MA30059/MA40059: Mathematical Methods 2

Lecturer: Kirill Cherednichenko, University of Bath

Monday 3rd February, 2025

Contents

1 An Overview and Motivation **3**

1.1 The Three Parts of This Unit: Elliptic PDEs, Parabolic PDEs, and Calculus of Variations . . . 3

1.1.1 Part One of the Unit . . . 3

1.1.2 Part Two of the Unit . . . 3

1.1.3 Part Three of the Unit . . . 3

1.2 Different Perspectives on Partial Differential Equations and How They Link This and Other Units . . . 4

2 Notation and terminology **5**

3 Elliptic Equations (Laplace Operator) **8**

3.1 Harmonic Functions: Mean-Value Property, Maximum Principle . . . 8

3.1.1 Mean-value theorem . . . 9

3.1.2 Maximum principle . . . 10

3.2 Harmonic functions in a ball in $\mathbb{R}^3$ by separation of variables . . . 11

3.3 Harnack inequality (Not examinable) . . . 15

4 Fundamental Solution, Potentials, and Green's Representation Theorem **16**

4.1 Distributions . . . 16

4.2 Fundamental Solution for the Laplace Operator in $\mathbb{R}^d$ . . . 19

4.3 Green's Function for the Laplace Operator in a Bounded Region . . . 21

5 Solution to the Dirichlet Problem in Specific Regions by the Method of Images **22**

5.1 Construction of the Green's Function for a Ball ($d \geq 3$) . . . 22

5.2 Poisson's Formula for the Solution of the Dirichlet Problem in a Ball . . . 25

5.3 Green's Function for Specific Regions by the Method of Images . . . 27

6 Physical Derivation of the Heat Equation (Not examinable) **31**

7 Initial Boundary-Value Problem for the Heat Equation on a Bounded Interval **33**

7.1 Separation of Variables and Solution of the Homogeneous Equation . . . 34

7.2 Green's Function for the Heat Equation on a Bounded Interval . . . 35

7.2.1 Green's function as a response to an instantaneous point source . . . 35

7.2.2 Physical interpretation (Not examinable) . . . 39
7.3 The Inhomogeneous Equation of Heat Conduction . . . 39
7.3.1 Mathematical applied mathematician's derivation . . . 40
7.3.2 Physical applied mathematician's derivation (Not examinable) . . . 41
7.4 General Initial Boundary-Value Problem . . . 42

8 Initial Boundary-Value Problem for the Heat Equation in a Bounded Region of $\mathbb{R}^d$ 43
8.1 Maximum Principle for the Heat Equation . . . 43
8.2 Uniqueness Theorem for the Initial Boundary-Value Problem and Other Consequences of the Maximum Principle . . . 44
8.2.1 Uniqueness theorem . . . 45
8.2.2 Large-time behaviour of solutions with zero boundary data . . . 45
8.2.3 Stabilisation property . . . 46

9 Cauchy Problem for the Heat Equation on the Whole Space $\mathbb{R}^d$ 48
9.1 Maximum Principle and Uniqueness Theorem . . . 48
9.2 Fourier Transform and Its Basic Property . . . 49
9.3 Derivation of the Poisson's Formula . . . 50
9.4 Justification of Poisson's Formula (Not examinable) . . . 51
9.5 Some Important Remarks . . . 52

10 Calculus of Variations: Motivation 53
10.1 The Brachistochrone Problem . . . 53
10.2 The Navigation Problem (Not examinable) . . . 54
10.3 The Hanging Chain Problem (Not examinable) . . . 56
10.4 One-Dimensional Nonlinear Elasticity (Not examinable) . . . 57

11 Global and Local Minimisers of Integral Functionals 58
11.1 Notation and Terminology . . . 59
11.2 Global Minimisers and Motivation for New Analysis . . . 60

12 Necessary and Sufficient Conditions for a Minimum 63
12.1 Euler-Lagrange Equation: Necessary Condition for a Local Minimum . . . 64
12.2 Natural Boundary Conditions . . . 68
12.3 Smoothness of local minima . . . 71
12.4 Special Cases for Integrating the Euler-Lagrange Equation . . . 72
12.5 Sufficient Condition for a Local Minimum . . . 72

13 Method of Lagrange Multipliers for Constrained Problems (Not examinable) **76**

14 Appendix A: Some Useful Results from Calculus 78
14.1 Intermediate Value Theorem . . . 78
14.2 Derivative Mean-Value Theorem . . . 78
14.3 Integral Mean-Value Theorem . . . 78
14.4 Different Versions of the Taylor Formula . . . 78
14.5 Integration by Parts . . . 79

15 Appendix B: Solutions to Selected Exercises **80**
15.1 Exercise 4.4 80
15.2 Exercise 4.3(a) 81

1 An Overview and Motivation

1.1 The Three Parts of This Unit: Elliptic PDEs, Parabolic PDEs, and Calculus of Variations

1.1.1 Part One of the Unit

The first part of the unit concerns the elliptic PDEs, and in particular the Laplace's equation. Given an open set $\Omega \subset \mathbb{R}^d$, Laplace equation for a function $u : \Omega \to \mathbb{R}$ is

$$\Delta u(x) = 0, \qquad x \in \Omega, \tag{1}$$

where

$$\Delta := \sum_{j=1}^{d} \frac{\partial^2}{\partial x_j^2} \tag{2}$$

is the "Laplacian" or "Laplace operator".[1] The inhomogeneous equation with the operator Δ, i.e. an equation obtained from (1) by setting some function f (not necessarily vanishing) on the right-hand side:

$$-\Delta u(x) = f, \qquad x \in \Omega, \tag{3}$$

is referred to as the Poisson's equation.[2]

1.1.2 Part Two of the Unit

We will continue by exploring parabolic PDEs, and in particular the heat, or diffusion, equation. With Ω again an open set in $\mathbb{R}^d$ and $t \in [0, T]$ for some $T > 0$ (possibly $T = \infty$), the heat equation for a function $u = u(x, t)$, $x \in \Omega$, $t \in [0, T]$ is

$$\frac{\partial u}{\partial t}(x, t) = \Delta u(x, t), \qquad x \in \Omega, \quad t \in [0, T].$$

1.1.3 Part Three of the Unit

Here we will look at the calculus of variations, an approach to the study of a class of optimisation problems. The basic problem here is the following: let X be a class of real-valued functions defined on $\Omega \subset \mathbb{R}^d$ and introduce a functional I (i.e. an $\mathbb{R}$-valued "function of functions") defined on X by

$$I(u) = \int_\Omega f\big(x, u(x), \nabla u(x)\big), \qquad u \in X, \tag{4}$$

[1] As we know from the study of Sturm-Liouville problems, the properly defined notion of an operator must involve, in addition to a differential expression, such as (2), a set of functions ("domain") on which this expression is considered.

[2] There is a good reason why the minus sign appears in front of Δ in (3). It is fundamentally the same as why we wrote $Lu = -(pu')' + \dots$ for the Sturm-Liouville operator in MA30044, namely to make sure that the operator is non-negative. Of course, of you see the equation $\Delta u = f$, you can immediately turn it into an equation fo the form (3), by writing $-\Delta u = -f$.

where $f : \Omega \times \mathbb{R} \times \mathbb{R}^d$. The problem is then to determine $u \in X$ such that[3] $I(u) \leq I(v)$ for all $v \in X$, or alternatively for all v in a sufficiently small neighbourhood of u. (We equip X with a metric so that the notion of a neighbourhood of u makes sense.)

1.2 Different Perspectives on Partial Differential Equations and How They Link This and Other Units

In some sense, the unifying theme of the unit is partial differential equations (we shall see that although variational principles such as those mentioned in Section 1.1.3 are not differential equations, they are intimately linked to them). You have met several PDEs over the last few years, but we shall (hopefully) study PDEs in a different manner to how you may have done so up to this point:

1. Up to now, the majority of times you have met a PDE you have tried to find an explicit expression for the solution. Unfortunately, *for the vast majority of PDE problems there is not an explicit expression for the solution.* Since we still want to study PDEs (and find out information about the real-world phenomenon that they model), this means we need to be more flexible; that is, we often have to give up trying to find explicit solutions and instead try to find other information about the particular PDEs of interest.

2. Up to now you have usually assumed that everything was "legal", that is, the PDE had a solution, and there was only one of them. In many cases one can make physical arguments why this should be the case (assuming that the PDE is really representing the physical reality), but sometimes there are surprises (in fact, later in the unit we will see an example where physical intuition lets us down). Therefore, in this unit we will be a bit more careful and try to show that, for example, the solutions to the PDE problems that we are interested in are unique.

There is a spectrum of ways in which mathematicians study PDEs. Some important points on this spectrum are the following:

(a) Attempt to prove that a solution exists, is unique, and that the problem involving the PDE is "well-posed" in some sense.[4]

(b) Attempt to find an approximation to the solution (either numerical or asymptotic).

(c) Attempt to prove qualitative knowledge about the solution (e.g. the solution satisfies a maximum principle) or additional properties (e.g. the PDE can be put in variational form).

(d) Attempt to find an explicit expression for the solution.

These different approaches are listed (roughly) in order of generality, e.g. the class of PDEs for which you can prove that the solution is unique is much wider than the class of PDEs for which you can find an explicit expression for the solution.

Obviously this is a slightly crude distinction, with many of the approaches overlapping:

- If you can prove that a PDE has a maximum principle (c) then, in some cases, you can prove uniqueness (a) (we will do this in this unit for the Laplace and heat equations).

- If you can find an explicit expression for the solution to a PDE (d), then you can prove that a solution exists (a) (you essentially did this in MA20223 and MA30044).

[3] Here we consider minimising I. If instead we want to maximise it, we can use the fact that $\min I = ? \min(?I)$.

[4] A problem for a PDE is well-posed if the problem has a solution, the solution is unique, and the solution depends continuously on the data given in the problem.

- If you can put a PDE in variational form (c) then, in some cases, you can prove that a solution exists (a) (this is beyond the scope of the unit).

It is perhaps instructive to look back at the units you have done and see which of the different approaches you have used.

- MA20223 (Vector calculus and PDEs): approach (d) on Laplace's equation and the heat and wave equations via separation of variables.

- MA30044 (Mathematical Methods 1): approach (d) on Laplace's equation and the heat and wave equations via separation of variables and transform methods, approach (d) via characteristics (the idea of characteristics can also be used for approach (a) via the so-called "Cauchy-Kovalevskaya theorem").

In this unit we will mainly be concentrating on approaches (a) and (c), expanding your knowledge of Laplace's equation and the heat equation, and introducing the Calculus of Variations (mainly from a physical, rather than PDE, point of view). Finally, we will conclude this with a look at other units you may have done, and some units you may do in the future, in light of the discussion above, and the approaches (a)–(d).

- MA20222 (Numerical analysis): approach (b) to certain ODEs.

- MA30062 (Analysis of nonlinear ordinary differential equations) approaches (c) and (a) to ODEs.

- MA30170 (Numerical solution of elliptic PDEs) and MA40171 (MA40171: Numerical solution of evolution equations): approach (b) to certain PDEs (including Poisson's equation, the heat equation, and the wave equation).

- MA40048 (Analytical and geometric theory of differential equations) approach (c) to ODEs via the Lagrangian and Hamiltonian frameworks

- MA40255 (Viscous fluid mechanics) and MA40049 (Elasticity) contain more examples of the modelling process of converting physical processes into differential equations.

2 Notation and terminology

- We will often use the Einstein's convention about summing with respect to the repeated indices and write, for example $x_j y_j$ instead of $\sum_{j=1}^{d} x_j y_j$ (where the range for j would be clear from the context).

- The symbol $:=$ stands for "the object on the side of : is the notation for the object on the side of =". For example,
$$\partial_j f(x) := \lim_{\varepsilon \to 0} \frac{f(x + \varepsilon e_j) - f(x)}{\varepsilon}$$
should be read as "$\partial_j f(x)$ is the notation for the limit on the right-hand side (i.e. the partial derivative of f with respect to the j-th coordinate, evaluated at the point x.)

- For $d \in \mathbb{N}$, we denote by $\mathbb{R}^d$ the d-dimensional Euclidean space. For $x \in \mathbb{R}^d$, we denote by $|x|$ the Euclidean norm of x :

$$|x| := \sqrt{x_1^2 + \cdots + x_d^2} = \sqrt{\sum_{j=1}^{d} x_j^2} = \sqrt{x_j x_j}.$$

- For $X \subset \mathbb{R}^d$, the closure of X (i.e. the set of all limit points of sequences in X) is dented by $\overline{X}$.

- In what follows, we denote by $B_\rho(x_0)$ the open ball of radius $\rho > 0$ centred at $x_0 \in \mathbb{R}^d$:

$$B_\rho(x_0) := \big\{x \in \mathbb{R}^d : |x - x_0| < \rho\big\},$$

  and by S_ρ the boundary of this ball:

$$S_\rho(x_0) := \partial B_\rho(x_0) = \big\{x \in \mathbb{R}^d : |x - x_0| = \rho\big\}.$$

  Combining this with the previous notation, we conclude, in particular, that $\overline{B_\rho(x_0)}$ the closed ball or radius ρ centred at x_0, i.e. the set $\{x \in \mathbb{R}^d : |x - x_0| \leq \rho\}$.

- In line with the above notation, for $\rho > 0$, we will denote by $|B_\rho|$ and $|S_\rho|$ the volume of a ball of radius ρ and the area[5] of a sphere of radius ρ (which do not depend on the centre of the ball or the sphere). In particular, one has

$$|B_1| = \frac{|B_\rho|}{\rho^d} = \begin{cases} 2, & d = 1, \\ \pi, & d = 2, \\ \dfrac{4}{3}\pi, & d = 3 \end{cases}, \qquad |S_1| = \frac{|S_\rho|}{\rho^{d-1}} = \begin{cases} 2, & d = 1, \\ 2\pi, & d = 2, \\ 4\pi, & d = 3. \end{cases}$$

- By the *interior* of a set $X \subset \mathbb{R}^d$, we mean the union of all open balls that are contained in X. Clearly, this is an open set, as the union of open sets.

- The *boundary* ∂X of a set $X \subset \mathbb{R}^d$ is the difference of the closure and the interior of X. So, for example, we can write $S_\rho(x_0) = \partial B_\rho(x_0)$.

- For a function u defined in a "neighbourhood" of a point x_0 that belongs to an interior of X (e.g. a ball containing x_0 and contained in X_0,), we denote by $\nabla u(x_0)$ the *gradient* of u at x_0, i.e. the vector of its partial derivatives with respect to $x_1, x_2, \dots, x_d$:

$$\nabla u(x_0) := \left(\frac{\partial u}{\partial x_1}(x_0), \frac{\partial u}{\partial x_2}(x_0), \dots, \frac{\partial u}{\partial x_d}(x_0)\right).$$

  If x_0 varies, we obtain a vector-valued function ∇u.

  When u depends on more than one variable (say, $u = u(x, y)$) and we would like to emphasise that the gradient is related to a specific variable (say, y), we write $\nabla_y u$, i.e.

$$\nabla_y u(x, y) := \left(\frac{\partial u)}{\partial y_1}(x, y), \frac{\partial u}{\partial y_2}(x, y), \dots, \frac{\partial u}{\partial y_d}(x, y)\right).$$

[5] More precisely, the volume of co-dimension one.

- For a function u differentiable at $x \in \mathbb{R}^d$ and a unit vector l, the *derivative of u in the direction of l* ("directional derivative") at the point x is defined by

$$\frac{\partial u}{\partial l} := \nabla u \cdot l,$$

  where the value ∇u is taken at the point x.

  When u depends on more than one variable (say, $u = u(x, y)$) and we would like to emphasise that the directional derivative is related to a specific variable (say, y), we use the notation $\partial u / \partial l_y$, i.e.

$$\frac{\partial u}{\partial l_y} := \nabla_y u \cdot l.$$

- For a bounded region $\Omega \subset \mathbb{R}^d$, we denote by $C^l(\Omega)$ the set of (real-valued) functions on Ω that have all derivatives of order l in Ω (and hence all lower orders, too), and by $C^l(\overline{\Omega})$ we denote the subset of $C^l(\Omega)$ consisting of functions whose derivatives of order l can be continuously extended to points of $\partial\Omega$. (alternatively, this this the set of restrictions to $\overline{\Omega}$ of functions that are in $C^l(\widetilde{\Omega})$ for any open set $\widetilde{\Omega}$ that contains $\overline{\Omega}$.) The (linear) set $C^l(\overline{\Omega})$ is a Banach space (i.e. a complete normed space) with the norm[6]

$$\|\cdot\|_{C^l(\overline{\Omega})} := \max_{\alpha_1 + \dots + \alpha_d \leq l} \quad \max_{x \in \overline{\Omega}} \left| \partial_1^{\alpha_1} \dots \partial_d^{\alpha_d} u(x) \right| \tag{5}$$

  By $C^\infty(\Omega)$ we define the intersection of the sets $C^l(\Omega)$ over all l, and by $C_0^\infty(\Omega)$ its subset of functions that vanish outside some compact subset of Ω. Similarly, $C_0^\infty(\mathbb{R}^d)$ consists of functions that are infinitely differentiable on $\mathbb{R}^d$ and vanish outside some ball, and $C^\infty(\overline{\Omega})$ denotes the intersection of the sets $C^l(\overline{\Omega})$ over all l.

  By $C_0^1([a, b])$ we denote the space of functions u that are continuously differentiable on $[a, b] \subset \mathbb{R}$ and such that $u(a) = u(b) = u'(a) = u'(b) = 0$.

- For a function $u \in C(\overline{\Omega})$, we denote by $u|_{\partial\Omega}$ the *trace* of u on the boundary $\partial\Omega$, i.e. simply the restriction of u to $\partial\Omega$.

- We will say that a bounded surface (such as the boundary of a set) in $\mathbb{R}^d$ is C^1 or *continuously differentiable*, if it can be covered by a union of balls such that its intersection with each ball in the union is the graph of a C^1 function in an appropriate Euclidean frame in the above space $\mathbb{R}^d$.

- By dS we denote an element of a surface area, so that, for example, for a surface surface $\mathcal{S}$, and a function u defined on $\mathcal{S}$,

$$\int_{\mathcal{S}} u dS$$

[6] A useful point to remember is that as all notions of norm on a Euclidean space are equivalent, i.e. for any two norms $\|\cdot\|_1$, $\|\cdot\|_2$ there are constants $c_2 > c_1 > 0$ such that $c_1\|x\|_2 \leq \|x\|_1 \leq c_2\|x\|_2$ for all x. This implies, in particular, that in (5) the first maximum, which is here a norm in a $\binom{l}{d-1}$-dimensional space, can be replaced by any other norm in this space, for example we could use the following equivalent norm in $C^l(\overline{\Omega})$:

$$\|\cdot\|_{C^l(\overline{\Omega})} := \sum_{\alpha_1 + \dots + \alpha_d \leq l} \quad \max_{x \in \overline{\Omega}} \left| \partial_1^{\alpha_1} \dots \partial_d^{\alpha_s} u(x) \right|.$$

is the integral of u over $\mathcal{S}$. In particular, if $\mathcal{S}$ is the boundary $\partial\Omega$ of a domain Ω, and u is defined on $\overline{\Omega}$, then

$$\int_{\partial\Omega} u dS$$

is the integral of of u over the boundary $\partial\Omega$.

If the expression under the integral depends of more than one variable and we would like to specify the variable of integration (say, y), we will often attach the variable of integration as the subscript (say, dS_y) so, for example, in

$$\int_{\partial\Omega} u(x-y) dS_y$$

the integration is carried out in the variable y over the surface $\partial\Omega$ (and the integral is a function of x). On occasion, we will do so also for functions of one variable, as in

$$\int_{\partial\Omega} u(x) dS_x.$$

3 Elliptic Equations (Laplace Operator)

3.1 Harmonic Functions: Mean-Value Property, Maximum Principle

Definition 3.1. *Consider a bounded region (open connected set) Ω in $\mathbb{R}^d$. A function $u \in C^2(\Omega) \cap C(\overline{\Omega})$ that satisfies the Laplace's equation*[7]

$$\Delta u(x) = 0, \qquad x \in \Omega, \tag{6}$$

is said to be **harmonic** *in Ω.*

Exercise 3.2. *For an interval $(a, b) \subset \mathbb{R}$, determine all functions that are harmonic in (a, b).*

Remark 3.3. *Note that (6) makes sense for any function $u \in C^2(\Omega)$, for example the function $\log |x|$ satisfies (6) in $\Omega = B_1(0) \setminus \{0\}$ is two dimensions (when $B_1(0)$ is a disc of radius one), and $|x|^{-1}$ satisfies (6) in $\Omega = B_1(0) \setminus \{0\}$ is three dimensions (when $B_1(0)$ is a ball of radius one), but these functions are discontinuous at zero, which is a point on the boundary of the set. (Notice, by the way, that the disc/ball in this example can be replaced by any open set containing zero.)*

One could adopt a wider definition of a harmonic function, where the function is not necessarily continuous up to the boundary of Ω. In the some of the properties below we refer to $\partial\Omega$ (such as in maximum principle), hence our assumption that $u \in C(\overline{\Omega})$. However, it is good practice to keep your eyes peeled for when some of the assumptions we make can be dropped and the statement still remains valid!

Remark 3.4. *A mathematically exciting and physically relevant challenge is to extend the notion of a solution to (6) so that it makes sense for functions that are not twice differentiable (and even not differentiable at all!) The physical motivation behind this is that real-world physical processes are full of singularities, which need to be understood and quantified. Such non-smooth solutions to*

[7] Recall the notation (2) for the Laplace differential operator.

PDEs are referred to as "weak", "very weak", "distributional", depending on the degree of non-smoothness. An analysis of such solutions is largely beyond the content of this unit, although we will have a glimpse of distributions when we discuss the δ-function and fundamental solutions (something you already came across in MA30044.)

Incidentally, as we will see in Corollary 4.7, harmonic functions are infinitely differentiable, and this statement can be extended to functions that are harmonic in the weak or distributional sense. This is due to a "smoothening" property of the inverse of the Laplace operator, due to its ellipticity. Parabolic equations, studied in the second part of this course, also have the same property ("heat distributions are smooth"), but not hyperbolic equations ("waves can break").

We next prove several fundamental properties of harmonic functions. Throughout Section 3.1, we assume that $\Omega \subset \mathbb{R}^d$ is a bounded region and u is harmonic in Ω.

3.1.1 Mean-value theorem

Theorem 3.5 (First Mean-Value Theorem)**.** *Suppose $x_0 \in \Omega$, $B_\rho(x_0) \subset \Omega$, $\rho > 0$. Then*

$$u(x_0) = \frac{1}{|S_\rho|} \int_{S_\rho(x_0)} u(x) dS_x. \tag{7}$$

Proof. Without loss of generality, we can assume that $\overline{B}_\rho \subset \Omega$. Indeed, once the property (7) is established for balls such that $\overline{B}_\rho \subset \Omega$, it extends by continuity to the general case (i.e. including balls touching the boundary of Ω.)

Under the assumption, we have $u \in C^2(\overline{B}_\rho(x_0))$. Setting $v = 1$ in the Green's identity (see Section 14.5)

$$\int_{B_\rho(x_0)} (v\Delta u - u\Delta v) dx = \int_{\partial B_\rho(x_0)} \left(v\frac{\partial u}{\partial n} - u\frac{\partial v}{\partial n} \right) dS,$$

we obtain

$$\begin{aligned} 0 = \int_{S_\rho(x_0)} \frac{\partial u}{\partial n} dS &= \int_{\partial B_\rho(x_0)} \nabla u \cdot n \, dS = \int_{|y|=\rho} (\nabla u)(x_0 + y) \cdot \frac{y}{\rho} dS_y \\ &= \rho^{d-1} \int_{|y|=1} (\nabla u)(x_0 + \rho y) \cdot y dS_y = \rho^{d-1} \frac{\partial}{\partial \rho} \int_{|y|=1} u(x_0 + \rho y) dS_y. \end{aligned}$$

Therefore,

$$\int_{|y|=1} u(x_0 + \rho y) dS_y$$

is constant in ρ, and hence must necessarily be equal to its value at $\rho = 0$, which is $|S_1|u(x_0)$. It follows that

$$u(x_0) = \frac{1}{|S_1|} \int_{|y|=1} u(x_0 + \rho y) dS_y = \frac{1}{|S_\rho|} \int_{|y|=\rho} u(x_0 + y) dS_y = \frac{1}{|S_\rho|} \int_{|x-x_0|=\rho} u(x) dS_x,$$

as claimed. □

Theorem 3.6 (Second Mean-Value Theorem)**.** *Under the same assumptions as Theorem 3.5, one has*

$$u(x_0) = \frac{1}{|B_\rho|} \int_{B_\rho(x_0)} u(x) dx \tag{8}$$

Proof. The formula (8) is obtained immediately from (7) by setting $\rho = \sigma$ in

$$u(x_0)|S_1| = \int_{S_\sigma(x_0)} u(x)dS_x$$

and integrating with respect to σ over the interval $[0, \rho]$. □

3.1.2 Maximum principle

Theorem 3.7 (Maximum Principle)**.** *Suppose that u attains its maximum or minimum at a point $x_0 \in \Omega$. Then u is constant in Ω.*

Proof. Consider $r > 0$ such that $\overline{B}_r(x_0) \subset \Omega$. By Second Mean-Value Theorem, we have

$$u(x_0) = \frac{1}{|B_r|}\int_{|y|\leq r} u(x_0 + y)dy$$

Denoting $M = u(x_0) = \max_{x\in\Omega} u(x)$, we obtain

$$0 = u(x_0) - M = \frac{1}{|B_r|}\int_{|y|\leq r} \big(u(x_0 + y) - M\big)dy \leq 0, \tag{9}$$

as $u(x_0 + y) \leq M$, by the definition of M. It follows from the "sandwich" (9) that $u(x_0 + y) = M$ for all y, $|y| \leq r$. As Ω is connected, we can show that $u(x) = M$ for every $x \in \Omega$, by repeating the above argument as many times as necessary along a path that connects x and x_0 and lies completely within Ω, see Fig. 1.

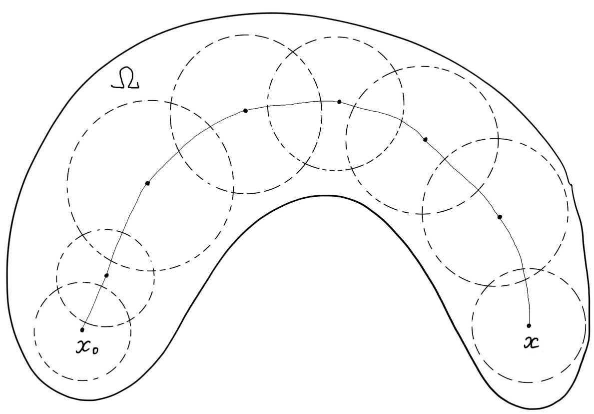

Figure 1: Connecting points $x_0, x \in \Omega$ by a finite sequence of overlapping balls. Why is it sufficient that x_0 and x can be connected by a continuous curve?

□

Corollary 3.8. 1. *("Weak Maximum Principle") The value of u at any point $x \in \Omega$ is bounded by the minimum and the maximum of u on the boundary of Ω :*

$$\min_{y\in\partial\Omega} u(y) \leq u(x) \leq \max_{y\in\partial\Omega} u(y) \qquad \forall x \in \Omega.$$

2. *("Principle of the Maximum Modulus") The maximum of the modulus of u over $\overline{\Omega}$ is attained on the boundary of Ω :*

$$\max_{x\in\overline{\Omega}}\left|u(x)\right| = \max_{x\in\partial\Omega}\left|u(x)\right|.$$

Proof. The proof is an (easy) exercise. For part 1, try to argue by contradiction, and part 2 follows from part 1 immediately. □

Remark 3.9. *In the proof of the maximum principle we only used the fact that (8) holds for u for all x_0 and ρ such that $\overline{B_\rho(x_0)} \subset \Omega$. Therefore, it also holds for any function that satisfies the mentioned property, even if it is not harmonic.*

Theorem 3.10 (Uniqueness for the Dirichlet Problem)**.** *Suppose $\varphi \in C(\partial\Omega)$, $f \in C(\Omega)$. Consider the problem ("Dirichlet problem") of finding $u \in C^2(\Omega) \cap C(\overline{\Omega})$, such that*

$$\begin{cases} -\Delta u = f \quad \text{in } \Omega, \\ u|_{\partial\Omega} = \varphi. \end{cases} \tag{10}$$

The above problem has no more than one solution u.

Proof. Suppose that u_1, u_2 are solutions to (10). Then $u = u_1 - u_2$ is harmonic in Ω, and u vanishes on $\partial\Omega$. By the principle of maxium modulus, we have

$$\max_{x\in\overline{\Omega}}\left|u(x)\right| = \max_{x\in\partial\Omega}\left|u(x)\right| = 0,$$

hence $u = 0$ everywhere in Ω, and hence $u_1 = u_2$. □

Remark 3.11. *The equation $-\Delta u = f$, i.e. the equation in Ω in (10) is often referred to as the Poisson's equation. Accordingly, the problem (10) is often referred as the "Dirichlet problem for the Poisson's equation" as well as the "Dirichlet problem for the Laplace operator".*

Exercise 3.12. *Produce a new version of Theorem 3.10, using the weak maximum principle rather than the principle of maxium modulus.*

Remark 3.13. *The assumption of boundedness of Ω in Definition 3.1 is made for convenience: in the present section we study properties of functions u satisfying (6) in bounded regions Ω. Of course, is if fine to use the term "harmonic function" for u that satisfies (6) in an unbounded region Ω, for example $\Omega = \mathbb{R}^d$, $\mathbb{R}^d \setminus \{0\}$, an infinite "strip" (in $\mathbb{R}^2$) or a layer in $\mathbb{R}^3$, an sector or cone, the exterior of a bounded region, or a region of the kind shown in Fig. 2 (or the interior of its complement, if connected), you name it! However, when analysing the properties of harmonic functions in an unbounded region Ω, we always bear in mind that the function satisfies the properties discussed in this section on its any bounded region contained in it (e.g. any ball contained in Ω.)*

3.2 Harmonic functions in a ball in $\mathbb{R}^3$ by separation of variables

The electric potential of a charged particle positioned at a point $x_0 \in \mathbb{R}^3$ (with a unit charge, for simplicity) is given by (cf. (30) with $d = 3$)

$$V_{\text{point}}(x) = \frac{1}{4\pi|x - x_0|}, \qquad x \in \mathbb{R}^3. \tag{11}$$

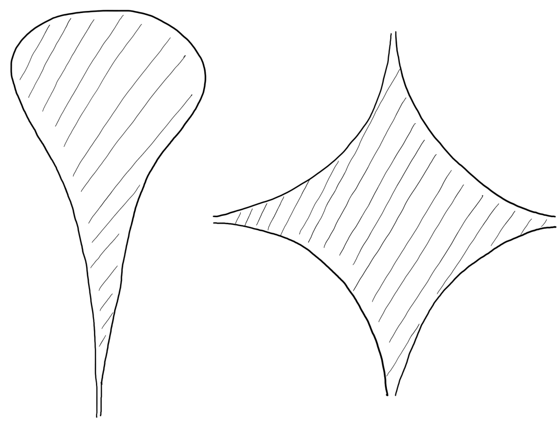

Figure 2: Unbounded regions with "cusps" going to infinity. Note that the interior of the complement of the region on the left is a region, and the interior of the complement of the region on the right is a union of four regions.

A direct calculation shows that it satisfies Laplace's equation everywhere away from the point x_0. By superposition, i.e. the observation that the potential of a charge distribution is the sum of the potentials of individual point charges, the potential of a charge distribution of density $f = f(x)$ is

$$V(x) = \int_{\mathbb{R}^3} V_{\text{point}}(x-y)f(y)dy = \frac{1}{4\pi}\int_{\mathbb{R}^3} \frac{f(y)}{|x-y|}dy, \tag{12}$$

and so

$$-\Delta V = f.$$

The expressions (11), (12) are examples of functions that are harmonic away from a region of $\mathbb{R}^3$ (which in this case represents the region occupied by the charges). In particular, if all the charges are located in the exterior of the unit ball $B_1(0)$, then we obtain functions that are harmonic in the ball.

Suppose now that u is harmonic in $B_1(0)$, in particular

$$\Delta u = 0 \quad \text{in} \quad B_1(0). \tag{13}$$

The fact that the function u satisfies Laplace's equation (13) can be written in spherical coordinates, see Fig. 3. For this, consider the function $\widetilde{u}$ defined in the set $[0,\infty)\times[0,2\pi)\times[0,\pi]$ and related to u by the formula

$$\widetilde{u}(r,\phi,\theta) = u(r\sin\theta\cos\phi, r\sin\theta\sin\phi, r\cos\theta).$$

For simplicity, assume that the function $\widetilde{u}$ is independent of ϕ : $\widetilde{u} = \widetilde{u}(r,\theta)$. As was noted in MA30044, we then have

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\widetilde{u}}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\widetilde{u}}{\partial\theta}\right) = 0, \tag{14}$$

Making another change variable $\xi = \cos\theta$, we wrote (14) in an equivalent form

$$\frac{\partial}{\partial r}\left(r^2\frac{\partial\widehat{u}}{\partial r}\right) + \frac{\partial}{\partial\xi}\left((1-\xi^2)\frac{\partial\widehat{u}}{\partial\xi}\right) = 0, \tag{15}$$

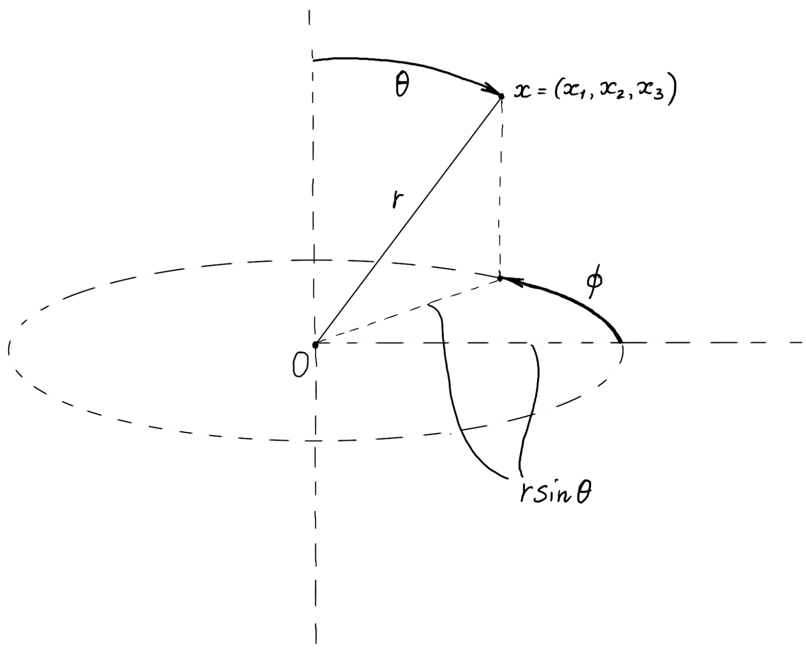

Figure 3: SPHERICAL COORDINATES r, ϕ, θ OF A POINT $x \in \mathbb{R}^3$.

where $\widehat{u}(r, \cos\theta) = \widetilde{u}(r, \theta)$. Looking for solutions of (15) in the form

$$\widehat{u}(r, \xi) = R(r)\Xi(\xi)$$

and dividing (15) through by $R\Xi$, we obtain

$$\big(R(r)\big)^{-1}\big(r^2 R'(r)\big)' = -\big(\Xi(\xi)\big)^{-1}\big((1-\xi^2)\Xi'(\xi)\big)' = \lambda, \qquad \lambda \in \mathbb{R},$$

i.e. the equation "separates" into two ODEs, in r only and in ξ only. The equation for Ξ is the Legendre eigenvalue equation and the requirement of regular behaviour for all θ yields $\lambda = n(n+1)$, $n = 0, 1, 2, \ldots$ The equation for R then yields

$$R(r) = C_n r^n + D_n r^{-(n+1)}, \qquad C_n, D_n \in \mathbb{R}.$$

As a result, the two series

$$\sum_{n=0}^{\infty} C_n r^n P_n(\cos\theta), \qquad \sum_{n=0}^{\infty} D_n r^{-(n+1)} P_n(\cos\theta) \tag{16}$$

provide the general form of functions that are independent of ϕ are harmonic in $\mathbb{R}^3$ and $\mathbb{R}^3 \setminus \{0\}$, respectively (extending the notion of a harmonic function to unbounded regions, see Remark 3.13).

Exercise 3.14 (Priority: low)**.** *Revise the related calculation and incorporate the dependence on ϕ in it. What functions of ϕ enter the general series representation?*

Exercise 3.15 (Priority: medium)**.** *Using the expressions for the Legendre polynomials you know from MA30044, write the first two harmonic functions entering the series (16), namely*

$$rP_1(\cos\theta), \qquad r^2P_2(\cos\theta), \qquad r \in [0,\infty), \quad \theta \in [0,\pi].$$

in terms of the Euclidean variables[8] x_1, x_2, x_3. *What property do these two expressions share? If after a minute's thought the pattern does not emerge, try doing the same for*

$$r^3P_3(\cos\theta).$$

Try to formulate a general statement about the form of

$$r^nP_n(\cos\theta), \qquad n \in \mathbb{N},$$

when expressed in terms of Euclidean coordinates.

Exercise 3.16 (Priority: high)**.** *Determine the pairs* (r,θ) *on which the maximum and minimum values of the function*

$$rP_1(\cos\theta), \qquad r \in [0,1], \quad \theta \in [0,\pi]$$

are attained. Investigate whether your result is in line with the maximum principle on the ball $B_1(0)$.

Exercise 3.17. *1. Consider the solution to the problem (cf. (10))*

$$\begin{cases} -\Delta u = f \quad \text{in} \quad B_1(0), \\ u|_{\partial\Omega} = \varphi, \end{cases} \tag{17}$$

i.e. the Dirichlet problem (10) with $\Omega = B_1(0)$, *and assume that the representation of the boundary values* ϕ *in spherical coordinates is independent of the angle* ϕ, *i.e.* $\varphi = \varphi(\theta)$. *More specifically, set*

$$\varphi(\theta) = P_m(\cos\theta), \qquad \theta \in [0,\pi], \tag{18}$$

for some fixed $m \in \mathbb{N}\cup\{0\}$. *(We can start with* $n = 0$.*) By uniqueness of the solution to the Dirichlet problem for the Laplace's operator, Theorem 3.10, the solution is given by the first expression in (16) with some* C_n, $n = 0, 1, \dots$ *— determine them. Explore their behaviour with respect to* r, θ *and check directly that they satisfy the maximum principle.*

2. Among functions given by the second expression in (16) there is also one that satisfies the boundary condition (18) – determine it. Why is this function not harmonic in $B_1(0)$, *even though it satisfies the Laplace's equation?*

Remark 3.18. *We must not underestimate the value of solutions in the form of the second series in (16). Consider the electric potential* u *in a ball (made of some solid dielectric rather than air) of radius 1 that contains a conducting inclusion centred at zero, see Fig. 4. The function* u *satisfies the Laplace's equation in the region (which is a spherical shell) occupied by the dielectric, subject to zero condition on the boundary of the inclusion and some boundary condition on the boundary of the ball. In other words, we have a Dirichlet problem for the Laplace's equation in a sphere with a "puncture". The solution is then the sum of the expressions (16), where the coefficients* C_n, D_n *are found by writing the two boundary conditions as series in terms of* $P_n(\cos\theta)$, $n = 0, 1, \dots$.

Exercise 3.19. *Follow through the programme for deriving series representation for solutions to the Dirichlet problem (10) in two dimensions, for the case when* $\Omega = B_1(0)$ *is a disc of radius one centred at zero.*

---

[8] Assuming that the vertical axis on Fig. 3 is x_3, recall that $r\cos\theta = x_3$ and $r^2 = x_1^2 + x_2^2 + x_3^2$.

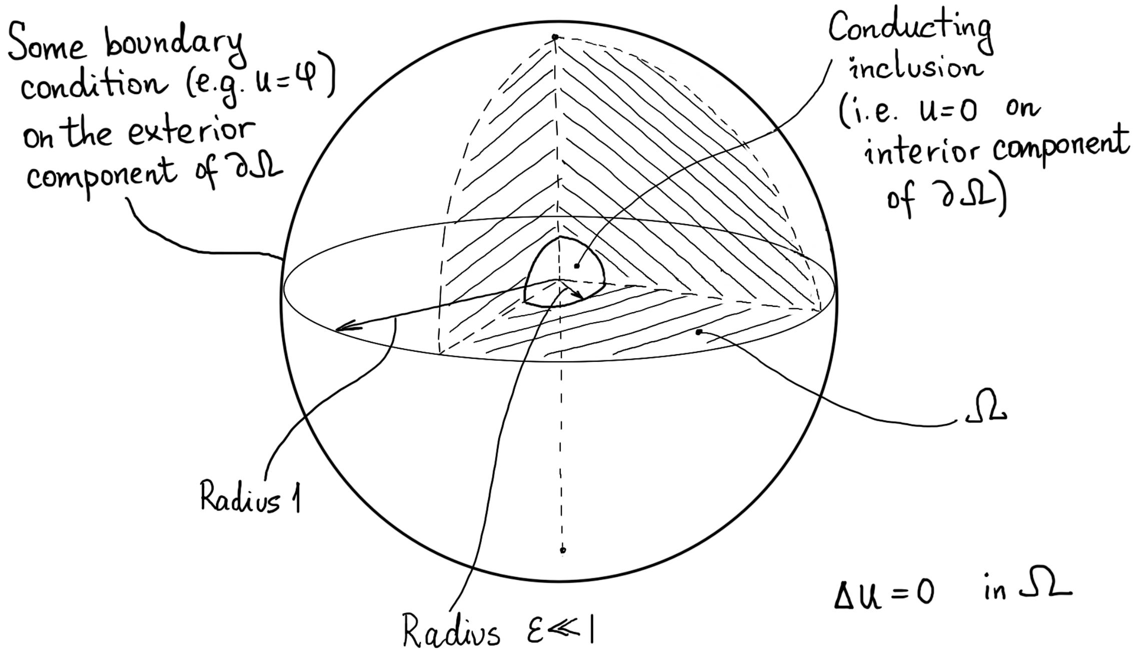

Figure 4: Harmonic functions in a punctured domain. Harmonic functions in such domains model the electrostatic potential of a charge outside the shell, where the "puncture" models a conducting inclusion.

3.3 Harnack inequality (Not examinable)

Theorem 3.20 (Harnack inequality)**.** *Consider a region Ω' such that $\overline{\Omega'} \subset \Omega$. Then there exists a constant $C = C(\Omega, \Omega')$ such that for all functions u that are non-negative in Ω one has*

$$\min_{x\in\overline{\Omega'}} u(x) \geq C \max_{x\in\overline{\Omega'}} u(x). \tag{19}$$

Proof. Consider a function u that is non-negative in Ω and choose $\rho > 0$ small enough so that for all $\widehat{x} \in \Omega'$ one has $B_{4\rho}(\widehat{x}) \subset \Omega$. (This can always be achieved due to the fact that $\overline{\Omega'} \subset \Omega$.)

For any $x_1, x_2 \in B_\rho(\widehat{x})$, we have the following estimates. First, notice that

$$u(x_1) = \frac{1}{|B_\rho|}\int_{|y|\leq\rho} u(x_1+y)dy,$$

and combining this with the fact that $B_\rho(x_1) \subset B_{2\rho}(\widehat{x})$, we obtain

$$u(x_1) = \frac{1}{|B_\rho|}\int_{B_{2\rho(\widehat{x})}} u, \tag{20}$$

as the function u is non-negative. Furthermore,

$$u(x_2) = \frac{1}{|B_{3\rho}|} \int_{|y|\le 3\rho} u(x_2+y)dy,$$

and combining this with the fact that $B_{3\rho}(x_1) \supset B_{2\rho}(\widehat{x})$, we obtain

$$u(x_2) = \frac{1}{|B_{3\rho}|} \int_{B_{2\rho}(\widehat{x})} u, \tag{21}$$

using again the fact that u is non-negative.

It follows from (20), (21) that

$$u(x_1) \le \frac{|B_{3\rho}|}{|B_\rho|} u(x_2). \tag{22}$$

Now, suppose that $x_1, x_2 \in \overline{\Omega'}$. Let us cover $\overline{\Omega'}$ by a finite set of balls of radius ρ. (We can always do so, as Ω, and hence $\overline{\Omega'}$, is bounded.) Construct a "path" $y_0 = x_1$, y_1, y_2, y_3, ..., y_{n-1}, $y_n = x_2$ such that for each $j = 0, \dots, n-1$, the point y_{j+1} lies in the interior of the ball centred at y_j, see Figure 5. Then, applying the estimate (22) repeatedly, we obtain

$$u(x_1) \le \frac{|B_{3\rho}|}{|B_\rho|} u(y_1) \le \frac{|B_{3\rho}|^2}{|B_\rho|^2} u(y_2) \le \cdots \le \frac{|B_{3\rho}|^n}{|B_\rho|^n} u(x_2) \tag{23}$$

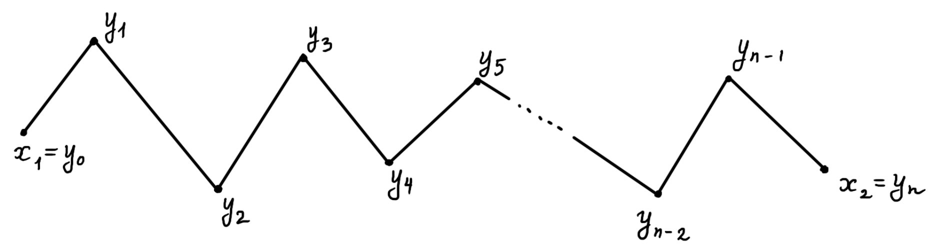

Figure 5: A "PATH" CONNECTING POINTS x_1 AND x_2 WITHOUT LEAVING THE SET $\overline{\Omega'}$.

It follows that for all $x_1, x_2 \in \overline{\Omega'}$, one has

$$u(x_1) \le \widetilde{C}(\Omega, \Omega')u(x_2), \tag{24}$$

where the constant $\widetilde{C}(\Omega, \Omega')$ depends on Ω, Ω'. Taking the maximum over x_1 on the left-hand side of (24) and the minimum over x_2 on its right-hand side, we obtain (19), where $C = 1/\widetilde{C}(\Omega, \Omega')$. □

4 Fundamental Solution, Potentials, and Green's Representation Theorem

4.1 Distributions

You came across distributions in MA20220, where they were used to provide a solution concept for ordinary differential equations with discontinuous source terms (i.e. right-hand sides) of inhomogeneous equations. A particular distribution on $\mathbb{R}$, the so-called δ-function, was mentioned.

Linear ODEs with the δ-function on the right-hand side (the so-called "fundamental solutions") can be used to determine solutions to inhomogeneous versions of the same ODEs, with arbitrary right-hand sides. In what follows we explore this idea for differential equations in $\mathbb{R}^d$, $d \geq 2$.

Definition 4.1. *A* **distribution** *is a linear continuous map* $F : C_0^\infty(\mathbb{R}^d) \to \mathbb{R}$.

Continuity means, as usual, that if ϕ_n converge to ϕ in some sense, then the values $F(\phi_n)$ converge to $F(\phi)$. Let us use the symbol $\Rightarrow$ for convergence in $C_0^\infty(\mathbb{R}^d)$ (which we are still to define), then $\phi_n \Rightarrow \phi$ implies $F(\phi_n) \to F(\phi)$. By $\phi_n \Rightarrow \phi$, we mean that: (a) The is a compact set K such that ϕ_n vanishes outside K *for all* n, and (b) any (fixed) derivative of ϕ_n converges, as $n \to \infty$ as uniformly on K, to the corresponding derivative of ϕ.

In the context of distributions, functions in $C_0^\infty(\mathbb{R}^d)$ are often referred to as *test functions.*

Every function u such that for every compact K the integral

$$\int_K u$$

exists and is finite gives rise to a distribution: the corresponding functional $[u]$ is defined by

$$[u](\phi) = \int_{\mathbb{R}^d} u\phi, \qquad \phi \in C_0^\infty(\mathbb{R}^d), \tag{25}$$

which clearly exists and is finite, as the integration only needs to be carried out over the (compact) support of ϕ. One example of a distribution for which there is no function u such that (25) holds is the so-called δ-function:

$$\delta(\phi) = \phi(0), \qquad \phi \in C_0^\infty(\mathbb{R}^d). \tag{26}$$

Exercise 4.2 (Priority: High)**.** *Show that the functional F_u defined by (25) and the functional δ defined by (26) are continuous with respect to the convergence $\Rightarrow$ of test functions, in the sense described above.*

Exercise 4.3 (Priority: Medium)**.** *A sequence of distributions $F^{(n)}$ is said to converge to a distribution F, if for any test function ϕ one has*

$$F^{(n)}(\phi) \to F(\phi) \quad \text{as} \quad n \to \infty. \tag{27}$$

Show that the sequence $F^{(n)} := [u_n]$, where $u_n = u_n(x)$, $x \in \mathbb{R}$, given by:

$$\text{(a)} \qquad u_n(x) = \frac{1}{2}\sqrt{\frac{n}{\pi}}\mathrm{e}^{-nx^2/4},$$

$$\text{(b)} \qquad u_n(x) = \frac{1}{\pi x}\sin(nx),$$

$$\text{(c)} \qquad u_n(x) = \frac{1}{\pi}\frac{n}{(nx)^2+1},$$

$$\text{(d)} \qquad u_n(x) = \frac{1}{n\pi x^2}\sin^2(nx)$$

converges to the delta-function δ on $\mathbb{R}$ as $n \to \infty$.

For each of the sequences above, write $1/\varepsilon$ instead of n, so $\varepsilon \to 0$ along the sequence. Are any of the formulae you have obtained familiar to you? Sketch the plots of functions in each of these families.

Many common operations (obviously addition and multiplication by a real number, but also multiplication by a smooth function, differentiation, “change of independent variable”) can be carried out on distributions. The key to extending these operations to such “general” functions, which are not necessarily functions in the usual sense, is that the set of all test functions is closed with respect to these operations.

For example, for every $x_0 \in \mathbb{R}^d$, one can consider the “shift of the independent variable by x_0” (although there is no independent variable in the usual sense!) by the formula

$$F_{x_0}(\phi) = F\big(\phi(\cdot + x_0)\big), \qquad \phi \in C_0^\infty(\mathbb{R}^d),$$

so that, in particular, the functional

$$\delta_{x_0} = \delta(\phi(\cdot + x_0)) = \phi(x_0)$$

is the x_0-shift of the δ-function (26), or the “δ-function at x_0.”

The derivative $\partial_j F$ of F is defined to be the functional

$$\partial_j F(\phi) = -F(\partial_j \phi), \qquad \phi \in C_0^\infty(\mathbb{R}^d), \tag{28}$$

where we use the shorthand $\partial_j \phi$ for the partial derivative of ϕ with respect to the j-th coordinate:

$$\partial_j \phi(x) := \frac{\partial \phi(x)}{\partial x_j}, \qquad x \in \mathbb{R}^d.$$

The definition (28) is, of course, a generalisation of the usual integration by parts rule. (To see this, set $F = F_u$, as in (25), where $u \in C^1(\mathbb{R}^d)$.)

The definition (28) can be iterated, so derivatives of *any* order are defined for F. Distributions are “infinitely differentiable”, even though some of them are not even representable by functions! This is because they “act” on infinitely differentiable functions and the property of differentiability can be “moved over” onto them from the objects they act upon (i.e. test functions). This is a common mathematical trick, usually referred to as “definition by duality”.

In particular, we now have the Laplace operator on distributions, via the formula

$$(\Delta F)(\phi) = F(\Delta \phi),$$

and the equation (6) defining harmonic functions (ignoring the condition $u \in C^2(\Omega) \cap C(\overline{\Omega})$ for a moment) can be understood “in the distributional sense”: we say that (6) holds for a distribution u if

$$u(\Delta \phi) = 0 \qquad \forall \phi \in C_0^\infty(\mathbb{R}^d). \tag{29}$$

As was already mentioned in Remark 3.4, the inverse of the Laplace operator “smoothens” functions. It turns out that for a distribution u that satisfies (29) there exists a function $\widetilde{u} \in C_0^\infty(\mathbb{R}^d)$ such that $u = F_{\widetilde{u}}$ in the sense of (25). In other words, any distributional solution of (6) is infinitely smooth is the sense of usual derivatives! This deep property of the Laplace operator is referred to as *hypoellipticity.*

4.2 Fundamental Solution for the Laplace Operator in $\mathbb{R}^d$

Consider the function

$$\Phi(x) = \begin{cases} \dfrac{1}{(d-2)|S_1||x|^{d-2}}, & x \in \mathbb{R}^d \setminus \{0\}, \qquad d \geq 3, \\ \dfrac{1}{2\pi} \log \dfrac{1}{|x|} = -\dfrac{1}{2\pi} \log |x|, & x \in \mathbb{R}^2 \setminus \{0\}, \end{cases} \tag{30}$$

We have come across the above expression for $d = 3$ when we discussed the electrostatic potential of a unit point charge, see (30).

Exercise 4.4. *Show that it satisfies the equation*

$$-\Delta\Phi = \delta, \tag{31}$$

where δ is the Dirac's delta-function. The equation (31) is understood in the sense that

$$-\int_{\mathbb{R}^d} \Phi\Delta\phi = \phi(0) \qquad \forall \phi \in C_0^\infty(\mathbb{R}^d). \tag{32}$$

[Hint: You only need to integrate over the support[9] of ϕ, and you can use the fact that

$$\phi(x) = \phi(0) + x \cdot \psi(x), \qquad x \in \mathbb{R}^d, \tag{33}$$

where the vector function ψ is smooth (i.e. its components are smooth functions), which can be easily deduced from Corollary 14.7.]

In the remainder of this section, $\Omega \subset \mathbb{R}^d$ a bounded region in $\mathbb{R}^d$ with C^1 boundary.[10]

Definition 4.5. *For a bounded region Ω with C^1 boundary,*[11] *consider functions $\varphi \in C(\partial\Omega)$ and $f \in C(\overline{\Omega})$.*

1. *The* **single-layer potential** *on $\partial\Omega$ with density φ is defined by the formula*[12]

$$\mathcal{S}[\varphi](x) = \int_{\partial\Omega} \Phi(x-y)\varphi(y)dS_y, \quad x \in \mathbb{R}^d \setminus \partial\Omega. \tag{34}$$

2. *The* **double-layer potential** *on $\partial\Omega$ with density φ is defined by the formula*

$$\mathcal{D}[\varphi](x) = \int_{\partial\Omega} \frac{\partial\Phi(x-y)}{\partial n_y}\varphi(y)dS_y, \quad x \in \mathbb{R}^d \setminus \partial\Omega.$$

3. *The* **Newtonian potential** *on Ω with density f is defined by*

$$\mathcal{N}[f](x) = \int_{\Omega} \Phi(x-y)f(y)dy, \quad x \in \mathbb{R}^d.$$

[9]The support of a function in $\mathbb{R}^d$ is the closure of the set of points where it does not vanish.

[10]More generally, a boundary consisting of a finite number of C^1 pieces will do. One often refers to such boundaries as "piecewise continuously differentiable".

[11]The regularity assumption about $\partial\Omega$ can be relaxed, namely one can assume that the boundary is "piecewise C^1", i.e. it consists of a finite number of C^1 "pieces" that make "corners" at those points of $\partial\Omega$ for which no unique normal vector can be defined.

[12]It can be shown that the single-layer potential is continuous across Ω, so the formula (34) is actually valid in all of $\mathbb{R}^d$.

Theorem 4.6 (Representation of twice differentiable functions in terms of potentials)**.** *Suppose that Ω is a bounded region with C^1 boundary. For any $u \in C^2(\overline{\Omega})$ the following formula holds:*

$$u(x) = \mathcal{S}\left[\frac{\partial u}{\partial n}\right](x) - \mathcal{D}[u](x) - \mathcal{N}[\Delta u](x), \qquad x \in \Omega. \tag{35}$$

Proof. We will give the proof for $d \geq 3$, i.e. when the fundamental solution is given by the first formula in (30).

For $x \in \Omega$, and $\rho > 0$ such that $\overline{B_\rho(x)} \subset \Omega$, consider the domain $\Omega_\rho := \Omega \setminus \overline{B_\rho(x)}$. Using the Green's identity (see Corollary 14.10), we can write

$$\int_{\Omega_\rho} \big(\Phi(x-y)\Delta u(y) - u(y)\Delta_y \Phi(x-y)\big)dy = \int_{\partial\Omega_\rho} \left(\Phi(x-y)\frac{\partial u(y)}{\partial n_y} - u(y)\frac{\partial \Phi(x-y)}{\partial n_y}\right)dS_y.$$

Rearranging the terms and using the fact that $\Delta_y \Phi(x-y) = 0$, $y \in \Omega_\rho$, it follows that

$$\begin{aligned} \int_{\Omega_\rho} \Phi(x-y)\Delta u(y)dy - \int_{\partial\Omega} \Phi(x-y)\frac{\partial u(y)}{\partial n_y}dS_y + \int_{\partial\Omega} u(y)\frac{\partial \Phi(x-y)}{\partial n_y}dS_y \\ = \int_{S_\rho(x)} \left(\Phi(x-y)\frac{\partial u(y)}{\partial n_y} - u(y)\frac{\partial \Phi(x-y)}{\partial n_y}\right)dS_y. \end{aligned} \tag{36}$$

We would now like to pass to the limit in (36) as $\rho \to 0$. To this end, notice first that there is a constant C such that

$$\begin{aligned} \left|\int_{S_\rho(x)} \Phi(x-y)\frac{\partial u(y)}{\partial n_y}dS_y\right| &\leq \int_{S_\rho(x)} \left|\Phi(x-y)\frac{\partial u(y)}{\partial n_y}\right|dS_y \\ &\leq \|u\|_{C^1(\overline{\Omega})}\frac{1}{(d-2)|S_1|\rho^{d-2}}|S_1|\rho^{d-1} = C\rho \to 0, \qquad \rho \to 0. \end{aligned} \tag{37}$$

Furthermore,

$$\begin{aligned} \int_{S_\rho(x)} u(y)\frac{\partial \Phi(x-y)}{\partial n_y}dS_y &= -\frac{1}{(d-2)|S_1|}\int_{S_\rho(x)} u(y)(r^{2-d})'\Big|_{r=\rho}dS_y \\ &= \frac{1}{|S_1|\rho^{d-1}}\int_{|x-y|=\rho} u(y)dS_y \\ &= \frac{1}{|S_1|\rho^{d-1}}\int_{|y|=\rho} u(x+y)dS_y \\ &= \frac{1}{|S_1|}\int_{|y|=1} u(x+\rho y)dS_y \to u(x), \qquad \rho \to 0. \end{aligned} \tag{38}$$

It follows from (37), (38) that the right-hand side of (36) converges to $u(x)$ as $\rho \to 0$. On the other hand, each term on its left-hand side clearly converges to the corresponding term in (35). $\square$

Corollary 4.7. *Any harmonic function in Ω is infinitely differentiable everywhere in Ω.*

Proof. For any ball B such that $\overline{B} \subset \Omega$, we have

$$u(x) = \int_{\partial B} \Phi(x-y)\frac{\partial u(y)}{\partial n_y} dS_y - \int_{\partial B} u(y)\frac{\partial \Phi(x-y)}{\partial n_y} dS_y, \qquad x \in B.$$

The claim follows by observing that both integrals in the last representation are infinitely differentiable on the interior of B. $\square$

Exercise 4.8. *Repeat the argument of the proof of the above representation result (Theorem 4.6) for case case of the fundamental solution (30) for the case of two dimensions.*

4.3 Green's Function for the Laplace Operator in a Bounded Region

Suppose that the boundary of Ω is piecewise continuously differentable, and consider the problem (cf. (10))

$$\begin{cases} -\Delta u = 0 \quad \text{in } \Omega, \\ u|_{\partial\Omega} = \varphi. \end{cases} \tag{39}$$

We have assumed that the right-hand side of the equation in Ω is zero. Notice that this does not lead to loss of generality, as in the case of a general right-hand side f (as in (10)), the Newtonian potential $\mathcal{N}[f]$ can be subtracted from the solution, and the remainder will necessarily satisfy (39).

Theorem 3.10 implies that the solution to (39), if it exists, is unique.

Definition 4.9. *Consider the function $G = G(x, y)$, $x \in \Omega, y \in \overline{\Omega}$, such that*

$$G(x, y) = \Phi(x-y) + g(x, y), \qquad x \in \Omega, \quad y \in \overline{\Omega}, \qquad x \neq y,$$

and the following conditions are satisfied for all $x \in \Omega$:

1. $g(x, \cdot) \in C^2(\overline{\Omega})$;
2. $\Delta_y g(x, y) = 0, \quad y \in \overline{\Omega}$;
3. $G(x, y) = 0, \quad y \in \partial\Omega$.

Then the function G is referred to as the **Green's function** *for the Laplace operator on Ω.*

Consider a region $\Omega \subset \mathbb{R}^d$ and suppose that a function described in the above definition exists. Then, assuming that (39) has a solution $u \in C^2(\overline{\Omega})$, we have by Green's identity (see Corollary 14.10)

$$0 = \int_{\Omega} \big(g(x, y)\Delta u(y) - u(y)\Delta_y g(x, y)\big)dy = \int_{\partial\Omega} \left(g(x, y)\frac{\partial u(y)}{\partial n_y} - u(y)\frac{\partial g(x, y)}{\partial n_y}\right)dS_y, \tag{40}$$

where we have used the fact that u is harmonic as well as the second property of the Green's function in Definition 4.9.

Therefore, invoking Theorem 4.6, the solution u to (39) can be represented as

$$u(x) = \mathcal{S}\left[\frac{\partial u}{\partial n}\right](x) - \mathcal{D}[u](x)$$

$$= \int_{\partial\Omega} \Big(\Phi(x-y) + g(x,y)\Big) \frac{\partial u(y)}{\partial n_y} dS_y - \int_{\partial\Omega} \frac{\partial}{\partial n_y} \Big(\Phi(x-y) + g(x,y)\Big) u(y) dS_y$$
$$= -\int_{\partial\Omega} \frac{\partial G}{\partial n_y}(x,y) u(y) dS_y = -\int_{\partial\Omega} \frac{\partial G}{\partial n_y}(x,y)\varphi(y) dS_y,$$

where for the first and second equality we have used (40) and the third property of G in Definition 4.9, respectively.

As a result, we have obtained the following statement.

Theorem 4.10. *Suppose that $\Omega \subset \mathbb{R}^d$ is a bounded region with piecewise continuously differentiable boundary, that the Green's function for Ω exists, and that $u \in C^2(\overline{\Omega})$ is the solution to the Dirichlet problem (39). Then one has*

$$u(x) = -\int_{\partial\Omega} \frac{\partial G}{\partial n_y}(x,y)\varphi(y) dS_y, \qquad x \in \Omega. \tag{41}$$

The above formula recovers the harmonic function in a bounded region Ω from its boundary values. One can say that we have thus found an explicit formula for the "harmonic lift" onto Ω of an arbitrary function φ on the boundary $\partial\Omega$.

5 Solution to the Dirichlet Problem in Specific Regions by the Method of Images

5.1 Construction of the Green's Function for a Ball ($d \geq 3$)

Suppose that $\Omega = B_R(0) =: B_R \subset \mathbb{R}^d$, $d \geq 3$, is a ball of radius $R > 0$, and recall (see Definition 4.9) that

$$G(x,y) = \Phi(x-y) + g(x,y), \qquad x \in B_R, \quad y \in \overline{B}_R, \qquad x \neq y,$$

where $g(x,\cdot) \in C^2(\overline{B_R})$ for all $x \in B_R$, and also

$$\Delta_y g(x,y) = 0, \qquad G(x,y)\Big|_{|y|=R} = 0. \tag{42}$$

Let us look for the term g of the Green's function G in the form

$$g(x,y) = -\frac{1}{(d-2)|S_1|} \frac{e(x)}{\big|\mathcal{K}(x) - y\big|^{d-2}}, \qquad x \in B_R \setminus \{0\}, \quad y \in \overline{B}_R \tag{43}$$

where $\mathcal{K}(x)$ is obtained by "inversion of x through $S_R = \partial B_R$" (or "Kelvin transform"), see Fig. 6, i.e. $\mathcal{K}(x) = cx$, $c > 0$, and $\big|\mathcal{K}(x)\big||x| = R^2$, equivalently

$$\mathcal{K}(x) := \frac{R^2}{|x|^2} x, \qquad x \neq 0, \tag{44}$$

and the function $e = e(x)$ is to be determined. (We will treat the point $x = 0$ separately.) Then automatically $g(x,\cdot) \in C^2(\overline{B_R})$, as the point $\mathcal{K}(x)$ is outside the ball B_R, whenever $x \in B_R$. Also, the first condition in (42) clearly holds: the function $|y|^{2-d}$, $y \in \mathbb{R} \setminus \{0\}$, is satisfies the Laplace's equation away from the point $y = 0$, as we have seen before, and modifying it with $e, \mathcal{K}$, which are

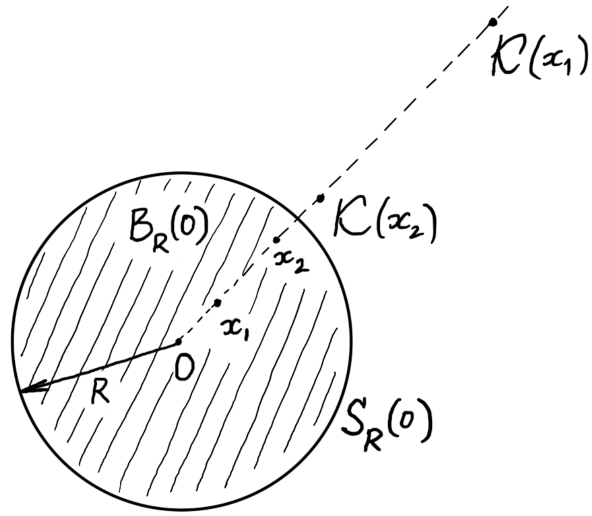

Figure 6: Kelvin transform (44), or inversion, with respect to the sphere $S_R(0)$. Check that for two points x_1, x_2 lying in the same half-line originating at zero, the Kelvin transform of the point x_1, which is closer to zero than x_2, is further away from zero than the Kelvin transform of the point x_2. Verify also that the image of the ball $B_R(0)$ under $\mathcal{K}$ is the exterior of its closure $\overline{B_R(0)}$.

constant in y, in the manner proposed in (43) again results in a function satisfying the Laplace's equation with respect to the variable y, only away from the point $y = \mathcal{K}(x)$, which lies outside the ball B_R.

Let us now determine the function e so that the second condition in (42) holds. This will completely determine the function g, and hence G, so that all three conditions in the definition of a Green's function are satisfied. This final requirement can be written in the following form:

$$\frac{1}{(d-2)|S_1|}\left(\frac{1}{|x-y|^{d-2}} - \frac{e(x)}{\big|\mathcal{K}(x)-y\big|^{d-2}}\right)\Bigg|_{|y|=R} = 0,$$

from which it follows that e must satisfy the condition

$$e(x) = \frac{\big|\mathcal{K}(x)-y\big|^{d-2}}{|x-y|^{d-2}}, \qquad |y| = R. \tag{45}$$

It may seem that (45) is impossible, as its right-hand side appears to depend on y, while its left-hand side is y-independent. The next few lines show that this appearance is deceptive.

Indeed, consider the expression in (45) as a function of $y \in S_R(0)$:

$$\begin{aligned}
|x-y|^2 &= |x|^2 - 2x\cdot y + R^2,\\
\big|\mathcal{K}(x)-y\big|^2 &= \big|\mathcal{K}(x)\big| - 2\mathcal{K}(x)\cdot y + R^2 = \frac{R^4}{|x|^4}|x|^2 - 2\frac{R^2}{|x|^2}x\cdot y + R^2\\
&= \frac{R^2}{|x|^2}\big(|x|^2 - 2x\cdot y + R^2\big) = \frac{R^2}{|x|^2}|x-y|^2,
\end{aligned}$$

so, in fact,

$$\frac{\left|\mathcal{K}(x)-y\right|^{d-2}}{|x-y|^{d-2}} = \frac{R^{d-2}}{|x|^{d-2}} \tag{46}$$

is y-independent! It follows that if we set

$$e(x) = \frac{R^{d-2}}{|x|^{d-2}},$$

then the condition (45) is satisfied automatically. Combining this with (43), we obtain

$$g(x,y) = -\frac{1}{(d-2)|S_1|}\frac{R^{d-2}}{|x|^{d-2}}\frac{1}{\left|\mathcal{K}(x)-y\right|^{d-2}},$$

and hence

$$G(x,y) = \frac{1}{(d-2)|S_1|}\left(\frac{1}{|x-y|^{d-2}} - \frac{R^{d-2}}{|x|^{d-2}}\frac{1}{\left|\mathcal{K}(x)-y\right|^{d-2}}\right), \tag{47}$$
$$x \in B_R \setminus \{0\}, \quad y \in \overline{B}_R \qquad x \neq y.$$

So far formula (47) has been derived under the assumption $x \neq 0$. However, the singularity brought into (47) turns out to be "removable": indeed, we can write

$$\begin{aligned}\frac{R^{d-2}}{|x|^{d-2}}\frac{1}{\left|\mathcal{K}(x)-y\right|^{d-2}} &= \frac{R^{d-2}}{|x|^{d-2}}\frac{1}{\left|\frac{R^2}{|x|^2}x - y\right|^{d-2}} \\ &= \frac{1}{R^{d-2}}\left|\frac{x}{|x|} - \frac{|x|}{R^2}y\right|^{2-d} \xrightarrow{x\to 0,\ x\neq 0} \frac{1}{R^{d-2}}\left|\frac{x}{|x|}\right|^{-1} = \frac{1}{R^{d-2}} \qquad \forall y \in \overline{B}_R.\end{aligned}$$

Remark 5.1. *The case $x=0$ can be treated in a different way. By the definition of the Green's function, we have*

$$G(0,y) = \frac{1}{(d-2)|S_1||y|^{d-2}} + g(0,y), \qquad y \in \overline{B}_R,$$

so if we set

$$g(0,y) = -\frac{1}{(d-2)|S_1|R^{d-2}}, \qquad y \in \overline{B}_R,$$

then all three conditions in the definition of the Green's function are satisfied. (In fact, there is no other choice for $g(0,y)$, $y \in B_R$, as it is a harmonic function, hence it is determined uniquely from its boundary values!) In other words,

$$G(0,y) = \frac{1}{(d-2)|S_1|}\left(\frac{1}{|y|^{d-2}} - \frac{1}{R^{d-2}}\right), \qquad x=0, \quad y \in \overline{B}_R \setminus \{0\}. \tag{48}$$

In summary, the final formula for the Green's function for the ball $B_R(0)$ is as follows:

$$G(x,y) = \frac{1}{(d-2)|S_1|}\left(\frac{1}{|x-y|^{d-2}} - \frac{1}{(h(x,y)R)^{d-2}}\right), \qquad x \in B_R, \quad y \in \overline{B}_R \qquad x \neq y, \tag{49}$$

where

$$h(x,y) := \begin{cases} \left|\frac{x}{|x|} - \frac{|x|}{R^2}y\right|, & x \in B_R \setminus \{0\}, \quad y \in \overline{B}_R, \\ 1, & x = 0, \quad y \in B_R. \end{cases} \tag{50}$$

5.2 Poisson's Formula for the Solution of the Dirichlet Problem in a Ball

We would like to substitute the expression (49) into the formula (41) for the case $\Omega = B_R(0)$, $R > 0$. It is more convenient, however, to use the formula (47) and then just notice that as $x \to 0$, $x \neq 0$, the obtained expression has a removable singularity, so hopefully the formula we find at the end of the calculation can simply be extended to $x = 0$. Proceeding with this plan, we notice that for $y \in S_R(0)$ one has

$$\frac{\partial}{\partial n_y}\left(|y-x|^{2-d}\right) = \frac{\partial}{\partial y_j}\left(\left(|y-x|^2\right)^{\frac{2-d}{2}}\right)n_j = \frac{2-d}{2}\left(\left(|y-x|^2\right)^{\frac{2-d}{2}-1}\right)2(y_j-x_j)n_j$$

$$= \frac{2-d}{|y-x|^d}\frac{(y_j-x_j)y_j}{R} = \frac{2-d}{|y-x|^d}\frac{|y|^2-x\cdot y}{R} = \frac{2-d}{|y-x|^d}\frac{R^2-x\cdot y}{R},$$

where $n_y = (n_1, \dots, n_d) = y/R$, $y \in S_1(0)$, is the unit outward-pointing normal vector to $S_1(0)$. Similarly, we have

$$\frac{\partial}{\partial n_y}\left(\left|y-\mathcal{K}(x)\right|^{2-d}\right) = \frac{2-d}{\left|y-\mathcal{K}(x)\right|^d}\frac{R^2-\mathcal{K}(x)\cdot y}{R}.$$

It follows that

$$-\frac{\partial}{\partial n_y}G(x,y) = \frac{1}{|S_1|R}\left(\frac{R^2-x\cdot y}{|y-x|^d} - \frac{R^{d-2}}{|x|^{d-2}}\frac{R^2-\mathcal{K}(x)\cdot y}{\left|\mathcal{K}(x)-y\right|^d}\right)$$

$$= \frac{1}{|S_1|R|y-x|^d}\left(R^2-x\cdot y - \frac{R^{d-2}}{|x|^{d-2}}\frac{|y-x|^d}{\left|\mathcal{K}(x)-y\right|^d}\left(R^2-\mathcal{K}(x)\cdot y\right)\right), \qquad y \in S_1(0).$$

Combining this with the fact that (see (46))

$$\left|\mathcal{K}(x)-y\right| = \frac{R}{|x|}|x-y|, \qquad |y| = R,$$

we infer that

$$-\frac{\partial}{\partial n_y}G(x,y) = \frac{1}{|S_1|R|y-x|^d}\left(R^2-x\cdot y - \frac{|x|^2}{R^2}\left(R^2-\mathcal{K}(x)\cdot y\right)\right)$$

$$= \frac{R^2-|x|^2}{|S_1|R|y-x|^d}, \qquad y \in S_R(0). \tag{51}$$

Notice that as $x \to 0$, the function found is continuous for all $y \in S_R(0)$, so there is no need to consider the case $x = 0$ separately. (As an alternative check, the expression (48) can be used directly, to yield the constant value $1/(|S_1|R^{d-1})$, i.e. (51) — recall that $|y| = R$ on $S_R(0)$.)

Finally, substituting this formula into (41), we obtain

$$u(x) = \frac{R^2-|x|^2}{|S_1|R}\int_{S_R(0)}\frac{\varphi(y)}{|x-y|^d}dS_y, \qquad x \in B_R(0). \tag{52}$$

This is the so-called *Poisson's formula*, see also Section 5.3.

Remark 5.2 (For deeper understanding of the Poisson's formula). *Note that the integral (52) is divergent for $x = (x_1, \dots, x_d) \in S_R(0)$. Indeed, denote $\widetilde{x} = (x_1, \dots, x_{d-1}) \in \mathbb{R}^{d-1}$ (i.e. we keep $d-1$ variables out of d). Then the integral of the function $|\widetilde{x} - z|^{-d}$, $z \in \mathbb{R}^{d-1}$, over the $(d-1)$-dimensional ball $\widetilde{B}_1(\widetilde{x}) \subset \mathbb{R}^{d-1}$ centred at $\widetilde{x}$ is divergent (notice how the variable $z \in \widetilde{B}_1(\widetilde{x})$ plays the role of the variable $y \in S_R(0)$ in (52), where we simply set $\varphi(y) = 1$):*

$$\int_{\widetilde{B}_1(\widetilde{x})} \frac{dz}{|\widetilde{x} - z|^d} = \infty.$$

[*To see this, you may wish to set $d = 2$ and notice that*

$$\int_{x_1 - 1}^{x_1 + 1} \frac{dz}{(x_1 - z)^2} = \infty, \qquad x_1 \in \mathbb{R}.]$$

Therefore the formula (52) does not make sense for $x \in S_R(0)$, as the integral on the right-hand side is divergent, except for points $x \in S_R(0)$ for which φ vanishes with some of its derivatives (remember the Hadamard's form of the Taylor formula! — Corollary 14.7.) However, the limit of $u(x)$ given by (52) exists and, as x tends towards a point $\widehat{x} \in S_R(0)$, is given by $\varphi(\widehat{x})$ (as it should, since $\varphi(\widehat{x})$ is a boundary value of the function u.) This is not too surprising as we have the coefficient $R^2 - |x|^2$ at the front of the integral and in the limit as $x \to \widehat{x}$ one may well have $0 \cdot \infty = \phi(\widehat{x})$, which is what the formula (52) is telling us!

As an application of the existence of the solution to the Dirichlet problem in a ball, we prove the following statement. (Not examinable)

Theorem 5.3 (Inverse Mean-Value Theorem). *Suppose that $u \in C(\overline{\Omega})$ and for any ball $B_r(x_0) \subset \Omega$ one has*

$$u(x_0) = \frac{1}{|B_r|} \int_{B_r(x_0)} u$$

Then u is harmonic in Ω.

Proof. Fix $\rho > 0$, $\widehat{x} \in \Omega$ and consider a ball $B_\rho(\widehat{x}) \subset \Omega$. Denote by φ the function on $S_\rho(\widehat{x})$ obtaind by taking the boundary values of u on $S_\rho(\widehat{x})$:

$$u|_{S_\rho(\widehat{x})} = \varphi.$$

Consider the solution $v \in C^2(\overline{B_\rho(\widehat{x})})$ to the problem

$$\begin{cases} \Delta v = 0 \quad \text{in} \quad \Omega, \\ v|_{S_\rho(\widehat{x})} = \varphi. \end{cases}$$

By the above, this problem has a solution, found by the Poisson's formula. Consider the function $u - v$ on $B_\rho(\widehat{x})$. It satisfies the conditions of the Second Mean-Value Theorem (Theorem 3.6) for harmonic functions in $B_\rho(\widehat{x})$ and hence the Maximum Principle (Theorem 3.7) for it holds, as discussed in Remark 3.9: it attains its maximum in $B_\rho(\widehat{x})$ on its boundary, i.e. on $S_\rho(\widehat{x})$. At the same time, we have

$$(u - v)|_{S_\rho(\widehat{x})} = 0$$

by construction, and therefore $u = v$ everywhere on $B_\rho(\widehat{x})$. Due to the arbitrary choice of ρ and $\widehat{x}$, the functions u and v coincide in the whole of Ω, and since v is harmonic in Ω, so is u. □

5.3 Green's Function for Specific Regions by the Method of Images

The second example of a region Ω for which can find an explicit expression for the Green's function is a half-plane. However, a slight complication in this case is that Ω is unbounded. When Ω is an unbounded domain, all the ideas involving Green's functions remain valid (including a straightforward generalisation of Definition (4.9)) if we impose a suitable "boundary condition" at infinity on G, i.e. $G(x,y) \to 0$ as $y \to \infty$, for every $x \in \Omega$ We also need some (natural) constraints on how the given boundary conditions for u behave at infinity — these are best illustrated by example (see the discussion beneath equation (58) below).

Example 5.4 (Green's function for the half-plane)**.** *Let*

$$\Omega = \big\{(x_1, x_2) \in \mathbb{R}^2 : x_2 > 0\big\}.$$

Given $x = (x_1, x_2) \in \Omega$, put $\mathcal{R}(x) = (x_1, -x_2)$ (reflection across the x_1-axis, i.e. the boundary of the half-plane). Then

$$G(x,y) = \frac{1}{2\pi}\left(\log\frac{1}{|x-y|} - \log\frac{1}{\big|\mathcal{R}(x)-y\big|}\right), \qquad x \in \Omega, \quad y \in \overline{\Omega}, \quad x \neq y. \tag{53}$$

We would like to verify that the three conditions in Definition 4.9 hold for the expression (53). Indeed, we immediately notice that for each $x \in \Omega$, the function

$$g(x,\cdot) = G(x,\cdot) - \frac{1}{2\pi}\log\frac{1}{|x-\cdot|} = -\frac{1}{2\pi}\log\frac{1}{\big|\mathcal{R}(x)-\cdot\big|},$$

is an element of $C^2(\overline{\Omega})$, as for each $x \in \Omega$ the point $\mathcal{R}(x)$ lies in the lower half-plane, and so $\mathcal{R}(x) \neq y$ for any $y \in \Omega$. Now, how can we see that second condition in the definition of the Green's function

$$\Delta_y g(x,y) = 0, \qquad x \in \Omega, \quad y \in \overline{\Omega},$$

is satisfied? We can do a direct calculation:

$$\frac{\partial}{\partial y_j}\log\big|\mathcal{R}(x)-y\big| = \frac{1}{\big|\mathcal{R}(x)-y\big|}\left\{-\frac{\big(\mathcal{R}(x)-y\big)_j}{\big|\mathcal{R}(x)-y\big|}\right\}, \qquad j = 1,2, \tag{54}$$

$$\begin{aligned}\frac{\partial^2}{\partial y_j^2}\log\big|\mathcal{R}(x)-y\big| &= \frac{1}{\big|\mathcal{R}(x)-y\big|^2} \\ &\quad + \Big\{-\big(\mathcal{R}(x)-y\big)_j\Big\}\left\{-\frac{2}{\big|\mathcal{R}(x)-y\big|^3}\right\}\left\{-\frac{\big(\mathcal{R}(x)-y\big)_j}{\big|\mathcal{R}(x)-y\big|}\right\}, \qquad j = 1,2, \end{aligned} \tag{55}$$

and hence

$$\Delta_y \log\big|\mathcal{R}(x)-y\big| = \sum_{j=1}^{d}\frac{\partial^2}{\partial y_j^2}\log\big|\mathcal{R}(x)-y\big| = \frac{2}{\big|\mathcal{R}(x)-\cdot\big|^2} - 2\frac{\sum_{j=1}^2\big(\mathcal{R}(x)-y\big)_j^2}{\big|\mathcal{R}(x)-y\big|^4} = 0, \tag{56}$$

as required.

However, if you have checked in the past that

$$\Delta \log \frac{1}{|x|} = 0, \qquad x \neq 0, \tag{57}$$

for example when checking that the fundamental solution Φ *in two dimensions (see the second formula in (30)) satisfies the Laplace's equation away from the origin, you will probably have a nagging feeling that the calculation (54)–(56) is very similar, only that* y *now plays the role of* x *in (57), and of course replacing* y *by* $\mathcal{R}(x) - y$ *does not affect the outcome! We already came across this kind of situation when we had to verify the same property for the function* g *in (43), see the discussion just below (44).*

Finally, the third condition in the definition of a Green's function is satisfied automatically, as for any $x \in \Omega$ *one has*

$$|x - y| = \big|\mathcal{R}(x) - y\big| \qquad \forall y \in \partial\Omega$$

by symmetry, and so (53) vanishes for $y \in \partial\Omega$.

What does the formula (41) give us, assuming we can apply it in the case of a half-plane? For this, we evaluate the normal derivative of the Green's function, bearing in mind that the normal to the boundary at its point y *is given by* $n_y = (0, -1)$:

$$\begin{aligned}
\frac{\partial G}{\partial n_y}(x, y) &= \nabla_y G(x, y) \cdot n_y = -\frac{\partial G}{\partial y_2}(x, y) \\
&= \frac{1}{2\pi}\left(\frac{1}{|x-y|}\left\{-\frac{(x-y)_2}{|x-y|}\right\} - \frac{1}{\big|\mathcal{R}(x) - y\big|}\left\{-\frac{\big(\mathcal{R}(x) - y\big)_2}{\big|\mathcal{R}(x) - y\big|}\right\}\right) \\
&= -\frac{1}{2\pi}\left(\frac{x_2 - y_2}{(x_1 - y_1)^2 + (x_2 - y_2)^2} - \frac{-x_2 - y_2}{(x_1 - y_1)^2 + (-x_2 - y_2)^2}\right) \\
&\overset{y_2 = 0}{=} \quad -\frac{1}{\pi}\frac{x_2}{(x_1 - y_1)^2 + x_2^2}, \qquad x \in \Omega, \quad y_1 \in \mathbb{R}.
\end{aligned}$$

Hence, recalling (41), we obtain an analogue of the Poisson's formula (52) for the case of a half-plane:

$$u(x) = \frac{x_2}{\pi}\int_{-\infty}^{\infty} \frac{\varphi(y_1)}{(x_1 - y_1)^2 + x_2^2} dy_1, \qquad x = (x_1, x_2) \in \Omega, \tag{58}$$

which is the desired formula for the solution to the Dirichlet problem, in terms of the boundary data (i.e. the harmonic lift of the boundary data φ *to the half-plane* Ω*.)*

Now it remains to notice that in order for the integral in (58) to make sense and yield finite values for $u(x)$*, we should assume that the boundary function* φ *is "sufficiently nice"; e.g. piecewise continuous and bounded on* $\mathbb{R}$ *suffices.*

Example 5.5 (Green's function for the half-space in three dimensions)**.** *Suppose that*

$$\Omega = \big\{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_3 > 0\big\}.$$

Given $x = (x_1, x_2, x_3) \in \Omega$*, put* $\mathcal{R}(x) = (x_1, x_2, -x_3)$ *(reflection across the* (x_1, x_2)*-plane, i.e. the boundary of the half-space). Then, proceeding as above, while bearing in mind the first formula in*

(30), we obtain

$$G(x,y) = \frac{1}{2\pi}\left(\frac{1}{|x-y|} - \frac{1}{\big|\mathcal{R}(x)-y\big|}\right), \qquad x \in \Omega, \quad y \in \overline{\Omega}, \quad x \neq y.$$

The normal derivative on the boundary $\{x_3 = 0\}$ *is*

$$\begin{aligned}
\frac{\partial G}{\partial n_y}(x,y) &= \nabla_y G(x,y)\cdot n_y = -\frac{\partial G}{\partial y_3}(x,y) \\
&= -\frac{1}{4\pi}\left(-\frac{1}{|x-y|^2}\left\{-\frac{(x-y)_3}{|x-y|}\right\} + \frac{1}{\big|\mathcal{R}(x)-y\big|^2}\left\{-\frac{\big(\mathcal{R}(x)-y\big)_3}{\big|\mathcal{R}(x)-y\big|}\right\}\right) \\
&= -\frac{1}{4\pi}\left(\frac{x_3-y_3}{\big\{(x_1-y_1)^2+(x_2-y_2)^2+(x_3-y_3)^2\big\}^{3/2}}\right. \\
&\qquad\qquad \left. - \frac{-x_3-y_3}{\big\{(x_1-y_1)^2+(x_2-y_2)^2+(-x_3-y_3)^2\big\}^{3/2}}\right) \\
&\overset{y_3=0}{=} \; -\frac{1}{2\pi}\frac{x_3}{\big\{(x_1-y_1)^2+(x_2-y_2)^2+x_3^2\big\}^{3/2}}, \qquad x \in \Omega, \quad y_1 \in \mathbb{R},
\end{aligned}$$

and so (cf. (41),) the Poisson's formula for the half-space is as follows:

$$u(x) = \frac{x_3}{2\pi}\int_{\mathbb{R}^2} \frac{\varphi(y_1,y_2)}{\big\{(x_1-y_1)^2+(x_2-y_2)^2+x_3^2\big\}^{3/2}} dy_1, \qquad x = (x_1,x_2,x_3) \in \Omega,$$

Example 5.6 (Green's function for the disc)**.** *Let* $\Omega = B_R(0) \subset \mathbb{R}^2$, *for some* $R > 0$. *As in Section 5.1, consider the Kelvin transform ("inversion through the circle")* $\mathcal{K}$ *defined by*[13]

$$\mathcal{K}(x) = \frac{R^2}{|x|}x, \qquad x \in \Omega.$$

Then, for $y \in \overline{\Omega}$ *and* $x \in \Omega$, $x \neq y$, *one has*

$$G(x,y) = \begin{cases} \dfrac{1}{2\pi}\left(\log\dfrac{1}{|y|} - \log\dfrac{1}{R}\right), & x = 0 \\ \dfrac{1}{2\pi}\left\{\log\dfrac{1}{|x-y|} - \log\left(\dfrac{R}{|x|}\dfrac{1}{\big|\mathcal{K}(x)-y\big|}\right)\right\}, & x \in \Omega\setminus\{0\} \end{cases}$$

$$= \frac{1}{2\pi}\left(\log\frac{1}{|x-y|} - \log\frac{1}{h(x,y)R}\right), \qquad x \in \Omega,$$

[13]Note that (cf. (46))

$$\frac{|x|}{R}\big|y - \mathcal{K}(x)\big| = |y-x|, \qquad y \in \partial\Omega, \ x \in \Omega.$$

where $h(x, y)$ is defined by (50).

It is easy to check that

$$\frac{\partial G}{\partial n_y}(x, y) = -\frac{R^2 - |x|^2}{2\pi R|x-y|^2}, \qquad y \in \partial\Omega, \quad x \in \Omega. \tag{59}$$

Substituting (59) into the formula (41), we find the the solution to the Dirichlet problem fo Laplace's equation on the disc $B_R(0)$ is given by

$$u(x) = \frac{R^2 - |x|^2}{2\pi R} \int_{|y|=R} \frac{\varphi(y)}{|x-y|^2} dS_y, \qquad x \in \overline{B}_R, \tag{60}$$

or, rewriting the integral in polar coordinates, $u(r\cos\phi, r\sin\phi) =: \widetilde{u}(r, \phi)$,

$$\widetilde{u}(r, \phi) = \frac{R^2 - r^2}{2\pi} \int_0^{2\pi} \frac{\widetilde{\phi}(\alpha)}{r^2 + R^2 - 2rR\cos(\phi - \alpha)} d\alpha, \qquad r \in [0, R], \quad \phi \in [0, 2\pi). \tag{61}$$

where $\widetilde{\varphi}(\alpha) := \varphi(\cos\alpha, \sin\alpha)$, $\alpha \in [0, 2\pi]$.

The formula (60), equivalently (61), is sometimes referred to as Poisson's formula for the disc, with the kernel in the integral, i.e. the expression

$$\frac{1}{|x-y|^2} = \frac{1}{r^2 + R^2 - 2rR\cos(\phi - \alpha)}, \qquad x = (r\cos\phi, r\sin\phi), \quad y = (R\cos\alpha, R\sin\alpha),$$

called the Poisson's kernel (for the disc).

So far the only domains for which we have found explicit expressions for the Green's function have been domains for which you have already found explicit expressions for the solution using separation of variables or transform methods. The next example is for a domain in which you have not yet found an explicit expression for the solution.

Example 5.7 (Green's function for the quarter-plane)**.** *Let*

$$\Omega = \big\{(x_1, x_2) \in \mathbb{R}^2 : x_1 > 0, x_2 > 0\big\}.$$

Consider the mappings

$$\mathcal{R}_1(x) := (x_1, -x_2), \qquad \mathcal{R}_2(x) := (-x_1, x_2), \qquad \mathcal{R}_{12}(x) := (-x_1, -x_2), \qquad x \in \Omega,$$

i.e. reflection about the x_1-axis, the reflection about the x_2-axis, the reflection with respect to the origin (equivalently, sequential reflection about the x_1- and x_2-axes), see Fig. 7.

Then

$$G(x, y) = \frac{1}{2\pi}\left(\log\frac{1}{|x-y|} - \log\frac{1}{\big|\mathcal{R}_1(x) - y\big|} - \log\frac{1}{\big|\mathcal{R}_2(x) - y\big|} + \log\frac{1}{\big|\mathcal{R}_{12}(x) - y\big|}\right), \qquad x \in \Omega, \quad y \in \overline{\Omega}.$$

Remark 5.8 (Method of Images)**.** *More generally, for some regions Ω, the Green's function can be sought in the form*

$$G(x, y) = \Phi(x - y) + \sum_{j=1}^{N} \alpha_j(x)\Phi\big(\mathcal{R}_j(x) - y\big) + \beta(x), \qquad x \in \Omega, \quad y \in \overline{\Omega}. \tag{62}$$

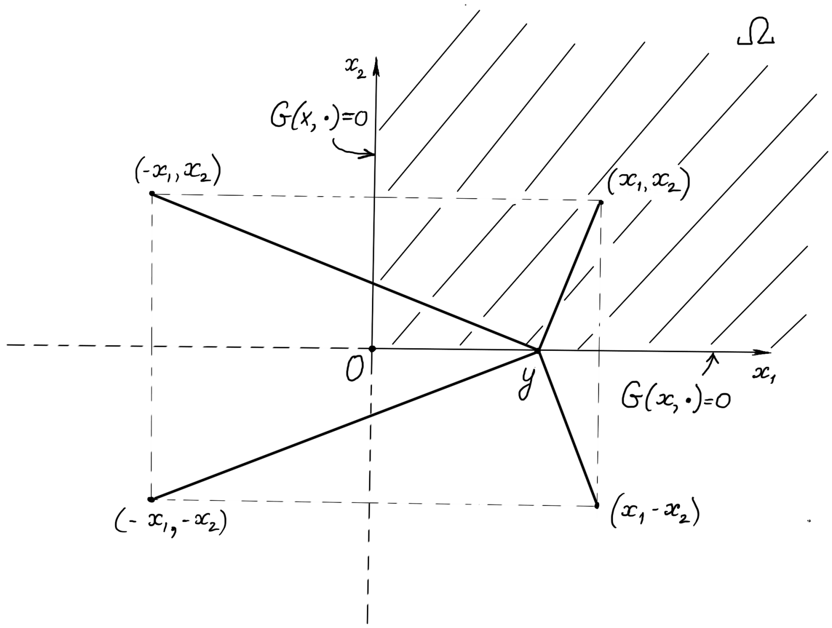

Figure 7: Green's function for a quadrant ("infinite $\pi/2$ corner"). Adding together the three "images" of the fundamental solution in the coordinate axes, taken with appropriate signs, produces a harmonic function, which when added to the fundamental solution itself makes the function $G(x,\cdot)$ vanish on the boundary of the quadrant for each $x \in \Omega$.

In the formula (62), the functions $\mathcal{R}_j$ are the reflections about lines and inversions through circles. The three properties of a Green's function (see Definition (4.9)) can satisfied by a suitable choice of $\mathcal{R}_j$, α_j, $j = 1, \ldots, N$, and β. Indeed:

(1) The requirement that $\mathcal{R}_j(x) \notin \Omega$ (satisfied by construction) ensures that for

$$g(x,y) := G(x,y) - \Phi(x-y) \in C^2(\overline{\Omega}), \qquad x \in \Omega, \quad y \in \overline{\Omega},$$

one has $g(x,\cdot) \in C^2(\overline{\Omega})$ for all $x \in \Omega$.

(2) One has $\Delta g(x,\cdot) = 0$ for all $x \in \Omega$.

(3) The functions α_j, $j = 1, \ldots, N$, and β are chosen so that for each $x \in \Omega$ one has

$$G(x,y) = 0, \qquad y \in \partial\Omega.$$

Identifying the regions for which this is possible requires some group theory (in particular, reflection groups).

6 Physical Derivation of the Heat Equation (Not examinable)

Consider a homogeneous rod $[a,b]$, thermally insulated along the surface and sufficiently thin so we can assume that any moment of time the temperature is constant at all points of every cross-section. If the ends of the rod are maintained at constant temperatures u_1, u_2, then a linear distribution of temperature is established:

$$u(x) = u_1 + \frac{u_2 - u_1}{b - a} x. \tag{63}$$

At the same time, heat flows from the warmer to the cooler end. The amount of heat Q flowing through a cross-sectional area S in a unit of time is given by the experimental formula

$$Q = -k \frac{u_2 - u_1}{b - a} S = -k \frac{\partial u}{\partial x} S, \tag{64}$$

where k is the coefficient of thermal conductivity, which depends on the material of the rod. Propagation of heat in the rod may be described by a function $u = u(x,t)$ representing the temperature of a section x at time t. Let us find the equation that u must satisfy. For this, we first formulate the physical laws that describe heat propagation.

1. Fourier's Law. If the temperature of a body is non-uniform (unlike in (63)), heat currents arise in it, directed from points of higher temperature to points of lower temperature. The amount of heat flowing through section x in the time interval $(t, t + dt)$ is equal to

$$dQ = -k \frac{\partial u}{\partial x} dt, \tag{65}$$

which is a generalisation of (64). Integrating (65), we obtain

$$Q(t_2) - Q(t_1) = -S \int_{t_1}^{t_2} k \frac{\partial u}{\partial x}(x,t) dt,$$

for the amount of heat that has flown through the section x in the time interval (t_1, t_2). If the rod is heterogeneous, then the coefficient k is a function of space x.

2. The amount of heat that must be given to a homogeneous body in order to increase its temperature by $u(x,t_2) - u(x,t_1)$ is equal to

$$Q(t_2) - Q(t_1) = c\rho V \big(u(x,t_2) - u(x,t_1)\big) u,$$

where c is a constant called "specific heat", ρ is the density, V is the volume. If the change of temperature is a function of x, or if the rod is heterogeneous, then

$$Q = S \int_a^b c(x)\rho(x) \big(u(x,t_2) - u(x,t_1)\big) dx,$$

where a, b are the end-points of the rod.

3. Inside the rod, heat may be produced or absorbed (e.g. by passage of current, as a result of chemical reaction, etc.) The generation of heat may be characterised by the density of heat sources at the point x at time t. As a result, in a length of rod $(x, x + dx)$ in a time interval $(t, t + dt)$, there is an amount of heat generated

$$dQ = SF(x,t) dx dt. \tag{66}$$

For the spatial interval (a, b) and time interval (t_1, t_2), the amount of heat is obtained by integrating (66):

$$Q(t_2) - Q(t_1) = S \int_{t_1}^{t_2} \int_a^b F(x,t) dx dt. \tag{67}$$

Now we can obtain the equation of heat conduction, by considering the balance of heat for given spatial and time intervals (a,b), (t_1,t_2) :

$$\int_{t_1}^{t_2}\left(k(x)\frac{\partial u}{\partial x}(x,t)\bigg|_{x=b} - k\frac{\partial u}{\partial x}(x,t)\bigg|_{x=a}\right)dt + \int_a^b\int_{t_1}^{t_2} F(x,t)dt = \int_a^b c(x)\rho(x)\big(u(x,t_2)-u(x,t_1)\big)dx. \tag{68}$$

In order to obtain the differential form of of this balance law, let us first use the integral mean-value theorem in each term of (68) to obtain

$$\begin{aligned}\left(k\frac{\partial u}{\partial x}(x,t)\bigg|_{x=b} - k\frac{\partial u}{\partial x}(x,t)\bigg|_{x=a}\right)\bigg|_{t=t^{(1)}}(t_2-t_1) + F\big(x^{(1)},t^{(2)}\big)(b-a)(t_2-t_1)\\ = c(x)\rho(x)\big(u(x,t_2)-u(x,t_1)\big)\big|_{x=x^{(2)}}(b-a).\end{aligned} \tag{69}$$

Now using the derivative mean-value theorem (for the first term on the left-hand side of (69) and for its right-hand side, we obtain

$$\begin{aligned}\frac{\partial}{\partial x}\left(k(x)\frac{\partial u}{\partial x}(x,t)\right)\bigg|_{x=x^{(3)},t=t^{(1)}}(b-a)(t_2-t_1) + F\big(x^{(1)},t^{(2)}\big)(b-a)(t_2-t_1)\\ = c(x)\rho(x)\frac{\partial u}{\partial t}\bigg|_{x=x^{(3)},t=t^{(3)}}(b-a)(t_2-t_1),\end{aligned} \tag{70}$$

where $x^{(j)}\in(x_1,x_2)$, $t^{(j)}\in(t_1,t_2)$, $j=1,2,3$. Dividing (70) by $(b-a)(t_2-t_1)$ and passing to the limit as $x_1,x_2\to x$, $t_1,t_2\to t$, we obtain, for each x,t the equation

$$\frac{\partial}{\partial x}\left(k(x)\frac{\partial u}{\partial x}(x,t)\right) + F(x,t) = c(x)\rho(x)\frac{\partial u}{\partial t}(x,t),$$

or

$$c\rho u_t = (ku_x)_x + F.$$

This is the classical heat equation. If the functions c, ρ k are constant long the rod (i.e. the rod is made from a homogeneous material), then dividing through by $c\rho$ and denoting

$$D := \frac{k}{c\rho}, \qquad F(x,t) := \frac{f(x,t)}{c\rho},$$

we obtain

$$u_t = Du_{xx} + f(x,t), \tag{71}$$

which is the classical form of the heat equation, which we will be using henceforth.

7 Initial Boundary-Value Problem for the Heat Equation on a Bounded Interval

We would like to address the solution of the heat equation (cf. (71))

$$u_t = Du_{xx} + f(x,t), \qquad x\in[a,b],\ t\geq 0, \tag{72}$$

subject to the initial condition

$$u(x,0) = \varphi(x), \qquad x \in [a,b], \tag{73}$$

and boundary conditions

$$u(a,t) = \psi_a(t), \qquad u(b,t) = \psi_b(t), \qquad t \geq 0, \tag{74}$$

where φ, ψ_a, ψ_b are continuous functions. The first observation in tackling this problem is that, due to linearity, the solution u is the sum of the solutions to the three problems obtained by setting two of the three pieces of data (the right-hand side f, the initial profile φ, and the boundary data ψ_a, ψ_b) to zero and leaving the third piece of data arbitrary. Two of these three problems are considered in turn in Sections 7.1, 7.3. It turns out that it suffices to consider these two problems only, as the general problem is immediately reduced to a problem with $\psi_a = \psi_b = 0$; to see this, you may wish to take a peek at Section 7.4.

7.1 Separation of Variables and Solution of the Homogeneous Equation

Let us begin by considering the homogeneous problem first:

$$u_t = Du_{xx}, \qquad x \in [a,b],\ t \geq 0, \tag{75}$$

subject to the initial condition

$$u(x,0) = \varphi(x), \qquad x \in [a,b], \tag{76}$$

and boundary conditions

$$u(a,t) = u(b,t) = 0, \qquad t \geq 0. \tag{77}$$

For this, let us investigate whether there are functions of the form "with separated variables"

$$u(x,t) = X(x)T(t) \tag{78}$$

that satisfy (75) and (77). Substituting (78) into (75) and dividing both sides of the resulting equality by $DX(x)T(t)$, we obtain

$$\frac{1}{D}\frac{T'}{T} = \frac{X''}{X} = -\lambda, \tag{79}$$

where λ must be a constant, as the first and the second expression in (79) are functions of t and x only! It follows that

$$-X'' = \lambda X, \tag{80}$$

$$T' = -D\lambda T. \tag{81}$$

The boundary conditions (77) give

$$X(a) = X(b) = 0. \tag{82}$$

The equation (80) together with boundary conditions (82) forms a Sturm-Liouville problem. In MA30044, the eigenvalues of the corresponding Sturm-Liouville operator were found to be

$$\lambda_n = \left(\frac{n\pi}{b-a}\right)^2, \quad n = 1, 2, 3, \ldots,$$

with the corresponding eigenfunctions (up to a constant factor) given by

$$X_n(x) = \sin\left(n\pi\frac{x-a}{b-a}\right), \qquad x \in [a,b].$$

The general solution to (81) corresponding to these values $\lambda = \lambda_n$ is

$$T_n(t) = T_n(0)\mathrm{e}^{-D\lambda_n t}.$$

Putting together the above information, we see that the functions

$$u_n(x,t) = C_n\mathrm{e}^{-D\lambda_n t}\sin\left(n\pi\frac{x-a}{b-a}\right), \qquad x \in [a,b],\ t \geq 0,$$

where $C_n := T_n(0)$, $n \in \mathbb{N}$, are arbitrary, are solutions of (75) satisfying the boundary conditions (77). Let us try to find coefficients C_n, $n \in \mathbb{N}$, such that the formal series

$$u(x,t) = \sum_{n=1}^{\infty} u_n(x,t) = \sum_{n=1}^{\infty} C_n\mathrm{e}^{-D\lambda_n t}\sin\left(n\pi\frac{x-a}{b-a}\right) \tag{83}$$

satisfies (76):

$$\varphi(x) = u(x,0) = \sum_{n=1}^{\infty} u_n(x,0) = \sum_{n=1}^{\infty} C_n\sin\left(n\pi\frac{x-a}{b-a}\right), \qquad x \in [a,b],$$

i.e. C_n are the Fourier coefficients of the function φ on the interval (a,b) :

$$C_n = \frac{2}{b-a}\int_a^b \varphi(\xi)\sin\left(n\pi\frac{\xi-a}{b-a}\right)d\xi, \qquad n \in \mathbb{N}. \tag{84}$$

Notice that (83) satisfies the boundary condition (77) since each term of it satisfies (77). In summary, the formula (83), where C_n are given by (84) is a formal solution of the problem (75), (76), (77). It can be shown that the series found is convergent, has derivatives of all orders for $t > 0$ and is continuous for $t \geq 0$. We skip the proof of this claim (at least for the time being) and come back to it when we study the heat equation in the whole space $\mathbb{R}^d$ (notice $d = 1$ in the context of the present section).

7.2 Green's Function for the Heat Equation on a Bounded Interval

Here we introduce the Green's function for the heat equation on a bounded interval, subject to zero boundary conditions, and derive equation that it satisfies. Based on the equation derived, we then give a physical interpretation to the Green's function for the heat equation as the response function for an instantaneous point source of heat.

7.2.1 Green's function as a response to an instantaneous point source

Let us transform the solution (83) substituting the integrals for C_n :

$$u(x,t) = \sum_{n=1}^{\infty} u_n(x,t) = \sum_{n=1}^{\infty} C_n\mathrm{e}^{-D\lambda_n t}\sin\left(n\pi\frac{x-a}{b-a}\right)$$

$$= \sum_{n=1}^{\infty} \left(\frac{2}{b-a} \int_a^b \varphi(\xi) \sin\left(n\pi \frac{\xi - a}{b-a} \right) d\xi \right) \mathrm{e}^{-D\lambda_n t} \sin\left(n\pi \frac{x-a}{b-a} \right)$$

$$= \int_a^b \left\{ \frac{2}{b-a} \sum_{n=1}^{\infty} \mathrm{e}^{-D\lambda_n t} \sin\left(n\pi \frac{x-a}{b-a} \right) \sin\left(n\pi \frac{\xi - a}{b-a} \right) \right\} \varphi(\xi) d\xi. \tag{85}$$

Remark 7.1. *This rearrangement of the order of summation and integration is always valid for* $t > 0$, *because the series in brackets converges uniformly in* ξ *for* $t > 0$.

Let us define

$$G(x, \xi, t) = \frac{2}{b-a} \sum_{n=1}^{\infty} \mathrm{e}^{-D\lambda_n t} \sin\left(n\pi \frac{x-a}{b-a} \right) \sin\left(n\pi \frac{\xi - a}{b-a} \right), \qquad x,\, \xi \in [a, b], \quad t \geq 0. \tag{86}$$

The function G, with which we can represent the solution u as

$$u(x, t) = \int_a^b G(x, \xi, t) \varphi(\xi) d\xi, \tag{87}$$

is referred to as the instantaneous point-source function (as it represents the temperature effect of an instantaneous point source of heat), or the Green's function for the heat equation. Formally, it solves the equation

$$G_t(x, \xi, t) = DG_{xx}(x, \xi, t) + \delta(t)\delta(x - \xi), \tag{88}$$

where δ is the delta-function you came across in MA20220 and MA30044.

Note: the remainder of Section 7.2.1 is not examinable

Consider the inhomogeneous problem for a Sturm-Liouville operator L, subject to some boundary conditions at the end-points of the interval $[a, b]$:

$$Lv + sv = f, \qquad s \geq 0, \tag{89}$$

where we assume for simplicity that the weight function is identically equal to unity, $r(x) = 1$, $x \in [a, b]$.

In MA30044, you looked at expanding the right-hand side f in a series with respect to (normalised) eigenfunctions of L, i.e. ϕ_n such that

$$L\phi_n = \mu_n \phi_n, \qquad \int_a^b |\phi_n|^2 = 1.$$

We will assume that the sequence $\{\phi_n\}$ is orthogonal. This can always be ensured by applying an orthogonalisation procedure to the (eigenfunctions that correspond to the same eigenvalue; eigenfunctions that correspond to different eigenvalues are orthogonal automatically. An expansion like this has the form

$$f = \sum_{n=1}^{\infty} c_n \phi_n, \qquad c_n = \int_a^b f\phi_n.$$

Write the solution v to (89) in the form of a similar series:

$$v = \sum_{n=1}^{\infty} d_n \phi_n.$$

Substituting this into (89) yields

$$\sum_{n=1}^{\infty} d_n \mu_n \phi_n + s \sum_{n=1}^{\infty} d_n \phi_n = \sum_{n=1}^{\infty} c_n \phi_n,$$

from which we obtain

$$d_n = \frac{c_n}{\mu_n + s}, \qquad n = 1, 2, \dots .$$

and hence

$$v = \sum_{n=1}^{\infty} \frac{c_n \phi_n}{\mu_n + s}.$$

Now, what does this formula give for $f(x) = \delta(x - \xi)$, for a fixed $\xi \in (a, b)$? We have

$$c_n = c_n(\xi) = \int_a^b \delta(x - \xi) \phi_n(x) dx = \phi_n(\xi),$$

and so

$$\delta(x - \xi) = \sum_{n=1}^{\infty} \phi_n(\xi) \phi_n(x).$$

Hence, the corresponding solution $v(\cdot) = v(\cdot, \xi, s)$ to

$$Lv(\cdot, \xi, s) + sv(\cdot, \xi, s) = \delta(\cdot - \xi)\theta(s), \tag{90}$$

where

$$\theta(s) = \begin{cases} 1, & s \geq 0, \\ 0, & s < 0 \end{cases}$$

is the "step function", can be written as follows:

$$v(x, \xi, s) = \sum_{n=1}^{\infty} \frac{\phi_n(\xi) \phi_n(x)}{\mu_n + s}. \tag{91}$$

Notice that up to this point we have not used any explicit expression for L. Now, suppose that L is the harmonic Sturm-Liouville operator, i.e. $Lv = Dv_{xx}$, subject to zero boundary conditions at a, b. Its eigenvalues are

$$\mu_n = D\left(\frac{n\pi}{b - a}\right)^2, \qquad n = 1, 2, \dots .$$

and the corresponding normalised eigenfunctions are

$$\phi_n(x) = \sqrt{\frac{2}{b - a}} \sin\left(n\pi \frac{x - a}{b - a}\right), \qquad x \in [a, b],$$

so

$$v(x,\xi,s) = \frac{2}{b-a}\sum_{n=1}^{\infty}\left\{D\left(\frac{n\pi}{b-a}\right)^2 + s\right\}^{-1}\sin\left(n\pi\frac{x-a}{b-a}\right)\sin\left(n\pi\frac{\xi-a}{b-a}\right),$$

Next, notice that the function v, see (91) is the Laplace transform of (cf. (86))

$$G(x,\xi,t) = \sum_{n=1}^{\infty} \mathrm{e}^{-\mu_n t}\phi_n(x)\phi_n(\xi)$$

with respect to the variable t :

$$v(x,\xi,s) = \int_0^{\infty} G(x,\xi,t)\mathrm{e}^{-st}dt =: \mathcal{L}[G](x,\xi,s). \tag{92}$$

Starting with the equation (90) for the function v, let us try to obtain an equation on G that is similar in spirit to (90), in that it has a delta-function on the right-hand side, but also depends on time t. To this end, we rewrite the left-hand side of the equation for v in the form

$$-Dv_{xx} + sv = -D\mathcal{L}[G_{xx}] + \mathcal{L}[G_t] + G(x,\xi,0), \tag{93}$$

and its right-hand side in the form[14]

$$\delta(\cdot-\xi)\theta(s) = s\mathcal{L}\big[\delta(\cdot-\xi)\theta(t)\big](s) = \mathcal{L}\big[\delta(\cdot-\xi)\delta(t)\big](s) + \delta(\cdot-\xi). \tag{94}$$

Here, both for (93) and (94) we use the formula

$$\mathcal{L}[g'](s) = -g(0) + s\mathcal{L}[g](s), \tag{95}$$

which is valid for a continuously differentiable function $g = g(t)$, $t \geq 0$, such that $\mathcal{L}[g]$ is well defined, but here we apply it formally in (94), as the function $\theta(t)$ is not differentiable at $t = 0$. (You can recall the formula (95) from MA20220 or re-derive it by using the definition of $\mathcal{L}$, see (92).)

Putting together (90), (93), and (94), we obtain

$$-D\mathcal{L}[G_{xx}] + \mathcal{L}[G_t] + G(x,\xi,0) = \mathcal{L}\big[\delta(x-\xi)\delta(t)\big] + \delta(x-\xi).$$

Using the fact that

$$G(x,\xi,0) = \sum_{n=1}^{\infty} \mathrm{e}^{-\mu_n t}\phi_n(x)\phi_n(\xi)\Big|_{t=0} = \sum_{n=1}^{\infty}\phi_n(x)\phi_n(\xi)\delta(x-\xi),$$

we infer that

$$-D\mathcal{L}[G_{xx}] + \mathcal{L}[G_t] = \mathcal{L}\big[\delta(x-\xi)\delta(t)\big], \tag{96}$$

and finally, taking the inverse Laplace transform of both sides of (96) yields the claimed equation (88).

It is worth noting that while the above argument is formal, it can be justified rigorously, which however requires tools from functional analysis that are beyond the standard university course of functional analysis.

[14] For the first equality in (94), we use the formal equality $\theta(s) = s\mathcal{L}[\theta](s)$, which we do not currently have the tools to justify rigorously. For the second equality in (94), we use (95) while assuming that $\theta(0) = 1$.

7.2.2 Physical interpretation (Not examinable)

Let us see that the Green's function $G(x, \xi, t)$, considered as a function of x, represents the distribution of of temperature in the rod $a \leq x \leq b$ at the time t if the temperature at the initial time $t = 0$ equals zero and if at this time a certain amount of heat is instantaneously released at the point $x = \xi$ and the ends of the rod are maintained at zero temperature.

The expression "the amount of heat Q released at the point ξ" means, as usual, that we are concerned with the heat released in a "small" interval around the point ξ. The temperature change $\psi_\varepsilon = \psi_\varepsilon(\eta)$, $\eta \in [a, b]$, caused by this heat release is equal to zero outside the interval $(\xi - \varepsilon, \xi + \varepsilon)$, $\varepsilon > 0$, in which the heat is released and inside this interval ψ_ε is assumed to be a positive differentiable function for which

$$c\rho \int_{\xi-\varepsilon}^{\xi+\varepsilon} \psi_\varepsilon = Q, \tag{97}$$

where c is the "specific heat", which is a material constant quantifying the link between the change in temperature and the corresponding change in internal energy, or "amount of heat") and ρ is the mass density (see Section 6).

As discussed above, the process of distribution of temperature is described by (87), with φ replaced by[15] ψ_ε :

$$u_\varepsilon(x, t) = \int_a^b G(x, \eta, t)\psi_\varepsilon(\eta) d\eta, \tag{98}$$

Let us perform a passage to the limit as $\varepsilon \to 0$. Using the intermediate value theorem (see Section 14.1) for $G(x, \xi, t)$ as a function of ξ for fixed values of x, t, and taking into account the equality (97), we obtain[16]

$$u_\varepsilon(x, t) = \int_{\xi-\varepsilon}^{\xi+\varepsilon} G(x, \eta, t)\psi_\varepsilon(\eta) d\eta = G(x, \eta_*, t) \int_{\xi-\varepsilon}^{\xi+\varepsilon} \psi_\varepsilon(\eta) d\eta = G(x, \xi_*, t) \frac{Q}{c\rho},$$

where $\eta_* \in [\xi - \varepsilon, \xi + \varepsilon]$. Using the continuity of the function $G(x, \xi, t)$ at the point ξ for $t > 0$, we therefore have

$$\lim_{\varepsilon \to 0} u_\varepsilon(x, t) = \frac{Q}{c\rho} G(x, \xi, t).$$

It follows that $G(x, \xi, t)$ represents the temperature effect of an instantaneous point source of output $Q = c\rho$, located at time $t = 0$ at the point ξ of the interval (a, b).

7.3 The Inhomogeneous Equation of Heat Conduction

Let us examine the inhomogenenous equation of heat conduction

$$u_t(x, t) = D u_{xx}(x, t) + f(x, t), \qquad x \in [a, b],\ t \geq 0, \tag{99}$$

[15] Notice how we replaced the variable of integration in (98) by η, as ξ is now a fixed point as above, due to the fact that we would like to investigate the physical meaning of the value $G(x, \xi, t)$.

[16] Indeed, using the assumption about ψ_ε being positive, we notice that

$$\min_{\eta \in [a,b]} G(x, \eta, t) \leq \frac{\int_{\xi-\varepsilon}^{\xi+\varepsilon} G(x, \eta, t)\psi_\varepsilon(\eta) d\eta}{\int_{\xi-\varepsilon}^{\xi+\varepsilon} \psi_\varepsilon(\eta) d\eta} \leq \max_{\eta \in [a,b]} G(x, \eta, t),$$

which immediately sets us up for using the intermediate value theorem with initial condition

$$u(x,0) = 0 \tag{100}$$

and boundary conditions

$$u(a,t) = u(b,t) = 0. \tag{101}$$

7.3.1 Mathematical applied mathematician's derivation

We shall search for a solution as a Fourier series in $\sin(n\pi(x-a)/(b-a))$, $n \in \mathbb{N}$, which we know to be the eigenfunctions of the harmonic Sturm-Liouville operator (cf. the first term on the right-hand side of (99)), subject to zero boundary conditions at a, b (cf. (101)):

$$u(x,t) = \sum_{n=1}^{\infty} u_n(t) \sin\left(n\pi \frac{x-a}{b-a}\right), \qquad x \in [a,b], \quad t \geq 0. \tag{102}$$

In order to determine the functions $u_n = u_n(t)$, let us represent the function $f = f(x,t)$ as a series

$$f(x,t) = \sum_{n=1}^{\infty} f_n(t) \sin\left(n\pi \frac{x-a}{b-a}\right), \qquad x \in [a,b], \quad t \geq 0,$$

where

$$f_n(t) = \frac{2}{b-a} \int_a^b f(\xi,t) \sin\left(n\pi \frac{\xi-a}{b-a}\right) d\xi, \qquad n \in \mathbb{N}. \tag{103}$$

Substituting (102) into (99), we have

$$\sum_{n=1}^{\infty} \sin\left(n\pi \frac{x-a}{b-a}\right) \left\{ \left(\frac{n\pi}{b-a}\right)^2 D u_n(t) + \dot{u}_n(t) - f_n(t) \right\} = 0.$$

This equation will be satisfied if all the coefficients in the bracket vanish, i.e.

$$\dot{u}_n(t) = -D \left(\frac{n\pi}{b-a}\right)^2 u_n(t) + f_n(t), \qquad n \in \mathbb{N}. \tag{104}$$

From the initial condition (100), we have

$$u(x,0) = \sum_{n=1}^{\infty} u_n(0) \sin\left(n\pi \frac{x-a}{b-a}\right) = 0,$$

and therefore

$$u_n(0) = 0 \qquad \forall n. \tag{105}$$

Solving (104) subject to (105), we obtain

$$u_n(t) = \int_0^t \mathrm{e}^{-D\lambda_n(t-\tau)} f_n(\tau) d\tau, \qquad n = 1, 2, \ldots$$

Substituting these expressions into (102), we obtain the solution to the original problem in the form

$$u(x,t) = \sum_{n=1}^{\infty} \int_0^t \mathrm{e}^{-D\lambda_n(t-\tau)} f_n(\tau) d\tau \sin\left(n\pi \frac{x-a}{b-a}\right).$$

Combining the expressions (103) with the last formula, we can transform the solution to

$$
\begin{aligned}
u(x,t) &= \int_0^t \int_a^b \left\{ \frac{2}{b-a} \sum_{n=1}^{\infty} \mathrm{e}^{-D\lambda_n(t-\tau)} \sin\left(n\pi \frac{x-a}{b-a}\right) \sin\left(n\pi \frac{\xi-a}{b-a}\right) f(\xi,\tau) d\xi d\tau \right. \\
&= \int_0^t \int_a^b G(x,\xi,t-\tau) f(\xi,\tau) d\xi d\tau, \qquad x \in [a,b],\ t \geq 0,
\end{aligned}
\tag{106}
$$

where $G(x,\xi,t)$ coincides with the Green's function defined by the formula (86).

7.3.2 Physical applied mathematician's derivation (Not examinable)

Exercise 7.2. *Following an argument similar to that discussed in Section 7.2.2, show that the expression*

$$
\frac{Q}{c\rho} G(x,\xi_0,t-\tau_0), \qquad \tau_0 > 0, \tag{107}
$$

may be interpreted as representing the effect at the point x at time t of a single instantaneous heat source, releasing an amount of heat Q, located at the point ξ_0 at time τ_0.

If the function $(Q/(c\rho))G(x,\xi,t)$ is known, then it is possible to express the effect of the sources distributed continuously with density $F(x,t) = c\rho f(x,t)$, by the formula (106). Indeed, the temperature effect of the heat sources acting in the space-time region $[\xi_1,\xi_2) \times [\tau_1,\tau_2)$ containing a point (ξ,τ) is given by the expression

$$
G(x,\xi,t-\tau) f(\xi,\tau)(\xi_2-\xi_1)(\tau_2-\tau_1), \tag{108}
$$

since in this case, as follows from the meaning of the source density $F(x,t)$, one has

$$
Q = F(\xi,\tau)(\xi_2-\xi_1)(\tau_2-\tau_1),
$$

and therefore

$$
\frac{Q}{c\rho} = f(\xi,\tau)(\xi_2-\xi_1)(\tau_2-\tau_1). \tag{109}
$$

Combining (107) and (109), we obtain (108), as claimed.

If the sources are distributed continuously, then, we can sum the effects of the sources over the entire space-time region $\xi \in [a,b]$, $\tau \in [0,t]$, each expressed by (108):

$$
\sum_j G\big(x,\xi^{(j)},t-\tau^{(j)}\big) f\big(\xi^{(j)},\tau^{(j)}\big)\big(\xi_2^{(j)}-\xi_1^{(j)}\big)\big(\tau_2^{(j)}-\tau_1^{(j)}\big)
$$

By then passing to the limit as

$$
\max_j\big(\xi_2^{(j)}-\xi_1^{(j)}\big) \to 0, \qquad \max_j\big(\tau_2^{(j)}-\tau_1^{(j)}\big) \to 0,
$$

we obtain precisely the formula (106). (Recall the definition of the integral as the limit of the sum of areas of thin rectangles constructed under the graph of the function.) Thus, having understood the physical significance of the Green's function $G(x,\xi,t)$, it is possible at once to write down (106) for the function that gives the solution to the inhomogeneous equation.

We have considered the inhomogeneous equation with zero initial conditions. If the initial condition differs from zero, then the solution of the homogeneous equation with the given initial condition $u(x,0) = \varphi(x)$ found in Section 7.1 should be added to this solution.

7.4 General Initial Boundary-Value Problem

Consider the general boundary-value problem, cf. (72)–(74):

$$
\begin{aligned}
&u_t(x,t) = Du_{xx}(x,t) + f(x,t), \qquad x \in [a,b],\ t \ge 0,\\
&u(x,0) = \varphi(x), \qquad x \in [a,b],\\
&u(a,t) = \psi_a(t), \quad u(b,t) = \psi_b(t), \qquad t \ge 0.
\end{aligned}
$$

Let us introduce a new unknown function $v = v(x,t)$, representing the deviation of u from a certain known function $U = U(x,t)$. This function v will be defined as the solution to the equation

$$
\begin{aligned}
&v_t(x,t) = Dv_{xx}(x,t) + \widetilde{f}(x,t), \qquad x \in [a,b],\ t \ge 0,\\
&\widetilde{f}(x,t) := f(x,t) - \big(U_t(x,t) - DU_{xx}(x,t)\big) \qquad x \in [a,b],\ t \ge 0,
\end{aligned}
$$

subject to the additional conditions

$$
\begin{aligned}
&v(x,0) = \varphi(x) - U(x), \qquad x \in [a,b],\\
&v(a,t) = \psi_a(t) - U(a,t), \quad v(b,t) = \psi_b(t) - U(b,t) \qquad t \ge 0.
\end{aligned}
$$

Let us choose the auxiliary function $U(x,t)$ so that

$$
\psi_a(t) - U(a,t) = \psi_b(t) - U(b,t) = 0, \qquad t \ge 0
$$

for example

$$
U(x,t) = \psi_a(t) + \frac{x-a}{b-a}\big(\psi_a(t) - \psi_b(t)\big), \qquad x \in [a,b],\ t \ge 0, \tag{110}
$$

for which

$$
\widetilde{f}(x,t) := f(x,t) - \psi_a'(t) - \frac{x-a}{b-a}\big(\psi_a'(t) - \psi_b'(t)\big) \qquad x \in [a,b],\ t \ge 0.
$$

Thus, the determination of the solution u to the general initial boundary-value problem is reduced to a determination of the solution v to an initial-boundary-value problem with zero boundary conditions. The method for finding the function v is based on combining the procedures described in Sections 7.1, 7.3.

Example 7.3. *Consider the problem of determining the temperature in a rod represented by the interval $[a,b]$ whose ends are maintained at constant temperatures u_a, u_b :*

$$
\begin{aligned}
&u_t(x,t) = Du_{xx}(x,t), \qquad x \in [a,b], \quad t \ge 0,\\
&u(x,0) = \varphi(x), \quad x \in [a,b],\\
&u(a,t) = u_a, \quad u(b,t) = u_b.
\end{aligned}
$$

The function U defined by (110) is time-independent, $U(x,t) = \widetilde{U}(x)$, where

$$
\widetilde{U}(x) = u_a + \frac{x-a}{b-a}(u_b - u_a),
$$

and v is the deviation from this constant temperature:

$$
\begin{aligned}
v_t(x,t) &= Dv_{xx}(x,t), \qquad x \in [a,b],\ t \geq 0,\\
v(x,0) &= \varphi(x) - \widetilde{U}(x), \quad x \in [a,b],\\
v(a,t) &= v(b,t) = 0.
\end{aligned}
$$

It can be found easily by the method of separation of variables, as in Section 7.1. As a result, we obtain the following formula for the solution (cf. (85))

$$
\begin{aligned}
u(x,t) = u_a + \frac{x-a}{b-a}(u_b-u_a) + v(x,t) = u_a + \frac{x-a}{b-a}(u_b-u_a)\\
+ \frac{2}{b-a}\sum_{n=1}^{\infty} \mathrm{e}^{-D\lambda_n t} \sin\left(n\pi\frac{x-a}{b-a}\right)\int_a^b \sin\left(n\pi\frac{\xi-a}{b-a}\right)\left(\varphi(\xi) - u_a - \frac{\xi-a}{b-a}(u_b-u_a)\right)d\xi.
\end{aligned}
$$

8 Initial Boundary-Value Problem for the Heat Equation in a Bounded Region of $\mathbb{R}^d$

In Sections 8, 9, we assume that $D = 1$ in (71). This can always be arranged by a suitable time rescaling $\widehat{t} = Dt$ and adjusting the right-hand side of the equation accordingly, $\widehat{f}(x,t) = D^{-1}f(x,t)$.

8.1 Maximum Principle for the Heat Equation

Definition 8.1. *Suppose that Ω is a bounded region in $\mathbb{R}^d$ and $T > 0$. The* **parabolic boundary** *of the cylinder*[17] *$Q_T := \Omega \times (0,T]$ is the set*

$$
(\Omega \times \{0\}) \cup \big(\partial\Omega \times [0,T]\big), \tag{111}
$$

see Fig. 8.

Theorem 8.2 (Maximum Principle)**.** *Suppose that a function $u = u(x,t)$ has two continuous derivatives with respect to x and one continuous derivative with respect to t on the cylinder Q_T, is continuous on $\overline{Q}_T = \overline{\Omega} \times [0,T]$, and satisfies the equation*

$$
u_t(x,t) - \Delta_x u(x,t) = 0, \qquad (x,t) \in Q_T.
$$

Then u attains its maximum and minimum values on the parabolic boundary of Q.

Proof. Note first that since u is continuous on $\overline{\Omega} \times [0,T]$, it has a minimum and a maximum on this set. It suffices to consider the maximum, as for the minimum we can apply the same argument to the function $-u$.

For each $\varepsilon > 0$, define the function $v^\varepsilon(x,t) = u(x,t) + \varepsilon|x|^2$. Then

$$
v_t^\varepsilon - \Delta v^\varepsilon = u_t - \Delta u - \varepsilon\Delta\big(|x|^2\big) = -2d\varepsilon < 0. \tag{112}
$$

[17] Note that the cylinder Q_T is defined in such a way that its parabolic boundary is the set $\overline{Q}_T \setminus Q_T$.

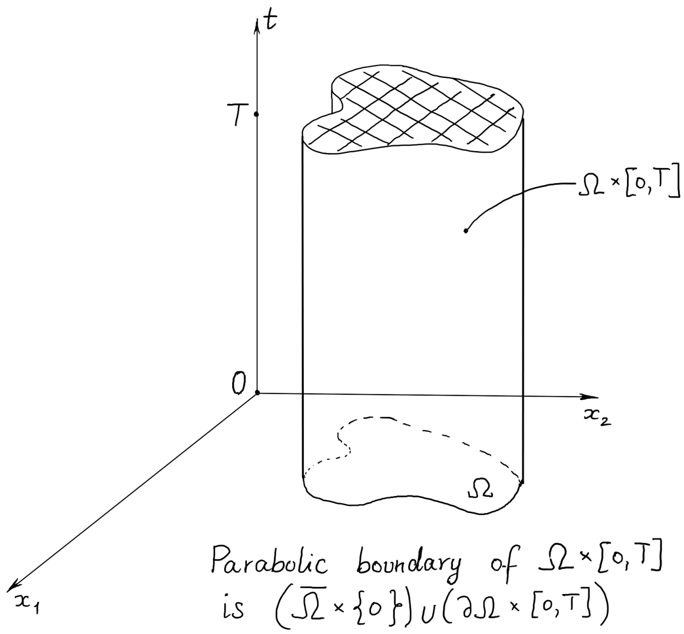

Figure 8: PARABOLIC BOUNDARY OF A CYLINDER IN SPACE-TIME.

The function v^ε must attain its maximum on the parabolic boundary S_T, otherwise we would have $v^\varepsilon_t - \Delta v^\varepsilon \geq 0$ at the point of maximum, which contradicts (112). It follows that

$$u(x,t) \leq v^\varepsilon(x,t) \leq \max_{\overline{Q}_T} v^\varepsilon = \max_{S_T} v^\varepsilon \leq \max_{S_T} u + \varepsilon\big(\max_{x\in\overline{\Omega}} |x|\big)^2.$$

As ε is arbitrarily small, we infer

$$u(x,t) \leq \max_{S_T} u,$$

as claimed. $\square$

Corollary 8.3. *("Principle of Maximum of Modulus") Under the conditions of Theorem 8.2, one has*

$$\max_{Q_T} \big|u(x,t)\big| \leq \max_{S_T} \big|u(x,t)\big|$$

Similarly to the situation we encountered in Section 3.1.2, when studying harmonic functions we proved the uniqueness theorem for the Dirichlet problem for Laplace's equation, we can use the above maximum principle to prove a uniqueness statement for the solution of an initial boundary value problem for the heat equation, as follows.

8.2 Uniqueness Theorem for the Initial Boundary-Value Problem and Other Consequences of the Maximum Principle

Definition 8.4. *Suppose that $T > 0$ and*

$$f = f(x,t), \qquad (x,t) \in Q_T,$$

$$\varphi = \varphi(x), \qquad x \in \Omega,$$
$$\psi = \psi(x,t), \qquad x \in \partial\Omega, \quad t \in [0,T],$$

are continuous. An **initial boundary-value problem** *for the heat equation is a problem of determining* $u = u(x,t)$ *that is continuous on* $\overline{Q}_T$*, has continuous second derivatives in* x *and a continuous derivative in* t *on* Q_T*, and satisfies*

$$\begin{cases} u_t(x,t) - \Delta_x u(x,t) = f(x,t), & (x,t) \in Q_T, \\ u(x,0) = \varphi(x), & x \in \Omega, \\ u(x,t) = \psi(x,t), & x \in \partial\Omega, \quad t \in [0,T]. \end{cases} \tag{113}$$

8.2.1 Uniqueness theorem

Theorem 8.5 (Uniqueness for Initial Boundary-Value Problem)**.** *The initial boundary value problem (113) cannot have two different solutions.*

Proof. Suppose that u_1 and u_2 are solutions of (113). Then $u := u_1 - u_2$ satisfies the conditions of the maximum principle (8.2) and $u(x,t) = 0$, $(x,t) \in S_T$. Therefore, one has $u(x,t) = 0$ for all $(x,t) \in \overline{Q}_T$, and hence $u_1 = u_2$. $\square$

8.2.2 Large-time behaviour of solutions with zero boundary data

The maximum principle sometimes allows one to infer information about the asymptotic behaviour of the solution to the problem (113) at $t \to \infty$, for example when $\psi = 0$ (i.e. the boundary is maintained at zero temperature, with no heat sources inside Ω at any moment of time), as stated in the following theorem.

Theorem 8.6 (Exponential Time Decay)**.** *Let* $T > 0$*, and for given initial* φ *as in Definition 8.4 (i.e. continuous on* Ω*) consider the initial boundary value problem (cf. (113))*

$$\begin{cases} u_t(x,t) - \Delta_x u(x,t) = 0, & (x,t) \in Q_T, \\ u(x,0) = \varphi(x), & x \in \Omega, \\ u(x,t) = 0, & x \in \partial\Omega, \quad t \in [0,T]. \end{cases} \tag{114}$$

If u *solves (114), then the following estimate holds for some* $A, b > 0$ *that are independent of* T *:*

$$\big|u(x,t)\big| \le A\mathrm{e}^{-bt}, \qquad (x,t) \in \overline{Q}_T. \tag{115}$$

Proof. Consider the function
$$v(x,t) = A\mathrm{e}^{-bt}\Pi_{j=1}^{d}\cos(cx_j),$$
where the parameters $A, b, c > 0$ will be determined for the argument below. By a direct calculation, we see that
$$v_t - \Delta v = (-b + dc^2)v,$$
and for a given c (to be chosen in the next step) we choose b so that the right-hand side of the last equation vanishes, i.e. $b = dc^2$.

Furthermore, for sufficiently small c, we have

$$\prod_{j=1}^{d} \cos(cx_j) > 0, \qquad x \in \overline{\Omega}. \tag{116}$$

Let us fix such a value c. Denote by δ the minimum over $x \in \overline{\Omega}$ of the left-hand side of (116) – note that this is positive and depends on c only. Then

$$v(x,0) = A \prod_{j=1}^{d} \cos(cx_j) \geq A\delta \geq \max_{x \in \overline{\Omega}} \bigl|\varphi(x)\bigr|,$$

where the last inequality can be achieved by a suitable choice of A, for example

$$A = \delta^{-1} \max_{x \in \overline{\Omega}} |\varphi(x)|$$

(which thus depends on c.)

Now consider the functions

$$w_{\pm}(x,t) := v(x,t) \pm u(x,t), \qquad (x,t) \in \overline{Q}_T.$$

They have the regularity properties listed in Definition 8.4. Moreover, one has

$$(w_{\pm})_t - \Delta w_{\pm} = 0, \qquad w_{\pm}|_{S_T} \geq 0.$$

It follows, by the maximum principle (Theorem 8.2) that $w_{\pm}$ are non-negative on $\overline{Q}_T$, and therefore

$$\bigl|u(x,t)\bigr| \leq v(x,t) \leq A\mathrm{e}^{-bt}, \qquad x \in \overline{Q}_T,$$

as claimed. $\square$

Corollary 8.7. *If $T = \infty$, then the solution to the problem (114) tends to zero exponentially fast as $t \to \infty$.*

Remark 8.8. *As can be seen from the proof of Theorem 8.6, the constant b depends only on the size of the region Ω (the larger the diameter of Ω, the smaller value for b), while the constant A is based on the choice of c in (116) and the maximum of the initial function φ over Ω. The decay constant b in (115) can be achieved to be as close to dc_*^2 as possible, where c_* is the smallest positive value of c for which the infimum of the product in (116) is zero:*

$$c_* = \min\left\{ c > 0 : \min_{x \in \overline{\Omega}} \prod_{j=1}^{d} \cos(cx_j) = 0 \right\}$$

8.2.3 Stabilisation property

Consider an initial value problem (113) with f and ψ independent of time t :

$$\begin{cases} u_t(x,t) - \Delta_x u(x,t) = f(x), \qquad (x,t) \in Q_T, \\ u(x,0) = 0, \qquad x \in \Omega, \\ u(x,t) = \psi(x), \qquad x \in \partial\Omega, \quad t \in [0,T]. \end{cases} \tag{117}$$

What will happen to the solution $u = u(x,t)$ when $t \to \infty$?

Theorem 8.9 (Stabilisation to the Solution of a Boundary-Value Problem for the Poisson's Equation)**.** *Suppose that $u = u(x,t)$ is a solution of the initial boundary value problem (117) and $v \in C^2(\Omega) \cap C(\overline{\Omega})$ solves the Dirichlet problem (cf. (10))*

$$\begin{cases} -\Delta v = f \quad \text{in } \Omega, \\ v|_{\partial\Omega} = \psi. \end{cases} \tag{118}$$

Then one has

$$\big|u(x,t) - v(x)\big| \leq A\mathrm{e}^{-bt}, \qquad (x,t) \in \overline{Q}_T, \tag{119}$$

where $A, b > 0$ are the constants determined from Theorem 8.6, where the initial condition is taken to be $\varphi = -v$.

Proof. Clearly, the function $\widetilde{u}(x,t) = v(x)$, $x \in \overline{\Omega}$, $t \in [0,T]$, solves the initial boundary value problem

$$\begin{cases} \widetilde{u}_t(x,t) - \Delta_x \widetilde{u}(x,t) = f(x), \qquad (x,t) \in Q_T, \\ \widetilde{u}(x,0) = v(x), \qquad x \in \Omega, \\ \widetilde{u}(x,t) = \psi(x), \qquad x \in \partial\Omega, \quad t \in [0,T]. \end{cases}$$

Subtracting (8.2.3) from (117), we obtain

$$\begin{cases} (u - \widetilde{u})_t(x,t) - \Delta_x (u - \widetilde{u})(x,t) = 0, \qquad (x,t) \in Q_T, \\ (u - \widetilde{u})(x,0) = -v(x), \qquad x \in \Omega, \\ (u - \widetilde{u})(x,t) = 0, \qquad x \in \partial\Omega, \quad t \in [0,T], \end{cases} \tag{120}$$

and applying Theorem 8.6 we obtain the required bound (119). □

Remark 8.10. *For a boundary-value problem with f and ψ independent of t as in the above result and arbitrary initial function φ as in the exponential decay property (Theorem 8.6), i.e.*

$$\begin{cases} u_t(x,t) - \Delta_x u(x,t) = f(x), \qquad (x,t) \in Q_T, \\ u(x,0) = \varphi(x), \qquad x \in \Omega, \\ u(x,t) = \psi(x), \qquad x \in \partial\Omega, \quad t \in [0,T]. \end{cases} \tag{121}$$

we can write the solution as the sum $u = u_1 + u_2$, where u_1 solves an initial boundary-value problem with $\phi = 0$ (of the type discussed in the stabilisation property, Theorem 8.9) and u_2 solves an initial boundary-value problem with $f = 0$, $\psi = 0$ (of the type discussed in the exponential decay property, Theorem 8.6) — we used such an idea when we treated the general initial boundary-value problem on an interval in Section 7. This will immediately yield a stabilisation property for the problem (121), i.e. the estimate (119), where the "asymptotics" v still solves the problem (118), while the constants A, b are determined from Theorem 8.6 on the basis of the initial condition $\varphi - v$, rather than $-v$ as in (120).

In a nutshell, no matter what the initial condition is, the solution to the heat equation with time-independent boundary data ψ (external heat device) and the right-hand side f (internal sources of heat) will converge to the solution of (118) exponentially fast (with the parameters A, b dependent on the initial condition).

9 Cauchy Problem for the Heat Equation on the Whole Space $\mathbb{R}^d$

In what follows $d \in \mathbb{N}$.

Definition 9.1. *Suppose that $T > 0$ and*

$$f = f(x,t), \quad x \in \mathbb{R}^d, \quad t \in [0,T],$$
$$\varphi = \varphi(x), \quad x \in \mathbb{R}^d,$$

are continuous. A **Cauchy problem** *for the heat equation on $\mathbb{R}^d$ is a problem of determining a function $u = u(x,t)$ that is continuous and bounded in $\mathbb{R}^d \times [0,T]$, has continuous second derivatives in x and a continuous derivative in t in $\mathbb{R}^d \times (0,T]$, and satisfies*

$$\begin{cases} u_t(x,t) - \Delta_x u(x,t) = f(x,t), \quad x \in \mathbb{R}^d, \quad t \in (0,T], \\ u(x,0) = \varphi(x), \quad x \in \mathbb{R}^d. \end{cases}$$

9.1 Maximum Principle and Uniqueness Theorem

In this section, we use the following notation:

$$M_+ := \sup_{x\in\mathbb{R}^d,\, t\in[0,T]} u(x,t) \geq \sup_{x\in\mathbb{R}^d,\, t=0} u(x,t) =: N_+,$$
$$M_- := \inf_{x\in\mathbb{R}^d,\, t\in[0,T]} u(x,t) \leq \inf_{x\in\mathbb{R}^d,\, t=0} u(x,t) =: N_-.$$

Theorem 9.2. *Suppose that a function $u = u(x,t)$ satisfies the regularity (i.e. smoothness, boundedness) properties listed in Definition 9.1 and*

$$u_t(x,t) = \Delta_x u(x,t), \qquad x \in \mathbb{R}^d, \quad t \in (0,T]. \tag{122}$$

Then $M_+ = N_+$, $M_- = N_-$.

Proof. (Not examinable) Replacing u by $-u$, the proof of the property $M_- = N_-$ is reduced to the proof of the property $M_+ = N_+$, so it suffices to show the latter.

Consider the function

$$V(x,t,\varepsilon) := u(x,t) - \varepsilon\big(2dt + |x|^2\big), \qquad x \in \mathbb{R}^d, \ \ t \in [0,T], \varepsilon \geq 0.$$

Note that, due to (122), one has

$$V_t - \Delta V = 0 \quad \text{on} \ \ \mathbb{R}^d \times [0,T].$$

For $R > 0$, consider the cylinder $\overline{B_R(0)} \times [0,T] =: Q_T^R$ and note that V, as a function of x, t satisfies the assumptions of the maximum principle for the corresponding initial boundary-value problem on Q_T^R (Theorem 8.2.) Therefore, it attains its maximum on the parabolic boundary of Q_T. Now, on the lower base $\overline{B_R(0)} \cup \{0\}$ of Q_T^R, we have

$$V(x,t,\varepsilon)\big|_{t=0} = u(x,0) - \varepsilon|x|^2 \leq u(x,0) \leq N_+.$$

On the lateral surface of Q_T^R, we have

$$V(x,t,\varepsilon)\big|_{|x|=R} = u(x,t)\big|_{|x|=R} - \varepsilon(2dt + R^2) \le M_+ - \varepsilon R^2.$$

In what follows, we will keep $\varepsilon > 0$ and R linked in such a was that

$$M_+ - \varepsilon R^2 \le N_+, \tag{123}$$

and then let $R \to \infty$.

Consider a point $(x^*, t^*) \in \mathbb{R}^d \times [0,T]$. For any $\varepsilon > 0$, it will become an element of the cylinder Q_T^R, for sufficiently large R, such that (123) is satisfied and will continue staying inside Q_T^R for larger R. But this means that

$$V(x,t,\varepsilon) \le N_+$$

on the parabolic boundary of Q_T^R, for all such R and $\varepsilon \in (0,1]$. Passing to the limit as $\varepsilon \to 0$ in the bound

$$V(x^*,t^*,\varepsilon) = u(x^*,t^*) - \varepsilon\big(2dt^* + |x^*|^2\big) \le N_+,$$

we obtain

$$V(x^*,t^*,\varepsilon) \le N_+.$$

Since (x^*,t^*) was arbitrary, we conclude that $M_+ \le N_+$. $\square$

Corollary 9.3 (Uniqueness Theorem for the Cauchy Problem)**.** *If $u = u(x,t)$, $v = v(x,t)$ solve Cauchy problems with the same data $f = f(x,t)$, $\varphi = \varphi(x)$, then $u = v$ on $\mathbb{R}^d \times [0,T]$.*

Remark 9.4. *If one drops the condition of boundedness in the definition of a Cauchy problem then the uniqueness theorem stops being valid.*

9.2 Fourier Transform and Its Basic Property

For a sufficiently smooth function $u = u(x)$ that decays sufficiently fast as $|x| \to \infty$, we define the Fourier Transform:

$$\mathcal{F}[u](k) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}k\cdot x} u(x)dx, \qquad k \in \mathbb{R}^d,$$

where $k \cdot x$ is the standard inner product in $\mathbb{R}^d$, i.e. $k \cdot x := \sum_{j=1}^d k_j x_j$.

It then turns out that $\mathcal{F} = \mathcal{F}(k)$ is smooth and decays as $|k| \to \infty$. If one defines the Inverse Fourier Transform by the formula

$$\mathcal{F}^{-1}[u](x) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}k\cdot x} u(k)dk, \qquad x \in \mathbb{R}^d,$$

then it turns out that $\mathcal{F}\mathcal{F}^{-1} = \mathcal{F}^{-1}\mathcal{F} = I$, that is $\mathcal{F}^{-1}$ and $\mathcal{F}$ are the inverses of each other. (We do not show this in the present unit.)

Let us note one property of Fourier Transform. For a fixed $j = 1, 2 \ldots, d$, we have

$$\begin{aligned}\mathcal{F}\left[\partial_j u\right](k) &= \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}k\cdot x} \frac{\partial u(x)}{\partial x_j} dx = -\frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \frac{\partial}{\partial x_j}\big(\mathrm{e}^{-\mathrm{i}k\cdot x}\big) u(x)dx \\ &= \frac{\mathrm{i}k_j}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}k\cdot x} u(x)dx = \mathrm{i}k_j \mathcal{F}[u](k), \qquad k \in \mathbb{R}^d.\end{aligned}$$

This shows that

$$\mathcal{F}\partial_j = \mathrm{i}k_j \mathcal{F}, \qquad j = 1, 2, \ldots, d.$$

9.3 Derivation of the Poisson's Formula

Suppose that φ is a bounded continuous function in $\mathbb{R}^d$, and consider the following problem for $u = u(x,t)$:

$$\begin{cases} u_t(x,t) = \Delta_x u(x,t), & x \in \mathbb{R}^d, \quad t \geq 0, \\ u(x,0) = \varphi(x), & x \in \mathbb{R}^d. \end{cases} \tag{124}$$

Taking the Fourier Transform with respect to the variable x of both sides of the equation, we have

$$\frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}k\cdot x} u_t(x,t)dx = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}k\cdot x} \Delta_x u(x,t)dx,$$

or equivalently, after taking the time derivative outside the integral on the left and integrating by parts on the right,

$$\frac{\partial}{\partial t}\left(\frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}k\cdot x} u(x,t)dx\right) = -|k|^2 \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}k\cdot x} u(x,t)dx. \tag{125}$$

Let us use the notation $F(k,t) := \mathcal{F}_x[u](k,t)$, $k \in \mathbb{R}^d$, $t \geq 0$. Then (125) can be written as

$$F_t + |k|^2 F = 0,$$

and hence

$$F(k,t) = F(k,0)\mathrm{e}^{-|k|^2 t}, \qquad k \in \mathbb{R}^d, \quad t \geq 0.$$

Furthermore,

$$F(k,0) = \mathcal{F}_x[u](k,0) = \mathcal{F}[\varphi](k) =: \psi(k), \qquad k \in \mathbb{R}^d,$$

and therefore

$$F(k,t) = \psi(k)\mathrm{e}^{-|k|^2 t}, \qquad k \in \mathbb{R}^d, \quad t \geq 0.$$

It follows that

$$\begin{aligned} u(x,t) &= \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \psi(k)\mathrm{e}^{-|k|^2 t}\mathrm{e}^{\mathrm{i}k\cdot x}dk = \frac{1}{(2\pi)^{d/2}} \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \varphi(y)\mathrm{e}^{-\mathrm{i}k\cdot y}dy\mathrm{e}^{-|k|^2 t}\mathrm{e}^{\mathrm{i}k\cdot x}dk \\ &= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \varphi(y) \int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}k\cdot y - |k|^2 t + \mathrm{i}k\cdot x} dk dy. \end{aligned} \tag{126}$$

Finally notice that[18]

$$\begin{aligned} -\mathrm{i}k\cdot y - |k|^2 t + \mathrm{i}k\cdot x &= -|k|^2 t + \mathrm{i}k\cdot(x-y) \\ &= -\left\{|k\sqrt{t}|^2 + 2k\sqrt{t}\cdot\frac{\mathrm{i}(y-x)}{2\sqrt{t}} + \left(\frac{\mathrm{i}(y-x)}{2\sqrt{t}}\right)^2\right\} - \frac{(y-x)^2}{4t} \\ &= -\left\{k\sqrt{t} + \frac{\mathrm{i}(y-x)}{2\sqrt{t}}\right\} - \frac{(y-x)^2}{4t}, \end{aligned}$$

[18]Here we use the notation $a\cdot b := a_j b_j$ for $a, b \in \mathbb{C}^d$, and in particular $a^2 := a_j a_j$ for $a \in \mathbb{C}^d$ and $c^2 = |c|^2$ for $x \in \mathbb{R}^d$.

and hence

$$\int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}k\cdot y-|k|^2 t+\mathrm{i}k\cdot x}dk = \int_{\mathbb{R}^d} \exp\biggl\{-\Bigl(k\sqrt{t}+\frac{\mathrm{i}(y-x)}{2\sqrt{t}}\Bigr)^2\biggr\}\exp\Bigl(-\frac{(y-x)^2}{4t}\Bigr)dk$$

$$= \exp\Bigl(-\frac{|y-x|^2}{4t}\Bigr)\int_{\mathbb{R}^d} \exp\biggl\{-\Bigl(k\sqrt{t}+\frac{\mathrm{i}(y-x)}{2\sqrt{t}}\Bigr)^2\biggr\}dk$$

$$= \exp\Bigl(-\frac{|y-x|^2}{4t}\Bigr)\frac{1}{(\sqrt{t})^d}\int_{\mathbb{R}^d} \exp\biggl\{-\Bigl(\widetilde{k}+\frac{\mathrm{i}(y-x)}{2\sqrt{t}}\Bigr)^2\biggr\}d\widetilde{k}$$

$$= \exp\Bigl(-\frac{|y-x|^2}{4t}\Bigr)\frac{1}{(\sqrt{t})^d}(\sqrt{\pi})^d = \Bigl(\frac{\pi}{t}\Bigr)^{d/2}\exp\Bigl(-\frac{|y-x|^2}{4t}\Bigr).$$

Combining this with (126), we obtain the final formula

$$u(x,t) = \frac{1}{(2\sqrt{\pi t})^d}\int_{\mathbb{R}^d}\exp\Bigl(-\frac{|x-y|^2}{4t}\Bigr)\varphi(y)dy, \qquad t>0. \tag{127}$$

This is Poisson's formula for the solution to the Cauchy problem for the heat equation on $\mathbb{R}^d$. The formula (127) can be written as

$$u(\cdot,t) = K(\cdot,t)*\varphi, \qquad t>0. \tag{128}$$

where $*$ stands for the convolution:

$$(f*g)(x) := \int_{\mathbb{R}^d} f(x-y)g(y)dy, \qquad x\in\mathbb{R}^d, \tag{129}$$

and

$$K(x,t) = \frac{1}{(2\sqrt{\pi t})^d}\exp\Bigl(-\frac{|x|^2}{4t}\Bigr), \qquad x\in\mathbb{R}^d,\quad t>0.$$

The function K is referred to as the heat kernel, and the formula (128) is the multidimensional version of the formula for the solution to the Cauchy problem for the one-dimensional heat equation.

9.4 Justification of Poisson's Formula (Not examinable)

Suppose that φ is continuous and bounded. The the function $u=u(x,t)$ given by the formula (127) is the solution to the Cauchy problem for the heat equation subject to the initial condition $u(x,0)=\varphi(x)$, $x\in\mathbb{R}$. For each $\delta>0$, the integral on the right-hand side of (127) converges uniformly on in the set

$$\bigl\{(x,t) : x\in\mathbb{R}^d,\ t\geq\delta\bigr\}. \tag{130}$$

Similarly, for any $\alpha_1,\dots,\alpha_d$, β, the integrals

$$\int_{\mathbb{R}^d}\frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}}\cdots\frac{\partial^{\alpha_d}}{\partial x_d^{\alpha_d}}\frac{\partial^\beta}{\partial t^\beta}\exp\Bigl(-\frac{|x-y|^2}{4t}\Bigr)\varphi(y)dy$$

converge uniformly on the set (130). Furthermore, it is verified directly that

$$\Bigl(\frac{\partial}{\partial t}-\Delta_x\Bigr)\frac{1}{(2\sqrt{\pi t})^d}\int_{\mathbb{R}^d}\exp\Bigl(-\frac{|y-x|^2}{4t}\Bigr) = 0$$

It follow that the function u solves

$$u_t - u_{xx} = 0, \qquad x \in \mathbb{R}^d,\ t > 0.$$

Next, we verify that u is continuous up to $t = 0$ and that $\lim_{t\searrow 0} u(x,t) = \varphi(x)$, $x \in \mathbb{R}^d$. To this end, denote $z = (y-x)/(2\sqrt{t})$. Then we have

$$u(x,t) = \frac{1}{(\sqrt{\pi})^d} \int_{\mathbb{R}^d} \mathrm{e}^{-|z|^2} \varphi\big(x + 2\sqrt{t}z\big)dz. \tag{131}$$

This integral on the right-hand side of (131) converges uniformly for $x \in \mathbb{R}^d$, $t \in [0,T]$, for any $T > 0$, and

$$\lim_{t\searrow 0} u(x,t) = \lim_{t\to 0} \frac{1}{(\sqrt{\pi})^d} \int_{\mathbb{R}^d} \exp -|z|^2 \varphi(x + 2\sqrt{t}z)dz = \varphi(x), \qquad x \in \mathbb{R}^d = u(x,0), \qquad x \in \mathbb{R}^d,$$

as required.

Finally, the solution $u = u(x,t)$ is bounded on $\mathbb{R}^d \times [0,T]$, as by the maximum principle (Theorem 9.2) we have

$$\sup_{x\in\mathbb{R}^d,\ t\in[0,T]} \big|u(x,t)\big| \le \sup_{x\in\mathbb{R}^d} \big|\varphi(x)\big| < \infty.$$

9.5 Some Important Remarks

1. (Smoothness of the solution to the heat equation.) If the initial datum φ is continuous and bounded, then for $t > 0$ the solution $u = u(x,t)$ is infinitely smooth.

2. (Infinite speed of heat propagation). If $\varphi \ge 0$ and $\varphi = 0$ outside a ball B_R for some $R > 0$, then $u(x,t) > 0$ for all $x \in \mathbb{R}^d$, $t > 0$.

3. (Nonlocal dependence on the initial data.) The value of the solution u at each fixed point $(x_0, t_0) \subset \mathbb{R}^d \times (0,\infty)$ depends on the values of φ at all points of $x \in \mathbb{R}^d$.

4. (Duhamel's Principle for the solution of the inhomogeneous Cauchy problem for the heat equation.) Consider the problem (cf. (124))

$$\begin{cases} u_t(x,t) = \Delta_x u(x,t) + f(x,t), & x \in \mathbb{R}^d,\ t \ge 0, \\ u(x,0) = \varphi(x), & x \in \mathbb{R}^d. \end{cases}$$

We can write $u = u_1 + u_2$, where u_1, u_2 solve inhomogeneous problems

$$\begin{cases} (u_1)_t(x,t) = \Delta_x u_1(x,t) + f(x,t), & x \in \mathbb{R}^d,\ t \ge 0, \\ u_1(x,0) = 0, & x \in \mathbb{R}^d. \end{cases} \tag{132}$$

$$\begin{cases} (u_2)_t(x,t) = \Delta_x u_2(x,t), & x \in \mathbb{R}^d,\ t \ge 0, \\ u_2(x,0) = \varphi(x), & x \in \mathbb{R}^d. \end{cases}$$

For each $s \geq 0$, consider the solution[19] $U = U(x,t;s)$ to the problem

$$\begin{cases} U_t(x,t;s) = \Delta_x U(x,t;s), & x \in \mathbb{R}^d,\ t \geq 0, \\ U(x,0;s) = f(x,s), & x \in \mathbb{R}^d. \end{cases} \tag{133}$$

Then the function

$$u_1(x,t) = \int_0^t U(x,t-s;s)ds, \qquad x \in \mathbb{R}^d,\ t \geq 0. \tag{134}$$

solves the problem (132). Indeed, one clearly has $u_1(x,0) = 0$, and

$$\begin{aligned}(u_1)_t(x,t) &= U(x,0;t) + \int_0^t U_t(x,t-s;s)ds = f(x,t) + \int_0^t \Delta_x U(x,t-s;s)ds \\ &= f(x,t) + \Delta_x \int_0^t U(x,t-s;s)ds = f(x,t) + \Delta_x u_1(x,t),\end{aligned}$$

as required.

10 Calculus of Variations: Motivation

In this part of the course we will often denote by u_x the derivative u' of a function $u = u(x)$ of a one-dimensional variable x.

10.1 The Brachistochrone Problem

Formulated by Galieo and first "solved" by Johann Bernoulli in 1696. To determine the shape of the curve, joining two prescribed points in space, down which a particle will slide without friction in the least time.

Assume that we have a wire in the (x,y)-plane (see Fig. 9), with x-axis horizontal and y-axis vertical directed downwards, modelled by the graph of a function $y = u(x)$ joining the origin O with the point $F = (x_1, y_1)$, where $x_1 > 0$, $y_1 > 0$.

The particle, whose mass is denoted by m, starts at the origin O with zero velocity and slides towards $A = (x_1, y_1)$ alongs a curve connecting O and A. Denote by $t(x)$ be the time to reach the point $P = (x, u(x))$, by $s(x)$ the arc-length of the curve OP, and by $v(x)$ the speed of the particle at the point P.

Conservation of energy gives the equation

$$\frac{1}{2}m\big(v(x)\big)^2 = mgu(x), \qquad x \in [0, x_1]$$

where g is the gravitational acceleration. It follows that $v(x) = \sqrt{2gu(x)}$, and[20]

$$\frac{dt}{dx} = \frac{dt}{ds} \cdot \frac{ds}{dx} = \left(\frac{ds}{dt}\right)^{-1} \cdot \frac{ds}{dx} = \frac{\sqrt{u_x^2 + 1}}{\sqrt{2gu}}.$$

[19] Here s is can be treated as a parameter, on which the function U of x, t depends through the initial condition in (133). Notice how the Duhamel's formula (134) "combines" the variable t and the parameter s.

[20] Recall that the arc-length element for the graph of a function $u = u(x)$ is given by $ds = \sqrt{u_x^2 + 1}dx$, so that $ds/dx = \sqrt{u_x^2 + 1}$.

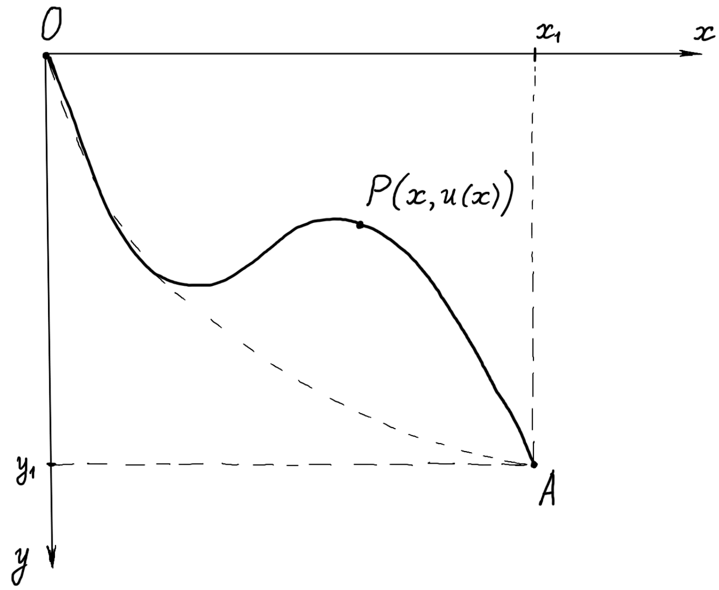

Figure 9: BRACHISTOCHRONE PRONLEM. A point P slides along a curve from O to A.

Therefore, the total travel time[21] is

$$T(u) = \int_0^{x_1} \frac{dt}{dx} dx = \int_0^{x_1} \sqrt{\frac{u_x^2 + 1}{2gu}}. \tag{135}$$

Note that $T(u)$ depends on the whole function u, and the problem is to find u that minimises[22] $T(u)$ subject to the conditions

$$u(0) = 0, \quad u(x_1) = y_1.$$

Bernoulli circulated this problem to the mathematician of Europe as a challenge. Newton heard about it on the afternoon of 29 January 1697, on returning at 4pm, tired after his day's work as Master of the Royal Mint. He "solved" the problem, together with another geometrical one posed by Bernoulli, by 4am the next morning, and the same day communicated his solutions anonymously to teh Royal Society. Bernoulli "recognised the lion by his claws" but also said that neither problem should take "a man capable of (solving) it more than half-an-hour's careful thought."

10.2 The Navigation Problem (NOT EXAMINABLE)

Consider a "river" with parallel banks distance x_1 apart, and suppose that a boat is travelling with constant speed 1 relative to still water, see Fig. 10. Denote by $r = r(x) > 0$ the downstream

[21] Note that it does not depend on the mass of the particle!

[22] Since $2g$ is a constant, it suffices to minimise

$$\int_0^{x_1} \sqrt{\frac{u_x^2 + 1}{u}}. \tag{136}$$

current a distance x away from the bank passing through the origin O, and assume that $r(x) < 1$, $x \in [0, x_1]$. (Below we will see the reason for this assumption.)

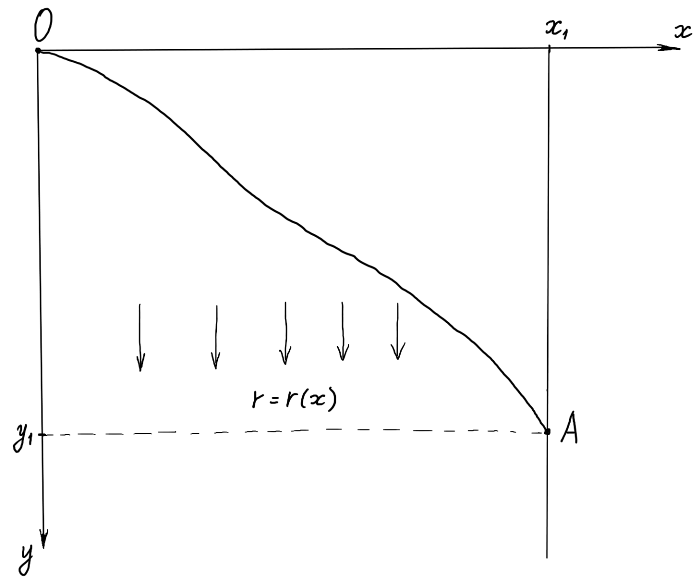

Figure 10: Navigation problem: mathematical setup.

Find the path joining the points O and $A = (x_1, y_1)$ corresponding to the minimum crossing time. The crossing time is given by (cf. the first equation in (135))

$$T(u) = \int_0^{x_1} \frac{dt}{dx} dx = \int_0^{x_1} \frac{1}{dx/dt} dx = \int_0^{x_1} \frac{1}{\cos(\theta(x))} dx, \tag{137}$$

where $\theta(x)$ is the steering angle, i.e. the angle between the rudder and (say) the x-axis, so the velocity vector of the boat relative to still water is[23]

$$\left(\frac{dx}{dt}, \frac{dy}{dt}\right) = \big(\cos\theta, \sin\theta\big) \tag{138}$$

In order to write the expression (137) in terms of the function $u = u(x)$, it remains to express the function θ in terms of the function u. To this end, notice that due to the observation (138) and the fact that the y-component of the speed of the boat is the sum of the y-component of the speed relative to still water and the speed of the river flow, i.e.

$$\frac{du}{dt} = \frac{dy}{dt} + r,$$

[23]Written in terms of the time variable, the equality (138) reads

$$\big(x'(t), y'(t)\big) = \big(\cos(\theta(x(t))), \sin(\theta(x(t)))\big), \qquad t \in \big[0, T(u)\big],$$

where $\{(x(t), y(t)), t \in [0, T(u)]\}$ is the "imagined" path in still water, i.e. the actual path in the river minus the displacement in the y direction due to the river flow.

we have

$$u_x = \frac{du}{dx} = \frac{du}{dt}\left(\frac{dx}{dt}\right)^{-1} = \frac{\sin\big(\theta(x)\big) + r(x)}{\cos\big(\theta(x)\big)},$$

and hence

$$(u_x \cos\theta - r)^2 = \sin^2\theta = 1 - \cos^2\theta,$$

or equivalently

$$(u_x^2 + 1)\cos^2\theta - 2u_x r \cos\theta + r^2 - 1 = 0.$$

Solving this for $\cos\theta$, we obtain

$$\cos\theta = \frac{u_x r \pm \sqrt{(u_x r)^2 + (u_x^2+1)(1-r^2)}}{u_x^2+1}. \tag{139}$$

Now we can see that if it is not the case that $r(x) < 1$, $x \in [0, x_1]$, then the problem may not have a solution (the river is too fast for one to be able to control the boat successfully). Continuing to assume that $r(x) < 1$, $x \in [0, x_1]$, we notice that in (139) we should take the $+$ sign, as taking $-$ would mean that $\cos\theta < 0$, equivalently $|\theta| > \pi/2$, i.e. the boat is steered in the direction of the departure river bank, which could not lead to a minimum time $T(u)$.

Plugging the expression (139) into (137) yields

$$\begin{aligned} T(u) &= \int_0^{x_1} \frac{u_x^2+1}{u_x r + \sqrt{(u_x r)^2 + (u_x^2+1)(1-r^2)}} \\ &= \int_0^{x_1} \frac{(u_x^2+1)(u_x r - \sqrt{(u_x r)^2 + (u_x^2+1)(1-r^2)})}{(u_x r)^2 - \big\{(u_x r)^2 + (u_x^2+1)(1-r^2)\big\}} \\ &= \int_0^{x_1} \frac{\sqrt{u_x^2 + 1 - r^2} - r u_x}{1-r^2}. \end{aligned}$$

In other words, the above problem of crossing the river corresponds to minimising

$$T(u) = \int_0^{x_1} \frac{\sqrt{u_x^2 + 1 - r^2} - r u_x}{1-r^2}$$

subject to the conditions

$$u(0) = 0, \quad u(x_1) = y_1.$$

Assuming that $0 < r(x) < 1$ for all $x \in [0, x_1]$, any point A on the opposite bank can be reached.

We can also consider the problem when the point A is not specified, that is we seek to cross the river in the shortest time without caring about the arrival point. This corresponds to minimising $T(u)$ subject to the single end condition $u(0) = 0$.

10.3 The Hanging Chain Problem (Not examinable)

Consider a perfectly flexible, inextensible, uniform chain or rope of length $l > 0$, suspended between two fixed points that are at the same height and distance $d \le l$ apart, see Fig. 11. What configuration of the chain minimises the total potential energy?

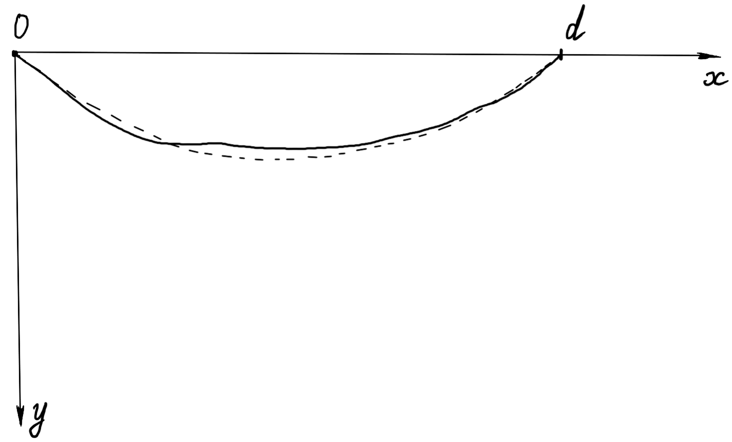

Figure 11: THE HANGING CHAIN PROBLEM: MATHEMATICAL SETUP.

To answer this question, we need to minimise

$$E(u) = -\int_0^d u\sqrt{u_x^2 + 1},$$

subject to the conditions $u(0) = u(d) = 0$ and the constraint

$$\int_0^d \sqrt{u_x^2 + 1} = l,$$

imposed by the fact that the rope is inextensible.

10.4 One-Dimensional Nonlinear Elasticity (NOT EXAMINABLE)

Consider an elastic rod made of a homogeneous material, occupying the interval $[0, l]$, see Fig. 12. Its elastic energy can be modelled by the integral

$$E(u) = \int_0^l f(u_x),$$

where F is the "stored-energy density function" of the material. We seek to minimise I subject to given end-point conditions, e.g.

$$u(0) = 0, \qquad u(l) = \lambda > 0.$$

We are interested in deformations u that are invertible (to avoid interpenetration of matter). A way to achieve this mathematically is to assume that $f(v) \to \infty$ as $v \to +0$, $f(v) = \infty$ for $v \leq 0$. then any u with $E(u) < \infty$ satisfies $u' > 0$ at every point $x \in [0, l]$.

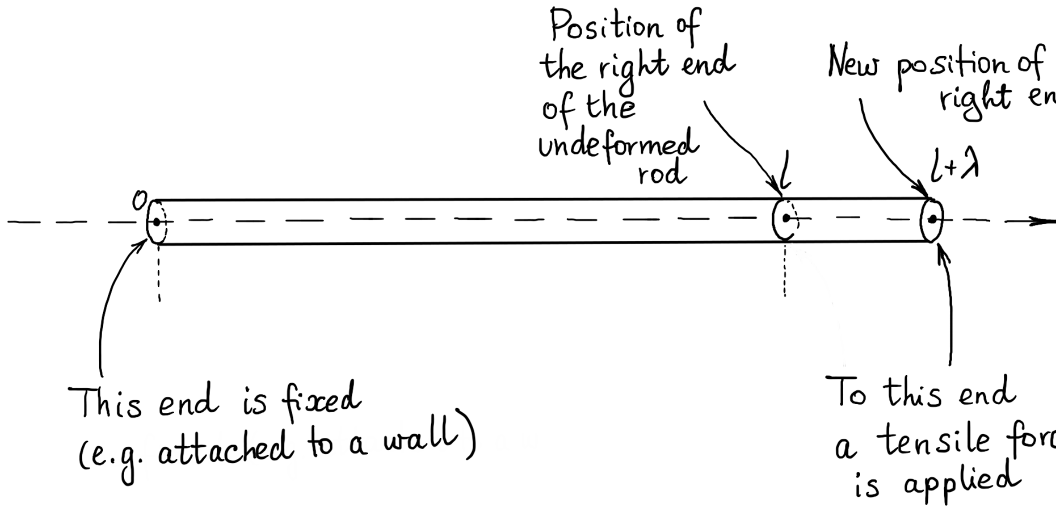

Figure 12: Nonlinearly elastic rod: mathematical setup.

Why do we minimise energy? Because of the Second Law of Thermodynamics, which implies that the corresponding dynamic equations have a *Lyapunov function.* For example, we could consider the dynamic viscoelastic evolution

$$u_{tt} = \big(f'(u_x)\big)_x + u_{xxt},$$

for which we formally have

$$\frac{d}{dt}\int_0^l \left(\frac{u_t^2}{2} + f(u_x)\right)dx = -\int_0^l u_{xt}^2.$$

This suggests a strong connection between calculus of variations and dynamical stability.

11 Global and Local Minimisers of Integral Functionals

Suppose that $f : [a,b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is continuous, $-\infty < a < b < \infty$. Consider the problem of minimising

$$I(u) = \int_a^b f\big(x, u(x), u_x(x)\big)\,dx = \int_a^b f(x, u, u_x), \tag{140}$$

subject to given boundary conditions

$$u(a) = \alpha, \quad u(b) = \beta, \qquad \alpha, \beta \in \mathbb{R}.$$

The function f is often referred to as the Lagrangian, or stored energy density (if the expression (140) has the physical meaning of energy).

11.1 Notation and Terminology

We will use the notation[24] $C^1_{\rm pw}([a,b])$ for the (linear) space of functions u that are continuous on $[a,b]$ and such that for some a_j, $j=1,\dots,N$, such that $a=a_0<a_1<a_2<\cdots<a_N=b$, one has $u\in C^1([a_j,a_{j+1}])$, $j=0,\dots N-1$. In other words, the interval $[a,b]$ can be split into a finite set of intervals such that on each of them the function u is continuously differentiable. In particular, the graph of a function $u\in C^1_{\rm pw}([a,b])$ may have a finite set[25] of "corners", at the points $a_1,\dots,a_{N-1}$, see Fig. 13.

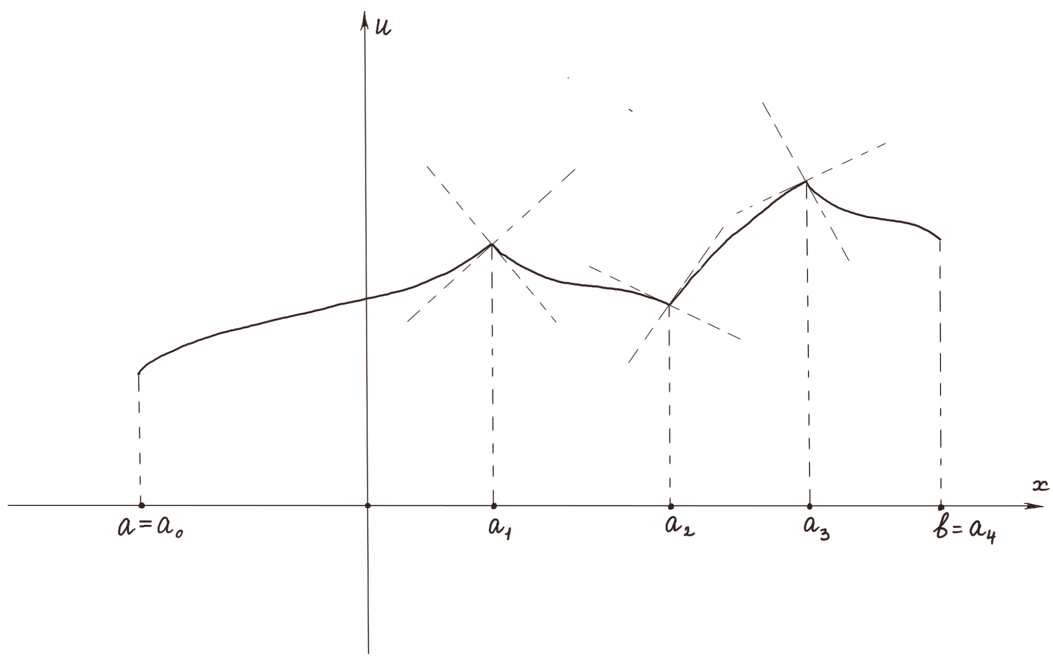

Figure 13: THE GRAPH OF A TYPICAL FUNCTION IN THE SPACE $C^1_{\rm pw}([a,b])$. The crossing dashed lines at corner points are the tangent lines to each side of the corner, thus representing the (unequal) one-sided derivatives of the function.

We equip the space $C^1_{\rm pw}([a,b])$ with the norm[26]

$$\|u\|_1 := \max_{x\in[a,b]}\big|u(x)\big| + \max_{j\in\{0,\dots,N-1\}}\ \max_{x\in[a_j,a_{j+1}]}\big|u_x(x)\big|, \qquad u\in C^1_{\rm pw}([a,b]). \tag{141}$$

[24]The subscript "pw" stands for "piecewise", so functions in $C^1_{\rm pw}([a,b])$ are "piecewise C^1" in the interval $[a,b]$.

[25]If the tgraph of u does not have any corners, then the derivative of u is continuous at all points of $[a,b]$ and so $u\in C^1[a,b]$.

[26]Recall that a norm on a linear space X over $\mathbb{R}$ is a scalar function $\|\cdot\|$ on X with the following properties: (a) $\|\lambda u\|=|\lambda|\|u\|$ for all $\lambda\in\mathbb{R}$, $u\in X$; (b) $\|u\|>0$ for $u\neq 0$; (c) $\|u+v\|\le\|u\|+\|v\|$ for all $u,v\in X$.

Consider the set of *admissible functions*

$$\mathcal{A} := \left\{u \in C^1_{\text{pw}}([a,b]) : u(a) = \alpha, u(b) = \beta\right\}$$

and the mapping $I : \mathcal{A} \to \mathbb{R}$ defined by

$$I(u) = \sum_{j=0}^{N-1} \int_{a_j}^{a_{j+1}} f\big(x, u(x), u_x(x)\big)dx, \tag{142}$$

where a_j, $j = 0, 1, \ldots, N$, are the points of a partition of the interval $[a,b]$ such that $u \in C^1([a_j, a_{j+1}])$, $j = 0, \ldots N-1$.

In what follows, we simply write (cf. (140))

$$I(u) = \int_a^b f\big(x, u(x), u_x(x)\big)dx = \int_a^b f(x, u, u_x), \tag{143}$$

implying the sum (142) for a given element $u \in \mathcal{A}$. We will also refer to a mapping[27] from $C^1_{\text{pw}}([a,b])$ to $\mathbb{R}$, in particular (143), as a **functional**.

11.2 Global Minimisers and Motivation for New Analysis

Definition 11.1. *A function $u \in \mathcal{A}$ is said to be a* **global minimiser** *of the functional (143) on the set $\mathcal{A}$, if $I(w) \geq I(u)$ for all $w \in \mathcal{A}$, and it is said to be a* **strict global minimiser** *if $I(w) > I(u)$ for all $w \in \mathcal{A}$, $w \neq u$.*

Remark 11.2. *Even if f attains its infimum, i.e. there exists $(x_{\min}, u_{\min}, v_{\min}) \in [a,b] \times \mathbb{R} \times \mathbb{R}$ such that*

$$f(x_{\min}, u_{\min}, v_{\min}) = \inf_{[a,b]\times\mathbb{R}\times\mathbb{R}} f =: f_{\min} > -\infty, \tag{144}$$

there need be no global minimiser of I on $\mathcal{A}$, even if there are function sequences $\{u^{(j)}\}$ satisfying[28]

$$I(u^{(j)}) \to \inf_{\mathcal{A}} I$$

that are bounded, and even convergent, with respect to the norm (141) — see examples below. This cannot happen for continuous functions on a finite-dimensional space, so it is a special feature of the function space setting. This observation motivates a separate subject, the Calculus of Variations, which explores the conditions under which we may or may not claim the existence minimisers (global, local – see below, etc) on an infinite-dimensional space. The examples of Section 10 show that such conditions are also practically relevant.

[27] In general, a functional are a mapping from a function space to a number field ($\mathbb{R}$, $\mathbb{C}$).

[28] Note that under (144) the functional I is bounded below:

$$\inf_{\mathcal{A}} I \geq \int_a^b f_{\min} = (b-a) f_{\min}.$$

Example 11.3 (O. Bolza)**.** *Consider the integral*

$$I(u) = \int_0^1 \big(u^2 + (u_x^2 - 1)^2\big) \tag{145}$$

on the set

$$\mathcal{A} = \big\{u \in C^1_{\mathrm{pw}}([0,1]) : u(0) = u(1) = 0\big\}. \tag{146}$$

It turns out that

$$\inf_{\mathcal{A}} I = 0. \tag{147}$$

Indeed, consider the function sequence $u^{(j)}$ *whose typical element is shown on Fig. 14.*

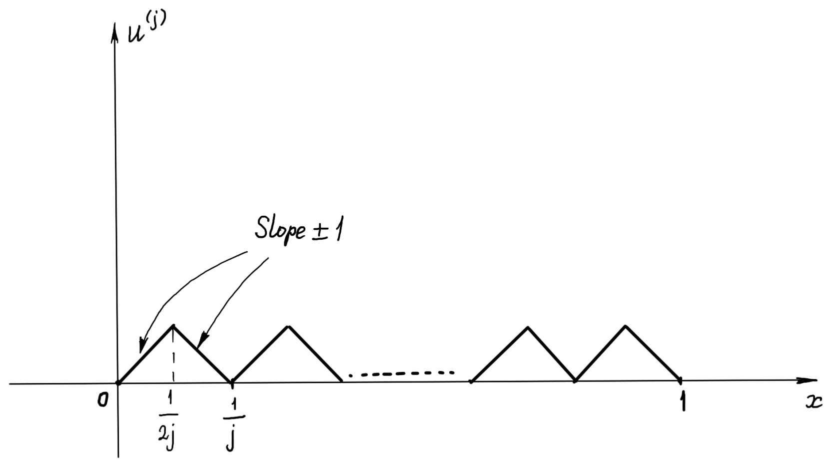

Figure 14: A SEQUENCE $\{u^{(j)}\}$ SHOWING THAT THE INFIMUM OF (145) OVER THE SET (146) EQUALS ZERO, SEE (147).

One has

$$I\big(u^{(j)}\big) = \int_0^1 \big(u^{(j)}\big)^2 \int_0^1 \left(\frac{1}{2j}\right)^2 \leq \frac{1}{4j^2} \to 0 \quad \text{as} \quad j \to \infty,$$

but there is no $u \in \mathcal{A}$ *with* $I(u) = 0$*, since* $I(u) = 0$ *implies*

$$\big(u(x)\big)^2 + \big(\big(u_x(x)\big)^2 - 1\big)^2 = 0 \qquad \forall x \in [0,1],$$

and hence at all points x *one has* $u(x) = 0$ *and* $|u_x(x)| = 1$ *at the same time, which is impossible.*

Example 11.4 (K. Weierstrass)**.** *Consider the functional*

$$I(u) = \int_0^1 x^2 u_x^2 dx,$$

subject to the boundary conditions

$$u(0) = 0, \qquad u(1) = 1. \tag{148}$$

We claim that

$$\inf_{\mathcal{A}} I = 0,$$

where

$$\mathcal{A} = \big\{ u \in C^1_{\mathrm{pw}}([0,1]) : u(0) = 0, u(1) = 1 \big\}.$$

Indeed, consider the sequence $\{u^{(j)}\}$ shown on Fig. 15. By a direct calculation, one has

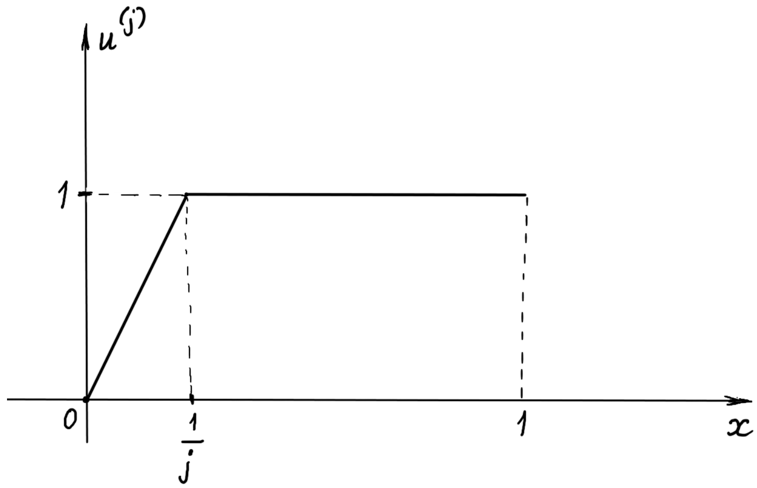

Figure 15: THE INFIMUM OF THE FUNCTIONAL IN EXAMPLE 11.4 IS NOT ATTAINED. The sequence shown is minimising: the limit of the values of the functional I on the elements of this sequence is zero. However, there is no function $u \in \mathcal{A}$ such that $I(u) = 0$.

$$I\big(u^{(j)}\big) = \int_0^{1/j} x^2 j^2 dx = \frac{1}{3j} \to 0 \qquad \text{as} \quad j \to \infty.$$

However, $I(u) = 0$ implies that $u_x = 0$, and hence u is constant, which is impossible in view of the boundary conditions (148).

Example 11.5. *Consider the functional*

$$I(u) = \int_0^1 \sqrt{u^2 + u_x^2}\, dx,$$

subject to the boundary conditions

$$u(0) = 0, \qquad u(1) = 1. \tag{149}$$

Here the set $\mathcal{A}$ is the same as in Example 11.4. Consider the sequence $\{u^{(j)}\} \subset \mathcal{A}$ shown in Fig. 16. One has

$$I\big(u^{(j)}\big) = \int_{1-1/j}^1 \sqrt{\left(x - 1 + \frac{1}{j}\right)^2 j^2 + j^2}\, dx = \int_0^1 \sqrt{y^2 + j^2}\, \frac{dy}{j} = \int_0^1 \sqrt{\frac{y^2}{j^2} + 1}\, dy \to 1 \qquad \text{as } j \to \infty.$$

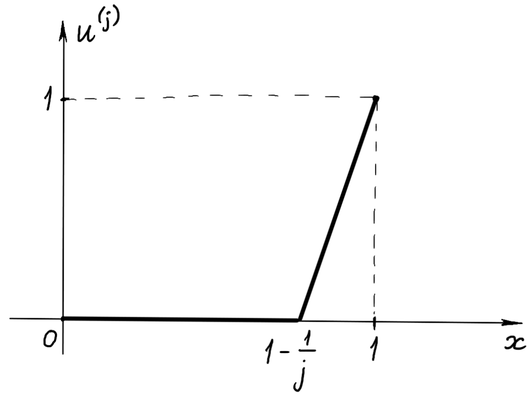

Figure 16: A MINIMISING SEQUENCE FOR THE FUNCTIONAL IN EXAMPLE 11.5.

However, for all $u \in \mathcal{A}$, the following estimate holds:

$$\begin{aligned} I(u) &= \int_0^1 \left(\sqrt{u^2+u_x^2} - |u_x| + |u_x|\right) dx \\ &\geq \int_0^1 \left(\sqrt{u^2+u_x^2} - |u_x|\right) dx + \int_0^1 u_x dx = \int_0^1 \left(\sqrt{u^2+u_x^2} - |u_x|\right) dx + 1, \end{aligned}$$

which implies that if $I(u) = 1$ then $\sqrt{u^2+u_x^2} = |u_x|$ and hence $u = 0$, which is impossible, due to the boundary conditions (149).

12 Necessary and Sufficient Conditions for a Minimum

In what follows $\|\cdot\|_1$ is the norm on $C^1_{\mathrm{pw}}([a,b])$ defined by (141).

Definition 12.1. *A function $u \in \mathcal{A}$ is said to be a* **local minimiser** *of I in $\mathcal{A}$ if there exists $\varepsilon > 0$ such that $I(w) \geq I(u)$ whenever $w \in \mathcal{A}$ with $\|w - u\|_1 \leq \varepsilon$.*

Example 12.2. *Consider the functional*

$$I(u) = \int_0^1 f(u_x), \qquad f(v) = (v^2-1)^2\left(v^2 + \frac{1}{8}\right), \tag{150}$$

subject to the end-point conditions

$$u(0) = u(1) = 0.$$

Then $u(x) = 0$, $x \in [0,1]$, is a local minimiser. Indeed if $\|w - u\|_1$ is small, then by Corollary 14.6 we can write

$$f\big(w_x(x)\big) = f\big(u_x(x)\big) + f'\big(u_x(x)\big)\big(w_x(x) - u_x(x)\big) + \frac{f''\big(\theta(x)\big)}{2}\big(w_x(x) - u_x(x)\big)^2, \tag{151}$$

where $\theta(x)$ is situated between $u_x(x)$ and $w_x(x)$. Since for the function f in (150) one has $f'\big(u_x(x)\big) = f'(0) = 0$, see Fig. 17, we have

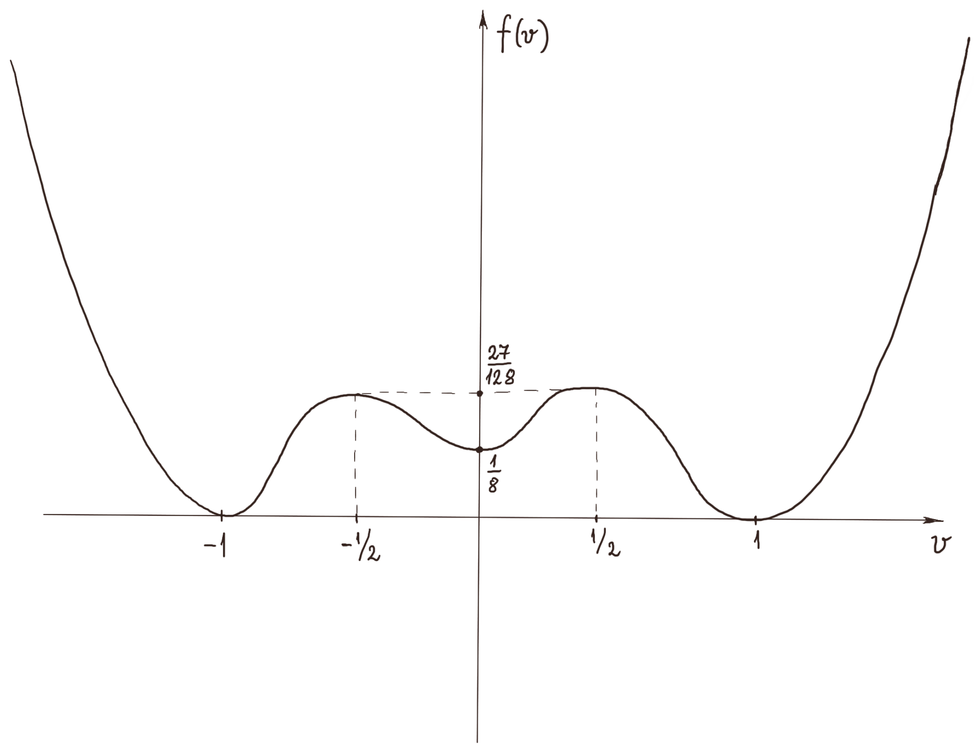

Figure 17: The function f in (150). The Lagrangian f in Example 12.2 depends on the values of the derivative u_x only (represented by the argument v of f). Notice that f is a polynomial of degree 6, with two symmetric "wells" at the points $v = \pm 1$ and one additional "raised" well at $v = 0$.

$$f\big(w_x(x)\big)dx = f\big(u_x(x)\big) + \frac{f''\big(\theta(x)\big)}{2}\big(w_x(x) - u_x(x)\big)^2 \geq f\big(u_x(x)\big), \quad x \in [0, 1], \tag{152}$$

and hence

$$I(w) = \int_0^1 f\big(w_x(x)\big)dx \geq \int_0^1 f\big(u_x(x)\big)dx = I(u),$$

as required for u to be a local minimum.

Remark 12.3. *The set of functions w satisfying $\|w - u\|_1 \leq \varepsilon$ is referred to as the ε-neighbourhood of u. So Definition 12.1 can be rephrased to say that u is a local minimiser if the values of I in some ε-neighbourhood of u are not below the value $I(u)$. It is important that neighbourhoods are considered in the set $\mathcal{A}$ for which the minimisation problem is posed.*

12.1 Euler-Lagrange Equation: Necessary Condition for a Local Minimum

Definition 12.4. *Suppose that f_u, f_v are all continuous functions of x, u, v. For a given $u \in \mathcal{A}$, the functional $\delta I(u)$ that maps each $\phi \in C_0^1([a, b])$ to the value*

$$\delta I(u)[\phi] := \int_a^b \big(f_u(x, u, u_x)\phi + f_v(x, u, u_x)\phi_x\big) \tag{153}$$

is referred to as the **first variation***, or just the* **variation***, of the functional I at u.*

Remark 12.5. *1. In the definition above we assumed some regularity of the function f in (143) (f_u, f_v are continuous functions). As it happens, for many practically relevant problems, such as the brachistochrone problem of Section 10.1, this is not the case – notice that the integrand in (135), (136)) and its derivative f_u vanish at $u = 0$ (and actually, all functions u we consider vanish at $x = 0$, so the condition required by Definition 12.4 definitely does not hold!). So we will actually extend the above definition to a more general class of functionals I, by referring to the expression on the right-hand side of (153) as the first variation of I even if f_u, f_v have points of discontinuity, as long as this expression makes sense for functions $u \in C^1_{\mathrm{pw}}([a,b])$, $\phi \in C^1_0([a,b])$ (for example, as an improper integral). For the purposes of this unit, it will suffice to assume that this is always the case, i.e.*

$$\int_a^b \big(f_u(x,u,u_x)\phi + f_v(x,u,u_x)\phi_x\big)$$

is well defined for $u \in C^1_{\mathrm{pw}}([a,b])$, $\phi \in C^1_0([a,b])$.

2. By the same token, while the theorem below (Theorem 12.6) is formulated only for the case of regular integrands f, its conclusion has wider applicability and we will use it to find minima in problems for which f_u, f_v are discontinuous, in particular in the brachistochrone problem.

Theorem 12.6 (Necessary condition for a local minimiser)**.** *Consider $f = f(x,u,v)$ and suppose that $f_u, f_v \in C^1([a,b]\times\mathbb{R}\times\mathbb{R})$. Suppose that $u \in \mathbb{R}$ is a local minimiser of I in $\mathcal{A}$. Then the first variation of I at u vanishes:*

$$\delta I(u)[\phi] = 0 \qquad \forall \phi \in C^1_0([a,b]), \tag{154}$$

where

$$f_u = f_u\big(x, u(x), u_x(x)\big), \quad f_v = f_v\big(x, u(x), u_x(x)\big).$$

We will refer to (154) as the **Euler-Lagrange identity.**

Proof. Suppose that $\phi \in C^1_0([a,b])$. Then $u + t\phi \in \mathcal{A}$, $t \in \mathbb{R}$, and so $I(u+t\phi)$ has a local minimum at $t = 0$. For $t \neq 0$, we have, by the derivative mean-value theorem (Theorem 14.2),

$$\begin{aligned}\frac{I(u+t\phi)-I(u)}{t} &= \int_a^b \frac{f(x,u+t\phi,u_x+t\phi_x)-f(x,u,u_x)}{t}dx \\ &= \int_a^b \big\{f_u(x,u+\tau t\phi,u_x+\tau t\phi_x)\phi + f_v(x,u+\tau t\phi,u_x+\tau t\phi_x)\phi_x\big\}dx = 0,\end{aligned} \tag{155}$$

where $\tau(x,t) \in [0,1]$.

Indeed, for each $x \in [a,b]$ consider the function

$$\omega(s) := f\big(x, u(x)+s\phi(x), u_x + s\phi_x(x)\big), \qquad s \in [0,t].$$

Due to the assumption that the partial derivatives f_u, f_v are continuous, the chain rule implies that ω is continuously differentiable on $[0,t]$ and

$$\omega'(s) = f_u\big(x,u(x)+s\phi(x),u_x+s\phi_x(x)\big)\phi(x) + f\big(x,u(x)+s\phi(x),u_x+s\phi_x(x)\big)\phi_x(x), \qquad s \in [0,t].$$

Applying Theorem 14.2 to ω on the interval $[0,t]$, we obtain the existence of $\xi = \xi(x,t) \in [0,t]$ such that

$$\frac{f(x,u+t\phi,u_x+t\phi_x)-f(x,u,u_x)}{t} = \frac{\omega(t)-\omega(0)}{t} = \omega'(\xi) = f\big(x,u+\xi\phi,u_x+\xi\phi_x\big), \qquad x \in [a,b].$$

Finally, denoting $\tau(x,t) = t^{-1}\xi(x,t)$, we obtain (155).

Passing to the limit in (155) as $t \to 0$ and using the bounded convergence theorem, we deduce that

$$\lim_{t\to 0} \frac{I(u+t\phi) - I(u)}{t} = \int_a^b (f_u\phi + f_v\phi_x)dx,$$

and the result follows. $\square$

Theorem 12.7 ("Fundamental Theorem of Calculus of Variations")**.** *Suppose that $h, g \in C([a,b])$. If the following integral identity holds:*

$$\int_a^b \big(h\phi + g\phi_x\big) = 0 \qquad \forall \phi \in C_0^1([a,b]),$$

then g is differentiable and $g_x = h$. Hence, in particular, one has $g \in C^1([a,b])$.

Proof. Consider the function

$$H(x) := \int_a^x h + c, \qquad x \in [a,b],$$

where the constant c is to be chosen later. Then $H_x = h$, and

$$\int_a^b h\phi = \int_a^b H_x\phi = [H\phi]_a^b - \int_a^b H\phi_x = -\int_a^b H\phi_x,$$

since $\phi(a) = \phi(b) = 0$. Hence,

$$\int_a^b (h\phi + g\phi_x) = \int_a^b (-H+g)\phi_x = 0 \qquad \forall \phi \in C_0^1\big([a,b]\big) \tag{156}$$

Now take

$$\phi(x) = \int_a^x (-H+g), \qquad x \in [a,b].$$

Then $\phi(a) = 0$ and

$$\phi(b) = \int_a^b \Big(-\int_a^x h - c + g(x) \Big) dx = 0$$

for a suitable c. Under this choice of c, we have $\phi \in C_0^1([a,b])$. Substituting this ϕ into the identity (156) yields

$$\int_a^b (-H+g)^2 = 0,$$

hence $-H + g = 0$, and therefore

$$g(x) = \int_a^x h + c, \qquad x \in [a,b],$$

as claimed. $\square$

Corollary 12.8. *Under the conditions of Theorem 12.6, if $u \in \mathcal{A}$ is a local minimiser of I in $\mathcal{A}$, then the derivative*

$$\frac{d}{dx} f_v\big(x, u(x), u_x(x)\big)$$

exists for $x \in [a, b]$, and

$$f_u\big(x, u(x), u_x(x)\big) - \frac{d}{dx} f_v\big(x, u(x), u_x(x)\big) = 0, \qquad x \in [a, b]. \tag{157}$$

The equation (157) is referred to as the **Euler-Lagrange equation.** *Integrating (157), we obtain its equivalent "integrated" form*

$$f_v = c + \int_a^x f_u, \tag{158}$$

where f_v and f_u stand for the functions $f_v(x, u(x), u_x(x))$, $x \in [a, b]$, and $f_u(x, u(x), u_x(x))$, $x \in [a, b]$, respectively, and the constant c is automatically the value of f_v at $x = a$.

If the the function u has corner points $a_1, \dots, a_{N-1}$, see the discussion at the beginning of Section 11.1, then the equation (157) holds on all intervals $[a_j, a_{j+1}]$, where $j = 0, 1, \dots, N-1$, $a_0 = a$, $a_N = b$.

Remark 12.9. *Note that while a solution to (157) may have corner points, they must comply with the condition that $f_v\big(x, u(x), u_x(x)\big)$ is differentiable. In particular, $f_v\big(x, u(x), u_x(x)\big)$ must be continuous. Consider, for example, the case $f(v) = v^2/2$. Then $f_v(u_x(x) = u_x(x)$, which means that u_x is continuous and disallows any corner points!*

Remark 12.10. *In line with Remark 12.5, we will us the Euler-Lagrange equation (157) to find minimiser for problems that satisfies weaker versions of the conditions of Theorem 12.7, e.g. when the functions h, g have discontinuities (such as in applying it to the brachistochrone problem). In order to make this rigorous, we would need to work in wider function spaces, such as the space of Lebesgue integrable functions $L^1(a, b)$, which are covered in courses/units on real analysis.*

Example 12.11. *Consider the functional discussed in Example 12.2, i.e. the functional*

$$I(u) = \int_0^1 \big(u_x^2 - 1\big)^2 \left(u_x^2 + \frac{1}{8}\right)$$

on the set

$$\mathcal{A} = \big\{u \in C^1_{\mathrm{pw}}([0, 1]) : u(0) = u(1) = 0\big\}.$$

Since the Lagrangian $f(x, u, v)$ is a function of v only:

$$f(v) = (v^2 - 1)^2 \left(v^2 + \frac{1}{8}\right),$$

the Euler-Lagrange equation has the form

$$\frac{d}{dx} f_v\big(u_x(x)\big) = 0, \tag{159}$$

and hence

$$f_v(u_x(x)) = c, \qquad x \in [0, 1], \tag{160}$$

for some constant c. Noticing that

$$f_v = f'(v) = 6v(v^2-1)\left(v^2 - \frac{1}{4}\right),$$

we can write (160) explicitly:

$$u_x(x)\Big(\big(u_x(x)\big)^2 - 1\Big)\left(\big(u_x(x)\big)^2 - \frac{1}{4}\right) = \widetilde{c}, \tag{161}$$

where $\widetilde{c} = c/6$. *Therefore, any function* $u \in C^1_{\mathrm{pw}}([0,1])$ *satisfying the conditions* $u(0) = u(1) = 0$, *whose derivative* u_x *is a piecewise constant function taking values in the set of solutions to the algebraic equation*

$$v(v^2-1)(v^2-1/4) = \widetilde{c},$$

for some fixed constant $\widetilde{c}$, *solves the Euler-Lagrange equation (159) and is hence a candidate for a local minimum.*

Which of these functions actually are local minimisers needs further investigation. For some of them the sufficient condition (Theorem 12.22) discussed in Section 12.5 may hold. However, if we require that the local minimiser be continuously differentiable, i.e. $u \in C^1([0,1])$, *then we can proceed with the analysis of (161) one step further. Indeed, by the derivative mean-value theorem (Theorem 14.2), if* $u \in C^1([0,1])$ *then there exists a point* $\xi \in (0,1)$ *such that*

$$u_x(\xi) = \frac{u(1) - u(0)}{1 - 0} = 0,$$

since $u(1) = u(0) = 0$. *It follows that the only possible value for* $\widetilde{c}$ *in (161) is zero, as*

$$u_x(\xi)\Big(\big(u_x(\xi)\big)^2 - 1\Big)\left((u_x(\xi))^2 - \frac{1}{4}\right) = 0.$$

Now the equation (161) takes the form

$$u_x(x)\Big(\big(u_x(x)\big)^2 - 1\Big)\left((u_x(x))^2 - \frac{1}{4}\right) = 0,$$

from which we obtain five options

$$u_x(x) = 0, \qquad u_x(x) = \pm 1, \qquad u_x(x) = \pm\frac{1}{2}, \tag{162}$$

of which one and only one must hold for all $x \in [0,1]$. *The only such function in* $C^1([0,1])$ *that satisfies the boundary conditions* $u(0) = u(1) = 0$ *is* $u(x) = 0$, $x \in [0,1]$.

12.2 Natural Boundary Conditions

Suppose that as before $-\infty < a < b < \infty$ and we replace the set $\mathcal{A}$ of admissible functions by

$$\widetilde{\mathcal{A}} = \big\{u \in C^1_{\mathrm{pw}}([a,b]) : u(a) = \alpha\big\}, \qquad \alpha \in \mathbb{R},$$

so that the end-point $x = b$ is free (i.e. no boundary condition is specified there).

Theorem 12.12. *Suppose the partial derivatives f_u, f_v exist and are continuous on $[a,b]\times\mathbb{R}\times\mathbb{R}$. If $u \in \widetilde{\mathcal{A}}$ be a local minimiser of I in $\widetilde{\mathcal{A}}$, then*

$$\int_a^b (f_u\phi + f_v\phi_x)dx = 0 \qquad \forall \phi \in C^1([a,b]) \ \text{ with } \phi(a) = 0. \tag{163}$$

The identity (163) holds if and only if[29]

$$f_v\big(x, u(x), u_x(x)\big) = -\int_x^b f_u\big(s, u(s), u_x(s)\big)ds \qquad \forall x \in [a,b]. \tag{164}$$

Proof. The identity (163) is obtained in the same way as (154) is obtained in the proof of Theorem 12.6, by considering $I(u + t\phi)$ as a function of $t \in \mathbb{R}$.

For the second part of the statement, notice that in the identity (163) $\phi(a) = 0$ and ϕ takes arbitrary values at b. In particular, since we can choose $\phi \in C_0^1([a,b])$, from (158) we have

$$\begin{aligned} f_v\big(x, u(x), u_x(x)\big) &= c + \int_a^x f_u\big(s, u(s), u_x(s)\big)ds \\ &= -c_1 - \int_x^b f_u\big(s, u(s), u_x(s)\big)ds \qquad \forall x \in [a,b], \end{aligned} \tag{165}$$

for some constants c, c_1. Let us show that the constant c_1 vanishes, so (164) holds.

Lemma 12.13. *Suppose that $z \in C^1_{\mathrm{pw}}([a,b])$ with $z(b) = 0$. Then*

$$\int_a^b (\phi_x z + \phi z_x)dx = 0 \qquad \forall \phi \in C^1([a,b]) \ \text{ with } \phi(a) = 0. \tag{166}$$

Proof. We can proceed either via integration by parts[30] on each sub-interval of $[a,b]$ where z is C^1, or "by approximation". For the latter, consider a sequence $z^{(j)} \in C^1([a,b])$ such that $z^{(j)} \to z$ in $C^1_{\mathrm{pw}}([a,b])$. Then

$$\int_a^b \big(\phi_x z^{(j)} + \phi z_x^{(j)}\big)dx = \int_a^b \big(\phi z^{(j)}\big)_x = \phi(b)z^{(j)}(b).$$

Passing to the limit as $j \to \infty$, we obtain (166). □

Setting[31]

$$z(x) = -\int_x^b f_u, \qquad x \in [a,b],$$

[29] At this point we can immediately see that an additional condition at $x = b$ must be satisfied. What is it? You may wish to take a peek at Corollary 12.14 to verify your answer.

[30] The integration by parts argument can be implemented via the following observation:

$$0 = \int_a^b (f_u\phi + f_v\phi_x)dx = \int_a^b \left(f_u - \frac{d}{dx}f_v\right)\phi + [f_v\phi]_a^b = c_1\phi(b),$$

where the integrals are understood as the sums of the corresponding integrals over intervals where f_u is continuous (and so f_v is continuously differentiable), similar to (142).

[31] Here, for brevity, we omit the arguments of f_u, f_v.

in (166) and using (165), we obtain

$$0 = \int_a^b \left(\phi_x\left(-\int_x^b f_u\right) + \phi f_u\right)dx = \int_a^b \big((f_v + c_1)\phi + f_u\phi\big)dx,$$

and hence

$$0 = \int_a^b (f_u\phi + f_v\phi_x)dx = c_1\int_a^b \phi_x = c_1\phi(b),$$

from which the result follows by choosing any ϕ with $\phi(b) = 0$. $\square$

Corollary 12.14. *Under the conditions of Theorem 12.12, one has (in addition to the Euler-Lagrange equation (157))*

$$f_v\big(b, u(b), u_x(b)\big) = 0. \tag{167}$$

Similarly, if the end-point $x = a$ is free, the condition $u(a) = \alpha$ is replaced by

$$f_v\big(a, u(a), u_x(a)\big) = 0.$$

Remark 12.15. *Notice that for the specific case when f is quadratic, $f(v) = v^2/2$, one has $f_v = v$, and so the condition (167) takes the form $u_x(b) = 0$, i.e. the Neumann condition discussed in the context of Sturm-Liouville problems in MA30044 and also in some of the problem sheets for this unit.*

Example 12.16. *Let us continue discussing the functional of Examples 12.2, 12.11, i.e.*

$$I(u) = \int_0^1 f(u_x), \qquad f(v) = (v^2 - 1)^2\left(v^2 + \frac{1}{8}\right),$$

but now consider it on the set

$$\widetilde{\mathcal{A}} = \big\{u \in C^1_{\mathrm{pw}}([0,1]) : u(0) = 0\big\}.$$

Due to Corollary 12.14, any local minimiser of I on $\widetilde{\mathcal{A}}$ must satisfy (161) together with the condition (167), which takes the form

$$u_x(1)\Big(\big(u_x(1)\big)^2 - 1\Big)\left(\big(u_x(1)\big)^2 - \frac{1}{4}\right) = 0,$$

i.e. one of the following conditions must hold:

$$u_x(1) = 0, \qquad u_x(1) = \pm 1, \qquad u_x(1) = \pm\frac{1}{2}, \tag{168}$$

In addition, the constant $\widetilde{c}$ in (161) vanishes, and hence as each point $x \in [0,1]$ we have one of the following five options (cf. (162))

$$u_x(x) = 0, \qquad u_x(x) = \pm 1, \qquad u_x(x) = \pm\frac{1}{2}.$$

Arguing as in Example 12.11, we infer that any function $u \in C^1_{\mathrm{pw}}([0,1])$ satisfying the condition $u(0) = 0$, whose derivative u_x is a piecewise constant function taking values in the set

$\{0, -1, 1, -1/2, 1/2\}$ is a candidate for a local minimum. As in the case of Example 12.11, the question about which of these functions actually are local minimisers of I in $\widetilde{\mathcal{A}}$ requires further investigation. However, if we impose that the local minimiser is continuously differentiable, i.e. $u \in C^1([0,1])$, then the choice is reduced to five functions:

$$u(x) = 0, \quad x \in [0,1], \tag{169}$$

$$u(x) = x, \quad x \in [0,1], \qquad\qquad u(x) = -x, \quad x \in [0,1], \tag{170}$$

$$u(x) = \frac{x}{2}, \quad x \in [0,1], \qquad\qquad u(x) = -\frac{x}{2}, \quad x \in [0,1]. \tag{171}$$

Using an argument similar to that of Example 12.2, it can be seen that the functions (169), (170) are local minimisers of I in $\widetilde{\mathcal{A}}$, while the functions (171) are not.

Exercise 12.17. *Explain why the functions (170) are global minimisers of I on $\widetilde{\mathcal{A}}$, while the function (169) is not.*

12.3 Smoothness of local minima

Note that in general we cannot claim that

$$\begin{aligned}\frac{d}{dx} f_v\big(x, u(x), u_x(x)\big) = f_{vx}\big(x, u(x), u_x(x)\big)& \\ + f_{vu}\big(x, u(x), u_x(x)\big)u_x(x) + f_{vv}\big(x, u(x), u_x(x)\big)u_{xx},&\end{aligned} \tag{172}$$

as the chain rule would suggest. (We do not even know if u_{xx} exists!) However, sometimes one can claim that u is actually smoother than an arbitrary C^1_{pw} function, by performing a straightforward check on the Lagrangian f.

Theorem 12.18. *Suppose that $f = f(x, u, v)$ is twice continuously differentiable on $[a, b] \times \mathbb{R} \times \mathbb{R}$, and in addition $f_{vv}(x, u, v) > 0$ for all x, u, v. If $u \in C^1_{\mathrm{pw}}([a, b])$ solves the Euler-Lagrange equation (157), then $u \in C^2([a, b])$ and u_{xx} is expressed via the formula (172), i.e.*

$$u_{xx} = g(x, u, u_x), \qquad\qquad g(x, u, v) := \frac{f_u(x, u, v) - f_{xv}(x, u, v) - f_{uv}(x, u, v)v}{f_{vv}(x, u, v)}. \tag{173}$$

(Note how we replaced the left-hand side of (172) by f_u, in view of the Euler-Lagrange equation (157).)

Proof. Omitted; not examinable. □

Corollary 12.19. *Suppose $f = f(x, u, v)$ is infinitely differentiable on $[a, b] \times \mathbb{R} \times \mathbb{R}$, with $f_{vv} > 0$. If $u \in C^1_{\mathrm{pw}}([a, b])$ solves the Euler-Lagrange equation (157), then $u \in C^\infty([a, b])$.*

Proof. By induction, based on the equation (173). □

12.4 Special Cases for Integrating the Euler-Lagrange Equation

When the function $f = f(x, u, v)$ in (143) does not depend on either x or u (or both), then the Euler-Lagrange equation (157) can usually be reduced to a first-order equation.

1. If $f = f(x, v)$, then the Euler-Lagrange equation takes the form

$$\frac{d}{dx} f_v\big(x, u(x), u_x(x)\big) = 0, \qquad x \in [a, b],$$

and hence

$$f_v(x, u_x(x)) = c, \qquad x \in [a, b],$$

for some constant c. This kind of situation occurred in Example 12.11.

2. If $u \in C^2([a, b])$ is a solution of (157), then, using the chain rule, we obtain

$$\begin{aligned}
&\frac{d}{dx}\big\{ f\big(x, u(x), u_x(x)\big) - u_x(x) f_v\big(x, u(x), u_x(x)\big)\big\} \\
&\qquad = f_x\big(x, u(x), u_x(x)\big) + f_u\big(x, u(x), u_x(x)\big) u_x(x) + f_v\big(x, u(x), u_x(x)\big) u_{xx}(x) \\
&\qquad\qquad - u_{xx}(x) f_v\big(x, u(x), u_x(x)\big) - u_x(x) \frac{d}{dx} f_v\big(x, u(x), u_x(x)\big) \\
&\qquad = f_x\big(x, u(x), u_x(x)\big) + \left(f_u\big(x, u(x), u_x(x)\big) - \frac{d}{dx} f_v\big(x, u(x), u_x(x)\big) \right) u_x(x) \\
&\qquad = f_x\big(x, u(x), u_x(x)\big),
\end{aligned}$$

where for the last equality we used the Euler-Lagrange equation (157). In particular, if $f = f(u, v)$ (such as the case for the Brachistochrone problem, for example, for which $f(u, v) = \sqrt{(1 + v^2)/u}$), then $f_x = 0$, and so

$$f\big(u(x), u_x(x)\big) - u_x(x) f_v\big(u(x), u_x(x)\big) = c$$

for some constant c.

For example, this property can be used when the conditions of Theorem 12.18 are satisfed, as they guarantee that $u \in C^2([a, b])$.

Remark 12.20. *The identity*

$$\frac{d}{dx}\big\{ f\big(x, u(x), u_x(x)\big) - u_x(x) f_v\big(x, u(x), u_x(x)\big)\big\} = f_x\big(x, u(x), u_x(x)\big) \qquad \forall u \in C^2([a, b])$$

is often referred to as the Beltrami identity.

12.5 Sufficient Condition for a Local Minimum

Definition 12.21. *Suppose that f_{uu}, f_{uv}, f_{vv} are continuous functions of x, u, v. For given $u \in \mathcal{A}$, the functional $\delta^2 I(u)$ that maps each $\phi \in C_0^1([a, b])$ to the value*

$$\delta^2 I(u)[\phi, \phi] := \int_a^b \Big(f_{uu} \phi^2 + 2 f_{uv} \phi \phi_x + f_{vv} \phi_x^2 \Big) \tag{174}$$

is referred to as the **second variation** *of the functional I at u. Following our convention, in (174) we assume the arguments $x, u(x), u_x(x)$ of the functions f_{uu}, f_{uv}, f_{vv}.*

Theorem 12.22. *Suppose that f_{uu}, f_{uv}, f_{vv} are continuous functions of x, u, v and the first variation $\delta I(u)$ vanishes, equivalently $u \in \mathcal{A}$ is a solution of the Euler-Lagrange equation (157). If there exists $\gamma > 0$ such that for the second variation $\delta^2 I(u)$ the following estimate holds:*

$$\delta^2 I(u)[\phi, \phi] \geq \gamma \int_a^b \left(\phi^2 + \phi_x^2\right) \qquad \forall \phi \in C_0^1([a,b]), \tag{175}$$

then u is a local minimiser of I in $\mathcal{A}$.

Proof. Omitted; not examinable. □

Remark 12.23. *In verifying the bound (175), the following inequality is sometimes helpful (see e.g. Question 3, Problem Sheet 9):*

$$\int_a^b u^2 \leq \left(\frac{b-a}{\pi}\right)^2 \int_a^b u_x^2 \qquad \forall u \in C_0^1([a,b]). \tag{176}$$

This is the so-called Friedrichs inequality for functions in $C_0^1([a,b])$. Its proof is actually an instructive exercise in calculus of variations. Indeed, suppose $v \neq 0$ minimises

$$F(u) := \frac{\int_a^b u_x^2}{\int_a^b u^2}$$

in $C_0^1([a,b])$. Then for all $\phi \in C_0^1([a,b])$ we have

$$0 = \frac{d}{dt} F(v + t\phi)\bigg|_{t=0} = \frac{d}{dt}\left(\frac{\int_a^b (v_x + t\phi_x)^2}{\int_a^b (v + t\phi)^2}\right)\Bigg|_{t=0} = \frac{2\int_a^b v_x \phi_x \cdot \int_a^b v^2 - \int_a^b v_x^2 \cdot 2\int_a^b v\phi}{\left(\int_a^b v^2\right)^2}.$$

Hence

$$\int_a^b \left(v_x \phi_x - \lambda v \phi\right) = 0 \qquad \forall \phi \in C_0^1([a,b]),$$

where $\lambda := F(v)$. By the Fundamental Theorem of Calculus of Variations (Theorem 12.7), the function v satisfies the eigenvalue problem

$$-v_{xx} = \lambda v \qquad \text{on } [a,b], \qquad\qquad v(a) = v(b) = 0. \tag{177}$$

The smallest eigenvalue is

$$\lambda = \left(\frac{\pi}{b-a}\right)^2, \tag{178}$$

which provides the bound (176). This only shows that the eigenfunction of (177) corresponding to the eigenvalue (178) is a possible local minimiser with the smallest value of F among all local minima, so a bit more work needs to be done to show that this is actually a global minimiser of F.

Note: The remainder of Section 12.5 is not examinable

Example 12.24. *(The solution to the Brachistochrone Problem) For the variation of the functional (cf. (135), (136))*

$$I(u) = \int_0^{x_1} \sqrt{\frac{u_x^2 + 1}{u}},$$

we have

$$\delta I(u)[\phi] = \int_0^{x_1} \left\{ \frac{u_x \phi_x}{\sqrt{u(u_x^2+1)}} - \frac{\sqrt{(u_x^2+1)}}{2u^{3/2}} \phi \right\} = 0 \qquad \forall \phi \in C_0^1([a,b]),$$

Formally applying Theorem 12.7, see Remark 12.10, we obtain that $u_x/\sqrt{u(u_x^2+1)}$ is differentiable and

$$-\left(\frac{u_x}{\sqrt{u(u_x^2+1)}} \right)_x - \frac{\sqrt{(u_x^2+1)}}{2u^{3/2}} = 0.$$

Notice that the Lagrangian

$$f(u,v) = \sqrt{\frac{v^2+1}{u}}$$

is x-independent, and so by applying the second remark in Section 12.4, we obtain

$$f(u, u_x) - u_x f_v(u, u_x) = c, \tag{179}$$

fr some constant c. Let us write (179) in more detail for our specific f :

$$\sqrt{\frac{u_x^2+1}{u}} - u_x \frac{u_x}{\sqrt{u(u_x^2+1)}} = c,$$

or equivalently

$$c^2 u(u_x^2 + 1) = 1. \tag{180}$$

This is a first-order ODE solvable by separating the variables x, u. Indeed, we can write (180) in the form

$$u_x = \pm\sqrt{\frac{1-c^2u}{c^2u}},$$

where we choose the $+$ sign, as from the physical interpretation of the solution a trajectory part of which has a negative slope can be deformed into one with a positive slope throughout, providing a lower value of the travel time $T(u)$ — compare the solid and dashed lines in Fig. 9. (We skip a rigorous proof of this claim.) Therefore, one has

$$\sqrt{\frac{c^2u}{1-c^2u}} du = dx, \tag{181}$$

and, upon integration, we obtain

$$x = \int_0^x dy = \int_0^u \sqrt{\frac{c^2v}{1-c^2v}} dv = \frac{1}{c^2} \int_0^{c^2u} \sqrt{\frac{w}{1-w}} dw, \tag{182}$$

where for the second equality we have used (181) and the boundary condition $u(0) = 0$ (effectively determining the constant of integration — remember that we will still need to determine the constant *c, for which we will use the second boundary condition $u(x_1) = y_1$, see Fig. 9), and for the third equality have made the change of variable $v = -c^2 w$.*

The integral on the right-hand side of (182) can be calculated by making the trigonometric change of variable $w = \sin^2\tau$, for which $dw = 2\sin\tau\cos\tau d\tau$. As a result, we have (denoting by t the upper limit of integration in τ, following the change of variable)

$$\begin{aligned} x &= \frac{1}{c^2}\int_0^{c^2 u}\sqrt{\frac{\sin^2\tau}{\cos^2\tau}}2\sin\tau\cos\tau d\tau = \frac{2}{c^2}\int_0^t \sin^2\tau d\tau \\ &= \frac{2}{c^2}\int_0^t \frac{1-\cos(2\tau)}{2}d\tau = \frac{1}{c^2}\left[\tau - \frac{\sin(2\tau)}{2}\right]_0^t = \frac{1}{c^2}\left(t - \frac{\sin(2t)}{2}\right), \end{aligned}$$

where $\sin^2 t = c^2 u$, equivalently $u = (1-\cos(2t))/(2c^2)$.

Modulo the (still undetermined) constant c, we have obtained a parametric representation of the sought curve:

$$\begin{cases} x(t) = \dfrac{1}{2c^2}\big(2t - \sin(2t)\big), \\ u(t) = \dfrac{1}{2c^2}\big(1-\cos(2t)\big), \end{cases}$$

or, relabelling $2t$ by θ,

$$\begin{cases} x = \dfrac{1}{2c^2}(\theta - \sin\theta), \\ u = \dfrac{1}{2c^2}(1-\cos\theta). \end{cases} \tag{183}$$

For each value c^2, the curve described by (183), which is referred to as a **cycloid**, *coincides with with the trajectory of a point on the rim of a circle (e.g. bicycle wheel) rolling without friction. The diameter of this circle is linked to the value c^2 : it is simply the maximum value of u, which is attained at $\theta = \pi$ and is equal to $c^{-2} = u(\pi/(2c^2))$. The curve (183) appears in several classical dynamics problems.*

In order to determine the constant c we need to satisfy the boundary condition (see Fig. 9) $u(x_1) = y_1$, which corresponds to solving the following transcendental system for the pair (c, θ):

$$\begin{cases} x_1 = \dfrac{1}{2c^2}(\theta - \sin\theta), \\ y_1 = \dfrac{1}{2c^2}(1-\cos\theta). \end{cases} \tag{184}$$

Denote by (c_*, θ_*) its solution. The total travel time for the particle is (see (135))

$$T(u) = \frac{1}{\sqrt{2g}}\int_0^{x_1}\sqrt{\frac{u_x^2+1}{u}} \tag{185}$$

and since, by (183), one has

$$u_x = \frac{du}{dx} = \frac{du}{d\theta}\left(\frac{dx}{d\theta}\right)^{-1} = \frac{\sin\theta}{1-\cos\theta} = \cot\left(\frac{\theta}{2}\right),$$

we infer

$$\int_0^{x_1}\sqrt{\frac{u_x^2+1}{u}}dx=\int_0^{\theta_*}\sqrt{\frac{u_x^2+1}{u}}\frac{dx}{d\theta}d\theta=\frac{1}{c_*\sqrt{2}}\int_0^{\theta_*}\sqrt{\frac{\cot^2(\theta/2)+1}{1-\cos\theta}}(1-\cos\theta)d\theta=\frac{\theta_*}{c_*}.$$

Combining this with (185) yields the minimal time

$$T_{\min}=\frac{\theta_*}{c^*\sqrt{2g}}.$$

13 Method of Lagrange Multipliers for Constrained Problems (Not examinable)

Suppose we would like to minimise a functional

$$I(u)=\int_a^b f(x,u,u_x)dx,$$

in the class $C^1([a,b])$, subject to the boundary conditions

$$u(a)=\alpha,\quad u(b)=\beta,\qquad \alpha,\beta\in\mathbb{R}, \tag{186}$$

and some *integral constrains*

$$J_j(u):=\int_a^b g_j(x,u,u_x)dx=c_j,\qquad j=1,\dots,N.$$

Theorem 13.1. *Suppose that $u\in C^1([a,b])$ is a local minimiser of the above constrained problem and there exists $\phi\in C_0^1([a,b])$ such that*

$$\int_a^b\big((g_j)_u(x,u,u_x)+(g_j)_v(x,u,u_x)\phi_x\big)dx\neq 0.$$

Then there exists $\lambda_j\in\mathbb{R}$, $j=0,1,\dots N$, such that for

$$h:=f(x,u,v)+\sum_{j=1}^N\lambda_j g_j(x,u,v)$$

one has

$$\int_a^b\big(h_u(x,u,u_x)+h_v(x,u,u_x)\phi_x\big)dx=0\qquad\forall\phi\in C_0^1([a,b]).$$

Proof. By assumption there is are functions ψ_j, $j=1,\dots,N$, such that

$$\int_a^b\big((g_j)_u(x,u,u_x)\psi_j+(g_j)_v(x,u,u_x)(\psi_j)_x\big)dx=1,\qquad j=1,\dots,N.$$

For $\varepsilon,t_1,\dots,t_N$ such that $|\varepsilon|+|t_1|+\dots+|t_N|\le\varepsilon_0$, $\varepsilon_0>0$, define the functions

$$\Phi(\varepsilon,t):=I(u+\varepsilon\phi+t_1\psi_1+\dots+t_N\psi_N),$$

$$\Psi(\varepsilon, t_j) := J_j(u + \varepsilon\psi + t_j\psi_j), \qquad j = 1, \ldots, N.$$

For $j = 1, \ldots, N$, we have $(\Psi_j)_{t_j}(0,0) = 0$. Therefore, applying the Implicit Function Theorem, and, for sufficiently small ε_j, $j = 1, \ldots, N$, we obtain functions $\tau_j = \tau_j(\varepsilon)$ such that

$$|\varepsilon| + \big|\tau_1(\varepsilon)\big| + \cdots + \big|\tau_N(\varepsilon)\big| \leq \varepsilon_0,$$

and

$$\Psi_j\big(\varepsilon, \tau_j(\varepsilon)\big) = c_j \qquad \forall \varepsilon \in (-\varepsilon_j, \varepsilon_j), \qquad j = 1, \ldots, N,$$

and therefore

$$\tau_j'(0) = -(\Psi_j)_\varepsilon(0,0).$$

Furthermore, the functions $u = u(x)$ and

$$v_j(x) := u(x) + \varepsilon\phi(x) + t_j\psi_j(x), \quad x \in [a,b], \qquad j = 1, \ldots, N$$

all satisfy the same boundary conditions at $x = a$ and $x = b$, see (186), and the C^1-distances

$$\|v_j - u\|_{C^1([a,b])} \leq \varepsilon\|\phi\| + |t_j|\|\psi_j\|_{C^1([a,b])}, \qquad j = 1, \ldots, N,$$

all tend to zero as $\varepsilon \to 0$, $t_1 \to 0, \ldots t_n \to 0$. Thus, we have

$$\Phi\big(\varepsilon, \tau_j(\varepsilon)\big) \geq \Phi(0,0) \qquad \text{for sufficiently small } \varepsilon \qquad \forall j = 1, \ldots, N$$

and it follows that

$$\Phi_\varepsilon(0,0) + \Phi_{t_j}(0,0)\tau_j'(0) = 0, \qquad j = 1, \ldots, N.$$

Defining the Lagrange multipliers

$$\lambda_j = -\Phi_{t_j}(0,0), \qquad j = 1, \ldots, N,$$

which are all independent of ϕ, we obtain

$$\Phi_\varepsilon(0,0) + \sum_{j=1}^{N} (\Psi_j)_\varepsilon(0,0) = 0,$$

as claimed. □

Example 13.2. *Consider the problem of minimisation of the functional*

$$I(u) = \int_a^b \big(|u'|^2 + q|u|^2\big),$$

among $u \in C^1([a,b])$ satisfying the conditions $u(a) = u(b) = 0$ as well as the constraint

$$\int_a^b u^2 = 1.$$

then any minimiser of this problem is in $C^2([a,b])$ and satisfies

$$-u'' + qu = \lambda u, \qquad u(a) = u(b) = 0.$$

Thus the Lagrange multiplier λ is an eigenvalue of the Sturm-Liouville operator

$$L = -\frac{d^2}{dx^2} + q$$

subject to zero boundary conditions, the minimiser is an eigenfunction of this operator corresponding to the eigenvalue λ, and it is easily seen that λ is the smallest eigenvalue of L with these boundary conditions.

14 Appendix A: Some Useful Results from Calculus

14.1 Intermediate Value Theorem

Theorem 14.1. *Suppose that $f : [a,b] \to \mathbb{R}$ is continuous. Then for any value*

$$\theta \in \Big[\min_{[a,b]} f, \max_{[a,b]} f\Big],$$

there exists $x_\theta \in [a,b]$ such that $f(x_\theta) = \theta$.

14.2 Derivative Mean-Value Theorem

Theorem 14.2. *Consider a closed interval $[a,b] \subset \mathbb{R}$. For any function $u \in C^1([a,b])$, there exists $\xi(a,b)$ such that*

$$u(b) - u(a) = u'(\xi)(b-a)$$

Geometric interpretation: There exists a point ξ such that the tangent to the graph of u is parallel to the straight line in $\mathbb{R}^2$ connecting the points $(a, u(a))$, $(b, u(b))$.

Mechanical interpretation: For a point moving during a time interval $[a,b]$ with no sudden speed changes, there is a moment of time between the start and the finish when the speed equals the average speed over the whole interval.

14.3 Integral Mean-Value Theorem

Theorem 14.3. *For a function $u \in C([a,b])$, there exists $\xi \in (a,b)$ such that*

$$\int_a^b u = u(\xi)(b-a).$$

Geometric interpretation: The area under the graph of u equals the area of a rectangle with base $[a,b]$ and height in the interval $(\min_{[a,b]} u, \max_{[a,b]u})$.

14.4 Different Versions of the Taylor Formula

Theorem 14.4 (Peano)**.** *Suppose that $l \in \mathbb{N}$. For any function $u \in C^l(a,b)$ and $x \in (a,b)$, one has*

$$u(y) = \sum_{j=0}^{l-1} \frac{u^{(j)}(x)}{j!}(y-x)^j + o\big((y-x)^{l-1}\big), \qquad (a,b) \ni y \to x,$$

Theorem 14.5. *Suppose that $l \in \mathbb{N}$. For any function $u \in C^l([a,b])$ and $x \in (a,b)$, one has*

$$\begin{aligned} u(y) &= u(x) + u_x(x)(y-x) + \frac{u''(x)}{2!}(y-x)^2 \cdots + \frac{u^{(l-1)}(x)}{(l-1)!}(y-x)^{l-1} \\ &\quad + \frac{1}{(l-1)!}\left(\int_0^1 u^{(l)}\big(y + t(x-y)\big)t^{l-1}dt\right)(y-x)^l \\ &= \sum_{j=0}^{l-1} \frac{u^{(j)}(x)}{j!}(y-x)^j + \frac{1}{(l-1)!}\left(\int_0^1 u^{(l)}\big(y+t(x-y)\big)t^{l-1}dt\right)(y-x)^l \qquad y \in [a,b]. \end{aligned} \tag{187}$$

corollary

Corollary 14.6 (Lagrange)**.** *Combining (187) with Theorem 14.3, we obtain another form of the remainder in (187), namely for each* $y \in [a,b]$*, there exists* θ *(that depends on* x, y*) situated between the points* x *and* y*, such that*

$$u(y) = \sum_{j=0}^{l-1} \frac{u^{(j)}(x)}{j!}(y-x)^j + \frac{u^{(l)}(\theta)}{(l-1)!}(y-x)^l.$$

Corollary 14.7 (Hadamard)**.** *For a function* $u \in C^\infty(a,b)$ *one has, for any* $l \in \mathbb{N}$*,*

$$u(y) = \sum_{j=0}^{l-1} \frac{u^{(j)}(x)}{j!}(y-x)^j + (y-x)^l \omega_l(y,x), \qquad x, y \in (a,b),$$

for some $\omega_l \in C^\infty((a,b) \times (a,b))$*.*

Alternatively, one can say that for for a smooth u, the function

$$\frac{1}{(y-x)^l}\left(u(y) - \sum_{j=0}^{l-1} \frac{u^{(j)}(x)}{j!}(y-x)^j\right), \qquad x, y \in (a,b),$$

is infinitely smooth with respect to *both* x and y.

14.5 Integration by Parts

Theorem 14.8. *Suppose that* $\Omega \subset \mathbb{R}^d$ *is a bounded open set with a* C^1 *boundary* $\partial\Omega$*, and* $u, v \in C^1(\overline{\Omega})$*. Then, for any* $j \in \{1, \ldots, d\}$ *one has*[32]

$$\int_\Omega u \frac{\partial v}{\partial x_j} dx = \int_{\partial\Omega} uvn_j dS - \int_\Omega \frac{\partial u}{\partial x_j} v dx, \tag{188}$$

where $n_j = n_j(x)$ *is the* j*-th component of the exterior normal to* $\partial\Omega$*.*

Corollary 14.9 (Gauss Theorem)**.** *For a set* Ω *as described in Theorem 14.8 and a vector function* $u = (u_1, \ldots u_d)$*, one has*

$$\int_\Omega \operatorname{div} u = \int_{\partial\Omega} u \cdot n dS, \tag{189}$$

where $n = (n_1, \ldots, n_d)$ *is the exterior normal to* $\partial\Omega$*, and* $\operatorname{div} u$ *is the* **divergence** *of* u*, defined as the trace of its Jacobi matrix:*

$$\operatorname{div} u := \sum_{j=1}^d \frac{\partial u_j}{\partial x_j}.$$

[32]Note that when one writes (188), one says that the formula of integration by parts is applied to the functions u, v. For example is one says "Apply the formula of integration by parts to $\sin(x_1 x_2)$ and $\partial g/\partial x_2$, with $j = 1$" for $\Omega \subset \mathbb{R}^2$ and some function $g = g(x_1, x_2)$, it is implied that (188) is written with $u = \sin(x_1 x_2)$, $v = \partial g/\partial x_2$, $j = 1$, so one is dealing with the integral

$$\int_\Omega \sin(x_1 x_2) \frac{\partial^2 g}{\partial x_1 \partial x_2} dx.$$

The expression ion the right-hand side of (189) is a quantity with an immediate physical interpretation, the "total flux of u across Ω". So the Gauss Theorem links this to the information about u inside Ω, more precisely, it tells us that the total flux across the boundary depends only on the "divergence of u.

Corollary 14.10 (Green's identity)**.** *For a set Ω as described in Theorem 14.8 and functions $u, v \in C^2(\overline{\Omega})$, one has*

$$\int_\Omega (v\Delta u - u\Delta v)dx = \int_{\partial\Omega}\left(v\frac{\partial u}{\partial n} - u\frac{\partial v}{\partial n}\right)dS,$$

15 Appendix B: Solutions to Selected Exercises

15.1 Exercise 4.4

In order to verify (32), choose $R > 0$ such that ϕ vanishes outside the ball of radius $R/2$ and notice that

$$-\int_{\mathbb{R}^d}\Phi\Delta\phi = -\lim_{\varepsilon\searrow 0}\int_{B_R(0)\setminus\overline{B_\varepsilon(0)}}\Phi\Delta\phi. \tag{190}$$

Using the Green's identity for the integral under the limit, we obtain:

$$\begin{aligned}\int_{B_R(0)\setminus\overline{B_\varepsilon(0)}}\Phi\Delta\phi &= \int_{B_R(0)\setminus\overline{B_\varepsilon(0)}}\phi\Delta\Phi + \int_{S_\varepsilon(0)}\left(\Phi(y)\frac{\partial\phi(y)}{\partial n_y} - \phi(y)\frac{\partial\Phi(y)}{\partial n_y}\right)dS_y\\ &= \int_{S_\varepsilon(0)}\left(\Phi(y)\frac{\partial\phi(y)}{\partial n_y} - \phi(y)\frac{\partial\Phi(y)}{\partial n_y}\right)dS_y,\end{aligned} \tag{191}$$

where we have used the fact that ϕ vanishes in the vicinity of $S_R(0)$ and Φ satisfies the Laplace's equation away from zero. Now replacing ϕ in the second term under the integral in (191) by the formula (33), we obtain

$$\begin{aligned}\int_{S_\varepsilon(0)}\left(\Phi(y)\frac{\partial\phi(y)}{\partial n_y} - \phi(y)\frac{\partial\Phi(y)}{\partial n_y}\right)dS_y &= \int_{S_\varepsilon(0)}\Phi(y)\frac{\partial\phi(y)}{\partial n_y}dS_y - \int_{S_\varepsilon(0)} y\cdot\psi(y)\frac{\partial\Phi(y)}{\partial n_y}dS_y\\ &\quad - \phi(0)\int_{S_\varepsilon(0)}\frac{\partial\Phi(y)}{\partial n_y}dS_y\end{aligned} \tag{192}$$

It remains to estimate the first two integrals on the right-hand side of (192) by quantities that go to zero as $\varepsilon \to$ and notice that

$$\int_{S_\varepsilon(0)}\frac{\partial\Phi(y)}{\partial n_y}dS_y = -\frac{1}{(d-2)|S_1|}\int_{S_\varepsilon(0)}(r^{2-d})'\big|_{r=\varepsilon}dS_y = -\frac{1}{(d-2)|S_1|}(2-d)\varepsilon^{1-d}\int_{S_\varepsilon(0)}dS_y = 1. \tag{193}$$

The required estimates for the mentioned two integrals are as follows:

$$\begin{aligned}\left|\int_{S_\varepsilon(0)}\Phi(y)\frac{\partial\phi(y)}{\partial n_y}dS_y\right| &\le \|\phi\|_{C^1(\overline{B_R(0)})}\int_{S_\varepsilon(0)}\Phi(y)dS_y\\ &= \|\phi\|_{C^1(\overline{B_R(0)})}\frac{\varepsilon^{2-d}}{(d-2)|S_1|}\int_{S_\varepsilon(0)}dS_y = \frac{\varepsilon}{d-2}\|\phi\|_{C^1(\overline{B_R(0)})},\end{aligned} \tag{194}$$

and

$$\begin{aligned}\left|\int_{S_\varepsilon(0)} y\psi(y)\frac{\partial\Phi(y)}{\partial n_y}dS_y\right| &\le \frac{\|\phi\|_C}{(d-2)|S_1|}\int_{S_\varepsilon(0)}|y|\Big|(r^{2-d})'\big|_{r=\varepsilon}\Big|dS_y \\ &= \frac{\|\phi\|_C}{(d-2)|S_1|}\int_{S_\varepsilon(0)}\varepsilon(d-2)\varepsilon^{1-d}dS_y = \varepsilon\|\phi\|_{C(\overline{B_R(0)})}.\end{aligned} \tag{195}$$

Combining (192) with (193), (194), (195), we obtain

$$\int_{S_\varepsilon(0)}\left(\Phi(y)\frac{\partial\phi(y)}{\partial n_y} - \phi(y)\frac{\partial\Phi(y)}{\partial n_y}\right)dS_y = -\phi(0) + o(1), \qquad \varepsilon\to 0.$$

Finally, plugging this into (191), we obtain

$$\int_{B_R(0)\setminus\overline{B_\varepsilon(0)}}\Phi\Delta\phi = -\phi(0),$$

which in view (190) yields (32).

Exercise 15.1. *Repeat the above argument for the case $d = 2$.*

15.2 Exercise 4.3(a)

According to (27), we need to check that for any $\phi\in C_0^\infty(\mathbb{R})$, one has

$$\lim_{n\to\infty}\frac{1}{2}\sqrt{\frac{n}{\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{-nx^2/4}\phi(x)dx = \phi(0).$$

Since $\phi\in C_0^\infty(\mathbb{R})$, there exists $R>0$ such that $\phi(x)=0$ for $x\in\mathbb{R}\setminus[-R,R]$. Furthermore, according to Corollary 14.7, one has

$$\phi(x) = \phi(0) + x\omega(x), \qquad x\in[-R,R], \tag{196}$$

where $\omega\in C^\infty([-R,R])$ Hence,

$$\begin{aligned}\frac{1}{2}\sqrt{\frac{n}{\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{-nx^2/4}\phi(x)dx &= \frac{1}{2}\sqrt{\frac{n}{\pi}}\int_{-R}^{R}\mathrm{e}^{-nx^2/4}\big(\phi(0)+x\omega(x)\big)dx \\ &= \phi(0)\frac{1}{2}\sqrt{\frac{n}{\pi}}\int_{-R}^{R}\mathrm{e}^{-nx^2/4}dx + \frac{1}{2}\sqrt{\frac{n}{\pi}}\int_{-R}^{R}\mathrm{e}^{-nx^2/4}x\omega(x)dx.\end{aligned}$$

It remains to notice that

$$\frac{1}{2}\sqrt{\frac{n}{\pi}}\int_{-R}^{R}\mathrm{e}^{-nx^2/4}dx = \frac{1}{\sqrt{\pi}}\int_{-R\sqrt{n}/2}^{R\sqrt{n}/2}\mathrm{e}^{-z^2}dz \overset{n\to\infty}{\longrightarrow} \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{-z^2}dz = 1,$$

and since ω is smooth on $[-R,R]$ and, in particular, bounded, we have

$$\left|\frac{1}{2}\sqrt{\frac{n}{\pi}}\int_{-R}^{R}\mathrm{e}^{-nx^2/4}x\omega(x)dx\right| \le \Big(\max_{[-R,R]}|\omega|\Big)\frac{2}{\sqrt{n\pi}}\int_{-R\sqrt{n}/2}^{R\sqrt{n}/2}\mathrm{e}^{-z^2}|z|dz \overset{n\to\infty}{\longrightarrow} 0,$$

as

$$1/\sqrt{n}\overset{n\to\infty}{\longrightarrow}0, \qquad\qquad \int_{-R\sqrt{n}/2}^{R\sqrt{n}/2}\mathrm{e}^{-z^2}|z|dz \overset{n\to\infty}{\longrightarrow} 2\int_0^\infty\mathrm{e}^{-z^2}zdz < \infty.$$

References

[1] M. Abramowitz, I. A. Stegun, 1972. *Handbook of Mathematical Finctions,* National Bureau of Standards.

[2] V. I. Arnold, 2004. *Lectures on Partial Differential Equations*, Springer.

[3] B. M. Budak, A. A. Samarskii, A. N. Tikhonov, 1964. *A Collection of Problems on Mathematical Physics,* Pergamon Press.

[4] G. Buttazzo, M. Giaquinta, S. Hildebrandt, 1998. *One-Dimensional Variational Problems,* Oxford University Press

[5] R. Dautray, J.-L. Lions, 1990. *Mathematical Analysis and Numerical Methods for Science and Technology, Vol 1: Physical Origins and Classical Methods,* Springer. Available online: https://link-springer-com.ezproxy1.bath.ac.uk/book/10.1007%2F978-3-642-61527-6

[6] I. S. Gradshteyn, I. M. Ryzhik, 2007. *Table of Integrals, Series, and Products*, Elsevier.

[7] S. Howison, 2005. *Practical Applied Mathematics,* Cambridge Texts in Applied Mathematics, Cambridge University Press.

[8] M. Kot, 2014. *A First Course in the Calculus of Variations,* Student Mathematical Library, Vol. 72, American Mathematical Society, Providence, Rhode Island.

[9] D. G. Luenberger, 1969. *Optimization by Vector Space Methods,* Jon Wiley & Sons.

[10] J. Ockendon, S. Howison, A. Lacey, A. Movchan, 2003. *Applied Partial Differential Equations,* Oxford University Press.

[11] M. Renardy, R. C. Rogers, 1993. *An Introduction to Partial Differential Equations,* Texts in Applied Mathematics, Springer.

[12] K. F. Riley, M. P. Hobson, S. J. Bence, 2006. *Mathematical Methods for Physics and Engineering.* Cambridge University Press.

[13] W. Rudin, 1966. *Real and Complex Analysis,* McGraw-Hill Series in Higher Mathematics, McGraw Hill Book Company.

[14] W. A. Strauss, 2008. *Partial Differential Equations: An Introduction,* Wiley.

[15] A. N. Tikhonov, A. A. Samarskii, 1963. *Equations of Mathematical Physics,* Dover.

[16] V. S. Vladimirov, 1971. *Equations of Mathematical Physics,* Marcel Dekker Inc, New York.

[17] A. Zettl, 2012. *Sturm-Liouville Theory,* Mathematical Surveys and Monographs, Vol. 121, American Mathematical Society.