}^{\infty} r^{|n|}e^{in\theta} \tag{31}$$

The series converges uniformly for $0 \le r < 1-\delta$ for any $\delta > 0$.

Proof. We write

$$2\pi P(r,\theta) = \sum_{-\infty}^{-1} r^{-n}e^{in\theta} + 1 + \sum_{1}^{\infty} r^n e^{in\theta}.$$

If $r \le 1-\delta$ then the terms in the second series are bounded by

$$|r^n e^{in\theta}| = r^n \le (1-\delta)^n.$$

The series $\sum_1^\infty (1-\delta)^n$ is a geometric series, which converges, with $\sum_1^\infty (1-\delta)^n = \frac{1-\delta}{\delta}$. Hence the Weierstrass M-test implies that $\sum_1^\infty r^n e^{in\theta}$ converges uniformly for $r \le 1-\delta$. Similar arguments apply to the first sum.

We compute the sum

$$\begin{aligned}
2\pi P(r,\theta) &= \sum_{-\infty}^{-1} r^{-n}e^{in\theta} + 1 + \sum_{1}^{\infty} r^n e^{in\theta} \\
&= \sum_{1}^{\infty} r^{n}e^{-in\theta} + 1 + \sum_{1}^{\infty} r^n e^{in\theta} \\
&= \sum_{1}^{\infty} \left(re^{-i\theta}\right)^n + 1 + \sum_{1}^{\infty} \left(re^{i\theta}\right)^n \\
&= \frac{re^{-i\theta}}{1-re^{-i\theta}} + 1 + \frac{re^{i\theta}}{1-re^{i\theta}} \\
&= 1 + \frac{2r\cos\theta + 2r^2}{1-2r\cos\theta + r^2} \\
&= \frac{1-r^2}{1-2r\cos\theta + r^2}.
\end{aligned}$$

Q.E.D.

1.3. Lemma.

$$\int_{-\pi}^{\pi} P(r,\theta)d\theta = 1.$$

Proof. The series (31) for $P(r,\theta)$ converges uniformly, so we have

$$\int_{-\pi}^{\pi} P(r,\theta)d\theta = \sum_{-\infty}^{\infty} \frac{1}{2\pi} r^{|n|} \int_{-\pi}^{\pi} e^{in\theta}d\theta.$$

The only nonzero term in the series is the one with $n=0$, so we find

$$\int_{-\pi}^{\pi} P(r,\theta)d\theta = \frac{1}{2\pi}r^0 2\pi = 1.$$

Q.E.D.

1.4. Lemma. The function $P(r,\theta)$ is harmonic, i.e. if $u : \mathcal{R} \to \mathbb{R}$ is the function on the unit disc that satisfies

$$u(r\cos\theta, r\sin\theta) = P(r,\theta)$$

then u is C^2 and $u_{xx} + u_{yy} = 0$.

Proof. Since $re^{\pm i\theta} = x \pm iy$ we can write the Poisson kernel in Cartesian coordinates as

$$P(r,\theta) = u(x,y) = 1 + \sum_1^\infty (x+iy)^n + \sum_1^\infty (x-iy)^n,$$

i.e. $u(x,y) = \lim_{N\to\infty} u_N(x,y)$ where

$$u_N(x,y) = 1 + \sum_1^N (x+iy)^n + \sum_1^N (x-iy)^n.$$

Each term $(x \pm iy)^n$ is harmonic, because

$$\frac{\partial (x+iy)^n}{\partial x} = n(x+iy)^{n-1}, \qquad \frac{\partial^2 (x+iy)^n}{\partial x^2} = n(n-1)(x+iy)^{n-2}$$
$$\frac{\partial (x+iy)^n}{\partial y} = in(x+iy)^{n-1}, \qquad \frac{\partial^2 (x+iy)^n}{\partial y^2} = i^2 n(n-1)(x+iy)^{n-2}$$
$$= -n(n-1)(x+iy)^{n-2}$$

implies

$$\frac{\partial^2 (x+iy)^n}{\partial x^2} + \frac{\partial^2 (x+iy)^n}{\partial y^2} = 0,$$

and similarly for $(x-iy)^n$.

This implies that u_N is harmonic. to show that the limit $u = \lim_{N\to\infty} u_N$ is harmonic we must show that the partial derivatives $(u_N)_{xx}$ and $(u_N)_{yy}$ converge uniformly for $|x+iy| \le 1-\delta$ for any $\delta > 0$. Since $(u_N)_{xx}$ is given by

$$\frac{\partial^2 u_N}{\partial x^2} = \sum_1^N n(n-1)(x+iy)^{n-2} + n(n-1)(x-iy)^{n-2}$$

we can use the Weierstrass M-test: assume $|x+iy| \le 1-\delta$, then

$$\left| n(n-1)(x+iy)^{n-2} + n(n-1)(x-iy)^{n-2} \right| \le n(n-1)\left(|x+iy|^{n-2} + |x-iy|^{n-2} \right)$$
$$\le 2n(n-1)(1-\delta)^{n-2} \stackrel{\text{def}}{=} M_n.$$

The series $\sum_1^\infty M_n$ converges, and therefore we are allowed to differentiate the serie for u term by term to get

$$\frac{\partial^2 u}{\partial x^2} = \sum_1^\infty n(n-1)(x+iy)^{n-2} + n(n-1)(x-iy)^{n-2}.$$

Similarly one has

$$\frac{\partial^2 u}{\partial y^2} = -\sum_1^\infty n(n-1)(x+iy)^{n-2} + n(n-1)(x-iy)^{n-2},$$

which implies $u_{xx} + u_{yy} = 0$ on $\mathcal{R}$. *Q.E.D.*

1.5. Lemma. If $g : \mathbb{R} \to \mathbb{C}$ is a continuous 2π-periodic function, then the function $u : \mathcal{R} \to \mathbb{C}$ given by (27) is harmonic.

Proof. The representation of u in polar coordinates is

$$v(r,\theta) = u(r\cos\theta, r\sin\theta) = \int_{-\pi}^{\pi} P(r,\theta-\phi)g(\phi)\,d\phi.$$

Therefore the Laplacian of u in polar coordinates is given by

$$v_{rr}+\frac{1}{r}v_r+\frac{1}{r^2}v_{\theta\theta} = \int_{-\pi}^{\pi}\left\{\frac{\partial^2 P(r,\theta-\phi)}{\partial r^2}+\frac{1}{r}\frac{\partial P(r,\theta-\phi)}{\partial r}+\frac{1}{r^2}\frac{\partial^2 P(r,\theta-\phi)}{\partial \theta^2}\right\}g(\phi)d\phi.$$

Since the Poisson kernel is harmonic we have

$$\frac{\partial^2 P(r,\theta-\phi)}{\partial r^2}+\frac{1}{r}\frac{\partial P(r,\theta-\phi)}{\partial r}+\frac{1}{r^2}\frac{\partial^2 P(r,\theta-\phi)}{\partial \theta^2}=0,$$

which implies that $v_{rr}+\frac{1}{r}v_r+\frac{1}{r^2}v_{\theta\theta}=0$, i.e. u is harmonic. *Q.E.D.*

1.6. Lemma. If $g : \mathbb{R} \to \mathbb{C}$ is a continuous 2π-periodic function, then

$$u(r,\theta) = \int_{-\pi}^{\pi} P(r,\theta-\phi)g(\phi)\,d\phi$$

satisfies

$$\lim_{r\nearrow 1} u(r,\theta) = g(\theta).$$

The convergence is uniform in θ.

Proof. Let $\epsilon > 0$ be given.

Since $\int_{-\pi}^{\pi} P(r,\phi)d\phi = 1$ we have, substituting $\phi = \theta - \varphi$,

$$u(r,\theta)-g(\theta) = \int_{-\pi}^{\pi} P(r,\theta-\phi)\big\{g(\phi)-g(\theta)\big\}\,d\phi = \int_{-\pi}^{\pi} P(r,\varphi)\big\{g(\theta-\varphi)-g(\theta)\big\}\,d\varphi.$$

For any $\delta > 0$ we split the integral in two parts:

$$|u(r,\theta)-g(\theta)| \le \int_{|\varphi|<\delta} P(r,\varphi)\big|g(\theta-\varphi)-g(\theta)\big|\,d\varphi+\int_{\delta<|\varphi|<\pi} P(r,\varphi)\big|g(\theta-\varphi)-g(\theta)\big|\,d\varphi.$$

From the explicit expression for $P(r,\varphi)$ we see that $0 < P(r,\varphi) \le P(r,\delta)$ if $\delta \le |\varphi| \le \pi$. The function g is continuous and hence bounded, so we have

$$|g(\theta-\varphi)-g(\theta)| \le 2\|g\|_\infty \text{ for all } \theta,\varphi,$$

and we can estimate the second integral by

$$\int_{\delta<|\varphi|<\pi} P(r,\varphi)\big|g(\theta-\varphi)-g(\theta)\big|\,d\varphi \le P(r,\delta)\,2\|g\|_\infty \int_{\delta<\phi<\pi} d\phi.$$

Since $\int_{\delta<|\phi|<\pi} d\phi = 2\pi - 2\delta < 2\pi$, we get

$$\int_{\delta<|\varphi|<\pi} P(r,\varphi)\big|g(\theta-\varphi)-g(\theta)\big|\,d\varphi \le 4\pi P(r,\delta)\|g\|_\infty.$$

It follows from $\lim_{r\nearrow 1} P(r,\delta) = 0$ that for our given $\epsilon > 0$ there is an $r_\epsilon \in (0,1)$ such that

$$P(r,\delta) < \frac{\epsilon}{8\pi\|g\|_\infty} \text{ for all } r \in (r_\epsilon, 1).$$

Since g is continuous it is also uniformly continuous, and thus there exists a $\delta > 0$ such that

$$|g(\theta - \varphi) - g(\theta)| < \frac{\epsilon}{2} \text{ for all } \theta, \varphi \in \mathbb{R} \text{ with } |\varphi| < \delta.$$

It follows that we can bound the first integral by

$$\int_{|\varphi|<\delta} P(r,\varphi)\big|g(\theta-\varphi) - g(\theta)\big|\,d\varphi \le \frac{\epsilon}{2}\int_{|\varphi|<\delta} P(r,\varphi)\,d\varphi < \frac{\epsilon}{2}.$$

Combine the bounds for the two integrals and we find that if $r_\epsilon < r < 1$ we have

$$|u(r,\theta) - g(\theta)| < \frac{\epsilon}{8\pi\|g\|_\infty}\cdot 4\pi\|g\|_\infty + \frac{\epsilon}{2} = \epsilon$$

for all $\theta \in \mathbb{R}$. *Q.E.D.*

2. Problems

2.1. The Poisson kernel in cartesian coordinates. Consider the vector $\boldsymbol{x} = \binom{r\cos\theta}{r\sin\theta}$, with $r < 1$, and let $\boldsymbol{e}_1 = \binom{1}{0}$. Show that

$$P(r,\theta) = \frac{1 - \|\boldsymbol{x}\|^2}{\|\boldsymbol{x} - \boldsymbol{e}_1\|^2}$$

3. Minimizers of the Dirichlet integral and the need for generalized functions

One strategy to construct a solution $u : \mathcal{R} \to \mathbb{R}$ of

$$\Delta u = 0 \text{ on } \mathcal{R}, \qquad u = g \text{ on } \partial\mathcal{R}$$

for some given function $g : \partial\mathcal{R} \to \mathbb{R}$ is to show find a function $u : \mathcal{R} \to \mathbb{R}$ that minimizes the Dirichlet integral

$$\mathcal{D}[u] = \frac{1}{2}\iint_{\mathcal{R}} |\nabla u(x)|^2\,dx_1 \ldots dx_n$$

among all functions with $u = g$ on $\partial\mathcal{R}$. In this section we discuss the problem of finding such a minimizer.

The following notation will be convenient:

$$C^1_g \stackrel{\text{def}}{=} \big\{u \mid u : \mathcal{R} \to \mathbb{R} \text{ is } C^1, u = g \text{ on } \partial\mathcal{R}\big\}.$$

3.1. Minimizing sequences. Since $\mathcal{D}[u] \ge 0$ for any function u, there is a greatest lower bound for the values $\mathcal{D}[u]$ can have if $u = g$ on $\partial\mathcal{R}$. We denote this number by

$$D_- \stackrel{\text{def}}{=} \inf\big\{\mathcal{D}[u] \mid u \in C^1_g\big\}.$$

If we can find a function with $u \in C^1_g$ for which $\mathcal{D}[u] = D_-$ then we are done, because u is a minimizer. The definition of "infimum" implies that for each $k \in \mathbb{N}$ there is a C^1 function $u_k \in C^1_g$ for which

$$D_- \le \mathcal{D}[u_k] \le D_- + \frac{1}{k}.$$

This implies

$$\lim_{k\to\infty} \mathcal{D}[u_k] = D_-.$$

Any sequence of functions with this property is called a minimizing sequence.

3.2. A minimizing sequence is a Cauchy sequence. Let $u, v :\in C^1_g$ be functions with

$$\mathcal{D}[u], \mathcal{D}[v] \le D_- + \epsilon.$$

Then

$$\iint_{\mathcal{R}} |\nabla u - \nabla v|^2 \, dx_1 \dots dx_n \le 2\epsilon. \tag{32}$$

If $u_k \in C^1_g$ is a sequence of functions with $\mathcal{D}[u_k] \to D_-$ then

$$\iint_{\mathcal{R}} |\nabla u_k - \nabla u_l|^2 \, dx_1 \dots dx_n \le \frac{2}{k}$$

for all $l \ge k \ge 1$.

Proof. By expanding squares one finds

$$|\nabla u - \nabla v|^2 + |\nabla u + \nabla v|^2 = 2|\nabla u|^2 + 2|\nabla v|^2$$

and thus

$$\left|\frac{\nabla u - \nabla v}{2}\right|^2 + \left|\frac{\nabla u + \nabla v}{2}\right|^2 = \frac{1}{2}\big(|\nabla u|^2 + |\nabla v|^2\big)$$

Integrate over $\mathcal{R}$:

$$\begin{aligned}\iint_{\mathcal{R}} \left|\frac{\nabla u - \nabla v}{2}\right|^2 dx_1 \dots dx_n &= \iint_{\mathcal{R}} \tfrac{1}{2}|\nabla u|^2 + \tfrac{1}{2}|\nabla v|^2 dx - \iint_{\mathcal{R}} \left|\frac{\nabla u + \nabla v}{2}\right|^2 dx \\ &= \mathcal{D}[u] + \mathcal{D}[v] - 2\mathcal{D}\left[\tfrac{u+v}{2}\right] \\ &\le \mathcal{D}[u] + \mathcal{D}[v] - 2D_- \\ &\le 2\epsilon.\end{aligned}$$

Q.E.D.

3.3. Poincaré's inequality. If $\mathcal{R}$ is a bounded domain then there is a constant $C_{\mathcal{R}}$ such that for all C^1 functions $u : \mathcal{R} \to \mathbb{R}$ with $u = 0$ on $\partial\mathcal{R}$ one has

$$\iint_{\mathcal{R}} u^2 \, dx \le C_{\mathcal{R}} \iint_{\mathcal{R}} |\nabla u|^2 \, dx.$$

Proof.

$$u(x)^2 = \int_0^x 2u(s)u'(s)ds$$

implies

$$\int_0^\ell u(x)^2 dx = \int_0^\ell \int_0^x 2u(s)u'(s) \, ds \, dx = \int_0^\ell \int_s^\ell 2u(s)u'(s) \, dx \, ds$$

Q.E.D.

3.4. A minimizing sequence is a Cauchy sequence in L^2. It follows from the Poincaré inequality that

$$\int_{\mathcal{R}} \big(u_k - u_\ell\big)^2 dx \le C_{\mathcal{R}} \int_{\mathcal{R}} |\nabla u_k - \nabla u_\ell|^2 dx \le \frac{2C_{\mathcal{R}}}{k}$$

for all $\ell \ge k \ge 1$.

4. Generalized functions and Distributions

Let $\mathcal{R} \subset \mathbb{R}^n$ be an open set. Ordinary functions $f : \mathcal{R} \to \mathbb{R}$ are defined by specifying the value $f(x)$ at each point $x \in \mathcal{R}$. If the function f is continuous then it is completely determined once one knows all the possible averages

$$T_f[\varphi] = \int_{\mathcal{R}} f(x)\varphi(x)\,dx.$$

The notion of distribution generalizes that of ordinary function by no longer requiring that $f(x)$ be defined for all x or even for any x, but instead by requiring that for any "test function" φ the average $T_f[\varphi]$ is defined.

The following definitions make these ideas more precise.

4.1. Test functions. A function $\varphi : \mathcal{R} \to \mathbb{R}$ is called a test function if

(1) φ is infinitely often differentiable, and
(2) there exists a compact set $K \subset \mathcal{R}$ with $\varphi(x) = 0$ for all $x \in \mathcal{R} \setminus K$.

We define $C_c^\infty(\mathcal{R})$ to be the set of all test functions on $\mathcal{R}$.

4.2. Lemma. $C_c^\infty(\mathcal{R})$ is a vector space. For every $p \in \mathcal{R}$ and $\epsilon > 0$ with $\overline{B_\epsilon(p)} \subset \mathcal{R}$ there is a test function φ with $\varphi(x) > 0 \iff x \in B_\epsilon(p)$

Proof. A possible choice for the function φ is

$$\varphi(x) = \Phi\Big(\frac{|x-p|}{\epsilon}\Big) \text{ where } \Phi(t) \stackrel{\text{def}}{=} \begin{cases} e^{-1/(1-t^2)} & 0 \le t < 1 \\ 0 & t \ge 1 \end{cases}$$

We omit the proof that φ is C^∞. *Q.E.D.*

4.3. Distributions on $\mathcal{R}$. A distribution of order m on $\mathcal{R}$ is a linear map $T : C_c^\infty(\mathcal{R}) \to \mathbb{R}$ with the property that there exists a $C > 0$ such that for all test functions φ one has

(33) $$\big|T[\varphi]\big| \le C \sup_{i_1+\cdots+i_n \le m} \sup_{x\in\mathcal{R}} \left| \frac{\partial^{i_1+\cdots+i_m}\varphi}{(\partial x_1)^{i_1}\cdots(\partial x_n)^{i_n}}(x) \right|$$

4.4. Example — ordinary functions as distributions. Let $f : \mathcal{R} \to \mathbb{R}$ be a bounded Riemann integrable function. Then

$$T_f[\varphi] \stackrel{\text{def}}{=} \int_{\mathcal{R}} f(x)\varphi(x)\,dx$$

defines a distribution. Two different Riemann integrable functions f and g can end up defining the same distribution, but if $f, g : \mathcal{R} \to \mathbb{R}$ are continuous at a point $p \in \mathcal{R}$, and if $T_f = T_g$ then $f(p) = g(p)$.

4.5. Example — Dirac's δ-function. Let $n = 1$, and $\mathcal{R} = (-1, 1)$. Then

$$\delta[\varphi] = \varphi(0)$$

defines a distribution on $\mathcal{R}$. There is no Riemann integrable function $f : (-1, 1) \to \mathbb{R}$ such that $\delta = T_f$.

4.6. The derivative of a distribution. By definition the partial derivative of a distribution $T : C_c^\infty(\mathcal{R}) \to \mathbb{R}$ with respect to x_i is again a distribution. It is defined by

$$D_i T[\varphi] \stackrel{\text{def}}{=} -T\Big[\frac{\partial\varphi}{\partial x_i}\Big] \tag{34}$$

Unlike ordinary functions, distributions can always be differentiated.

This definition is consistent with the definition of derivatives for ordinary functions, namely, if $f : \mathcal{R} \to \mathbb{R}$ is a C^1 function, then

$$D_i T_f = T_{\partial f/\partial x_i}.$$

4.7. Example — derivative of the Heaviside function. The Heaviside function is defined by

$$H(x) = \begin{cases} 1 & x > 0 \\ 0 & x < 0 \end{cases}$$

It defines a distribution T_H on $\mathbb{R}$ by

$$T_H[\varphi] \stackrel{\text{def}}{=} \int_\infty^\infty H(x)\varphi(x)dx = \int_0^\infty \varphi(x)dx.$$

The Heaviside function is constant on all of $\mathbb{R}$ except at $x = 0$ where it has a jump discontinuity. The derivative of H is therefore zero at all $x \neq 0$, and the Heaviside function is not differentiable at $x = 0$. The derivative of the distribution T_H is well defined. To see how it is defined let $\varphi \in C_c^\infty(\mathbb{R})$ by a test function. Thus $\varphi : \mathbb{R} \to \mathbb{R}$ is infinitely often differentiable, and there is an $L > 0$ such that $\varphi(x) = 0$ when $|x| \geq L$. Then

$$\begin{aligned}
(DT_H)[\varphi] &= -T_H[\varphi'] && \text{definition of } DT_H \\
&= -\int_0^\infty \varphi'(x)dx && \text{def of } T_H \\
&= -\int_0^L \varphi'(x)dx && \varphi(x) = 0 \text{ for } x \geq L \\
&= -\big(\varphi(L) - \varphi(0)\big) && \text{Fundamental Thm of calculus} \\
&= \varphi(0).
\end{aligned}$$

By definition of Dirac's δ-function $\delta[\varphi] = \varphi(0)$ for all test functions φ. Hence $(DT_H)[\varphi] = \delta[\varphi]$ for all test functions, which means that DT_H and δ are the same distribution: we have shown that

$$DT_H = \delta.$$

The derivative of the Heaviside function in the sense of distributions is Dirac's δ-function.

4.8. Example. Let $\mathcal{R} = (-1, 1)$ and consider the function

$$f(x) = \begin{cases} e^{-x} & 0 < x < 1 \\ 0 & -1 < x < 0 \end{cases}$$

Let T_f be the corresponding distribution. We show that

$$DT_f + T_f = \delta,$$

where δ is Dirac's δ-distribution (see § 4.5 above.)

To prove that the two distributions $DT_f + T_f$ and δ are equal we must show that

$$\big(DT_f + T_f\big)[\varphi] = \delta[\varphi] \text{ for all } \varphi \in C_c^\infty(\mathcal{R}).$$

By definition

$$\delta[\varphi] = \varphi(0).$$

On the other hand we have

$$\begin{aligned}
\big(DT_f + T_f\big)[\varphi] &= (DT_f)[\varphi] + (T_f)[\varphi] && \text{definition of } DT_f + T_f \\
&= -T_f[\varphi'] + T_f[\varphi] && \text{definition of } DT_f \\
&= T_f[-\varphi' + \varphi] && T_f \text{ is linear} \\
&= \int_{-1}^{1} f(x)\big(-\varphi'(x) + \varphi(x)\big)dx && \text{def. of } T_f
\end{aligned}$$

The function $f(x)$ vanishes for $x < 0$ and equals e^{-x} for $x > 0$, so the integral is

$$\big(DT_f + T_f\big)[\varphi] = \int_{-1}^{1} f(x)\big(-\varphi'(x) + \varphi(x)\big)dx = \int_0^1 e^{-x}\big(-\varphi'(x) + \varphi(x)\big)dx. \tag{35}$$

Now integrate by parts

$$\begin{aligned}
\int_0^1 e^{-x}\varphi'(x)dx &= \big[e^{-x}\varphi(x)\big]_{x=0}^{1} - \int_0^1 \frac{de^{-x}}{dx}\varphi(x)dx \\
&= e^{-1}\varphi(1) - \varphi(0) + \int_0^1 e^{-x}\varphi(x)dx
\end{aligned}$$

Recall that $\varphi \in C_c^\infty(\mathcal{R})$ is a test function in the interval $\mathcal{R} = (-1, 1)$, so that $\varphi(1) = 0$. It follows that

$$\int_0^1 e^{-x}\varphi'(x)dx = -\varphi(0) + \int_0^1 e^{-x}\varphi(x)dx.$$

Apply this to equation (35):

$$\big(DT_f + T_f\big)[\varphi] = \int_0^1 e^{-x}\big(-\varphi'(x) + \varphi(x)\big)dx = \varphi(0).$$

Thus $\big(DT_f + T_f\big)[\varphi] = \delta[\varphi]$ for all test functions $\varphi \in C_c^\infty(\mathcal{R})$, and we have shown that $DT_f + T_f = \delta$.

4.9. Harmonic distributions. A C^2 function $f : \mathcal{R} \to \mathbb{R}$ is called harmonic if $\Delta f = 0$. Similarly a distribution $T : C_c^\infty(\mathcal{R}) \to \mathbb{R}$ is harmonic if

$$D_1(D_1T) + \cdots + D_n(D_nT) = 0.$$

5. Minimizing the Dirichlet integral

5.1. A minimizing sequence converges in the sense of distributions. Let $u_k \in C_g^1$ be a minimizing sequence of the Dirichlet integral $\mathfrak{D}[u_k]$, as in § 3.1. Then we have shown in § 3.4 that u_k is a Cauchy sequence in the L^2 norm. If $\varphi \in C_c^\infty(\mathcal{R})$ is a test function then the sequence of numbers

$$a_k = \int_{\mathcal{R}} u_k(x)\varphi(x)\,dx$$

is a Cauchy sequence of real numbers. Indeed, we have

$$\begin{aligned}
|a_k - a_\ell| &= \left| \int_{\mathcal{R}} \big(u_k(x) - u_\ell(x)\big)\varphi(x)\,dx \right| \\
&\le \int_{\mathcal{R}} |u_k(x) - u_\ell(x)|\varphi(x)\,dx \\
&\le \left(\int_{\mathcal{R}} (u_k - u_\ell)^2 dx\right)^{1/2} \left(\int_{\mathcal{R}} \varphi(x)^2 dx\right)^{1/2} && \text{(Cauchy-Schwarz } \le\text{)} \\
&\le \frac{2C_{\mathcal{R}}}{k} \|\varphi\|_2 && \text{(by § 3.4)}
\end{aligned}$$

for all $\ell \ge k \ge 1$. Every Cauchy sequence of real numbers converges ("$\mathbb{R}$ is complete"), so $\lim a_k$ exists. We can therefore define

$$T_\infty[\varphi] = \lim_{k\to\infty} \int_{\mathcal{R}} u_k(x)\varphi(x)\,dx.$$

5.2. Lemma. T_∞ is a distribution.

To prove this one has to verify linearity, i.e. $T_\infty[c_1\varphi_1 + c_2\varphi_2] = c_1T_\infty[\varphi_1] + c_2T_\infty[\varphi_2]$, and the condition (33). This was done in lecture

5.3. Theorem. T_∞ is a harmonic distribution.

A proof was given in lecture.

6. The maximum principle

If $u : [a, b] \to \mathbb{R}$ is a C^2 function with $u''(x) \ge 0$ for all x, then the graph of u is convex, and u cannot attain it maximum in the interior of the interval $[a, b]$, i.e. u must attain its maximum at a or b. The maximum principle generalizes this fact to solutions of certain partial differential equations.

6.1. Assumptions in this section. Let $\mathcal{R} \subset \mathbb{R}^n$ be a bounded open set, and let $c : \bar{\mathcal{R}} \to \mathbb{R}$ be a continuous function. We consider functions that satisfy the inequality

$$\Delta u(x) - c(x)u(x) \ge 0 \tag{36}$$

for all $x \in \mathcal{R}$. This includes all functions that satisfy the partial differential equation

$$\Delta u(x) - c(x)u(x) = 0 \quad (x \in \mathcal{R}).$$

The maximum principle is useful because it allows one to compare solutions to the PDE $\Delta u - c(x)u = 0$ with other functions that satisfy the inequality (36).

6.2. Theorem (maximum principle). Assume $c(x) \geq 0$ for all $x \in \mathcal{R}$, and assume also that $u : \bar{\mathcal{R}} \to \mathbb{R}$ is a continuous function that is C^2 on $\mathcal{R}$. Let M be the maximum value of u on the boundary of the domain, i.e.

$$M = \max_{y \in \partial\mathcal{R}} u(y).$$

If u satisfies (36), then $u(x) \leq \max\{M, 0\}$ for all $x \in \mathcal{R}$.

A different formulation of the conclusion of this theorem is to say that u does not attain an interior positive maximum.

A proof was given in lecture.

6.3. A special case of the maximum principle. Let $u : \bar{\mathcal{R}} \to \mathbb{R}$ be a continuous function that is C^2 in the interior $\mathcal{R}$. If u satisfies $\Delta u(x) - c(x)u(x) \geq 0$ on $\mathcal{R}$, and if $u(x) \leq 0$ for all $x \in \partial\mathcal{R}$, then $u(x) \leq 0$ for all $x \in \mathcal{R}$. In other words, if $u \leq 0$ on the boundary, then $u \leq 0$ on the whole domain.

This follows immediately from the maximum principle in § 6.2, because the assumption that $u \leq 0$ on $\partial\mathcal{R}$ implies that $M \leq 0$.

6.4. An example. Let $u : (0, \infty) \to \mathbb{R}$ be a solution of

$$\frac{d^2u}{dx^2} = (2 + \sin x)u$$

and assume that $u(x) \leq 1$ for all $x > 0$. Using the maximum principle we will find a number $\theta \in (0, 1)$ and show that $u(x) \leq \theta$ for all $x \geq 1$.

From here on abbreviate $c(x) = 2 + \sin x$, and note that $c(x) \geq 1 \geq 0$ for all $x \in \mathbb{R}$.

In our proof we first let a be any number with $a > 1$ and find a function $v : (a-1, a+1) \to \mathbb{R}$ that satisfies

$$v'' - c(x)v \leq 0 \text{ for all } x \in (a-1, a+1)$$

and

$$v(a-1) = v(a+1) = 1.$$

We try a function of the form $v(x) = 1 - \alpha + \alpha(x-a)^2$ with $0 < \alpha < 1$. Substituting leads to

$$\begin{aligned} v''(x) - c(x)v(x) &= +2\alpha - c(x)\big(1 - \alpha + \alpha(x-a)^2\big) \\ &= 2\alpha - (1-\alpha)c(x) - \alpha c(x)(x-a)^2 && \text{(use } c(x) \geq 0\text{)} \\ &\leq 2\alpha - (1-\alpha)c(x) && \text{(use } c(x) \geq 1\text{)} \\ &\leq 2\alpha - (1-\alpha) \\ &= 3\alpha - 1. \end{aligned}$$

If we choose $\alpha = \frac{1}{3}$ then we have shown that

$$v(x) = \frac{2}{3} + \frac{1}{3}(x-a)^2$$

satisfies $v'' - c(x)v(x) \leq 0$ on $(a-1, a+1)$, and $v(a \pm 1) = 1$.

Now that we have the function v we can apply the maximum principle to

$$w(x) = u(x) - v(x)$$

on the domain $\mathcal{R} = (a-1, a+1)$. The function w satisfies

$$w''(x) - c(x)w(x) = \underbrace{u''(x) - c(x)u(x)}_{=0} - \{\underbrace{v''(x) - c(x)v(x)}_{\leq 0}\} \geq 0$$

for all $x \in (a-1, a+1)$. On the boundary of the domain $\mathcal{R} = (a-1, a+1)$ we have

$$w(a \pm 1) = u(a \pm 1) - v(a \pm 1) = u(\pm 1) - 1 \leq 0,$$

because we are given that $u(x) \leq 1$ for all x.

The maximum principle 6.2 now implies that $w(x) \leq 0$ for all $x \in (a-1, a+1)$. In particular $w(a) \leq 0$, or,

$$u(a) \leq v(a) = \frac{2}{3}.$$

The only assumption we made on the number a was $a > 1$, so we have shown that if $u : (0, \infty) \to \mathbb{R}$ is a solution to $u'' = (2 + \sin x)u$ with $u(x) \leq 1$ for all $x > 0$, then $u(x) \leq \frac{2}{3}$ for all $x > 1$.

APPENDIX A

Analysis reference

1. Integrating and differentiating sequences and series

1.1. Fundamental Theorem of Calculus. If $f : [a, b] \to \mathbb{R}$ is continuous, then

$$F(x) \stackrel{\text{def}}{=} \int_a^x f(\xi)\, d\xi$$

is a differentiable function, and

$$F'(x) = \frac{d}{dx} \int_a^x f(\xi)\, d\xi = f(x).$$

2. Switching and differentiating integrals

2.1. Switching integrals. If $f : [a, b] \times [c, d] \to \mathbb{R}$ is a Riemann integrable function of two variables, then

$$\int_a^b \int_c^d f(x, y)\, dy\, dx = \int_c^d \int_a^b f(x, y)\, dx\, dy$$

2.2. Integrals depending on a parameter. If $f : [a, b] \times (t_0, t_1) \to \mathbb{R}$ is continuous, then the function $F : (t_0, t_1) \to \mathbb{R}$ given by

$$F(t) \stackrel{\text{def}}{=} \int_a^b f(x, t)\, dx$$

is continuous.

2.3. Differentiating the parameter in an integral. If $f, \frac{\partial f}{\partial t} : [a, b] \times (t_0, t_1) \to \mathbb{R}$ are continuous, then

$$\frac{d}{dt} \int_a^b f(x, t)\, dx = \int_a^b \frac{\partial f}{\partial t}(x, t)\, dx$$

Proof. Let

$$F(t) = \int_a^b f(x, t)\, dx.$$

Choose $T \in (t_0, t_1)$. Then

$$\begin{aligned} F(t) &= F(T) + \int_a^b \left\{ f(x, t) - f(x, T) \right\} dx \\ &= F(T) + \int_a^b \int_T^t \frac{\partial f}{\partial t}(x, \tau)\, d\tau\, dx. \end{aligned}$$

Since $\partial f/\partial t$ is continuous we are allowed to switch the order of integration:

$$F(t) = F(T) + \int_T^t \int_a^b \frac{\partial f}{\partial t}(x,\tau)\,dx\,d\tau.$$

The Fundamental Theorem of Calculus now implies

$$F'(t) = \frac{d}{dt}\int_T^t \int_a^b \frac{\partial f}{\partial t}(x,\tau)\,dx\,d\tau = \int_a^b \frac{\partial f}{\partial t}(x,t)\,dx.$$

Q.E.D.

3. Norms and inner products

3.1. Definition (real inner product). A real inner product on a real vector space V is a real valued function on $V \times V$, usually written as (x, y) or $\langle x, y\rangle$ such that the following properties hold for all $x, y, z \in V$ and $a \in \mathbb{R}$:

- $\langle x, y\rangle = \langle y, x\rangle$
- $\langle ax, y\rangle = a\langle x, y\rangle$
- $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$
- $\langle x, x\rangle > 0$ if $x \neq 0$

3.2. Definition (complex inner product). A complex inner product on a complex vector space V is a complex valued function on $V \times V$, usually written as (x, y) or $\langle x, y\rangle$ such that the following properties hold for all $x, y, z \in V$ and $a \in \mathbb{C}$:

- $\langle x, y\rangle = \overline{\langle y, x\rangle}$ where $\overline{z}$ is the complex conjugate of $z \in \mathbb{C}$
- $\langle ax, y\rangle = a\langle x, y\rangle$
- $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$
- $\langle x, x\rangle > 0$ if $x \neq 0$

3.3. Example: inner product on a function space. Let $\mathcal{C}_{\text{per}}$ be the space of 2π-periodic continuous functions $f : \mathbb{R} \to \mathbb{C}$. Then

$$\langle f, g\rangle \stackrel{\text{def}}{=} \frac{1}{2\pi}\int_0^{2\pi} f(t)\overline{g(t)}\,dt$$

defines an inner product on $\mathcal{C}_{\text{per}}$.

3.4. Definition (norm). If V is a real or complex vector space then a function $x \in V \mapsto \|x\| \in \mathbb{R}$ is called a norm on V if

- $\|x\| > 0$ for all $x \in V$ with $x \neq 0$
- $\|ax\| = |a|\,\|x\|$ for all $x \in V$ and $a \in \mathbb{R}$ or $a \in \mathbb{C}$
- $\|x + y\| \leq \|x\| + \|y\|$ (the triangle inequality)

3.5. Definition (distance between vectors). The distance between two vectors $x, y \in V$ in a normed vector space is $d(x, y) \stackrel{\text{def}}{=} \|x - y\|$.

3.6. Theorem about the norm in an inner product space. If V is a real or complex inner product space with inner product $\langle x, y\rangle$. Then

$$\|x\| = \sqrt{\langle x, x\rangle}.$$

is a norm on V. In addition to the norm properties (see definition 3.4) it also satisfies the Cauchy-Schwarz inequality

$$|\langle x, y\rangle| \le \|x\| \, \|y\|$$

for all $x, y \in V$.

Note that $\langle x, x\rangle$ is never negative, so the square root is always defined.

3.7. Definition (orthogonality). Two vectors $x, y \in V$ are said to be orthogonal if $\langle x, y\rangle = 0$. Notation: $x \perp y$ means x, y are orthogonal.

In particular, the zero vector is orthogonal to every other vector, because $\langle x, 0\rangle = 0$ for all $x \in V$.

4. Convergence of sequences and series of functions

4.1. Uniform convergence. Let V be the vector space of bounded functions from some set X to the real numbers:

$$V = \big\{ f : X \to \mathbb{R} \mid \sup_{x \in X} |f(x)| < \infty \big\}.$$

Then

$$\|f\|_\infty \stackrel{\text{def}}{=} \sup_{x \in X} |f(x)|$$

defines a norm on V, often called the "sup–norm". By definition a sequence of functions $f_n : X \to \mathbb{R}$ converges uniformly to a function $f : X \to \mathbb{R}$ if $\lim_{n\to\infty} \|f_n - f\|_\infty = 0$.

4.2. Theorem — Weierstrass M-test. Let $f_n : [a, b] \to \mathbb{C}$ be a sequence of functions for which we can find real numbers $M_n > 0$ such that

- $|f_n(x)| \le M_n$ for all $x \in [a, b]$, and
- $\sum_{n=1}^\infty M_n < \infty$.

Then the series $\sum_{n=1}^\infty f_n(x)$ converges uniformly on $[a, b]$.

Proof. Let $s(x) = \sum_{n=1}^\infty f_n(x)$ and $s_N(x) = \sum_{n=1}^N f_n(x)$. We have to show that $s_N(x)$ converges uniformly to $s(x)$.

Let $\epsilon > 0$ be given. For any $x \in [a, b]$ and any $N \in \mathbb{N}$ we have

$$\begin{aligned}
|s(x) - s_N(x)| &= \left| s(x) - \sum_{n=1}^{N} f_n(x) \right| \\
&= \left| \sum_{n=1}^{\infty} f_n(x) - \sum_{n=1}^{N} f_n(x) \right| \\
&= \left| \sum_{n=N+1}^{\infty} f_n(x) \right| \\
&\le \sum_{n=N+1}^{\infty} |f_n(x)| \\
&\le \sum_{n=N+1}^{\infty} M_n
\end{aligned}$$

Since the series $\sum M_n$ is known to converge, an $N_\epsilon \in \mathbb{N}$ exists for which $\sum_{n=N_\epsilon}^{\infty} M_n < \epsilon$. It then follows that for all $N \ge N_\epsilon$ and all $x \in [a, b]$ one has

$$|s(x) - s_N(x)| < \epsilon.$$

This implies $s_N(x) \to s(x)$ uniformly. *Q.E.D.*

Example 1. The series $\sum_{n=1}^{\infty} \frac{\sin nx}{n^2}$ converges uniformly for all $x \in \mathbb{R}$ because the n^{th} term is bounded by $\left|\frac{\sin nx}{n^2}\right| \le M_n$ with $M_n = \frac{1}{n^2}$, and the series $\sum_1^{\infty} \frac{1}{n^2}$ converges. Each term is continuous and the series converges uniformly. Therefore the function $f(x) = \sum_1^{\infty} \frac{\sin nx}{n^2}$ is a continuous function.

Example 2. The sawtooth function is defined by $f(x) = \frac{\pi - x}{2}$ $(0 < x < 2\pi)$ and $f(x + 2\pi) = f(x)$ for all $x \in \mathbb{R}$. Its Fourier series is

$$f(x) = \sum_{n=1}^{\infty} \frac{\sin nx}{n}.$$

The best upper bound we can find for the n^{th} term is $M_n = \frac{1}{n}$. The series $\sum_n M_n = \sum_n \frac{1}{n}$ does not converge, so we cannot apply the Weierstrass M-test. The sawtooth function f is discontinuous at $x = 2k\pi$ $(k \in \mathbb{Z})$, so the series cannot converge uniformly.

4.3. The integral and derivative of a limit.