Andrej Bóna[*], Michael A. Slawinski[†]

Ray Theory: Characteristics and Asymptotics (draft)

July 24, 2007

[*]Senior Lecturer, Department of Exploration Geophysics, Curtin University of Technology, Perth, Australia

[†]Professor, Department of Earth Sciences, Memorial University, St. John's, Canada

Contents

List of Figures v

Acknowledgements vii

Preface ix

1 Characteristic equations of first-order linear partial differential equations . . 1
Preliminary remarks 1
1.1 Motivational example 1
1.1.1 General and particular solutions 1
1.1.2 Characteristics 2
1.2 Directional derivatives 3
1.3 Taylor expansion of solutions 5
1.4 Incompatibility of side conditions 8
1.4.1 Motivation: Linear equations in two dimensions 8
1.4.2 Generalization: Semilinear equations in higher dimensions 10
1.4.3 Relation between incompatible side conditions and directional derivatives . 11
1.5 System of linear first-order equations 12
Closing remarks 14
Exercises 14

2 Characteristic equations of second-order linear partial differential equations 27
Preliminary remarks 27
2.1 Motivational examples 27
2.1.1 Equation with directional derivative 28
2.1.2 Wave equation in one spatial dimension 32
2.1.3 Heat equation in one spatial dimension 36
2.1.4 Laplace equation 38
2.2 General formulation 38
2.2.1 Semilinear equations 38
2.2.2 Systems of semilinear equations 42
2.2.3 Quasilinear equations 44
2.3 Physical applications of semilinear equations 46

2.3.1 Laplace equation . . . 46
2.3.2 Heat equation . . . 47
2.3.3 Wave equation . . . 48
2.4 Physical applications of systems of semilinear equations . . . 48
2.4.1 Elastodynamic equations . . . 48
2.4.2 Maxwell equations . . . 49
Closing remarks . . . 50
Exercises . . . 50

3 Characteristic equations of first-order nonlinear partial differential equations . . . 57
Preliminary remarks . . . 57
3.1 General formulation . . . 57
3.2 Side conditions . . . 63
3.3 Physical applications . . . 63
3.3.1 Elasticity . . . 63
3.3.2 Electromagnetism . . . 72
Closing remarks . . . 73
Exercises . . . 73

4 Asymptotic solutions of differential equations . . . 79
Preliminary remarks . . . 79
4.1 Asymptotic series . . . 79
4.2 Choice of asymptotic sequence . . . 84
4.3 Asymptotic differential equations . . . 85
4.4 Eikonal equation . . . 86
4.5 Solution of eikonal equation . . . 88
4.6 Transport equation . . . 89
4.7 Solution of transport equation . . . 91
4.8 Higher-order transport equations . . . 96
4.9 Asymptotic solution of elastodynamic equations . . . 98
Closing remarks . . . 99
Exercises . . . 99

5 Caustics . . . 109
Preliminary remarks . . . 109
5.1 Singularities of transport equation . . . 109
5.2 Caustics as envelopes of characteristics . . . 109
5.3 Phase change on caustics . . . 111
5.3.1 Waves in isotropic homogeneous media . . . 111
5.3.2 Method of stationary phase . . . 113
5.3.3 Phase change . . . 115
Closing remarks . . . 116
Exercises . . . 116

6 Symbols of linear differential operators . . . 123
Preliminary remarks . . . 123
6.1 Motivational example . . . 123
6.2 General formulation . . . 124
6.2.1 Wavefront of distribution . . . 124
6.2.2 Principal symbol . . . 126
6.2.3 Symbol . . . 127
6.3 Physical applications . . . 127
6.3.1 Support of singularities . . . 127
6.3.2 Laplace equation . . . 127
6.3.3 Heat equation . . . 127
6.3.4 Wave equation . . . 128
Closing remarks . . . 129
Exercises . . . 129

7 Relations among discussed methods . . . 131
Preliminary remarks . . . 131
7.1 Characteristics and asymptotic solutions . . . 131
Closing remarks . . . 132
Exercises . . . 132

Appendices . . . 133
A Integral theorems . . . 133
A.1 Divergence theorem . . . 133
A.2 Curl theorem . . . 137
B Fourier transform . . . 141
B.1 Some spaces of functions defined on closed interval . . . 141
B.2 Fourier series . . . 143
B.3 Fourier transform . . . 146
C Distributions . . . 151
C.1 Dirac's delta . . . 151
C.2 Definition of distributions . . . 154
C.3 Operations on distributions . . . 155
C.4 Convolution . . . 157
D Elastodynamic equations . . . 158
D.1 Cauchy's equations of motion . . . 158
D.2 Stress-strain equations . . . 162
D.3 Elastodynamic equations . . . 162
E Scalar and vector potentials in elastodynamics . . . 164
E.1 Equations of motion . . . 164
E.2 Scalar and vector potentials . . . 165
E.3 Wave equations . . . 166
E.4 Equations of motion versus wave equations . . . 167
F Maxwell equations . . . 169

F.1 Formulation . . . 169
F.2 Scalar and vector potentials . . . 172
G Oscillatory flow in incompressible fluid . . . 176
G.1 Physical setup . . . 176
G.2 Conservation laws . . . 176
G.3 Hydrostatic equilibrium . . . 178
G.4 Density stratification . . . 179
G.5 Equations of motion . . . 180
G.6 Incompressible fluid . . . 181
G.7 On boundary conditions . . . 182
G.8 Characteristics . . . 183
H Transport equation on manifolds . . . 183
H.1 Volume forms . . . 183
H.2 Half-densities . . . 184
H.3 Hamilton equations . . . 185
H.4 Transport Equation . . . 185

List of symbols . . . 187

References . . . 189

List of Figures

1.1 Solution of equation (1.1) with side condition (1.3) 2

2.1 Illustration of Cauchy data 29
2.2 Characteristic curves for wave equation. 33
2.3 Characteristic curve for heat equation. 37

3.1 The Monge cone 58
3.2 Envelope of a family of lines 60
3.3 Characteristics of the eikonal equation 70
3.4 Level curves of the eikonal function 71
3.5 Eikonal function 72

5.1 Lines $z = -\xi x + 1/\xi$ and their envelope given by $\left[-1/\xi^2, 2/\xi\right]$ 121
5.2 Propagation of a parabolic wavefront 122

A.1 A rectangular box used to formulate the divergence theorem 134
A.2 Two connected rectangular boxes used to formulate the divergence theorem 136
A.3 Rectangles used to formulate the curl theorem 138
A.4 Two connected rectangles used to formulate the curl theorem 139
B.1 The first ten terms of the Fourier series of function $f(x) = x$ 147
C.1 Several members of sequence (C.3), which defines Dirac's delta 152

Acknowledgements

Vassily M. Babich, Nelu Bucataru, David Dalton, Michael Rochester

Preface

la physique ne nous donne pas seulement l'occasion de résoudre des problèmes; elle nous aide à en trouver les moyens, et cela de deux manières. Elles nous fait pressentir la solution; elle nous suggère des raisonnements.[1]

Henri Poincaré (1905) *La valeur de la science*

In these lecture notes we strive to explain the understanding of the underpinnings of ray theory. These notes are intended for senior undergraduate and graduate students interested in the modern treatment of ray theory expressed in mathematical language. We assume that the reader is familiar with linear algebra, differential and integral calculus, vector calculus as well as tensor analysis.

To investigate seismic wave propagation, we often use the concepts of rays and wavefronts. These concepts result from studying the elastodynamic equations using the method of characteristics or using the high-frequency approximation. Characteristics of the elastodynamic equations are given by the eikonal function whose level sets are wavefronts. Characteristic equations of the eikonal equation are the Hamilton ray equations whose solutions are rays. Hence, rays are bicharacteristics of the elastodynamic equations. Characteristics are entities that are associated with differential equations in a way that is invariant under a change of coordinates. This property illustrates the fact that characteristics possess information about the physical essence of a given phenomenon.

Several key aspects of the method of characteristics for studying partial differential equations were introduced in the second half of the eighteenth century by Paul Charpit and Joseph-Louis Lagrange, and further elaborated at the beginning of the nineteenth century by Augustin-Louis Cauchy and Gaspard Monge.[2] Also, in the second half of the twentieth century this method has been significantly extended, as we discuss in these lecture notes.

Each chapter begins with a section called *Preliminary remarks*, where we provide the motivation for the specific concepts discussed therein, outline the structure of the chapter and

[1] physics not only provides us with the opportunity to solve problems, but also helps us to find the methods to get these solutions; this being achieved in two ways. Physics gives us a feeling for the solution and also suggests the path of reasoning.

[2] Readers interested in a mathematical description of the development of the method of characteristics might refer to Kline, M., (1972) Mathematical thought from ancient to modern times: Oxford University Press, Vol. II, pp. 531 – 538.

provide links to other chapters. Each chapter ends with a section called *Closing remarks*, which emphasizes the importance of the discussed concepts and show their relevance to other chapters. Each chapter is followed by *Exercises* and their solutions. Often, these exercises supply steps that are omitted from the exposition in the text. Such exercises are referred to in the main text.

1

Characteristic equations of first-order linear partial differential equations

Preliminary remarks

In this chapter we introduce the concept of characteristics by studying first-order linear partial differential equations. The understanding of characteristics in this context will help us to construct characteristics of more complex differential equations that we discuss later in these notes.

We begin this chapter by a motivational example in which we introduce the concept of characteristics. Subsequently, we use the directional derivative to find characteristics and define them rigorously. Then, we relate the characteristics to compatibility between the differential equation and its side conditions. We complete the discussions of this chapter by considering systems of first-order linear equations.

1.1 Motivational example

1.1.1 General and particular solutions

Consider the partial differential equation given by

$$\frac{\partial f\left(x_1, x_2\right)}{\partial x_1} = 0. \tag{1.1}$$

The general solution of equation (1.1) is

$$f\left(x_1, x_2\right) = g\left(x_2\right), \tag{1.2}$$

where g is an arbitrary function of x_2.

Often we wish to obtain a more specific solution. To do so, we impose extra conditions that the solution must satisfy. We refer to these conditions as side conditions. This terminology avoids distinctions between the initial and boundary conditions, which can be misleading in cases of differential equations whose variables are not associated with time or position. We can exemplify the use of a side condition in the following way.

Since the domain of f is the x_1x_2-plane, we can specify that, for instance,

$$f\left(x_1, x_2\right)|_{\gamma} = x_1^2 \tag{1.3}$$

along curve γ that is given by $x_2 = 2x_1$. We wish to express the side condition along line γ. Using $x_1 = x_2/2$, we can express f in terms of x_2 alone. Hence, we can write side condition (1.3) as

$$f\left(\frac{x_2}{2}, x_2\right) = \left(\frac{x_2}{2}\right)^2.$$

Using the right-hand side of this equation in solution (1.2), we see that $f(x_1, x_2) = g(x_2) = (x_2/2)^2$ is a solution of both equations (1.1) and (1.3). We can directly verify this solution since $\partial (x_2/2)^2 / \partial x_1 = 0$ and, along $x_2 = 2x_1$, $(x_2/2)^2 = x_1^2$. This solution is shown in Figure 1.1. It consists of a surface in the three-dimensional space spanned by x_1, x_2, y, where the range of f is along the y-coordinate. This surface is constant along the x_1-axis.

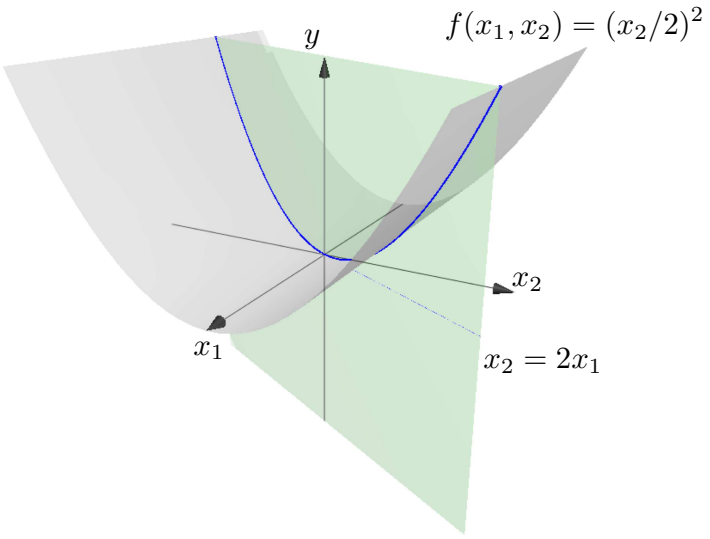


Fig. 1.1. Solution of equation (1.1) with side condition (1.3)

1.1.2 Characteristics

Using the concept of a side condition along a given line, we can illustrate that there are lines along which we cannot arbitrarily specify $f(x_1, x_2)$. To do so, let us specify that

$$f(x_1, x_2)|_\gamma = x_1^2, \tag{1.4}$$

where γ is the line given by $x_2 = C$ with C denoting a constant. Substituting $x_2 = C$ into solution (1.2), we obtain

$$f(x_1, x_2) = g(C) = A, \tag{1.5}$$

where A is a constant. Since, according to expression (1.5), $f(x_1, x_2)$ is constant along $x_2 = C$, it cannot be equal to x_1^2, as required by expression (1.4). These exceptional lines are the characteristics of equation (1.1). We see that along these curves we cannot set the side conditions

freely, since the behaviour of the solution along these curves is constrained by the differential equation itself.

1.2 Directional derivatives

The approach that we used to find characteristics of first-order partial differential equations with constant coefficients can be used to investigate first-order partial differential equations with variable coefficients.

Let us study the general differential first-order equation in n variables given by the following expression.

$$A_1(x_1, x_2, \ldots, x_n)\frac{\partial f}{\partial x_1} + A_2(x_1, x_2, \ldots, x_n)\frac{\partial f}{\partial x_2} + \cdots + A_n(x_1, x_2, \ldots, x_n)\frac{\partial f}{\partial x_n} = B(x_1, x_2, \ldots, x_n) f + C(x_1, x_2, \ldots, x_n) \quad (1.6)$$

This equation states that the solution, f, changes along direction $[A_1(x), A_2(x), \ldots, A_n(x)]$ at the rate of $B(x) f + C(x)$, as indicated by discussion of equation (1.32). In other words, the behaviour of solutions is prescribed along the curves whose tangent vectors are $A(x)$. Thus, these curves satisfy the following system of ordinary differential equations.

$$\begin{aligned}
\frac{\mathrm{d}x_1}{\mathrm{d}s} &= A_1(x_1(s), x_2(s), \ldots, x_n(s)), \\
\frac{\mathrm{d}x_2}{\mathrm{d}s} &= A_2(x_1(s), x_2(s), \ldots, x_n(s)), \\
&\vdots \\
\frac{\mathrm{d}x_n}{\mathrm{d}s} &= A_n(x_1(s), x_2(s), \ldots, x_n(s))
\end{aligned}$$

The original equation can be expressed as a derivative along the characteristics:

$$D_A f \equiv A \cdot \nabla f \equiv \frac{\mathrm{d}}{\mathrm{d}s} f(x(s)) = B(x(s)) f(x(s)) + C(x(s)). \quad (1.7)$$

This is a restatement of the fact that equation (1.6) describes the behaviour of the solutions only along the characteristics. The behaviour of the solutions in directions transverse to – meaning not tangent to – the characteristics must be given by extra conditions: the side conditions.

The reduction of partial differential equations into ordinary differential equations along the characteristics is a general property of the first-order partial differential equations. This property plays an important role in our studies and it results in the Hamilton equations, which are the ordinary differential equations discussed in Chapter 3.

The following example illustrates the construction of characteristics and their use to find solutions of first-order linear partial differential equations with variable coefficients.

Example 1.1. Let us consider

$$\frac{\partial f(x_1, x_2)}{\partial x_1} + x_2 \frac{\partial f(x_1, x_2)}{\partial x_2} = 0. \quad (1.8)$$

Since we can write equation (1.8) as

$$[1, x_2] \cdot \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2} \right] = 0,$$

we recognize that it is the directional derivative of f in the direction $[1, x_2]$. Following definition (1.34), we write equation (1.8) as

$$\mathrm{D}_{[1,x_2]} f(x_1, x_2) = 0, \tag{1.9}$$

which means that f does not change along the curve whose tangent vector is $[1, x_2]$; once we fix the value of f at a single point on this curve, the value of f is determined for all other points along the curve. Hence, this curve represents the characteristics of equation (1.8).

We can write the slope of the tangent to the characteristic as

$$\frac{\mathrm{d}x_2}{\mathrm{d}x_1} = \frac{x_2}{1} = x_2, \tag{1.10}$$

which we refer to as the characteristic equation. Equation (1.10) is an ordinary differential equation whose solution is the family of characteristic curves given by

$$x_2(x_1) = C \exp x_1, \tag{1.11}$$

with C being a constant that corresponds to the x_2-intercept of a given characteristic.

Solving expression (1.11) for C, we obtain $C = x_2 \exp(-x_1)$. Using the fact that, herein, f does not change along the characteristics and each characteristic is specified by the value of C, we can write the general solution of equation (1.8) as

$$f(x_1, x_2) = g(x_2 \exp(-x_1)), \tag{1.12}$$

where g is a differentiable function of one variable. We note that formally the differential equation expressed by equation (1.8) requires f to be differentiable. However, a nondifferentiable function g still accommodates solution (1.12) if we interpret the differential equation as the directional derivative (1.9).[1] To obtain a particular solution of equation (1.8), we can specify the value of g at a single point of each characteristic.

We note that although in general the solution is prescribed along the characteristics, it need not be constant, as illustrated in Exercise 1.2.

To obtain a particular solution of equation (1.8), let us specify the value of f at $x_1 = 0$, for each characteristic. This means that we specify the value of g along the x_2-axis. For instance, along this line, we let $f(0, x_2) = x_2^2$. In other words, since for each point along the x_2-axis, $x_2 = C$, at $x_1 = 0$, we set $f = C^2$, for each characteristic.

Now, we return to the general solution of equation (1.8). At $x_1 = 0$, solution (1.12) is

$$f(0, x_2) = g(x_2).$$

Since we let $f = x_2^2$ at $x_1 = 0$, this implies that $g(x_2) = x_2^2$. It means that g is a rule according to which we square the argument. Following solution (1.12), we can write this rule for all values

[1] The concept of nondifferentiable solutions is extended in Section **??** in the context of weak derivatives.

of x_1 to get

$$f(x_1, x_2) = [x_2 \exp(-x_1)]^2 = \frac{x_2^2}{\exp(2x_1)}.$$

This is a particular solution of equation (1.8) with the side condition given by

$$f(x_1, x_2)|_\gamma = x_2^2, \tag{1.13}$$

where γ is a noncharacteristic line given by $x_1 = 0$, which is the x_2-axis.

Now, let us specify the value of f at $x_2 = 0$. In other words, let us specify the value of g for each point along the x_1-axis. For instance, along this line, we let $f(x_1, 0) = x_1^2$. We return to the general solution of equation (1.8). At $x_2 = 0$, solution (1.12) is

$$f(x_1, 0) = g(0).$$

This means that once we set the value at $x_2 = 0$, it remains the same for all points along the x_1-axis. Thus, since $g(0)$ is a constant, we cannot set it to be equal to x_1^2, which represents a function whose value changes with x_1.

To understand this result, let us return to the characteristics of equation (1.8) by recalling expression (1.11), namely, $x_2(x_1) = C \exp x_1$. We realize that $x_2 = 0$, which is the x_1-axis, is one of the characteristics; it corresponds to $C = 0$. Since equation (1.8) requires f to be constant along the characteristics, we cannot specify f to be x_1^2 along the x_1-axis.

Equation (1.8) together with side conditions (1.13) is referred to as a Cauchy problem. In general, a Cauchy problem consists of finding the solution of a differential equation that also satisfies the side conditions that are given along a hypersurface and consist of the values of all the derivatives of the order lower than the order of the differential equation.

1.3 Taylor expansion of solutions

Since a differential equation can be viewed as a relation among derivatives of its solution, it is natural to ask if one can use this relation to determine all the derivatives of the solution at a point to consider the Taylor expansion of the solution. In this section we use the first-order linear equations to explore this idea. An approach analogous to the one presented in this section can be used to find solutions of higher-order differential equation, as we do on page 40.

Consider the general first-order linear differential equation,

$$\sum_{i=1}^{n} A_i(x) \frac{\partial f}{\partial x_i} = B(x) f + C(x),$$

together with the following side condition along the hypersurface $x = x(s_1, \ldots, s_{n-1})$:

$$f(x(s_1, \ldots, s_{n-1})) = f_0(s_1, \ldots, s_{n-1}).$$

To find the first derivatives along hypersurface $x(s_1, \ldots, s_{n-1})$, we can differentiate the side condition with respect to the parameters s to obtain $n-1$ linear equations for the n derivatives, $\partial f / \partial x_i$. By evaluating the original differential equation at a point along the hypersurface, we

obtain another linear equation for the first derivatives. We can write these equations as

$$\begin{bmatrix} \frac{\partial x_1}{\partial s_1} & \frac{\partial x_2}{\partial s_1} & \cdots & \frac{\partial x_n}{\partial s_1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_1}{\partial s_{n-1}} & \frac{\partial x_2}{\partial s_{n-1}} & \cdots & \frac{\partial x_n}{\partial s_{n-1}} \\ A_1 & A_2 & \cdots & A_n \end{bmatrix} \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \frac{\partial f_0}{\partial s_1} \\ \vdots \\ \frac{\partial f_0}{\partial s_{n-1}} \\ Bf_0 + C \end{bmatrix}. \tag{1.14}$$

The last equation of the system is linearly independent of the first $n-1$ equations if and only if vector A is transverse to the hypersurface. In such a case, the above system is invertible, and we can find the first derivatives of the solution at any point of the hypersurface. Subsequently, we can differentiate the first derivatives with respect to s to obtain linear expressions for all the second derivatives except the second derivative in a transverse direction to the hypersurface. To complete this system of equations, we consider the derivative of the original differential equation in the transverse direction, which completes the system of equations for the second derivatives at x_0. We can proceed in a similar manner to obtain all the derivatives of the solution at this point. Having found all the derivatives, we can construct the following Taylor series for a function of n variables.

$$\sum_{\alpha_1=0}^{\infty} \cdots \sum_{\alpha_n=0}^{\infty} \frac{1}{\alpha_1!\alpha_2!\cdots\alpha_n!} \frac{\partial^{\alpha_1+\alpha_2+\cdots+\alpha_n} f(x_0)}{\partial x_1^{\alpha_1}\partial x_2^{\alpha_2}\cdots\partial x_n^{\alpha_n}} (x_1 - x_{01})^{\alpha_1} (x_2 - x_{02})^{\alpha_2} \cdots (x_n - x_{0n})^{\alpha_n},$$

which can be conveniently written using the multiindex notation as

$$\sum_{|\alpha|\geq 0} \frac{1}{\alpha!} \frac{\partial^{|\alpha|} f(x_0)}{\partial x^{\alpha}} (x - x_0)^{\alpha},$$

where α is a multiindex $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$, $|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_n$ and $x^{\alpha} = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$; to get familiar with this notation, the reader may refer to Exercises 4.4, 4.5 and 4.6. The convergence of this series to the solution of the Cauchy problem is guaranteed if the functions involved in the differential equation and the side conditions are analytic in a neighbourhood of x_0, as stated by Cauchy-Kovalevskaya theorem.[2] In the following example we return to equation (1.37) to find its solution using the Taylor series expansion.

Example 1.2. We want to find the solution of the Cauchy problem consisting of equation (1.37) together with Cauchy data given by

$$x_2 \frac{\partial f(x_1, x_2)}{\partial x_1} - x_1 \frac{\partial f(x_1, x_2)}{\partial x_2} = x_2,$$
$$f(0, x_2) = x_2^2.$$

Let us choose a point that is along the hypersurface given by the x_2-axis, say $x_0 = [0, 1]$, and find all the derivatives of the solution at this point. From the Cauchy data we see that the zeroth

[2] The proof of this theorem can be found, for example, in Courant and Hilbert (1989, Volume 2, pp. 48-54). Also, there exists a stronger version of this theorem, Holmgren's theorem, that does not require the analyticity of the side conditions. For more details, the reader might refer to Courant and Hilbert (1989, Volume 2, pp. 237-239.).

derivative at x_0 is $f(0,1) = 1$. The first derivatives can be found using system (1.14), which herein is

$$\begin{bmatrix} 0 & 1 \\ x_2 & -x_1 \end{bmatrix} \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 2x_2 \\ x_2 \end{bmatrix}.$$

Solving for the first derivatives, we get

$$\begin{aligned} \frac{\partial f}{\partial x_1}(0, x_2) &= 1, \\ \frac{\partial f}{\partial x_2}(0, x_2) &= 2x_2. \end{aligned} \tag{1.15}$$

Evaluating at x_0, we obtain

$$\begin{aligned} \frac{\partial f}{\partial x_1}(0,1) &= 1, \\ \frac{\partial f}{\partial x_2}(0,1) &= 2. \end{aligned}$$

The second derivatives can be obtained by differentiating expressions (1.15) with respect to x_2 and the original differential equation with respect to x_1.

$$\begin{aligned} \frac{\partial^2 f}{\partial x_2 \partial x_1}(0, x_2) &= 0, \\ \frac{\partial^2 f}{\partial x_2^2}(0, x_2) &= 2, \\ x_2 \frac{\partial^2 f(x_1, x_2)}{\partial x_1^2} - \frac{\partial f(x_1, x_2)}{\partial x_2} - x_1 \frac{\partial^2 f(x_1, x_2)}{\partial x_2 \partial x_1} &= 0. \end{aligned} \tag{1.16}$$

Evaluating at x_0, we obtain

$$\begin{aligned} \frac{\partial^2 f}{\partial x_2 \partial x_1}(0,1) &= 0, \\ \frac{\partial^2 f}{\partial x_2^2}(0,1) &= 2, \\ \frac{\partial^2 f}{\partial x_1^2}(0,1) &= 2, \end{aligned}$$

where in evaluating the last expression we used the values of derivatives obtained above. The third derivatives that contain derivatives with respect to x_2 are all zero, since they result from differentiating the second derivatives (1.16), which are all constants. The third derivative with respect to x_1 can be obtained by differentiating twice the original differential equation with respect to x_1.

$$x_2 \frac{\partial^3 f(x_1, x_2)}{\partial x_1^3} = 2 \frac{\partial^2 f(x_1, x_2)}{\partial x_1 \partial x_2} + x_1 \frac{\partial^3 f(x_1, x_2)}{\partial x_2 \partial x_1^2}. \tag{1.17}$$

Solving for $\partial^3 f / \partial x_1^3$ at x_0, we use the above values and the equality of the mixed partial derivatives to get

$$\frac{\partial^3 f}{\partial x_1^3}(0,1) = 0.$$

All the higher derivatives are zero, as can be seen from differentiating expression (1.17) with respect to x_1.

The initial terms of the Taylor series at point x_0 are

$$f(0,1) + \frac{\partial f}{\partial x_1}(0,1)\, x_1 + \frac{\partial f}{\partial x_2}(0,1)\,(x_2 - 1) + \frac{1}{2}\left(\frac{\partial^2 f}{\partial x_1^2}(0,1)\, x_1^2 + 2\frac{\partial^2 f}{\partial x_1 \partial x_2}(0,1)\, x_1 (x_2 - 1) + \frac{\partial^2 f}{\partial x_2^2}(0,1)\,(x_2 - 1)^2\right) + \cdots.$$

Using the above results, we write this series as

$$1 + x_1 + 2(x_2 - 1) + \frac{1}{2}\left(2x_1^2 + 2(x_2 - 1)^2\right) = x_1 + x_1^2 + x_2^2,$$

which is the solution of the Caquchy problem.

1.4 Incompatibility of side conditions

To obtain the Taylor-expansion solution discussed in Section 1.3 we require that the side conditions not be given along a curve that is parallel to vector A, which is the direction of the differentiation in the original differential equation. This directional derivative is discussed in Section 1.2, and allows us to obtain the characteristics. We have learnt that the behaviour of the solution along the characteristics is prescribed by the equation itself, and, hence, we cannot arbitrarily set the side conditions along the characteristic curves. This conclusion suggests a new way of looking at the characteristics and, consequently, leads us to another method for obtaining them. In this method we look at curves along which arbitrary side conditions lead to an incompatibility with the differential equation. Even though for the linear first-order equation these two methods are equivalent, this new approach is more general, as we will see by studying the second-order equations in Chapter 2.

1.4.1 Motivation: Linear equations in two dimensions

Let us return to the study of the general differential first-order equation in two variables given by expression (1.6), namely,

$$A_1(x_1, x_2)\frac{\partial f}{\partial x_1} + A_2(x_1, x_2)\frac{\partial f}{\partial x_2} = B(x_1, x_2)\, f + C(x_1, x_2). \tag{1.18}$$

According to the new approach, we want to find curves γ along which we cannot arbitrarily set the side conditions. To do so, let us find conditions under which the side conditions given along $\gamma(s) = [x_1(s), x_2(s)]$ are compatible with the differential equation itself. In other words, we will determine under which conditions the solutions can satisfy both the differential equation and the side condition.

If f is given along $\gamma(s) = [x_1(s), x_2(s)]$ by a side condition, its derivative $f'(s)$ along this curve is known. This derivative can be expressed in terms of the partial derivatives along x_1

and x_2 as

$$f'\left(x_1\left(s\right),x_2\left(s\right)\right)=\frac{\partial f}{\partial x_1}x_1'\left(s\right)+\frac{\partial f}{\partial x_2}x_2'\left(s\right). \tag{1.19}$$

Thus, we have two linear equations for unknowns $\partial f/\partial x_1$ and $\partial f/\partial x_2$: one from the differential equation (1.18) and one from the side condition (1.19), namely,

$$\begin{bmatrix} x_1'\left(s\right) & x_2'\left(s\right) \\ A_1\left(x_1,x_2\right) & A_2\left(x_1,x_2\right) \end{bmatrix}\begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix}=\begin{bmatrix} f'\left(s\right) \\ B\left(x_1,x_2\right)f+C\left(x_1,x_2\right) \end{bmatrix}. \tag{1.20}$$

This system cannot be solved uniquely for the two unknown derivatives if and only if the determinant of the coefficient matrix is zero, namely,

$$A_2\left(x_1\left(s\right),x_2\left(s\right)\right)x_1'\left(s\right)-A_1\left(x_1\left(s\right),x_2\left(s\right)\right)x_2'\left(s\right)=0. \tag{1.21}$$

Since $x_i'\left(s\right)$ stands for $\mathrm{d}x_i/\mathrm{d}s$, this condition is equivalent to

$$\mathrm{d}x_1A_2\left(x_1,x_2\right)=\mathrm{d}x_2A_1\left(x_1,x_2\right), \tag{1.22}$$

which is the characteristic equation.

In this case, equations (1.18) and (1.19) are either incompatible with one another or equivalent to one another.

If the two equations are equivalent to one another, then the characteristic equation (1.22) is satisfied and the equations are scalar multiples of one another. This equivalence can be written as the following compatibility condition.

$$\left[A_1\left(x_1\left(s\right),x_2\left(s\right)\right),A_2\left(x_1\left(s\right),x_2\left(s\right)\right),B\left(x_1\left(s\right),x_2\left(s\right)\right)f+C\left(x_1\left(s\right),x_2\left(s\right)\right)\right]=\zeta\left[x_1'\left(s\right),x_2'\left(s\right),f'\left(s\right)\right], \tag{1.23}$$

where ζ is a proportionality constant. In other words, the compatibility condition tells us whether or not the side condition given along a characteristic is compatible with the differential equation. In other words, equation (1.23) restates equation (1.7), since

$$f'=\frac{\partial f}{\partial x_1}x_1'\left(s\right)+\frac{\partial f}{\partial x_2}x_2'\left(s\right)=\frac{1}{\zeta}\left(A_1\frac{\partial f}{\partial x_1}+A_2\frac{\partial f}{\partial x_2}\right)=Bf+C.$$

The compatible side conditions along the characteristics do not add any information about the solutions; they are useless as side conditions.

Example 1.3. To illustrate the above results, let us revisit equation (1.8), namely,

$$\frac{\partial f\left(x_1,x_2\right)}{\partial x_1}+x_2\frac{\partial f\left(x_1,x_2\right)}{\partial x_2}=0, \tag{1.24}$$

and find the family of characteristic curves for this equation. Examining equation (1.24) and in view of equation (1.18), we see that $A_1\left(x_1,x_2\right)=1$, $A_2\left(x_1,x_2\right)=x_2$ and $B\left(x_1,x_2\right)=C\left(x_1,x_2\right)=0$. Firstly, we see that equation (1.22) becomes equation (1.10). Hence, the family of characteristic curves is given by equation (1.11), namely,

$$x_2(x_1) = C \exp x_1. \tag{1.25}$$

Secondly, for equation (1.24) we write the compatibility equation as

$$\frac{1}{x_1'(s)} = \frac{x_2(s)}{x_2'(s)} = \frac{0}{f'(s)}.$$

From this expression we can infer that

$$\frac{\mathrm{d}x_2}{\mathrm{d}x_1} = x_2$$

and

$$f'(s) = 0,$$

where the latter equation results from the assumption that $|x_1'| < \infty$, which is discussed below. The first equation is equivalent to characteristic equation (1.10). The second equation is equivalent to the statement that the solution, f, in this particular case, does not change along the characteristic curves, which is consistent with results in Section 1.2. We can see that if the side condition is given by a constant function along the characteristic, the solution of the differential equation would not be uniquely determined; any function that is constant along the characteristics and has a proper value along the side condition would be a solution. The compatible side condition along characteristics does not restrict the possible solutions, and thus does not serve the purpose of a side condition.

In the case of $f'(s) \neq 0$, we see that $x_1'(s) = \pm\infty$ and $x_2(s) = 0$. This results in a curve that coincides with the x_1-axis, which can be obtained by setting $C = 0$ in equation (1.25).

1.4.2 Generalization: Semilinear equations in higher dimensions

In this section we turn our attention to semilinear first-order partial differential equations in the n-dimensional case.

We want to study characteristic surfaces of the following equation.

$$A_1(x)\frac{\partial f}{\partial x_1} + A_2(x)\frac{\partial f}{\partial x_2} + \cdots + A_n(x)\frac{\partial f}{\partial x_3} = B(x, f). \tag{1.26}$$

A typical side condition for this equation would be given as a fixed value of f along a hypersurface. In the n-dimensional case, this hypersurface can be parametrized, at least locally, by $n-1$ parameters, say $s_1, \ldots, s_{n-1}$, namely,

$$\begin{aligned} x_1 &= x_1(s_1, \ldots, s_{n-1}), \\ x_2 &= x_2(s_1, \ldots, s_{n-1}), \\ &\vdots \\ x_n &= x_n(s_1, \ldots, s_{n-1}). \end{aligned}$$

Hence, we can write the side condition as

$$f\left(x\left(s_1,\ldots,s_{n-1}\right)\right)=f_0\left(s_1,\ldots,s_{n-1}\right).$$

Having stated the side condition, we can ask if it is compatible with the differential equation. More precisely, we can write a system of equations for $\partial f/\partial x_1, \partial f/\partial x_2, \ldots, \partial f/\partial x_n$, namely

$$\underbrace{\begin{bmatrix} \frac{\partial x_1}{\partial s_1} & \frac{\partial x_2}{\partial s_1} & \cdots & \frac{\partial x_n}{\partial s_1} \\ \frac{\partial x_1}{\partial s_2} & \frac{\partial x_2}{\partial s_2} & \cdots & \frac{\partial x_n}{\partial s_2} \\ \vdots & \vdots & \ddots & \vdots \\ A_1 & A_2 & \cdots & A_n \end{bmatrix}}_{M} \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \frac{\partial f_0}{\partial s_1} \\ \frac{\partial f_0}{\partial s_2} \\ \vdots \\ B \end{bmatrix}, \tag{1.27}$$

and ask how many solutions does the system possess. This system has a unique solution only if the determinant of the $n \times n$ matrix, denoted by M, is nonzero. If the determinant is equal to zero, there are either none or infinitely many solutions.

The determinant of M is zero if and only if the rows of matrix M are linearly dependent vectors. Since we consider a hypersurface, the first $n-1$ rows are linearly independent of each other. The only possible linear dependence of the rows can be expressed as

$$[A_1, A_2, \ldots, A_n] = \zeta_1 \left[\frac{\partial x_1}{\partial s_1}, \cdots, \frac{\partial x_n}{\partial s_1}\right] + \zeta_2 \left[\frac{\partial x_1}{\partial s_2}, \cdots, \frac{\partial x_n}{\partial s_2}\right] + \cdots + \zeta_{n-1} \left[\frac{\partial x_1}{\partial s_{n-1}}, \cdots, \frac{\partial x_n}{\partial s_{n-1}}\right]. \tag{1.28}$$

This expression states that vector A must be tangent to the characteristic surface; a characteristic surface is composed of characteristic curves whose tangents are parallel to A. This result is consistent with the fact that differential equation (1.26) determines the rate of change of the solution along direction A.

1.4.3 Relation between incompatible side conditions and directional derivatives

We have discussed a method that specifies hypersurfaces along which we are not free to set the side conditions. As discussed in Section 1.2, the differential equation itself governs the behaviour of the solutions along specific curves. Thus, these curves cannot be part of the hypersurface along which we specify the side conditions. We refer to these curves as characteristic curves.

In the two-dimensional case, the hypersurfaces are curves. In this case the characteristic hypersurfaces coincide with the characteristic curves, since the side conditions cannot be specified along these curves. In this case, both methods give the same result, as expected and as illustrated in Exercise 1.2.

In the three-dimensional case, the characteristic hypersurface is a surface that is composed of characteristic curves. This can be illustrated by expression (1.28). These curves are given by the direction discussed in the directional derivative approach.

In the higher-dimensional cases, the situation is analogous to the three-dimensional case. Therein, the characteristic hypersurface is composed of the characteristic curves as well.

1.5 System of linear first-order equations

A general system of m linear first-order partial differential equations in n variables can be written as

$$\sum_{j=1}^{n}\sum_{k=1}^{m} A_{ijk}(x)\frac{\partial f_k}{\partial x_j} = B_i(x), \quad \text{where} \qquad i = 1, 2, \ldots, m.$$

For the sake of further discussion, we will rewrite this system using the following notation.

$$\sum_{k=1}^{m} \mathcal{A}_{ik}(x) f_k = B_i(x), \tag{1.29}$$

where

$$\mathcal{A}_{ik}(x) := \sum_{j=1}^{n} A_{ijk}(x)\frac{\partial}{\partial x_j}.$$

Equation (1.29) can be written in the matrix form as

$$\begin{bmatrix} \mathcal{A}_{11}(x) & \mathcal{A}_{12}(x) & \cdots & \mathcal{A}_{1m}(x) \\ \mathcal{A}_{21}(x) & \ddots & & \\ \vdots & & \ddots & \vdots \\ \mathcal{A}_{m1}(x) & & \cdots & \mathcal{A}_{mm}(x) \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{bmatrix} = \begin{bmatrix} B_1(x) \\ B_2(x) \\ \vdots \\ B_m(x) \end{bmatrix}^T.$$

To decouple this system, we can transform matrix $\mathcal{A}$ into upper diagonal form. To do so, we can use Gaussian elimination. The system of equations in the new form can be expressed as

$$\begin{bmatrix} \mathcal{A}_{11}(x) & \mathcal{A}_{12}(x) & \cdots & \mathcal{A}_{1m}(x) \\ 0 & \mathcal{A}^2_{22}(x) & \cdots & \mathcal{A}^2_{2m}(x) \\ \vdots & & \ddots & \vdots \\ 0 & & \cdots & \mathcal{A}^m_{mm}(x) \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{bmatrix} = \begin{bmatrix} B_1(x) \\ D^1 B_2(x) \\ \vdots \\ D^{m-1} B_m(x) \end{bmatrix}^T, \tag{1.30}$$

where $\mathcal{A}^k_{ij}(x)$ denotes a linear differential operator of degree k and D^l denotes a differential operator of degree l.

The system of equations expressed in form (1.30) is essentially a decoupled system. In other words, it is possible to solve the last equation of the system, which is a linear partial differential equation of degree m. After this solution is found, one can substitute it into the second-last equation of the system, which becomes a linear partial differential equation of degree $m-1$. One can continue this recursive method until $f_1(x)$ is found. An example of such a system is discussed in the following example, as well as in Exercise 1.12.

Example 1.4. Find the solution of the following system of partial differential equations.

$$\begin{aligned} \frac{\partial f_1}{\partial x_1} + \frac{\partial f_1}{\partial x_2} + \frac{\partial f_2}{\partial x_1} - \frac{\partial f_2}{\partial x_2} &= 0 \\ \frac{\partial f_1}{\partial x_1} - \frac{\partial f_1}{\partial x_2} + 2\frac{\partial f_2}{\partial x_1} + \frac{\partial f_2}{\partial x_2} &= x_1. \end{aligned}$$

The system can be written as

$$\begin{bmatrix} \frac{\partial}{\partial x_1}+\frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_1}-\frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial x_1}-\frac{\partial}{\partial x_2} & 2\frac{\partial}{\partial x_1}+\frac{\partial}{\partial x_2} \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} 0 \\ x_1 \end{bmatrix}.$$

Acting on the first row by operator $\partial/\partial x_1 - \partial/\partial x_2$, acting on the second row by operator $\partial/\partial x_1 + \partial/\partial x_2$ and subtracting the results, we can replace the second equation by the result, namely,

$$\begin{bmatrix} \frac{\partial}{\partial x_1}+\frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_1}-\frac{\partial}{\partial x_2} \\ 0 & \left(\frac{\partial}{\partial x_1}+\frac{\partial}{\partial x_2}\right)\left(2\frac{\partial}{\partial x_1}+\frac{\partial}{\partial x_2}\right)-\left(\frac{\partial}{\partial x_1}-\frac{\partial}{\partial x_2}\right)^2 \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} 0 \\ \left(\frac{\partial}{\partial x_1}+\frac{\partial}{\partial x_2}\right) x_1 \end{bmatrix}.$$

The second equation of this system, which can be written as

$$\left(\frac{\partial^2}{\partial x_1^2}+5\frac{\partial^2}{\partial x_1 \partial x_2}\right) f_2 = 1, \tag{1.31}$$

is decoupled from unknown function f_1.

The solution of equation (1.31) by the method of characteristics is given in Exercise 2.7. The solution is

$$f_2(x_1,x_2) = \left(x_1-\frac{1}{5}x_2\right)\frac{1}{5}x_2 + g\left(x_1-\frac{1}{5}x_2\right) + h\left(\frac{1}{5}x_2\right).$$

After we substitute this solution into the first equation of the original system, the first equation becomes

$$\begin{aligned} \frac{\partial f_1}{\partial x_1}+\frac{\partial f_1}{\partial x_2} &= -\frac{1}{5}x_2 - g'\left(x_1-\frac{1}{5}x_2\right) + \frac{1}{5}x_1 - \frac{2}{25}x_2 - \frac{1}{5}g'\left(x_1-\frac{1}{5}x_2\right) + \frac{1}{5}h'\left(\frac{1}{5}x_2\right) \\ &= \frac{1}{5}x_1 - \frac{7}{25}x_2 - \frac{6}{5}g'\left(x_1-\frac{1}{5}x_2\right) + \frac{1}{5}h'\left(\frac{1}{5}x_2\right). \end{aligned}$$

The solution of this equation is given in Exercise 1.7 and in this case can be written as

$$f_1(y_1,y_1-y_2) = \int \frac{1}{5}y_1 - \frac{7}{25}(y_1-y_2) - \frac{6}{5}g'\left(y_1-\frac{1}{5}(y_1-y_2)\right) + \frac{1}{5}h'\left(\frac{1}{5}(y_1-y_2)\right) \mathrm{d}y_1 + c(y_2),$$

where $y_1 = x_1$ and $y_2 = x_1 - x_2$. We can integrate to obtain

$$f_1(y_1,y_1-y_2) = \frac{1}{10}y_1^2 - \frac{7}{50}y_1^2 + \frac{7}{25}y_1y_2 - \frac{3}{2}g\left(y_1-\frac{1}{5}(y_1-y_2)\right) + h\left(\frac{1}{5}(y_1-y_2)\right) + c(y_2),$$

which, in the original coordinates, is

$$f_1(x_1,x_2) = \frac{6}{25}x_1^2 - \frac{7}{25}x_1x_2 - \frac{3}{2}g\left(x_1-\frac{1}{5}x_2\right) + h\left(\frac{1}{5}x_2\right)x_1 + c(x_1-x_2).$$

To summarize, we recall the solution for f_2,

$$f_2(x_1,x_2) = \left(x_1-\frac{1}{5}x_2\right)\frac{1}{5}x_2 + g\left(x_1-\frac{1}{5}x_2\right) + h\left(\frac{1}{5}x_2\right).$$

Closing remarks

In this chapter, we have introduced the concept of the characteristic curves for linear partial differential equations. The key point of this introduction is the fact that we cannot arbitrarily set side conditions along the characteristics curves, since the differential equation itself prescribes restrictions along these curves.

In this chapter we saw that the characteristics are given by the differential equation itself. This is true also for higher-order equations as long as they are linear or even semilinear. It is not so for quasilinear equations, as we will see in Section 2.2.3.

Exercises

Exercise 1.1. Solve the following linear partial differential equation using the directional-derivative approach.

$$\frac{\partial f\left(x_1, x_2\right)}{\partial x_1} + c\frac{\partial f\left(x_1, x_2\right)}{\partial x_2} = 0, \tag{1.32}$$

where c is a constant. Suggest a manner of imposing the side conditions in order to obtain a particular solution; note that if $c = 0$, this equation reduces to equation (1.1). Equation (1.32) is referred to as the transport equation; justify this name.

Solution 1.1. We can rewrite equation (1.32) as

$$[1, c] \cdot \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}\right] \equiv [1, c] \cdot \nabla f = 0,$$

where the dot denotes the scalar product. We recognize that this is the directional derivative of f in direction $[1, c]$. Let us write equation (1.32) as

$$\mathrm{D}_{[1,c]} f\left(x_1, x_2\right) = 0, \tag{1.33}$$

where

$$\mathrm{D}_X := X \cdot \nabla \tag{1.34}$$

stands for the directional-derivative operator along vector X.

Equation (1.33) implies that $f\left(x_1, x_2\right)$ does not change along direction $[1, c]$. In other words, the equation states that f is constant along this direction. This means that, if we choose f at any point on a curve whose tangent is $[1, c]$, equation (1.32) determines the values of f for all the remaining points along this curve. As introduced in Section 1.1.2, the curves along which we cannot specify the side conditions are the characteristics. In the present case, the family of characteristics is composed of lines whose tangent is $[1, c]$; in other words,

$$x_2 - cx_1 = C, \tag{1.35}$$

with C being a constant that corresponds to the x_2-intercept of a given characteristic.

Since according to equation (1.32) f does not change along $x_2 - cx_1 = C$, once we choose the value of f at a given point, it will remain unchanged along these lines. Since each line is

distinguished from the others by the value of $x_2 - cx_1$, we can write the general solution of equation (1.32) as

$$f(x_1, x_2) = g(x_2 - cx_1), \tag{1.36}$$

where g is an arbitrary function.

To obtain a particular solution, we can, for instance, specify the value of f along the line $x_1 = 0$. In other words, we specify it for all points along the x_2-axis. Since $c \neq \infty$, the x_2-axis is not a characteristic, and, hence, we can specify an arbitrary function along this line. However, if $c = 0$ we cannot specify the value of f along this line, since the lines parallel to the x_1-axis are characteristic, as shown in the context of equation (1.1).

If x_1 in equation (1.32) represents time and x_2 represents position, we can view this equation as describing a physical system in which quantity f is being transported with speed c along the x_2-axis. Hence, this equation is referred to as the transport equation.

Exercise 1.2. Find the characteristics of equation

$$x_2 \frac{\partial f}{\partial x_1} - x_1 \frac{\partial f}{\partial x_2} = x_2, \tag{1.37}$$

using both the directional-derivative method and the incompatibility-of-side-conditions method.

Solution 1.2. To express equation (1.37) in terms of directional derivatives, we write

$$[x_2, -x_1] \cdot \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}\right] = x_2,$$

which, in view of expression (1.34), we can rewrite as

$$\mathrm{D}_{[x_2, -x_1]} f(x_1, x_2) = x_2.$$

This means that $f = x_2$ along the curve whose tangent vector is $[x_2, -x_1]$. We can write the tangent to this curve as

$$\frac{\mathrm{d}x_2}{\mathrm{d}x_1} = -\frac{x_1}{x_2}, \tag{1.38}$$

which is an ordinary differential equation. Separating the variables and integrating, we get

$$\int x_1 \mathrm{d}x_1 = -\int x_2 \mathrm{d}x_2,$$

which results in

$$x_1^2 + x_2^2 = C, \tag{1.39}$$

where C is the constant resulting from integration. Equation (1.39) gives the family of the characteristic curves, which are circles centered at the origin of the x_1x_2-plane whose radii are $\sqrt{C}$. To use the incompatibility-of-side-conditions method to find the characteristic curves, we examine equation (1.37) in the context of equation (1.18). We see that $a_1(x_1, x_2) = x_2$, $a_2(x_1, x_2) = -x_1$, $b(x_1, x_2) = 0$ and $c(x_1, x_2) = x_2$. Thus, we can rewrite equation (1.23) as

$$\frac{\mathrm{d}x_1}{x_2} = -\frac{\mathrm{d}x_2}{x_1} = \frac{\mathrm{d}f}{x_2}. \tag{1.40}$$

Considering the first equality, we can write

$$\frac{\mathrm{d}x_2}{\mathrm{d}x_1} = -\frac{x_1}{x_2},$$

which is equation (1.38) whose solution is given by expression (1.39), as expected.

Exercise 1.3. Find the solutions of equation (1.37) along its characteristics.

Solution 1.3. We can rewrite equation (1.40) as two equations, namely,

$$\frac{\mathrm{d}x_2}{\mathrm{d}x_1} = -\frac{x_1}{x_2}$$

and

$$\mathrm{d}f = \mathrm{d}x_1.$$

The corresponding solutions are

$$x_1^2 + x_2^2 = C_1 \tag{1.41}$$

and

$$f = x_1 + C_2, \tag{1.42}$$

respectively. Equation (1.41) describes the family of the characteristic curves and equation (1.42) describes a plane.
If we view equation (1.41) as an equation for a right circular cylinder, then the intersection of this cylinder with the plane is the graph of the solution along the characteristic, which – in this case – is an ellipse.

Exercise 1.4. Using equations (1.41) and (1.42), find and verify the general solution of equation (1.37).

Solution 1.4. From equation (1.41) we see that any constant can be written as a function of $x_1^2 + x_2^2$. Hence we can write equation (1.42) as

$$f = x_1 - e\left(x_1^2 + x_2^2\right).$$

Inserting f into equation (1.37), we can verify that it is a solution, namely,

$$\begin{aligned} x_2\frac{\partial}{\partial x_1}\left[x_1 - e\left(x_1^2 + x_2^2\right)\right] - x_1\frac{\partial}{\partial x_2}\left[x_1 - e\left(x_1^2 + x_2^2\right)\right] &= x_2\left[1 - \frac{\partial e\left(x_1^2 + x_2^2\right)}{\partial x_1}\right] + x_1\frac{\partial e\left(x_1^2 + x_2^2\right)}{\partial x_2} \\ &= x_2\left[1 - e'\frac{\partial\left(x_1^2 + x_2^2\right)}{\partial x_1}\right] + x_1 e'\frac{\partial\left(x_1^2 + x_2^2\right)}{\partial x_2} = x_2 - 2x_1x_2e' + 2x_1x_2e' = x_2, \end{aligned}$$

as required.

Exercise 1.5. Solve equation (1.37) using the directional-derivative approach.

Solution 1.5. We can rewrite equation (1.37) as

$$[x_2, -x_1] \cdot \nabla f = x_2,$$

which is equivalent to

$$\mathrm{D}_{[x_2,-x_1]}f = x_2. \tag{1.43}$$

The characteristics of equation (1.43) are curves whose tangent vectors are $[x_2, -x_1]$. Hence, these curves are solutions of the following system of ordinary differential equations.

$$\begin{aligned}\frac{\mathrm{d}x_1}{\mathrm{d}s} &= x_2\\ \frac{\mathrm{d}x_2}{\mathrm{d}s} &= -x_1.\end{aligned}$$

The solutions of this system are circles centered at the origin, namely

$$x_1^2 + x_2^2 = C_1, \tag{1.44}$$

which is equation (1.41). In parametric form, these solutions can be written as

$$\begin{aligned}x_1(s) &= C_1 \sin s\\ x_2(s) &= C_1 \cos s.\end{aligned} \tag{1.45}$$

The only thing remaining to show is how the function changes along these circles. This can be inferred from the right-hand side of equation (1.43). Expressing this equation in terms of parameter s, we write

$$\frac{\mathrm{d}f}{\mathrm{d}s} = x_2(s).$$

In view of expressions (1.45), we see that $f(s) = \int x_2(s)\,\mathrm{d}s = C_1 \int \cos s \mathrm{d}s = C_1 \sin s + C_2 = x_1(s) + C_2$. Since this is true along a characteristic for all s, we can write

$$f(x_1, x_2) = x_1 + C_2,$$

which is equation (1.42), as expected. The integration constant C_2 depends on the choice of the characteristic curve along which we integrate. Hence it is a function of these curves. Since, according to equation (1.44), we parametrize characteristic curves by quantity $x_1^2 + x_2^2$, we can write the above equation as

$$f(x_1, x_2) = x_1 + g\left(x_1^2 + x_2^2\right).$$

This is the general solution of equation (1.37), as verified in Exercise 1.4.

Exercise 1.6. Find the general solution of the following equation.

$$a_1\frac{\partial f}{\partial x_1} + a_2\frac{\partial f}{\partial x_2} + a_3\frac{\partial f}{\partial x_3} = b,$$

where a_i and b are constants, such that $a_3 \neq 0$.

Solution 1.6. We begin by writing this differential equation as a directional derivative, namely,

$$[a_1, a_2, a_3] \cdot \nabla f = b.$$

Hence, the characteristic curves are the solutions of

$$x_1'(s) = a_1,$$
$$x_2'(s) = a_2,$$
$$x_3'(s) = a_3.$$

These solutions are

$$x_1(s) = a_1 s + c_1,$$
$$x_2(s) = a_2 s + c_2,$$
$$x_3(s) = a_3 s + c_3,$$

where c_i are the integration constants. Herein, the characteristic curves are straight lines. Along these lines, the original partial differential equation becomes an ordinary differential equation, namely,

$$\frac{\mathrm{d}f}{\mathrm{d}s}(x_1(s), x_2(s), x_3(s)) = b.$$

The solution of this ordinary differential equation is

$$f(x_1(s), x_2(s), x_3(s)) = bs + f_0, \tag{1.46}$$

where f_0 is the integration constant that depends on the choice of a characteristic line along which we integrate. We can distinguish between different characteristic lines by setting $c_3 = 0$ and varying the values of c_1 and c_2. This way we change the coordinates x_1, x_2 and x_3 to s, c_1 and c_2. These new coordinates are related to the original ones by

$$s = \frac{1}{a_3} x_3$$
$$c_1 = x_1 - \frac{a_1}{a_3} x_3$$
$$c_2 = x_2 - \frac{a_2}{a_3} x_3.$$

Since the integration constant f_0 depends on the choice of the characteristic line, we can consider it to be an arbitrary function of c_1 and c_2. Thus, in the new coordinates, solution (1.46) becomes

$$f = bs + f_0(c_1, c_2).$$

Transforming this expression into the original coordinates, we obtain

$$f(x_1, x_2, x_3) = \frac{b}{a_3} x_3 + f_0\left(x_1 - \frac{a_1}{a_3} x_3, x_2 - \frac{a_2}{a_3} x_3\right),$$

which is the general solution of the original partial differential equation.

Exercise 1.7. Find the general solution of the following equation.

$$\frac{\partial f}{\partial x_1} + \frac{\partial f}{\partial x_2} = g(x_1, x_2),$$

where g is a function of x_1 and x_2.

Solution 1.7. We can rewrite this equation along the characteristic line $\left[x_1(s_1) + x_1^0(s_2), x_2(s_1) + x_2^0(s_2)\right]$ as

$$\frac{\partial f}{\partial s_1} = g\left(x_1(s_1, s_2), x_2(s_1, s_2)\right),$$

where s_1 is a parameter along the characteristic line and s_2 specifies the characteristic line. We can choose $s_2 = x_1 - x_2$ and $s_1 = x_1$. The solution of this equation is

$$f = \int g\left(x_1(s_1, s_2), x_2(s_1, s_2)\right) \mathrm{d}s_1 + h(s_2),$$

where h is an arbitrary differentiable function.

Exercise 1.8. Find the general solution of

$$a_1 \frac{\partial f}{\partial x_1} + a_2 \frac{\partial f}{\partial x_2} = \sin x_1, \tag{1.47}$$

where a_1 and a_2 are nonzero constants, by using the fact that first-order partial differential equations become ordinary differential equations along the characteristic curves.

Solution 1.8. To consider an ordinary differential equation along characteristic curves $[x_1(s), x_2(s)]$, we write the left-hand side of equation (1.47) as

$$\frac{\mathrm{d}f\left(x_1(s), x_2(s)\right)}{\mathrm{d}s} = \frac{\partial f}{\partial x_1}\frac{\mathrm{d}x_1}{\mathrm{d}s} + \frac{\partial f}{\partial x_2}\frac{\mathrm{d}x_2}{\mathrm{d}s},$$

Comparing this expression with equation (1.47), we see that

$$\frac{\mathrm{d}x_1}{\mathrm{d}s} = a_1 \tag{1.48}$$

and

$$\frac{\mathrm{d}x_2}{\mathrm{d}s} = a_2, \tag{1.49}$$

which are the characteristic equations of equation (1.47) whose solutions are the characteristic curves. Returning to equation (1.47), we write it along these curves as

$$\frac{\mathrm{d}f\left(x_1(s), x_2(s)\right)}{\mathrm{d}s} = \sin x_1(s). \tag{1.50}$$

To solve it, we will integrate this equation along the characteristic curves. Since to integrate along these curves we must integrate with respect to s, we first solve equations (1.48) and (1.49) to get

$$x_1 = a_1 s + x_1^0 \tag{1.51}$$

and

$$x_2 = a_2 s + x_2^0, \tag{1.52}$$

respectively. This expressions describe a family of characteristic curves, with $\left(x_1^0, x_2^0\right)$ being the point through which a particular curve passes. Herein, these curves are straight lines. Now, we rewrite equation (1.50) as

$$\mathrm{d}f\left(x_1\left(s\right),x_2\left(s\right)\right)=\sin\left(a_1s+x_1^0\right)\mathrm{d}s.$$

Integrating both sides, we obtain

$$f\left(x_1\left(s\right),x_2\left(s\right)\right)=-\frac{1}{a_1}\cos\left(a_1s+x_1^0\right)+g, \tag{1.53}$$

where g is the integration constant whose value depends on a particular line. Thus, we have obtained the solution of equation (1.47) along a characteristic line. Now, we wish to write the solution for the entire x_1x_2-plane. In other words, we wish to write expression (1.53) in terms of x_1 and x_2 only. To do so, we return to solutions (1.51) and (1.52). Since $\left(x_1^0,x_2^0\right)$ specifies a point in the x_1x_2-plane that lies on a given characteristic line and $a_2\neq0$, we can choose this point in such a way that $x_2^0=0$. This way, we choose x_1^0 to identify the characteristic lines; in such a case, the integration constant, g, is a function of x_1^0. Thus, letting $x_2^0=0$ and solving equations (1.51) and (1.52) for x_1^0 in terms of x_1 and x_2, we get

$$x_1^0=x_1-\frac{a_1}{a_2}x_2. \tag{1.54}$$

Using expressions (1.51) and (1.54) in solution (1.53), we write

$$f\left(x_1,x_2\right)=-\frac{1}{a_1}\cos\left(x_1\right)+g\left(x_1-\frac{a_1}{a_2}x_2\right), \tag{1.55}$$

which is the general solution of equation (1.47).

To verify solution (1.55), we return to equation (1.47) to get

$$a_1\frac{\partial f}{\partial x_1}+a_2\frac{\partial f}{\partial x_2}=a_1\left(\frac{1}{a_1}\sin x_1+g'\right)+a_2\left(-\frac{a_1}{a_2}g'\right)=\sin x_1, \tag{1.56}$$

as required. Herein, g' denotes the derivative of g with respect to its argument.

Exercise 1.9. Find the general solution of

$$a_1\frac{\partial f}{\partial x_1}+a_2\frac{\partial f}{\partial x_2}=\sin x_1, \tag{1.57}$$

where a_1 and a_2 are nonzero constants, by a convenient change of coordinates.

Solution 1.9. We would like to express the original differential equation in such a way that the left-hand side is a derivative with respect to a single variable. Hence, we consider

$$\frac{\partial}{\partial y_1}f\left(x_1\left(y_1,y_2\right),x_2\left(y_1,y_2\right)\right)=\frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial y_1}+\frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial y_1}. \tag{1.58}$$

Examining this expression together with the left-hand side of equation (1.57), we see that

$$\frac{\partial x_1}{\partial y_1}=a_1$$

and

$$\frac{\partial x_2}{\partial y_1}=a_2,$$

which implies that

$$x_1 = a_1 y_1 + A\left(y_2\right)$$

and

$$x_2 = a_2 y_1 + B\left(y_2\right).$$

These are equations that relate the original coordinates, x_1 and x_2, to the new coordinates, y_1 and y_2. We require these two equations be linearly independent; otherwise, functions A and B are arbitrary. Consequently, we can conveniently set $A\left(y_2\right) = a_1 y_2$ and $B\left(y_2\right) = -a_2 y_2$. Hence,

$$x_1 = a_1\left(y_1 + y_2\right) \tag{1.59}$$

and

$$x_2 = a_2\left(y_1 - y_2\right). \tag{1.60}$$

Using expressions (1.58) and (1.59), we write equation (1.57) in the new coordinates as

$$\frac{\partial f}{\partial y_1} = \sin\left[a_1\left(y_1 + y_2\right)\right].$$

Integrating, we obtain the solution given by

$$f\left(y_1, y_2\right) = -\frac{1}{a_1}\cos\left[a_1\left(y_1 + y_2\right)\right] + h\left(y_2\right),$$

where h is the integration constant whose constancy is with respect to y_1. To express this solution in the original coordinates, we note – in view of expression (1.59) – that the argument of the trigonometric function is x_1. Also, we solve equations (1.59) and (1.60) to get

$$y_2 = \frac{1}{2}\left(\frac{x_1}{a_1} - \frac{x_2}{a_2}\right).$$

Thus, we write the general solution as

$$f\left(x_1, x_2\right) = -\frac{1}{a_1}\cos x_1 + h\left(\frac{1}{2}\left(\frac{x_1}{a_1} - \frac{x_2}{a_2}\right)\right). \tag{1.61}$$

To verify solution (1.61), we return to equation (1.57) to get

$$a_1\frac{\partial f}{\partial x_1} + a_2\frac{\partial f}{\partial x_2} = a_1\left[\frac{1}{a_1}\sin x_1 + \frac{1}{2a_1}h'\right] + a_2\left[-\frac{1}{2a_2}h'\right] = \sin x_1,$$

as required. Herein, h' denotes the derivative of h with respect to its argument.

Exercise 1.10. Compare and discuss solutions (1.55) and (1.61).

Solution 1.10. As shown in Exercises 1.8 and 1.9, both solutions (1.55) and (1.61) satisfy the same differential equation, namely,

$$a_1\frac{\partial f}{\partial x_1} + a_2\frac{\partial f}{\partial x_2} = \sin x_1,$$

where a_1 and a_2 are nonzero constants. The apparent difference between these two solutions are the arguments of functions g and h. However, both these arguments represent the same family of characteristic curves. As shown in Exercises 1.8 and 1.9, in the first case we can write the argument as

$$x_1^0 = x_1 - \frac{a_1}{a_2} x_2,$$

while in the second case we can write it as

$$y_2 = \frac{1}{2}\left(\frac{x_1}{a_1} - \frac{x_2}{a_2}\right),$$

which, using the fact that a_1 is a constant, we can rewrite as

$$2a_1 y_2 = x_1 - \frac{a_1}{a_2} x_2 = x_1^0,$$

as expected. Thus, in both cases, the argument of g and h is an expression defining a given characteristic line. In Exercise 1.8, by setting $x_2^0 = 0$, we identified each line of the family of the characteristic lines by its intercept with the x_2-axis, which in such a case is given by x_1^0, while in Exercise 1.9, by requiring the linear independence of the two equations that relate the coordinates, we identified each of the characteristic lines by the value of y_2.

Exercise 1.11. Find the characteristic surfaces for the following equation.

$$x_2 \frac{\partial g}{\partial x_1} - x_1 \frac{\partial g}{\partial x_2} + x_3^2 = 1$$

Solution 1.11. We can write this equation as

$$[x_2, -x_1, 0] \cdot \nabla g = 1 - x_3^2.$$

The characteristic curves parametrized by s are the solutions of

$$\begin{aligned} x_1'(s) &= x_2, \\ x_2'(s) &= -x_1, \\ x_3'(s) &= 0. \end{aligned}$$

Since $x_3' = 0$, the solutions are restricted to the planes parallel to the x_1x_2-plane and hence we can study the solutions only in this plane. Then we can consider x_2 as a function of x_1. The first two equations become

$$\frac{\mathrm{d}x_2}{\mathrm{d}x_1} = -\frac{x_1}{x_2}.$$

This is a separable equation, whose solution is

$$x_1^2 + x_2^2 = C^2,$$

where C^2 is the integration constant. We conclude that the characteristic surfaces are composed of circles that are parallel to the x_1x_2-plane, and whose radius is C.

Exercise 1.12. Find the general solution of the following system of equations.

$$\begin{aligned}\frac{\partial f_1}{\partial x_1} - 2\frac{\partial f_1}{\partial x_2} + 2\frac{\partial f_2}{\partial x_1} + \frac{\partial f_2}{\partial x_2} &= x_1\\ \frac{\partial f_1}{\partial x_1} - \frac{\partial f_1}{\partial x_2} - \frac{\partial f_2}{\partial x_1} + \frac{\partial f_2}{\partial x_2} &= 0.\end{aligned}$$

Solution 1.12. We will reduce the system to the upper triangular form by applying

$$\frac{\partial}{\partial x_1} - \frac{\partial}{\partial x_2}$$

to the first equation, applying

$$\frac{\partial}{\partial x_1} - 2\frac{\partial}{\partial x_2}$$

to the second equation and subtracting the results. Thus, we write

$$\left(\frac{\partial}{\partial x_1} - \frac{\partial}{\partial x_2}\right)\left(\frac{\partial f_1}{\partial x_1} - 2\frac{\partial f_1}{\partial x_2} + 2\frac{\partial f_2}{\partial x_1} + \frac{\partial f_2}{\partial x_2}\right) - \left(\frac{\partial}{\partial x_1} - 2\frac{\partial}{\partial x_2}\right)\left(\frac{\partial f_1}{\partial x_1} - \frac{\partial f_1}{\partial x_2} - \frac{\partial f_2}{\partial x_1} + \frac{\partial f_2}{\partial x_2}\right) = 1$$

and hence obtain

$$\left(\frac{\partial}{\partial x_1} - \frac{\partial}{\partial x_2}\right)\left(2\frac{\partial f_2}{\partial x_1} + \frac{\partial f_2}{\partial x_2}\right) - \left(\frac{\partial}{\partial x_1} - 2\frac{\partial}{\partial x_2}\right)\left(-\frac{\partial f_2}{\partial x_1} + \frac{\partial f_2}{\partial x_2}\right) = 1$$

After simplifying this, we obtain the following second-order equation

$$3\frac{\partial^2 f_2}{\partial x_1^2} - 4\frac{\partial^2 f_2}{\partial x_1 \partial x_2} + \frac{\partial^2 f_2}{\partial x_2^2} = 1.$$

The characteristic equation of this partial differential equation is

$$3\left(x_1'\right)^2 - 4x_1'x_2' + \left(x_2'\right)^2 = 0.$$

Instead of considering this elimination of f_1 from the system, we can also eliminate f_2 from the system. To do so, we apply

$$-\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_2}$$

to the first equation, apply

$$2\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_2}$$

to the second equation and subtract the results. We obtain

$$\left(-\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_2}\right)\left(\frac{\partial f_1}{\partial x_1} - 2\frac{\partial f_1}{\partial x_2}\right) - \left(2\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_2}\right)\left(\frac{\partial f_1}{\partial x_1} - \frac{\partial f_1}{\partial x_2}\right) = -1.$$

This simplifies to

$$3\frac{\partial^2 f_1}{\partial x_1^2} - 5\frac{\partial f_1}{\partial x_1 \partial x_2} + \frac{\partial^2 f_1}{\partial x_2^2} = 1.$$

The characteristic equation of this partial differential equation is

$$3\left(x_1'\right)^2 - 5x_1'x_2' + \left(x_2'\right)^2 = 0.$$

After division by $\left(x_2'\right)^2$ the above equation becomes

$$3\left(\frac{\mathrm{d}x_1}{\mathrm{d}x_2}\right)^2 - 5\frac{\mathrm{d}x_1}{\mathrm{d}x_2} + 1 = 0.$$

The solutions of this algebraic equation are

$$\frac{\mathrm{d}x_1}{\mathrm{d}x_2} = \frac{5 \pm \sqrt{25-12}}{6} = \frac{5 \pm \sqrt{13}}{6}.$$

For simplicity, we denote these by $a_1$and a_2. Hence, the characteristic curves are the straight lines given by

$$x_1 = a_i x_2 + c_i,$$

where c_i are the integration constants with $i = 1, 2$. The change of coordinates along these lines results in

$$\frac{\partial^2 f_1}{\partial y_1 \partial y_2} = -1,$$

where $y_i = x_1 - a_i x_2$ are the new coordinates. In these coordinates the solution is obtained by the following integrations

$$\frac{\partial f_1}{\partial y_2} = -y_1 + g\left(y_2\right)$$

$$f_1\left(y_1, y_2\right) = -y_1 y_2 + G\left(y_2\right) + H\left(y_1\right),$$

where G and H are arbitrary differentiable functions. In the original coordinates this solution is

$$f_1\left(x_1, x_2\right) = -\left(x_1 - a_1 x_2\right)\left(x_1 - a_2 x_2\right) + G\left(x_1 - a_2 x_2\right) + H\left(x_1 - a_1 x_2\right).$$

Inserting this solution to the second equation of the original system, we obtain

$$\begin{aligned}\frac{\partial f_2}{\partial x_1} + \frac{\partial f_2}{\partial x_2} &= \frac{\partial f_1}{\partial x_2} - \frac{\partial f_1}{\partial x_1} \\ &= a_1\left(x_1 - a_2 x_2\right) + a_2\left(x_1 - a_1 x_2\right) - a_2 G'\left(x_1 - a_2 x_2\right) - a_1 H'\left(x_1 - a_1 x_2\right).\end{aligned}$$

If we denote the right-hand side of this equation by $R\left(x_1, x_2\right)$, the equation becomes

$$\frac{\partial f_2}{\partial x_1} + \frac{\partial f_2}{\partial x_2} = R\left(x_1, x_2\right).$$

To solve this equation, we can consider the characteristic lines of this equation, which are

$$x_1\left(s\right) = s + b_1$$

$$x_2\left(s\right) = s + b_2.$$

Along these lines the equation becomes

$$\frac{\mathrm{d}f_2}{\mathrm{d}s} = R\left(x_1\left(s\right), x_2\left(s\right)\right)$$

and its solution is given by the following integral,

$$f_2\left(x_1\left(s\right), x_2\left(s\right)\right) = \int R\left(x_1\left(s\right), x_2\left(s\right)\right) \mathrm{d}s + C,$$

where the integration constant, C, depends on the choice of the characteristic line along which we integrate. We can identify the lines by the choice of b_1 while setting b_2 equal to zero. The integral is

$$\int a_1\left(x_1\left(s\right) - a_2 x_2\left(s\right)\right) + a_2\left(x_1\left(s\right) - a_1 x_2\left(s\right)\right)$$
$$-a_2 G'\left(x_1\left(s\right) - a_2 x_2\left(s\right)\right) - a_1 H'\left(x_1\left(s\right) - a_1 x_2\left(s\right)\right) \mathrm{d}s =$$

$$\int a_1\left(s + b_1 - a_2 s\right) + a_2\left(s + b_1 - a_1 s\right) - a_2 G'\left(s + b_1 - a_2 s\right) - a_1 H'\left(s + b_1 - a_1 s\right) \mathrm{d}s =$$

$$\frac{1}{2} s^2 \left(a_1 - a_2\right)^2 + s\left(a_1 b_1 + a_2 b_1\right) - \frac{a_2}{1 - a_2} G\left(s + b_1 - a_2 s\right) - \frac{a_1}{1 - a_1} H\left(s + b_1 - a_1 s\right).$$

Hence,

$$f_2\left(x_1\left(s, b_1\right), x_2\left(s, b_1\right)\right) = \frac{1}{2} s^2 \left(a_1 - a_2\right)^2 + s\left(a_1 b_1 + a_2 b_1\right)$$
$$- \frac{a_2}{1 - a_2} G\left(s + b_1 - a_2 s\right) - \frac{a_1}{1 - a_1} H\left(s + b_1 - a_1 s\right) + C\left(b_1\right).$$

Since $s = x_2$ and $b_1 = x_1 - x_2$, the solution is

$$f_2\left(x_1, x_2\right) = \frac{1}{2} x_2^2 \left(a_1 - a_2\right)^2 + x_2\left(a_1 + a_2\right)\left(x_1 - x_2\right)$$
$$- \frac{a_2}{1 - a_2} G\left(2x_1 - x_2 - a_2 x_2\right) - \frac{a_1}{1 - a_1} H\left(2x_1 - x_2 - a_1 x_2\right) + C\left(x_1 - x_2\right),$$

where $C,\ G$ and H are functions and $a_i = \left(5 \pm \sqrt{13}\right)/6$.

2

Characteristic equations of second-order linear partial differential equations

Preliminary remarks

In Chapter 1 we saw that characteristic curves appear in two contexts. In the first context, we could solve the partial differential equation along the characteristics by reducing it to an ordinary differential equation. In the second context, the characteristics restrict the hypersurfaces along which the Cauchy data is meaningful for solving the equation by the Taylor expansion. In this chapter, we will see that it is not always possible to solve a partial differential equation along the characteristics by reducing it to a simpler form. However, the second context, the one in which the characteristics restrict the Cauchy data, is valid for higher-order equation.

The side conditions of the first-order semilinear equations are of the same form: they are given along a hypersurface, as we saw in Chapter 1. However, several forms of side conditions can be associated with higher-order equations. For most of this chapter, we will restrict our attention to the side conditions in the form of Cauchy data, which are used in the Taylor expansion of solutions.

We begin this chapter with three examples of second-order partial differential equations in two variables for which we find the characteristic curves. In these examples, we use the methods analogous to the ones used in Chapter 1. Subsequently, we formulate the general method to be applied to any second-order semilinear partial differential equation as well as the general method for systems of second-order semilinear partial differential equations. Also, examining this method, we will see that certain aspects of our approach extend to the second-order quasilinear partial differential equations. Having formulated the general method, we apply this approach to three equations of mathematical physics: the Laplace equation, the heat equation and the wave equation. We also apply this approach to two systems of equations: the elastodynamic equations and the Maxwell equations.

2.1 Motivational examples

In this section we will see that, unlike in the case of first-order linear partial differential equations, it is not always possible to find the equations of characteristics for second-order linear partial differential equations using the directional derivatives. In such a case, however, we can still resort to the incompatibility-of-side-conditions method introduced in Section 1.4.

2.1.1 Equation with directional derivative

Directional derivative

Consider the following differential equation.

$$\frac{\partial^2 f}{\partial x_1^2} + 2\frac{\partial^2 f}{\partial x_1 \partial x_2} + \frac{\partial^2 f}{\partial x_2^2} = 0 \tag{2.1}$$

We can rewrite this equation as

$$\left(\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_2}\right)\left(\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_2}\right) f = 0,$$

which can be expressed as

$$\mathrm{D}_{[1,1]}\mathrm{D}_{[1,1]} f = 0.$$

We see that direction $[1, 1]$ is a special direction for equation (2.1). The lines parallel to this direction are the characteristic curves of this equation. We can express these curves, which herein are straight lines, as

$$x_1 - x_2 = C, \tag{2.2}$$

where C is a constant. Herein, a solution of this equation is not generally constant along the characteristic curves. Indeed, as shown in Exercise 2.1, the solution is of the following form.

$$f(x_1, x_2) = (x_1 + x_2)\, g(x_1 - x_2) + h(x_1 - x_2), \tag{2.3}$$

where g and h are arbitrary functions of a single variable. Along the characteristic curves, g and h are constant, but the solution f is not. Following an argument analogous to the one on page 4, functions g and h need not be differentiable depending on the interpretation of the differential equation.

There are second-order differential equations that we cannot express in terms of directional derivatives. An example of such equations is the heat equation discussed in Section 2.1.3, below. However, we can still find their characteristics by the method introduced in Section 1.4: the incompatibility-of-side-conditions method.

Incompatibility of side conditions

To illustrate the incompatibility-of-side-conditions method for the second-order differential equations, let us consider the following problem.

We are given equation (2.1), namely,

$$\frac{\partial^2 f}{\partial x_1^2} + 2\frac{\partial^2 f}{\partial x_1 \partial x_2} + \frac{\partial^2 f}{\partial x_2^2} = 0$$

together with side conditions

$$f(x_1(s), x_2(s)) = f_0(s), \tag{2.4}$$

$$\mathrm{D}_N f(x_1(s), x_2(s)) = f_N(s), \tag{2.5}$$

where $[x_1(s), x_2(s)]$ is a curve parametrized by s, N is a vector normal to this curve, say $[x_2', -x_1']$, and f_0 and f_N are two given functions. Function f_0 specifies the value of f along the curve while function f_N specifies the directional derivative in the direction normal to this curve. Since f_N is the derivative along the normal to the curve and f_0 can be used to find the derivative along the curve, these two functions specify the derivative of f in any direction at any point on this curve, as illustrated in Exercise 2.3.

The Cauchy data alone do not provide us with the information about the second derivative in the direction transverse to the hypersurface along which the data are given. To find this derivative, we can invoke the differential equation itself. The differential equation will not provide the information about the second derivative in the transverse direction if the Cauchy data are given along the characteristics. In such cases, the Cauchy data might contradict the differential equation. The requirement that the side conditions do not contradict the differential equation is given by the compatibility condition.

By checking if side conditions are compatible with the differential equation, we also find the characteristic curves. Thus, we want to check whether or not the second derivatives along curve $[x_1(s), x_2(s)]$ satisfy the differential equation. Hence, we wish to determine the second derivatives along this curve using the given information. To do so, we start by expressing the first derivatives in terms of f_0 and f_N.

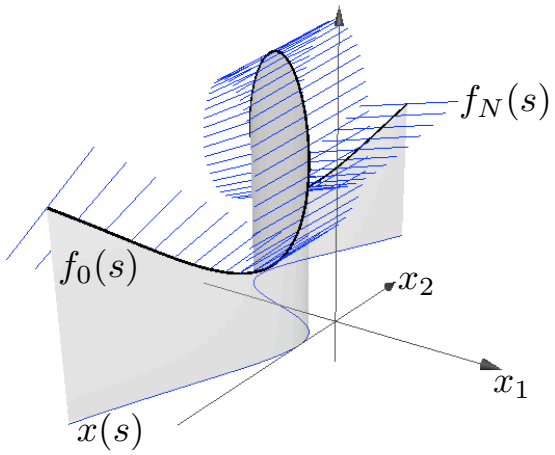


Fig. 2.1. The side condition along curve $[x_1(s), x_2(s)]$ specified by value of the solution, f_0, and the value of the normal derivative, f_N, along this curve.

Taking the derivative of equation (2.4) with respect to s, we get

$$x_1'(s)\frac{\partial f}{\partial x_1}(x_1(s), x_2(s)) + x_2'(s)\frac{\partial f}{\partial x_2}(x_1(s), x_2(s)) = f_0'(s).$$

In view of expression (1.34), we can rewrite the expression for the normal derivative given by equation (2.5) as

$$\left[x_2'(s), -x_1'(s)\right] \cdot \left[\frac{\partial f}{\partial x_1}\left(x_1(s), x_2(s)\right), \frac{\partial f}{\partial x_2}\left(x_1(s), x_2(s)\right)\right] = f_N(s).$$

The last two equations form a system of linear algebraic equations for the first derivatives, $\partial f/\partial x_1$ and $\partial f/\partial x_2$, along curve $[x_1(s), x_2(s)]$. This system can be written as

$$\begin{bmatrix} x_1'(s) & x_2'(s) \\ x_2'(s) & -x_1'(s) \end{bmatrix} \begin{bmatrix} \frac{\partial f}{\partial x_1}\left(x_1(s), x_2(s)\right) \\ \frac{\partial f}{\partial x_2}\left(x_1(s), x_2(s)\right) \end{bmatrix} = \begin{bmatrix} f_0'(s) \\ f_N(s) \end{bmatrix}.$$

Solving this system, we get

$$\frac{\partial f}{\partial x_1}\left(x_1(s), x_2(s)\right) = \frac{f_0'(s)\, x_1'(s) + f_N(s)\, x_2'(s)}{\left(x_1'\right)^2 + \left(x_2'\right)^2}, \tag{2.6}$$

$$\frac{\partial f}{\partial x_2}\left(x_1(s), x_2(s)\right) = \frac{-f_N(s)\, x_1'(s) + f_0'(s)\, x_2'(s)}{\left(x_1'\right)^2 + \left(x_2'\right)^2}. \tag{2.7}$$

We wish to find the second derivatives of f. For convenience, we denote the right-hand sides of the above expressions by a_1 and a_2, to write

$$\begin{aligned} \frac{\partial f}{\partial x_1}\left(x_1(s), x_2(s)\right) &= a_1(s), \\ \frac{\partial f}{\partial x_2}\left(x_1(s), x_2(s)\right) &= a_2(s). \end{aligned} \tag{2.8}$$

Differentiating these two equations with respect to s, we obtain

$$x_1'(s) \frac{\partial^2 f}{\partial x_1^2}\left(x_1(s), x_2(s)\right) + x_2'(s) \frac{\partial^2 f}{\partial x_1 \partial x_2}\left(x_1(s), x_2(s)\right) = a_1'(s), \tag{2.9}$$

$$x_1'(s) \frac{\partial^2 f}{\partial x_1 \partial x_2}\left(x_1(s), x_2(s)\right) + x_2'(s) \frac{\partial^2 f}{\partial x_2^2}\left(x_1(s), x_2(s)\right) = a_2'(s). \tag{2.10}$$

These are two equations for three unknowns, $\partial^2 f/\partial x_1^2$, $\partial^2 f/\partial x_1 \partial x_2$ and $\partial^2 f/\partial x_2^2$. The third equation, which is necessary to solve for the second derivatives, is the original differential equation. We can write the three equations as

$$\begin{bmatrix} x_1'(s) & x_2'(s) & 0 \\ 0 & x_1'(s) & x_2'(s) \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} a_1'(s) \\ a_2'(s) \\ 0 \end{bmatrix}, \tag{2.11}$$

where the last equation is equation (2.1). If the determinant of the coefficient matrix of this system is zero, the system has no unique solution.

Note that herein we are not interested in the compatibility condition since we are looking only for the characteristic curves, and not checking if the side conditions along these curves are compatible with the differential equation.

Setting the determinant to zero, we write

$$\left(x_1'(s)\right)^2 - 2x_1'(s)\,x_2'(s) + \left(x_2'(s)\right)^2 = 0,$$

which is

$$\left(x_1'(s) - x_2'(s)\right)^2 = 0.$$

Hence,

$$x_1'(s) = x_2'(s), \tag{2.12}$$

which means that

$$x_1(s) - x_2(s) = C,$$

where C is a constant. This is equation (2.2) parametrized by s.

The above equation describes the curves along which we cannot uniquely determine the second derivatives from the side conditions that are given along any of these curves. These are the characteristic curves. This implies that we are not able to solve uniquely the original differential equation with such side conditions that are given along the characteristic curves. There are two possible cases: either there are no solutions or there are infinitely many solutions. The case of no solutions results from the side conditions that contradict the differential equation. The case of infinitely many solutions results from the side conditions along which the differential equation does not add information about the second derivatives.

In view of the last two sections we see that in the case of equation (2.1) we can obtain the characteristic curves either by recognizing a special direction associated with this equation or by determining the curves along which we cannot obtain the unique solution for the second derivatives.

Compatibility conditions

Let us examine the case of infinitely many possible second derivatives. This case is tantamount to the original differential equation being a linear combination of the equations resulting from the side conditions. Herein, the condition for such a linear combination is called the compatibility condition. There are many ways of establishing such a condition, such as the Gaussian elimination method. To illustrate another way of obtaining the compatibility condition, let us revisit equation (2.1).

Examining system (2.11), we see that the condition for linear dependence can be written as

$$b_1\left[x_1', x_2', 0, a_1'\right] + b_2\left[0, x_1', x_2', a_2'\right] = [1, 2, 1, 0],$$

which is a system of four equations for six unknowns. From the first equation it follows that

$$b_1 = \frac{1}{x_1'}. \tag{2.13}$$

From the third equation it follows that

$$b_2 = \frac{1}{x_2'}. \tag{2.14}$$

Using these equalities, we see that the second equation implies that $x_1' = x_2'$, which is the characteristic equation given by expression (2.12). Then, the fourth equation yields $a_1' + a_2' = 0$, which is the sought after compatibility condition.

Thus, the compatibility condition along the characteristic curve, $x_1 - x_2 = C$, is given by

$$a_1' + a_2' = 0. \tag{2.15}$$

As shown in Exercise 2.4, this implies that

$$f_0(s) = x_1(s)\, k + l_1, \tag{2.16}$$

or equivalently,

$$f_0(s) = x_2(s)\, k + l_2,$$

where k, l_1 and l_2 are constants. We see that the side condition given by expression (2.16) along the characteristic curves does not provide any information that is not already contained in the original differential equation. The original differential equation states that the second derivative of a function in the direction of the characteristic curves is zero. This means that the function is increasing linearly in this direction, which is exactly the statement of expression (2.16). Hence such a side condition gives redundant information about the solutions. Note that we have not restricted the information given by f_N, which can be arbitrary.

2.1.2 Wave equation in one spatial dimension

Directional derivative

Consider the following differential equation.

$$\frac{\partial^2 f(x_1, x_2)}{\partial x_1^2} = c^2 \frac{\partial^2 f(x_1, x_2)}{\partial x_2^2}, \tag{2.17}$$

where c is a constant. We can interpret this equation as the wave equation if the independent variables x_1 and x_2 represent time and space, respectively. We can rewrite this equation as

$$\left(c\frac{\partial}{\partial x_2} + \frac{\partial}{\partial x_1}\right)\left(c\frac{\partial}{\partial x_2} - \frac{\partial}{\partial x_1}\right) f(x_1, x_2) = 0, \tag{2.18}$$

as shown in Exercise 2.6. Furthermore, we can write equation (2.18) as

$$\left\{[1, c] \cdot \left[\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}\right]\right\}\left\{[-1, c] \cdot \left[\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}\right]\right\} f(x_1, x_2) = 0,$$

where each term in braces is a directional-derivative operator. Thus, we write

$$\mathrm{D}_{[1,c]}\mathrm{D}_{[-1,c]} f(x_1, x_2) = 0,$$

where D_X denotes the directional derivative along vector X. In view of the equality of mixed partial derivatives, we can write the above equation as

$$\mathrm{D}_{[-1,c]}\mathrm{D}_{[1,c]}f\left(x_1,x_2\right)=0.$$

We see that any function that is constant along direction $[1,c]$ or $[-1,c]$ is a solution of equation (2.17). Thus

$$f\left(x_1,x_2\right)=g\left(x_2-cx_1\right)+h\left(x_2+cx_1\right) \tag{2.19}$$

is a solution of equation (2.17), and the straight lines given by

$$x_2\mp cx_1=C_{\pm} \tag{2.20}$$

are the characteristics, where C_+ parametrizes the characteristics with the slope $+c$ and C_- parametrizes the characteristics with the slope $-c$.

For the wave equation, as shown in Exercise 2.5, the side conditions specify the displacement along a space-time curve at the initial time, as well as the rate of change of displacement with time. If the space-time curve is a characteristic, we know – in view of expression (2.19) – that the displacement remains unchanged along this line. Hence, we are not free to specify its change. Pictorially, we represent characteristics (2.20) in Figure 2.2.

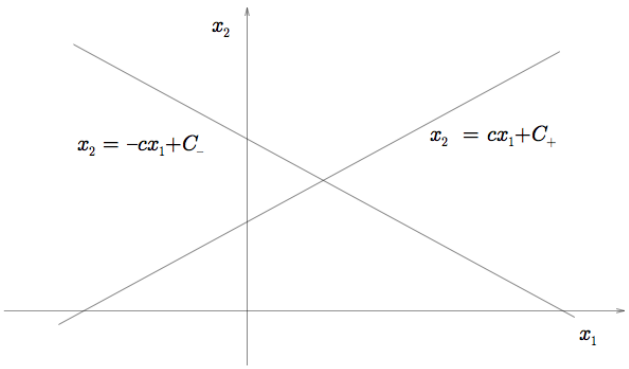


Fig. 2.2. Characteristic curves for wave equation.

Incompatibility of side conditions

Let us find the characteristics using the approach introduced in Section 1.4.

Consider equation (2.17) together with the side conditions along curve $\left[x_1\left(s\right),x_2\left(s\right)\right]$ that are given by

$$f\left(x_1\left(s\right), x_2\left(s\right)\right) = f_0\left(s\right), \tag{2.21}$$

$$\mathrm{D}_X f\left(x_1\left(s\right), x_2\left(s\right)\right) = f_X\left(s\right), \tag{2.22}$$

where X is a vector that is transverse to curve $[x_1(s), x_2(s)]$, say $[x_2', -x_1']$, and f_0 and f_X are given functions. Function f_0 specifies the value of f along curve $[x_1(s), x_2(s)]$ whereas function f_X specifies the directional derivative in the direction transverse to this curve, which we chose herein to be normal to the curve. The physical insight into this formulation of side conditions is discussed in Exercise 2.5.

Following the approach described in detail in Section 2.1.1, we obtain two equations for three unknowns, $\partial^2 f/\partial x_1^2, \partial^2 f/\partial x_1 \partial x_2$ and $\partial^2 f/\partial x_2^2$. These two equations are identical to equations (2.9) and (2.10). The third equation, which is necessary to solve for the second derivatives, is the original differential equation, herein equation (2.17). We can write the three equations as

$$\begin{bmatrix} x_1'(s) & x_2'(s) & 0 \\ 0 & x_1'(s) & x_2'(s) \\ 1 & 0 & -c^2 \end{bmatrix} \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} a_1'(s) \\ a_2'(s) \\ 0 \end{bmatrix}. \tag{2.23}$$

If the determinant of this system is zero, there is no unique solution. The determinant equal to zero means that

$$\left(c x_1'(s)\right)^2 = \left(x_2'(s)\right)^2,$$

which implies that

$$x_2'(s) = \pm c x_1'(s).$$

Hence,

$$x_2 \mp c x_1 = C_\pm \tag{2.24}$$

are curves along which we cannot uniquely determine the second derivatives from the differential equation and the side conditions that are given along these curves. This implies that we are not able to uniquely solve the original differential equation with the side conditions along any of these curves using the Taylor expansion. As expected, equation (2.24) is identical to equation (2.20).

We can follow this method further and look at the compatibility condition. In this case, the side conditions are compatible with the differential equation only if the equations in system (2.23) are linearly dependent on each other. This requirement translates into

$$a_1' x_1' = a_2' x_2',$$

which, considering the expression for the characteristics, can be written as

$$a_1' = \pm c a_2'. \tag{2.25}$$

After integrating this equation with respect to s, we obtain

$$a_1 = \pm c a_2 + k_\pm,$$

where $k_\pm$ depends on the characteristic along which we integrated the expression (2.25); it depends on $C_\pm$. Thus, we can write the compatibility conditions at point (x_1, x_2) as

$$a_1(x_1, x_2) \mp c a_2(x_1, x_2) = k_\pm(C_\pm(x_1, x_2)),$$

or, after substituting for a_1 and a_2 from expressions (2.8),

$$\frac{\partial f}{\partial x_1} \mp c\frac{\partial f}{\partial x_2} = k_\pm(C_\pm). \tag{2.26}$$

We can exploit the compatibility condition in a manner similar to the one used in Section 1.4 for linear first-order equations, and obtain solutions along the characteristics. To do so, we consider side conditions along a noncharacteristic curve. To compare our results with a more standard approach in other books, for example Folland (1995); McOwen (2003, p. 75)[1], we set them to be:

$$\begin{aligned} f(0, x_2) &= g(x_2), \\ \frac{\partial f}{\partial x_1}(0, x_2) &= h(x_2). \end{aligned} \tag{2.27}$$

Comparing these conditions with the compatibility condition (2.26), we conclude that

$$k_\pm(C(0, x_2)) = h(x_2) \mp c g'(x_2).$$

Since $k_\pm$ is constant along each characteristic, we can express its dependence on x_1 and x_2 as

$$k_\pm(x_1, x_2) = h(x_2 \mp c x_1) \mp c g'(x_2 \mp c x_1).$$

Substituting this result into equation (2.26), we get

$$D_{[1, \mp c]} f(x_1, x_2) = \frac{\partial f}{\partial x_1}(x_1, x_2) \mp c\frac{\partial f}{\partial x_2}(x_1, x_2) = h(x_2 \mp c x_1) \mp c g'(x_2 \mp c x_1).$$

This is a linear system for $\partial f/\partial x_1$ and $\partial f/\partial x_2$ whose solution is

$$\begin{aligned} \frac{\partial f}{\partial x_1}(x_1, x_2) &= \frac{1}{2}\left(h(x_2 - c x_1) + h(x_2 + c x_1) - c g'(x_2 - c x_1) + c g'(x_2 + c x_1)\right), \\ \frac{\partial f}{\partial x_2}(x_1, x_2) &= \frac{1}{2c}\left(h(x_2 + c x_1) - h(x_2 - c x_1) + c g'(x_2 + c x_1) + c g'(x_2 - c x_1)\right). \end{aligned}$$

Integrating with respect to x_1 and x_2 we obtain the solution of the original differential equation:

$$f = \frac{1}{2c}\int_{x_2 - c x_1}^{x_2 + c x_1} h(\xi)\,\mathrm{d}\xi + \frac{1}{2}\left(g(x_2 + c x_1) + g(x_2 - c x_1)\right).$$

To emphasize the fact that for obtaining the characteristics of second-order equations we need the side conditions containing the values of both the function and the transverse derivative along a hypersurface, we examine solution (2.19), namely,

[1] Note that in McOwen (2003) the term "compatibility condition" is used to express selfconsistency of the side conditions, and is not interchangeable with our use of the same term.

$$f(x_1, x_2) = g(x_2 - cx_1) + h(x_2 + cx_1).$$

We can set f along a characteristic curve, say $cx_1 = x_2$, to any function, say f_0. In such a case, solution (2.19) would be

$$f(x_1, x_2) = g(x_2 - cx_1) + f_0(x_2 + cx_1) - g(0).$$

At the first glance it might appear that this contradicts the defining property of a characteristic curve, namely, that we cannot arbitrarily set the side conditions along such a curve. However, we must remember that side conditions herein, given by expressions (2.21) and (2.22) correspond to both f_0 and f_X, which are functions that do not depend on one another. We see that the derivative of f in the direction of $[c, -1]$ along the characteristic $cx_1 = x_2$ is given by

$$\begin{aligned}[c,-1] \cdot \left[\frac{\partial f}{\partial x_1}(x_1, cx_1), \frac{\partial f}{\partial x_2}(x_1, cx_1)\right] &= [c,-1] \cdot \left[-cg'(0) + cf_0'(2cx_1), g'(0) + f_0'(2cx_1)\right] \\ &= \left(c^2 - 1\right) f_0'(2cx_1) - \left(c^2 + 1\right) g'(0),\end{aligned}$$

which must be equal to f_X; hence we cannot set the side conditions arbitrarily along these curves. The restrictions on the side conditions along the characteristics are compatibility conditions (2.26).

2.1.3 Heat equation in one spatial dimension

Consider the following differential equation

$$\frac{\partial^2 f}{\partial x_1^2} = c\frac{\partial f}{\partial x_2}, \tag{2.28}$$

where c is a constant. We can interpret this equation as the heat equation if the independent variables x_1 and x_2 represent space and time, respectively. We do not know how to write this equation using the directional derivative. However, we can investigate the possibility of obtaining the second derivatives from given side conditions.

Let us consider equation (2.28) together with the side conditions along curve $[x_1(s), x_2(s)]$ that are given by

$$f(x_1(s), x_2(s)) = f_0(s), \tag{2.29}$$

$$\mathrm{D}_X f(x_1(s), x_2(s)) = f_X(s), \tag{2.30}$$

where X is a vector that is not tangent to curve $[x_1(s), x_2(s)]$, say $[x_2', -x_1']$, and f_0 and f_X are functions, which are also given. Function f_0 specifies the value of f along curve $[x_1(s), x_2(s)]$ while function f_X specifies the directional derivative in the direction transverse to this curve, which herein we chose to be normal to the curve.

As shown in Sections 2.1.1 and 2.1.2, by differentiating equations (2.29) and (2.30) along the curve $[x_1(s), x_2(s)]$, we obtain two equations for three unknowns, $\partial^2 f/\partial x_1^2, \partial^2 f/\partial x_1 \partial x_2$ and $\partial^2 f/\partial x_2^2$, namely equations (2.9) and (2.10). The third equation is the original differential equation, herein equation (2.28). We can write the three equations as

$$\begin{bmatrix} x_1'(s) & x_2'(s) & 0 \\ 0 & x_1'(s) & x_2'(s) \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} f_0'' - \frac{\partial f}{\partial x_1} x_1'' - \frac{\partial f}{\partial x_2} x_2'' \\ f_X' - \frac{\partial f}{\partial x_1} x_2'' + \frac{\partial f}{\partial x_2} x_1'' \\ c \frac{\partial f}{\partial x_2} \end{bmatrix}. \tag{2.31}$$

If the determinant of the coefficient matrix of this system is equal to zero, the system has no unique solution. The determinant is equal to zero if

$$\left(x_2'(s)\right)^2 = 0,$$

which is the characteristic equation of equation (2.28). Hence,

$$x_2(s) = C \tag{2.32}$$

are curves along which we cannot determine uniquely the second derivatives from the side conditions that are given along these curves. This implies that we are not able to solve uniquely the original differential equation with the side conditions along any of these curves.

For the heat equation, the side condition represents the temperature along a space-time curve and a change of temperature with time. In this case the characteristic curve is a line $x_2 = C$. This means that we are not free to specify the side conditions along this line. Physically, this means that we cannot specify initial temperature together with its temporal change. Pictorially, we represent characteristics (2.32) in Figure 2.3.

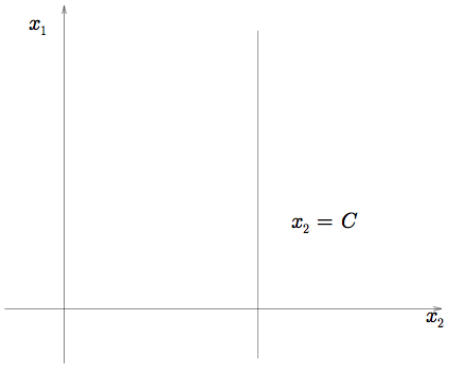


Fig. 2.3. Characteristic curve for heat equation.

If we try to find the solution of the heat equation following the method we used for the wave equation in Section 2.1.2, we fail. The reason for the failure is the fact that the middle equation of system (2.31) is always linearly independent of the other two equations. Along the

characteristics, which corresponds to the determinant of the system being zero, the first and third equations are multiples of one another. As shown in Exercise 2.2, following this method we will obtain only a trivial result: the solution has to satisfy the original heat equation. The reason for this is that we cannot integrate the second-order equation along the characteristics to obtain a first-order equation, as we did with the wave equation in equation (2.25).

2.1.4 Laplace equation

2.2 General formulation

In this section, we will generalize the method based on the incompatibility of side conditions, which allowed us to obtain the characteristic curves for all three equations discussed in Section 2.1. As we will see, this method works not only for linear equations but also for semilinear second-order partial differential equations and for systems of such equations.

2.2.1 Semilinear equations

Let us follow the procedure described in Section 2.1.1 for the general case of a semilinear second-order equation, namely,

$$\sum_{i,j=1}^{n} A_{ij}(x) \frac{\partial^2 f}{\partial x_i \partial x_j} = B\left(x, f, \frac{\partial f}{\partial x}\right), \tag{2.33}$$

where the right-hand side can depend on the unknown function and its first derivatives in a nonlinear way. We consider the side condition along hypersurface $x(s_1, \ldots, s_{n-1})$, as follows.

$$f(x(s_1, \ldots, s_{n-1})) = f_0(s_1, \ldots, s_{n-1}), \tag{2.34}$$

$$\mathrm{D}_X f(x(s_1, \ldots, s_{n-1})) = f_X(s_1, \ldots, s_{n-1}), \tag{2.35}$$

where X is a vector that is not tangent to hypersurface $x(s_1, \ldots, s_{n-1})$, and f_0 and f_X are the given functions. Function f_0 specifies the value of f along hypersurface $x(s_1, \ldots, s_{n-1})$, whereas function f_X specifies the value of directional derivative in the direction transverse to this surface.

To find the second derivatives of f along hypersurface $x(s_1, \ldots, s_{n-1})$, we begin by computing the first derivatives of f along this hypersurface. To do this, we differentiate expression (2.34) with respect to all $n-1$ parameters s_j, namely,

$$\sum_{i=1}^{n} \frac{\partial x_i(s_1, \ldots, s_{n-1})}{\partial s_j} \frac{\partial f}{\partial x_i}(x(s_1, \ldots, s_{n-1})) = \frac{\partial f_0(s_1, \ldots, s_{n-1})}{\partial s_j}.$$

This gives us $n-1$ equations for n unknowns, $\partial f / \partial x_i$. To obtain a system of n equations for n unknowns, we also use equation (2.35), which can be written as

$$\sum_{i=1}^{n} X_i(s_1, \ldots, s_{n-1}) \frac{\partial f}{\partial x_i}(x(s_1, \ldots, s_{n-1})) = f_X(s_1, \ldots, s_{n-1}).$$

We can solve this system to obtain

$$\frac{\partial f}{\partial x_i}\left(x\left(s_1,\ldots,s_{n-1}\right)\right)=a_i\left(s_1,\ldots,s_{n-1}\right), \tag{2.36}$$

where $a_i\left(s_1,\ldots,s_{n-1}\right)$ are the n solutions of the system. Having side conditions (2.34) and (2.35), we can always obtain the unique solution for a_i, as illustrated in Exercise 2.3. Now we are ready to find the equations for the second derivatives. To do so, we differentiate n equations (2.36) with respect to all $n-1$ parameters s_k, namely,

$$\sum_{j=1}^{n}\frac{\partial x_j\left(s_1,\ldots,s_{n-1}\right)}{\partial s_k}\frac{\partial^2 f}{\partial x_i\partial x_j}\left(x\left(s_1,\ldots,s_{n-1}\right)\right)=\frac{\partial a_i\left(s_1,\ldots,s_{n-1}\right)}{\partial s_k}. \tag{2.37}$$

This expression provides us with $n\left(n-1\right)$ equations for $\left(n+1\right)n/2$ variables, $\partial^2 f/\partial x_i\partial x_j$. However, among these $n\left(n-1\right)$ equations there are only $\left(n+1\right)n/2-1$ independent equations, since the right-hand side of equation (2.37) contains mixed derivatives. In the next paragraphs, we will give detailed explanation of this fact by introducing more convenient coordinates.

We choose a new coordinate system $\tilde{x}_1,\tilde{x}_2,\ldots,\tilde{x}_n$ in such a way, that, at least locally, the hypersurface is expressed in these coordinates as $\tilde{x}_n=0$ and the transverse vector along which we know the derivative of f along the hypersurface is expressed in the new coordinates as $[0,\ldots,0,1]$. This way, we can replace the $n-1$ parameters s_j by $n-1$ coordinates $\tilde{x}_1,\tilde{x}_2,\ldots,\tilde{x}_{n-1}$. In the new coordinates, the original differential equation and the side conditions are expressed as

$$\sum_{i,j=1}^{n}\tilde{A}_{ij}\left(\tilde{x}\right)\frac{\partial^2\tilde{f}}{\partial\tilde{x}_i\partial\tilde{x}_j}=\tilde{B}\left(\tilde{x},\tilde{f},\frac{\partial\tilde{f}}{\partial\tilde{x}}\right) \tag{2.38}$$

and

$$\begin{aligned}\tilde{f}\left(\tilde{x}_1,\ldots,\tilde{x}_{n-1},0\right)&=\tilde{f}_0\left(\tilde{x}_1,\ldots,\tilde{x}_{n-1}\right),\\ \mathrm{D}_{[0,\ldots,0,1]}\tilde{f}\left(\tilde{x}_1,\ldots,\tilde{x}_{n-1},0\right)&=\tilde{f}_X\left(\tilde{x}_1,\ldots,\tilde{x}_{n-1}\right),\end{aligned} \tag{2.39}$$

respectively, where $\tilde{f}\left(\tilde{x}_1,\ldots,\tilde{x}_n\right)=f\left(x_1\left(\tilde{x}_1,\ldots,\tilde{x}_n\right),\ldots,x_n\left(\tilde{x}_1,\ldots,\tilde{x}_n\right)\right)$. The form of equation (2.38) is the same as the form of equation (2.33). However, the form of the corresponding side conditions is more convenient herein.

In the new coordinates, the expressions for the first derivatives become

$$\begin{aligned}\frac{\partial\tilde{f}}{\partial\tilde{x}_i}\left(\tilde{x}_1,\ldots,\tilde{x}_{n-1},0\right)&=\frac{\partial\tilde{f}_0}{\partial\tilde{x}_i}\left(\tilde{x}_1,\ldots,\tilde{x}_{n-1}\right),\quad\text{for } i=1,\ldots,n-1,\\ \frac{\partial\tilde{f}}{\partial\tilde{x}_n}\left(\tilde{x}_1,\ldots,\tilde{x}_{n-1},0\right)&=\tilde{f}_X\left(\tilde{x}_1,\ldots,\tilde{x}_{n-1}\right).\end{aligned}$$

To find the equations for the second derivative, we differentiate these equations with respect to the parameters of the hypersurface, which are $\tilde{x}_1,\ldots,\tilde{x}_{n-1}$, to obtain

$$\frac{\partial^2\tilde{f}}{\partial\tilde{x}_j\partial\tilde{x}_i}\left(\tilde{x}_1,\ldots,\tilde{x}_{n-1},0\right)=\frac{\partial^2\tilde{f}_0}{\partial\tilde{x}_j\partial\tilde{x}_i}\left(\tilde{x}_1,\ldots,\tilde{x}_{n-1}\right), \tag{2.40}$$

where $i,j=1,\ldots,n-1$, and

$$\frac{\partial^2 \tilde{f}}{\partial \tilde{x}_j \partial \tilde{x}_n}\left(\tilde{x}_1, \ldots, \tilde{x}_{n-1}, 0\right) = \frac{\partial \tilde{f}_X}{\partial \tilde{x}_j}\left(\tilde{x}_1, \ldots, \tilde{x}_{n-1}\right), \tag{2.41}$$

where $j = 1, \ldots, n-1$. Expression (2.40) consists of $(n-1)^2$ equations. However, there are only $n(n-1)/2$ independent equations, due to the equality of the mixed partial derivatives. Expression (2.41) consists of $n-1$ independent equations. Combining these expressions, we obtain $(n+1)n/2-1$ equations for $(n+1)n/2$ unknowns, $\partial^2 \tilde{f}/\partial \tilde{x}_i \partial \tilde{x}_j$. We are still one equation short of completing this system.

We can complete this system by taking the original linear equation in the new coordinates, namely equation (2.38), and thus obtaining enough equations to solve for the second partial derivatives of $\tilde{f}$ along the hypersurface. The complete system is

$$\begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 \\ \tilde{A}_{11}(y) & \tilde{A}_{12}(y) & \cdots & \tilde{A}_{nn-1}(y) & \tilde{A}_{nn}(y) \end{bmatrix} \begin{bmatrix} \frac{\partial^2 \tilde{f}}{\partial \tilde{x}_1^2} \\ \frac{\partial^2 \tilde{f}}{\partial \tilde{x}_1 \partial \tilde{x}_2} \\ \vdots \\ \frac{\partial^2 \tilde{f}}{\partial \tilde{x}_n \partial \tilde{x}_{n-1}} \\ \frac{\partial^2 \tilde{f}}{\partial \tilde{x}_n^2} \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 \tilde{f}_0}{\partial \tilde{x}_1^2}\left(\tilde{x}_1, \ldots, \tilde{x}_{n-1}\right) \\ \frac{\partial^2 \tilde{f}_0}{\partial \tilde{x}_1 \partial \tilde{x}_2}\left(\tilde{x}_1, \ldots, \tilde{x}_{n-1}\right) \\ \vdots \\ \frac{\partial \tilde{f}_X}{\partial \tilde{x}_{n-1}}\left(\tilde{x}_1, \ldots, \tilde{x}_{n-1}\right) \\ \tilde{B}\left(y, \tilde{f}(y), \frac{\partial \tilde{f}}{\partial \tilde{x}}(y)\right) \end{bmatrix}, \tag{2.42}$$

where y stands for $(\tilde{x}_1, \ldots, \tilde{x}_{n-1}, 0)$, the $n-1$ penultimate equations are equations (2.41), and the form of the last component on the right-hand side is due to the fact that we evaluate equation (2.38) along the hypersurface.

If the determinant of the matrix of system (2.42) is nonzero, we can solve for the second derivatives of the solution along the side conditions. If the coefficients of the differential equation and the side conditions are analytic, we can compute all the higher derivatives, and write the Taylor series as discussed in Section 1.3. To find these derivatives, we can proceed in a similar manner to the one for the second derivatives. We can use f_0 to compute any derivative of the solution in the direction of the side conditions. The derivatives in the transverse direction to the side conditions can be calculated by differentiating the original differential equation in the transverse direction. In the situation where the determinant of the matrix of system (2.42) is zero, we cannot uniquely solve for the derivatives and construct the solution using the Taylor series. In the following paragraphs we turn our attention to the situation of the zero determinant of the above system.

To find the characteristic surface is to find a hypersurface along which we cannot uniquely solve for the second partial derivatives. The above system has no unique solution only if the determinant of the coefficient matrix equals to zero. Since the determinant of this matrix is $\tilde{A}_{nn}(\tilde{x}_1, \ldots, \tilde{x}_{n-1}, 0)$, we can express this condition as

$$\tilde{A}_{nn}\left(\tilde{x}_1, \ldots, \tilde{x}_{n-1}, 0\right) = 0. \tag{2.43}$$

To be able to make use of this equation in the context of the original problem, we express condition (2.43) in terms of the original coordinates.

The compatibility condition along the characteristic surface states that the equations in system (2.42) are linearly dependent, namely,

$$\tilde{A}_{11}(y)\frac{\partial^2\tilde{f}_0}{\partial\tilde{x}_1^2}(\tilde{x}_1,\ldots,\tilde{x}_{n-1})+\tilde{A}_{12}(y)\frac{\partial^2\tilde{f}_0}{\partial\tilde{x}_1\partial\tilde{x}_2}(\tilde{x}_1,\ldots,\tilde{x}_{n-1})+\cdots+\tilde{A}_{nn-1}(y)\frac{\partial\tilde{f}_X}{\partial\tilde{x}_{n-1}}(\tilde{x}_1,\ldots,\tilde{x}_{n-1})$$
$$=\tilde{B}\left(y,\tilde{f}(y),\frac{\partial\tilde{f}}{\partial\tilde{x}}(y)\right).$$

For convenience, we assume that the hypersurface can be expressed by $x_n=\psi(x_1,\ldots,x_{n-1})$. In this case, the new coordinates are given by

$$\tilde{x}_i=x_i,\quad i=1,\ldots,n-1,$$
$$\tilde{x}_n=x_n-\psi(x_1,\ldots,x_{n-1}),$$

where $\tilde{x}_n=0$, as required. Expressing the partial derivatives with respect to the original coordinates in terms of the new coordinates, namely,

$$\frac{\partial}{\partial x_i}=\frac{\partial}{\partial\tilde{x}_i}-\frac{\partial\psi}{\partial x_i}\frac{\partial}{\partial\tilde{x}_n},\quad i=1,\ldots,n,$$

we can write the second derivatives as

$$\begin{aligned}\frac{\partial^2}{\partial x_j\partial x_i}&=\left(\frac{\partial}{\partial\tilde{x}_j}-\frac{\partial\psi}{\partial x_j}\frac{\partial}{\partial\tilde{x}_n}\right)\left(\frac{\partial}{\partial\tilde{x}_i}-\frac{\partial\psi}{\partial x_i}\frac{\partial}{\partial\tilde{x}_n}\right)\\&=\frac{\partial^2}{\partial\tilde{x}_j\partial\tilde{x}_i}-\frac{\partial}{\partial\tilde{x}_j}\left(\frac{\partial\psi}{\partial x_i}\frac{\partial}{\partial\tilde{x}_n}\right)-\frac{\partial\psi}{\partial x_j}\frac{\partial^2}{\partial\tilde{x}_n\partial\tilde{x}_i}+\frac{\partial\psi}{\partial x_j}\frac{\partial}{\partial\tilde{x}_n}\left(\frac{\partial\psi}{\partial x_i}\frac{\partial}{\partial\tilde{x}_n}\right)\\&=\frac{\partial^2}{\partial\tilde{x}_j\partial\tilde{x}_i}-\frac{\partial\psi}{\partial x_j}\frac{\partial^2}{\partial\tilde{x}_n\partial\tilde{x}_i}-\frac{\partial\psi}{\partial x_i}\frac{\partial^2}{\partial\tilde{x}_j\partial\tilde{x}_n}+\frac{\partial\psi}{\partial x_j}\frac{\partial\psi}{\partial x_i}\frac{\partial^2}{\partial\tilde{x}_n^2}.\end{aligned}$$

From this expression we infer that the second derivatives in the original equations are related to the second derivatives in the new coordinates as follows.

$$\sum_{i,j=1}^{n}A_{ij}\frac{\partial^2}{\partial x_i\partial x_j}=\sum_{i,j=1}^{n}A_{ij}\left(\frac{\partial^2}{\partial\tilde{x}_j\partial\tilde{x}_i}-\frac{\partial\psi}{\partial x_j}\frac{\partial^2}{\partial\tilde{x}_n\partial\tilde{x}_i}-\frac{\partial\psi}{\partial x_i}\frac{\partial^2}{\partial\tilde{x}_j\partial\tilde{x}_n}+\frac{\partial\psi}{\partial x_j}\frac{\partial\psi}{\partial x_i}\frac{\partial^2}{\partial\tilde{x}_n^2}\right)$$

We want to find the expression for $\tilde{A}_{nn}$. To do so, we have to look for all the terms in the above expression that contain $\partial^2/\partial\tilde{x}_n^2$. These terms are

$$\tilde{A}_{nn}\frac{\partial^2}{\partial\tilde{x}_n^2}=\left[\sum_{i,j=1}^{n-1}A_{ij}\left(\frac{\partial\psi}{\partial x_i}\right)\left(\frac{\partial\psi}{\partial x_j}\right)-2\sum_{i=1}^{n-1}A_{in}\left(\frac{\partial\psi}{\partial x_i}\right)+A_{nn}\right]\frac{\partial^2}{\partial\tilde{x}_n^2},$$

where, in the first two summations, we used the fact that $\partial\psi/\partial x_i=0$ for $i=n$. Following equation (2.43), we require the term in brackets to be zero. Hence, we write explicitly:

$$\sum_{i,j=1}^{n-1}A_{ij}(x_1,\ldots,x_{n-1},\psi)\left(\frac{\partial\psi}{\partial x_i}\right)\left(\frac{\partial\psi}{\partial x_j}\right)$$
$$-2\sum_{i=1}^{n-1}A_{in}(x_1,\ldots,x_{n-1},\psi)\left(\frac{\partial\psi}{\partial x_i}\right)+A_{nn}(x_1,\ldots,x_{n-1},\psi)=0.\quad(2.44)$$

This equation is a first-order nonlinear partial differential equation for ψ in $n-1$ independent variables, $x_1, \ldots, x_{n-1}$. The solutions of this equation are functions whose graphs form the characteristic hypersurfaces in the space spanned by $x_1, \ldots, x_n$. Along these hypersurfaces we cannot uniquely solve for the second partial derivatives from the knowledge of the side conditions and the differential equation.

To illustrate this result, we consider the following differential equation in two independent variables.

$$A_{11}\frac{\partial^2 f}{\partial x_1^2} + A_{12}\frac{\partial^2 f}{\partial x_1 \partial x_2} + A_{21}\frac{\partial^2 f}{\partial x_2 \partial x_1} + A_{22}\frac{\partial^2 f}{\partial x_2^2} = B_1\frac{\partial f}{\partial x_1} + B_2\frac{\partial f}{\partial x_2} + Cf + D.$$

If we consider the characteristic curve to be parametrized by x_1, namely $x_2 = \psi(x_1)$, we obtain the characteristic equation given by

$$A_{11}\left(\frac{\mathrm{d}\psi}{\mathrm{d}x_1}\right)^2 - 2A_{12}\frac{\mathrm{d}\psi}{\mathrm{d}x_1} + A_{22} = 0, \tag{2.45}$$

where we used the equality of mixed partial derivatives. This is the equation for the derivative of ψ, which represents the slope of the characteristic curve. Depending on the coefficients, there can be two, one or no real solutions for this slope. This specifies the number of distinct families of characteristic curves.

An example of this approach is illustrated in Exercise 2.8. Physical applications of this approach are illustrated in Section 2.3.

2.2.2 Systems of semilinear equations

In this section, we will use an analogous approach to the one formulated in the preceding section to investigate systems of equations.

Consider a system of m semilinear second-order partial differential equations given by

$$\sum_{k=1}^{m}\sum_{i,j=1}^{n} A_{ijkl}(x)\frac{\partial^2 f_k}{\partial x_i \partial x_j} = B_l\left(x, f, \frac{\partial f}{\partial x}\right),$$

where $l = 1, \ldots, m$ and x stands for $x_1, \ldots, x_n$. We want to find a hypersurface along which it is impossible to determine uniquely the second derivatives from the side conditions that are given along this hypersurface. Let the hypersurface be given by

$$x_i = x_i(s_1, \ldots, s_{n-1}).$$

We express the side conditions as

$$\begin{aligned} f_k(x_1(s_1, \ldots, s_{n-1}), \ldots, x_n(s_1, \ldots, s_{n-1})) &= f_k^0(s_1, \ldots, s_{n-1}), \\ \mathrm{D}_X f_k(x_1(s_1, \ldots, s_{n-1}), \ldots, x_n(s_1, \ldots, s_{n-1})) &= f_k^x(s_1, \ldots, s_{n-1}), \end{aligned}$$

where $k = 1, \ldots, m$. Using the same change of coordinates as discussed in Section 2.2.1 on page 39, we can rewrite these conditions as

$$\tilde{f}_k\left(\tilde{x}_1,\ldots,\tilde{x}_{n-1},0\right)=\tilde{f}_k^0\left(\tilde{x}_1,\ldots,\tilde{x}_{n-1}\right), \tag{2.46}$$

$$\frac{\partial\tilde{f}_k}{\partial\tilde{x}_n}\left(\tilde{x}_1,\ldots,\tilde{x}_{n-1},0\right)=\tilde{f}_k^X\left(\tilde{x}_1,\ldots,\tilde{x}_{n-1}\right), \tag{2.47}$$

where $k=1,\ldots,m$. Differentiating expressions (2.46) with respect to $\tilde{x}_1,\ldots,\tilde{x}_{n-1}$, we obtain

$$\frac{\partial\tilde{f}_k}{\partial\tilde{x}_i}\left(\tilde{x}_1,\ldots,\tilde{x}_{n-1},0\right)=\frac{\partial\tilde{f}_k^0}{\partial\tilde{x}_i}\left(\tilde{x}_1,\ldots,\tilde{x}_{n-1}\right),\quad \text{for } i=1,\ldots,n-1.$$

Differentiating these equations again and differentiating expressions (2.47), we get

$$\frac{\partial^2\tilde{f}_k}{\partial\tilde{x}_j\partial\tilde{x}_i}\left(\tilde{x}_1,\ldots,\tilde{x}_{n-1},0\right)=\frac{\partial^2\tilde{f}_k^0}{\partial\tilde{x}_j\partial\tilde{x}_i}\left(\tilde{x}_1,\ldots,\tilde{x}_{n-1}\right),$$
$$\frac{\partial^2\tilde{f}_k}{\partial\tilde{x}_j\partial\tilde{x}_n}\left(\tilde{x}_1,\ldots,\tilde{x}_{n-1},0\right)=\frac{\partial\tilde{f}_k^X\left(\tilde{x}_1,\ldots,\tilde{x}_{n-1}\right)}{\partial\tilde{x}_j},$$

where $i,j=1,\ldots,n-1$. This is a set of $m\left(n-1\right)^2+m\left(n-1\right)$ equations. However, due to the equality of mixed partial derivatives there are only $mn\left(n-1\right)/2+m\left(n-1\right)$ independent equations for $mn\left(n+1\right)/2$ unknowns. There are still m equations missing to complete the system. These equations are provided by the original system of m differential equations. Now, similarly to the case of a single equation, which was discussed in Section 2.2.1, we can write the complete system for the second derivatives as

$$\begin{bmatrix}
1 & 0 & \ldots & 0 & \ldots & 0 & \ldots & 0\\
0 & 1 & \ldots & 0 & \ldots & 0 & \ldots & 0\\
\vdots & \vdots & \ddots & \vdots & & \vdots & & \vdots\\
0 & 0 & \ldots & 1 & \ldots & 0 & \ldots & 0\\
\vdots & \vdots & & \vdots & \ddots & \vdots & & \vdots\\
\tilde{A}_{1111}\left(y\right) & \tilde{A}_{1211}\left(y\right) & \ldots & \tilde{A}_{ijk1}\left(y\right) & \ldots & \tilde{A}_{nn11}\left(y\right) & \ldots & \tilde{A}_{nnm1}\left(y\right)\\
\vdots & \vdots & & \vdots & & \vdots & \ddots & \vdots\\
\tilde{A}_{11m1}\left(y\right) & \tilde{A}_{12m1}\left(y\right) & \ldots & \tilde{A}_{ijkm}\left(y\right) & \ldots & \tilde{A}_{nn1m}\left(y\right) & \ldots & \tilde{A}_{nnmm}\left(y\right)
\end{bmatrix}
\begin{bmatrix}
\frac{\partial^2\tilde{f}_1}{\partial\tilde{x}_1^2}\left(y\right)\\
\frac{\partial^2\tilde{f}_1}{\partial\tilde{x}_1\partial\tilde{x}_2}\left(y\right)\\
\vdots\\
\frac{\partial^2\tilde{f}_k}{\partial\tilde{x}_j\partial\tilde{x}_i}\left(y\right)\\
\vdots\\
\frac{\partial^2\tilde{f}_1}{\partial\tilde{x}_n\partial\tilde{x}_n}\left(y\right)\\
\vdots\\
\frac{\partial^2\tilde{f}_m}{\partial\tilde{x}_n\partial\tilde{x}_n}\left(y\right)
\end{bmatrix}$$

$$=\begin{bmatrix}
\frac{\partial^2\tilde{f}_1^0}{\partial\tilde{x}_1^2}\left(\tilde{x}_1,\ldots,\tilde{x}_{n-1}\right)\\
\frac{\partial^2\tilde{f}_1^0}{\partial\tilde{x}_1\partial\tilde{x}_2}\left(\tilde{x}_1,\ldots,\tilde{x}_{n-1}\right)\\
\vdots\\
\frac{\partial^2\tilde{f}_k^0}{\partial\tilde{x}_j\partial\tilde{x}_i}\left(\tilde{x}_1,\ldots,\tilde{x}_{n-1}\right)\\
\vdots\\
\tilde{B}_1\left(y,\tilde{f}\left(y\right),\frac{\partial\tilde{f}}{\partial\tilde{x}}\left(y\right)\right)\\
\vdots\\
\tilde{B}_m\left(y,\tilde{f}\left(y\right),\frac{\partial\tilde{f}}{\partial\tilde{x}}\left(y\right)\right)
\end{bmatrix},$$

where y stands for $\left(\tilde{x}_1,\ldots,\tilde{x}_{n-1},0\right)$. The determinant of the coefficient matrix of this system is given by the determinant of the lower-right $m\times m$ block of the matrix, namely,

$$\det\left[\tilde{A}_{nnkl}\left(\tilde{x}_1,\ldots,\tilde{x}_{n-1},0\right)\right],$$

where n is fixed and $k,l=1,2,\ldots,m$. Following the same procedure as in Section 2.2.1 on page 41, we express each entry $\tilde{A}_{nnkl}$ as

$$\begin{aligned}\tilde{A}_{nnkl}\left(\tilde{x}_1,\ldots,\tilde{x}_{n-1},0\right)=&\sum_{i,j=1}^{n-1}A_{ijkl}\left(x_1,\ldots,x_{n-1},\psi\right)\left(\frac{\partial\psi}{\partial x_i}\right)\left(\frac{\partial\psi}{\partial x_j}\right)\\&-2\sum_{i=1}^{n-1}A_{inkl}\left(x_1,\ldots,x_{n-1},\psi\right)\left(\frac{\partial\psi}{\partial x_i}\right)+A_{nnkl}\left(x_1,\ldots,x_{n-1},\psi\right).\end{aligned}$$

Using these expressions for all $k,l=1,\ldots,m$, we express the fact that the determinant is zero as

$$\det\left[\sum_{i,j=1}^{n-1}A_{ijkl}\left(\frac{\partial\psi}{\partial x_i}\right)\left(\frac{\partial\psi}{\partial x_j}\right)-2\sum_{i=1}^{n-1}A_{inkl}\left(\frac{\partial\psi}{\partial x_i}\right)+A_{nnkl}\right]=0, \tag{2.48}$$

where $k,l=1,2,\ldots,m$ and coefficients A_{ijkl} depend on $x_1,\ldots,x_{n-1},\psi$. This determinant is a polynomial of degree m. Thus, in general, we obtain m different hypersurfaces that satisfy this equation.

2.2.3 Quasilinear equations

As we have seen in the preceding sections, the highest-order derivatives play the key role in finding the characteristics. As long as the equation is linear with respect to the highest-order derivatives, we are able to use the methods from the preceding sections.

In view of these remarks, we consider a quasilinear second-order equation, which has the general form given by

$$\sum_{i,j=1}^{n}A_{ij}\left(x,f,\frac{\partial f}{\partial x}\right)\frac{\partial^2 f}{\partial x_i\partial x_j}=B\left(x,f,\frac{\partial f}{\partial x}\right),$$

where x stands for $x_1,\ldots,x_n$ and $\partial f/\partial x$ stands for $\partial f/\partial x_i$ with $i\in\{1,\ldots,n\}$. Both the coefficients in front of the second derivatives and the right-hand side of the equation are, generally, nonlinear functions of x, function f and first derivatives. This equation is a generalization of equation (2.33).

Using the same approach as the one discussed in Section 2.2.1, we consider a special coordinate system $\tilde{x}_1,\ldots,\tilde{x}_n$ in which the side conditions can be expressed by equation (2.39), namely

$$\begin{aligned}\tilde{f}\left(\tilde{x}_1,\ldots,\tilde{x}_{n-1},0\right)&=\tilde{f}_0\left(\tilde{x}_1,\ldots,\tilde{x}_{n-1}\right),\\ \mathrm{D}_{[0,\ldots,0,1]}\tilde{f}\left(\tilde{x}_1,\ldots,\tilde{x}_{n-1},0\right)&=\tilde{f}_X\left(\tilde{x}_1,\ldots,\tilde{x}_{n-1}\right).\end{aligned}$$

In these coordinates, the original equation is

$$\sum_{i\le j}\tilde{A}_{ij}\left(\tilde{x},\tilde{f},\frac{\partial\tilde{f}}{\partial\tilde{x}}\right)\frac{\partial^2\tilde{f}}{\partial\tilde{x}_i\partial\tilde{x}_j}=\tilde{B}\left(\tilde{x},\tilde{f},\frac{\partial\tilde{f}}{\partial\tilde{x}}\right).$$

Following the steps discussed in Section 2.2.1, we obtain the system of algebraic equations for the second derivatives along the hypersurface $x_n = 0$. This system is

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ & & \ddots & \\ \tilde{A}_{11}(y) & \tilde{A}_{12}(y) & \cdots & \tilde{A}_{nn}(y) \end{bmatrix} \begin{bmatrix} \frac{\partial^2 \tilde{f}}{\partial \tilde{x}_1^2} \\ \frac{\partial^2 \tilde{f}}{\partial \tilde{x}_1 \partial \tilde{x}_2} \\ \vdots \\ \frac{\partial^2 \tilde{f}}{\partial \tilde{x}_n^2} \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 \tilde{f}_0}{\partial \tilde{x}_1^2}(\tilde{x}_1, \ldots, \tilde{x}_{n-1}) \\ \frac{\partial^2 \tilde{f}_0}{\partial \tilde{x}_1 \partial \tilde{x}_2}(\tilde{x}_1, \ldots, \tilde{x}_{n-1}) \\ \vdots \\ \tilde{B}(y) \end{bmatrix},$$

where y stands for $\left(\tilde{x}_1, \ldots, \tilde{x}_{n-1}, 0, \tilde{f}_0, \partial \tilde{f}/\partial \tilde{x}\,(\tilde{x}_1, \ldots, \tilde{x}_{n-1}, 0)\right)$, and where the form of the last entry on the right-hand side is due to the fact that we evaluate the original equation along the hypersurface. This system has no unique solution if

$$\tilde{A}_{nn}\left(\tilde{x}_1, \ldots, \tilde{x}_{n-1}, 0, \tilde{f}_0, \frac{\partial \tilde{f}}{\partial \tilde{x}}(\tilde{x}_1, \ldots, \tilde{x}_{n-1}, 0)\right) = 0. \tag{2.49}$$

We wish to express condition (2.49) in the original coordinates. For convenience, we assume that the hypersurface can be expressed by $x_n = \psi(x_1, \ldots, x_{n-1})$. In this case, the new coordinates are given by

$$\begin{aligned} \tilde{x}_i &= x_i, \quad i = 1, \ldots, n-1, \\ \tilde{x}_n &= x_n - \psi(x_1, \ldots, x_{n-1}). \end{aligned}$$

The second derivatives in the original equations are related to the second derivatives in the new coordinates as follows.

$$\sum_{i,j=1}^{n} A_{ij}\left(x, f, \frac{\partial f}{\partial x}\right) \frac{\partial^2}{\partial x_i \partial x_j} = \sum_{i,j=1}^{n} A_{ij}\left(x, f, \frac{\partial f}{\partial x}\right)\left(\frac{\partial^2}{\partial \tilde{x}_j \partial \tilde{x}_i} - \frac{\partial \psi}{\partial x_j}\frac{\partial^2}{\partial \tilde{x}_n \partial \tilde{x}_i} - \frac{\partial \psi}{\partial x_i}\frac{\partial^2}{\partial \tilde{x}_j \partial \tilde{x}_n} + \frac{\partial \psi}{\partial x_j}\frac{\partial \psi}{\partial x_i}\frac{\partial^2}{\partial \tilde{x}_n^2}\right)$$

We want to find the expression for $\tilde{A}_{nn}$. To do so, we have to look for all the terms in the above expression that contain $\partial^2/\partial \tilde{x}_n^2$. These terms are

$$\tilde{A}_{nn}\frac{\partial^2}{\partial \tilde{x}_n^2} = \left[\sum_{i,j=1}^{n-1} A_{ij}\left(\frac{\partial \psi}{\partial x_i}\right)\left(\frac{\partial \psi}{\partial x_j}\right) - 2\sum_{i=1}^{n-1} A_{in}\left(\frac{\partial \psi}{\partial x_i}\right) + A_{nn}\right]\frac{\partial^2}{\partial \tilde{x}_n^2},$$

where, in the first two summations, we used the fact that $\partial \psi/\partial x_i = 0$ for $i = n$. If we evaluate this expression on the hypersurface and set it to zero, we obtain the following equation.

$$\sum_{i,j=1}^{n-1} A_{ij}\left(x_1,\ldots,x_{n-1},\psi,f_0,\frac{\partial f}{\partial x}\left(x_1,\ldots,x_{n-1},\psi\right)\right)\left(\frac{\partial\psi}{\partial x_i}\right)\left(\frac{\partial\psi}{\partial x_j}\right)$$
$$-2\sum_{i=1}^{n-1} A_{in}\left(x_1,\ldots,x_{n-1},\psi,f_0,\frac{\partial f}{\partial x}\left(x_1,\ldots,x_{n-1},\psi\right)\right)\left(\frac{\partial\psi}{\partial x_i}\right)$$
$$+A_{nn}\left(x_1,\ldots,x_{n-1},\psi,f_0,\frac{\partial f}{\partial x}\left(x_1,\ldots,x_{n-1},\psi\right)\right)=0. \quad (2.50)$$

If we take normal vector X to the hypersurface $x_n=\psi\left(x_1,\ldots,x_{n-1}\right)$ to be

$$X=\left[\frac{\partial\psi}{\partial x_1},\ldots,\frac{\partial\psi}{\partial x_{n-1}},-1\right],$$

we can express the normal derivative as

$$\mathrm{D}_X=\left[\frac{\partial\psi}{\partial x_1},\ldots,\frac{\partial\psi}{\partial x_{n-1}},-1\right]\cdot\left[\frac{\partial}{\partial x_1},\ldots,\frac{\partial}{\partial x_n}\right],$$

from where we see that

$$\frac{\partial}{\partial x_n}=\sum_{i=1}^{n-1}\frac{\partial\psi}{\partial x_i}\frac{\partial}{\partial x_i}-\mathrm{D}_X.$$

Thus, we can write the partial derivatives in the arguments of equation (2.50) expressed along the hypersurface as

$$\frac{\partial f}{\partial x_i}\left(x_1,\ldots,x_{n-1},\psi\right)=\frac{\partial f_0}{\partial x_i}\left(x_1,\ldots,x_{n-1}\right),\qquad \text{for } i=1,\ldots,n-1,$$
$$\frac{\partial f}{\partial x_n}\left(x_1,\ldots,x_{n-1},\psi\right)=\sum_{i=1}^{n-1}\frac{\partial\psi}{\partial x_i}\frac{\partial f_0}{\partial x_i}-f_X.$$

Equation (2.50) depends on the side conditions themselves, not only on the hypersurface along which we specify them. We can use this equation to check if the given side conditions allow us to compute the unique second-order derivatives along an a priori given hypersurface. However, we cannot obtain the characteristic surface using this equation.

2.3 Physical applications of semilinear equations

In this section, we will apply the method discussed in Section 2.2.1 to a steady-state problem, diffusion problem and wave-propagation problem. For that purpose, we will study the following three linear equations: Laplace equation, heat equation and wave equation. Also, we will consider a semilinear wave equation.

2.3.1 Laplace equation

Let us consider the Laplace equation in three spatial dimensions, namely,

$$\nabla^2 f\left(x\right):=\frac{\partial^2 f}{\partial x_1^2}+\frac{\partial^2 f}{\partial x_2^2}+\frac{\partial^2 f}{\partial x_3^2}=0. \quad (2.51)$$

We can write this equation in the notation used in expression (2.33) by setting $n = 3$, $A_{ij}(x) = \delta_{ij}$ and $B(x, f, \partial f/\partial x) = 0$; this is a linear equation, which can be viewed as a particular case of a semilinear equation – hence, the above-developed methods apply.

If we choose to parametrize the characteristic surface as

$$x_3 = \psi(x_1, x_2),$$

then, following expression (2.44), we can write the equation for the surface as

$$\left(\frac{\partial \psi}{\partial x_1}\right)^2 + \left(\frac{\partial \psi}{\partial x_2}\right)^2 + 1 = 0.$$

This equation has no real solution. This means that any surface in the three dimensional space is a noncharacteristic surface.

Note that this conclusion is independent of the particular choice of the parametrization. This will become clearer after studying Chapter 6, where we will discuss the characteristics in the context of the Fourier transform. In the context of this transform, we can view equation (2.51) as equation describing a sphere of zero radius.

Since the characteristic surface is determined by the highest derivatives alone, we would have come to the same conclusion if B was nonzero.

2.3.2 Heat equation

We can write the heat equation in three spatial dimensions using the notation of equation (2.33) as follows; $n = 4$, $A_{ij}(x) = \delta_{ij}$, for $i, j = 1, 2, 3$, $A_{44}(x) = 0$, and $B_i(x, f, \partial f/\partial x) = 0$, for $i = 1, 2, 3$, $B_4(x, f, \partial f/\partial x) = 1/k$, with all other coefficients zero. The resulting equation is the linear equation given by

$$\frac{\partial^2 f}{\partial x_1^2} + \frac{\partial^2 f}{\partial x_2^2} + \frac{\partial^2 f}{\partial x_3^2} = \frac{1}{k}\frac{\partial f}{\partial x_4}, \tag{2.52}$$

with k denoting conductivity and x_4 corresponding to time. In this case, the equation for the characteristic surface $x_4 = \psi(x_1, x_2, x_3)$ satisfies

$$\left(\frac{\partial \psi}{\partial x_1}\right)^2 + \left(\frac{\partial \psi}{\partial x_2}\right)^2 + \left(\frac{\partial \psi}{\partial x_3}\right)^2 = 0.$$

This means that

$$\frac{\partial \psi}{\partial x_1} = \frac{\partial \psi}{\partial x_2} = \frac{\partial \psi}{\partial x_3} = 0,$$

and hence, function ψ does not depend on x_1, x_2 or x_3. Thus, we conclude that $\psi(x_1, x_2, x_3)$ is constant, which means that the characteristic hypersurfaces in $x_1x_2x_3x_4$-space for the heat equation are hyperplanes given by x_4 being constant.

This result is consistent with our analysis described in Section 2.1.3. If we specify the temperature along the hyperplane parallel to the time axis, we can no longer specify the temporal change of the temperature.

We can see that in spite of the similarity between the highest-order of derivatives of the heat equation (2.52) and Laplace equation (2.51), the characteristics are different. This results from the fact that the heat equation deals with one more independent variable, x_4.

2.3.3 Wave equation

If we set $n = 3$, $A_{11}(x) = A_{22}(x) = 1$, $A_{33}(x) = -1/v^2$ and we set all the other coefficients to zero in equation (2.33), we obtain the wave equation in two spatial dimensions, namely,

$$\frac{\partial^2 f}{\partial x_1^2} + \frac{\partial^2 f}{\partial x_2^2} = \frac{1}{v^2}\frac{\partial^2 f}{\partial x_3^2}, \tag{2.53}$$

with x_3 corresponding to time. Equation (2.53) is a linear case of a general semilinear equation (2.33); the semilinear equation allows us to consider several extensions of the linear wave equation. If $B(x, f, \partial f/\partial x) = C(x) f$, we can study a linear frequency dispersion. Also, if $B(x, f, \partial f/\partial x) = B_i(x)\,\partial f/\partial x_i$, we can study a linear dissipation. Furthermore, if $B(x, f, \partial f/\partial x) = D(x)$ for all x, we can study wave phenomena associated with the wave source.[2]

Let us assume that the characteristic surface can be expressed as

$$x_3 = \psi(x_1, x_2).$$

Following equation (2.44), we write

$$\left(\frac{\partial \psi}{\partial x_1}\right)^2 + \left(\frac{\partial \psi}{\partial x_2}\right)^2 = \frac{1}{v^2}. \tag{2.54}$$

This is the condition for a surface $\psi(x_1, x_2)$ in the $x_1x_2x_3$-space along which we cannot set side conditions that would allow us to solve uniquely for the second derivatives of $f(x_1, x_2, x_3)$. In other words, equation (2.54) is the characteristic equation of equation (2.53). This equation is called the eikonal equation.

Unlike in the case of characteristics for the Laplace equation, real solutions of the characteristic equation for the wave equation exist and are discussed in Chapter 3.

2.4 Physical applications of systems of semilinear equations

In this section, we will apply the method discussed in Section 2.2.2 to elastodynamics and to electromagnetism.

2.4.1 Elastodynamic equations

Consider the elastodynamic equations, which are derived in Appendix D, namely,

$$\rho(x)\frac{\partial^2 u_i(x,t)}{\partial t^2} = \sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3}\frac{\partial c_{ijkl}(x)}{\partial x_j}\frac{\partial u_k(x,t)}{\partial x_l} + \sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3} c_{ijkl}(x)\frac{\partial^2 u_k(x,t)}{\partial x_j \partial x_l},$$

[2] Readers interested in these extensions might refer to McOwen (2003, pp. 95 – 97).

where $i = 1, 2, 3$.

Writing the general system of linear second-order equations as stated in Section 2.2.2, namely,

$$\sum_{k=1}^{m} \sum_{i,j=1}^{n} A_{ijkl}(x) \frac{\partial^2 f_k}{\partial x_i \partial x_j} = B_l\left(x, f, \frac{\partial f}{\partial x}\right),$$

we can write the elastodynamic equations by setting $m = 3, n = 4$, letting $x_4 = t$ and considering

$$\begin{aligned} f_i(x) &= u_i(x,t), \quad \text{for } i = 1,2,3 \\ A_{ijkl}(x) &= c_{ljki}(x), \quad \text{for } i,j,k,l = 1,2,3 \\ A_{i4kl}(x) &= A_{4ikl}(x) = 0, \quad \text{for } i,k,l = 1,2,3,4 \\ A_{44kl}(x) &= -\rho(x)\delta_{kl}, \quad \text{for } k,l = 1,2,3,4 \\ B_l\left(x, f, \frac{\partial f}{\partial x}\right) &= 0, \quad \text{for } l = 1,2,3,4. \end{aligned}$$

The middle equality results from the fact that there are no mixed partial derivatives containing time. Using equation (2.48), we obtain

$$\det\left[\sum_{i,j=1}^{3} c_{ljki}(x)\left(\frac{\partial \psi}{\partial x_i}\right)\left(\frac{\partial \psi}{\partial x_j}\right) - \rho(x)\delta_{kl}\right] = 0, \tag{2.55}$$

where $k, l = 1, 2, 3$ and $t \equiv x_4 = \psi(x_1, x_2, x_3)$ represents the characteristic surface.

In general, this equation has three distinct solutions, which are the eikonal equations corresponding to the three types of waves that can propagate in an elastic medium.

2.4.2 Maxwell equations

Consider the Maxwell equations in their potential form, which are derived in Appendix F, namely,

$$\begin{aligned} \frac{\partial^2 A_i}{\partial x_1^2} + \frac{\partial^2 A_i}{\partial x_2^2} + \frac{\partial^2 A_i}{\partial x_3^2} - \frac{1}{c^2}\frac{\partial^2 A_i}{\partial t^2} &= -\frac{J_i}{c^2 \epsilon_0}, \text{ for } i = 1,2,3, \\ \frac{\partial^2 \phi}{\partial x_1^2} + \frac{\partial^2 \phi}{\partial x_2^2} + \frac{\partial^2 \phi}{\partial x_3^2} - \frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2} &= -\frac{\rho}{\epsilon_0}. \end{aligned}$$

Writing the general system of linear second-order equations as stated in Section 2.2.2, namely,

$$\sum_{k=1}^{m} \sum_{i,j=1}^{n} A_{ijkl}(x) \frac{\partial^2 f_k}{\partial x_i \partial x_j} = B_l\left(x, f, \frac{\partial f}{\partial x}\right)$$

we can write the Maxwell equations by setting $n, m = 4$, letting $x_4 = t$, and considering

$$\begin{aligned} f_k &= A_k, \quad \text{for } k = 1,2,3, \\ f_4 &= \phi \end{aligned}$$

and

$$
\begin{aligned}
A_{ijkl} &= \delta_{ij}\delta_{kl}, \quad \text{for } i,j,k,l = 1,2,3,\\
A_{44kl} &= -\frac{1}{c^2}\delta_{kl}, \quad \text{for } k,l = 1,\dots,4,\\
B_k &= -\frac{J_k}{c^2\epsilon_0}, \quad \text{for } k = 1,2,3,\\
B_4 &= -\frac{\rho}{\epsilon_0},
\end{aligned}
$$

with all other coefficients set to zero.

Using equation (2.48), we obtain

$$
\det\left[\sum_{i=1}^{3}\delta_{kl}\left(\frac{\partial\psi}{\partial x_i}\right)^2 - \frac{1}{c^2}\delta_{kl}\right] = 0,
$$

where $k,l = 1,2,3,4$. This equation has a quadruple root that is given by

$$
\sum_{i=1}^{3}\left(\frac{\partial\psi}{\partial x_i}\right)^2 = \frac{1}{c^2}, \tag{2.56}
$$

which is the eikonal equation for electromagnetic waves.

Closing remarks

In this chapter we have seen that the characteristic hypersurfaces depend only on the highest-order derivatives of the given semilinear or quasilinear differential equation. In Chapter 6, we will formulate this dependence by introducing the concept of the principal symbol.

In the case of the wave equation in the one spatial dimension we reduced it using the characteristics to a first-order equation. Such a reduction was not possible for the heat and Laplace equations. However, the characteristics allowed us to set properly the Cauchy data in order to solve the equations by the Taylor expansion.

In Chapter 1 characteristics are viewed as curves along which the behaviour of the solutions is determined by the equations; we cannot set side conditions along these curves freely. In this chapter the characteristics are viewed as surfaces along which we cannot uniquely determine the highest derivatives from the equation and side conditions along these surfaces. The construction of characteristics for nonlinear first-order equations – to which we turn our attention in the following chapter – is not clearly related to the side conditions. This is consistent with the fact that the characteristics for linear first and second-order equations do not need to be constructed from side conditions; we can use the directional derivative and the principal symbol, respectively.

Exercises

Exercise 2.1. Find the general solution of equation (2.1), namely,

$$
\frac{\partial^2 f}{\partial x_1^2} + 2\frac{\partial^2 f}{\partial x_1\partial x_2} + \frac{\partial^2 f}{\partial x_2^2} = 0.
$$

Solution 2.1. We can make the following change of variables.

$$y_1 = x_1 - x_2,$$
$$y_2 = x_1 + x_2.$$

We choose to make this change of variables in such a way that one of the new variables distinguishes between the characteristic curves, and the other one is a parameter along these curves. In this particular case, y_1 parametrizes the different characteristic curves, and y_2 is a parameter along these curves.

Using this coordinate transformation, we express the differential operators in the new variables as

$$\frac{\partial}{\partial x_1} = \frac{\partial y_1}{\partial x_1}\frac{\partial}{\partial y_1} + \frac{\partial y_2}{\partial x_1}\frac{\partial}{\partial y_2} = \frac{\partial}{\partial y_1} + \frac{\partial}{\partial y_2},$$
$$\frac{\partial}{\partial x_2} = \frac{\partial y_1}{\partial x_2}\frac{\partial}{\partial y_1} + \frac{\partial y_2}{\partial x_2}\frac{\partial}{\partial y_2} = -\frac{\partial}{\partial y_1} + \frac{\partial}{\partial y_2}.$$

These expressions result in the following three expressions for the second derivatives.

$$\frac{\partial^2}{\partial x_1^2} = \left(\frac{\partial}{\partial y_1} + \frac{\partial}{\partial y_2}\right)\left(\frac{\partial}{\partial y_1} + \frac{\partial}{\partial y_2}\right) = \frac{\partial^2}{\partial y_1^2} + 2\frac{\partial^2}{\partial y_1 \partial y_2} + \frac{\partial^2}{\partial y_2^2},$$
$$2\frac{\partial^2}{\partial x_1 \partial x_2} = 2\left(\frac{\partial}{\partial y_1} + \frac{\partial}{\partial y_2}\right)\left(-\frac{\partial}{\partial y_1} + \frac{\partial}{\partial y_2}\right) = 2\left(-\frac{\partial^2}{\partial y_1^2} + \frac{\partial^2}{\partial y_2^2}\right),$$
$$\frac{\partial^2}{\partial x_2^2} = \left(-\frac{\partial}{\partial y_1} + \frac{\partial}{\partial y_2}\right)\left(-\frac{\partial}{\partial y_1} + \frac{\partial}{\partial y_2}\right) = \frac{\partial^2}{\partial y_1^2} - 2\frac{\partial^2}{\partial y_1 \partial y_2} + \frac{\partial^2}{\partial y_2^2}.$$

Adding the three expressions, we obtain

$$\frac{\partial^2}{\partial x_1^2} + 2\frac{\partial^2}{\partial x_1 \partial x_2} + \frac{\partial^2}{\partial x_2^2} = 4\frac{\partial^2}{\partial y_2^2}.$$

Thus, in the new variables, the original differential equation is

$$\frac{\partial^2 f}{\partial y_2^2} = 0. \tag{2.57}$$

The solution of this equation can be obtained by integrating twice with respect to y_2. The result of such integrations is

$$f(y_1, y_2) = \int\int 0 \mathrm{d}y_2 \mathrm{d}y_2 = \int g(y_1)\,\mathrm{d}y_2 = y_2 g(y_1) + h(y_1).$$

Expressing this result in the original variables, we obtain the general solution of the differential equation, namely,

$$f(x_1, x_2) = (x_1 + x_2)\, g(x_1 - x_2) + h(x_1 - x_2).$$

Exercise 2.2. Follow the method of obtaining the solution for the wave equation in one spatial dimension for the heat equation in one spatial dimension.

Solution 2.2. Using the system of equations describing the behaviour of the solution along the characteristics , namely,

$$\begin{bmatrix} x_1'(s) & x_2'(s) & 0 \\ 0 & x_1'(s) & x_2'(s) \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} f_0'' - \frac{\partial f}{\partial x_1} x_1'' - \frac{\partial f}{\partial x_2} x_2'' \\ f_X' - \frac{\partial f}{\partial x_1} x_2'' + \frac{\partial f}{\partial x_2} x_1'' \\ c \frac{\partial f}{\partial x_2} \end{bmatrix},$$

we obtain .... FINISH!!

Exercise 2.3. Verify that the side conditions along surface $x_3 = \psi(x_1, x_2)$ that are given by

$$f(x_1, x_2, \psi(x_1, x_2)) = g(x_1, x_2) \tag{2.58}$$

and

$$\mathrm{D}_N f(x_1, x_2, \psi(x_1, x_2)) = h(x_1, x_2) \tag{2.59}$$

always allow us to obtain the first derivatives, namely, $\partial f/\partial x_1$, $\partial f/\partial x_2$ and $\partial f/\partial x_3$.

Solution 2.3. Differentiating condition (2.58) with respect to x_1, we get

$$\frac{\partial}{\partial x_1} f(x_1, x_2, \psi(x_1, x_2)) = \frac{\partial f}{\partial x_1} + \frac{\partial f}{\partial x_3} \frac{\partial \psi}{\partial x_1} = \frac{\partial g}{\partial x_1},$$

where we used the fact that the third argument, ψ, can be denoted by x_3. Similarly, differentiating condition (2.58) with respect to x_2, we get

$$\frac{\partial}{\partial x_2} f(x_1, x_2, \psi(x_1, x_2)) = \frac{\partial f}{\partial x_2} + \frac{\partial f}{\partial x_3} \frac{\partial \psi}{\partial x_2} = \frac{\partial g}{\partial x_2}.$$

A vector in the $x_1 x_2 x_3$-space that is normal to a level set of ψ, namely to a surface given by $\psi(x_1, x_2) - x_3 = const.$, is

$$N = \left[\frac{\partial(\psi(x_1, x_2) - x_3)}{\partial x_1}, \frac{\partial(\psi(x_1, x_2) - x_3)}{\partial x_2}, \frac{\partial(\psi(x_1, x_2) - x_3)}{\partial x_3} \right] = \left[\frac{\partial \psi}{\partial x_1}, \frac{\partial \psi}{\partial x_2}, -1 \right].$$

Thus, we can write condition (2.59) as

$$\begin{aligned} N \cdot \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3} \right] &= \left[\frac{\partial \psi}{\partial x_1}, \frac{\partial \psi}{\partial x_2}, -1 \right] \cdot \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3} \right] \\ &= \frac{\partial f}{\partial x_1} \frac{\partial \psi}{\partial x_1} + \frac{\partial f}{\partial x_2} \frac{\partial \psi}{\partial x_2} - \frac{\partial f}{\partial x_3} = h(x_1, x_2). \end{aligned}$$

Now, we have a system of three linear algebraic equations, which we can write as

$$\begin{bmatrix} 1 & 0 & \frac{\partial \psi}{\partial x_1} \\ 0 & 1 & \frac{\partial \psi}{\partial x_2} \\ \frac{\partial \psi}{\partial x_1} & \frac{\partial \psi}{\partial x_2} & -1 \end{bmatrix} \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \frac{\partial f}{\partial x_3} \end{bmatrix} = \begin{bmatrix} \frac{\partial g}{\partial x_1} \\ \frac{\partial g}{\partial x_2} \\ h \end{bmatrix}.$$

To uniquely solve this system for $\partial f/\partial x_1$, $\partial f/\partial x_2$ and $\partial f/\partial x_3$, we require that

$$\det \begin{bmatrix} 1 & 0 & \frac{\partial \psi}{\partial x_1} \\ 0 & 1 & \frac{\partial \psi}{\partial x_2} \\ \frac{\partial \psi}{\partial x_1} & \frac{\partial \psi}{\partial x_2} & -1 \end{bmatrix} \neq 0.$$

In other words, we require that

$$\left(\frac{\partial \psi}{\partial x_1}\right)^2 + \left(\frac{\partial \psi}{\partial x_2}\right)^2 + 1 \neq 0. \tag{2.60}$$

Since expression (2.60) is the squared length of a vector normal to surface $\psi(x_1, x_2) - x_3$, it is never zero. Hence, we can always uniquely solve for the first partial derivatives of f on this surface.

Exercise 2.4. What are the side conditions for equation (2.1) that allow infinitely many solutions?

Solution 2.4. Recall equation (2.15), namely,

$$a_1' + a_2' = 0, \tag{2.61}$$

where a_1 and a_2 are given by expressions (2.6) and (2.7), namely,

$$a_1(s) = \frac{f_0'(s) x_1'(s) + f_N(s) x_2'(s)}{(x_1')^2 + (x_2')^2},$$
$$a_2(s) = \frac{-f_N(s) x_1'(s) + f_0'(s) x_2'(s)}{(x_1')^2 + (x_2')^2}.$$

Since the characteristic curves satisfy $x_1' = x_2'$ and we are interested in the side condition along these curves, we reduce the above expressions;

$$a_1(s) = \frac{f_0'(s) + f_N(s)}{2x_1'},$$
$$a_2(s) = \frac{-f_N(s) + f_0'(s)}{2x_1'}.$$

Differentiating these two equations with respect to s, adding them together and using equation (2.61), we get

$$x_1' f_0'' = x_1'' f_0'.$$

This equation can be rewritten as

$$(\ln|f_0'|)' = (\ln|x_1'|)',$$

which implies that

$$\ln|f_0'| = \ln|x_1'| + C.$$

Exponentiating and using the properties of the absolute value, we get

$$f_0(s) = \pm x_1(s) e^C + l,$$

which can be written as

$$f_0(s) = x_1(s)\, k + l,$$

where k and l are arbitrary constants. This is equation (2.16), as expected.

Exercise 2.5. Set the side conditions for the one-dimensional wave equation written as

$$\frac{\partial^2 f(x,t)}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2 f(x,t)}{\partial t^2}, \tag{2.62}$$

to be the initial conditions giving the displacement f and velocity $\partial f/\partial t$ at time zero.

Solution 2.5. Viewing equation (2.62) as the wave equation, we see that the x-axis and the t-axis correspond to space and time, respectively. We set the hypersurface, which in this case is a line, to coincide with the x-axis. In such a case, the first side condition, namely,

$$f(x,0) = f_0(x),$$

gives the displacement f along this line, which is the displacement at time zero. The second side condition, namely,

$$\mathrm{D}_{[0,1]} f(x,0) = f_X(x),$$

gives a derivative in the direction that is not tangent to the hypersurface. Herein, we set this direction to be parallel to the t-axis. Hence, the second side condition provides the information about the rate of change along the x-axis in the direction of time; this gives the velocity of displacement at time zero. In view of expression (1.34), the second side condition is

$$[0,1] \cdot \left[\frac{\partial f}{\partial x}(x,0), \frac{\partial f}{\partial t}(x,0)\right] = \frac{\partial f}{\partial t}(x,0),$$

which is the velocity at time zero.

Exercise 2.6. Show that equation (2.18), namely,

$$\left(c\frac{\partial}{\partial x_2} + \frac{\partial}{\partial x_1}\right)\left(c\frac{\partial}{\partial x_2} - \frac{\partial}{\partial x_1}\right) f(x_1, x_2) = 0, \tag{2.63}$$

is equivalent to equation (2.17), namely,

$$\frac{\partial^2 f(x_1,x_2)}{\partial x_1^2} = c^2 \frac{\partial^2 f(x_1,x_2)}{\partial x_2^2}. \tag{2.64}$$

Solution 2.6. Using the linearity of differential operators, we can write equation (2.63) as

$$\left(c\frac{\partial}{\partial x_2} + \frac{\partial}{\partial x_1}\right)\left(c\frac{\partial}{\partial x_2} - \frac{\partial}{\partial x_1}\right) f = \left(c\frac{\partial}{\partial x_2} + \frac{\partial}{\partial x_1}\right)\left(c\frac{\partial f}{\partial x_2} - \frac{\partial f}{\partial x_1}\right) = 0.$$

Multiplying through, we get

$$c^2\frac{\partial^2 f}{\partial x_2^2} - c\frac{\partial^2 f}{\partial x_2 \partial x_1} + c\frac{\partial^2 f}{\partial x_1 \partial x_2} - \frac{\partial^2 f}{\partial x_1^2} = 0.$$

Since the two middle terms vanish due to the equality of mixed partial derivatives, we obtain

$$c^2 \frac{\partial^2 f}{\partial x_2^2} - \frac{\partial^2 f}{\partial x_1^2} = 0,$$

which is equation (2.64), as required.

Exercise 2.7. Find the general solution of the following equation using the method of characteristics.

$$\left(\frac{\partial^2}{\partial x_1^2} + 5 \frac{\partial^2}{\partial x_1 \partial x_2} \right) f(x_1, x_2) = 1.$$

Solution 2.7. Following the method of characteristics, we can write for the side-conditions along curve $[x_1(s), x_2(s)]$ the following system.

$$\begin{bmatrix} x_1'(s) & x_2'(s) & 0 \\ 0 & x_1'(s) & x_2'(s) \\ 1 & 5 & 0 \end{bmatrix} \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} a_1'(s) \\ a_2'(s) \\ 0 \end{bmatrix},$$

where

$$\frac{\partial f}{\partial x_1}(x_1(s), x_2(s)) = a_1(s)$$
$$\frac{\partial f}{\partial x_2}(x_1(s), x_2(s)) = a_2(s).$$

The determinant of the above matrix is zero if

$$(x_2'(s))^2 = 5 x_1'(s) x_2'(s).$$

The solutions of this equation are given by

$$x_2'(s) = 0,$$

or

$$x_2'(s) = 5 x_1'(s).$$

Hence, along the characteristic curves,

$$x_2(s) = C$$

or

$$x_2(s) = 5 x_1(s) + D,$$

where either of the values of C or D determines the choice of a particular characteristic curve.

Along these curves, the original differential equation can be written as

$$\frac{\partial^2}{\partial y_1 \partial y_2} f(x_1(y_1, y_2), x_2(y_1, y_2)) = 1, \tag{2.65}$$

where $y_1 = x_1 - (1/5) x_2$ and $y_2 = (1/5) x_2$. The general solution of equation (2.65) is

$$f\left(x_1\left(y_1,y_2\right),x_2\left(y_1,y_2\right)\right)=y_1y_2+g\left(y_1\right)+h\left(y_2\right),$$

for some functions g and h. The solution can be written in the original coordinates as

$$f\left(x_1,x_2\right)=\left(x_1-\frac{1}{5}x_2\right)\frac{1}{5}x_2+g\left(x_1-\frac{1}{5}x_2\right)+h\left(\frac{1}{5}x_2\right).$$

Exercise 2.8. Find the characteristic curves of

$$4\frac{\partial^2 f}{\partial x_1^2}+16x_1\frac{\partial^2 f}{\partial x_1\partial x_2}+7x_1^2\frac{\partial^2 f}{\partial x_2^2}=4\sin\left(x_2\right).$$

Solution 2.8. Following equation (2.45) and using the fact that $A_{12}+A_{21}=16x_1$, we write

$$4\left(\frac{\mathrm{d}x_2}{\mathrm{d}x_1}\right)^2-16x_1\frac{\mathrm{d}x_2}{\mathrm{d}x_1}+7x_1^2=0,$$

which is the characteristic equation. Solving for $\mathrm{d}x_2/\mathrm{d}x_1$, we obtain

$$\frac{\mathrm{d}x_2}{\mathrm{d}x_1}=\begin{cases}\frac{7}{2}x_1\\ \frac{1}{2}x_1\end{cases}.$$

The solutions of these two equations are

$$x_2=\frac{7}{4}x_1^2+C$$

and

$$x_2=\frac{1}{4}x_1^2+D.$$

3

Characteristic equations of first-order nonlinear partial differential equations

Preliminary remarks

In this chapter we derive the characteristic equations for nonlinear first-order partial differential equations. The importance of this formulation in the context of physics results from the fact that the eikonal equations derived in Sections 2.4.1 and 2.4.2 are nonlinear first-order partial differential equations.

This chapter begins with the derivation of the characteristic equations for first-order nonlinear partial differential equations. Subsequently, we exemplify the characteristic equation in contexts of elasticity and electromagnetism.

3.1 General formulation

Let us consider the general form of an n-dimensional nonlinear first-order partial differential equation for function ψ.

$$F\left(x_1, x_2, \ldots, x_n, \psi, \frac{\partial \psi}{\partial x_1}, \frac{\partial \psi}{\partial x_2}, \ldots, \frac{\partial \psi}{\partial x_n}\right) = 0, \tag{3.1}$$

where x_i are the n independent variables. The graph of a solution of such an equation is a hypersurface, $w = \psi(x_1, x_2, \ldots, x_n)$, in the $(n+1)$-dimensional xw-space of vectors $[x_1, \ldots, x_n, w]$. Letting $p_i := \partial\psi/\partial x_i$, we can rewrite the differential equation as a level-surface of a function of $2n+1$ variables given by

$$F(x_1, x_2, \ldots, x_n, w, p_1, p_2, \ldots, p_n) = 0. \tag{3.2}$$

We can use the fact that the differential equation provides, at least implicitly, a relation among the p_i. Specifically, at a fixed point $P = \left[x_1^P, \ldots, x_n^P, w^P\right]$, we can write

$$F\left(x_1^P, \ldots, x_n^P, w^P, p_1(\xi_1, \xi_2, \ldots, \xi_{n-1}), \ldots, p_n(\xi_1, \xi_2, \ldots, \xi_{n-1})\right) = 0, \tag{3.3}$$

where ξ_j are $n-1$ parameters relating the p_i. We wish to study the graph of a solution, $w = \psi(x_1, x_2, \ldots, x_n)$. It can be obtained as the envelope of its tangent hyperplanes in the xw-space.

A hyperplane is defined by a point through which it passes and by its normal vector. To see the possible tangent hyperplane to $\psi(x)$ at point P, we consider the normal vector to the graph of ψ in the xw-space, which is given by

$$\left[\left.\frac{\partial\psi}{\partial x_1}\right|_P, \left.\frac{\partial\psi}{\partial x_2}\right|_P, \ldots, \left.\frac{\partial\psi}{\partial x_n}\right|_P, -1\right] \equiv \left[p_1(\xi_1,\ldots,\xi_{n-1}), \ldots, p_n(\xi_1,\ldots,\xi_{n-1}), -1\right],$$

since $p := \nabla\psi$. The equation of this hyperplane is

$$\begin{aligned} 0 &= \left[\left.\frac{\partial\psi}{\partial x_1}\right|_P, \left.\frac{\partial\psi}{\partial x_2}\right|_P, \ldots, \left.\frac{\partial\psi}{\partial x_n}\right|_P, -1\right] \cdot \left[\left(x_1 - x_1^P\right), \left(x_2 - x_2^P\right), \ldots, \left(x_n - x_n^P\right), \left(w - w^P\right)\right] \\ &= \left.\frac{\partial\psi}{\partial x_1}\right|_P \left(x_1 - x_1^P\right) + \left.\frac{\partial\psi}{\partial x_2}\right|_P \left(x_2 - x_2^P\right) + \cdots + \left.\frac{\partial\psi}{\partial x_n}\right|_P \left(x_n - x_n^P\right) - \left(w - w^P\right) \\ &= p_1(\xi_1,\ldots,\xi_{n-1})\left(x_1 - x_1^P\right) + \cdots + p_n(\xi_1,\ldots,\xi_{n-1})\left(x_n - x_n^P\right) - \left(w - w^P\right). \end{aligned} \tag{3.4}$$

As ξ_j vary, this equation describes an $(n-1)$-parameter family of hyperplanes through point P that contains the hyperplane tangent to the graph of a solution. The envelope of these hyperplanes is referred to as the Monge cone, in honor of Gaspard Monge (1746–1818). Note that this is truly a cone only in a two-dimensional case where the family of hyperplanes is a one-parameter family of planes, as illustrated in Figure 3.1.

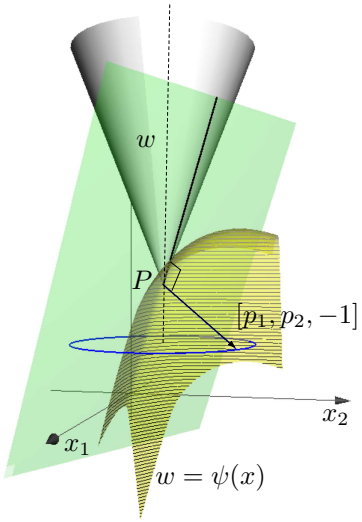


Fig. 3.1. The Monge cone at point $P = \left[x_1^P, x_2^P, w^P\right]$ constructed as an envelope of hyperplanes with normal vectors of the form $[p_1, p_2, -1]$, where p satisfies $F\left(x^P, p, w^P\right) = 0$. We show one of these hyperplanes that is tangent to both the cone and the graph of $\psi(x_1, x_2)$.

As a result of our work, we want to reduce the partial differential equation to a set of ordinary differential equations, which is analogous to studying a vector field. Consequently, we wish to study one particular line of the hyperplane given by equation (3.4). A natural choice of such line is the intersection of the given hyperplane with the envelope of all such hyperplanes; in other words, the intersection of this hyperplane with the Monge cone. To find this envelope,

we consider the set of points in the hyperplane where the given equation of the hyperplane is stationary with respect to the ξ_j. Thus, we are looking for points of the hyperplanes that remain stationary if we change parameters ξ_j infinitesimally.

To visualize this process, let us consider the following example, which can be illustrated in two dimensions.

Example 3.1. Consider equation

$$\frac{1}{\xi}x_1 + \xi x_2 - 1 = 0, \tag{3.5}$$

which represents a one-parameter family of lines in the x_1x_2-plane parametrized by ξ. We wish to find a particular point that is representative of each line and is also parametrized by ξ. A natural choice of such a point is the intersection of the given line with envelope of all such lines in the x_1x_2-plane. To find this envelope, which results from varying ξ, we consider the $x_1x_2\xi$-space. In other words, we consider the left-hand side of equation (3.5) as a function of three variables – x_1, x_2 and ξ – in a three-dimensional space; we denote this function by $f(x_1, x_2, \xi)$. We use the fact that the vectors normal to $f = 0$ at the points that project to the envelope on the x_1x_2-plane have vanishing ξ-component. It means that these vectors are parallel to the x_1x_2-plane. In view of equation (3.5) , this means that – in the $x_1x_2\xi$-space – the ξ-component of the gradient that is given by

$$\frac{\partial}{\partial \xi}\left(\frac{1}{\xi}x_1 + \xi x_2 - 1\right)$$

vanishes.

This means that the directional derivative of f in the direction of ξ is zero at the points that project on the envelope. Geometrically we can justify this by realizing that a point on the envelope belongs to the given tangent line and the lines that are infinitesimally close, as measured by ξ, to this line.

Returning to the vanishing component of the gradient, we get

$$-\frac{1}{\xi^2}x_1 + x_2 = 0, \tag{3.6}$$

which is another family of lines parametrized by the same parameter ξ. The point on the envelope parametrized by ξ belongs to this family of lines as well as to the original family of lines given by equation (3.5). To find this point we combine equations (3.5) and (3.6). In other words, for each particular value of ξ, equations (3.5) and (3.6) considered simultaneously represent a point in the x_1x_2-plane.

Herein, solving equations (3.5) and (3.6) for x_1 and x_2, we obtain equations for the envelope parametrized by ξ, namely,

$$\begin{aligned} x_1 &= \frac{1}{2}\xi, \\ x_2 &= \frac{1}{2\xi}. \end{aligned}$$

In this case, the envelope is a part of a hyperbola.

Having completed the example, let us return to our main discussion and consider the family of hyperplanes given by equation (3.4). To make the stationarity condition, let us take partial

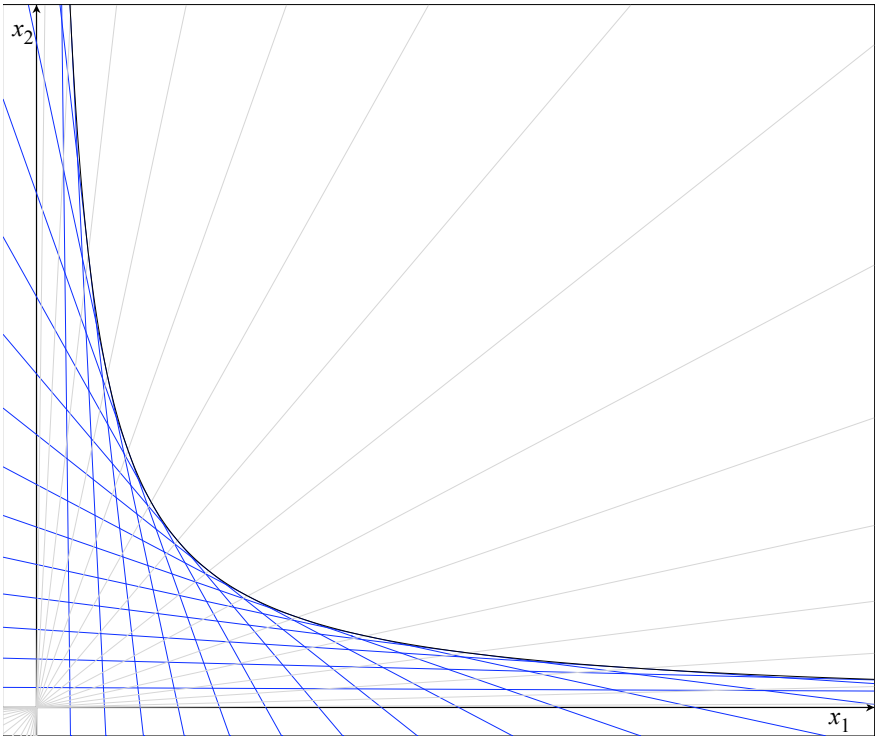


Fig. 3.2. A family of lines $\frac{1}{\xi}x_1 + \xi x_2 - 1 = 0$ and its envelope constructed by intersecting it with the lines $-\frac{1}{\xi^2}x_1 + x_2 = 0$ obtained by differentiating it with respect to parameter ξ.

derivatives of equation (3.4) with respect to the ξ_j to obtain

$$\frac{\partial p_1}{\partial \xi_j}\left(x_1 - x_1^P\right) + \frac{\partial p_2}{\partial \xi_j}\left(x_2 - x_2^P\right) + \cdots + \frac{\partial p_n}{\partial \xi_j}\left(x_n - x_n^P\right) = 0,$$

which are $n-1$ equations of hyperplanes passing through point P and have normal vectors given by $[\partial p_1/\partial \xi_j, \ldots, \partial p_n/\partial \xi_j, 0]$.

Herein, the intersection of these $n-1$ hyperplanes is a two-dimensional plane. Since these vectors are tangent to the surface given by $[p(\xi), -1]$– and hence the corresponding hyperplanes are orthogonal to one another – we can intersect the two-dimensional plane with the hyperplane given by equation (3.4), which we can rewrite as

$$p_1(\xi_1, \xi_2)\left(x_1 - x_1^P\right) + p_2(\xi_1, \xi_2)\left(x_2 - x_2^P\right) + p_3(\xi_1, \xi_2)\left(x_3 - x_3^P\right) = w - w^P,$$

to obtain a line.

Above, we have constructed a one-parameter family of lines at point P. Each line from this family belongs to one of the possible tangent planes to the graph of a solution of the original differential equation. We are going to find a curve in this graph whose tangent vector at a given point belongs to the one-parameter family of these lines at that point. If we parametrize this curve by a parameter s, we write the tangent vector to this curve, $[x_1(s), \ldots, x_n(s), w(s)]$, as $[\mathrm{d}x_1(s)/\mathrm{d}s, \ldots, \mathrm{d}x_n(s)/\mathrm{d}s, \mathrm{d}w(s)/\mathrm{d}s]$. Since this vector is in the intersection of the n hyperplanes, its components must satisfy

$$\frac{\partial p_1}{\partial \xi_j}\frac{\mathrm{d}x_1}{\mathrm{d}s} + \frac{\partial p_2}{\partial \xi_j}\frac{\mathrm{d}x_2}{\mathrm{d}s} + \cdots + \frac{\partial p_n}{\partial \xi_j}\frac{\mathrm{d}x_n}{\mathrm{d}s} = 0, \tag{3.7}$$

and

$$p_1\frac{\mathrm{d}x_1}{\mathrm{d}s} + p_2\frac{\mathrm{d}x_2}{\mathrm{d}s} + \cdots + p_n\frac{\mathrm{d}x_n}{\mathrm{d}s} = \frac{\mathrm{d}w}{\mathrm{d}s}.$$

To gain a geometrical insight, we can rewrite the n equations above in terms of scalar products as

$$\left[\frac{\partial p_1}{\partial \xi_j}, \frac{\partial p_2}{\partial \xi_j}, \ldots, \frac{\partial p_n}{\partial \xi_j}, 0\right] \cdot \left[\frac{\mathrm{d}x_1}{\mathrm{d}s}, \frac{\mathrm{d}x_2}{\mathrm{d}s}, \ldots, \frac{\mathrm{d}x_n}{\mathrm{d}s}, \frac{\mathrm{d}w}{\mathrm{d}s}\right] = 0, \tag{3.8}$$

and

$$[p_1, p_2, \ldots, p_n, -1] \cdot \left[\frac{\mathrm{d}x_1}{\mathrm{d}s}, \frac{\mathrm{d}x_2}{\mathrm{d}s}, \ldots, \frac{\mathrm{d}x_n}{\mathrm{d}s}, \frac{\mathrm{d}w}{\mathrm{d}s}\right] = 0, \tag{3.9}$$

respectively. We see that these n orthogonality conditions in the $(n+1)$-dimensional space determine vector $[\mathrm{d}x_1/\mathrm{d}s, \ldots, \mathrm{d}x_n/\mathrm{d}s, \mathrm{d}w/\mathrm{d}s]$ up to a scalar multiple.

To express the orthogonality conditions in terms of the original differential equation given by function F stated in expression (3.3), we differentiate equation (3.3) with respect to ξ_j to get

$$\frac{\partial F}{\partial p_1}\frac{\partial p_1}{\partial \xi_j} + \frac{\partial F}{\partial p_2}\frac{\partial p_2}{\partial \xi_j} + \cdots + \frac{\partial F}{\partial p_n}\frac{\partial p_n}{\partial \xi_j} = 0.$$

We can rewrite the above equations as

$$\left[\frac{\partial F}{\partial p_1}, \frac{\partial F}{\partial p_2}, \ldots, \frac{\partial F}{\partial p_n}, A\right] \cdot \left[\frac{\partial p_1}{\partial \xi_j}, \frac{\partial p_2}{\partial \xi_j}, \ldots, \frac{\partial p_n}{\partial \xi_j}, 0\right] = 0, \tag{3.10}$$

where component A is an arbitrary number. We can take advantage of this arbitrariness by letting

$$A = p_1\frac{\partial F}{\partial p_1} + p_2\frac{\partial F}{\partial p_2} + \cdots + p_n\frac{\partial F}{\partial p_n}. \tag{3.11}$$

This can be written as

$$\left[\frac{\partial F}{\partial p_1}, \frac{\partial F}{\partial p_2}, \ldots, \frac{\partial F}{\partial p_n}, A\right] \cdot [p_1, p_2, \ldots, p_n, -1] = 0. \tag{3.12}$$

Comparing equations (3.8) and (3.9) with equations (3.10) and (3.12), we see that vector

$$\left[\frac{\mathrm{d}x_1}{\mathrm{d}s}, \frac{\mathrm{d}x_2}{\mathrm{d}s}, \ldots, \frac{\mathrm{d}x_n}{\mathrm{d}s}, \frac{\mathrm{dw}}{\mathrm{d}s}\right]$$

is orthogonal to the same n vectors as vector

$$\left[\frac{\partial F}{\partial p_1}, \frac{\partial F}{\partial p_2}, \ldots, \frac{\partial F}{\partial p_n}, A\right].$$

Since these are $(n+1)$-dimensional vectors, the fact that they are perpendicular to the same n linearly independent vectors implies that these two vectors are parallel to one another. We can express this as

$$\left[\frac{\mathrm{d}x_1}{\mathrm{d}s}, \frac{\mathrm{d}x_2}{\mathrm{d}s}, \ldots, \frac{\mathrm{d}x_n}{\mathrm{d}s}, \frac{\mathrm{dw}}{\mathrm{d}s}\right] = \zeta\left[\frac{\partial F}{\partial p_1}, \frac{\partial F}{\partial p_2}, \ldots, \frac{\partial F}{\partial p_n}, p_1\frac{\partial F}{\partial p_1} + p_2\frac{\partial F}{\partial p_2} + \cdots + p_n\frac{\partial F}{\partial p_n}\right],$$

where we used expression (3.11) for A. The coefficient of proportionality, ζ, depends on the choice of point P. Hence, ζ is a function of x and w. This equality can be written as

$$\frac{\mathrm{d}x_i}{\mathrm{d}s} = \zeta\left(x, w\right)\frac{\partial F}{\partial p_i}, \tag{3.13}$$

and

$$\frac{\mathrm{d}w}{\mathrm{d}s} = \zeta\left(x, w\right)\left(p_1\frac{\partial F}{\partial p_1} + p_2\frac{\partial F}{\partial p_2} + \cdots + p_n\frac{\partial F}{\partial p_n}\right). \tag{3.14}$$

The form of ζ can be determined for a given problem by the choice of parameter s, which we commonly choose to be time. Equations (3.13) belong to the characteristic equations of equation (3.1).

Since these equations depend on the p_i, we need to know how the p_i change along the curve $[x\left(s\right), w\left(s\right)]$ in order to establish a determined system of equations. To obtain this information, we differentiate the original differential equation with respect to x_i. Using the chain rule, we can write

$$\frac{\partial}{\partial x_i}F\left(x, \psi\left(x\right), p_1\left(x\right), \ldots, p_n\left(x\right)\right) = \frac{\partial F}{\partial x_i} + \frac{\partial F}{\partial w}\frac{\partial \psi}{\partial x_i} + \sum_{j=1}^{n}\frac{\partial F}{\partial p_j}\frac{\partial p_j}{\partial x_i} = 0. \tag{3.15}$$

We can modify this equation in the following way. Since $p_i := \partial\psi/\partial x_i$, we can write

$$\frac{\partial p_j}{\partial x_i} = \frac{\partial^2 \psi}{\partial x_i \partial x_j}.$$

Invoking the equality of mixed partial derivatives, we see that

$$\frac{\partial p_j}{\partial x_i} = \frac{\partial p_i}{\partial x_j},$$

where we could explicitly write the right-hand side as $\partial p_i/\partial x_j = \partial^2\psi/\partial x_j\partial x_i$. Now, we can write equation (3.15) as

$$\frac{\partial F}{\partial x_i} + \frac{\partial F}{\partial w}\frac{\partial \psi}{\partial x_i} + \sum_{j=1}^{n}\frac{\partial F}{\partial p_j}\frac{\partial p_i}{\partial x_j} = 0.$$

Using characteristic equations (3.13), we can rewrite this equation as

$$\frac{\partial F}{\partial x_i} + \frac{\partial F}{\partial w}\frac{\partial \psi}{\partial x_i} + \frac{1}{\zeta}\sum_{j=1}^{n}\frac{\mathrm{d}x_j}{\mathrm{d}s}\frac{\partial p_i}{\partial x_j} = 0. \tag{3.16}$$

In view of the chain rule,

$$\sum_{j=1}^{n}\frac{\mathrm{d}x_j}{\mathrm{d}s}\frac{\partial p_i}{\partial x_j} = \frac{\mathrm{d}p_i\left(x\right)}{\mathrm{d}s},$$

we rewrite equation (3.16) as

$$\frac{\partial F}{\partial x_i} + \frac{\partial F}{\partial w}\frac{\partial \psi}{\partial x_i} + \frac{1}{\zeta}\frac{\mathrm{d}p_i}{\mathrm{d}s} = 0,$$

which can be restated as

$$\frac{\mathrm{d}p_i}{\mathrm{d}s} = -\zeta\left(\frac{\partial F}{\partial x_i} + \frac{\partial F}{\partial w}p_i\right). \tag{3.17}$$

These equations complete the system of $2n + 1$ equations for the same number of unknowns.

Note that in order to construct the graph of a solution we need only the curves given by $(x_1(s), \dots, x_n(s), w(s))$. However, to find these curves, we also need functions $p_j(s)$. Let us interpret geometrically this solution of the system of linear equations. Herein, we are dealing with an infinitesimal strip rather than with a curve. The curve itself is given by $(x_1(s), \dots, x_n(s), w(s))$ – and is referred to as a base or projected characteristic – while vector $[-p_1(s), \dots, -p_n(s), 1]$ is normal to the infinitesimal pieces of the tangent hyperplanes along this curve.

To summarize, we restate characteristic equations (3.13) and (3.14):

$$\frac{\mathrm{d}x_i}{\mathrm{d}s} = \zeta(x, w) \frac{\partial F}{\partial p_i} \tag{3.18}$$

and

$$\frac{\mathrm{d}w}{\mathrm{d}s} = \zeta(x, w) \sum_{j=1}^{n} p_j \frac{\partial F}{\partial p_j} \tag{3.19}$$

and equations (3.17):

$$\frac{\mathrm{d}p_i}{\mathrm{d}s} = -\zeta(x, w) \left(\frac{\partial F}{\partial x_i} + \frac{\partial F}{\partial w} p_i \right), \tag{3.20}$$

where $w = \psi(x_1, x_2, \dots, x_n)$ and $p_i = \partial\psi/\partial x_i$. These equations are known as the Hamilton equations. They are characteristic equations for a nonlinear first-order partial differential equation. The solutions of this system are characteristic curves. As shown in Exercise 3.1, the Hamilton equation applied to a first-order linear equation result in characteristics obtained in Section 1.1.2.

3.2 Side conditions

3.3 Physical applications

3.3.1 Elasticity

In Section 2.4.1 we derived the characteristic equation for the elastodynamic equations, namely equation (2.55):

$$\det \left[\sum_{i,j=1}^{3} c_{ljki}(x) \left(\frac{\partial \psi}{\partial x_i} \right) \left(\frac{\partial \psi}{\partial x_j} \right) - \rho(x) \delta_{kl} \right] = 0.$$

This is an algebraic equation for $p \equiv \nabla\psi$, which has in general three real roots. We denote these roots by v_i^{-2} where $i \in \{1, 2, 3\}$ corresponds to different roots. Subsequently, we write

$$p^2 = \frac{1}{v_i^2}, \tag{3.21}$$

which are the eikonal equations – the characteristic equations of the elastodynamic equations. As discussed in Section 3.1, we can solve the eikonal equation using equations (3.18), (3.20) and (3.19). To do so, following expression (3.21), we set $F = p^2 v^2$. Since F does not depend on w, the resulting equations are

$$\frac{\mathrm{d}x_i}{\mathrm{d}s} = \zeta \frac{\partial F}{\partial p_i},$$

$$\frac{\mathrm{d}p_i}{\mathrm{d}s} = -\zeta \frac{\partial F}{\partial x_i}$$

and

$$\frac{\mathrm{d}w}{\mathrm{d}s} = \zeta \left(p_1 \frac{\partial F}{\partial p_1} + p_2 \frac{\partial F}{\partial p_2} + \cdots + p_n \frac{\partial F}{\partial p_n} \right). \tag{3.22}$$

To facilitate a physical interpretation of equations and solutions, we want to parametrize the characteristics by time. Since $w = \psi$ corresponds to time, we require $\mathrm{d}w/\mathrm{d}s = 1$. Since $F = p^2 v^2$ is a homogeneous function of degree two in p, the term in parentheses in equation (3.22) is equal to $2F$. In view of expression (3.21), $F = 1$. Thus, we see that $\zeta = 1/2$. It is common to denote $F/2$ by H, which is called the ray-theory Hamiltonian; in such a case we write

$$\begin{aligned} \dot{x_i} &= \frac{\partial H}{\partial p_i} \\ & \qquad\qquad\qquad\qquad i \in \{1, 2, 3\}, \\ \dot{p_i} &= -\frac{\partial H}{\partial x_i} \end{aligned} \tag{3.23}$$

which are the Hamilton ray equations, and where $\mathrm{d}/\mathrm{d}t$ is denoted by the dot above a symbol.

In the following example we illustrate the use of equations (3.23) in a seismological context.

Example 3.2. [1]In this example, we study the Hamilton ray equations for a particular case of a wave that exhibits an elliptical velocity dependence with direction and a linear velocity dependence along one axis. This assumption as well as a consideration of a two-dimensional continuum allow us to illustrate the meaning of the Hamilton ray equations by obtaining analytic solutions.

The present example is inspired by the velocity models used in applied seismology. Since the depth is commonly associated with the z-axis, we will consider herein the physical space spanned by the x-axis and the z-axis, which correspond to the horizontal distance and depth, respectively.

Considering the xz-plane, we can write the wavefront velocity of a wave subjected to the elliptical anisotropy as

$$v(\vartheta) = \sqrt{v_x^2 \sin^2 \vartheta + v_z^2 \cos^2 \vartheta},$$

where v_x and v_z are the magnitudes of wavefront velocity along the x-axis and z-axis, respectively, and ϑ is the wavefront angle measured from the z-axis. We would also like to subject the wave to the linear increase of velocity along the z-axis. Since in this example, we wish to obtain analytic solutions of the Hamilton equations (3.18) and (3.20), we assume that the ratio of the magnitudes of wavefront velocity along the x-axis and z-axis is constant. For this purpose, we define the parameter given by

$$c := \frac{v_x}{v_z},$$

where c is a dimensionless quantity that is equal to unity in the isotropic case. Using this definition, we can rewrite the expression for the wavefront velocity as

$$v(\vartheta) = v_z \sqrt{c^2 \sin^2 \vartheta + \cos^2 \vartheta}.$$

[1]The example discussed in this section appeared in Rogister and Slawinski (2005).

Now, we can conveniently specify the linear dependence of velocity along the z-axis to write

$$v(\vartheta, z) = (a + bz)\sqrt{c^2 \sin^2 \vartheta + \cos^2 \vartheta}, \tag{3.24}$$

where a and b are constants whose units are the units of velocity and of the reciprocal of time, respectively. We refer to the velocity model described by this expression as the abc model.

In view of equation (3.21), the eikonal equation for two spatial dimensions is

$$p^2 := p_x^2 + p_z^2 = \frac{1}{v^2}. \tag{3.25}$$

Hence, in view of equation (3.24), the eikonal equation for the abc model is

$$p_x^2 + p_z^2 = \frac{1}{(a + bz)^2 \left[c^2 \sin^2 \vartheta + \cos^2 \vartheta\right]}.$$

To express ϑ in terms of vector $p = [p_x, p_z]$, we can write

$$\vartheta = \arctan \frac{p_x}{p_z}. \tag{3.26}$$

Inserting the expression for ϑ into the above equation and using trigonometric identities, we get

$$p_x^2 + p_z^2 = \frac{p_x^2 + p_z^2}{(a + bz)^2 \left[c^2 p_x^2 + p_z^2\right]}.$$

Simplifying, we obtain

$$(a + bz)^2 \left[c^2 p_x^2 + p_z^2\right] = 1, \tag{3.27}$$

where $p_x := \partial\psi/\partial x$ and $p_z := \partial\psi/\partial z$. This is the eikonal equation that corresponds to elliptical anisotropy and linear inhomogeneity, and whose solution is the eikonal function, ψ, with its level curves corresponding to wavefronts. Equation (3.27) can be viewed as an expression for level set of the inverse $H^{-1}(1/2)$ of Hamiltonian

$$H \equiv \frac{(a + bz)^2 \left[c^2 p_x^2 + p_z^2\right]}{2}. \tag{3.28}$$

Considering system (3.23) in two dimensions, we let $x := [x_1, x_2] \equiv [x, z]$ and $p := [p_1, p_2] \equiv [p_x, p_z]$. Using expression (3.28), we can explicitly write this system as

$$\begin{aligned}
\dot{x} &= \frac{\partial H(z, p_x, p_z)}{\partial p_x} = (a + bz)^2 c^2 p_x, \\
\dot{z} &= \frac{\partial H(z, p_x, p_z)}{\partial p_z} = (a + bz)^2 p_z \\
\dot{p}_x &= -\frac{\partial H(z, p_x, p_z)}{\partial x} = 0 \\
\dot{p}_z &= -\frac{\partial H(z, p_x, p_z)}{\partial z} = -b(a + bz)\left[c^2 p_x^2 + p_z^2\right]
\end{aligned} \tag{3.29}$$

Since we can write equation (3.27) as

$$c^2 p_x^2 + p_z^2 = \frac{1}{(a+bz)^2},$$

we can rewrite the last Hamilton ray equation as

$$\dot{p}_z = -\frac{b}{a+bz}.$$

Since $\dot{p}_x = 0$, it immediately follows that $p_x(t) = \mathfrak{p}$, where $\mathfrak{p}$ denotes a constant. Now, we can write the remaining three Hamilton ray equations as a system of ordinary differential equations to be solved for x, z and p_z. These equations are

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = [a+bz(t)]^2 c^2 \mathfrak{p}, \tag{3.30}$$

$$\frac{\mathrm{d}z(t)}{\mathrm{d}t} = [a+bz(t)]^2 p_z(t) \tag{3.31}$$

and

$$\frac{\mathrm{d}p_z(t)}{\mathrm{d}t} = -\frac{b}{a+bz(t)}. \tag{3.32}$$

To complete this system of differential equations, we need additional constraints. We choose to use the initial conditions, which correspond to the values of the unknowns at the initial time. In other words, we need $x(t)$, $z(t)$, $p_x(t)$ and $p_z(t)$ at $t = 0$. While we already have $p_x(0) = \mathfrak{p}$, let us set $x(0) = 0$ and $z(0) = 0$ and $p_z(0) = p_{z0}$. The initial condition for $p_z(t)$ is not independent from the initial condition for $p_x(t)$. They are related by eikonal equation (3.27). Solving this equation for p_z at $t = 0$, which corresponds to $z = 0$, we get

$$p_z(0) = \sqrt{\frac{1}{a^2} - c^2 [p_x(0)]^2}.$$

Since $p_x(0) = \mathfrak{p}$, the initial condition for p_z that obeys the eikonal equation is

$$p_z(0) = \sqrt{\frac{1}{a^2} - c^2 \mathfrak{p}^2}. \tag{3.33}$$

The above system of differential equations accompanied by the initial conditions has a clear meaning in the context of ray theory.

First of all, let us take a look at eikonal equation (3.27) and, in view of definition of p, rewrite it as

$$(a+bz)^2 \left[c^2 \left[\frac{\partial \psi(x,z)}{\partial x} \right]^2 + \left[\frac{\partial \psi(x,z)}{\partial z} \right]^2 \right] = 1. \tag{3.34}$$

We are looking for function $\psi(x,z)$ that satisfies equation (3.34) and the initial conditions. Rather than attempting to solve this nonlinear partial differential equation, we will study the system of ordinary differential equations that are the characteristic equations of equation (3.34). The solution of this system will provide us with information about the physical phenomenon that is governed by eikonal equation (3.34).

The first two equations of system (3.29) define the vector field in the xz-plane. The solutions, $x(t)$ and $z(t)$ are the integral curves, which describe the path of evolution of the propagation

of a signal under an elliptical velocity dependence with direction and under a linear velocity dependence along the z-axis. These curves correspond to rays. At a given point, $(x(t), z(t))$, we can express the direction of a given ray as

$$\theta = \arctan \frac{\mathrm{d}x}{\mathrm{d}z},$$

where θ is measured from the z-axis and is referred to as the ray angle. Also, at any point, we can express the magnitude of the velocity of the signal along the ray as

$$V = \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}z}{\mathrm{d}t}\right)^2},$$

where V is referred to as the ray velocity.

To examine the last two equations of system (3.29), we recall that contours of $\psi(x, z)$ correspond to wavefronts at given instants of time. Hence, $p_x := \partial\psi/\partial x$ and $p_z := \partial\psi/\partial z$ at a given point, $(x(t), z(t))$, are the components of slowness with which a wavefront propagates at this point. Also, vector $p = [p_x, p_z]$ at $(x(t), z(t))$ is normal to the wavefront at this point. Hence, it provides the direction of wavefront propagation at this point. We can express this direction by expression (3.26), with ϑ being the wavefront angle. Also, we can express the magnitude of the velocity of the wavefront propagation as

$$v = \frac{1}{\sqrt{p_x^2 + p_z^2}},$$

where v is referred to as the wavefront velocity.

Now, let us examine the initial conditions. Setting $[x(0), z(0)] = [0, 0]$, we fix the origin of the ray at the initial time. In other words, we locate the point source. In view of continuity of wavefronts and considering the inhomogeneity along the z-axis only, we know that

$$\mathfrak{p} = \frac{\sin\vartheta}{v(\vartheta, z)}.$$

Considering $z = 0$ and using expression (3.24), we can write

$$p_x(0) = \mathfrak{p} = \frac{\sin\vartheta_0}{a\sqrt{c^2\sin^2\vartheta_0 + \cos^2\vartheta_0}}, \tag{3.35}$$

where ϑ_0 denotes the take-off wavefront angle; in other words, the direction of the wavefront at the source.

We wish to solve system (3.29). Since we already know that $p_x = \mathfrak{p}$, we must solve equations (3.30), (3.31) and (3.32) for $x(t)$, $z(t)$ and $p_z(t)$. To do so, let us write the second equation of system (3.29) as

$$p_z = \frac{\frac{\mathrm{d}z}{\mathrm{d}t}}{[a + bz]^2}.$$

Differentiating with respect to t, we get

$$\frac{\mathrm{d}p_z}{\mathrm{d}t} = \frac{\frac{\mathrm{d}^2 z}{\mathrm{d}t^2}(a+bz) - 2b\left(\frac{\mathrm{d}z}{\mathrm{d}t}\right)^2}{(a+bz)^3}.$$

Equating this result with the third equation of system (3.29) and rearranging, we get

$$(a+bz)\frac{\mathrm{d}^2 z}{\mathrm{d}t^2} - 2b\left(\frac{\mathrm{d}z}{\mathrm{d}t}\right)^2 + b(a+bz)^2 = 0.$$

Letting $(a+bz)^2 = y$, we get

$$\frac{1}{2b}\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} - \frac{3}{4yb}\left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2 + by = 0.$$

Letting $y = a^2 \exp u$, we get

$$\frac{\mathrm{d}^2 u}{\mathrm{d}t^2} - \frac{1}{2}\left(\frac{\mathrm{d}u}{\mathrm{d}t}\right)^2 + 2b^2 = 0.$$

Letting $\mathrm{d}u/\mathrm{d}t = q$, we get

$$\frac{\mathrm{d}q}{\mathrm{d}t} - \frac{1}{2}q^2 + 2b^2 = 0.$$

We can rewrite this equation as

$$\mathrm{d}t = \frac{\mathrm{d}q}{\frac{1}{2}q^2 - 2b^2}.$$

Integrating both sides of this equation, we get

$$t + A_1 = -\frac{1}{b}\tanh^{-1}\frac{q}{2b},$$

where A_1 is an integration constant. Solving for q, we get

$$q = 2b\tanh\left(-b\left(t + A_1\right)\right).$$

To obtain u, we integrate and get

$$u = -2\ln\left(\cosh\left(b\left(t + A_1\right)\right)\right) + A_2,$$

where A_2 is an integration constant. Hence,

$$y = a^2 \exp\left[A_2\right]\cosh^{-2}\left(b\left(t + A_1\right)\right). \tag{3.36}$$

Since $z = \left(\sqrt{y} - a\right)/b$, we have the solution of the second equation of system (3.29), namely,

$$z\left(t\right) = \frac{1}{b}\frac{a\exp\left(\frac{A_2}{2}\right)}{\cosh\left(b\left(t + A_1\right)\right)} - \frac{a}{b}.$$

Also – in view of the second equation of system (3.29) – we have $p_z = \dot{z}/y$. Thus, we have the solution of the third equation of system (3.29); this solution is

$$p_z\left(t\right) = -\frac{\sinh\left(b\left(t + A_1\right)\right)}{a\exp\left(\frac{A_2}{2}\right)}.$$

We can write the remaining equation of system (3.29) as

$$\frac{\mathrm{d}x}{\mathrm{d}t} = yc^2\mathfrak{p},$$

where y is given by expression (3.36). Integrating, we get

$$x\left(t\right) = \frac{\mathfrak{p}a^2c^2\exp A_2}{b}\tanh\left(b\left(t+A_1\right)\right) + A_3,$$

where A_3 is an integration constant. Thus, we can concisely write the general solution of system (3.29) as

$$\begin{aligned} x\left(t\right) &= \frac{\mathfrak{p}a^2c^2\exp A_2}{b}\tanh\left(b\left(t+A_1\right)\right) + A_3, \\ z\left(t\right) &= \frac{1}{b}\frac{a\exp\left(\frac{A_2}{2}\right)}{\cosh\left(b\left(t+A_1\right)\right)} - \frac{a}{b}, \\ p_x\left(t\right) &= \mathfrak{p}, \\ p_z\left(t\right) &= -\frac{\sinh\left(b\left(t+A_1\right)\right)}{a\exp\left(\frac{A_2}{2}\right)} \end{aligned} \tag{3.37}$$

To get the particular solution, we must find the three integration constants. To find these integration constants, we invoke the initial conditions. At $t = 0$, we can rewrite the solutions stated in the first two expressions of set (3.37) as

$$x\left(0\right) = \frac{\mathfrak{p}a^2c^2\exp A_2}{b}\tanh\left(bA_1\right) + A_3 = 0,$$

$$z\left(0\right) = \frac{1}{b}\frac{a\exp\left(\frac{A_2}{2}\right)}{\cosh\left(bA_1\right)} - \frac{a}{b} = 0.$$

Also, considering $z = 0$ and using the second equation of system (3.29) combined with solution $z\left(t\right)$ given in set (3.37) and evaluating at $t = 0$, we can write

$$p_z\left(0\right) = \frac{\dot{z}\left(0\right)}{a^2} = -\frac{\sinh\left(bA_1\right)}{a\exp\left(\frac{A_2}{2}\right)} = \sqrt{\frac{1}{a^2} - \mathfrak{p}^2c^2},$$

where the right-hand side is given in expression (3.33).

Considering the last two equations, we have a system of two equations in two unknowns, A_1 and A_2. Solving, we obtain

$$A_1 = -\frac{1}{b}\tanh^{-1}\sqrt{1-\mathfrak{p}^2a^2c^2}$$

and

$$A_2 = -\ln\left[\mathfrak{p}^2a^2c^2\right].$$

Inserting $A_1$and A_2 into the equation for $x\left(0\right)$, we obtain

$$A_3 = \frac{\sqrt{1-\mathfrak{p}^2a^2c^2}}{\mathfrak{p}b}.$$

Examining A_1, A_2 and A_3, we see that the units of A_1 are the units of time, A_2 is dimensionless and the units of A_3 are the units of distance. This is consistent with positions of A_1, A_2 and A_3 in system (3.37).

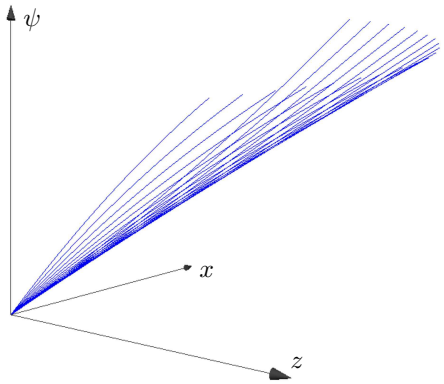


Fig. 3.3. Characteristics of the eikonal equation: Graph of the solution of the Hamilton ray equations (3.29), which is given in expression (3.38) with $a = 2000,\ b = 0.8$ and $c = 1.25$.

Having found A_1, A_2 and A_3, we can rewrite solutions (3.37) as

$$
\begin{aligned}
x(t) &= \frac{1}{\mathfrak{p}b}\left[\tanh\left(bt - \tanh^{-1}\sqrt{1-\mathfrak{p}^2a^2c^2}\right) + \sqrt{1-\mathfrak{p}^2a^2c^2}\right] \\
z(t) &= \frac{a}{b}\left[\frac{1}{\mathfrak{p}ac\cosh\left(bt - \tanh^{-1}\sqrt{1-\mathfrak{p}^2a^2c^2}\right)} - 1\right] \\
p_x(t) &= \mathfrak{p} \\
p_z(t) &= -\frac{1}{\mathfrak{p}a^2c}\sinh\left(bt - \tanh^{-1}\sqrt{1-\mathfrak{p}^2a^2c^2}\right)
\end{aligned}
\tag{3.38}
$$

where, in view of expression (3.35), we have

$$
\mathfrak{p} = \frac{\sin\vartheta_0}{a\sqrt{c^2\sin^2\vartheta_0 + \cos^2\vartheta_0}}
$$

with ϑ_0 being the take-off wavefront angle.

Thus, for abc model, we can choose the wavefront take-off angle, ϑ_0, and using the first two expressions of solutions (3.38) obtain the ray along which the signal generated at $(0,0)$ propagates.

Examination of the first two expressions of solutions (3.38) allows us to learn about the shape of rays for the abc model. We can write each of these expressions as

$$
\mathfrak{p}bx(t) - \sqrt{1-\mathfrak{p}^2a^2c^2} = \tanh\left(bt - \tanh^{-1}\sqrt{1-\mathfrak{p}^2a^2c^2}\right)
$$

and

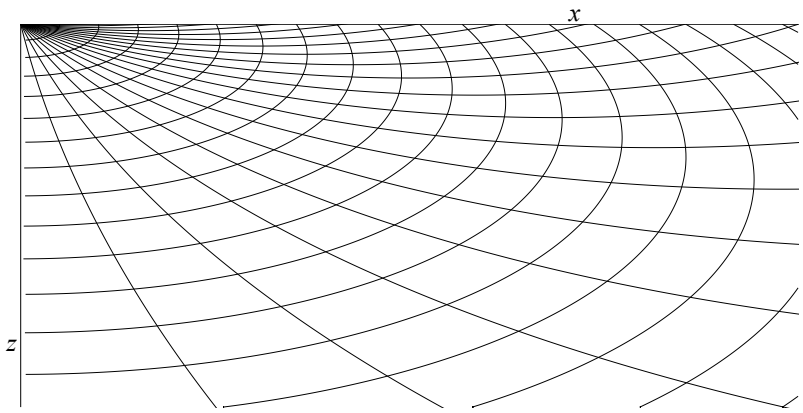


Fig. 3.4. Level curves of the eikonal function, shown in Figure 3.5, and the characteristics of the eikonal equation, shown in Figure 3.3, projected onto the xz-plane. They correspond to wavefronts and rays, respectively, for the abc model with $a = 2000$, $b = 0.8$ and $c = 1.25$.

$$\left[\frac{b}{a}z\left(t\right)+1\right]\mathfrak{p}ac=\frac{1}{\cosh\left(bt-\tanh^{-1}\sqrt{1-\mathfrak{p}^{2}a^{2}c^{2}}\right)},$$

respectively. Squaring these two equations, adding them together and using standard identities, we obtain

$$\frac{\left(x-\frac{\sqrt{1-\mathfrak{p}^{2}a^{2}c^{2}}}{\mathfrak{p}b}\right)^{2}}{\left(\frac{1}{\mathfrak{p}b}\right)^{2}}+\frac{\left(z+\frac{a}{b}\right)^{2}}{\left(\frac{1}{\mathfrak{p}bc}\right)^{2}}=1. \tag{3.39}$$

This is the equation of an ellipse with a centre on the line given by $z = -a/b$. In other words, in the abc model, rays are elliptical arcs. In view of $v\left(z\right) = a + bz$ – if, as usual, we assume $b > 0$ to describe the increase of velocity along the z-axis – we conclude that the centre of the ellipse corresponds to the level where the velocity vanishes. As shown in Figure 3.4, we have obtained rays and wavefronts. Note that although, at this stage, the last two expressions of solutions (3.37) are no longer necessary to obtain rays, traveltimes and wavefronts, we had to use all four equations to solve system (3.29).

For an isotropic version of this model, we set $c = 1$. In such a case equation (3.39) reduces to the equation of a circle whose radius is $1/pb$ and centred at $x = \sqrt{1-p^{2}a^{2}}/pb$ and $z = -a/b$. Derivation of such an equation is shown in Exercise 3.3, where we use a Hamiltonian that is different from expression (3.28), used herein. Examining equation (3.39) and Exercise 3.3, we could see that two different Hamiltonians that are both rooted in eikonal equation (3.25) result in the same rays, since they are the solutions of the characteristic equations of eikonal equation (3.25).

For a homogeneous version of this model, we set $b = 0$. In such a case the rays are straight lines.

Using solutions (3.38), we can obtain the graph of the solution of the original partial differential equation, namely, eikonal equation (3.27), as a parametric plot, $\left[x\left(t\right), z\left(t\right), t\right]$, for all $\mathfrak{p}$ that are consistent with the original equation, as shown in Figure 3.3. In the present case, following expression (3.35), if we set $\vartheta_0 \in (-\pi/2, \pi/2)$, we get

$$\mathfrak{p} \in \left(-\frac{1}{ac}, \frac{1}{ac}\right).$$

In the case of the present example, we can also obtain an explicit analytic form of the solution of the original partial differential equation, namely, eikonal equation (3.27). In other words, we can use our results to obtain $\psi(x, z)$. To do so, we proceed in the following way. Since $t = \psi(x, z)$, let us solve the first equation of set (3.38) for t to get

$$t(x; \mathfrak{p}) = \frac{\tanh^{-1}\left(\mathfrak{p}bx - \sqrt{1 - \mathfrak{p}^2 a^2 c^2}\right) + \tanh^{-1}\sqrt{1 - \mathfrak{p}^2 a^2 c^2}}{b}. \tag{3.40}$$

To express t in terms of x and z, and the parameters of a given abc model, we solve equation (3.39) for $\mathfrak{p}$ to get

$$\mathfrak{p}(x, z) = \frac{2x}{\sqrt{\left[x^2 + c^2 z^2\right]\left[(2a + bz)^2 c^2 + b^2 x^2\right]}}. \tag{3.41}$$

Thus, expression (3.40) with $\mathfrak{p}$ given by expression (3.41) is the solution, $\psi = t(x, z)$, of equation (3.27). This solution is shown in Figure 3.5.

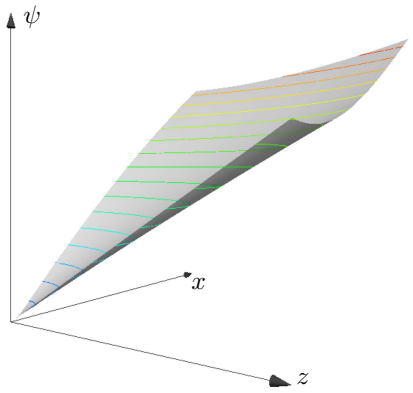


Fig. 3.5. Eikonal function: Graph of the solution of eikonal equation (3.27) with $a = 2000$, $b = 0.8$ and $c = 1.25$.

3.3.2 Electromagnetism

The eikonal equation associated with the electromagnetic waves is given by expression (2.56), namely,

$$\sum_{i=1}^{3}\left(\frac{\partial\psi}{\partial x_i}\right)^2 = \frac{1}{c^2},$$

where c is the speed of light in a vacuum. This equation results from Maxwell equations – given in expressions (F.2), (F.4), (F.6) and (F.9), which describe the electromagnetic phenomena in the idealization of free space: a space that does not react to these phenomena; consequently, we are dealing with an isotropic and homogeneous medium. Hence, $p = \nabla\psi$ is a constant equal to $1/c$.

Furthermore, F is constant with respect to all variables, and the Hamilton equations (3.18) and (3.20) are reduced to

$$\frac{\mathrm{d}x_i}{\mathrm{d}s} = 0$$

and

$$\frac{\mathrm{d}p_i}{\mathrm{d}s} = 0,$$

respectively, which means that both x and p are constants.

Closing remarks

In the first three chapters we have constructed characteristics for the linear first- and second-order equations and for nonlinear first-order equations. Extensions to the higher-order linear equations are straightforward following the method of the Taylor expansion of solutions for Cauchy problems. Extensions of the higher-order nonlinear equations are rendered difficult due to the fact that the relations among the p_i would depend on higher-order derivatives; in other words, we would not be able to relate them using only ξ_j. Remaining within the realm of the discussed equations is sufficient to study a wealth of differential equations in mathematical physics.

Exercises

Exercise 3.1. Consider equation (1.8), which we can rewrite as

$$\frac{\partial\psi\left(x_1, x_2\right)}{\partial x_1} + x_2\frac{\partial\psi\left(x_1, x_2\right)}{\partial x_2} = 0. \tag{3.42}$$

Using equations (3.18), show that the characteristic curves of equation (3.42) are given by expression (1.11), namely,

$$x_2 = C\exp x_1. \tag{3.43}$$

Solution 3.1. Denoting the gradient of ψ by p, we can write equation (3.42) as

$$p_1 + x_2 p_2 = 0.$$

With the Hamiltonian $F = p_1 + x_2 p_2$, equations (3.18) become

$$\frac{\mathrm{d}x_1}{\mathrm{d}s} = \zeta\frac{\partial F}{\partial p_1} = \zeta$$

and

$$\frac{\mathrm{d}x_2}{\mathrm{d}s} = \zeta\frac{\partial F}{\partial p_2} = \zeta x_2.$$

We can write the ratio of $\dot{x}_2$ and $\dot{x}_1$ as

$$\frac{\frac{\mathrm{d}x_2}{\mathrm{d}s}}{\frac{\mathrm{d}x_1}{\mathrm{d}s}} = \frac{\mathrm{d}x_2}{\mathrm{d}x_1} = x_2,$$

which is expression (1.10). Solving this ordinary differential equation, we obtain

$$x_2 = C \exp x_1,$$

where C denotes a constant. This is the characteristic curve given by expression (3.43), as required. Thus, solving the Hamilton equations for x_1 and x_2, we obtained the same result as we did using the directional derivative in Section 1.1.2.

Exercise 3.2. Find the solution of the following equation.

$$\frac{\partial f}{\partial x} - \left(\frac{\partial f}{\partial y}\right)^2 = x,$$

where the solution satisfies the following side condition.

$$f(0, y) = f_0(y)$$

Solution 3.2. This is a nonlinear first-order partial differential equation. Let the Hamiltonian be given by

$$F = \frac{\partial f}{\partial x} - \left(\frac{\partial f}{\partial y}\right)^2 - x,$$

which can be written as

$$F = p_x - p_y^2 - x.$$

Hence, we write the Hamilton equations as

$$\begin{aligned}
x'(s) &= \frac{\partial F}{\partial p_x} = 1, \\
y'(s) &= \frac{\partial F}{\partial p_y} = -2p_y, \\
p_x'(s) &= -\frac{\partial F}{\partial x} - \frac{\partial F}{\partial f} p_x = 1, \\
p_y'(s) &= -\frac{\partial F}{\partial y} - \frac{\partial F}{\partial f} p_y = 0, \\
f'(s) &= p_x \frac{\partial F}{\partial p_x} + p_y \frac{\partial F}{\partial p_y},
\end{aligned}$$

where we chose the parametrization by s to be such that the scaling parameter is equal to one. Integrating, we obtain the solutions of the first four equations; they are

$$\begin{aligned}
x(s) &= s + x_0, \\
y(s) &= -2p_y s + y_0, \\
p_x(s) &= s + p_{x0}, \\
p_y(s) &= p_{y0},
\end{aligned}$$

where p_{x0}, p_{y0}, x_0 and y_0 are constants. Hence, the fifth equation implies that the unknown function changes along the characteristic curves as

$$f'(s) = s + p_{x0} - 2p_{y0}^2.$$

Integrating, we obtain

$$f(s) = \frac{1}{2}s^2 + \left(p_{x0} - 2p_{y0}^2\right)s + f_0. \tag{3.44}$$

Considering the side condition, $f(0, y) = f_0(y)$, we see that

$$\begin{aligned} x(s) &= s \\ y(s) &= -2f_0' s + y_0 \\ p_x(s) &= s + \left(f_0'(y_0)\right)^2 \\ p_y(s) &= f_0'(y_0) \end{aligned}$$

and

$$f(s) = \frac{1}{2}s^2 - \left(f_0'(y_0)\right)^2 s + f_0(y_0). \tag{3.45}$$

We can parametrize the xy-plane by s and y_0. In general, we are unable to express explicitly the dependence of s and y_0 on x and y. Such an expression is possible for certain cases, as we will exemplify by considering the following particular form of f_0, namely, $f_0(y) = y$. In such a case, $s = x$ and $y_0 = y + 2x$. After we substitute for these in expression (3.45), we get

$$f(x, y) = \frac{1}{2}x^2 - x + (y + 2x) = \frac{1}{2}x^2 + x + y.$$

Exercise 3.3. Recall eikonal equation (3.25), and choose the corresponding Hamiltonian to be

$$H = \frac{1}{2}\left(p_x^2 + p_z^2 - \frac{1}{v^2}\right), \tag{3.46}$$

which is different from the one used in Section 3.2. Considering a vertically inhomogeneous medium and using the Hamilton equations, derive the general expression for a ray as a function given by $x = x(z)$. Then, use the derived expression to find the traveltime along the ray. Also, discuss rays in the vertically inhomogeneous medium where the velocity increases linearly with depth.

Solution 3.3. Inserting expression (3.46), which for a vertically inhomogeneous medium is given explicitly by

$$H(p_x, p_z, z) = \frac{1}{2}\left[p_x^2 + p_z^2 - \frac{1}{v(z)^2}\right],$$

into the Hamilton equations, we obtain

$$x' = \frac{\partial H}{\partial p_x} = p_x, \tag{3.47}$$

$$z' = \frac{\partial H}{\partial p_z} = p_z, \tag{3.48}$$

$$p_x' = -\frac{\partial H}{\partial x} = 0 \tag{3.49}$$

and

$$p_z' = -\frac{\partial H}{\partial z} = -\frac{\frac{\mathrm{d}v}{\mathrm{d}z}}{v^3}, \tag{3.50}$$

where the prime denotes a derivative with respect to parameter s.

From equation (3.49) we see that p_x is constant along a ray. Using this fact in equation (3.47), we infer that

$$x(s) = p_x s + x_0. \tag{3.51}$$

This expression states the x-component of a ray as a function of parameter s.

To obtain an expression given as $x = x(z)$, we recall eikonal equation (3.25), which we can write as

$$p_z = \pm\sqrt{\frac{1}{v(z)^2} + p_x^2}.$$

In view of this expression, equation (3.48) implies that

$$z' = \pm\sqrt{\frac{1}{v(z)^2} - p_x^2}.$$

This is a separable differential equation whose solution is

$$s = \pm\int_0^z \frac{1}{\sqrt{\frac{1}{v(\zeta)^2} - p_x^2}}\mathrm{d}\zeta,$$

where we set s in such a way that $z(0) = 0$. Using equation (3.51), we write

$$x(z) = \pm p_x \int_0^z \frac{1}{\sqrt{\frac{1}{v(\zeta)^2} - p_x^2}}\mathrm{d}\zeta + x_0, \tag{3.52}$$

which is the desired expression; it describes a ray in a vertically inhomogeneous medium.

To find the expression for the traveltime along the ray, we invoke the fact that we can write

$$t(x) = p_x x, \tag{3.53}$$

where we set $x(0) = 0$. In other words, since p_x is constant along a ray, the traveltime corresponding to the ray is the product of the x-component of slowness and the distance measured along the x-axis. To write the traveltime in terms of the z-axis, we use equations (3.52) and (3.53) to get

$$t(z) = p_x\left(\pm p_x \int_0^z \frac{1}{\sqrt{\frac{1}{v(\zeta)^2} - p_x^2}}\mathrm{d}\zeta + x_0\right).$$

Together, the two above equations give us the desired relation between the traveltime, t, and position, $[x, z]$.

To discuss rays in a vertically inhomogeneous medium where the velocity increases linearly with depth, we let $v(z) = a + bz$. In such a case, we write expression (3.52) as

$$x(z) = \pm p_x \int_0^z \frac{1}{\sqrt{\frac{1}{(a+b\zeta)^2} - p_x^2}}\mathrm{d}\zeta + x_0.$$

After integration, we get

$$x(z) = \pm \frac{p_x^2 (a+bz)^2 - 1}{b p_x (a+bz) \sqrt{\frac{1}{(a+bz)^2} - p_x^2}} + x_0,$$

which we can simplify to obtain

$$x(z) = \pm \sqrt{\frac{1}{b^2 p_x^2} - \left(z + \frac{a}{b}\right)^2} + x_0.$$

To geometrically interpret this result, we rewrite it as

$$(x - x_0)^2 + \left(z + \frac{a}{b}\right)^2 = \frac{1}{b^2 p_x^2}. \tag{3.54}$$

This is an equation of a circle whose centre is at $[x_0, -a/b]$ and whose radius is $1/bp_x$. Thus, in a vertically inhomogeneous medium where the velocity increases linearly with depth, rays are circular arcs. The radius is infinitely long if $b = 0$ or if $p_x = 0$. Thus, the rays are straight lines if the medium is homogeneous or if the ray coincides with the direction of the velocity gradient. As expected, equation (3.54) corresponds to equation (3.39) with $c = 1$; an ellipse is reduced to a circle.

4

Asymptotic solutions of differential equations

Preliminary remarks

One is often interested in the behaviour of a function around a particular point and not necessarily in the whole domain. There is a well-known apparatus for describing a differentiable function at a point; the Taylor series describes an infinitely differentiable function at a point.

In this chapter we will describe another way of using an infinite series to describe a functions at a point. We will use the asymptotic series. Unlike the Taylor series, the asymptotic series need not converge to a function. However, unlike in the case of the Taylor series, a given function need not be differentiable at the point of interest to be represented by an asymptotic series. As a matter of fact, this point might be a singularity of that function. Moreover, often they are better approximations of given functions than the Taylor series, since the latter ones might converge slowly while already the initial terms of the asymptotic series are much closer to the values of the functions they represent.

4.1 Asymptotic series

Since, in general, asymptotic series do not converge, there must be an advantage in using them for a function representation. This advantage is given by the measure of how we approximate the function. In other words, the notion of approximation of a function is changed as compared to, say, the Taylor series. To define the asumptotic approximation, we start by saying that two functions f and g are asymptotically equivalent to one another in the Poincaré sense at x_0 if

$$\lim_{x \to x_0} (x - x_0)^{-m} (f(x) - g(x)) = 0, \quad \forall m \in N,$$

where f and g are defined on some interval containing x_0. We denote this equvalency by $f \sim g$ as $x \to x_0$. If we want to compare two functions at infinty, we must change the definition as follows. We say that f and g are asymptotically equivalent to one another at infinity, if

$$\lim_{x \to \infty} x^m (f(x) - g(x)) = 0, \quad \forall m \in N.$$

Since we want to approximate functions by series that might not converge, we cannot directly use the definition of asymptotic equivalence to define the asymptotic expansion in an infinite series. We define asymptotic expansion as follows.

Definition 4.1. *For a function f we say that $\sum_{n=1}^{\infty} a_n (x-x_0)^n$ is its infinite asymptotic expansion (of Poinaré type) at x_0 if*

$$\lim_{x\to x_0} (x-x_0)^{-n} \left(f(x) - \sum_{j=0}^{n} a_j (x-x_0)^j \right) = 0, \quad \forall n \in N.$$

We can find the coefficients of the asymptotic series starting from computing a_0, then a_1 and so on. We compute a_0 from Definition 4.1 by setting $n=0$, namely,

$$\lim_{x\to x_0} (f(x) - a_0) = 0 \Rightarrow a_0 = \lim_{x\to x_0} f(x).$$

Similarly, for $n=1$, we compute a_1 using the computed value of a_0,

$$\lim_{x\to x_0} \frac{(f(x)-a_0) - a_1(x-x_0)}{(x-x_0)} = 0 \Rightarrow a_1 = \lim_{x\to x_0} \frac{f(x)-a_0}{(x-x_0)}.$$

For $n=2$,

$$\lim_{x\to x_0} \frac{(f(x)-a_0) - a_1(x-x_0) - a_2(x-x_0)^2}{(x-x_0)^2} = 0 \Rightarrow a_2 = \lim_{x\to x_0} \frac{(f(x)-a_0) - a_1(x-x_0)}{(x-x_0)^2}.$$

If $f \in C^2(I)$, where I is open and $x_0 \in I$, then $a_0 = f(x_0)$, $a_1 = f'(x_0)$ and $a_2 = f''(x_0)/2!$.

In general, if $f \in C^{\infty}(I)$, then the coefficients of its asymptotic series are given uniquely by

$$a_n = \frac{f^{(n)}(x_0)}{n!}$$

and the nth partial sum S_n becomes

$$S_n(x) = \sum_{j=0}^{n} \frac{f^{(j)}(x_0)}{j!(x-x_0)^j}.$$

If f is only continuous on I, then $a_0 = f(x_0)$, $a_1 = \lim_{x\to x_0} (f(x) - S_0(x))/(x-x_0)$, and in general, $a_n = \lim_{x\to x_0} (f(x) - S_{n-1}(x))/(x-x_0)$.

As we will see in Section 4.2, we will have to generalize the asymptotic sequence of the Poincaré type, $\sum_{n=1}^{\infty} a_n (x-x_0)^n$, to be used in our studies.

While a function has a unique asymptotic sequence, an asymptotic sequence corresponds to many functions, as exemplified in the following example.

Example 4.1. Consider the following function

$$g(x) = \begin{cases} e^{-\frac{1}{x}}, & x > 0 \\ 0, & x \le 0 \end{cases}.$$

We see that $g \in C^{\infty}(\mathbb{R})$ and $g^{(n)}(0) = 0$. Thus, as $x \to 0$ the asymptotic expansion of this function is a sequence of zeros. This implies that 0 and $g(x)$ have the same expansion around zero; thus we can write $g(x) \sim 0$. This implies that for all functions f that have an asymptotic expansion around zero, f and $f+g$ have the same asymptotic expansion around zero; $f(x) + g(x) \sim f(x)$.

Also, any sequence of real numbers defines an asymptotic expansion of a function, as shown in the following theorem.

Theorem 4.1. *For any sequence of real numbers* $\{a_n\}_{n=1}^{\infty} \in \mathbb{R}$*, there is a function* f *such that*

$$f(x) \sim \sum_{n=0}^{\infty} a_n (x - x_0)^n \text{ as } x \to x_0.$$

Proof. If the series $\sum_{n=0}^{N} a_n (x - x_0)^n$ converges uniformly to a function $f(x)$ as $N \to \infty$, then $f(x) \sim \sum_{n=0}^{\infty} a_n (x - x_0)^n$ as $x \to x_0$. If the series does not converge uniformly to a function, then we have to consider a convergent series that behaves around x_0 the same way as the original series. This series can be constructed by multiplying each term of the series by a proper cut-off function that is one in a neighbourhood of x_0 and zero outside some neighbourhood containing the first one. Since the asymptotic series only describes the function around a point, it follows that such a modified series with the cut-off functions is going to have the same asymptotic behaviour around point x_0 as the original series. Now we have to answer the question of uniform convergence of such a modified series. ?

It is useful to introduce a short-hand notation for the asymptotic behaviour. In particular, we express the fact that

$$\lim_{x \to x_0} \frac{f(x) - S_n(x)}{(x - x_0)^n} = 0, \quad \forall n \in \mathbb{N}$$

by writing

$$f(x) - S_n(x) = o\left((x - x_0)^n\right), \quad \forall n \in \mathbb{N}. \tag{4.1}$$

It is also convenient to introduce the following notation. Similarly, we express the fact that

$$\lim_{x \to x_0} \left| \frac{f(x) - S_n(x)}{(x - x_0)^{n+1}} \right| < K, \quad \forall n \in \mathbb{N},$$

where K is a constant, by writing

$$f(x) - S_n(x) = O\left((x - x_0)^{n+1}\right), \quad \forall n \in N. \tag{4.2}$$

Note that equations (4.1) and (4.2) are equivalent definitions for the existence of infinite asymptotic expansions. The proof of this is given in Exercise 4.1.

Similarly, we can express the asymptotic expansion of $f(x)$ at infinity, namely

$$f(x) \sim \sum_{n=0}^{\infty} a_n x^{-n},$$

by writing

$$f(x) - S_n(x) = o\left(x^{-n}\right),$$

which is equivalent to

$$f(x) - S_n(x) = O\left(x^{-(n+1)}\right).$$

Following the definition, series $\sum_{n=0}^{\infty} a_n x^{-n}$ is an asymptotic expansion of $f(x)$ as $x \to \infty$ if

$$\lim_{x\to\infty} x^n \left(f\left(x\right) - S_n\left(x\right)\right) = 0, \text{ for fixed } n,$$

even though, for fixed x, the following might happen:

$$\lim_{n\to\infty} x^n \left(f\left(x\right) - S_n\left(x\right)\right) = \infty.$$

Consider the following example.

Example 4.2. We will find the asymptotic expansion at infinity of

$$f\left(x\right) = xe^x \int_x^\infty t^{-1}e^{-t}\mathrm{d}t.$$

If $f\left(x\right)$ has an asymptotic expansion as $x \to \infty$, namely

$$f(x) \sim \sum_{n=0}^{\infty} a_n x^{-n},$$

then

$$a_0 = \lim_{x\to\infty} f\left(x\right).$$

Explicitly, we write

$$\lim_{x\to\infty} xe^x \int_x^\infty t^{-1}e^{-t}\mathrm{d}t = \lim_{x\to\infty} \frac{x\int_x^\infty t^{-1}e^{-t}\mathrm{d}t}{e^{-x}}.$$

Since $\lim_{x\to\infty} x\int_x^\infty t^{-1}e^{-t}\mathrm{d}t = 0$ and $\lim_{x\to\infty}\int_x^\infty t^{-1}e^{-t}\mathrm{d}t = 0$, as shown in Exercise 4.2, we can use the de l'Hôpital rule twice to evaluate the limit.

$$\lim_{x\to\infty} \frac{\int_x^\infty t^{-1}e^{-t}\mathrm{d}t - xx^{-1}e^{-x}}{-e^{-x}} = 1 + \lim_{x\to\infty} \frac{\int_x^\infty t^{-1}e^{-t}\mathrm{d}t}{-e^{-x}} = 1 + \lim_{x\to\infty} \frac{-x^{-1}e^{-x}}{e^{-x}} = 1.$$

Hence, $a_0 = 1$. To compute a_1 we use the formula that follows from the definition of the asymptotic sequence, namely,

$$a_n = \lim_{x\to\infty} x^n \left(f\left(x\right) - a_0 - a_1x^{-1} - a_2x^{-2} - \cdots - a_{n-1}x^{-(n-1)}\right). \tag{4.3}$$

Herein, using the de l'Hôpital rule, we compute

$$\begin{aligned}
a_1 &= \lim_{x\to\infty} \left(x^2e^x \int_x^\infty t^{-1}e^{-t}\mathrm{d}t - x\right) \\
&= \lim_{x\to\infty} \frac{x^2\int_x^\infty t^{-1}e^{-t}\mathrm{d}t - xe^{-x}}{e^{-x}} \\
&= \lim_{x\to\infty} \frac{2x\int_x^\infty t^{-1}e^{-t}\mathrm{d}t}{-e^{-x}} + \lim_{x\to\infty} \frac{-x^2x^{-1}e^{-x} - e^{-x} + xe^{-x}}{-e^{-x}} \\
&= -1.
\end{aligned}$$

Using formula (4.3), we could recursively find all coefficients a_n. This might not be the most convenient procedure to find the asymptotic expansion. Herein, we can use integration by parts to write

$$xe^x \int_x^\infty t^{-1}e^{-t}\mathrm{d}t = xe^x \int_x^\infty t^{-1}\left(-e^{-t}\right)'\mathrm{d}t,$$

and proceed to get

$$f(x) = -xe^x t^{-1}e^{-t}\Big|_x^\infty - xe^x \int_x^\infty t^{-2}e^{-t}\mathrm{d}t.$$

Continuing the process, we obtain

$$\begin{aligned} f(x) &= 1 + xe^x t^{-2}e^{-t}\Big|_x^\infty + 2!xe^x \int_x^\infty t^{-3}e^{-t}\mathrm{d}t \\ &= 1 + x^{-1} + 2!xe^x \int_x^\infty t^{-3}\left(-e^{-t}\right)'\mathrm{d}t \\ &= 1 + x^{-1} + 2!x^{-2} + \ldots + (-1)^n n!x^{-n} + (-1)^{n+1}(n+1)!xe^x \int_x^\infty t^{-n-1}e^{-t}\mathrm{d}t \\ &= \sum_{m=0}^{n} \frac{(-1)^m m!}{x^m} + (-1)^{n+1}(n+1)! \int_x^\infty t^{-n-1}e^{-t}\mathrm{d}t. \end{aligned}$$

To see that the sum in the above expression is the asymptotic expansion of function f, we have to show that the remainder vanishes as $x \to \infty$ faster than x^{n+1} goes to infinity. In other words, we have to show that

$$\lim_{x\to\infty} x^{n+1} \int_x^\infty t^{-n-1}e^{-t}\mathrm{d}t = 0,$$

which can be accomplished following the steps used in Exercise 4.2.

Therefore we can write

$$f(x) \sim \sum_{m=0}^{\infty} \frac{(-1)^m m!}{x^m}, \quad \text{as } x \to \infty.$$

Sometimes we are not interested in the entire asymptotic series, but we want to know the first few terms. This motivates the following definition.

Definition 4.2. *A function f has a finite asymptotic expansion as $x \to x_0$ to N terms if there are N coefficients $\{a_n\}_{n=0}^{N-1}$ such that the partial sums*

$$S_n = \sum_{j=0}^{n-1} a_j (x - x_0)^j$$

satisfy the following condition.

$$\lim_{x\to x_0} \left(f(x) - S_n(x)\right)(x - x_0)^n = 0, \quad n = 0, 1, \ldots, N-1.$$

As discussed in Section 4.2, we need to generalize the notion of the asymptotic sequence.

Definition 4.3. *An asymptotic sequence as $x \to x_0$ is a sequence of functions $\{\phi_n(x)\}_{n=0}^{\infty}$ such that $\phi_{n+1}(x) = o(\phi_n(x))$ as $x \to x_0$ for all n.*

Using such a sequence, we can develop asymptotic expansions of functions, as defined below.

Definition 4.4. *Let $\{\phi_n\}$ be an asymptotic sequence. We say that f has an infinite asymptotic expansion with respect to this sequence as $x \to x_0$ if there exist a sequence $\{a_n\} \in \mathbb{R}$ such that*

$$\lim_{x \to x_0} \frac{f(x) - \sum_{j=0}^{n} a_j \phi_j(x)}{\phi_n(x)} = 0,$$

which is equivalent to

$$f(x) - S_n(x) = o(\phi_n(x)), \forall n \in N.$$

This definition can be generalized to vector-valued functions.

Definition 4.5. *For a vector-valued function $f : R \to V$, where V is a normed vector space with norm $\| \cdot \| : V \to R$, function $f(x)$ has an asymptotic expansion with respect to $\{\phi_n\}$ as $x \to x_0$ if there exist a sequence $a_n \in V$ such that*

$$\lim_{x \to x_0} \frac{\| f(x) - \sum_{m=0}^{n} a_m \phi_m \|}{\phi_n} = 0, \quad \forall n \in N.$$

Example 4.3. As an example of the above definition, we consider the asymptotic expansion of a vector valued function $\tilde{u}(x, \omega)$ with respect to basis $\phi_n = \exp(i\omega\psi) / (i\omega)^n$ as $\omega \to \infty$.

$$\tilde{u}(x, \omega) \sim e^{\iota\omega\psi(x)} \left(A_0(x) + \frac{A_1(x)}{i\omega} + \sum_{n=2}^{\infty} \frac{A_n(x)}{(i\omega)^n} \right) \tag{4.4}$$

Series (4.4) is an asymptotic series of $\tilde{u}$ that we use below to study elastodynamic equations. The computation of the first three coefficients A_n is shown in Exercise 4.12.

4.2 Choice of asymptotic sequence

In the subsequent sections, we will study equations of the following type

$$Df = 0, \tag{4.5}$$

where

$$D = \sum_{k \le m} \frac{1}{(i\omega)^k} D_k,$$

with

$$D_k = \sum_{|\alpha| \le k} B_\alpha^k(x) \frac{\partial^{|\alpha|}}{\partial x^\alpha}.$$

If we want to solve such an equation by using an asymptotic series, the first thing we must decide is to choose the proper asymptotic sequence ϕ_n and the variable in which we are going to

make the expansion. Since the solution is a function of both x and ω, the functions ϕ_n are also depending on x and ω. If we choose to expand the solution in ω as $\omega \to \infty$, then the asymptotic sequence must satisfy

$$\lim_{\omega\to\infty} \frac{\phi_{n+1}(x,\omega)}{\phi_n(x,\omega)} = 0. \tag{4.6}$$

Considering the particular form of the differential equation, we require that $\partial^m \phi_n/\partial x^m$ and $(i\omega)^{-k}\phi_n$ are a finite sum of terms $a_k(x)\phi_k$. These conditions are satisfied if ϕ is of the form

$$\phi_n = \exp\left((i\omega)^m \psi(x)\right) \frac{1}{(i\omega)^k},$$

as shown in Exercise 4.13. We can choose k and m such that they satisfy condition (4.6), for example

$$\phi_n = \exp\left(i\omega\psi(x)\right) \frac{1}{(i\omega)^n}.$$

Prior to inserting the asymptotic series into the equations of motion, let us explain our motivation for using series (4.4). In particular, let us discuss the importance of the exponential term. What is the reason for considering an asymptotic series given by the product of the exponential term and the summation, which we explicitly write as

$$\sum_{n=0}^{N} A_n(x)\phi_n = \exp\left(i\omega\psi(x)\right) \sum_{n=0}^{N} \frac{A_n(x)}{(i\omega)^n}, \tag{4.7}$$

rather than the summation itself? An insight into the physical reason is given by the meanings of ψ and A, which are associated with wavefronts and amplitudes, respectively. This is, however, an insight achieved by examining the solutions to the equations of motion. The mathematical reason to begin with such series to obtain the solution can be explained as follows. The bases of our asymptotic expansion, which we will use in differential equations, are $\exp(i\omega\psi)/(i\omega)^n$. Consider the derivatives of expression (4.7). Each time we take a derivative of the exponential term with respect to x_i, we get $i\omega(\partial\psi/\partial x_i)\exp(i\omega\psi)$. Thus, the derivatives of expression (4.7) are linear combinations of $\exp(i\omega\psi)/(i\omega)^n$.

4.3 Asymptotic differential equations

In this section we will discuss the general form of asymptotic differential equations.

The general form of an asymptotic differential equation of order m can be written as

$$Df = 0, \tag{4.8}$$

where

$$D = \sum_{k\le m} \frac{1}{(i\omega)^k} D_k, \tag{4.9}$$

with

$$D_k = \sum_{|\alpha|\le k} B_\alpha^k(x) \frac{\partial^{|\alpha|}}{\partial x^\alpha}. \tag{4.10}$$

Following the discussion of the choice of the asymptotic sequence in Section 4.2, we look for solutions in the form given by

$$f(x,\omega) = e^{i\omega\psi(x)} \sum_{j=0}^{\infty} A_j(x) (i\omega)^{-j} \tag{4.11}$$

that satisfy equation (4.8). Substituting expression (4.11) into equation (4.8), we write

$$\sum_{k\leq m} (i\omega)^{-k} \sum_{|\alpha|\leq k} B_\alpha^k(x) \frac{\partial^{|\alpha|}}{\partial x^\alpha} \left(e^{i\omega\psi(x)} \sum_{j=0}^{\infty} A_j(x) (i\omega)^{-j} \right) = 0, \tag{4.12}$$

where the zero on the right-hand side is represented by its asymptotic series with the coefficients set to zero, namely,

$$0 = e^{i\omega\psi} \sum_{l=0}^{\infty} 0 (i\omega)^{-l}.$$

To solve equation (4.8), one has to investigate individual terms of the resulting asymptotic sequence. We will compute these terms by applying asymptotic differential operator (4.9) to solution (4.11).

Following the generalized Leibniz rule shown in Exercise 4.6, we write equation (4.12) as

$$\sum_{k\leq m} (i\omega)^{-k} \sum_{|\alpha|\leq k} B_\alpha^k(x) \sum_{\beta+\gamma=\alpha} \frac{\alpha!}{\beta!\gamma!} \frac{\partial^{|\beta|} e^{i\omega\psi}}{\partial x^\beta} \sum_{j=0}^{\infty} \frac{\partial^{|\gamma|} A_j}{\partial x^\gamma} (i\omega)^{-j} = 0. \tag{4.13}$$

Since the left-hand side of equation (4.13) is an asymptotic series, let us consider the terms with different powers of $(i\omega)$.

4.4 Eikonal equation

The coefficient corresponding to the zeroth power of $i\omega$ on the left-hand side of equation (4.13) results from differentiating k times only the exponential terms, which correspond to $\beta = \alpha$. For multiindex α we can write

$$\frac{\partial^{|\alpha|}}{\partial x^\alpha} e^{i\omega\psi(x)} = \frac{\partial^{|\alpha|-1}}{\partial x^{\alpha-(0,\ldots,0,1,0,\ldots,0)}} \frac{\partial}{\partial x_i} \left(e^{i\omega\psi(x)} \right) = \frac{\partial^{|\alpha|-1}}{\partial x^{\alpha-(0,\ldots,0,1,0,\ldots,0)}} e^{i\omega\psi(x)} (i\omega) \frac{\partial\psi}{\partial x_i},$$

where 1 in multiindex $(0,\ldots,0,1,0,\ldots,0)$ appears in the ith location. Differentiating with respect to x_j, we write

$$\frac{\partial^{|\alpha|}}{\partial x^\alpha} e^{i\omega\psi(x)} = \frac{\partial^{|\alpha|-2}}{\partial x^{\alpha-(0,\ldots,0,1,0,\ldots,0,1,0,\ldots,0)}} \frac{\partial}{\partial x_j} \left(e^{i\omega\psi(x)} (i\omega) \frac{\partial\psi}{\partial x_i} \right)$$

as

$$\frac{\partial^{|\alpha|}}{\partial x^{\alpha}} e^{i\omega\psi(x)} = \frac{\partial^{|\alpha|-2}}{\partial x^{\alpha-(0,\ldots,0,1,0,\ldots,0,1,0,\ldots,0)}} e^{i\omega\psi(x)} \left((i\omega)^2 \left(\frac{\partial \psi}{\partial x} \right)^{(0,\ldots,0,1,0,\ldots,0,1,0,\ldots,0)} + (i\omega) \frac{\partial^2 \psi}{\partial x_i \partial x_j} \right),$$

where 1 in multiindex $(0,\ldots,0,1,0,\ldots,0,1,0,\ldots,0)$ appears in the ith and jth locations. If i is equal to j, the multiindex has 2 in the ith location. Continuing the differentiation, we obtain

$$\frac{\partial^{|\alpha|} e^{i\omega\psi}}{\partial x^{\alpha}} = e^{i\omega\psi} \left((i\omega)^{|\alpha|} \left(\frac{\partial \psi}{\partial x} \right)^{\alpha} + (i\omega)^{|\alpha|-1} \left(\sum_{i,j \in \alpha} \frac{\partial^2 \psi}{\partial x_i \partial x_j} \left(\frac{\partial \psi}{\partial x} \right)^{\alpha-(0,\ldots,0,1,0,\ldots,0,1,0,\ldots,0)} \right) + \cdots \right), \quad (4.14)$$

where the sum is over all pairs of indices contained in multiindex α as many times as they appear, and the dots indicate the remaining terms with lower powers of $(i\omega)$. More explicitly, the number of times we have to count each term

$$\frac{\partial^2 \psi}{\partial x_i \partial x_j} \left(\frac{\partial \psi}{\partial x} \right)^{\alpha-(0,\ldots,0,1,0,\ldots,0,1,0,\ldots,0)}$$

in the summation is the number of times we can choose $(0,\ldots,0,1,0,\ldots,0,1,0,\ldots,0)$ from $(0,\ldots,0,\alpha_i,0,\ldots,0,\alpha_j,0,\ldots,0)$. If $i \neq j$, then this number is $\alpha_i \alpha_j$. If $i = j$, then this number is $\alpha_i (\alpha_i - 1)$. Using the Kronecker delta, we rewrite the summation as

$$\sum_{i,j=1}^{n} \frac{1}{2} \left(\alpha_i \alpha_j (1 - \delta_{ij}) + \alpha_i (\alpha_j - 1) \delta_{ij} \right) \frac{\partial^2 \psi}{\partial x_i \partial x_j} \left(\frac{\partial \psi}{\partial x} \right)^{\alpha-(0,\ldots,0,1,0,\ldots,0,1,0,\ldots,0)}, \quad (4.15)$$

where the factor of $1/2$ comes from the fact that we sum over both i and j and thus account for each term twice. The first term on the right-hand side of expression (4.14) is derived in Exercise 4.5.

Equating the coefficients corresponding to the zeroth power of $(i\omega)$ on both sides of equation (4.13), we see that

$$\sum_{k \leq m} \sum_{|\alpha|=k} B_{\alpha}^{k}(x) \left(\frac{\partial \psi}{\partial x} \right)^{|\alpha|} A_0(x) = 0. \quad (4.16)$$

In this equation there are two unknown functions, $\psi(x)$ and $A_0(x)$. If we consider only nonzero scalar functions A_0, then this equation reduces to

$$\sum_{k \leq m} \sum_{|\alpha|=k} B_{\alpha}^{k}(x) \left(\frac{\partial \psi}{\partial x} \right)^{|\alpha|} = 0. \quad (4.17)$$

This is a first-order nonlinear partial differential equation for $\psi(x)$. We refer to it as an eikonal equation. This name is justified by the following example.

Example 4.4. Consider the reduced wave equation discussed in Section ?, namely,

$$v^2 \nabla^2 u(x,\omega) = (i\omega)^2 u(x,\omega).$$

In this case, equation (4.17) reduces to

$$\sum_i \left(\frac{\partial \psi}{\partial x_i}\right)^2 - \frac{1}{v^2} = 0, \tag{4.18}$$

as shown in Exercise 4.8. This is eikonal equation (2.54), which we derived in Section 2.3.3 using the method of characteristics.

Recalling the definition of the symbol of a differential operator, as discussed in Chapter 6, we see that equation (4.16) can be written as

$$\sigma\left(D\right)\left(\mathrm{d}\psi\right) A_0\left(x\right) = 0. \tag{4.19}$$

If we consider only nonzero scalar functions A_0, then this equation reduces to

$$\sigma\left(D\right)\left(\mathrm{d}\psi\right) = 0. \tag{4.20}$$

Also, we can view the Eikonal equation, (4.19), as set of n equations of the type of equation (4.20). These $n$equations correspond to equating the eigenvalues of $\sigma\left(D\right)\left(d\psi\right)$ to zero. Solutions, $\psi\left(x\right)$, of first-order nonlinear differential equation (4.20) were discussed within the context of characteristics in Chapter 3 and are discussed in the next section in greater detail.

4.5 Solution of eikonal equation

In this section we discuss the solutions of the eikonal equation. The solution using the general method of characteristics for first-order nonlinear partial differential equations was given in Chapter 3.

To use the results from that chapeter in the context of eikonal equation (4.20), we express the eikonal equation in the notation of Chapter 3 as

$$0 = F\left(x_1, x_2, x_3, \psi, \frac{\partial \psi}{\partial x_1}, \frac{\partial \psi}{\partial x_2}, \frac{\partial \psi}{\partial x_3}\right) \equiv \sigma\left(D\right)\left(\mathrm{d}\psi\right).$$

For vector valued equations, we treat this as a set of n equations corresponding to the zero eigenvalues of $\sigma\left(D\right)\left(d\psi\right)$. In the case of the eikonal equation, $\sigma\left(D\right)\left(d\psi\right)$ does not depend explicitly on ψ. Consequently, the characteristic equations (3.18), (3.20) and (3.19) become

$$\begin{aligned}
\frac{\mathrm{d}x_i}{\mathrm{d}t} &= \zeta \frac{\partial \sigma\left(D\right)\left(p\right)}{\partial p_i}, \\
\frac{\mathrm{d}p_i}{\mathrm{d}t} &= -\zeta \frac{\partial \sigma\left(D\right)\left(p\right)}{\partial x_i} \\
\frac{\mathrm{d}\psi}{\mathrm{d}t} &= \zeta \sum_{i=1}^{3} \frac{\partial \sigma\left(D\right)\left(p\right)}{\partial p_i} p_i.
\end{aligned} \tag{4.21}$$

The solutions of this system depend on the original partial differential equation, which manifests itself in the form of $\sigma\left(D\right)\left(p\right)$. As an example, we return to the reduced wave equation in two spatial dimensions, in which case,

$$\sigma\left(D\right)\left(p\right)=p_1^2+p_2^2-\frac{1}{v^2}.$$

The solution is discussed in Exercise 3.3, in which we used $\zeta=1/2$.

4.6 Transport equation

In this section we will consider terms with $\left(i\omega\right)^{-1}$.

Examining equation (4.13), we focus our attention on all the terms with $\left(i\omega\right)^{-1}$. There are four types of such terms in the series.

Some of them result from setting $j=1$ and $|\beta|=|\alpha|=k$; they are

$$\sum_{k\leq m}\sum_{|\alpha|=k}B_\alpha^k\left(x\right)\left(\frac{\partial\psi}{\partial x}\right)^{|\alpha|}e^{i\omega\psi}A_1\left(x\right)\left(i\omega\right)^{-1}.$$

Others result from setting $j=0$ and $|\beta|=|\alpha|=k-1$; they are

$$\sum_{k\leq m}\sum_{|\alpha|=k-1}B_\alpha^k\left(x\right)\left(\frac{\partial\psi}{\partial x}\right)^{|\alpha|}e^{i\omega\psi}A_0\left(x\right)\left(i\omega\right)^{-1}.$$

More of them result from setting $j=0$, $|\beta|=|\alpha|=k$ and considering expression (4.15); they are

$$\sum_{k\leq m}\sum_{|\alpha|=k}B_\alpha^k\sum_{i,j=1}^{n}\frac{1}{2}\left(\alpha_i\alpha_j\left(1-\delta_{ij}\right)+\alpha_i\left(\alpha_j-1\right)\delta_{ij}\right)\frac{\partial^2\psi}{\partial x_i\partial x_j}\left(\frac{\partial\psi}{\partial x}\right)^{\alpha-(0,\ldots,0,1,0,\ldots,0,1,0,\ldots,0)}e^{i\omega\psi}A_0\left(i\omega\right)^{-1}.$$

The remaining terms result from setting $j=0$, $|\beta|+1=|\alpha|=k$; they are

$$\sum_{k\leq m}\sum_{|\alpha|=k}B_\alpha^k\sum_{i=1}^{n}\alpha_i\left(\frac{\partial\psi}{\mathrm{d}x}\right)^{\alpha-(0,\ldots,0,1,0,\ldots,0)}e^{i\omega\psi}\frac{\partial A_0}{\partial x_i}\left(i\omega\right)^{-1},$$

where 1 in $(0,\ldots,0,1,0,\ldots,0)$ appears in the ith location.

Gathering these four expressions and factoring out the common terms, we write

$$e^{i\omega\psi}\left(i\omega\right)^{-1}\left(\sum_{k\leq m}\sum_{|\alpha|=k}B_\alpha^k\left(x\right)\left(\frac{\partial\psi}{\partial x}\right)^{|\alpha|}A_1\left(x\right)+\sum_{k\leq m}\sum_{|\alpha|=k-1}B_\alpha^k\left(x\right)\left(\frac{\partial\psi}{\partial x}\right)^{|\alpha|}A_0+\right.$$
$$\sum_{k\leq m}\sum_{|\alpha|=k}B_\alpha^k\sum_{i,j=1}^{n}\frac{1}{2}\left(\alpha_i\alpha_j\left(1-\delta_{ij}\right)+\alpha_i\left(\alpha_j-1\right)\delta_{ij}\right)\frac{\partial^2\psi}{\partial x_i\partial x_j}\left(\frac{\partial\psi}{\partial x}\right)^{\alpha-(0,\ldots,0,1,0,\ldots,0,1,0,\ldots,0)}A_0+$$
$$\left.\sum_{k\leq m}\sum_{|\alpha|=k}B_\alpha^k\sum_{i=1}^{n}\alpha_i\left(\frac{\partial\psi}{\mathrm{d}x}\right)^{\alpha-(0,\ldots,0,1,0,\ldots,0)}\frac{\partial A_0}{\partial x_i}\right).$$

Since in an asymptotic series that is equal to zero all the coefficients are zero, we conclude that

$$\sum_{k\leq m}\sum_{|\alpha|=k} B_\alpha^k(x)\left(\frac{\partial\psi}{\partial x}\right)^{|\alpha|} A_1(x) + \left(\sum_{k\leq m}\sum_{|\alpha|=k-1} B_\alpha^k(x)\left(\frac{\partial\psi}{\partial x}\right)^{|\alpha|} + \right.$$
$$\sum_{k\leq m}\sum_{|\alpha|=k} B_\alpha^k \sum_{i,j=1}^{n}\frac{1}{2}\left(\alpha_i\alpha_j\left(1-\delta_{ij}\right)+\alpha_i\left(\alpha_j-1\right)\delta_{ij}\right)\frac{\partial^2\psi}{\partial x_i\partial x_j}\left(\frac{\partial\psi}{\partial x}\right)^{\alpha-(0,\ldots,0,1,0,\ldots,0,1,0,\ldots,0)} +$$
$$\left.\sum_{k\leq m}\sum_{|\alpha|=k} B_\alpha^k \sum_{i=1}^{n}\alpha_i\left(\frac{\partial\psi}{\mathrm{d}x}\right)^{\alpha-(0,\ldots,0,1,0,\ldots,0)}\frac{\partial}{\partial x_i}\right) A_0(x) = 0. \quad (4.22)$$

Example 4.5. Consider the reduced wave equation discussed in Section ?, namely,

$$v^2\nabla^2 u(x,\omega) = (i\omega)^2 u(x,\omega). \quad (4.23)$$

In this case, equation (4.22) reduces to

$$\left(\sum_i\left(\frac{\partial\psi}{\partial x_i}\right)^2 - \frac{1}{v^2}\right)A_1(x) + \left(\sum_i\frac{\partial^2\psi}{\partial x_i^2} + \sum_i 2\frac{\partial\psi}{\partial x_i}\frac{\partial}{\partial x_i}\right)A_0(x) = 0, \quad (4.24)$$

as shown in Exercise 4.9. Furthermore, if we assume that ψ satisfies eikonal equation (4.18), this equation becomes

$$\left(\sum_i\frac{\partial^2\psi}{\partial x_i^2} + \sum_i 2\frac{\partial\psi}{\partial x_i}\frac{\partial}{\partial x_i}\right)A_0(x) = 0,$$

which is a first-order linear partial differential equation for $A_0(x)$.

Using the symbols of the differential operators discussed in Chapter 6, we rewrite equation (4.22) as

$$\sigma(D)(\mathrm{d}\psi)A_1(x) +$$
$$\left(\sum_{k\leq m}\sum_{|\alpha|=k-1} B_\alpha^k(x)\left(\frac{\partial\psi}{\partial x}\right)^{|\alpha|} + \sum_{k\leq m}\frac{1}{2}\sum_{i,j=1}^{n}\frac{\partial^2\sigma(D_k)(p)}{\partial p_i\partial p_j}\frac{\partial^2\psi}{\partial x_i\partial x_j} + \sum_{k\leq m}\sum_{i=1}^{n}\frac{\partial\sigma(D_k)(p)}{\partial p_i}\frac{\partial}{\partial x_i}\right)A_0(x) = 0, \quad (4.25)$$

where $p = \mathrm{d}\psi$. The details of this rewriting are shown in Exercise 4.10.

If $\psi(x)$ satisfies eikonal equation (4.17) and if $A_j(x)$ are scalars or if we assume that $A_1(x)$ is in the kernel of $\sigma(D)(d\psi)$, the first term of the above equation is zero . Subsequently, equation (4.25) reduces to

$$\left(\sum_{k\leq m}\sum_{|\alpha|=k-1} B_\alpha^k(x)\left(\frac{\partial\psi}{\partial x}\right)^{|\alpha|} + \sum_{k\leq m}\frac{1}{2}\sum_{i,j=1}^{n}\frac{\partial^2\sigma(D_k)(p)}{\partial p_i\partial p_j}\frac{\partial^2\psi}{\partial x_i\partial x_j} + \sum_{k\leq m}\sum_{i=1}^{n}\frac{\partial\sigma(D_k)(p)}{\partial p_i}\frac{\partial}{\partial x_i}\right)A_0(x) = 0, \quad (4.26)$$

which is a first-order partial differential equation for $A_0(x)$.

4.7 Solution of transport equation

In this section we discuss the solutions of the reduced transport equation (4.26). To do so, we write equation (4.26) as

$$\sum_{k\leq m}\sum_{i=1}^{n}\frac{\partial\sigma\left(D_k\right)(p)}{\partial p_i}\frac{\partial}{\partial x_i}A_0(x)=-\left(\sum_{k\leq m}\sum_{|\alpha|=k-1}B_\alpha^k(x)\left(\frac{\partial\psi}{\partial x}\right)^{|\alpha|}+\sum_{k\leq m}\frac{1}{2}\sum_{i,j=1}^{n}\frac{\partial^2\sigma\left(D_k\right)(p)}{\partial p_i\partial p_j}\frac{\partial^2\psi}{\partial x_i\partial x_j}\right)A_0(x). \tag{4.27}$$

We can solve this equation by the method of characteristics, which is discussed in Chapter 1. In view of the solutions of the eikonal equation discussed in Section 4.5, the characteristics of the transport and the eikonal equations are the same. Along these characteristics, given by equations (4.21), the left-hand side of equation (4.27) is

$$\frac{1}{\zeta}\sum_{i=1}^{n}\frac{\mathrm{d}}{\mathrm{d}t}x_i(t)\frac{\partial}{\partial x_i}A_0(x(t)),$$

where $x(t)$ is a characteristic. For vector valued functions, we view this expressions as a set of n equations for each eigenvalue of $\sigma(D)(p)$. I such a case we do not get the solution of the transport equation as a vector A_0, but only as its length. The direction of A_0 has to be determined directly from the eikonal equation, (4.19), as the direction of the eigenvector corresponding to the given eigenvalue.

In view of the chain rule, we write the transport equation along the characteristics as

$$\frac{\mathrm{d}}{\mathrm{d}t}A_0(x(t))=-\zeta\left(\sum_{k\leq m}\sum_{|\alpha|=k-1}B_\alpha^k(x)\left(\frac{\partial\psi}{\partial x}\right)^{|\alpha|}+\sum_{k\leq m}\frac{1}{2}\sum_{i,j=1}^{n}\frac{\partial^2\sigma\left(D_k\right)(p)}{\partial p_i\partial p_j}\frac{\partial^2\psi}{\partial x_i\partial x_j}\right)A_0(x(t)).$$

Integrating along a characteristic, as shown in Exercise 4.11, we obtain

$$A_0(x(t))=\exp\left\{-\int_{t_0}^{t}\zeta\sum_{k\leq m}\sum_{|\alpha|=k-1}B_\alpha^k(x)\left(\frac{\partial\psi}{\partial x}\right)^{|\alpha|}+\sum_{k\leq m}\frac{\zeta}{2}\sum_{i,j=1}^{n}\frac{\partial^2\sigma\left(D_k\right)(p)}{\partial p_i\partial p_j}\frac{\partial^2\psi}{\partial x_i\partial x_j}\mathrm{d}s\right\}A_0(x(0)),$$

where $A_0(x(0))$ is the zeroth order amplitude at time $t=0$. The above expression leads to the solution of the transport equation.

If $\psi(x(t))$ is a solution of the eikonal equation, then, in view of the Hamilton equations (4.21), we write the second part of the above integrand as

$$\sum_{k\leq m}\frac{\zeta}{2}\sum_{i,j=1}^{n}\frac{\partial^2\sigma\left(D_k\right)(p)}{\partial p_i\partial p_j}\frac{\partial^2\psi}{\partial x_i\partial x_j}=\sum_{k\leq m}\frac{\zeta}{2}\sum_{i,j=1}^{n}\frac{\partial}{\partial p_j}\frac{\partial\sigma\left(D_k\right)(p)}{\partial p_i}\frac{\partial^2\psi}{\partial x_i\partial x_j}=\sum_{k\leq m}\frac{\zeta}{2}\sum_{i,j=1}^{n}\frac{\partial\frac{1}{\zeta}\dot{x}_i}{\partial p_j}\frac{\partial^2\psi}{\partial x_i\partial x_j},$$

where $\dot{x}\equiv\mathrm{d}x/\mathrm{d}t$. Also, since $p_j=\partial\psi/\partial x_j$, we can state it as

$$\sum_{k\leq m}\frac{\zeta}{2}\sum_{i,j=1}^{n}\frac{\partial^2\sigma\left(D_k\right)(p)}{\partial p_i\partial p_j}\frac{\partial^2\psi}{\partial x_i\partial x_j}=\frac{\zeta}{2}\sum_{i,j=1}^{n}\frac{\partial\frac{1}{\zeta}\dot{x}_i}{\partial p_j}\frac{\partial p_j}{\partial x_i},$$

which, using the chain rule, we concisely write as

$$\frac{\zeta}{2}\nabla\cdot\left(\frac{1}{\zeta}\dot{x}\left(x\left(t\right)\right)\right).$$

Consequently, solution $A_0\left(x\left(t\right)\right)$ becomes

$$A_0\left(x\left(t\right)\right)=\exp\left\{-\int_{t_0}^{t}\zeta\sum_{k\leq m}\sum_{|\alpha|=k-1}B_\alpha^k\left(x\right)\left(\frac{\partial\psi}{\partial x}\right)^{|\alpha|}\mathrm{d}s\right\}\exp\left\{-\int_{t_0}^{t}\frac{\zeta}{2}\nabla\cdot\left(\frac{1}{\zeta}\dot{x}\left(x\left(s\right)\right)\right)\mathrm{d}s\right\}A_0\left(x\left(0\right)\right). \tag{4.28}$$

The form of solution $A_0\left(x\left(t\right)\right)$ depends on the form of solution of the eikonal equation, $\psi\left(x\left(t\right)\right)$.

Let us explain the meaning of the second exponential. Since the integrand in the second exponent consists of a divergence, we wish to invoke the the divergence theorem stated on 133. To do so, let us consider the volume integral given by

$$\iiint_{V(t)}\nabla\cdot\dot{x}\left(x\left(s\right)\right)\mathrm{d}V, \tag{4.29}$$

where volume $V\left(t\right)$ is spanned by area S_0 that is normal to the characteristics, $x\left(s\right),s\in\left[t_0,t\right]$, along which this area propagates. Using Theorem A.1, we write this integral as

$$\iiint_{V(t)}\nabla\cdot\dot{x}\mathrm{d}V=\iint_{S(t)}\dot{x}\cdot n\mathrm{d}S, \tag{4.30}$$

where n is the outward normal to the surface of the ray tube, $S\left(t\right)$, and also is the outward normal to the tube ends. Along the tube, n is normal to $\dot{x}$; hence, $\dot{x}\cdot n=0$. On the two tube ends, n is parallel to $\dot{x}$; hence, since n is a unit vector, $\dot{x}\cdot n=\pm\left|\dot{x}\right|$. Consequently, we rewrite equation (4.30) as

$$\iiint_{V(t)}\nabla\cdot\dot{x}\mathrm{d}V=-\iint_{S_0}\left|\dot{x}\right|\mathrm{d}S+\iint_{S_t}\left|\dot{x}\right|\mathrm{d}S, \tag{4.31}$$

where the negative sign results from the fact that at S_0, $\dot{x}$ and n have opposite directions. To study the propagation along the ray tube, we differentiate both sides of equation (4.31) with respect to t to write

$$\frac{\mathrm{d}}{\mathrm{d}t}\iiint_{V(t)}\nabla\cdot\dot{x}\mathrm{d}V=-\frac{\mathrm{d}}{\mathrm{d}t}\iint_{S_0}\left|\dot{x}\right|\mathrm{d}S+\frac{\mathrm{d}}{\mathrm{d}t}\iint_{S_t}\left|\dot{x}\right|\mathrm{d}S. \tag{4.32}$$

Invoking the definition of derivative, we write the left-hand side of equation (4.32) as

$$\lim_{\mathrm{d}t\to0}\frac{\iiint_{V(t+\mathrm{d}t)}\nabla\cdot\dot{x}\mathrm{d}V-\iiint_{V(t)}\nabla\cdot\dot{x}\mathrm{d}V}{\mathrm{d}t}=\lim_{\mathrm{d}t\to0}\frac{\iiint_{V(t+\mathrm{d}t)-V(t)}\nabla\cdot\dot{x}\mathrm{d}V}{\mathrm{d}t}. \tag{4.33}$$

Let us consider the numerator. According to the Mean-value Theorem for integrals, there exists a surface, $S_{t+\mathrm{d}t}$, such that

$$\iiint_{V(t+\mathrm{d}t)-V(t)}\nabla\cdot\dot{x}\mathrm{d}V=\mathrm{d}t\iint_{S_{t+dt}}\nabla\cdot\dot{x}\mathrm{d}S.$$

Hence, returning to expression (4.33), we get

$$\lim_{\mathrm{d}t\to 0}\frac{\iiint\limits_{V(t+\mathrm{d}t)-V(t)}\nabla\cdot\dot{x}\mathrm{d}V}{\mathrm{d}t}=\lim_{\mathrm{d}t\to 0}\frac{\mathrm{d}t\iint\limits_{S_{t+dt}}\nabla\cdot\dot{x}\mathrm{d}S}{\mathrm{d}t}=\lim_{\mathrm{d}t\to 0}\iint\limits_{S_{t+dt}}\nabla\cdot\dot{x}\mathrm{d}S=\iint\limits_{S_t}\nabla\cdot\dot{x}\mathrm{d}S.$$

Thus, we write the left-hand side of equation (4.32) as

$$\frac{\mathrm{d}}{\mathrm{d}t}\iiint\limits_{V(t)}\nabla\cdot\dot{x}\mathrm{d}V=\iint\limits_{S_t}\nabla\cdot\dot{x}\mathrm{d}S. \tag{4.34}$$

Examining the right-hand side of equation (4.32), we see that the integral over S_0 is a constant and, hence, its derivative is zero. Let us consider the integral over S_t. According to the Mean-value Theorem for integrals, there exists a point, $y\left(t\right)$, on surface S_t such that

$$\iint\limits_{S_t}\left|\dot{x}\right|\mathrm{d}S=S_t\left|\dot{x}\left(y\left(t\right)\right)\right|.$$

Using this expression and invoking the product rule and the chain rule, we write

$$\frac{\mathrm{d}}{\mathrm{d}t}\iint\limits_{S_t}\left|\dot{x}\right|\mathrm{d}S=\frac{\mathrm{d}}{\mathrm{d}t}\left(S_t\left|\dot{x}\left(y\left(t\right)\right)\right|\right)=\frac{\mathrm{d}S_t}{dt}\left|\dot{x}\left(y\left(t\right)\right)\right|+S_t\nabla\left|\dot{x}\left(y\left(t\right)\right)\right|\cdot\dot{y}\left(t\right), \tag{4.35}$$

which is the right-hand side of equation (4.32). Equating the left-hand and right-hand sides of equation (4.32), which are given in expressions (4.34) and (4.35), respectively, we write equation (4.32) as

$$\iint\limits_{S_t}\nabla\cdot\dot{x}\mathrm{d}S=\frac{\mathrm{d}S_t}{dt}\left|\dot{x}\left(y\left(t\right)\right)\right|+S_t\nabla\left|\dot{x}\left(y\left(t\right)\right)\right|\cdot\dot{y}\left(t\right), \tag{4.36}$$

which is the time derivative of integral (4.29). To find solution of equation (4.36), we restrict the parametrization by t to the arclength parametrization. In such a case, $|\dot{x}|=1$ and the second term on the right-hand side of equation (4.36) is zero. Hence, equation (4.36) becomes

$$\iint\limits_{S_t}\nabla\cdot\dot{x}\mathrm{d}S=\frac{\mathrm{d}S_t}{dt}. \tag{4.37}$$

To use this result in the examination of solution (4.28), we wish to remove the surface integral. To do so, let us consider the left-hand side of equation (4.37). According to the Mean-value Theorem for integrals, there exists a point, $z\left(t\right)$, on surface S_t such that

$$\iint\limits_{S_t}\nabla\cdot\dot{x}\mathrm{d}S=S_t\nabla\cdot\dot{x}\left(z\left(t\right)\right).$$

Thus we write equation (4.37) as

$$S_t\nabla\cdot\dot{x}\left(z\left(t\right)\right)=\frac{\mathrm{d}S_t}{dt}.$$

In view of the chain rule, we rewrite it as

$$\frac{\frac{\mathrm{d}S_t}{\mathrm{d}t}}{S_t} = \frac{\mathrm{d}}{\mathrm{d}t}\ln S_t = \nabla\cdot\dot{x}\left(y\left(t\right)\right).$$

For an infinitesimally small surface, S_t, we can replace point $y\left(t\right)$ with $x\left(t\right)$. Thus, we write

$$\nabla\cdot\dot{x}\left(x\left(t\right)\right) = \frac{\mathrm{d}}{\mathrm{d}t}\ln S_t, \tag{4.38}$$

which is another form of equation (4.37), and consequently, equation (4.36) for the arclength parametrization.

If parameter t correspond to the arclength, then change in the surface area corresponds to the change in the volume along the ray tube. We can express this change using the Jacobian of the change of coordinates: from the Cartesian coordinates to the coordinates given by the characteristics. To do so, we write expression (4.38) as

$$\nabla\cdot\dot{x}\left(x\left(t\right)\right) = \frac{\mathrm{d}}{\mathrm{d}t}\ln \Delta V_t, \tag{4.39}$$

where ΔV_t is the volume generated by surface S_t displaced by Δt, where Δt is constant; in other words, $\Delta V_t = S_t \Delta t$. To justify this step, we note that

$$\frac{\mathrm{d}}{\mathrm{d}t}\ln \Delta V_t = \frac{\mathrm{d}}{\mathrm{d}t}\ln\left(S_t\Delta t\right) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\ln S_t + \ln \Delta t\right) = \frac{\mathrm{d}}{\mathrm{d}t}\ln S_t.$$

Thus, divergence $\nabla\cdot\dot{x}\left(x\left(t\right)\right)$ can be expressed as the change of the logarithm of the volume of a ray tube with respect to the arclength. Using n ray coordinates, $\xi_1, \xi_2, \ldots, \xi_{n-1}, t$, we can write this volume as

$$\Delta V_t = \det\begin{bmatrix} \frac{\partial x_1}{\partial \xi_1} & \frac{\partial x_1}{\partial \xi_2} & \cdots & \frac{\partial x_1}{\partial t} \\ \frac{\partial x_2}{\partial \xi_1} & \frac{\partial x_2}{\partial \xi_2} & \cdots & \frac{\partial x_2}{\partial t} \\ & & \ddots & \\ \frac{\partial x_n}{\partial \xi_1} & \frac{\partial x_n}{\partial \xi_2} & \cdots & \frac{\partial x_n}{\partial t} \end{bmatrix} \mathrm{d}\xi_1\mathrm{d}\xi_2\ldots\mathrm{d}t = J\mathrm{d}\xi_1\mathrm{d}\xi_2\ldots\mathrm{d}t,$$

where J stands for the determinant of the above matrix. Using this expression, we restate equation (4.39) as

$$\nabla\cdot\dot{x}\left(x\left(t\right)\right) = \frac{\mathrm{d}}{\mathrm{d}t}\ln\left(J\mathrm{d}\xi_1\mathrm{d}\xi_2\ldots\mathrm{d}t\right) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\ln J + \ln \mathrm{d}\xi_1\mathrm{d}\xi_2\ldots\mathrm{d}t\right).$$

Since $\ln \mathrm{d}\xi_1\mathrm{d}\xi_2\ldots\mathrm{d}t$ is a constant, we obtain

$$\nabla\cdot\dot{x}\left(x\left(t\right)\right) = \frac{\mathrm{d}}{\mathrm{d}t}\ln J, \tag{4.40}$$

which relates the divergence of the tangents of the characteristics to the change of volume associated with these characteristics.

Let us examine solution (4.28) in the context of expression (4.38). In the second exponent, we have $\nabla\cdot\left(\left(1/\zeta\right)\dot{x}\right)$. Invoking the product rule, we write

$$\nabla\cdot\left(\frac{1}{\zeta}\dot{x}\right) = \left(\nabla\frac{1}{\zeta}\right)\cdot\dot{x} + \frac{1}{\zeta}\nabla\cdot\dot{x}. \tag{4.41}$$

Inserting expression (4.38), we get

$$\nabla \cdot \left(\frac{1}{\zeta} \dot{x} \right) = \frac{\mathrm{d}}{\mathrm{d}t} \frac{1}{\zeta} + \frac{1}{\zeta} \frac{\mathrm{d}}{\mathrm{d}t} \ln S_t,$$

where, in view of the chain rule, we rewrote also the first term on the right-hand side. To write the right-hand side of this equation as a total derivative, we multiply both sides by ζ to get

$$\zeta \nabla \cdot \left(\frac{1}{\zeta} \dot{x} \right) = \zeta \frac{\mathrm{d}}{\mathrm{d}t} \frac{1}{\zeta} + \frac{\mathrm{d}}{\mathrm{d}t} \ln S_t,$$

which, in view of the derivative of a logarithm, can be written as

$$\zeta \nabla \cdot \left(\frac{1}{\zeta} \dot{x} \right) = \frac{\mathrm{d}}{\mathrm{d}t} \ln \frac{1}{\zeta} + \frac{\mathrm{d}}{\mathrm{d}t} \ln S_t,$$

which, in view of the linearity of the differential operator and the algebra of logarithms, can be written as

$$\zeta \nabla \cdot \left(\frac{1}{\zeta} \dot{x} \right) = \frac{\mathrm{d}}{\mathrm{d}t} \left(\ln \frac{1}{\zeta} + \ln S_t \right) = \frac{\mathrm{d}}{\mathrm{d}t} \left(- \ln \zeta + \ln S_t \right) = \frac{\mathrm{d}}{\mathrm{d}t} \ln \frac{S_t}{\zeta}. \tag{4.42}$$

Also, we can express this formula in terms of the Jacobian using expressions (4.41) and (4.40) as

$$\zeta \nabla \cdot \left(\frac{1}{\zeta} \dot{x} \right) = \frac{\mathrm{d}}{\mathrm{d}t} \left(\ln \frac{1}{\zeta} + \ln J \right) = \frac{\mathrm{d}}{\mathrm{d}t} \left(- \ln \zeta + \ln J \right) = \frac{\mathrm{d}}{\mathrm{d}t} \ln \frac{J}{\zeta}. \tag{4.43}$$

Substituting result (4.42) into solution (4.28), we write

$$\begin{aligned} A_0 \left(x \left(t \right) \right) &= A_0 \left(x \left(t_0 \right) \right) \exp \left\{ - \int_{t_0}^{t} \zeta \sum_{k \leq m} \sum_{|\alpha| = k-1} B_\alpha^k \left(x \right) \left(\frac{\partial \psi}{\partial x} \right)^{|\alpha|} \mathrm{d}s \right\} \exp \left\{ - \frac{1}{2} \int_{t_0}^{t} \frac{\mathrm{d}}{\mathrm{d}s} \ln \frac{S_s}{\zeta} \mathrm{d}s \right\} \\ &= A_0 \left(x \left(t_0 \right) \right) \exp \left\{ - \int_{t_0}^{t} \zeta \sum_{k \leq m} \sum_{|\alpha| = k-1} B_\alpha^k \left(x \right) \left(\frac{\partial \psi}{\partial x} \right)^{|\alpha|} \mathrm{d}s \right\} \sqrt{\frac{S_0 \zeta \left(x \left(t \right) \right)}{S_t \zeta \left(x \left(t_0 \right) \right)}}. \end{aligned} \tag{4.44}$$

Also, substituting result (4.43) into solution (4.28), we write

$$A_0 \left(x \left(t \right) \right) = A_0 \left(x \left(t_0 \right) \right) \exp \left\{ - \int_{t_0}^{t} \zeta \sum_{k \leq m} \sum_{|\alpha| = k-1} B_\alpha^k \left(x \right) \left(\frac{\partial \psi}{\partial x} \right)^{|\alpha|} \mathrm{d}s \right\} \sqrt{\frac{J_0 \zeta \left(x \left(t \right) \right)}{J_t \zeta \left(x \left(t_0 \right) \right)}}. \tag{4.45}$$

Solution (4.44) requires infinitesimally small surface S_0. Both solutions (4.44) and (4.45) require the arclength parametrization. The arclength parametrization means that $\| \mathrm{d}x / \mathrm{d}s \| = 1$; hence, following equation (4.21),

$$\zeta = \left\| \frac{\partial \sigma \left(D \right) \left(p \right)}{\partial p} \right\|^{-1}. \tag{4.46}$$

To see the meaning of this expression, let us consider the following example.

Example 4.6. Let us consider an isotropic medium and $F = p^2 \left(x \right) - 1 / v^2 \left(x \right)$. The choice of parametrization fixes the expression for ζ. Following equation (4.46), where $\sigma \left(D \right) \left(p \right) = F$, we see that herein

$$\zeta = \frac{1}{\left|\frac{\partial F}{\partial p_i}\right|} = \frac{1}{2\left|p\right|}.$$

Along the characteristics, $F = 0$, and, hence, $|p| = 1/v\left(x\right)$, where v is the velocity of propagation. Thus, $\zeta = v\left(x\right)/2$. Using the reduced wave equation (4.54) , we see that the exponent in expression (4.44) is zero, as shown in Exercise 4.9. Thus, we can write expression (4.44) as

$$A_0\left(x\left(t\right)\right) = A_0\left(t_0\right)\sqrt{\frac{S_0 v\left(x\left(t\right)\right)}{S_t v\left(x\left(0\right)\right)}}.$$

Solutions (4.44) and (4.45) describe the evolution of amplitude along the characteristics. This amplitude depends only on the radicand in expressions (4.44) and (4.45); the exponential does not depend on the characteristics since they are given only by symbol $\sigma\left(D\right)\left(p\right)$. This solution demonstrates that the square of the amplitude depends on the geometrical spreading of the characteristics; the greater the increase in the raytube cross-section, the greater the decrease of the square of the amplitude, provided that ζ remains the same. The case of the vanishing denominator, which implies that the amplitude tends to infinity, is discussed in Chapter 5.

4.8 Higher-order transport equations

To obtain the entire asymptotic solution of a given differential equation, we need to compute all coefficients $A_j\left(x\right)$. In the preceding section we formulated the way of looking for the first of these coefficients, $A_0\left(x\right)$. We are going to exemplify the procedure for finding these coefficients using the reduced wave equation, below; this example can be viewed as a continuation of Example 4.5.

Example 4.7. Let us consider the reduced wave equation given in expression (4.23) and write it as

$$\frac{v^2}{\left(i\omega\right)^2}\nabla^2 u\left(x,\omega\right) = u\left(x,\omega\right).$$

The asymptotic solution is of the form

$$u\left(x,\omega\right) = \sum_{j=0}^{\infty} e^{i\omega\psi(x)}\frac{A_j\left(x\right)}{\left(i\omega\right)^j}.$$

Substituting this expression into the reduced wave equation and considering the three-dimensional case, we get

$$\frac{v^2}{\left(i\omega\right)^2}\sum_{i=1}^{3}\frac{\partial}{\partial x_i}\left(e^{\iota\omega\psi(x)}\left(\iota\omega\frac{\partial\psi}{\partial x_i}\sum_{j=0}^{\infty}\frac{A_j\left(x\right)}{\left(\iota\omega\right)^j} + \sum_{j=0}^{\infty}\frac{\frac{\partial A_j(x)}{\partial x_i}}{\left(\iota\omega\right)^j}\right)\right) = \sum_{j=0}^{\infty} e^{\iota\omega\psi(x)}\frac{A_j\left(x\right)}{\left(\iota\omega\right)^{j-2}}. \tag{4.47}$$

Continuing the differentiation, we write the left-hand side is of this equation as

$$v^2\sum_{i=1}^{3} e^{\iota\omega\psi(x)}\sum_{j=0}^{\infty}\left(\frac{\partial^2\psi}{\partial x_i^2}\frac{A_j\left(x\right)}{\left(\iota\omega\right)^{j+1}} + \frac{\partial\psi}{\partial x_i}\frac{\frac{\partial A_j(x)}{\partial x_i}}{\left(\iota\omega\right)^{j+1}} + \frac{\frac{\partial^2 A_j(x)}{\partial x_i^2}}{\left(\iota\omega\right)^{j+2}} + \frac{\partial\psi}{\partial x_i}\left(\frac{\partial\psi}{\partial x_i}\frac{A_j\left(x\right)}{\left(\iota\omega\right)^j} + \frac{\frac{\partial A_j(x)}{\partial x_i}}{\left(\iota\omega\right)^{j+1}}\right)\right).$$

The procedure for finding the coefficients consists of equating the corresponding terms of $\iota\omega$ on both sides of equation (4.47).

Considering the terms with $(\iota\omega)^0$, we obtain

$$v^2 \sum_{i=1}^{3} \left(\frac{\partial \psi}{\partial x_i} \right)^2 A_0(x) = A_0(x),$$

which, for nonzero scalar $A_0(x)$, reduces to equation (4.18), as expected.

Considering the terms with $(\iota\omega)^{-1}$, we obtain

$$v^2 \sum_{i=1}^{3} e^{\iota\omega\psi(x)} \left(\frac{\partial^2 \psi}{\partial x_i^2} A_0(x) + \frac{\partial \psi}{\partial x_i} \frac{\partial A_0(x)}{\partial x_i} + \frac{\partial \psi}{\partial x_i} \left(\frac{\partial \psi}{\partial x_i} A_1(x) + \frac{\partial A_0(x)}{\partial x_i} \right) \right) = e^{\iota\omega\psi(x)} A_1(x),$$

which simplifies to

$$\left(\sum_{i=1}^{3} \left(\frac{\partial \psi}{\partial x_i} \right)^2 - \frac{1}{v^2} \right) A_1(x) + \sum_{i=1}^{3} \left(\frac{\partial^2 \psi}{\partial x_i^2} + 2 \frac{\partial \psi}{\partial x_i} \frac{\partial}{\partial x_i} \right) A_0(x) = 0,$$

which is equation (4.24), as expected.

Considering terms with $(\iota\omega)^{-k}$, where $k \geq 2$, we obtain

$$v^2 \sum_{i=1}^{3} e^{\iota\omega\psi(x)} \left(\frac{\partial^2 \psi}{\partial x_i^2} A_{k-1}(x) + \frac{\partial \psi}{\partial x_i} \frac{\partial A_{k-1}(x)}{\partial x_i} + \frac{\partial^2 A_{k-2}(x)}{\partial x_i^2} + \frac{\partial \psi}{\partial x_i} \left(\frac{\partial \psi}{\partial x_i} A_k(x) + \frac{\partial A_{k-1}(x)}{\partial x_i} \right) \right) = e^{\iota\omega\psi(x)} A_k(x),$$

which simplifies to

$$\left(\sum_{i=1}^{3} \left(\frac{\partial \psi}{\partial x_i} \right)^2 - \frac{1}{v^2} \right) A_k(x) + \sum_{i=1}^{3} \left(\frac{\partial^2 \psi}{\partial x_i^2} + 2 \frac{\partial \psi}{\partial x_i} \frac{\partial}{\partial x_i} \right) A_{k-1}(x) + \sum_{i=1}^{3} \frac{\partial^2 A_{k-2}(x)}{\partial x_i^2} = 0.$$

This equation allows us to compute $A_k(x)$, provided that we know all $A_l(x)$ for $l < k$. Hence, we can iteratively compute any term in the asymptotic expansion of the solution. This equation is referred to as the kth-order transport equation.

Since ψ satisfies the eikonal equation, the first term of the above equation is zero, and the kth-order transport equation reduces to

$$\sum_{i=1}^{3} \left(\frac{\partial^2 \psi}{\partial x_i^2} + 2 \frac{\partial \psi}{\partial x_i} \frac{\partial}{\partial x_i} \right) A_{k-1}(x) + \sum_{i=1}^{3} \frac{\partial^2 A_{k-2}(x)}{\partial x_i^2} = 0.$$

Having solved the lower-order transport equations, we know $A_{k-2}(x)$. Thus,

$$\sum_{i=1}^{3} \left(\frac{\partial^2 \psi}{\partial x_i^2} + 2 \frac{\partial \psi}{\partial x_i} \frac{\partial}{\partial x_i} \right) A_{k-1}(x) = - \sum_{i=1}^{3} \frac{\partial^2 A_{k-2}(x)}{\partial x_i^2}.$$

4.9 Asymptotic solution of elastodynamic equations

Following our discussions of asymptotic solutions of the reduced wave equation, we will study the asymptotic solutions of the elastodynamic equations derived in Section D.

Recall equations (D.9); namely,

$$\rho(\mathbf{x})\frac{\partial^2 u_i(\mathbf{x},t)}{\partial t^2}=\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3}\left(\frac{\partial c_{ijkl}(\mathbf{x})}{\partial x_j}\frac{\partial u_k}{\partial x_l}+c_{ijkl}(\mathbf{x})\frac{\partial^2 u_k}{\partial x_j\partial x_l}\right), \tag{4.48}$$

where $i\in\{1,2,3\}$. We wish equations (4.48) to be expressed in terms of $\mathbf{x}$ and ω in order to consider the limit of ω tending to infinity. For this purpose we will perform the Fourier transform discussed in Section B.3. Transforming equations (4.48) with t and ω being the variables of transformation, we get

$$(\iota\omega)^2\rho(\mathbf{x})\tilde{u}_i(\mathbf{x},\omega)=\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3}\left(\frac{\partial c_{ijkl}(\mathbf{x})}{\partial x_j}\frac{\partial \tilde{u}_k}{\partial x_l}+c_{ijkl}(\mathbf{x})\frac{\partial^2 \tilde{u}_k}{\partial x_j\partial x_l}\right), \tag{4.49}$$

where

$$\tilde{u}_i(\mathbf{x},\omega):=\frac{1}{2\pi}\int_{-\infty}^{\infty}u_i(\mathbf{x},t)\exp(-i\omega t)\,\mathrm{d}t,$$

and $i\in\{1,2,3\}$.

Following results of Sections 4.5 and4.7, we would like to write the solutions of the eikonal and transport equations for the Cauchy equations of motion. To do so, we first express equations (4.49) using notation form equation (4.8), namely

$$\frac{1}{(\iota\omega)^2}\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3}\left(\frac{\partial c_{ijkl}(\mathbf{x})}{\partial x_j}\frac{\partial \tilde{u}_k}{\partial x_l}+c_{ijkl}(\mathbf{x})\frac{\partial^2 \tilde{u}_k}{\partial x_j\partial x_l}\right)-\rho(\mathbf{x})\tilde{u}_i(\mathbf{x},\omega)=0$$

which can be written as

$$\left(\frac{1}{(\iota\omega)^2}D_2+D_0\right)\tilde{u}=0,$$

where

$$(D_2)_{il}\tilde{u}_l=\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3}\left(\frac{\partial c_{ijkl}(\mathbf{x})}{\partial x_j}\frac{\partial}{\partial x_k}+c_{ijkl}(\mathbf{x})\left(\frac{\partial^2}{\partial x_j\partial x_k}\right)\right)\tilde{u}_l$$
$$(D_0)_{il}\tilde{u}_l=-\rho\delta_{il}\tilde{u}_l.$$

Using the general for of the eikonal equation, (4.16), we can write it for the Cauchy equations of motion as

$$\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3}c_{ijkl}(\mathbf{x})\frac{\partial\psi}{\partial x_j}\frac{\partial\psi}{\partial x_k}-\rho\delta_{il}A_{0l}=0.$$

The solutions of the eikonal equation is discussed in Section 3.3.1.

Similarly, using the general form of the zero order transport equation, (4.22), we can write is in this case as

$$\left(\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3} c_{ijkl}\left(\mathbf{x}\right)\frac{\partial\psi}{\partial x_j}\frac{\partial\psi}{\partial x_k}-\rho\delta_{il}\right)A_{1l}+$$
$$\left(\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3}\frac{\partial c_{ijkl}\left(\mathbf{x}\right)}{\partial x_j}\frac{\partial\psi}{\partial x_k}+\right.$$
$$\frac{1}{2}\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3} c_{ijkl}\left(\mathbf{x}\right)\frac{\partial^2\psi}{\partial x_j\partial x_k}+$$
$$\left.\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3} c_{ijkl}\left(\mathbf{x}\right)\left(\frac{\partial\psi}{\mathrm{d}x_j}\right)\frac{\partial}{\partial x_k}\right)A_{0l}\left(x\right)=0.$$

To find the solutions of the transport equation, we assume that A_1 satisfies the eikonal equation and we follow the general solution given by expression (4.45), namely

$$A_0\left(x\left(t\right)\right)=A_0\left(x\left(t_0\right)\right)\exp\left\{-\int_{t_0}^{t}\zeta\sum_{k\leq m}\sum_{|\alpha|=k-1}B_\alpha^k\left(x\right)\left(\frac{\partial\psi}{\partial x}\right)^{|\alpha|}\mathrm{d}s\right\}\sqrt{\frac{J_0\zeta\left(x\left(t\right)\right)}{J_t\zeta\left(x\left(t_0\right)\right)}}.$$

We rewrite this expression using the particular B of the elastodynamic equations, namely

$$A_0\left(x\left(t\right)\right)=\sqrt{\frac{J_0\zeta\left(x\left(t\right)\right)}{J_t\zeta\left(x\left(t_0\right)\right)}}\exp\left\{-\int_{t_0}^{t}\zeta\frac{\partial c_{ijkl}}{\partial x_j}\frac{\partial\psi}{\partial x_k}\mathrm{d}s\right\}A_0\left(x\left(t_0\right)\right).$$

Closing remarks

Exercises

Exercise 4.1. Show that conditions (4.1) and (4.2) give equivalent definitions of the existence of infinite asymptotic expansions.

Solution 4.1. Condition (4.1), namely,

$$\lim_{x\to x_0}\frac{f\left(x\right)-S_n\left(x\right)}{\left(x-x_0\right)^n}=0,\quad\forall n\in\mathbb{N}$$

implies that

$$\lim_{x\to x_0}\left|\frac{f(x)-S_n(x)}{(x-x_0)^n}\right|<K\left|x-x_0\right|,\quad\forall n\in\mathbb{N},$$

for any positive constant K. Thus condition (4.1) implies condition (4.2), namely,

$$\lim_{x\to x_0}\left|\frac{f\left(x\right)-S_n\left(x\right)}{\left(x-x_0\right)^{n+1}}\right|<K,\quad\forall n\in\mathbb{N}.$$

Condition (4.2) implies that

$$f(x)-S_n(x)=O\left((x-x_0)^{n+1}\right)\Rightarrow$$
$$\lim_{x\to x_0}\frac{f(x)-S_n(x)-a_{n+1}(x-x_0)^{n+1}}{(x-x_0)^{n+1}}=\lim_{x\to x_0}\frac{f(x)-S_n(x)}{(x-x_0)^{n+1}}-a_{n+1}=0,$$

which means that

$$\lim_{x\to x_0}\frac{f(x)-S_n(x)}{(x-x_0)^{n}}=\lim_{x\to x_0}a_{n+1}(x-x_0)=0,$$

as desired.

Exercise 4.2. Evaluate

$$\lim_{x\to\infty}x\int_x^\infty t^{-1}e^{-t}\mathrm{d}t.$$

Solution 4.2. We rewrite this limit as

$$\lim_{x\to\infty}\frac{\int_x^\infty t^{-1}e^{-t}\mathrm{d}t}{x^{-1}}. \tag{4.50}$$

To evaluate this limit using the de l'Hôpital rule, we have to show that $\lim_{x\to\infty}\int_x^\infty t^{-1}e^{-t}\mathrm{d}t=0$. This is true since the integral is bounded for any $x>0$. Returning to limit (4.50), we write

$$\lim_{x\to\infty}\frac{\int_x^\infty t^{-1}e^{-t}\mathrm{d}t}{x^{-1}}=\lim_{x\to\infty}\frac{-x^{-1}e^{-x}}{-x^{-2}}=0.$$

Exercise 4.3. Show that

$$\phi_n=e^{-\omega^m\psi(x)}\frac{1}{(-\omega^2)^n}$$

is also a good asymptotic sequence for asymptotic expansion of solutions of equation (4.5).

Solution 4.3. Solution ?

Exercise 4.4. Verify that the fourth derivative of a composed function satisfies the following general chain rule:

$$\frac{\mathrm{d}^k}{\mathrm{d}x^k}f\circ g\equiv\frac{\mathrm{d}^k}{\mathrm{d}x^k}f\left(g\left(x\right)\right)=\sum\frac{k!}{l_1!l_2!\cdots l_k!}f^{(l_1+l_2+\cdots+l_k)}\circ g\left(\frac{g'}{1!}\right)^{l_1}\left(\frac{g''}{2!}\right)^{l_2}\cdots\left(\frac{g^{(k)}}{k!}\right)^{l_k}, \tag{4.51}$$

where the sum is over all nonnegative integer solutions of

$$l_1+2l_2+\cdots+kl_k=k.$$

Solution 4.4. The fourth derivative can be written as

$$\begin{aligned}\frac{\mathrm{d}^4}{\mathrm{d}x^4}f\circ g&=\frac{\mathrm{d}^3}{\mathrm{d}x^3}f'\circ g\cdot g'=\frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(f''\circ g\cdot g'^2+f'\circ g\cdot g''\right)\\&=\frac{\mathrm{d}}{\mathrm{d}x}\left(f'''\circ g\cdot g'^3+f''\circ g\cdot 2g'g''+f''\circ g\cdot g'g''+f'\circ g\cdot g'''\right)\\&=f''''\circ g\cdot g'^4+6f'''\circ g\cdot g'^2g''+3f''\circ g\cdot g''^2+4f''\circ g\cdot g'g'''+f'\circ gg''''.\end{aligned} \tag{4.52}$$

To verify that this derivative satisfies the general formula, first we have to find all the solutions for

$$l_1 + 2l_2 + 3l_3 + 4l_4 = 4.$$

We see that l_4 can be either one or zero.

If $l_4 = 1$, then $l_1 = l_2 = l_3 = 0$.

If $l_4 = 0$, then either $l_3 = 0$ or $l_3 = 1$. If $l_3 = 1$, then $l_1 = 1$. If $l_3 = 0$, then $l_2 = 2$, $l_2 = 1$ or $l_2 = 0$. If $l_2 = 2$ then $l_1 = 0$. If $l_2 = 1$, then $l_1 = 2$. If $l_2 = 0$, then $l_1 = 4$.

Thus, quadruple (l_1, l_2, l_3, l_4) can have values $(0,0,0,1)$, $(1,0,1,0)$, $(0,2,0,0)$, $(2,1,0,0)$, $(4,0,0,0)$. Substituting these quadruples into the general formula, we obtain

$$\frac{4!}{1!} f' \circ g \left(\frac{g''''}{4!}\right) + \frac{4!}{1!1!} f'' \circ g \left(\frac{g'}{1!}\right)\left(\frac{g'''}{3!}\right) + \frac{4!}{2!} f'' \circ g \left(\frac{g''}{2!}\right)^2 + \frac{4!}{2!1!} f''' \circ g \left(\frac{g'}{1!}\right)^2 \left(\frac{g''}{2!}\right) + \frac{4!}{4!} f'''' \circ g \left(\frac{g'}{1!}\right)^4,$$

which is equal to expression (4.52).

Exercise 4.5. Find the term containing the highest power of $i\omega$ in

$$\frac{\partial^{|\alpha|}}{\partial x^\alpha} \exp\left\{i\omega\psi\left(x\right)\right\},$$

where α is a general multiindex.

Solution 4.5. This derivative is a short-hand expression for

$$\frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \frac{\partial^{\alpha_2}}{\partial x_2^{\alpha_2}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} \exp\left\{i\omega\psi\left(x\right)\right\}.$$

Using expression (4.51) in Exercise 4.4, we can write this derivative as

$$\frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_{n-1}}}{\partial x_{n-1}^{\alpha_{n-1}}} \sum \frac{\alpha_n!}{l_1! l_2! \cdots l_{\alpha_n}!} (i\omega)^{l_1 + l_2 + \cdots + l_{\alpha_n}} \exp\left\{i\omega\psi\right\} \left(\frac{\frac{\partial\psi}{\partial x_n}}{1!}\right)^{l_1} \left(\frac{\frac{\partial^2\psi}{\partial x_n^2}}{2!}\right)^{l_2} \cdots \left(\frac{\frac{\partial^{\alpha_n}\psi}{\partial x_n^{\alpha_n}}}{\alpha_n!}\right)^{l_{\alpha_n}}.$$

Herein, the highest power of $(i\omega)$ results from $(l_1, l_2, \ldots, l_{\alpha_n}) = (\alpha_n, 0, \ldots, 0)$. Considering only this term of the sum, we write

$$\frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_{n-1}}}{\partial x_{n-1}^{\alpha_{n-1}}} (i\omega)^{\alpha_n} \exp\left\{i\omega\psi\right\} \left(\frac{\partial\psi}{\partial x_n}\right)^{\alpha_n}.$$

Differentiating this term with respect to x_{n-1}, we would get mixed terms resulting from the product rule. Since we are interested only in the highest power of $i\omega$, we consider only the part of the product rule in which we differentiate the exponential α_{n-1} times. Using the chain rule, we write the above expression as

$$\frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_{n-2}}}{\partial x_{n-2}^{\alpha_{n-2}}} \sum \frac{\alpha_{n-1}!}{l_1! l_2! \cdots l_{\alpha_{n-1}}!} (i\omega)^{l_1 + l_2 + \cdots + l_{\alpha_{n-1}}} \exp\left\{i\omega\psi\right\} \left(\frac{\frac{\partial\psi}{\partial x_{n-1}}}{1!}\right)^{l_1} \left(\frac{\frac{\partial^2\psi}{\partial x_{n-1}^2}}{2!}\right)^{l_2} \cdots \left(\frac{\frac{\partial^{\alpha_{n-1}}\psi}{\partial x_{n-1}^{\alpha_{n-1}}}}{\alpha_{n-1}!}\right)^{l_{\alpha_n}} (i\omega)^{\alpha_n} \left(\frac{\partial\psi}{\partial x_n}\right)^{\alpha_n}.$$

Again, taking into account only the highest power of $i\omega$, we consider only $(l_1, l_2, \ldots, l_{\alpha_{n-1}}) = (\alpha_{n-1}, 0, \ldots, 0)$, which yields

$$\frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_{n-2}}}{\partial x_{n-2}^{\alpha_{n-2}}} (i\omega)^{\alpha_{n-1}} (i\omega)^{\alpha_n} \exp\{i\omega\psi\} \left(\frac{\partial \psi}{\partial x_{n-1}}\right)^{\alpha_{n-1}} \left(\frac{\partial \psi}{\partial x_n}\right)^{\alpha_n}.$$

Following this pattern, we obtain the term with the highest power of $i\omega$, namely,

$$(i\omega)^{\alpha_1+\alpha_2+\cdots+\alpha_n} \exp\{i\omega\psi\} \left(\frac{\partial \psi}{\partial x_1}\right)^{\alpha_1} \left(\frac{\partial \psi}{\partial x_2}\right)^{\alpha_2} \cdots \left(\frac{\partial \psi}{\partial x_n}\right)^{\alpha_n}.$$

Exercise 4.6. Prove the generalized Leibniz rule for multiindex α, namely,

$$\frac{\partial^{|\alpha|}}{\partial x^\alpha} (f \cdot g) = \sum_{\beta+\gamma=\alpha} \frac{\alpha!}{\beta!\gamma!} \frac{\partial^{|\beta|} f}{\partial x^\beta} \frac{\partial^{|\gamma|} g}{\partial x^\gamma},$$

where $\alpha! = \alpha_1!\alpha_2! \ldots \alpha_n!$.

Solution 4.6. The left-hand side of the generalized Leibniz rule can be written as

$$\frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \frac{\partial^{\alpha_2}}{\partial x_2^{\alpha_2}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} (f \cdot g) = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \frac{\partial^{\alpha_2}}{\partial x_2^{\alpha_2}} \cdots \frac{\partial^{\alpha_{n-1}}}{\partial x_{n-1}^{\alpha_{n-1}}} \sum_{\beta_n+\gamma_n=\alpha_n} \frac{\alpha_n!}{\beta_n!\gamma_n!} \frac{\partial^{\beta_n} f}{\partial x_n^{\beta_n}} \frac{\partial^{\gamma_n} g}{\partial x_n^{\gamma_n}},$$

where we used the standard Leibniz rule, namely,

$$\frac{\mathrm{d}^k}{\mathrm{d}x^k} (f \cdot g) = \sum_{l+m=k} \frac{k!}{l!m!} \frac{\mathrm{d}^l f}{\mathrm{d}x^l} \frac{\mathrm{d}^m g}{\mathrm{d}x^m}.$$

Continuing the differentiation, we obtain

$$\sum_{\beta_1+\gamma_1=\alpha_1} \sum_{\beta_2+\gamma_2=\alpha_2} \cdots \sum_{\beta_n+\gamma_n=\alpha_n} \frac{\alpha_1!}{\beta_1!\gamma_1!} \frac{\alpha_2!}{\beta_2!\gamma_2!} \cdots \frac{\alpha_n!}{\beta_n!\gamma_n!} \frac{\partial^{\beta_1}\partial^{\beta_2}\cdots\partial^{\beta_n} f}{\partial x_1^{\beta_1}\partial x_2^{\beta_2}\cdots\partial x_n^{\beta_n}} \frac{\partial^{\gamma_1}\partial^{\gamma_2}\cdots\partial^{\gamma_n} g}{\partial x_1^{\gamma_1}\partial x_2^{\gamma_2}\cdots\partial x_n^{\gamma_n}}.$$

This expression can be written using the multiindex notation as the right-hand side of the generalized Leibniz rule,

$$\sum_{\beta+\gamma=\alpha} \frac{\alpha!}{\beta!\gamma!} \frac{\partial^{|\beta|} f}{\partial x^\beta} \frac{\partial^{|\gamma|} g}{\partial x^\gamma}.$$

Exercise 4.7. Following the general formulation on page 86, derive eikonal equation (4.18) from the reduced wave equation,

$$\frac{1}{(i\omega)^2} \sum_{i=1}^{3} \frac{\partial^2}{\partial x_i^2} u(x, \omega) - \frac{1}{v^2} u(x, \omega) = 0.$$

Solution 4.7. Following the discussion in Section 4.2, we consider an asymptotic solution of the form

$$u(x, \omega) = e^{i\omega\psi(x)} \sum_{j=0}^{\infty} A_j(x) \frac{1}{(i\omega)^j}.$$

Using this expression in the reduced wave equation, we write

$$\left(\frac{1}{(i\omega)^2}\sum_{i=1}^{3}\frac{\partial^2}{\partial x_i^2}-\frac{1}{v^2}\right)\left(e^{i\omega\psi(x)}\sum_{j=0}^{\infty}A_j\left(x\right)\frac{1}{(i\omega)^j}\right)=0.$$

Applying the product rule, we get

$$\frac{1}{(i\omega)^2}\left(\sum_{i=1}^{3}\left(\frac{\partial^2}{\partial x_i^2}e^{i\omega\psi(x)}\right)\sum_{j=0}^{\infty}A_j\left(x\right)\frac{1}{(i\omega)^j}+2\sum_{i=1}^{3}\frac{\partial}{\partial x_i}e^{i\omega\psi(x)}\frac{\partial}{\partial x_i}\sum_{j=0}^{\infty}A_j\left(x\right)\frac{1}{(i\omega)^j}+\right.$$
$$\left.e^{i\omega\psi(x)}\sum_{i=1}^{3}\frac{\partial^2}{\partial x_i^2}\sum_{j=0}^{\infty}A_j\left(x\right)\frac{1}{(i\omega)^j}\right)-\frac{1}{v^2}e^{i\omega\psi(x)}\sum_{j=0}^{\infty}A_j\left(x\right)\frac{1}{(i\omega)^j}=0.$$

Differentiating the exponentials, we get

$$\frac{e^{i\omega\psi(x)}}{(i\omega)^2}\left(\sum_{i=1}^{3}\left((i\omega)^2\left(\frac{\partial\psi}{\partial x_i}\right)^2+(i\omega)\frac{\partial^2\psi}{\partial x_i^2}\right)\sum_{j=0}^{\infty}A_j\frac{1}{(i\omega)^j}+2\sum_{i=1}^{3}(i\omega)\frac{\partial\psi}{\partial x_i}\sum_{j=0}^{\infty}\frac{\partial A_j}{\partial x_i}\frac{1}{(i\omega)^j}+\right.$$
$$\left.\sum_{i=1}^{3}\sum_{j=0}^{\infty}\frac{\partial^2 A_j}{\partial x_i^2}\frac{1}{(i\omega)^j}\right)-\frac{1}{v^2}e^{i\omega\psi(x)}\sum_{j=0}^{\infty}A_j\frac{1}{(i\omega)^j}=0 \quad (4.53)$$

The terms with the zeroth power of $(i\omega)$ in this asymptotic series are

$$e^{i\omega\psi(x)}A_0\left(x\right)\left(\sum_{i=1}^{3}\left(\frac{\partial\psi}{\partial x_i}\right)^2-\frac{1}{v^2}\right).$$

Viewing the right-hand side of equation (4.53) as an asymptotic series and equating the coefficients of the zeroth power of $(i\omega)$, we obtain

$$A_0\left(x\right)\left(\sum_{i=1}^{3}\left(\frac{\partial\psi}{\partial x_i}\right)^2-\frac{1}{v^2}\right)=0.$$

If we consider only nonzero functions A_0, then this equation reduces to the required equation,

$$\sum_{i=1}^{3}\left(\frac{\partial\psi}{\partial x_i}\right)^2=\frac{1}{v^2}.$$

Exercise 4.8. Following equation (4.17), show that the zeroth-order approximation of the solution of the three-dimensional reduced wave equation,

$$\frac{1}{(i\omega)^2}\nabla^2 u\left(x,\omega\right)-\frac{1}{v^2}u\left(x,\omega\right)=0,$$

results in the eikonal equation,

$$\sum_{i=1}^{3}\left(\frac{\partial\psi}{\partial x_i}\right)^2=\frac{1}{v^2}.$$

Solution 4.8. Considering the general form of an asymptotic differential equation given by expression (4.12), we see that the coefficients $B_\alpha\left(x\right)$ of the reduced wave equation are

$$
\begin{aligned}
B_{(2,0,0)} &= 1, \\
B_{(0,2,0)} &= 1, \\
B_{(0,0,2)} &= 1, \\
B_{(0,0,0)} &= -\frac{1}{v^2}.
\end{aligned}
$$

Substituting these coefficients into equation (4.17), we obtain the required equation,

$$
\sum_i \left(\frac{\partial \psi}{\partial x_i} \right)^2 - \frac{1}{v^2} = 0.
$$

Exercise 4.9. Show that in the case of the reduced wave equation, namely,

$$
\frac{v^2}{(i\omega)^2} \nabla^2 u(x, \omega) = u(x, \omega). \tag{4.54}
$$

equation (4.22) reduces to equation (4.24).

Solution 4.9. As stated in Exercise 4.8, the coefficients B for the reduced wave equation are

$$
\begin{aligned}
B_{(2,0,0)} &= 1, \\
B_{(0,2,0)} &= 1, \\
B_{(0,0,2)} &= 1, \\
B_{(0,0,0)} &= -\frac{1}{v^2}.
\end{aligned}
$$

Substituting these coefficients into equation (4.22), we obtain

$$
\left(\sum_i \left(\frac{\partial \psi}{\partial x_i} \right)^2 - \frac{1}{v^2} \right) A_1(x) + \left(\sum_i \frac{\partial^2 \psi}{\partial x_i^2} + \sum_i 2 \frac{\partial \psi}{\partial x_i} \frac{\partial}{\partial x_i} \right) A_0(x) = 0, \tag{4.55}
$$

which is the required equation.

The reduced wave equation (4.54) contains only the derivatives of order of power of $(i\omega)^{-1}$, namely, the zeroth and second derivatives. This means that only the principal symbol appears in transport equation (4.55). In particular, $\sigma(D_{k-1}) = 0$.

Exercise 4.10. Using the symbols of the differential operators discussed in Chapter 6, rewrite equation (4.22).

Solution 4.10. The symbol of the differential operator D is

$$
\sigma(D)(p) = \sum_{k \leq m} \sigma(D_k)(p),
$$

where

$$
\sigma(D_k)(p) = \sum_{|\alpha| = k} B_\alpha(x) p^\alpha.
$$

Letting $p = \mathrm{d}\psi$, which we can write in components as

$$(p_1, p_2, \ldots, p_n) = \left(\frac{\partial \psi}{\partial x_1}, \frac{\partial \psi}{\partial x_2}, \ldots, \frac{\partial \psi}{\partial x_n} \right),$$

we write

$$\sigma(D)(\mathrm{d}\psi) = \sum_{k \le m} \sum_{|\alpha|=k} B_\alpha(x) \left(\frac{\partial \psi}{\partial x} \right)^{|\alpha|}.$$

Similarly,

$$\sum_{k \le m} \sigma(D_{k-1})(\mathrm{d}\psi) = \sum_{k \le m} \sum_{|\alpha|=k-1} B_\alpha(x) \left(\frac{\partial \psi}{\partial x} \right)^{|\alpha|}.$$

Differentiation of the symbol with respect to its argument yields

$$\sum_{k \le m} \left. \frac{\partial \sigma(D_k)(p)}{\partial p_i} \right|_{p=\mathrm{d}\psi} = \sum_{k \le m} \sum_{|\alpha|=m} B_\alpha \alpha_i \left(\frac{\partial \psi}{\mathrm{d}x} \right)^{\alpha - (0,\ldots,0,1,0,\ldots,0)},$$

where 1 in $(0, \ldots, 0, 1, 0, \ldots, 0)$ appears in the ith location. Similarly, considering the second derivatives of the symbol, we obtain

$$\sum_{k \le m} \left. \frac{\partial^2 \sigma(D_k)(p)}{\partial p_i \partial p_j} \right|_{p=\mathrm{d}\psi} = \sum_{k \le m} \sum_{|\alpha|=k} B_\alpha \left(\alpha_i \alpha_j (1 - \delta_{ij}) + \alpha_i (\alpha_j - 1) \delta_{ij} \right) \frac{\partial^2 \psi}{\partial x_i \partial x_j} \left(\frac{\partial \psi}{\partial x} \right)^{\alpha - (0,\ldots,0,1,0,\ldots,0,1,0,\ldots,0)},$$

where 1 in multiindex $(0, \ldots, 0, 1, 0, \ldots, 0, 1, 0, \ldots, 0)$ appears in the ith and jth locations. Summing these four results, we obtain the desired expression, namely,

$$\sigma(D)(\mathrm{d}\psi) A_1(x) + \left(\sum_{k \le m} \sigma(D_{k-1})(\mathrm{d}\psi) + \sum_{k \le m} \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 \sigma(D_k)(p)}{\partial p_i \partial p_j} \frac{\partial^2 \psi}{\partial x_i \partial x_j} + \sum_{k \le m} \sum_{i=1}^{n} \frac{\partial \sigma(D_k)(p)}{\partial p_i} \frac{\partial}{\partial x_i} \right) A_0(x) = 0,$$

where $p = \mathrm{d}\psi$.

Exercise 4.11. Write the solution of

$$\frac{\mathrm{d}}{\mathrm{d}t} \ln A_0(x(t)) = -\zeta \left(\sum_{k \le m} \sigma(D_{k-1})(\mathrm{d}\psi) + \sum_{k \le m} \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 \sigma(D_k)(p)}{\partial p_i \partial p_j} \frac{\partial^2 \psi}{\partial x_i \partial x_j} \right).$$

Solution 4.11. Integrating both sides with respect to t between 0 and t, we get

$$\ln A_0(x(t)) - \ln A_0(x(0)) = \int_0^t -\zeta \left(\sum_{k \le m} \sigma(D_{k-1})(\mathrm{d}\psi) + \sum_{k \le m} \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 \sigma(D_k)(p)}{\partial p_i \partial p_j} \frac{\partial^2 \psi}{\partial x_i \partial x_j} \right) \mathrm{d}s,$$

which is

$$\ln \frac{A_0\left(x\left(t\right)\right)}{A_0\left(x\left(0\right)\right)} = \int_0^t -\zeta \left(\sum_{k\leq m} \sigma\left(D_{k-1}\right)\left(\mathrm{d}\psi\right) + \sum_{k\leq m} \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 \sigma\left(D_k\right)\left(p\right)}{\partial p_i \partial p_j} \frac{\partial^2 \psi}{\partial x_i \partial x_j} \right) \mathrm{d}s.$$

Exponentiating and multiplying by $A_0\left(x\left(0\right)\right)$ we obtain

$$A_0\left(x\left(t\right)\right) = A_0\left(x\left(0\right)\right) \int_0^t -\zeta \left(\sum_{k\leq m} \sigma\left(D_{k-1}\right)\left(\mathrm{d}\psi\right) + \sum_{k\leq m} \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 \sigma\left(D_k\right)\left(p\right)}{\partial p_i \partial p_j} \frac{\partial^2 \psi}{\partial x_i \partial x_j} \right) \mathrm{d}s.$$

Exercise 4.12. Applying Definition 4.5 to expression (4.4), we obtain

$$\lim_{\omega\to\infty} \left(\frac{\left(i\omega\right)^N}{\exp\left\{i\omega\psi\left(x\right)\right\}} \left(\tilde{u}\left(x,\omega\right) - \exp\left\{i\omega\psi\left(x\right)\right\} \sum_{n=0}^{N} \frac{A_n\left(x\right)}{\left(i\omega\right)^n} \right) \right) = 0, \tag{4.56}$$

allows us to determine uniquely all A_n. Determine A_0, A_1 and A_2.

Solution 4.12. For $N = 0$ also $n = 0$, and we write expression (4.56) as

$$\lim_{\omega\to\infty} \left(\frac{1}{\exp\left\{i\omega\psi\left(x\right)\right\}} \left(\tilde{u}\left(x,\omega\right) - \exp\left\{i\omega\psi\left(x\right)\right\} A_0\left(x\right) \right) \right) = 0,$$

which means that

$$A_0\left(x\right) = \lim_{\omega\to\infty} \frac{\tilde{u}\left(x,\omega\right)}{\exp\left\{i\omega\psi\left(x\right)\right\}}. \tag{4.57}$$

For $N = 1$, $n = 0, 1$, and we write expression (4.56) as

$$\lim_{\omega\to\infty} \left(\frac{i\omega}{\exp\left\{i\omega\psi\left(x\right)\right\}} \left(\tilde{u}\left(x,\omega\right) - \exp\left\{i\omega\psi\left(x\right)\right\} \left(A_0\left(x\right) + \frac{A_1\left(x\right)}{i\omega} \right) \right) \right) = 0.$$

We can rewrite it as

$$\lim_{\omega\to\infty} \left(i\omega \left(\frac{\tilde{u}\left(x,\omega\right)}{\exp\left\{i\omega\psi\left(x\right)\right\}} - \left(A_0\left(x\right) + \frac{A_1\left(x\right)}{i\omega} \right) \right) \right) = 0$$

to get

$$\lim_{\omega\to\infty} \left(i\omega \left(\frac{\tilde{u}\left(x,\omega\right)}{\exp\left\{i\omega\psi\left(x\right)\right\}} - A_0\left(x\right) \right) - A_1\left(x\right) \right) = 0.$$

This means that

$$A_1\left(x\right) = \lim_{\omega\to\infty} \left(i\omega \left(\frac{\tilde{u}\left(x,\omega\right)}{\exp\left\{i\omega\psi\left(x\right)\right\}} - A_0\left(x\right) \right) \right), \tag{4.58}$$

where A_0 is known from equation (4.57). For $N = 2$, $n = 1, 2, 3$, and we write expression (4.56) as

$$\lim_{\omega\to\infty} \left(\frac{\left(i\omega\right)^2}{\exp\left\{i\omega\psi\left(x\right)\right\}} \left(\tilde{u}\left(x,\omega\right) - \exp\left\{i\omega\psi\left(x\right)\right\} \left(A_0\left(x\right) + \frac{A_1\left(x\right)}{i\omega} + \frac{A_2\left(x\right)}{\left(i\omega\right)^2} \right) \right) \right) = 0.$$

We can rewrite it as

$$\lim_{\omega\to\infty} \left(\left(i\omega\right)^2 \left(\frac{\tilde{u}\left(x,\omega\right)}{\exp\left\{i\omega\psi\left(x\right)\right\}} - A_0\left(x\right) \right) - i\omega A_1\left(x\right) - A_2\left(x\right) \right),$$

which means that

$$A_2(x) = \lim_{\omega \to \infty} \left((i\omega)^2 \left(\frac{\tilde{u}(x,\omega)}{\exp\{i\omega\psi(x)\}} - A_0(x) \right) - i\omega A_1(x) \right), \tag{4.59}$$

where A_0 and A_1 are known from equations (4.57) and (4.58), respectively. Continuing this process we obtain uniquely all A_n.

Exercise 4.13. Find the general form of an asymptotic sequence for solving differential equations of type given by expression (4.5).

Solution 4.13. ?

5

Caustics

Preliminary remarks

Caustics are intrinsic entities of the asymptotic ray theory, and can be viewed as its limitation. Caustics correspond to points at which the amplitude of the signal is infinite – a nonphysical result. Yet, they provide insight into the physical phenomena of focusing, which is associated with high amplitudes.

We begin this chapter by relating the singularities of transport equations to caustics. Subsequently, we formulate caustics as the envelopes of characteristics in a manner similar to the one used in the context of the Monge cone. We conclude the chapter by discussing the phase shift of a signal that has passed through, or touched, the caustic.

5.1 Singularities of transport equation

In Section 4.7, we obtained the solution of transport equation (4.45). This solution is

$$A_0\left(x\left(t\right)\right) = A_0\left(x\left(0\right)\right) \exp\left\{ -\int_{t_0}^{t} \zeta \sum_{k \leq m} \sigma\left(D_{k-1}\right)\left(\mathrm{d}\psi\right) \mathrm{d}s \right\} \sqrt{\frac{J_0 \zeta\left(x\left(t\right)\right)}{J_t \zeta\left(x\left(0\right)\right)}}. \tag{5.1}$$

where A_0 denotes the amplitude, $\zeta = \left|\partial\sigma\left(D\right)\left(p\right)/\partial p_i\right|$, and J denotes the Jacobian of the change from the Cartesian to "ray" coordinates. In this chapter, we will study the singularities of this solution, namely, the points for which the Jacobian vanishes. Referring to the alternative expression for the solution given by formula (4.44), we see that the vanishing of the Jacobian is equivalent to the vanishing of the area of the raytube cross-section. Intuitively, we expect such vanishing to happen at points where the neighbouring characteristics converge. Such a convergence corresponds to an envelope of the characteristics, which is called a caustic, which we will discuss in the next section.

5.2 Caustics as envelopes of characteristics

In this section, we will show that caustics are the envelopes of the characteristics by constructing such an envelope. We will follow the construction analogous to the one discussed in Section 3.1 on page 58, where we described the Monge cone as the envelope of a family of planes.

Consider a hypersurface that is transverse to the characteristic curves, and on which we define the side conditions. Let us parametrize it by $\xi_1, \xi_2, \ldots, \xi_{n-1}$. The characteristics passing through this surface are denoted by $\gamma(\xi, t)$, where t denotes a parameter along the characteristic. The envelope is formed by intersections of characteristics that are infinitesimally close to each other. In other words, for a point, x, to be on the envelope, there must be characteristics, $\gamma(\xi, t)$, that pass through that point such that

$$x = \gamma(\xi, t) = \gamma(\xi + \mathrm{d}\xi, t),$$

where $\mathrm{d}\xi$ is an infinitesimally small vector in the $\xi_1\xi_2\ldots\xi_{n-1}$-space. This is a stationarity condition with respect to a direction on the hypersurface. This direction is given by a vector in the kernel of the derivative of map $\gamma : \xi_1, \xi_2, \ldots, \xi_{n-1}, t \to \mathbb{R}^n$. In other words, this vector corresponds to the zero eigenvalue of the derivative of the map $\gamma : \xi_1, \xi_2, \ldots, \xi_{n-1}, t \to \mathbb{R}^n$. This map is the coordinate change from the "ray" coordinates to the Cartesian ones. Thus, a point is on the envelope if for its coordinates, (ξ, t), the Jacobian matrix of γ is degenerate. At such points the Jacobian of γ is zero. In the preceding section, we saw that such points corresponds to caustics.

Let us give an example of caustics as ray envelopes in homogeneous media.

Example 5.1. To find the envelope of rays in a two-dimensional anisotropic homogeneous medium with wavefront velocity $v(N)$, we consider an initial wavefront S parametrized by its arclength ξ. Using the Huygens principle, the wavefront $S(\xi)$ can be propagated in time as

$$\gamma(\xi, t) = S(\xi) + N(\xi)\, v(N(\xi))\, t, \tag{5.2}$$

where $N(\xi)$ is the unit normal vector to the wavefront S at point $S(\xi)$. For the subsequent times t, we can view curves $\gamma(\xi, t)$ as wavefronts at those times. While these wavefronts are still parametrized by ξ, this is no longer the arclength parametrization, which means that the length of the tangent vector $\partial\gamma/\partial\xi$ is no longer equal to one, as can be seen from the fact that wavefronts can be "shrinking" or "expanding". The points where this vector vanishes are called the singularities of the wavefront. At such points the wavefront folds into itself, as illustrated in Exercise 5.4. These points form the envelope of the rays and must satisfy equation

$$\frac{\partial\gamma(\xi, t)}{\partial\xi} = 0,$$

which, in view of expression (5.2), can be written as

$$\frac{\partial S}{\partial\xi} + \frac{\partial N}{\partial\xi} v(N(\xi))\, t + N(\xi)\, \frac{\partial v}{\partial N}\frac{\partial N}{\partial\xi} t = 0.$$

Since $v(N)$ is zeroth-degree homogeneous in N, it follows from the homogeneous-function theorem that $N(\xi)\, \partial v/\partial N = 0$.[1] Hence, the above equation becomes

$$\frac{\partial S}{\partial\xi} + \frac{\partial N}{\partial\xi} v(N(\xi))\, t = 0.$$

[1] Readers interested in this theorem might refer to, e.g., Courant and Hilbert (1989, Vol. 2 p. 11)

Furthermore, due to the arclength parametrization of S by ξ, the equation becomes

$$l - \kappa l v t = 0,$$

where l is the unit tangent to S given by $\partial S/\partial \xi$ and κ is the curvature: the rate of change of the normal vector to the curve; herein, $\kappa = -\partial N/\partial \xi$. Thus, we conclude that the points on the envelope are given by

$$t = \frac{1}{\kappa v}.$$

In other words, the envelope is given by

$$\epsilon(\xi) = S(\xi) + \frac{1}{\kappa v\left(N(\xi)\right)} N(\xi). \tag{5.3}$$

Thus, the singular points of all wavefronts at different times, illustrated in Figure 5.2 below, belong to the envelope of rays emanating from the original wavefront.

The reader might be interested in expressing the above results in the language of classical differential geometry. Since the curvature is the reciprocal of the radius of curvature, $\kappa = 1/\rho$, we could write expression (5.3) as $\epsilon(\xi) = S(\xi) + \rho(\xi) n(\xi)$, which is the locus of the centres of curvatures of the initial wavefront, and is called the evolute. In other words, the evolute is the locus of the wavefront-osculating circles with radius ρ. Furthermore, the envelope of the normals to a curve is called an involute. Thus, a wavefront is the involute of the caustic, the rays are the wavefront normals, and the envelope of these rays is the evolute of the wavefront.

An example of constructing the envelopes in isotropic homogeneous media is given in Exercise 5.2. A particular example of such an envelope is discussed in Exercise 5.4.

5.3 Phase change on caustics

As indicated in Section 4.7, the solution of the transport equation might have a singularity, which is a caustic. We cannot continue this solution along the characteristics past the caustics. Below, we will show that the phase along a characteristic changes as the characteristic touches a caustic.

5.3.1 Waves in isotropic homogeneous media

Let us consider the wave equation in three spatial dimensions; namely,

$$\frac{\partial^2 u(x,t)}{\partial x_1^2} + \frac{\partial^2 u(x,t)}{\partial x_2^2} + \frac{\partial^2 u(x,t)}{\partial x_3^2} - \frac{1}{v^2}\frac{\partial^2 u(x,t)}{\partial t^2} = 0.$$

Expressing this equation in spherical coordinates; namely,

$$\begin{aligned} x_1 &= r\sin\phi\cos\vartheta, \\ x_2 &= r\sin\phi\sin\vartheta, \\ x_3 &= r\cos\vartheta, \end{aligned}$$

we write the wave equation as

$$\frac{\partial^2 u}{\partial r^2}+\frac{2}{r}\frac{\partial u}{\partial r}+\frac{1}{r^2}\left(\frac{\partial^2 u}{\partial \phi^2}+\cot\phi\frac{\partial u}{\partial \phi}+\frac{1}{\sin^2\phi}\frac{\partial^2 u}{\partial \vartheta^2}\right)-\frac{1}{v^2}\frac{\partial^2 u}{\partial t^2}=0.$$

Since v is a constant, the wave equation is spherically symmetric. If we consider the side conditions that are symmetric as well, the solutions do not depend on either ϕ or ϑ. Thus the wave equation becomes

$$\frac{\partial^2 u(r,t)}{\partial r^2}+\frac{2}{r}\frac{\partial u(r,t)}{\partial r}-\frac{1}{v^2}\frac{\partial^2 u(r,t)}{\partial t^2}=0. \tag{5.4}$$

To find solutions of this equation, we let $u(r,t)=w(r,t)/r$, which, after taking the derivatives, leads to

$$\frac{\partial^2 w}{\partial r^2}-\frac{1}{v^2}\frac{\partial^2 w}{\partial t^2}=0.$$

Formally, in a manner analogous to the d'Alembert solution, we write

$$w(r,t)=F(r+vt)+G(r-vt),$$

where functions F and G are constrained by the side conditions. Since $w(r,t)=ru(r,t)$, we write the solution of equation (5.4) as

$$u(r,t)=\frac{1}{r}\left(F(r+vt)+G(r-vt)\right).$$

To discuss this solution, let us express it as the Fourier transform with t and ω being the transformation variables. We write

$$\tilde{u}(r,\omega)=\frac{1}{r}\int_{-\infty}^{\infty}\left(F(r+vt)+G(r-vt)\right)e^{-\iota\omega t}\mathrm{d}t.$$

Substituting $r\pm vt=s$, which implies that $\mathrm{d}t=\mp\mathrm{d}s/v$, we rewrite it as

$$\tilde{u}(r,\omega)=\frac{1}{vr}\int_{-\infty}^{\infty}\left[F(s)\exp\left\{-\iota\omega\frac{s-r}{v}\right\}+G(s)\exp\left\{-\iota\omega\frac{r-s}{v}\mathrm{d}s\right\}\right].$$

We can split the integral into the sum of two integrals and factor out the terms that do not depend on s to get

$$\tilde{u}(r,\omega)=\frac{1}{vr}\left(\exp\left\{\iota\omega\frac{r}{v}\right\}\int_{-\infty}^{\infty}F(s)\exp\left\{-\iota\frac{\omega}{v}s\right\}\mathrm{d}s+\exp\left\{-\iota\omega\frac{r}{v}\right\}\int_{-\infty}^{\infty}G(s)\exp\left\{-\iota\left(-\frac{\omega}{v}\right)s\right\}\mathrm{d}s\right).$$

The two integrals are the Fourier transforms. For the first one, the transformation changes variable s to ω/v, and for the second one, s to $-\omega/v$. Thus, we write

$$\tilde{u}(r,\omega)=\frac{1}{vr}\left(\exp\left\{\iota\omega\frac{r}{v}\right\}\tilde{F}\left(\frac{\omega}{v}\right)+\exp\left\{-\iota\omega\frac{r}{v}\right\}\tilde{G}\left(-\frac{\omega}{v}\right)\right).$$

Let us consider the solution with $\tilde{F}=0$ and $\tilde{G}$ being a constant; namely,

$$\tilde{u}(r,\omega) = \frac{1}{vr}\exp\left\{-\iota\omega\frac{r}{v}\right\}\tilde{G}.$$

To see the meaning of this solution, we take the inverse Fourier transform to write

$$u(r,t) = \frac{A}{r}\delta\left(t-\frac{r}{v}\right),$$

which describes a spherical wavefront propagating from the point source at the origin.

If we consider a surface, S, from which the signal is propagating, we can construct the corresponding solution by adding the point sources from each location, y, of S. Thus we can write,

$$\tilde{u}(x,\omega) = \int_S \frac{e^{-\iota\omega|x-y|}}{|x-y|}c(y)\,\mathrm{d}y, \tag{5.5}$$

where $c(y)$ defines the amplitude of the wave originating at point y.

5.3.2 Method of stationary phase

In this section, we will introduce a method to evaluate integral (5.5) in the context of ray theory. To do so, we will study the behaviour for large ω of an integral exhibiting a more general form; namely,

$$I = \int_S a(y)\,e^{\iota\omega\psi(y)}\mathrm{d}y. \tag{5.6}$$

To evaluate this integral, we consider the vector field given by

$$X = \sum_{i=1}^{m}\frac{\partial\psi}{\partial y_i}\frac{\partial}{\partial y_i},$$

where m is the dimension of the surface S. The action of X on the exponential term in expression (5.6) is

$$Xe^{\iota\omega\psi} = \sum_{i=1}^{m}\frac{\partial\psi}{\partial y_i}\frac{\partial}{\partial y_i}e^{\iota\omega\psi} = \iota\omega p^2 e^{\iota\omega\psi},$$

where $p = \mathrm{d}\psi$. If $p^2 \neq 0$, we solve for the exponential term to get

$$e^{\iota\omega\psi} = \frac{Xe^{\iota\omega\psi}}{\iota\omega p^2}.$$

Inserting this result into equation (5.6), we get

$$I = \frac{1}{\iota\omega}\int_S \frac{a(y)}{p^2}Xe^{\iota\omega\psi}\mathrm{d}y.$$

Using the product rule, we obtain

$$I = \frac{1}{\iota\omega}\int_S X\left(\frac{a(y)}{p^2}e^{\iota\omega\psi}\right) - X\left(\frac{a(y)}{p^2}\right)e^{\iota\omega\psi}\mathrm{d}y = \frac{1}{\iota\omega}\left(\int_S X\left(\frac{a(y)}{p^2}e^{\iota\omega\psi}\right)\mathrm{d}y - \int_S X\left(\frac{a(y)}{p^2}\right)e^{\iota\omega\psi}\mathrm{d}y\right).$$

Let us consider the first integral. If we can express it as the integral over the integral lines of the vector field X and the integral over the quotient S/X, then the integral over the integral lines of X of the directional derivative $X(f)$ for a function f that vanishes at both "ends" of the integral lines of X this integral vanishes by the fundamental theorem of calculus. Thus,

$$I = -\frac{1}{\iota\omega}\int\limits_S X\left(\frac{a(y)}{p^2}\right)e^{\iota\omega\psi}\mathrm{d}y.$$

Letting $b_1(y) := X\left(a(y)/p^2\right)$, we write

$$I = -\frac{1}{\iota\omega}\int\limits_S b_1(y)\,e^{\iota\omega\psi}\mathrm{d}y.$$

Repeating the process of making the substitution for the exponential term and integrating by parts n times, we obtain

$$I = \left(-\frac{1}{\iota\omega}\right)^n\int\limits_S b_n(y)\,e^{\iota\omega\psi}\mathrm{d}y.$$

If all b_i are compactly supported and continuous, all of the integrals are bounded. Thus, we can say that

$$I = O\left(\frac{1}{(-\iota\omega)^n}\right). \tag{5.7}$$

This means that if $\mathrm{d}\psi \neq 0$ then I vanishes as fast as any negative power of ω, as $\omega \to \infty$. Thus, for large ω, the only nonzero contribution to I can come from points where $\mathrm{d}\psi = 0$. This property justifies the name of the discussed method for evaluation of such integrals for large ω as the method of stationary phase. Thus, in the next section we will evaluate I for all stationary values of $\psi(y)$.

To evaluate for stationary values of ψ, we invoke the Morse Lemma.

Lemma 1. *If $\mathrm{d}\psi(y_0) = 0$ and $\partial^2\psi/\partial y_i\partial y_j$ is nondegenerate, then there exist coordinates $z_1,\ldots,z_m$ in a neighbourhood of y_0 such that $\psi = \psi(y_0) + \frac{1}{2}\left(z_1^2 + z_2^2 + \ldots + z_l^2 - z_{l+1}^2 - \ldots - z_m^2\right)$.*

If view of this lemma , we express integral (5.6) in terms of coordinates z_i as

$$I = \int\limits_S a(z)\exp\left\{i\omega\left(\psi(y_0) + \frac{1}{2}\left(z_1^2 + z_2^2 + \ldots + z_l^2 - z_{l+1}^2 - \ldots - z_m^2\right)\right)\right\}\det\left[\frac{\partial y}{\partial z}\right]\mathrm{d}z,$$

where $\det[\partial y/\partial z]$ is the Jacobian. Factoring out the term independent of z_i, we write

$$I = e^{i\omega\psi(y_0)}\int\limits_S a(z)\exp\left\{\frac{i\omega}{2}\left(z_1^2 + z_2^2 + \ldots + z_l^2 - z_{l+1}^2 - \ldots - z_m^2\right)\right\}\det\left[\frac{\partial y}{\partial z}\right]\mathrm{d}z.$$

We express $a(z)\det[\partial y/\partial z]$ as

$$a(z)\det\left[\frac{\partial y}{\partial z}\right] = b_0 + b_1(z)\,z_1 + b_2(z)\,z_2 + \ldots + b_l(z)\,z_l - b_{l+1}(z)\,z_{l+1} - \ldots - b_m(z)\,z_m,$$

for some functions $b_i(z)$. Denoting $Q(z) := z_1^2 + z_2^2 + \ldots + z_l^2 - z_{l+1}^2 - \ldots - z_m^2$, we write the integral as

$$I = e^{i\omega\psi(y_0)}\left[b_0\int\limits_S e^{\frac{\iota\omega}{2}Q(z)}\mathrm{d}z + \int\limits_S\left(\sum_{i=1}^{l} b_i\left(z\right)z_i - \sum_{i=l+1}^{m} b_i\left(z\right)z_i\right)e^{\frac{\iota\omega}{2}Q(z)}\mathrm{d}z\right],$$

which can be written as

$$I = e^{i\omega\psi(y_0)}\left[b_0\int\limits_S e^{\frac{\iota\omega}{2}Q(z)}\mathrm{d}z + \frac{1}{i\omega}\int\limits_S\sum_{i=1}^{m} b_i\left(z\right)\frac{\partial}{\partial z_i}e^{\frac{\iota\omega}{2}Q(z)}\mathrm{d}z\right].$$

Following the method used to arrive at equation (5.7), we obtain

$$I = e^{i\omega\psi(y_0)}b_0\int\limits_S e^{\frac{i\omega}{2}Q(z)}\mathrm{d}z + O\left(\frac{1}{(i\omega)^n}\right),$$

which means that

$$I \sim e^{i\omega\psi(y_0)}b_0\int\limits_S e^{\frac{i\omega}{2}Q(z)}\mathrm{d}z \tag{5.8}$$

in the high-frequency limit. In other words, I is asymptotically equivalent to the right-hand side of expression (5.8), as explained in Section 4.1.

To evaluate the integral on the right-hand side, we consider the following lemma.

Lemma 2. *It is true that*

$$\int\limits_S e^{\frac{i\omega}{2}Q(z)}\mathrm{d}z = \left(\frac{2\pi}{\omega}\right)^{\frac{1}{2}}e^{\pm\iota\frac{\pi}{4}}.$$

In view of this lemma and expression (5.8), we sum over all y such that $\mathrm{d}\psi\left(y\right) = 0$ to obtain

$$\int\limits_S a\left(y\right)e^{\iota\omega\psi(y)}\mathrm{d}y \sim \left(\frac{2\pi}{\omega}\right)^{\frac{m}{2}}\sum_{y|\mathrm{d}\psi=0}\frac{e^{\iota\omega\psi(y)}a\left(y\right)}{\sqrt{\det\left[\frac{\partial^2\psi}{\partial y_i\partial y_j}\right]}}e^{\iota\frac{\pi}{4}(2l-m)}, \tag{5.9}$$

as $\omega \to \infty$. Now, we can proceed to evaluate integral (5.5) in the context of ray theory.

5.3.3 Phase change

In this section, we will apply the method of stationary phase to integral (5.5).

Comparing the right-hand side of expression (5.5) with the right-hand side of expression (5.6)), we see that

$$\psi\left(y\right) = -\left|x-y\right|$$

and

$$a\left(y\right) = \frac{c\left(y\right)}{\left|x-y\right|}.$$

Thus, we write

$$\tilde{u}\left(x,\omega\right) = \int\limits_S \frac{e^{-\iota\omega|x-y|}}{|x-y|}c\left(y\right)\mathrm{d}y \sim \left(\frac{2\pi}{\omega}\right)^{\frac{m}{2}}\sum_{y|\mathrm{d}|x-y|=0}\frac{c\left(y\right)e^{-\iota\omega|x-y|}}{|x-y|\sqrt{\det\left[\frac{\partial^2\psi}{\partial y_i\partial y_j}\right]}}e^{\iota\frac{\pi}{4}(2l-m)}.$$

At the points where the Hessian, $\det\left[\partial^2\psi/\partial y_i\partial y_j\right]$, is zero, the solution is infinity; these points belong to a caustic.

Closing remarks

Caustics are singularities of an the mathematical formulation of the asymptotic ray theory. They illustrate, however, important physical phenomena which can be observed in experimental studies.

Mathematical studies of the catastrophe theory provide a rigorous method for the analysis of ray singularities. Geometrical optics provides a good platform for the investigation of the consequences predicted by the catastrophe theory.[2]

Exercises

Exercise 5.1. Find the envelope of the family of straight lines in the xz-plane given generically by line

$$a\left(t\right)x+b\left(t\right)z=c\left(t\right). \tag{5.10}$$

Solution 5.1. Let us write a line within this family that neighbours line (5.10) as

$$a\left(t+\Delta t\right)x+b\left(t+\Delta t\right)z=c\left(t+\Delta t\right). \tag{5.11}$$

We wish to write expression (5.11) in a way that allows us for a convenient comparison with expression (5.10). In view of the mean-value theorem, there exists a real number, $\xi_a\in\left(t,t+\Delta t\right)$, such that

$$\left.\frac{\mathrm{d}a\left(t\right)}{\mathrm{d}t}\right|_{t=\xi_a}=\frac{a\left(t+\Delta t\right)-a\left(t\right)}{\Delta t}.$$

Solving for $a\left(t+\Delta t\right)$, we obtain

$$a\left(t+\Delta t\right)=a\left(t\right)+\left.\frac{\mathrm{d}a\left(t\right)}{\mathrm{d}t}\right|_{t=\xi_a}\Delta t.$$

Analogous expressions hold for b and c. Consequently, we can write expression (5.11) as

$$\left(a\left(t\right)+\left.\frac{\mathrm{d}a\left(t\right)}{\mathrm{d}t}\right|_{t=\xi_a}\Delta t\right)x+\left(b\left(t\right)+\left.\frac{\mathrm{d}b\left(t\right)}{\mathrm{d}t}\right|_{t=\xi_b}\Delta t\right)z=c\left(t\right)+\left.\frac{\mathrm{d}c\left(t\right)}{\mathrm{d}t}\right|_{t=\xi_c}\Delta t.$$

Rearranging, we rewrite it as

$$\left(a\left(t\right)x+b\left(t\right)x-c\left(t\right)\right)+\left.\frac{\mathrm{d}a\left(t\right)}{\mathrm{d}t}\right|_{t=\xi_a}\Delta tx+\left.\frac{\mathrm{d}b\left(t\right)}{\mathrm{d}t}\right|_{t=\xi_b}\Delta tz=\left.\frac{\mathrm{d}c\left(t\right)}{\mathrm{d}t}\right|_{t=\xi_c}\Delta t.$$

[2] Interested readers might refer to Arnold, V.I., (1991) The theory of singularities and its applications: Press Syndicate of the University of Cambridge, to Arnold, V.I., (1992) Catastrophe theory (3rd edition): Springer-Verlag, to Nye, J.F., (1999) Natural focusing and fine structure of light: Caustics and wave dislocations: Institute of Physics Publishing (IOP), Bristol and Philadelphia, and to Porteous, I.R., (1994) Geometric differentiation for the intelligence of curves and surfaces: Cambridge University Press.

In view of equation (5.10), the term in parentheses is zero. Also, since $\Delta t \neq 0$, we divide the remaining part by Δt to get

$$a'(\xi_a)\, x + b'(\xi_b)\, z = c'(\xi_c)\,,$$

where symbol $'$ stands for the derivative with respect to the argument. Since we are interested in the intersection of the neighbouring lines, we set $\Delta t \to 0$, which implies that $\xi_a \to t$, $\xi_b \to t$ and $\xi_c \to t$. Thus we write

$$a'(t)\, x + b'(t)\, z = c'(t)\,. \tag{5.12}$$

The condition for the two lines to intersect one another is that they both contain a given point, (x, z). To formulate this conditions, we solve a system consisting of equations (5.10) and (5.12), namely,

$$\begin{bmatrix} a(t) & b(t) \\ a'(t) & b'(t) \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} = \begin{bmatrix} c(t) \\ c'(t) \end{bmatrix}. \tag{5.13}$$

Following Cramer's rule, we write the solution, $[x(t), z(t)]$, as

$$x(t) = \frac{\det \begin{bmatrix} c(t) & b(t) \\ c'(t) & b'(t) \end{bmatrix}}{W}$$

and

$$z(t) = \frac{\det \begin{bmatrix} a(t) & c(t) \\ a'(t) & c'(t) \end{bmatrix}}{W},$$

where

$$W = \det \begin{bmatrix} a(t) & b(t) \\ a'(t) & b'(t) \end{bmatrix}.$$

Thus, the envelope of the family of lines given by expression (5.10) is solution $[x(t), z(t)]$. A particular example of such an envelope is discussed in Exercise 5.3 and shown in Figure 5.1.

Exercise 5.2. Use the results of Exercise 5.1 to show that we can view a caustic as the envelope of rays.

Solution 5.2. Let us consider wavefront $S(\xi) = [S_x(\xi), S_z(\xi)]$, where ξ is the arclength parameter. We would like to write the equations for the rays corresponding to this wavefront. Since these rays are straight lines, we can express them by

$$a(\xi)\, x + b(\xi)\, z = c(\xi)\,.$$

The normals to the wavefront, $N(\xi)$, are tangent to the rays. Hence, vectors $[a(\xi), b(\xi)]$ and $[N_1(\xi), N_2(\xi)]$ are orthogonal to one another, namely,

$$[a(\xi), b(\xi)] = [-N_2(\xi), N_1(\xi)]\,. \tag{5.14}$$

This allows us to write the equations for rays as

$$a(\xi)\,x + b(\xi)\,z \equiv \det\begin{bmatrix} x & z \\ -b(\xi) & a(\xi) \end{bmatrix} \equiv -\det\begin{bmatrix} x & z \\ N_1(\xi) & N_2(\xi) \end{bmatrix} = c(\xi).$$

Each ray passes through the wavefront $S(\xi) = [S_x(\xi), S_z(\xi)]$, and hence the ray must satisfy

$$-\det\begin{bmatrix} S_x(\xi) & S_z(\xi) \\ N_1(\xi) & N_2(\xi) \end{bmatrix} = c(\xi).$$

Combining the last two expressions, we obtain the equation for rays; namely,

$$a(\xi)\,x + b(\xi)\,z = -\det\begin{bmatrix} S_x(\xi) & S_z(\xi) \\ N_1(\xi) & N_2(\xi) \end{bmatrix} \equiv -\det[S(\xi), N(\xi)]. \tag{5.15}$$

We wish to obtain the equation for the envelope of rays. Following equation (5.13) and in view of expressions (5.14) and (5.15), we write

$$\begin{bmatrix} N_2(\xi) & -N_1(\xi) \\ \frac{\mathrm{d}}{\mathrm{d}\xi}N_2(\xi) & -\frac{\mathrm{d}}{\mathrm{d}\xi}N_1(\xi) \end{bmatrix}\begin{bmatrix} x \\ z \end{bmatrix} = \begin{bmatrix} \det[S(\xi), N(\xi)] \\ \frac{\mathrm{d}}{\mathrm{d}\xi}\det[S(\xi), N(\xi)] \end{bmatrix}. \tag{5.16}$$

To use Cramer's rule for solving system (5.16), we write the main determinant as

$$W = \det\begin{bmatrix} N_2(\xi) & -N_1(\xi) \\ N_2'(\xi) & -N_1'(\xi) \end{bmatrix} = \det\begin{bmatrix} N_1(\xi) & N_2(\xi) \\ N_1'(\xi) & N_2'(\xi) \end{bmatrix} \equiv \det[N(\xi), N'(\xi)],$$

Invoking one of the Serret-Frenet equations[3], namely,

$$N'(\xi) = -\kappa(\xi)\,T(\xi), \tag{5.17}$$

where κ denotes curvature of the wavefront and T denotes its tangent, we rewrite this determinant as

$$W = \det[N(\xi), -\kappa(\xi)\,T(\xi)] = -\kappa(\xi)\det[N(\xi), T(\xi)]. \tag{5.18}$$

Presently, we will show that the determinant on the right-hand side is equal to one. To do so, we write

$$\det[N(\xi), T(\xi)] = N_1(\xi)\,T_2(\xi) - N_2(\xi)\,T_1(\xi) = [N_1, N_2]\cdot[T_2, -T_1].$$

Recognizing that N and T are unit length vectors orthogonal to one another and that $[T_2, -T_1] = N$, we write

$$\det[N(\xi), T(\xi)] = N \cdot N = 1, \tag{5.19}$$

as required. Thus, returning to equation (5.18) and using result (5.19), we see that the main determinant is

$$W = -\kappa(\xi).$$

Continuing the search for the solution of system (5.16), we write

[3] Interested readers might refer to a standard book on differential geometry, e.g. , Spivak (1999, Volume 2, p. 34).

$$W_x = \det \begin{bmatrix} \det\left[S\left(\xi\right), N\left(\xi\right)\right], & -N_1\left(\xi\right) \\ \frac{\mathrm{d}}{\mathrm{d}\xi} \det\left[S\left(\xi\right), N\left(\xi\right)\right] & -N_1'\left(\xi\right) \end{bmatrix}.$$

First, let us consider

$$\frac{\mathrm{d}}{\mathrm{d}\xi} \det\left[S\left(\xi\right), N\left(\xi\right)\right] = \det\left[S'\left(\xi\right), N\left(\xi\right)\right] + \det\left[S\left(\xi\right), N'\left(\xi\right)\right].$$

Since $S'\left(\xi\right)$ is tangent to $S\left(\xi\right)$, we write

$$\frac{\mathrm{d}}{\mathrm{d}\xi} \det\left[S\left(\xi\right), N\left(\xi\right)\right] = \det\left[T\left(\xi\right), N\left(\xi\right)\right] + \det\left[S\left(\xi\right), N'\left(\xi\right)\right].$$

Using equation (5.19), we write

$$\frac{\mathrm{d}}{\mathrm{d}\xi} \det\left[S\left(\xi\right), N\left(\xi\right)\right] = -1 + \det\left[S\left(\xi\right), N'\left(\xi\right)\right]. \tag{5.20}$$

Consequently,

$$W_x = \det \begin{bmatrix} \det\left[S\left(\xi\right), N\left(\xi\right)\right] & -N_1\left(\xi\right) \\ -1 + \det\left[S\left(\xi\right), N'\left(\xi\right)\right] & -N_1'\left(\xi\right) \end{bmatrix}.$$

Proceeding to compute this determinant, we write

$$\begin{aligned} W_x &= -\det\left[S\left(\xi\right), N\left(\xi\right)\right] N_1'\left(\xi\right) + \left(-1 + \det\left[S\left(\xi\right), N'\left(\xi\right)\right]\right) N_1\left(\xi\right) \\ &= -\det \begin{bmatrix} S_x\left(\xi\right) & S_z\left(\xi\right) \\ N_1\left(\xi\right) & N_2\left(\xi\right) \end{bmatrix} N_1'\left(\xi\right) - N_1\left(\xi\right) + \det \begin{bmatrix} S_x\left(\xi\right) & S_z\left(\xi\right) \\ N_1'\left(\xi\right) & N_2'\left(\xi\right) \end{bmatrix} N_1\left(\xi\right). \end{aligned}$$

Performing algebraic manipulations and gathering terms with $S_x\left(\xi\right)$ and with $S_z\left(\xi\right)$, we get

$$W_x = -S_x\left(\xi\right)\left(N_2\left(\xi\right) N_1'\left(\xi\right) - N_2'\left(\xi\right) N_1\left(\xi\right)\right) + S_z\left(\xi\right)\left(N_1'\left(\xi\right) N_1\left(\xi\right) - N_1\left(\xi\right) N_1'\left(\xi\right)\right) - N_1\left(\xi\right).$$

Since the term associated with $S_z\left(\xi\right)$ is equal to zero, we write

$$W_x = -N_1\left(\xi\right) - S_x\left(\xi\right)\left(N_1'\left(\xi\right) N_2\left(\xi\right) - N_1\left(\xi\right) N_2'\left(\xi\right)\right).$$

Let us write the term in parentheses as a scalar product of vectors, namely,

$$W_x = -N_1\left(\xi\right) - S_x\left(\xi\right)\left[N_1'\left(\xi\right), N_2'\left(\xi\right)\right] \cdot \left[N_2\left(\xi\right), -N_1\left(\xi\right)\right].$$

Since, in view of expression (5.19), the second vector in the product is $-T\left(\xi\right)$, we write

$$W_x = -N_1\left(\xi\right) + S_x\left(\xi\right) N'\left(\xi\right) \cdot T\left(\xi\right).$$

Invoking Serret-Frenet equation (5.17), we get

$$W_x = -N_1\left(\xi\right) + S_x\left(\xi\right) \kappa\left(\xi\right) T\left(\xi\right) \cdot T\left(\xi\right).$$

Since T has the unit length, we conclude that

$$W_x = -N_1(\xi) - \kappa(\xi) S_x(\xi).$$

Similarly,

$$W_z = -N_2(\xi) - \kappa(\xi) S_z(\xi).$$

Having obtained W, W_x and W_z, we write the solution of system (5.16) as

$$\begin{aligned}\epsilon(\xi) &= \left[\frac{W_x}{W}, \frac{W_z}{W}\right] \\ &= \left[\frac{-N_1(\xi) - \kappa(\xi) S_x(\xi)}{-\kappa(\xi)}, \frac{-N_2(\xi) - \kappa(\xi) S_z(\xi)}{-\kappa(\xi)}\right] \\ &= \left[S_x(\xi) + \frac{1}{\kappa(\xi)} N_1(\xi), S_z(\xi) + \frac{1}{\kappa(\xi)} N_2(\xi)\right].\end{aligned}$$

Using properties of the vector addition, we write

$$\epsilon(\xi) = [S_x(\xi), S_z(\xi)] + \frac{1}{\kappa(\xi)} [N_1(\xi), N_2(\xi)].$$

Recognizing that the first term is the wavefront, we write

$$\epsilon(\xi) = S(\xi) + \frac{1}{\kappa(\xi)} N(\xi), \tag{5.21}$$

which is a special case of expression (5.3) for isotropic media with $v = 1$. Thus, we have illustrated the fact that caustics are envelopes of rays.

Exercise 5.3. Consider a family of straight lines given by

$$\xi^2 x + \xi z = 1.$$

Find the equation of their envelope, and plot it.

Solution 5.3. Following equation (5.13) with $a = \xi^2$ and $b = \xi$, we write the equations of the envelope as

$$\begin{bmatrix} \xi^2 & \xi \\ 2\xi & 1 \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

The solution of this system is

$$[x(\xi), z(\xi)] = \left(-\frac{1}{\xi^2}, \frac{2}{\xi}\right),$$

which is a curve shown in Figure 5.1.

Exercise 5.4. Consider a parametric representation of a wavefront, $S(\xi) = \left[\xi, \xi^2\right]$, in an isotropic homogeneous medium. Find the expression of the caustic generated by this wavefront, and plot it.

Solution 5.4. Following the Huygens principle, each point on the wavefront is a source. In an isotropic homogeneous medium, the rays are normal to the wavefront and their length is directly proportional to time. For simplicity, we set $v = 1$ in equation (5.3) to write

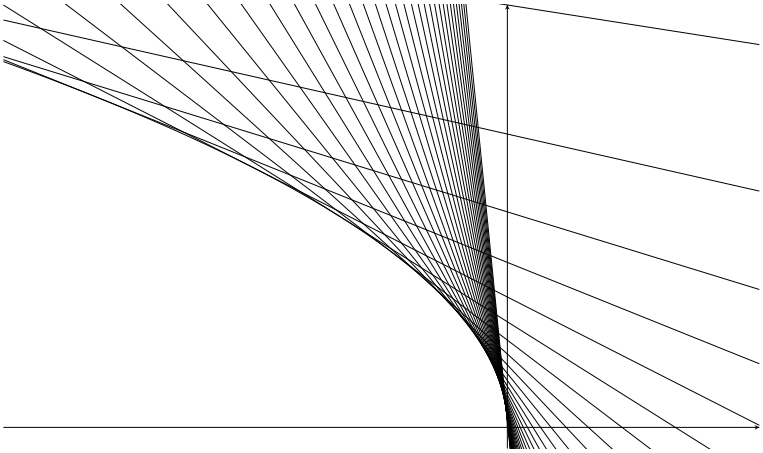


Fig. 5.1. Lines $z = -\xi x + 1/\xi$ and their envelope given by $\left[-1/\xi^2, 2/\xi\right]$

$$\epsilon\left(\xi\right) = S\left(\xi\right) + \frac{1}{\kappa\left(\xi\right)} N\left(\xi\right). \tag{5.22}$$

To find N, we use expression for the unit normal to S, namely,

$$\left[N_1, N_2\right] = \left[\frac{-\left(S'_z\right)^2}{\sqrt{\left(S'_x\right)^2 + \left(S'_z\right)^2}}, \frac{\left(S'_x\right)^2}{\sqrt{\left(S'_x\right)^2 + \left(S'_z\right)^2}}\right]$$

to write the normal of the given wavefront as

$$N\left(\xi\right) = \left[\frac{-2\xi}{\sqrt{1+4\xi^2}}, \frac{1}{\sqrt{1+4\xi^2}}\right].$$

To find κ, we invoke the standard expression for the curvature to write

$$\kappa = \frac{S'_x S''_z - S''_x S'_z}{\left(\left(S'\right)^2 + \left(S'_z\right)^2\right)^{\frac{3}{2}}} = \frac{2}{\left(1+4\xi^2\right)^{\frac{3}{2}}}.$$

Thus, inserting expressions for N and κ into equation (5.22), we get

$$\epsilon\left(\xi\right) = \left[\xi, \xi^2\right] + \frac{\left(1+4\xi^2\right)^{\frac{3}{2}}}{2}\left[\frac{-2\xi}{\sqrt{1+4\xi^2}}, \frac{1}{\sqrt{1+4\xi^2}}\right] = \left[\xi, \xi^2\right] + \left[-\xi\left(1+4\xi^2\right), \frac{1}{2}\left(1+4\xi^2\right)\right].$$

Adding the two vectors and simplifying, we obtain

$$\epsilon\left(\xi\right) = \left[-4\xi^3, 3\xi^2 + \frac{1}{2}\right].$$

The plots of the original wavefront, S, and the of the generated caustic, ϵ, which is the locus of singularities of all the subsequent *wavefronts*, are shown in Figure 5.2. Also, in view of the comment on page 111, the caustic is the evolute of the original wavefront.

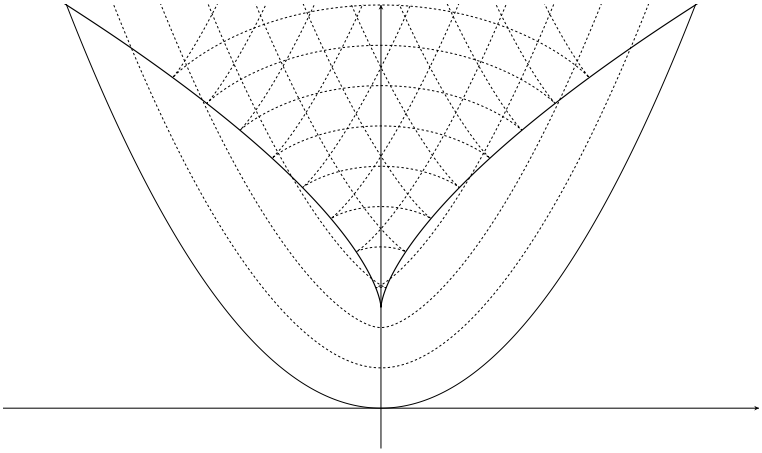


Fig. 5.2. Propagation of a parabolic wavefront, $S(\xi) = \left[\xi, \xi^2\right]$. The locus of singularities of all the wavefronts is the caustic; herein shown by the cusped curve.

6

Symbols of linear differential operators

Preliminary remarks

6.1 Motivational example

While studying partial differential equations, which we can formally write as

$$F\left(x, \frac{\partial^{\beta} u}{\partial x^{\beta}}\right) = 0,$$

where $|\beta| \leq m$, it is often convenient to consider a linear differential operator given by the following polynomial;

$$P\left(x, \frac{\partial}{\partial x}\right) = \sum_{\alpha} a_{\alpha}(x) \frac{\partial^{\alpha}}{\partial x^{\alpha}},$$

where

$$a_{\alpha}(x) = \frac{\partial F}{\partial\left(\frac{\partial^{\alpha} u}{\partial x^{\alpha}}\right)}\left(x, \frac{\partial^{\beta} u}{\partial x^{\beta}}\right).$$

To associate the differential operator with the corresponding algebraic polynomial, we replace

$$\left[\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}\right]$$

by

$$[\iota\xi_1, \ldots, \iota\xi_n],$$

where $\iota := \sqrt{-1}$, to get

$$P(x, \iota\xi) = \sum_{\alpha} a_{\alpha}(x) (\iota\xi)^{\alpha}.$$

$P(x, \iota\xi)$ is called the symbol of differential operator $P(x, \partial/\partial x)$. It is a polynomial of degree m in ξ whose coefficients depend on x. Let us also define

$$D_j := -\iota \frac{\partial}{\partial x_j};$$

in other words

$$D := -\iota \left[\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}\right].$$

This notation implies that the symbol of

$$P(x,D)=\sum_{\alpha}a_{\alpha}(x)D^{\alpha}$$

is

$$P(x,\xi)=\sum_{\alpha}a_{\alpha}(x)\xi^{\alpha}.$$

6.2 General formulation

6.2.1 Wavefront of distribution

In the following we follow Hörmander (1983); Gel'fand et al. (1969); Treves (1980) and Guillemin and Sternberg (1977).

Definition 6.1. *A distribution on a manifold X is a continuous linear function on $C_0^{\infty}(X)$.*

The support of a distribution is defined as follows.

Definition 6.2. *A point x does not lie in the support of a distribution ν, denoted by supp ν, if there is some neighbourhood $U$of x such that for a function $u\in C_0^{\infty}(X)$with supp $u\subset U$ the following holds: $\langle u,v\rangle=0$.*

Moreover, one can define the singular support of a distribution.

Definition 6.3. *The singular support of a distribution ν, denoted by sing suppν, is a closed subset of X, such that outside this set ν is a honest function from $C^{\infty}(X)$.*

A distribution is also called a generalized density (of order one), where a density is defined as follows.

Definition 6.4. *A density of order s is a map ρ from the frame bundle FX to the complex numbers, such that*

$$\rho(A\xi_1,\ldots,A\xi_n)=|\det A|^s\rho(\xi_1,\ldots,\xi_n),$$

for any linear transformation A. The space of densities of order s on X is denoted by $|\Lambda|^s(X)$ and the space of smooth densities by $C^{\infty}|\Lambda|^s(X)$.

The notion of a generalized density can be justified by realizing that a smooth density ρ of order one defines a distribution that acts on $u\in C_0^{\infty}(X)$ as

$$\rho(u)=\langle u,\rho\rangle=\int_X u\rho,$$

since $u\rho$ is a density of order one and can be integrated over X. On the other hand, one can define a generalized function as a continuous linear functional on $C_0^{\infty}|\Lambda|(X)$, i.e. a continuous linear functional on the space of smooth densities of compact support. The space of generalized densities is denoted by $C^{-\infty}|\Lambda|(X)$ and the space of generalized functions by $C^{-\infty}(X)$.

If f is a proper smooth map from X to a manifold Y (i.e. f maps compact sets to compact sets) and ρ is a generalized density on X, then $f_*\rho$ is a generalized density on Y such that

$$\langle u, f_*\rho\rangle = \langle f^*u, \rho\rangle\,. \tag{6.1}$$

Hence, if f is a proper smooth map from X to Y,

$$f^* : C_0^\infty(Y) \to C_0^\infty(X)$$

and

$$f_* : C^{-\infty}|\Lambda|(X) \to C^{-\infty}|\Lambda|(Y)\,.$$

Similarly, one can define a pairing of a compactly supported generalized density on smooth functions. These densities are denoted by $C_0^{-\infty}|\Lambda|(X)$ and are a continuous linear functionals on $C^\infty(X)$. In this case a smooth map f (not necessarily proper) from X to Y defines

$$f^* : C^\infty(Y) \to C^\infty(X)$$

and

$$f_* : C_0^{-\infty}|\Lambda|(X) \to C_0^{-\infty}|\Lambda|(Y)\,.$$

Since any density is also a generalized density, it is possible to restrict the definition of the push forward to densities. In some cases it is possible to compute the push forward of a density in a simple matter without the use of the pairing (6.1). In case f is a submersion $X \to Y$ and ρ is a smooth density with compact support on X, one can compute $f_*\rho$ by integration over fibers of f. More precisely, for each $y \in Y$ and each $x \in f^{-1}(y)$, one can identify spaces $|\Lambda|(T_xX)$ with $|\Lambda|\left(T_x\left(f^{-1}(y)\right)\right) \otimes |\Lambda|(T_yY)$ and thus the density ρ can be thought of as a density along the fiber with values in the densities at y. This way, one can integrate the density ρ along the fibers of f to get a density σ on Y, which is smooth. Hence, if f is a submersion, one can write

$$f_* : C_0^\infty|\Lambda|(X) \to C_0^\infty|\Lambda|(Y)$$

as

$$(f_*\rho)(y) = \int_{f^{-1}(y)} \rho.$$

In this case, using formula (6.1), one can extend the definition of a pull-back to generalized functions,

$$f^* : C^{-\infty}(Y) \to C^{-\infty}(X)\,.$$

Note that if the fibers of f are compact or ρ restricted to any fiber $f^{-1}(y)$ is compact even if ρ is only C^∞, i.e. not necessarily compactly supported, the above integral still makes sense and one can thus extend the definition of the push-forward to these cases.

To study pull-back of generalized functions in more general cases, one generalizes the notion of singular support of a distribution. To determine the singular support, one has to know where a distribution is smooth.

Definition 6.5. *A generalized density ρ is smooth at $(x,\alpha) \in T^*X$ if the following holds:*

Let S be a manifold and let $f : S \times X \to \mathbb{R}$ be any smooth function such that $df_s(x) = \alpha$, where $f_s(x) = f(s,x)$. If F is defined by $F(s,x) = (s, f(s,x))$, then there is a smooth function $b \in C_0^\infty(X)$ such that

1. $b(x) \neq 0$
2. $F_*(b\rho) \in C^\infty(S \times \mathbb{R})$ near s.

The next generalization of the singular support can be given by the following definition.

Definition 6.6. *The projective wave front set of a generalized density ρconsists of points $(x,\alpha) \in T^*M \setminus 0$ such that ρ is not smooth at (x,α).*

Roughly speaking, a distribution is smooth at (x,α) if for any function $f : X \to \mathbb{R}$ with $(df)(x) = \alpha$, the distribution $f_*\rho$ – which is well defined since $df = \alpha \neq 0$ and hence one can integrate over the fibers f^{-1} – is an honest compactly supported smooth function on $\mathbb{R}$.

If ρ is smooth at (x,α), then it is smooth also at $(x,c\alpha)$ for any real constant c. This can be verified by taking cf, which satisfies both conditions from the definition of the smoothness of ρ. This fact justifies the name *projective* wave front. To define wave front of a distribution, one can recall the following theorem.

Theorem 6.1. *A distribution u on $\mathbb{R}^n$ is a C^∞ function if and only if its Fourier transform is rapidly decaying at infinity (i.e. it belongs to the Schwartz space).*

Assume that ρ is smooth at (x,α). The theorem above implies that Fourier transform of $f_{s*}b\rho$ at τ vanishes to all orders as $\tau \to \pm\infty$, i.e.

$$\left\langle (f_{s*}b\rho)(x), e^{-i\tau x} \right\rangle = \left\langle b\rho, e^{-i\tau f_s} \right\rangle = O\left(|\tau|^{-N}\right) \qquad \text{for all } N \text{ as } \tau \to \pm\infty. \tag{6.2}$$

One can say that a distribution is forward smooth at (x,α) if the above condition (6.2) holds as $\tau \to +\infty$.

Definition 6.7. *The wave front set of a generalized density ρ consists of points $(x,\alpha) \in T^*M$ such that ρ is not forward smooth at (x,α).*

6.2.2 Principal symbol

Now, we can state the following definition.

Definition 6.8. *The principal symbol of*

$$P(x,D) = \sum_{|\alpha| \leq m} a_\alpha(x) D^\alpha$$

is the function given by

$$P_m(x,\xi) = \sum_{|\alpha| = m} a_\alpha(x) \xi^\alpha.$$

P_m is a polynomial homogeneous of degree m in ξ.

The principal symbol is important while studying highly oscillatory functions. This is stated the fundamental theorem of the asymptotic expansion.

Theorem 6.2. *If f is a smooth real-valued function, then – as $\lambda \to \infty$– we can write*

$$\begin{aligned} P(x, D) &= \exp[-\iota\lambda f] P(x, D) \exp[\iota\lambda f] \\ &= \lambda^m P_m(x, \mathrm{d}f) + O\left(\lambda^{m-1}\right). \end{aligned}$$

In general, the symbol tells us how a differential equation acts on functions that have their support contained in a small neighbourhood of point x. If a given function varies rapidly the highest-order derivatives are dominant and, hence, the principal symbol contains the most important information.

6.2.3 Symbol

The notion of the principal symbol leads to the definition of the symbol $\sigma(D)$ of a differential operator.

6.3 Physical applications

6.3.1 Support of singularities

Following the discussion of Section 6.2.1, we see ...?

6.3.2 Laplace equation

6.3.3 Heat equation

Revisiting equation (2.52), namely,

$$\frac{\partial^2 f}{\partial^2 x_1} + \frac{\partial^2 f}{\partial^2 x_2} + \frac{\partial^2 f}{\partial^2 x_3} = \frac{1}{k}\frac{\partial f}{\partial x_4},$$

we can see that the Fourier transform maps it into

$$\left(p_1^2 + p_2^2 + p_3^2 - \frac{1}{k}p_4\right)\hat{f}(p_1, p_2, p_3, p_4) = 0.$$

This implies that the support of $\hat{f}(p_1, p_2, p_3, p_4)$ lies in the set given by

$$p_1^2 + p_2^2 + p_3^2 - \frac{1}{k}p_4 = 0.$$

Any singularity of $f(x_1, x_2, x_3, x_4)$ reflects into infinitely large frequency in its Fourier transform. The "directional part" of the wavefront of the singularity is given by the direction of this infinite p. In the case of the heat equation, we see that

$$k\left(p_1^2 + p_2^2 + p_3^2\right) = p_4,$$

or the possible infinitely large ps are of the form

$$\left[p_1, p_2, p_3, k\left(p_1^2+p_2^2+p_3^2\right)\right].$$

The direction of this vector for large p is given by

$$\lim_{(p_1^2+p_2^2+p_3^2)\to\infty} \frac{\left[p_1, p_2, p_3, k\left(p_1^2+p_2^2+p_3^2\right)\right]}{\|\left[p_1, p_2, p_3, k\left(p_1^2+p_2^2+p_3^2\right)\right]\|} = [0,0,0,k].$$

Hence, the direction of the possible singularities of f are in the direction of $[0,0,0,1]$.

6.3.4 Wave equation

Let us consider wave equation (**??**), which in two spatial dimensions we can write as

$$\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} - \frac{1}{v^2}\frac{\partial^2 u}{\partial t^2} = 0. \tag{6.3}$$

The principal symbol is

$$P_2(x,t,\eta,\tau) = \eta_1^2 + \eta_2^2 - \frac{1}{v^2}\tau^2,$$

where, to emphasize the physical distinction between x and t, we let

$$\frac{\partial}{\partial x_j} = \iota\eta_j, \qquad j \in \{1,2\},$$

and

$$\frac{\partial}{\partial t} = \iota\tau.$$

Since P_2 is a polynomial, we can set it to zero and write

$$\eta_1^2 + \eta_2^2 = \frac{1}{v^2}\tau^2.$$

Since η_1, η_2 and τ are the components of a one-form, we can write this equation as

$$\left(\frac{\partial f}{\partial x_1}\right)^2 + \left(\frac{\partial f}{\partial x_2}\right)^2 = \frac{1}{v^2}\left(\frac{\partial f}{\partial t}\right)^2, \tag{6.4}$$

where f is a function of x that determines the one form with components $\partial f/\partial x_1, \partial f/\partial x_2, \partial f/\partial t$. This equation is identical to equation (6.6), discussed in Exercise 6.1, below, which is the characteristic equation of wave equation (6.3).?

If we let $f(x_1,x_2,t) = \psi(x_1,x_2) - t$, then the general form of characteristic equation (6.4) reduces to

$$\left(\frac{\partial \psi}{\partial x_1}\right)^2 + \left(\frac{\partial \psi}{\partial x_2}\right)^2 = \frac{1}{v^2},$$

which is the eikonal equation. Thus, the eikonal equation is the characteristic equation of wave equation (6.3) if f represents a level set.

Since wave equation (**??**) has no other terms than second-order derivatives, it exhibits a particular property that its principal symbol is also its complete symbol. In general, this is not the case. Consequently, seeking a characteristic equation by invoking the principal symbol allows

as to focus only on the highest-order derivatives, rather than examining the entire equation. This approach facilitates our work to be performed in the following section.

The fact that vanishing of the principal symbol results in characteristic equations allows us an insight into the meaning of these equations. The characteristic equation specifies the relation among the variables, where a given differential equation does not behave as expected based on its classification. For instance, along the characteristics, wave equation (**??**) does not behave as a second-order partial differential equation.

Closing remarks

At this point, we can clarify our terminology. The term "characteristic" in the mathematical context means that a given entity is connected to our object of interest in an invariant way. In the present case, this invariance refers to the coordinate transformations. More specifically, if a diffeomorphism transforms the original equation into a new equation, then it also transforms the characteristic equation of the original equation into the characteristic equation of the transformed equation. This is illustrated in Exercise 6.1.

Exercises

Exercise 6.1. Consider the wave equation given by

$$\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} = \frac{1}{v^2}\frac{\partial^2 u}{\partial t^2} \tag{6.5}$$

with its characteristic equation given by

$$\left(\frac{\partial f}{\partial x_1}\right)^2 + \left(\frac{\partial f}{\partial x_2}\right)^2 = \frac{1}{v^2}\left(\frac{\partial f}{\partial t}\right)^2. \tag{6.6}$$

Show that the relation between these two equations is preserved when they are both transformed into polar coordinates.

Solution 6.1. To express the wave equation and its characteristic equation in polar coordinates given by

$$x_1 = r\cos\varphi \tag{6.7}$$

$$x_2 = r\sin\varphi, \tag{6.8}$$

or, equivalently, by

$$r = \sqrt{x_1^2 + x_2^2} \tag{6.9}$$

$$\varphi = \arctan\frac{x_2}{x_1}, \tag{6.10}$$

we must express differential operators in terms of r and φ. Thus, we write

$$\frac{\partial}{\partial x_1} = \frac{\partial r}{\partial x_1}\frac{\partial}{\partial r} + \frac{\partial \varphi}{\partial x_1}\frac{\partial}{\partial \varphi}$$

and

$$\frac{\partial}{\partial x_2} = \frac{\partial r}{\partial x_2}\frac{\partial}{\partial r} + \frac{\partial \varphi}{\partial x_2}\frac{\partial}{\partial \varphi}.$$

Considering $\partial/\partial x_1$ and using expressions (6.9) and (6.10), we get

$$\begin{aligned}\frac{\partial}{\partial x_1} &= \frac{\partial\sqrt{x_1^2+x_2^2}}{\partial x_1}\frac{\partial}{\partial r} + \frac{\partial \arctan\frac{x_2}{x_1}}{\partial x_1}\frac{\partial}{\partial \varphi} \\ &= \frac{x_1}{\sqrt{x_1^2+x_2^2}}\frac{\partial}{\partial r} - \frac{x_2}{x_1^2+x_2^2}\frac{\partial}{\partial \varphi} = \frac{x_1}{r}\frac{\partial}{\partial r} - \frac{x_2}{r^2}\frac{\partial}{\partial \varphi}.\end{aligned}$$

Using expressions (6.7) and (6.8), we obtain

$$\frac{\partial}{\partial x_1} = \cos\varphi\frac{\partial}{\partial r} - \frac{\sin\varphi}{r}\frac{\partial}{\partial \varphi}. \tag{6.11}$$

Following a similar procedure for $\partial/\partial x_2$, we obtain

$$\frac{\partial}{\partial x_2} = \sin\varphi\frac{\partial}{\partial r} + \frac{\cos\varphi}{r}\frac{\partial}{\partial \varphi}. \tag{6.12}$$

Repeating similar procedures for the second derivatives, $\partial^2/\partial x_1^2$ and $\partial^2/\partial x_2^2$, we obtain the Laplace operator, $\nabla^2 := \left[\partial^2/\partial x_1^2 + \partial^2/\partial x_2^2\right]$, in polar coordinates, namely,

$$\nabla^2 := \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\left(\frac{\partial^2}{\partial \varphi^2} + \cot\varphi\frac{\partial}{\partial \varphi}\right).$$

Thus, we can write wave equation (6.5) in polar coordinates as

$$\nabla^2 u = \frac{\partial^2 u}{\partial r^2} + \frac{2}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\left(\frac{\partial^2 u}{\partial \varphi^2} + \cot\varphi\frac{\partial u}{\partial \varphi}\right) = \frac{1}{v^2}\frac{\partial^2 u}{\partial t^2}.$$

Hence, we can write the corresponding characteristic equation, which involves only the highest-order derivatives, as

$$\left(\frac{\partial f}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial f}{\partial \varphi}\right)^2 = \frac{1}{v^2}\left(\frac{\partial f}{\partial t}\right)^2. \tag{6.13}$$

Also, using expressions (6.11) and (6.12), we can write equation (6.6) in polar coordinates as

$$\left(\cos\varphi\frac{\partial f}{\partial r} - \frac{\sin\varphi}{r}\frac{\partial f}{\partial \varphi}\right)^2 + \left(\sin\varphi\frac{\partial f}{\partial r} + \frac{\cos\varphi}{r}\frac{\partial f}{\partial \varphi}\right)^2 = \frac{1}{v^2}\left(\frac{\partial f}{\partial t}\right)^2.$$

Squaring and simplifying, we obtain

$$\left(\frac{\partial f}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial f}{\partial \varphi}\right)^2 = \frac{1}{v^2}\left(\frac{\partial f}{\partial t}\right)^2,$$

which is equation (6.13), as expected.

7

Relations among discussed methods

Preliminary remarks

In this chapter, we discuss the relations among the method of characteristics, Fourier transform and asymptotic solutions of differential equations.

7.1 Characteristics and asymptotic solutions

We defined characteristic surfaces as surfaces along which we are not able to specify freely side conditions. In other words, along these surfaces the solutions cannot be arbitrary since they are governed by the differential equation. We are free to choose the side conditions along transverse directions to the characteristic surfaces and then the solution is given by propagating this side conditions along the characteristics. By choosing the side conditions to be a distribution, we can propagate along the characteristics also solutions that are non-differentiable, solutions whose Fourier transform is noncompact in some direction. In other words, we can propagate solutions that have "high frequency" content.

The eikonal function, ψ, in the asymptotic expansion is part of the choice of asymptotic sequence. In general, this choice is different for different equations; it is characteristic to the equation itself. As we saw in equation (4.16), function ψ is indeed the characteristic function of the differential equation.

Remark 1. The asymptotic approximation to a solution of a differential equation is

$$e^{i\omega\psi(x)}\left(A_0(x)+\frac{1}{i\omega}A_1(x)+\cdots\right),$$

and its Fourier transform is

$$A_0(x)\,\delta(\psi(x)-t)-A_1(x)\,\pi H(\psi(x)-t)+\cdots.$$

The asymptotic approximation as $\omega\to\infty$ describes well only the high frequency content of the solution. In other words, it describes the behaviour of discontinuities of the solution. As can be seen from the Fourier transform of the asymptotic solution, the coefficients A_i describe content of different types of discontinuities within the solution. This is related to Section 6.2.1...

Closing remarks

Exercises

Appendices

A Integral theorems

In this appendix, we state and justify two theorems that play an important role in mathematical physics and are used several times in this book. These are integral theorems, which are often introduced in courses on vector calculus. The two theorems that we will discuss are the divergence theorem and the curl theorem, which are three-dimensional and two-dimensional cases of Stokes's theorem, named in honour of George Gabriel Stokes (1819–1903). Also, in honour of Carl Friedrich Gauss (1777–1855), it is common to refer to the divergence theorem as Gauss' theorem.

Remark 2. The general form of Stokes's theorem is written in modern notation as

$$\int_{\Omega} \mathrm{d}\omega = \int_{\partial\Omega} \omega,$$

where Ω is a piecewise smooth n-dimensional manifold with $\partial\Omega$ being its boundary, and ω is an $n-1$ compactly supported continuously differentiable differential form on Ω. The Fundamental Theorem of Calculus, namely,

$$\int_{a}^{b} \frac{\mathrm{d}f}{\mathrm{d}x}\mathrm{d}x = f(b) - f(a).$$

is a one-dimensional case of Stokes's theorem.

A.1 Divergence theorem

Statement

Divergence theorem relates a surface integral in a vector field to the volume integral in this field. We use this theorem in Appendices D and F to derive differential equations governing physical systems.

Divergence theorem: The surface integral of a continuously differentiable vector field along a closed surface can be expressed as the integral of the divergence of this field over the volume enclosed by this surface. In other words,

$$\iint_S F \cdot N \mathrm{d}S = \iiint_V \nabla \cdot F \mathrm{d}V, \tag{A.1}$$

where N is the unit outward normal vector on surface S.

Since the flux of a vector field across a surface is defined as the surface integral of the normal component of this field, we see that the left-hand side of equation (A.1) is the flux of field F across surface S. Hence, the theorem states that the flux of F across S can be expressed as the volume integral of the divergence of F.

Verification

A general proof of the divergence theorem is quite involved. Below, we will provide an argument that makes the general statement of the theorem plausible.[1]

Let us consider vector field $F = [F_1, F_2, F_3]$. Let us also consider a rectangular box within this field. For convenience, we set the edges of the box to be parallel with the three coordinate axes. As shown in Figure A.1, each edge extends between a_1 and a_2, b_1 and b_2, c_1 and c_2, along the x_1-axis, the x_2-axis, the x_3-axis, respectively.

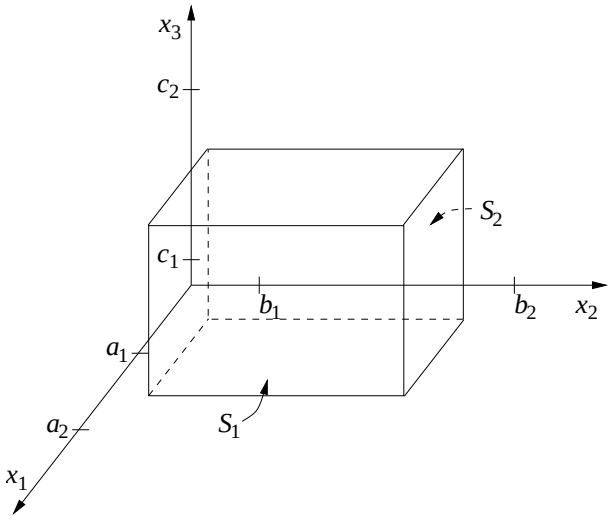


Fig. A.1. A rectangular box used to formulate the divergence theorem

We can parametrize side S_1 in Figure A.1 by

$$S_1 : (x_2, x_3) \mapsto (a_2, x_2, x_3)\,,$$

with $x_2 \in [b_1, b_2]$ and $x_3 \in [c_1, c_2]$. The unit outward normal on this side is $N_1 = [1, 0, 0]$. Hence, the flux across surface S_1 is

[1] Readers interested in a more general proof might refer to Marsden, J.E., and Tromba, A., (1976) Vector calculus: W.H. Freeman & Co., pp. 441 – 443.

$$\iint_{S_1} F \cdot N_1 \mathrm{d}S = \int_{c_1}^{c_2}\int_{b_1}^{b_2} [F_1, F_2, F_3] \cdot [1,0,0]\, \mathrm{d}x_2 \mathrm{d}x_3 = \int_{c_1}^{c_2}\int_{b_1}^{b_2} F_1\left(a_2, x_2, x_3\right) \mathrm{d}x_2 \mathrm{d}x_3.$$

We can parametrize the side opposite to S_1 as

$$S_2 : (x_2, x_3) \mapsto (a_1, x_2, x_3)\,,$$

where, again, $x_2 \in [b_1, b_2]$ and $x_3 \in [c_1, c_2]$, but the unit outward normal is $N_2 = [-1, 0, 0]$. In this case, the flux is

$$\iint_{S_2} F \cdot N_2 \mathrm{d}S = -\int_{c_1}^{c_2}\int_{b_1}^{b_2} F_1\left(a_1, x_2, x_3\right) \mathrm{d}x_2 \mathrm{d}x_3.$$

To consider the flux across these two faces, we sum the two integral expressions to get

$$\begin{aligned}\iint_{S_1} F \cdot N_1 \mathrm{d}S + \iint_{S_2} F \cdot N_2 \mathrm{d}S &= \int_{c_1}^{c_2}\int_{b_1}^{b_2} F_1\left(a_2, x_2, x_3\right) \mathrm{d}x_2 \mathrm{d}x_3 - \int_{c_1}^{c_2}\int_{b_1}^{b_2} F_1\left(a_1, x_2, x_3\right) \mathrm{d}x_2 \mathrm{d}x_3 \\ &= \int_{c_1}^{c_2}\int_{b_1}^{b_2} \left[F_1\left(a_2, x_2, x_3\right) - F_1\left(a_1, x_2, x_3\right)\right] \mathrm{d}x_2 \mathrm{d}x_3. \end{aligned} \tag{A.2}$$

In view of the Fundamental Theorem of Calculus, we can write the integrand of the right-hand side of equation (A.2) as

$$F_1\left(a_2, x_2, x_3\right) - F_1\left(a_1, x_2, x_3\right) = \int_{a_1}^{a_2} \frac{\partial F_1\left(x_1, x_2, x_3\right)}{\partial x_1} \mathrm{d}x_1.$$

Inserting the right-hand side of this equation into the right-hand side of equation (A.2), we get

$$\int_{c_1}^{c_2}\int_{b_1}^{b_2}\int_{a_1}^{a_2} \frac{\partial F_1\left(x_1, x_2, x_3\right)}{\partial x_1} \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3 = \iiint_V \frac{\partial F_1}{\partial x_1} \mathrm{d}V;$$

hence, equation (A.2) becomes

$$\iint_{S_1} F \cdot N_1 \mathrm{d}S + \iint_{S_2} F \cdot N_2 \mathrm{d}S = \iiint_V \frac{\partial F_1}{\partial x_1} \mathrm{d}V. \tag{A.3}$$

Considering the fluxes across the two faces perpendicular to the x_2-axis and the two faces perpendicular to the x_3-axis, we get expressions analogous to expression (A.3). Subsequently, adding the six surface and the three volume integrals, we write

$$\iint_S F \cdot N \mathrm{d}S = \iiint_V \left(\frac{\partial F_1}{\partial x_1} + \frac{\partial F_2}{\partial x_2} + \frac{\partial F_3}{\partial x_3}\right) \mathrm{d}V,$$

where S denotes the surface of the box and V denotes its volume. Examining this equation, we recognize that the term in parentheses is the divergence of F. Thus, we write

$$\iint_S F \cdot N \mathrm{d}S = \iiint_V \nabla \cdot F \mathrm{d}V,$$

which is equation (A.1), as required.

At this point, we have verified the divergence theorem for a rectangular box. However, as argued below, we can extend this result to an arbitrary shape. The motivation for our approach is the fact that different-size boxes allow us to approximate a volume of arbitrary shape.

Let us consider a rectangular box whose portion of the bottom face coincides with a portion of the top face of the original box, as shown in Figure A.2. We assume that the second box is also entirely contained within field F. Since we could demonstrate the validity of equation (A.1) for the new box by following the argument identical to the one described above, let us focus our attention on the portion of the plane that is common to both boxes, and let us denote this portion by S_c. We can write the flux through S_c that is associated with the lower box as

$$\iint_{S_c} [F_1, F_2, F_3] \cdot [0, 0, 1] \, \mathrm{d}S = \iint_{S_c} F_3 \mathrm{d}S$$

and the flux through S_c that is associated with the upper box as

$$\iint_{S_c} [F_1, F_2, F_3] \cdot [0, 0, -1] \, \mathrm{d}S = - \iint_{S_c} F_3 \mathrm{d}S,$$

where the only difference consists of the opposite directions of the normal vectors. If we consider the flux associated with both boxes, we see — by examining the right-hand sides of the above expressions — that the effects across the coinciding portions cancel one another. In view of the cancelling of the coinciding portions, we can conclude that the flux through the outer surface is equal to sum of the fluxes out of all interior pieces. Hence, if we build an arbitrary shape from many boxes, the result describes the flux across the outer surface.

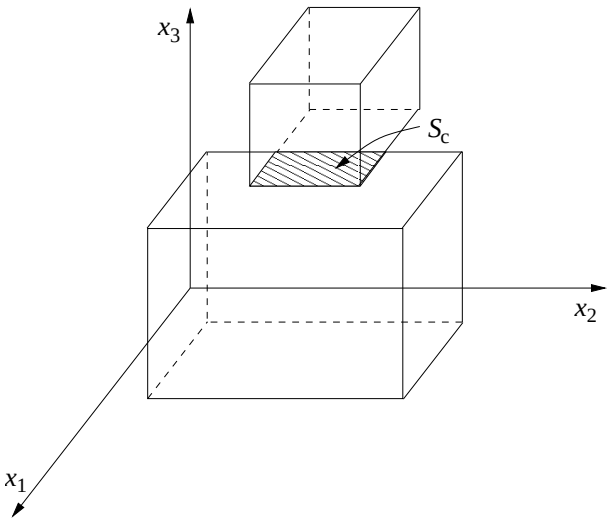


Fig. A.2. Two connected rectangular boxes used to formulate the divergence theorem

Studying the divergence theorem in the context of mathematical physics, we learn about the physical properties of the entities involved. For instance, if surface S is contained in the domain of the divergence-free vector field, $\nabla \cdot F = 0$, then there is no flux across this surface. A physical example of such a vector field is given in expression (F.4). Furthermore, using vector-calculus identities, one could show that $\nabla \cdot F = 0$ implies that $F = \nabla \times A$, where $A = A(x_1, x_2, x_3)$ is a scalar field.

Example 7.1. Consider the following vector field.

$$F = \left[\frac{x}{(x^2+y^2+z^2)^{3/4}}, \frac{y}{(x^2+y^2+z^2)^{3/4}}, \frac{z}{(x^2+y^2+z^2)^{3/4}} \right]$$

This vector field is not differentiable at the origin. Everywhere else, its divergence is zero. In accordance with the divergence theorem, flux of this vector field through a surface that does not enclose the origin is zero. However, a flux through a surface that encloses the origin is nonzero. This result emphasizes the differentiability requirement in the statement of Theorem A.1.

A.2 Curl theorem

Statement

In this section, we state and justify the curl theorem, which relates a line integral in a vector field to the surface integral in this field. We use this theorem to formulate Faraday's law in Appendix F.

The theorem that we discussed in Section A.1 relates the volume integral of a derivative of a vector field to the surface integral of this field over the closed surface bounding this volume. Herein, we will consider the theorem that relates the surface integral of a derivative of a vector field to the line integral of this field along the loop that bounds an area of this surface.

Curl theorem: The line integral of a continuously differentiable vector field along a closed loop can be expressed as the surface integral of the curl of this field along a surface bounded by this loop. In other words,

$$\int_C F \cdot n \, \mathrm{d}s = \iint_S [(\nabla \times F) \cdot N] \, \mathrm{d}S, \tag{A.4}$$

where $\mathrm{d}s$ is the element of length along the curve, n is the unit vector tangent to the curve and N is the unit normal vector to S.

Hence, the integral of the tangential component of F around boundary C is equal to the integral of the normal component of the curl of F on the surface enclosed by C.

In Section A.1, we invoked the definition of the flux of a vector field as the surface integral of the normal component of this field. Herein, we invoke the definition of the circulation of a vector field as the integral around the loop of the component of the field that is tangent to this loop. Thus, the above theorem states that the circulation of field F around loop C can be expressed as the surface integral of the component of the curl of this field that is normal to S.

Note that, unlike for the divergence theorem, herein the surface is not closed. Also, since the direction of the unit normal vector, N, changes the sign of the surface integral and the orientation of $\mathrm{d}s$ changes the sign of the curve integral, we must decide on the sign convention. We choose to orient the surface in such a way that vector N points in the direction of the thumb of the right hand with the fingers curled in the direction of $\mathrm{d}s$.

Verification

A general proof of the curl theorem is quite involved. Below, we will provide an argument that makes the general statement of the theorem plausible.

Let us consider vector field $F = [F_1, F_2, F_3]$. Let us also consider a rectangle within this field. For convenience, we position the rectangle in the x_1x_2-plane and set the edges of this rectangle to be parallel with two coordinate axes. As shown in Figure A.3, each edge extends between a_1 and a_2, b_1 and b_2, along the x_1-axis, the x_2-axis, respectively.

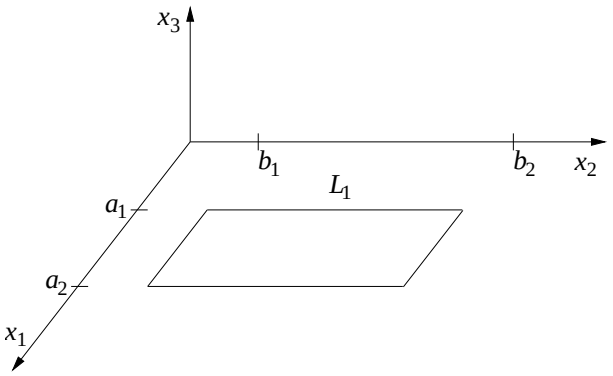


Fig. A.3. Rectangles used to formulate the curl theorem

We can parametrize the far edge in Figure A.3 by

$$L_1 : x_2 \mapsto (a_1, x_2, 0)\,,$$

with $x_2 \in [b_1, b_2, 0]$. In all integrations, we will consider the counterclockwise direction. Hence, the line integral of the component of the vector field that is tangent to L_1 is $\int_{b_2}^{b_1} F_2\,(a_1, x_2, 0)\,\mathrm{d}x_2$. Similarly, the line integral along the near edge is $\int_{b_1}^{b_2} F_2\,(a_2, x_2, 0)\,\mathrm{d}x_2$. Following the same approach for the other two edges of the rectangle and, then, summing all four segments, we get

$$\int\limits_C F \cdot n\,\mathrm{d}s = \int\limits_{b_2}^{b_1} F_2\,(a_1, x_2, 0)\,\mathrm{d}x_2 + \int\limits_{b_1}^{b_2} F_2\,(a_2, x_2, 0)\,\mathrm{d}x_2 + \int\limits_{a_1}^{a_2} F_1\,(x_1, b_1, 0)\,\mathrm{d}x_1 + \int\limits_{a_2}^{a_1} F_1\,(x_1, b_2, 0)\,\mathrm{d}x_1.$$

Changing the limits of integration, we can rewrite this equation as

$$\int\limits_C F \cdot n\,\mathrm{d}s = \int\limits_{b_1}^{b_2} \left[F_2\,(a_2, x_2, 0) - F_2\,(a_1, x_2, 0)\right]\mathrm{d}x_2 - \int\limits_{a_1}^{a_2} \left[F_1\,(x_1, b_2, 0) - F_1\,(x_1, b_1, 0)\right]\mathrm{d}x_1.$$

In view of the Fundamental Theorem of Calculus, we can write

$$\int_C F \cdot n \, \mathrm{d}s = \int_{b_1}^{b_2} \int_{a_1}^{a_2} \frac{\partial F_2\left(x_1, x_2, 0\right)}{\partial x_1} \mathrm{d}x_1 \mathrm{d}x_2 - \int_{a_1}^{a_2} \int_{b_1}^{b_2} \frac{\partial F_1\left(x_1, x_2, 0\right)}{\partial x_2} \mathrm{d}x_2 \mathrm{d}x_1.$$

Changing the order of integration, we can combine the two integrals to get

$$\int_C F \cdot n \, \mathrm{d}s = \int_{b_1}^{b_2} \int_{a_1}^{a_2} \left[\frac{\partial F_2\left(x_1, x_2, 0\right)}{\partial x_1} - \frac{\partial F_1\left(x_1, x_2, 0\right)}{\partial x_2} \right] \mathrm{d}x_1 \mathrm{d}x_2.$$

We recognize that the term in parentheses is the x_3-component of the curl of the vector field, $\left(\nabla \times F\left(x_1, x_2, x_3\right)\right)_3$. In view of the properties of the curl, we conclude that this component is normal to the plane containing the rectangle. Since we can always orient the coordinate system in such a way that its $x_1 x_2$-plane coincides with the rectangle, we can write

$$\int_C F \cdot n \, \mathrm{d}s = \iint_A \left[(\nabla \times F) \cdot N \right] \mathrm{d}a,$$

where N is the unit vector normal to rectangle A whose area element is $\mathrm{d}a = \mathrm{d}x_1 \mathrm{d}x_2$.

At this point, we have demonstrated the curl theorem for a plane segment bounded by a rectangle. We wish to extend it to a surface bounded by a loop in three dimensions. The motivation for our approach is the fact that small adjacent rectangles of different orientation allow us to approximate a surface of arbitrary shape.

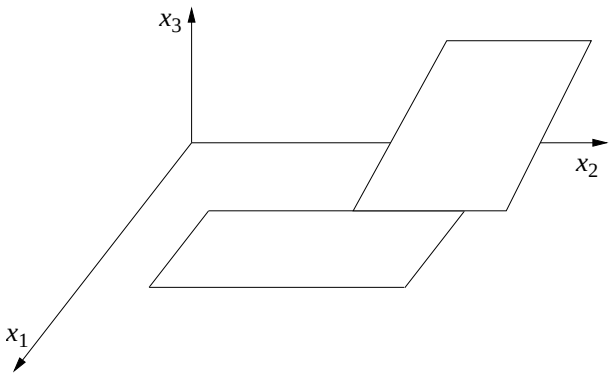


Fig. A.4. Two connected rectangles used to formulate the curl theorem

Let us consider two rectangles that touch one another along the portion of one edge, as shown in Figure A.4. Since we could demonstrate the validity of equation (A.4) for the new rectangle by following the argument identical to the one described above, let us focus our attention on the portion of the edge that is common to both rectangles. Since we chose the counterclockwise direction to describe the circulation along any curve, this means that for either rectangle the direction is opposite along the common portion. Consequently, along this portion, we get two integrals whose values differ by a sign. These integrals will cancel one another in the summation. In view of this cancelling, we conclude that the circulation through the outer loop is equal to the

sum of the circulations through all interior pieces. Furthermore, since the adjacent rectangles need not be coplanar, we can approximate an arbitrary surface bounded by a curve in three dimensions to obtain equation (A.4), as required.

Example 7.2. Consider the following vector field.

$$F = \left[\frac{y}{x^2 + y^2}, \frac{-x}{x^2 + y^2}, 0 \right]$$

This vector field is not differentiable along the z-axis. Everywhere else, its curl is zero. In accordance with the curl theorem, circulation of this vector field along a loop that does not enclose the z-axis is zero. However, a circulation along a loop that encloses the z-axis is nonzero. This result emphasizes the differentiability requirement in the statement of Theorem A.2.

B Fourier transform

In this appendix we look at a decomposition of functions to sines and cosines, or equivalently, to different powers of $e^{\imath x}$. We can motivate the need to study such decomposition by the fact that the function e^{ax} has the unique property: if we differentiate it, we obtain the same function up to the scalar multiple a. Thus, it is not surprising that such a function is important for solving differential euqations.

Solutions of differential equations are expressed in terms of functions. Hence, we start this appendix by studying different types – or spaces – of functions. Some functions are more relevant for such a study than others. An example of a function that may be preferable to look at whilst studying differential equations is a differentiable function. However, by far not all the functions are differentiable. The space of all differentiable functions is an example of a functional space. Another example of a functional space is the space of all integrable functions. It is clear that these two spaces contain very different functions and each space of functions may be useful for studying different problems.

Often, it is convenient to restrict a space of functions to some subspace that may have some more desirable properties than the original space. Then a natural question arises if it is possible to approximate an element of the bigger space by some element of the subspace. To be able to speak about approximations, one needs a notion of similarity of two functions in the big space. One can reformulate the question of similarity of two functions to the question of smallness of the difference of the two functions. This leads one to the notion of a norm. The norm of a function associates to a function its length, or size in some convenient manner. In some cases it is possible to construct a scalar product of two functions from the knowledge of this norm.

To illustrate the above facts, in the next section we will study some of the spaces of functions on a closed interval.

B.1 Some spaces of functions defined on closed interval

Since physical systems usually span only a limited space, in many physical applications we are interested only in functions defined on some closed domain. In cases of one dimensional problems, this reduces to study of functions on a closed interval.

An example of such a system is a vibrating string of length l. Mathematically, we can represent position of each point of the string at some fixed time t by a continuous function on closed interval $[0, l]$.

It is natural to ask what operations on functions from this space preserve this space. For example, if we add two functions that are continuous we obtain a function that is continuous as well. Similarly, if we multiply a continuous function by a real number, we obtain a continuous function. By similar arguments we can conclude that continuous functions on closed interval I form a vector space. We will denote this vector space by $C(I)$.

To study similarity of functions, we need to determine if a difference of two functions is small enough for our purposes. Thus, we pose the following question. What is the natural way how to define a length of a function?

If we consider only a real valued function, a natural candidate for the definition of this length, called a norm of a function, is given by the following formula.

$$\|f\| = \sqrt{\int_I f^2(x)\,\mathrm{d}x}.$$

If we consider also a complex valued functions, it is necessary to change the above formula in the following way.

$$\|f\| = \sqrt{\int_I f(x)\,\overline{f(x)}\mathrm{d}x} = \sqrt{\int_I |f(x)|^2\,\mathrm{d}x}, \tag{B.1}$$

where $\bar{f}$ represents complex conjugate of f and $|f|$ represents modulus of f. It is straightforward to see that this definition of a norm satisfies all the properties that are required for a norm; namely,

nondegeneracy: $\|f\| = 0$ only if $f = 0$,
triangular inequality: $\|f + g\| \leq \|f\| + \|g\|$,
absolute homogeneity: $\|af\| = |a|\,\|f\|$, for any number a.

The norm provides us with measuring the "size" of a function. We will restrict our attention only to the functions whose norm is a real number. In other words, we are interested in functions belonging to the space defined by

$$L^2(I) = \{f(x) \mid \|f\| < \infty\},$$

which is called the space of square integrable functions. It is clear that continuous functions on a closed interval are square integrable, thus we can write that $C(I) \subset L^2(I)$.

Now we can ask if it is possible to construct a scalar product, also known as dot product, on this vector space. While dealing with vectors in an n-dimensional vector space, we define the scalar product of vectors x and y as

$$\langle x, y\rangle \equiv x \cdot y = \sum_{i=1}^{n} x_i y_i, \tag{B.2}$$

where x_i and y_i are the components of vectors x and y in a fixed basis.

We can view an n-dimensional vector, $v = [v_1, v_2, \ldots, v_n]$, as a function $v : \{1, 2, \ldots, n\} \to \mathbb{R}$, where $v(i) = v_i$. Also, we can view function $f(x)$ as a vector with an infinite number of components. In other words, $f(x)$ exists in an infinite-dimensional space, where the value of each component is $f(x)$ for a given x from interval I. Since an integral is a summation, in view of expression (B.2), we define the scalar product of functions f and g in $L^2(I)$ as

$$\langle f, g\rangle = \int_I f(x)\,\overline{g(x)}\mathrm{d}x.$$

One can readily check that this has all the necessary properties of a scalar product, namely

commutativity: $\langle f, g\rangle = \langle g, f\rangle$
linearity: $\langle af + g, h\rangle = a\langle f, h\rangle + \langle g, h\rangle$ for any real a
positive definiteness: $\langle f, f\rangle \geq 0$ and is equal to zero only if $f = 0$.

The scalar product allows us to define perpendicularity of functions, and thus define orthogonal projections, which are useful for decomposition of functions.

This vector space has many of the properties of the vector space $\mathbb{R}^n$ with the standard scalar product $\langle x, y\rangle \equiv x \cdot y = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n$. For example, the Pythagorean theorem holds also for continuous functions on a closed interval.

It is important to note that the norm $\|\|$ is induced by the scalar product $\langle , \rangle$; namely, they are related by the following formula.

$$\|f\|^2 = \langle f, f\rangle .$$

We note that it is a general fact that a norm is induced by a scalar product if the norm satisfies the parallelogram law:

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

One of the main differences between $\mathbb{R}^n$ and $L^2(I)$ is the fact that $L^2(I)$ is infinite dimensional. This means that there are infinitely many linearly independent functions . We have to clarify what we mean by linear independence for infinitely many functions: We generalize the notion of linear independence of finite number of vectors by saying that a set of functions is linearly independent if any finite subset is linearly independent.

B.2 Fourier series

To motivate the discussions in the following section, we will look at a vibrating string with fixed ends.

Vibrating string.

We will consider a string of length L with fixed ends. The equation of motion for a vibrating string with small displacement is the following wave equation.

$$\frac{\partial^2 u}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2},$$

where u denotes a displacement of the string at the distance x from an end at the time t. Also, since the length of the string is L and the ends are fixed, the solutions must satisfy

$$u(0,t) = u(L,t) = 0$$

for all times t. Moreover, we supply also the initial displacement $u(x,0) = u_0(x)$ and the initial velocity $\partial u/\partial t|_{t=0} = v_0(x)$ as the side conditions.

We can try to solve this equation by separation of variables. Assume that we can write the solution in the following form, namely,

$$u(x,t) = \sum_{k=1}^{\infty} f_k(t) \sin\left(\frac{\pi k}{L}x\right).$$

Substituting this into the wave equation, we obtain

$$\frac{\mathrm{d}^2 f_k(t)}{\mathrm{d}t^2} = -\left(\frac{\pi k}{L}c\right)^2 f_k(t).$$

Solution of this equation is

$$f_k(t) = a_k \sin\left(\frac{\pi k}{L} ct\right) + b_k \cos\left(\frac{\pi k}{L} t\right).$$

Thus, the solution of the wave equation of a vibrating string is

$$u(x,t) = \sum_{k=1}^{\infty} \left(a_k \sin\left(\frac{\pi k}{L} ct\right) + b_k \cos\left(\frac{\pi k}{L} t\right)\right) \sin\left(\frac{\pi k}{L} x\right). \tag{B.3}$$

We still have to determine the coefficients a_k and b_k. These are determined from the side condition, namely from the requirements that

$$\begin{aligned} u(x,0) &= u_0(x) \\ \frac{\partial u}{\partial t}(x,0) &= v_0(x). \end{aligned}$$

After substituting zero for time in expression (B.3) and equating it to $u_0(x)$, we obtain that

$$\sum_{k=1}^{\infty} b_k \sin\left(\frac{\pi k}{L} x\right) = u_0(x).$$

After differentiating expression (B.3) with respect to time at $t = 0$ and equating it to $v_0(x)$, we obtain that

$$\sum_{k=1}^{\infty} a_k \frac{\pi k}{L} c \sin\left(\frac{\pi k}{L} x\right) = v_0(x).$$

The last two expressions are expressions of a function as a trigonometric series. The detailed discussion of these series and the method of determining the coefficients are given in the following section.

Fourier series

To illustrate the properties of $L^2(I)$, let us consider $I = [0, L]$. In this case, an important example of linearly independent and mutually orthogonal functions is given by

$$\begin{aligned} \phi_0 &= 1, \\ \phi_k &= \cos\left(\frac{2\pi}{L} kx\right), \\ \psi_k &= \sin\left(\frac{2\pi}{L} kx\right), \end{aligned}$$

where $k = 1, 2, 3, \ldots$. The proof of their mutual orthogonality and of their linear independence is left to the reader.

Now we can ask if it is possible to express any continuous function on this interval as a sum of some coefficients times these orthogonal functions. In other words, we ask if it is possible to find coefficients $a_0, a_1, \ldots$ and $b_1, b_2, \ldots$ for a given function $f \in L^2(I)$, such that f can be expressed as

$$f(x) = a_0 + \sum_{k=1}^{\infty} a_k \cos\left(\frac{2\pi}{L} kx\right) + \sum_{k=1}^{\infty} b_k \sin\left(\frac{2\pi}{L} kx\right), \tag{B.4}$$

where we must clarify what we mean by these infinite sums.

Let us assume that it is possible to express f in this manner and let us determine these coefficients. To do so, we can proceed analogously to finding components of a vector expressed in some orthogonal basis in $\mathbb{R}^n$. If the basis vectors were orthonormal, the coefficients would be the orthogonal projections of function (vector) f to the directions of functions (vectors) ϕ_k and ψ_k. These projections are given by the scalar products of f with the considered function. Since the norms of the vectors are not equal to one, we have to scale these projections by an appropriate factor. This factor is equal to the square of the norm of the vector. Hence,

$$a_k = \frac{\langle f, \phi_k \rangle}{\|\phi_k\|^2},$$
$$b_k = \frac{\langle f, \psi_k \rangle}{\|\psi_k\|^2}.$$

Since the norms of the vectors are

$$\|\phi_k\| = \sqrt{\int_0^L \left|\sin\left(\frac{2\pi}{L} kx\right)\right|^2 \mathrm{d}x} = \|\psi_k\| = \sqrt{\int_0^L \left|\cos\left(\frac{2\pi}{L} kx\right)\right|^2 \mathrm{d}x} = \sqrt{\frac{L}{2}}$$

and

$$\|\phi_0\| = \sqrt{\int_0^L 1 \mathrm{d}x} = \sqrt{L},$$

we can write the coefficients as

$$\begin{aligned} a_0 &= \frac{1}{L} \int_0^L f(x)\, \mathrm{d}x, \\ a_k &= \frac{2}{L} \int_0^L f(x) \cos\left(\frac{2\pi}{L} kx\right) \mathrm{d}x, \\ b_k &= \frac{2}{L} \int_0^L f(x) \sin\left(\frac{2\pi}{L} kx\right) \mathrm{d}x. \end{aligned} \tag{B.5}$$

Decomposition (B.4) of a function to the directions of ϕ_n and ψ_n is called the Fourier series and coefficients (B.5) are called Fourier coefficients.

Invoking

$$A\cos\theta + B\sin\theta = \sqrt{A^2 + B^2} \cos\left(\theta - \arctan\frac{B}{A}\right),$$

we can rewrite Fourier series (B.4) as

$$f(t) = a_0 + \sum_{k=1}^{\infty} d_k \cos\left(\frac{2\pi}{L} kx - \phi_x\right),$$

where

$$d_k = \sqrt{a_k^2 + b_k^2}$$

and

$$\phi_k = \arctan \frac{b_k}{a_k}.$$

Thus, we see that a function on interval I can be expanded as a sum of cosines. The kth cosine has amplitude d_k, angular frequency $2k\pi/L$, and phase ϕ_k. Sequence $\{d_n\}$ is a discrete amplitude spectrum of f and sequence $\{\phi_n\}$ is a discrete phase spectrum of f.

Note that if f is an odd function then all coefficients a_k are equal to zero and if f is an even function then all coefficients b_k vanish.

Now we have to answer the question of meaning of the infinite sums in the decomposition. To do this, we consider partial sums, namely,

$$f_n(x) = a_0 + \sum_{k=1}^{n} a_k \cos\left(\frac{2\pi}{L} kx\right) + \sum_{k=1}^{n} b_k \sin\left(\frac{2\pi}{L} kx\right)$$

and investigate the limit as $n \to \infty$. Naturally, the notion of a limit is closely related to the notion of a distance of two functions, or equivalently, to the notion of a norm. Since we want to express $f(x)$ as the infinite sum, we wish to establish the following equality.

$$\lim_{n\to\infty} \|f(x) - f_n(x)\| = 0.$$

This equality gives meaning to the the convergence of the Fourier series.

Example 7.3. Find the Fourier series for the function $f(x) = x$ on interval $x \in [-\pi, \pi]$.

First we compute the Fourier coefficients using formula (B.5). Since function $f(x)$ is odd, we see that all the coefficients a_k are equal to zero. Coefficient b_n can be calculated as follows.

$$\begin{aligned} b_k &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(kx)\,\mathrm{d}x = \frac{1}{\pi}\int_{-\pi}^{\pi} x\sin(kx)\,\mathrm{d}x \\ &= \frac{2}{\pi}\int_{0}^{\pi} x\sin(kx)\,\mathrm{d}x = \frac{2}{\pi}\left[-\frac{\cos(kx)}{k}\right]_0^{\pi} + \frac{2}{\pi k}\int_0^{\pi}\cos(kx)\,\mathrm{d}x \\ &= \frac{2}{k}(-1)^{k+1}. \end{aligned}$$

Thus we can write the Fourier series as

$$x = \sum_{k=1}^{\infty} \frac{2}{k}(-1)^{k+1}\sin(kx).$$

The partial sum of first twenty terms from the Fourier series together with the first ten terms are shown in Figure B.1.

It is important to note that the Fourier series is convergent for all functions that are square integrable, i.e. functions that belong to the space $L^2(I)$.

B.3 Fourier transform

In this section, we start the discussion by considering a more compact formula for the Fourier series and the Fourier coefficients. This new form of the Fourier series is going to be a starting point in generalizing the Fourier series, which is the main focus of this section.

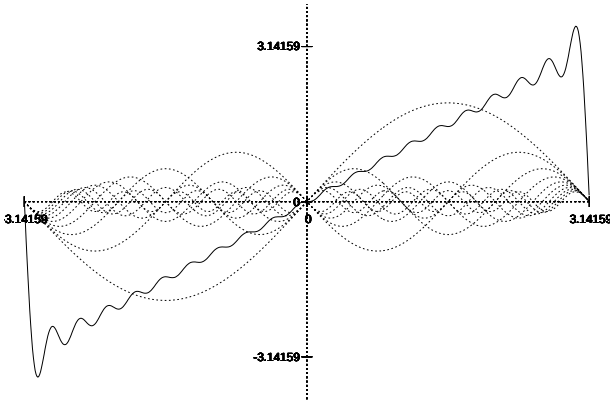


Fig. B.1. The first ten terms of the Fourier series of function $f(x) = x$ are shown in dotted lines and the partial sum of the first twenty terms is shown in a solid line.

We remind the reader the relation between the trigonometric functions and the complex exponential function, namely,

$$e^{ix} = \cos(x) + i\sin(x),$$

or equivalently,

$$\cos(x) = \frac{e^{ix} + e^{-ix}}{2}$$
$$\sin(x) = \frac{e^{ix} - e^{-ix}}{2i}.$$

Using these formulae, we can rewrite the Fourier series as

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{ikx},$$

where

$$c_0 = a_0, \quad c_k = \frac{1}{2}(a_n - ib_n), \quad c_{-k} = \frac{1}{2}(a_n + ib_n)$$

for $k = 1, 2, 3, \ldots$. In this new compact notation, the Fourier coefficients are given by

$$c_k = \frac{1}{L}\int_0^L f(x)\, e^{-ikx} \mathrm{d}x,$$

where $k = \ldots, -3, -2. - 1, 0, 1, 2, 3, \ldots$.

Next, we consider the Fourier series for a function defined on a closed interval $[-a, a]$. The Fourier series is defined in this case as

$$f(x) = \sum_{k=-\infty}^{\infty} c_{k,a} e^{ixk\frac{\pi}{a}},$$

where

$$c_{k,a} = \frac{1}{2\pi}\frac{\pi}{a}\int_{-a}^{a} f(\xi)\, e^{-ik\xi\frac{\pi}{a}} \mathrm{d}\xi.$$

We can combine these two expressions into a single expression

$$f(x) = \frac{1}{2\pi}\frac{\pi}{a}\sum_{k=-\infty}^{\infty} e^{ixk\frac{\pi}{a}} \int_{-a}^{a} f(\xi)\, e^{-ik\xi\frac{\pi}{a}} \mathrm{d}\xi.$$

Substituting $\Delta p := \pi/a$ and $p_k = k\Delta p = k\pi/a$, we can write the above as

$$f(x) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty} e^{ixp_k} \left(\int_{-a}^{a} f(\xi)\, e^{-ip_k\xi} \mathrm{d}\xi \right) \Delta p.$$

We observe that this expression resembles a Riemann sum defining the integral in a new variable p. Indeed, if we take a formal limit of the above expression as $a \to \infty$, we obtain that $\Delta p \to 0$ and

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ixp} \int_{-\infty}^{\infty} f(\xi)\, e^{-ip\xi} \mathrm{d}\xi \mathrm{d}p. \tag{B.6}$$

From this formal identity we could guess that if we define a transform

$$f(x) \mapsto \hat{f}(p) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)\, e^{-ipx} \mathrm{d}x, \tag{B.7}$$

then its inverse transform is given by

$$\hat{f}(p) \mapsto f(x) = \int_{-\infty}^{\infty} \hat{f}(p)\, e^{ixp} \mathrm{d}p. \tag{B.8}$$

Expression (B.7) defines the Fourier transform and expression (B.8) defines the inverse Fourier transform. We still have to specify the precise meaning of these two definitions and rigorously establish their inverse.

The Fourier transform defined by formula (B.7) transforms absolutly integrable functions to continuous functions, $f \in L^1(\mathbb{R}) \mapsto \hat{f} \in C(\mathbb{R})$. One can extend this definition to the square-integrable functions using the fact that compactly supported functions are dense in $L^2(\mathbb{R})$.

Now, we will discuss the Fourier transform for functions of more variables. If we are dealing with function $f(x)$, where $x = [x_1, x_2, \ldots, x_n]$, we can extend the Fourier transform to

$$\hat{f}(p) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} f(\xi) \exp\left[-i\xi \cdot p\right] \mathrm{d}\xi, \tag{B.9}$$

where $\xi = [\xi_1, \xi_2, \ldots, \xi_n]$ and $p = [p_1, p_2, \ldots, p_n]$. In this expression, $\int_{\mathbb{R}^n}$ stands for n integrals from $-\infty$ to ∞, $p \cdot x$ and $\xi \cdot p$ are the scalar products, and $\mathrm{d}\xi := \mathrm{d}\xi_1 \mathrm{d}\xi_2 \ldots \mathrm{d}\xi_n$ and $\mathrm{d}p := \mathrm{d}p_1 \mathrm{d}p_2 \ldots \mathrm{d}p_n$. The inverse transform can be then written as

$$f(x) = \int_{\mathbb{R}^n} \hat{f}(p) \exp\left[ix \cdot p\right] \mathrm{d}p. \tag{B.10}$$

Example 7.4. We often study the Fourier transform of function $f(t_1, t_2)$. In such a case, a pair of transforms of function f is given by expressions (B.9) and (B.10), which become

$$\hat{f}\left(\alpha_1,\alpha_2\right)=\frac{1}{\left(2\pi\right)^2}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}f\left(t_1,t_2\right)\exp\left[-i\left(\alpha_1t_1+\alpha_2t_2\right)\right]\mathrm{d}t_1\mathrm{d}t_2$$

and

$$f\left(t_1,t_2\right)=\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\hat{f}\left(\alpha_1,\alpha_2\right)\exp\left[i\left(\alpha_1t_1+\alpha_2t_2\right)\right]\mathrm{d}\alpha_1\mathrm{d}\alpha_2,$$

respectively.

Herein, $f\left(t_1,t_2\right)$ is a surface spanned by t_1 and t_2, while $\hat{f}\left(\alpha_1,\alpha_2\right)$ is a surface spanned by α_1 and α_2. In this context, we can view f and $\hat{f}$ as maps of one another in different domains.

There are several ways to state the Fourier transform pair. For instance, if we replace α by $-\alpha$ in the above double integrals, we would get

$$\hat{f}\left(\alpha\right)=\frac{1}{2\pi}\int\limits_{-\infty}^{\infty}f\left(t\right)\exp\left(i\alpha t\right)\mathrm{d}t$$

and

$$f\left(t\right)=\int\limits_{-\infty}^{\infty}\hat{f}\left(\alpha\right)\exp\left(-i\alpha t\right)\mathrm{d}\alpha.$$

Also, we could change the position of factor $1/2\pi$ to be, for example,

$$\hat{f}\left(\alpha\right)=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}f\left(t\right)\exp\left(i\alpha t\right)\mathrm{d}t$$

and

$$f\left(t\right)=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}\hat{f}\left(\alpha\right)\exp\left(-i\alpha t\right)\mathrm{d}\alpha,$$

where

$$\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2\pi}}=\frac{1}{2\pi},$$

as required. Notably, if we let $t=2\pi\tau$, we get

$$\hat{f}\left(\alpha\right)=\int\limits_{-\infty}^{\infty}f\left(\tau\right)\exp\left(2\pi i\alpha\tau\right)\mathrm{d}\tau$$

and

$$f\left(\tau\right)=\int\limits_{-\infty}^{\infty}\hat{f}\left(\alpha\right)\exp\left(-2\pi i\alpha\tau\right)\mathrm{d}\alpha.$$

To study the implications of the Fourier transform on a differential equation, we have to investigate the transform of a derivative of a function, namely,

$$\frac{\hat{\mathrm{d}f}}{\mathrm{d}x}=\int_{-\infty}^{\infty}\frac{\mathrm{d}f}{\mathrm{d}x}e^{-ipx}\mathrm{d}x.$$

To solve this integral, we use the method of integration by parts,

$$\begin{aligned}\widehat{\frac{\mathrm{d}f}{\mathrm{d}x}} &= \int_{-\infty}^{\infty} \frac{\mathrm{d}f}{\mathrm{d}x} e^{-ipx} \mathrm{d}x \\ &= \lim_{a\to\infty} \left[f(x)\, e^{-ipx} \right]_{x=-a}^{a} + ip \int_{-\infty}^{\infty} f(x)\, e^{-ipx} \mathrm{d}x.\end{aligned}$$

If function $f(x)$ vanishes at the positive and negative infinity, then the limit term in the above expression vanishes as well. The integral in the above expression is nothing else but the Fourier transform of $f(x)$, and hence, the Fourier transform of a derivative can be written as

$$\widehat{\frac{\mathrm{d}f}{\mathrm{d}x}} = ip\hat{f}.$$

This is an important formula. It states that the Fourier transform transforms the operation of differentiation into the operation of multiplication by the independent variable in the transformed domain.

What happens if we Fourier transform a higher-order derivative? The reader can check that the kth-order derivative transforms as follows.

$$\widehat{\left(\frac{\mathrm{d}^k f}{\mathrm{d}x^k}\right)} = (ip)^k\, \hat{f}.$$

This formula allows us to transform a linear differential equation into an algebraic equation. This algebraic equation is often easier to solve, as exemplified in the following example.

Example 7.5. Solve the following damped wave equation using the Fourier transform.

$$m\frac{\mathrm{d}^2 f}{\mathrm{d}t^2}(t) + \gamma \frac{\mathrm{d}f}{\mathrm{d}t}(t) + kf(t) = 0.$$

The Fourier transform transforms this equation into the following equation.

$$\begin{aligned}(ip)^2 m\hat{f}(p) + \gamma ip\hat{f}(p) + k\hat{f}(p) &= 0 \\ \left(-mp^2 + \gamma ip + k\right)\hat{f}(p) &= 0.\end{aligned}$$

From here it follows that $\hat{f}$ can be nonzero only where $-mp^2 - \gamma ip + k$ is equal to zero. Thus, we can write that

$$\hat{f}(p) = A_1 \delta(p_1) + A_2 \delta(p_2),$$

where $A_1,\ A_2$ are constants and $p_1,\ p_2$ are the roots of

$$-mp^2 + \gamma ip + k = 0,$$

namely,

$$p_{1,2} = \frac{\gamma i \mp \sqrt{4mk - \gamma^2}}{2m}$$

To obtain a solution of the original differential equation, we use the inverse transform to get

$$\begin{aligned} f(t) &= \frac{1}{2\pi}\int_{\infty}^{\infty} e^{itp}\hat{f}(p)\,\mathrm{d}p \\ &= \frac{1}{2\pi}\int_{\infty}^{\infty} e^{itp}\left(A_1\delta(p_1)+A_2\delta(p_2)\right)\mathrm{d}p \\ &= \frac{A_1}{2\pi}e^{itp_1}+\frac{A_2}{2\pi}e^{itp_2} \\ &= e^{-\frac{\gamma}{2m}t}\left(\frac{A_1}{2\pi}e^{-it\sqrt{4mk-\gamma^2}/2m}+\frac{A_2}{2\pi}e^{it\sqrt{4mk-\gamma^2}/2m}\right). \end{aligned}$$

Note that if $4mk=\gamma^2$ then we will not obtain a general solution $f(x)=e^{-\gamma t/2m}(A+Bt)$. To obtain this solution, we would need to consult more theory of ordinary differential equations.

C Distributions

Expression (2.19) does not require either g or h to be differentiable or even continuous. This is consistent with our physical intuition developed by observing wave phenomena. This is also consistent with our using the characteristics to construct solution f in the x_1x_2-plane; we set the value of f along a curve, γ, that is not a characteristic and that intercepts each characteristic curve at a single point. Herein, f is constant along each characteristic, so having set f along γ, we immediately obtain a solution for the entire x_1x_2-plane. In general, f need not be constant along the characteristics, as illustrated in solution (2.3), yet the same method for constructing the solution can be used. There is, however, an important question to consider.

Since, equation (2.17) is a second-order differential equations, to obtain a solution in the classical sense, one would require f to be at least twice differentiable, so that we can insert the expression for f into equation (2.17). This distinction between the requirement for differentiability of f in equation (2.17) and solution (2.19) was the cause of a long debate between Euler and d'Alembert in the middle of the eighteenth century. In this debate, Euler, in view of solution (2.19) and from physical considerations of wave propagation, considered it acceptable and necessary to admit nondifferentiable functions as solutions, while d'Alembert, in view of equation (2.17), strictly required differentiability. This problem was resolved by Sobolev and by Schwartz in the middle of the twentieth century by considering as possible solutions so called generalized functions, or distributions.

In this appendix, we will introduce this entities. We start our discussion with the definition of the Dirac delta function, which is probably the best known example of a distribution. We continue the exposition with the definition of distirbutions and conclude the appendix with a section on properties of and opperations on distributions.

C.1 Dirac's delta

In general, Dirac's delta is defined by its properties stating that

$$\delta(t)=0, \qquad t\neq 0,$$

and, for any well-behaved function, f,

$$f(t)|_{t=0} = \int_{-\infty}^{\infty} \delta(t) f(t) \, \mathrm{d}t. \tag{C.1}$$

In particular, setting $f(t) = 1$, we obtain

$$1 = \int_{-\infty}^{\infty} \delta(t) \, \mathrm{d}t, \tag{C.2}$$

since $f(t) = 1$ for all t, including $t = 0$.

We could suggest that integral (C.2) implies that δ is an infinitely high and infinitely thin spike. We can obtain the desired results as the limit of a sequence of functions.

For our study, let us consider the sequence given by

$$\delta_n(t) = \frac{\sin(nt)}{\pi t}. \tag{C.3}$$

Figure C.1 shows several members of this sequence. As n increases, the central lobe of this graph gets higher and thinner, as required.

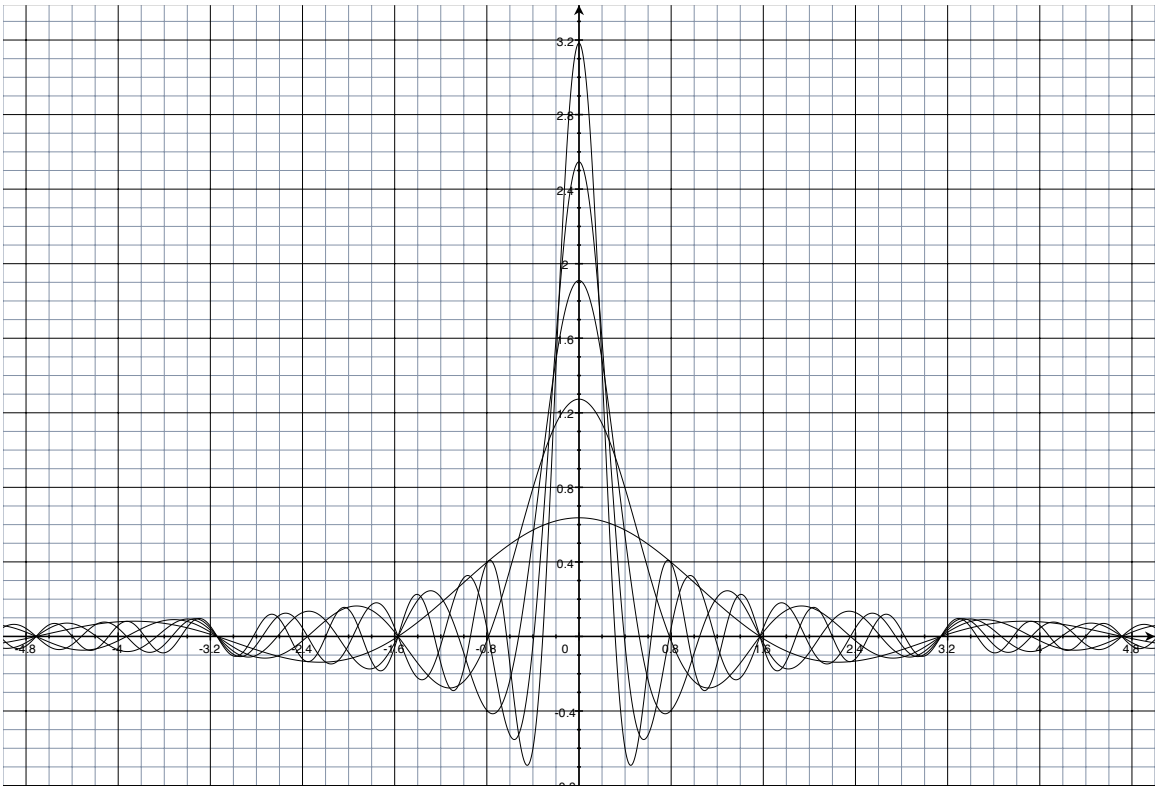


Fig. C.1. Several members of sequence (C.3), which defines Dirac's delta

To justify the use of sequence (C.3) as representing δ, let us consider the defining properties of δ. In view of property (C.1), we require that

$$\lim_{n\to\infty}\int_{-n}^{n}\frac{\sin(nt)}{\pi t}f(t)\,\mathrm{d}t = f(t)|_{t=0}\,.$$

Taking the limit, we get

$$\lim_{n\to\infty}\int_{-\infty}^{\infty}\frac{\sin(nt)}{\pi t}f(t)\,\mathrm{d}t = f(0)\,,$$

which justifies our use of sequence (C.3) for δ.

To put sequence (C.3) in the context of Fourier analysis, let us consider

$$\int_{-n}^{n}\cos(\alpha t)\,\mathrm{d}\alpha = \frac{\sin(\alpha t)}{t}\bigg|_{\alpha=-n}^{\alpha=n} = \frac{2\sin(nt)}{t}, \tag{C.4}$$

where, in the last step, we used the fact that sine is an antisymmetric function. Recalling Euler's relation, we can write

$$\int_{-n}^{n}e^{i\alpha t}\mathrm{d}\alpha = \int_{-n}^{n}\cos(\alpha t)\,\mathrm{d}\alpha + i\int_{-n}^{n}\sin(\alpha t)\,\mathrm{d}\alpha.$$

Again, since sine is an antisymmetric function, the second integral on the right-hand side vanishes and, recalling expression (C.4), we obtain

$$\int_{-n}^{n}e^{i\alpha t}\mathrm{d}\alpha = \frac{2\sin(nt)}{t}.$$

Hence, sequence (C.3) can be written as

$$\frac{\sin(nt)}{\pi t} = \frac{1}{2\pi}\int_{-n}^{n}e^{i\alpha t}\mathrm{d}\alpha.$$

Thus, in view of Figure C.1 and letting $n\to\infty$, we can formally write

$$\delta(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty}e^{i\alpha t}\mathrm{d}\alpha.$$

Using the transitional properties of δ, we can restate this expression as

$$\delta(t-x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}e^{i\alpha(t-x)}\mathrm{d}\alpha. \tag{C.5}$$

Dirac's delta and Fourier integral

To become familiar with expression (C.5) , consider the composition of the Fourier transform with its inverse to obtain

$$f(x) = \int_{-\infty}^{\infty} \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\alpha(t-x)} \mathrm{d}\alpha \right] f(t)\, \mathrm{d}t, \qquad x \in (-\infty, \infty).$$

Using expression (C.5) and considering all xs along the real axis, we can write

$$f(x) = \int_{-\infty}^{\infty} \delta(t-x)\, f(t)\, \mathrm{d}t, \qquad x \in (-\infty, \infty),$$

which, for $x = 0$, is expression (C.1).

Thus, we can regard Dirac's delta given by expression (C.5) in the context of the Fourier integral.

Before using Dirac's delta in our study of Fourier integral operators, let us investigate several of its aspects. We wish to emphasize that expression (C.1) is the key property defining Dirac's delta. As will become clear in Section C.2, Dirac's delta has a meaning only as part of an integrand. We can regard expression (C.1) as a linear operator $\delta(t)$ that operates — under the integral sign — on function $f(t)$ to give the value of $f(t)$ at $t = 0$. Since an operator that assigns a number to a function is called a functional, the integration of Dirac's delta with a function is a linear functional.

Dirac's delta and convolutions

Following the definition of convolution, which will be discussed in Section C.4, let us formally write

$$(\delta * f)(x) := \int_{-\infty}^{\infty} \delta(t-x)\, f(t)\, \mathrm{d}t, \qquad x \in (-\infty, \infty).$$

Examining thisexpression, we see that, for a well-behaved f, $\delta * f$ gives a number corresponding to $f|_{t=x}$, namely, $f(x)$.

C.2 Definition of distributions

Considering integral (C.2), we note that — within the realm of traditional functions — the value of the integral of a function that vanishes everywhere except at a single point is zero regardless of the value that the function assigns to that point. Furthermore, considering sequence (C.3), we notice that

$$\lim_{n\to\infty} \frac{\sin(nt)}{\pi t}$$

does not exist. It diverges for $t \neq k\pi$.

To obtain a rigorous formulation of Dirac's delta, we must consider the theory of distributions, where the key philosophical point is that "a function is not described by what it is but by its effect on other functions".

We define the effect of $h(t)$ on $f(t)$ by the value of the linear functional

$$\tilde{h}(f) := \int_{-\infty}^{\infty} h(t)\, f(t)\, \mathrm{d}t,$$

where we refer to $\tilde{h}$ as a distribution of h. Thus, instead of describing a function, h, on $\mathbb{R}^1$ by giving its values, $h(t)$, for appropriate points $t \in \mathbb{R}^1$, we describe it by the values of $\tilde{h}(f) = \int hf\mathrm{d}t$, for appropriate functions $f \in C_c^\infty$, where C_c^∞ is the space of infinitely differentiable and compactly supported functions, which means that f is nonzero on a closed and bounded interval but is zero everywhere else.

C.3 Operations on distributions

Since f is well-defined and compactly supported, we can perform regular operations on $\int hf\mathrm{d}t$, even though h might not be a standard function.

For instance, let us consider the derivative of h. In other words, let us look at the effect of the derivative of h on other functions, namely, $\int h'(t) f(t)\,\mathrm{d}t$. Integrating by parts, we obtain

$$\int_{-\infty}^{\infty} h'(t) f(t)\,\mathrm{d}t = h(t) f(t)|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} h(t) f'(t)\,\mathrm{d}t.$$

Since f is compactly supported, the first term on the right-hand side vanishes and we can state the effect of the derivative of h as

$$\int_{-\infty}^{\infty} h'(t) f(t)\,\mathrm{d}t = -\int_{-\infty}^{\infty} h(t) f'(t)\,\mathrm{d}t =: \partial\tilde{h}(f), \tag{C.6}$$

where $\partial\tilde{h}$ stands for the distributional derivative of h.

Formula (C.6) can be extended to derivatives of any order and we can write

$$\int_{-\infty}^{\infty} h^{(n)}(t) f(t)\,\mathrm{d}t = (-1)^n \int_{-\infty}^{\infty} h(t) f^{(n)}(t)\,\mathrm{d}t =: \partial^n\tilde{h}(f), \tag{C.7}$$

where $\partial^n\tilde{h}$ stands for the nth distributional derivative of h.

In formulae (C.6) and (C.7), we consider functions of a single variable. In general, formula (C.7) can be extended to functions of several variables. In such a case, we write

$$\partial_i^n\tilde{h}(f) := \int_{\mathbb{R}^m} \frac{\partial^n h(\mathbf{t})}{\partial t_i^n} f(\mathbf{t})\,\mathrm{d}\mathbf{t} = (-1)^n \int_{\mathbb{R}^m} h(\mathbf{t}) \frac{\partial^n f(\mathbf{t})}{\partial t_i^n}\mathrm{d}\mathbf{t},$$

where $\partial_i^n\tilde{h}$ stands for the nth distributional derivative of h with respect to t_i, and $\int_{\mathbb{R}^m}\mathrm{d}\mathbf{t}$ stands for the integral over all m variables from $-\infty$ to $+\infty$.

We note that the above formula has its analogue in the theory of convolutions. The derivative of a convolution of two C^1 functions is obtained by differentiating either one of these functions. Herein, if we consider f to be a C^∞ function, then all distributions are infinitely differentiable.

To illustrate expression (C.7), let us consider the nth derivative of Dirac's delta, $\partial^n\tilde{\delta}(f)$. We get

$$\partial^n\tilde{\delta}(f) = \int_{-\infty}^{\infty} \delta^{(n)}(t) f(t)\,\mathrm{d}t = (-1)^n \int_{-\infty}^{\infty} \delta(t) f^{(n)}(t)\,\mathrm{d}t = (-1)^n f^{(n)}(t)\Big|_{t=0},$$

where, in the last step, we used property (C.1). Considering the leftmost and the rightmost expressions, we see that $\delta^{(n)}$ can be regarded as a differential operator acting on f. Again, the above description emphasizes the fact that Dirac's delta has a meaning only as part of an integrand.

This brief illustration exemplified the fact that the theory of distributions generalizes the concept of differentiation of functions. Similar approach can also be used for the Fourier transform.

In general, we can write

$$\int_{-\infty}^{\infty} \hat{h}(t) f(t) \, \mathrm{d}t = \int_{-\infty}^{\infty} h(t) \hat{f}(t) \, \mathrm{d}t. \tag{C.8}$$

For instance, let us consider the Fourier transform of Dirac's delta. Following property (C.8), we can write

$$\int_{-\infty}^{\infty} \hat{\delta}(t) f(t) \, \mathrm{d}t = \int_{-\infty}^{\infty} \delta(t) \hat{f}(t) \, \mathrm{d}t.$$

Following property (C.1), we obtain

$$\int_{-\infty}^{\infty} \delta(t) \hat{f}(t) \, \mathrm{d}t = \hat{f}(t)\Big|_{t=0} = \hat{f}(0).$$

To interpret this result, consider the Fourier transform that is defined in expression (**??**), namely,

$$\hat{f}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(t) \exp(-ixt) \, \mathrm{d}t.$$

Setting $x = 0$, we get

$$\hat{f}(0) = \int_{-\infty}^{\infty} \frac{1}{2\pi} f(t) \, \mathrm{d}t.$$

Examining the first and last integrals in this formulation, we conclude that $\hat{\delta} = 1/2\pi$. In other words, the Fourier transform of Dirac's delta is a constant function for the entire domain. Furthermore, by the inversion theorem, it follows that $\delta = \widehat{1/2\pi}$, which allows us to regard Dirac's delta as the Fourier transform of a constant function.

We note that to obtain the Fourier transform we also wish to consider functions, f, that span from $-\infty$ to $+\infty$. Hence, f need not be compactly supported. Yet, we must require that f, together with all its derivatives, be rapidly decreasing at infinity. We refer to the resulting distributions as tempered distributions.

To conclude this section, we wish to emphasize that the integral used as the defining property of Dirac's delta, which is given in expression (C.1), is not a Riemann integral but rather an integral of a limit given by

$$\int_{-\infty}^{\infty} \delta(t) f(t) \, \mathrm{d}t = \lim_{n\to\infty} \int_{-\infty}^{\infty} \delta_n(t) f(t) \, \mathrm{d}t,$$

where δ is a distribution — and not a function — defined by sequence δ_n. Rigorous treatment of this integration requires the methods of the measure theory and the resulting concept of the Lebesgue integral.

C.4 Convolution

Since we have invoked the operation of convolution in our discussion of Dirac's delta, let us take a brief look at this operation.

The convolutions of two functions, $f(x)$ and $g(x)$, that are defined for all xs, is the function given by

$$h(x) = \int_{-\infty}^{\infty} f(\xi) g(x-\xi) \, \mathrm{d}\xi =: (f * g)(x).$$

In other words, a convolution of two functions of x is another function of x.

To see the properties of convolution in the context of the Fourier transform, let us consider

$$f(x) = u(x) \exp(-itx)$$

and

$$g(x) = v(x) \exp(-itx).$$

Convolving these two functions, we get

$$\int_{-\infty}^{\infty} u(\xi) \exp(-it\xi) v(x-\xi) \exp(-it(x-\xi)) \, \mathrm{d}\xi = \exp(-itx) \int_{-\infty}^{\infty} u(\xi) v(x-\xi) \, \mathrm{d}\xi.$$

To proceed, let us state — without a proof — the following theorem.

Theorem 7.1. *If $f(x)$ and $g(x)$ are bounded continuous and absolute-value integrable for all x, then*

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\xi) g(x-\xi) \, \mathrm{d}\xi \mathrm{d}x = \int_{-\infty}^{\infty} f(x) \, \mathrm{d}x \int_{-\infty}^{\infty} g(x) \, \mathrm{d}x.$$

Using this theorem, we can integrate the right-hand side of the above equation with respect to x to get

$$\int_{-\infty}^{\infty} \exp(-itx) \int_{-\infty}^{\infty} u(\xi) v(x-\xi) \, \mathrm{d}\xi \mathrm{d}x == \int_{-\infty}^{\infty} u(x) \exp(-itx) \, \mathrm{d}x \int_{-\infty}^{\infty} v(x) \exp(-itx) \, \mathrm{d}x.$$

We see that the functions resulting from these integrations are functions of t.

Examining the last equation, we recognize that the left-hand side is the Fourier transform of the convolution, while the right-hand side is the product of the Fourier transforms. In other, words, we can write

$$\widehat{(u * v)}(t) = \widehat{(u)}(t)\,\widehat{(v)}(t).$$

In this derivation, we have not imposed any particular conditions on u or v. Hence, this result is valid for any Fourier-transformable function. We can concisely write

$$\widehat{f * g} = \hat{f}\hat{g},$$

which states an important property of the Fourier transform that relates convolution and multiplication.

D Elastodynamic equations

In this appendix, we derive equations that describe the motion of displacement within a linearly elastic continuum. These equations are called the elastodynamic equations. The elastodynamic equations result from combining Cauchy's equations of motion with the constitutive equations of linear elasticity. Cauchy's equations of motion are rooted in the balance law of linear momentum, which states that the rate of change of momentum of a portion of continuum is equal to the sum of all forces acting upon it. The origin of the constitutive equations is in the empirical studies, according to which we can approximate the relation between forces and deformations in a linear fashion. These constitutive equations are the stress-strain equations of Hooke's law.

D.1 Cauchy's equations of motion

In this section, we present a derivation of Cauchy's equations of motion, which is rooted in the balance of momentum.

To study the momentum within a continuum, we have to state what we mean by velocity therein. This concept is more obvious in particle mechanics, but less so in continuum mechanics. We find it convenient to label a portion a continuum by coordinate X. This coordinate is attached to a particular element of the continuum. Although there are no particles in continuum mechanics, we can picture X as a label attached to a given particle. Thus, we can express position x of given element X at a given time t as

$$x = x(X, t),$$

where x is a coordinate system fixed in space. Equivalently, we can express element X at position x at time t as

$$X = X(x, t).$$

Functions $x(X, t)$ and $X(x, t)$ are the inverses of one another, namely $x(X(x, t), t) = x$ and $X(x(X, t), t) = X$.

Knowing how to express the position, we can discuss the velocity. Velocity of portion of continuum X is

$$\tilde{v}(X, t) := \frac{\mathrm{d}}{\mathrm{d}t} x(X, t).$$

If we want to talk about a velocity function in the spatial representation, we must say what we mean by this. We define the velocity in spatial coordinates as

$$v\left(x,t\right)=\tilde{v}\left(X\left(x,t\right),t\right).$$

This relation is true for any quantity expressed in the spatial and material coordinates; if a is a quantity in the spatial coordinates and $\tilde{a}$ is the same quantity in the material coordinates, they are related by

$$a\left(x,t\right)=\tilde{a}\left(X\left(x,t\right),t\right).$$

After defining the velocity, we are ready to express the balance of momentum: The temporal rate of change of momentum of a portion of a continuum is equal to the sum of forces acting on and within this part of continuum. Mathematically, we can express this statement as follows.

$$\frac{\mathrm{d}}{\mathrm{d}t}\iiint\limits_{V(t)}\rho\left(x,t\right)v\left(x,t\right)\mathrm{d}V\left(x\right)=\iint\limits_{S(t)}T\left(x,t\right)\mathrm{d}S\left(x\right)+\iiint\limits_{V(t)}F\left(x,t\right)\rho\left(x,t\right)\mathrm{d}V\left(x\right). \tag{D.1}$$

The volume integral on the left-hand side is the momentum of the portion of the continuum contained in volume $V\left(t\right)$, which we choose to be such that it contains always the same portion of continuum. If we express this volume in the material coordinates, $\tilde{V}$, we see that this is constant since it is expressed in coordinates X, which are associated with the portion of the continuum; coordinates X change in time together with the deformation of the volume. This fact will be useful in the computation of time derivative of an integral over this volume. The two integrals on the right-hand side describe all the forces that we consider as acting on and within the portion of continuum. T is the traction accounting for the surface forces, with S being the surface enclosing volume V, and F is a body force accounting for forces such as gravity. Thus, the left-hand side of the above equation is the rate of change of momentum, while the right-hand side is the sum of all forces.

Let us focus our attention on the left-hand side of equation (D.1). Since we chose V in such a way that it always contains the same portion of continuum, it follows that the mass in this volume does not change with time. Hence, we rewrite the left-hand side of equation (D.1) as

$$\frac{\mathrm{d}}{\mathrm{d}t}\iiint\limits_{V(t)}\rho\left(x,t\right)v\left(x,t\right)\mathrm{d}V\left(x\right)=\frac{\mathrm{d}}{\mathrm{d}t}\iiint\limits_{\tilde{V}}\tilde{\rho}\left(X,t\right)\tilde{v}\left(X,t\right)\mathrm{d}\tilde{V}\left(X,t\right)=\frac{\mathrm{d}}{\mathrm{d}t}\iiint\limits_{\tilde{V}}\tilde{v}\left(X,t\right)\mathrm{d}m\left(X\right),$$

where $\mathrm{d}m\left(X\right)=\tilde{\rho}\left(X,t\right)\mathrm{d}\tilde{V}\left(X,t\right)$ denotes the element of mass that does not depend on time. Since the domain of integration does not depend on time, we can interchange the differentiation and integration to write

$$\frac{\mathrm{d}}{\mathrm{d}t}\iiint\limits_{\tilde{V}}\tilde{\rho}\left(X,t\right)\tilde{v}\left(X,t\right)\mathrm{d}\tilde{V}\left(X,t\right)=\iiint\limits_{\tilde{V}}\frac{\mathrm{d}}{\mathrm{d}t}\tilde{v}\left(X,t\right)\mathrm{d}m\left(X\right).$$

This way we expressed the right hand side of equation (D.1) in the material coordinates. Since it is more convenient to work with the right-hand side of equation (D.1) in the spatial coordinates x that are fixed in space, rather than attached to the continuum, we re-express this integral in the spatial coordinates.

$$\begin{aligned}
\iiint_{\tilde{V}} \frac{\mathrm{d}}{\mathrm{d}t}\tilde{v}\left(X,t\right)\mathrm{d}m\left(X\right) &= \iiint_{\tilde{V}} \left\{\frac{\mathrm{d}}{\mathrm{d}t}\tilde{v}\left(X,t\right)\right\}\tilde{\rho}\left(X,t\right)\mathrm{d}\tilde{V}\left(X,t\right) \\
&= \iiint_{\tilde{V}} \left\{\frac{\partial}{\partial t}v\left(x\left(X,t\right),t\right)+\frac{\partial x}{\partial t}\cdot\frac{\partial}{\partial x}v\left(x\left(X,t\right),t\right)\right\}\rho\left(x\left(X,t\right),t\right)\mathrm{d}\tilde{V}\left(x\left(X,t\right),t\right) \\
&= \iiint_{\tilde{V}} \left\{\frac{\partial}{\partial t}v\left(x\left(X,t\right),t\right)+v\left(x\left(X,t\right),t\right)\cdot\frac{\partial}{\partial x}v\left(x\left(X,t\right),t\right)\right\}\rho\left(x\left(X,t\right),t\right)\mathrm{d}\tilde{V}\left(x\left(X,t\right),t\right) \\
&= \iiint_{V(t)} \left\{\frac{\partial}{\partial t}v\left(x,t\right)+v\left(x,t\right)\cdot\frac{\partial}{\partial x}v\left(x,t\right)\right\}\rho\left(x,t\right)\mathrm{d}V\left(x\right).
\end{aligned}$$

Using this result in equation (D.1), we can write

$$\iiint_{V(t)} \left\{\frac{\partial v\left(x,t\right)}{\partial t}+v\left(x,t\right)\cdot\frac{\partial}{\partial x}v\left(x,t\right)\right\}\rho\left(x,t\right)\mathrm{d}V = \iint_{S(t)} T\left(x,t\right)\mathrm{d}S+\iiint_{V(t)} F\left(x,t\right)\rho\left(x,t\right)\mathrm{d}V. \quad \text{(D.2)}$$

To combine integrals in equation (D.2), we wish to express all of them as volume integrals. At this point, we introduce an important equation of elasticity theory. We write the components of traction as

$$T_i = \sum_{j=1}^{3}\sigma_{ij}N_j, \qquad i \in \{1,2,3\}. \quad \text{(D.3)}$$

Herein, σ_{ij} are the components of the so-called stress tensor and N_j are the components of a unit vector normal to the element of surface within the continuum to which traction is applied. This expression can be viewed as the definition of the stress tensor, which allows us to study forces acting on an arbitrarily oriented plane in the continuum. The justification for writing traction in this way is rooted in balance of momentum (e.g. Slawinski (2003, pp. 40 – 46)).

Inserting expression (D.3) into equation (D.2) and invoking the divergence theorem stated on page 133, we can write equation (D.2) in terms of components as

$$\iiint_{V(t)} \left\{\frac{\partial v_i\left(x,t\right)}{\partial t}+\sum_{j=1}^{3}v_j\left(x,t\right)\frac{\partial}{\partial x_j}v_i\left(x,t\right)\right\}\rho\left(x,t\right)\mathrm{d}V = \iiint_{V(t)}\sum_{j=1}^{3}\frac{\partial\sigma_{ij}}{\partial x_j}\mathrm{d}V+\iiint_{V(t)} F_i\left(x\right)\rho\left(x,t\right)\mathrm{d}V,$$

where $\mathrm{d}V = \mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3$ and $i \in \{1,2,3\}$. Now, we combine the three volume integrals into a single integral to obtain

$$\iiint_{V(t)} \left(\sum_{j=1}^{3}\frac{\partial\sigma_{ij}}{\partial x_j}+\left(F_i-\frac{\partial v_i}{\partial t}-\sum_{k=1}^{3}v_k\frac{\partial v_i}{\partial x_k}\right)\rho\right)\mathrm{d}V = 0, \qquad i \in \{1,2,3\}. \quad \text{(D.4)}$$

To write the differential equations that govern the motion within the continuum, we use the fact that for equations (D.4) to be satisfied for an arbitrary volume, the integrands must be identically zero. Thus, we require

$$\sum_{j=1}^{3} \frac{\partial \sigma_{ij}}{\partial x_j} + \left(F_i - \frac{\partial v_i}{\partial t} - \sum_{k=1}^{3} v_k \frac{\partial v_i}{\partial x_k} \right) \rho = 0, \qquad i \in \{1,2,3\}.$$

Cauchy's equations of motion result from these differential equations upon using several approximations. In elasticity theory, we can often ignore the effect of body forces; in particular, the effect of elasticity on displacement is much greater than the effect of gravity. Moreover, we neglect the product of the velocity with the gradient of the velocity. Hence, we write

$$\sum_{j=1}^{3} \frac{\partial \sigma_{ij}}{\partial x_j} - \rho \frac{\partial v}{\partial t} = 0, \qquad i \in \{1,2,3\},$$

which are Cauchy's equations of motion.

To study Cauchy's equations of motion together with constitutive equations, which are discussed in the next section, we express velocity as the derivative of displacement. The displacement of portion of a continuum is the difference between its present and its original positions. We can express this difference as

$$\tilde{u}(X,t) = x(X,t) - x(X,0).$$

Taking the time derivative, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\tilde{u}(X,t) = \frac{\mathrm{d}}{\mathrm{d}t}x(X,t) = \tilde{v}(X,t),$$

where we used the fact that $x(X,0)$ does not depend on time. We would like to express this in the spatial coordinates so we can use it in Cauchy's equations of motion. To do so, we take a look at the displacement $u(x,t) := \tilde{u}(X(x,t),t)$ and its second derivative.

$$\begin{aligned} \frac{\partial^2 u}{\partial t^2}(x,t) &= \frac{\partial^2}{\partial t^2}\tilde{u}(X(x,t),t) = \frac{\partial}{\partial t}\left(\tilde{v} + \frac{\partial \tilde{u}}{\partial X}\frac{\partial X}{\partial t} \right)_{X=X(x,t)} \\ &= \frac{\partial \tilde{v}}{\partial t} + \frac{\partial \tilde{v}}{\partial X}\frac{\partial X}{\partial t} + \frac{\partial^2 \tilde{u}}{\partial X \partial t}\frac{\partial X}{\partial t} + \frac{\partial^2 \tilde{u}}{\partial X^2}\left(\frac{\partial X}{\partial t}\right)^2 + \frac{\partial \tilde{u}}{\partial X}\frac{\partial^2 X}{\partial t^2}. \end{aligned} \tag{D.5}$$

We observe that the firs two terms on the right-hand side are equal to $\partial v/\partial t$. Hence, after neglecting the last three terms in expression (D.5), we can write that

$$\frac{\partial v}{\partial t} = \frac{\partial^2 u}{\partial t^2},$$

and thus we can rewrite Cauchy's equations as

$$\sum_{j=1}^{3} \frac{\partial \sigma_{ij}}{\partial x_j} = \rho(x) \frac{\partial^2 u_i(x,t)}{\partial t^2}, \qquad i \in \{1,2,3\}. \tag{D.6}$$

D.2 Stress-strain equations

In this section, we formulate a mathematical description of a particular continuum, namely, the constitutive equation of a linearly elastic continuum. Constitutive equations relate forces and deformations. Herein, we wish to relate the stress tensor, σ_{ij}, to the displacement vector, u.

In our study of the elasticity theory, which focuses on wave propagation, we can limit our interests to small deformations. In other words, it is reasonable to assume that the displacements within the continuum due to wave propagation are small. In view of small displacements, we choose to describe the deformation by a tensor that is defined by

$$\varepsilon_{kl} = \frac{1}{2}\left(\frac{\partial u_k}{\partial x_l} + \frac{\partial u_l}{\partial x_k}\right), \qquad k, l \in \{1, 2, 3\}, \tag{D.7}$$

e.g., Slawinski (2003, pp. 15 – 19). We refer to it as the strain tensor; it is also called the infinitesimal strain tensor. To incorporate the study of deformations in Cauchy's equations of motion, we wish to relate to one another the two second-rank tensors, σ_{ij} and ε_{kl}, in a way that is physically justified and mathematically convenient. For elastic materials, as observed by Robert Hooke in the seventeenth century, the relation between the forces and deformations is well approximated by a linear relation. This approximation is accurate as long as the deformations are small. Thus, in a mathematical language, we can write

$$\sigma_{ij} = \sum_{k=1}^{3}\sum_{l=1}^{3} c_{ijkl}(x)\,\varepsilon_{kl}, \qquad i, j \in \{1, 2, 3\}, \tag{D.8}$$

where c_{ijkl} are the fourth-rank tensor.

Since, by its definition, the strain tensor is symmetric, namely, $\varepsilon_{kl} = \varepsilon_{lk}$, it follows that $c_{ijkl} = c_{ijlk}$.

In the context of elasticity theory, we refer to these equations as stress-strain equations and to the components of c_{ijkl} as the elasticity parameters. These are the constitutive equations of a linearly elastic anisotropic inhomogeneous continuum.

D.3 Elastodynamic equations

In this section, we will combine Cauchy's equations of motion (D.6) with stress-strain equations (D.8). Thus, we will obtain the equations that describe propagation of deformation in an elastic continuum.

Inserting the expression for the stress-tensor given in equation (D.8) into Cauchy's equations of motion (D.6), we write

$$\rho(x)\frac{\partial^2 u_i(x,t)}{\partial t^2} = \sum_{j=1}^{3}\frac{\partial}{\partial x_j}\sum_{k=1}^{3}\sum_{l=1}^{3} c_{ijkl}(x)\,\varepsilon_{kl}, \qquad i \in \{1, 2, 3\}.$$

Expressing the strain tensor in terms of the displacement vector as shown in expression (D.7), we get

$$\rho(x)\frac{\partial^2 u_i(x,t)}{\partial t^2} = \frac{1}{2}\sum_{j=1}^{3}\frac{\partial}{\partial x_j}\sum_{k=1}^{3}\sum_{l=1}^{3} c_{ijkl}(x)\left(\frac{\partial u_k}{\partial x_l} + \frac{\partial u_l}{\partial x_k}\right), \qquad i \in \{1, 2, 3\}.$$

Using the linearity of the differential operator and following the product rule, obtain the desired elastodynamic equations. They are

$$\rho(x)\frac{\partial^2 u_i(x,t)}{\partial t^2} = \frac{1}{2}\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3}\frac{\partial c_{ijkl}(x)}{\partial x_j}\left(\frac{\partial u_k(x,t)}{\partial x_l}+\frac{\partial u_l(x,t)}{\partial x_k}\right) + \frac{1}{2}\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3}c_{ijkl}(x)\left(\frac{\partial^2 u_k(x,t)}{\partial x_j \partial x_l}+\frac{\partial^2 u_l(x,t)}{\partial x_j \partial x_k}\right),$$

where $i \in \{1,2,3\}$. Using the fact that we can rename the summation indices, we can write these equations as

$$\rho(x)\frac{\partial^2 u_i(x,t)}{\partial t^2} = \sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3}\frac{1}{2}\left(\frac{\partial c_{ijkl}(x)}{\partial x_j}+\frac{\partial c_{ijlk}(x)}{\partial x_j}\right)\frac{\partial u_k(x,t)}{\partial x_l} + \sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3}\frac{1}{2}\left(c_{ijkl}(x)+c_{ijlk}(x)\right)\frac{\partial^2 u_k(x,t)}{\partial x_j \partial x_l}.$$

Furthermore, using the symmetry of the elasticity tensor, we can write these equations as

$$\rho(x)\frac{\partial^2 u_i(x,t)}{\partial t^2} = \sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3}\frac{\partial c_{ijkl}(x)}{\partial x_j}\frac{\partial u_k(x,t)}{\partial x_l} + \sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3}c_{ijkl}(x)\frac{\partial^2 u_k(x,t)}{\partial x_j \partial x_l}, \tag{D.9}$$

where $i \in \{1,2,3\}$.

These are linear second-order partial differential equations that describe propagation of displacement in an anisotropic inhomogeneous linearly elastic continuum. In Section 2.4.1, we study their characteristics hypersurfaces.

E Scalar and vector potentials in elastodynamics

An important aspect of study of wave propagation consists of examining the elastodynamic equations in the particular case of isotropy and homogeneity. In such a case, we can formulate the concept of wave propagation in terms of scalar and vector potentials. Such a formulation is the purpose of this appendix. We will then use the resulting system of differential equations to study their characteristic hypersurfaces and their meaning in elasticity theory in Section 2.4.1.

In the context of elasticity theory, in this appendix we will discuss a special case of the subject formulated in Appendix D. Consequently, rather than starting with the conservation of linear momentum, we will begin with Cauchy's equations of motion that we derived in Appendix D. In the context of mathematical physics, the subject discussed herein possesses several key similarities with our discussions in Appendix F, below, that deals with the Maxwell equations, which contain the theory of electrodynamics. These similarities exemplify the unifying aspects of mathematical formulations in physical sciences.

E.1 Equations of motion

To formulate the elastodynamic equations for an isotropic homogeneous continuum, consider Cauchy's equations of motion given in expression (D.6), namely,

$$\sum_{j=1}^{3} \frac{\partial \sigma_{ij}}{\partial x_j} = \rho \frac{\partial^2 u_i}{\partial t^2}, \qquad i \in \{1,2,3\}, \tag{E.1}$$

where $u(x,t)$ is the displacement vector and ρ is the mass density, which herein is a constant. In an elastic isotropic homogeneous medium, the stress tensor, σ_{ij}, can be expressed in terms of the strain tensor, ε_{ij}, by invoking equations (D.8), namely,

$$\sigma_{ij} = \sum_{k=1}^{3}\sum_{l=1}^{3} c_{ijkl}(x)\,\varepsilon_{kl}, \qquad i,j \in \{1,2,3\}, \tag{E.2}$$

which, in general, allows us to study anisotropic inhomogeneous continua. For a particular case of isotropy and homogeneity that we study herein, we will use the isotropic form of the elasticity tensor, c_{ijkl}, with its components not depending on position, x.

An isotropic tensor implies that for this tensor all systems of Cartesian coordinates are equivalent to each other. Herein, c_{ijkl} must have the same components for all such coordinate systems. In other words,

$$c_{ijkl} = \sum_{r=1}^{3}\sum_{s=1}^{3}\sum_{t=1}^{3}\sum_{w=1}^{3} A_{ir}A_{js}A_{kt}A_{lw}c_{rstw} = \tilde{c}_{ijkl},$$

where c and $\tilde{c}$ denote the same elasticity tensor in two Cartesian coordinate systems and A is the orthogonal transformation relating these coordinate systems. A second-rank tensor that transforms identically can be expressed by the Kronecker delta. It can be shown that the general form of an isotropic fourth-rank tensor is composed of Kronecker deltas to give

$$a_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \xi\delta_{ik}\delta_{jl} + \eta\delta_{il}\delta_{jk}, \qquad i,j,k,l \in \{1,2,3\}, \tag{E.3}$$

where λ, ξ and η are constants.[2] In other words, an isotropic fourth-rank tensor is stated in terms of three constants that do not depend on the choice of the coordinate system.

In elasticity theory, since the strain tensor is symmetric, namely, $\varepsilon_{kl} = \varepsilon_{lk}$, the isotropic elasticity tensor is

$$c_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \mu\delta_{ik}\delta_{jl}, \qquad i,j,k,l \in \{1,2,3\}\,, \tag{E.4}$$

where $\mu := (\xi + \eta)/2$. In the context of elasticity, λ and μ are called the Lamé parameters.

Inserting expressions (E.4) into equations (E.2), we obtain stress-strain equations for an isotropic homogeneous continuum; they are

$$\sigma_{ij} = \lambda\delta_{ij}\sum_{k=1}^{3}\varepsilon_{kk} + 2\mu\varepsilon_{ij}, \qquad i,j \in \{1,2,3\}\,. \tag{E.5}$$

Expressing the strain tensor in terms of the displacement vector using expression (D.7) and, then, inserting expression (E.5) into equation (E.1), we get the equations of motion in an elastic isotropic homogeneous medium, namely,

$$\rho\frac{\partial^2 u}{\partial t^2} = (\lambda+\mu)\,\nabla\,(\nabla\cdot u) + \mu\nabla^2 u. \tag{E.6}$$

E.2 Scalar and vector potentials

Now, we discuss the key step of our formulation. Following the Helmholtz theorem[3], which allows us to separate vector function u into its scalar potential, ϕ, and vector potential, A, we write the displacement as

$$u\,(x,t) = \nabla\phi + \nabla\times A. \tag{E.7}$$

In terms of u_1, u_2 and u_3, expression (E.7) constitutes three equations for four unknowns, namely, A_1, A_2, A_3 and φ.

Now, we will formulate another equation – which we will invoke in Section E.3 in the context of equations (E.7) – by using the fact that the divergence of vector A is arbitrary up to the gradient of function f. This means that we can change A in equation (E.7) by adding ∇f to it. If we choose f in such a way that $\nabla^2 f = -\nabla\cdot A$, we obtain a new vector potential, $\tilde{A}$, whose divergence vanishes. Namely,

$$\nabla\cdot\tilde{A} = \nabla\cdot(A+\nabla f) = \nabla\cdot A + \nabla^2 f = 0.$$

Since A is arbitrary, we conclude that the equation we seek is

$$\nabla\cdot A = 0. \tag{E.8}$$

Note that equation (E.8) is analogous to equation (F.18) with assumption that $\partial\phi/\partial t = 0$. This choice of the potential in electrodynamics is called the Coulomb calibration.

---

[2] Readers interested in the formulation of this form my refer to Synge, J.L., and Schild, A., (1949/1978) Tensor calculus: Dover, pp. 210 – 211.

[3] Readers interested in the Helmholtz theorem might refer to Arfken, G.B, and Weber, H.J., (2001) *Mathematical methods for physicists* ($5th$ edition): Harcourt/Academic Press, pp. 96 – 101.

Inserting expression (E.7) into equation (E.6) and rearranging, we get

$$\nabla\left[(\lambda+2\mu)\,\nabla^2\phi-\rho\frac{\partial^2\phi}{\partial t^2}\right]+\nabla\times\left[\mu\nabla^2 A-\rho\frac{\partial^2 A}{\partial t^2}\right]=0. \tag{E.9}$$

Examining equation (E.9), we see that, if we take the divergence of this equation, the second term disappears, while if we take the curl, the first term disappears. We will proceed in this way to obtain the two wave equations that correspond to P waves and S waves.

E.3 Wave equations

P waves

Taking the divergence of equation (E.9) and using the fact that the divergence of a gradient is the Laplace operator, ∇^2, while divergence of a curl disappears, we get

$$\nabla^2\left[(\lambda+2\mu)\,\nabla^2\phi-\rho\frac{\partial^2\phi}{\partial t^2}\right]=0.$$

Using the linearity of the differential operators and the fact that λ and μ are constants, we write as

$$\left[(\lambda+2\mu)\,\nabla^2\left(\nabla^2\phi\right)-\rho\frac{\partial^2}{\partial t^2}\left(\nabla^2\phi\right)\right]=0. \tag{E.10}$$

To examine the meaning of $\nabla^2\phi$ in terms of the displacement vector, let us take the divergence of equation (E.7). Again, using the fact that the divergence of a gradient is the Laplace operator, ∇^2, while divergence of a curl disappears, we obtain

$$\nabla^2\phi=\nabla\cdot u. \tag{E.11}$$

It is common to denote $\nabla\cdot u$ by φ, and to refer to it as dilatation. This name results from the fact that φ is a scalar quantity that we can write

$$\varphi:=\nabla\cdot u=\frac{\partial u_1}{\partial x_1}+\frac{\partial u_2}{\partial x_2}+\frac{\partial u_3}{\partial x_3},$$

which allows us to view φ as corresponding to the volume change of an infinitesimal cube. Using expression (E.11), we rewrite equation (E.10) as

$$\nabla^2\varphi-\frac{1}{\frac{\lambda+2\mu}{\rho}}\frac{\partial^2\varphi}{\partial t^2}=0. \tag{E.12}$$

This equation is a linear second-order partial differential equation for φ; it is a wave equation. Herein, this is the wave equation for dilatational waves, which are commonly referred to as the P waves.

S waves

Now, taking the curl of equation (E.9) and using the fact that the curl of the gradient disappears, we get

$$\nabla \times \nabla \times \left[\mu \nabla^2 A - \rho \frac{\partial^2 A}{\partial t^2}\right] = 0. \tag{E.13}$$

Let us invoke a vector-calculus identity to rewrite the left-hand side of equation (E.13) to obtain

$$\nabla \times \nabla \times \left[\mu \nabla^2 A - \rho \frac{\partial^2 A}{\partial t^2}\right] = \nabla \left(\nabla \cdot \left[\mu \nabla^2 A - \rho \frac{\partial^2 A}{\partial t^2}\right]\right) - \nabla^2 \left[\mu \nabla^2 A - \rho \frac{\partial^2 A}{\partial t^2}\right] = 0,$$

which we can rewrite as

$$\nabla \times \nabla \times \left[\mu \nabla^2 A - \rho \frac{\partial^2 A}{\partial t^2}\right] = \nabla \left(\nabla \cdot \left[\mu \nabla^2 A - \rho \frac{\partial^2 A}{\partial t^2}\right]\right) = \nabla^2 \left[\mu \nabla^2 A - \rho \frac{\partial^2 A}{\partial t^2}\right] = 0.$$

Considering the left-hand side of the above equation, we choose to write

$$\mu \nabla^2 \left(\nabla^2 A\right) - \rho \frac{\partial^2}{\partial t^2} \left(\nabla^2 A\right) = 0, \tag{E.14}$$

where we used the linearity of the differential operators and the fact that μ is constant. To examine the meaning of $\nabla^2 A$ in terms of the displacement vector, let us take the curl of equation (E.7). Again, using the fact that the curl of a gradient disappears, we obtain

$$\nabla \times u = \nabla \times \nabla \times A.$$

Following the vector-calculus identity, we write

$$\nabla \times u = \nabla \times \nabla \times A = \nabla \left(\nabla \cdot A\right) - \nabla^2 A.$$

In view of equation (E.8), we simplify the right-hand side to get

$$\nabla \times u = -\nabla^2 A.$$

Let us denote $\mathcal{A} := \nabla \times u$. Hence, we rewrite equation (E.14) as

$$\nabla^2 \mathcal{A} - \frac{1}{\frac{\mu}{\rho}} \frac{\partial^2 \mathcal{A}}{\partial t^2} = 0. \tag{E.15}$$

This equation is a linear second-order partial differential equation for $\mathcal{A}$; it is a wave equation. Herein, this is the wave equation for rotational waves, which are commonly referred to as the S waves. Since A is a vectorial quantity, equation (E.15) contains three equations to be solved for $\mathcal{A}_1$, $\mathcal{A}_2$ and $\mathcal{A}_3$.

E.4 Equations of motion versus wave equations

Thus, the three equations of motion given in expression (E.6) are equivalent to the four wave equations given by expressions (E.12) and (E.15). Within the wave equations, we notice the separation of physical quantities. Equation (E.12) deals with the scalar potential, which is associated with P waves, while equations (E.15) deal with the vector potential, which is associated with S waves.

If we solve the system given by expressions (E.12) and (E.15) for $\mathcal{A} = \nabla \times u$ and $\varphi = \nabla \cdot u$, we find u. Thus, at the end, we obtain $u\,(x,t)$, which is tantamount to solving equation (E.6).?.

Expressions (E.12) and (E.15) constitute a system of four linear partial differential equations for four unknowns, namely, $\mathcal{A}_1$, $\mathcal{A}_2$, $\mathcal{A}_3$ and φ. Herein, we have formulated this system to study its characteristics hypersurfaces in Section 2.4.1.

F Maxwell equations

F.1 Formulation

Fundamental equations

The electromagnetic field manifests its presence by the electric-field intensity, E, and by the magnetic induction, B. We measure the effect of this field by force F that is exerted on a particle bearing charge q and moving with velocity v. This force is

$$F = q\left(E + v \times B\right).$$

Equation of continuity for electric charges states that

$$\nabla \cdot J = -\frac{\partial \rho}{\partial t}, \tag{F.1}$$

with ρ denoting the electric-charge density, which is the amount of charge per unit volume. This equation states the conservation of charge and can be viewed as a definition of J, which is the electric current density: the rate at which charge flows through a unit area per second.

Let us consider the following equations involving E and B as well as ρ and J.

Coulomb's law

First, the flux of E through any closed surface, S, is proportional to the charge that is enclosed in this surface. This is Coulomb's law. Mathematically, we can write it as the integral equation given by

$$\iint_S E \cdot N \mathrm{d}S = \iiint_V \frac{\rho}{\epsilon_0} \mathrm{d}V,$$

where N is the unit vector normal to S and pointing away from volume V. Parameter ϵ_0 is a constant whose value can be determined experimentally by measuring the force between two unit charges. This value depends on our definition of unit charge.

Returning to the above integral equation and invoking the divergence theorem stated on page 133 as well as using the linearity of the integral operator, we get

$$\iiint_V \left(\nabla \cdot E - \frac{\rho}{\epsilon_0}\right) \mathrm{d}V = 0.$$

For the integral to vanish for an arbitrary volume, we require that the integrand vanishes. Thus, we obtain a differential equation given by

$$\nabla \cdot E = \frac{\rho}{\epsilon_0}. \tag{F.2}$$

In other words, the divergence of E is the charge density scaled by parameter ϵ_0.

No-monopole law

Having obtained equation (F.2), we search for an analogous equation for B. Since there is no experimental indication of magnetic monopoles, which for B would be analogous to the isolated electric charge for E, the net flux of magnetic induction through S is zero. We can refer to this statement as no-monopole law. Mathematically, we can write it as the integral equation given by

$$\iint_S B \cdot N \mathrm{d}S = 0; \tag{F.3}$$

in other words, flux of B through any closed surface is zero. To write equation (F.3) as a differential equation, we invoke the divergence theorem, stated on page 133, to get

$$\iiint_V \nabla \cdot B \, \mathrm{d}V = 0.$$

For the integral to vanish for an arbitrary volume, we require that

$$\nabla \cdot B = 0, \tag{F.4}$$

which is a differential equation that is tantamount to equation (F.3) stating that there are no magnetic monopoles.

Note that we could not use the argument of vanishing integrand in equation (F.3), since the surface is not arbitrary; it is closed.

Faraday's law

Now, we will formulate the law of the electromagnetic induction, which is called Faraday's law. This law states that the voltage around a loop, C, is equal to the negative of the time rate of change of the magnetic flux through this loop. Mathematically, we can write this law as the integral equation given by

$$\int_C E \cdot \mathrm{d}l = -\frac{\partial}{\partial t} \iint_S B \cdot N \mathrm{d}S, \tag{F.5}$$

where S is the surface bounded by curve C, $\mathrm{d}l$ is a vector element along this curve and N is the unit vector normal to S.

To invoke the curl theorem stated on page 137, which will allow us to combine both integrals in the above equation, we rewrite this equation as

$$\int_C E \cdot n \, \mathrm{d}s = -\frac{\partial}{\partial t} \iint_S B \cdot N \mathrm{d}S,$$

with n being the unit vector tangent to C and $\mathrm{d}s$ being the element of length along C. Following the curl theorem, we write

$$\int_C E \cdot n \, \mathrm{d}s = \iint_S [(\nabla \times E) \cdot N] \, \mathrm{d}S = -\frac{\partial}{\partial t} \iint_S B \cdot N \, \mathrm{d}S.$$

Considering the two surface integrals and using the linearity of the integral operator, we get

$$\iint_S \left[\nabla \times E + \frac{\partial B}{\partial t}\right] \cdot N \mathrm{d}S = 0.$$

For the integral to vanish for an arbitrary surface, we require the integrand to vanish. Since $N \neq 0$, we conclude that

$$\nabla \times E = -\frac{\partial B}{\partial t}, \tag{F.6}$$

which is tantamount to equation (F.5) stated herein as a differential equation.

Ampère's law

Having obtained equation (F.6), we search again for an analogous equation for B. This equation results from Ampère's law whose original form as the integral equation is

$$\int_C B \cdot \mathrm{d}l = \frac{1}{c^2 \epsilon_0} \iint_S J \cdot N \mathrm{d}S, \tag{F.7}$$

where $c^2\epsilon_0$ is a constant whose value can be determined experimentally by measuring the force between two unit currents, which are defined as a unit charge per second. Invoking the curl theorem, stated on page 137, and using the linearity of the integral operator as well as requiring the vanishing of the integral for an arbitrary surface, we can rewrite this integral equation as

$$\nabla \times B = \frac{J}{c^2 \epsilon_0}. \tag{F.8}$$

This equation, however, is not consistent with equation (F.1), which is a fundamental equation expressing the conservation of charge. We can see the contradiction between equations (F.1) and (F.8) by taking the divergence of the latter equation to get $\nabla \cdot J = 0$, which states that the total flux of current out of any closed surface is equal to zero, rather than to the rate of change of the charge inside this surface, as required by the conservation of charge. Maxwell solved this problem by adding a term to equation (F.8) to obtain

$$c^2 \nabla \times B = \frac{J}{\epsilon_0} + \frac{\partial E}{\partial t}. \tag{F.9}$$

To verify the consistency, we take the divergence of equation (F.9) and use equation (F.2) to get

$$\nabla \cdot J = -\epsilon_0 \frac{\partial}{\partial t} \nabla \cdot E = -\frac{\partial \rho}{\partial t},$$

which is equation (F.1), as required.

Examining equation (F.9) in the context of its derivation from equation (F.7), we see that the product of c^2 and the circulation of B around loop C is equal to the sum of the flux of the electric current through surface S and the temporal rate of change of the flux of E through this surface.

Speed of light

Examining constant ϵ_0 used in Coulomb's law and constant $c^2\epsilon_0$ used in Ampère's law, we realize that – although their values depend on our definition of the unit charge – their ratio is the same, namely, c^2. As originally remarked by Maxwell and confirmed by subsequent studies, c is the speed of light. Furthermore, the relation between the magnitudes of E and B is $\|E\| = c\,\|B\|$.

Maxwell equations

Equations (F.2), (F.4), (F.6) and (F.9) are the Maxwell equations. Formally, they contain the entire theory of electrodynamics, where concepts of electricity and magnetism are intimately linked. They also show that light belongs to this theory. To appreciate the conciseness of Maxwell's formulations, let us write all four equations; they are

$$\nabla \cdot E = \frac{\rho}{\epsilon_0}, \tag{F.10}$$

$$\nabla \cdot B = 0, \tag{F.11}$$

$$\nabla \times E = -\frac{\partial B}{\partial t} \tag{F.12}$$

and

$$c^2 \nabla \times B = \frac{J}{\epsilon_0} + \frac{\partial E}{\partial t}. \tag{F.13}$$

We note that if we consider a static case – that is the case where E and B do not depend on time and, hence, $\partial B/\partial t = \partial E/\partial t = 0$ – equations (F.10) and (F.12) deal only with electricity, while (F.11) and (F.13) deal only with magnetism; in other words, the Maxwell equations can be separated into the electrostatic and magnetostatic equations. In particular, from this argument we see that equation (F.8) is the magnetostatic analogue of equation (F.13), which incorporates time dependence. In general, from this argument we conclude that it is the time dependence that combines electrostatics and magnetostatics into a single coherent theory, namely, the theory of electrodynamics, which we can also refer to as the theory of electromagnetism, thus emphasizing the link between the electricity and magnetism.

F.2 Scalar and vector potentials

We wish to write the four Maxwell equations in a way that allows us to study them in the context of the characteristic hypersurfaces.

Consider equation (F.11). In the context of the properties of vector operators, there is a theorem stating that if the divergence of a vector field is zero, it follows that this field is a curl of a vector field. Thus, equation (F.11) implies that

$$B = \nabla \times A, \tag{F.14}$$

where A is a vector field. We can view this expression as a solution of equation (F.11), and we refer to A as a vector potential. A is not unique; in view of identity $\nabla \times \nabla\xi = 0$, we can rewrite it as $B = \nabla \times \breve{A}$, where $\breve{A} = A + \nabla\xi$ with ξ being an arbitrary scalar function.

Using expression (F.14), we can write equation (F.12) as

$$\nabla \times E = -\frac{\partial}{\partial t} \nabla \times A.$$

Exchanging the order of spatial and temporal derivatives, we rewrite it as

$$\nabla \times E = -\nabla \times \frac{\partial A}{\partial t},$$

to get

$$\nabla \times \left(E + \frac{\partial A}{\partial t} \right) = 0.$$

In the context of the properties of vector operators, there is a theorem stating that if the curl of a vector field is zero, it follows that this field is a gradient of a scalar function. Thus, we can write the term in parentheses as

$$E + \frac{\partial A}{\partial t} = -\nabla \phi;$$

we refer to ϕ as a scalar potential. Consequently, we can write a solution of equation (F.12) as

$$E = -\nabla \phi - \frac{\partial A}{\partial t}. \tag{F.15}$$

Examining equations (F.14) and (F.15), we see that vector potential A appears in the expressions for both B and E. Since A is not unique, we must constrain A in such a way that equations (F.14) and (F.15) are compatible with one another. If we let $\breve{A} = A + \nabla \xi$, then equation (F.15) becomes

$$E = -\nabla \phi - \frac{\partial A}{\partial t} = -\nabla \phi - \frac{\partial}{\partial t} \left(\breve{A} - \nabla \xi \right) = -\nabla \left(\phi - \frac{\partial \xi}{\partial t} \right) - \frac{\partial \breve{A}}{\partial t}.$$

Hence, in order to preserve the form of the equations, we set $\phi - \partial \xi / \partial t = \breve{\phi}$.

Let us consider the two remaining Maxwell equations. Using expression (F.15), we can write equation (F.10) as

$$\nabla \cdot \left(\nabla \phi + \frac{\partial A}{\partial t} \right) = -\frac{\rho}{\epsilon_0},$$

and rewrite it as

$$\nabla^2 \phi + \frac{\partial}{\partial t} \nabla \cdot A = -\frac{\rho}{\epsilon_0}. \tag{F.16}$$

This equation relates the two potentials to the source.

Consider the last Maxwell equation, namely equation (F.13). We can rewrite it as

$$c^2 \nabla \times B - \frac{\partial E}{\partial t} = \frac{J}{\epsilon_0}. \tag{F.17}$$

Using expressions (F.14) and (F.15), we can rewrite equation (F.17) in terms of the potentials as

$$c^2 \nabla \times (\nabla \times A) + \frac{\partial}{\partial t} \left(\nabla \phi + \frac{\partial A}{\partial t} \right) = \frac{J}{\epsilon_0}.$$

Invoking the identity given by

$$\nabla \times (\nabla \times A) = \nabla (\nabla \cdot A) - \nabla^2 A,$$

and using the linearity of the differential operator, we get

$$c^2 \nabla (\nabla \cdot A) - c^2 \nabla^2 A + \frac{\partial}{\partial t} \nabla \phi + \frac{\partial^2 A}{\partial t^2} = \frac{J}{\epsilon_0}.$$

Exchanging the order of spatial and temporal derivatives, we can combine the first and the third terms on the left-hand side to write

$$-c^2 \nabla^2 A + \nabla \left(c^2 \nabla \cdot A + \frac{\partial \phi}{\partial t} \right) + \frac{\partial^2 A}{\partial t^2} = \frac{J}{\epsilon_0}.$$

Since the divergence of A is arbitrary due to the freedom of adding the gradient of any function, as discussed on page 172, let us set

$$\nabla \cdot A = -\frac{1}{c^2} \frac{\partial \phi}{\partial t}. \tag{F.18}$$

This way, the term in parentheses disappears and we get

$$-c^2 \nabla^2 A + \frac{\partial^2 A}{\partial t^2} = \frac{J}{\epsilon_0},$$

which we can rewrite as

$$\nabla^2 A - \frac{1}{c^2} \frac{\partial^2 A}{\partial t^2} = -\frac{J}{c^2 \epsilon_0}; \tag{F.19}$$

this is a vector equation that contains three equations that relate A_1, A_2 and A_3 to J_1, J_2 and J_3.

To complete our formulations of the Maxwell equations in terms of potentials, let us substitute expression (F.18) into equation (F.16) to obtain

$$\nabla^2 \phi - \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} = -\frac{\rho}{\epsilon_0}. \tag{F.20}$$

Thus, the classical form of the Maxwell equations given by expressions (F.10), (F.11), (F.12) and (F.13), which comprise eight equations, are equivalent to the five equations given by expressions (F.19), (F.20) and calibration equation (F.18). It is remarkable to notice that expressions (F.19) and (F.20) possess such a similarity of form to one another. Furthermore, they both have a form of a wave equation. Also, we notice the separation of physical quantities. In equations (F.19), the vector potential, A, is related to the current, J, while in equation (F.20), the scalar potential, ϕ, is related to the charge density, ρ.

Expressions (F.19) and (F.20) are four linear second-order partial differential equations for four unknowns: A_1, A_2, A_3 and ϕ. Together with equation (F.18), which is a linear first-order partial differential equation that relates equations (F.19) and (F.20), we have a system of five equations. Herein, we have formulated this system as a background to study its characteristic hypersurfaces in Section 2.4.2.

If we solved the system given by expressions (F.19) and (F.20) for A and ϕ, we could get B using equation (F.14) and, subsequently, E using equation (F.15).

We conclude this appendix by remarking the analogies between equations (F.20) and (E.12), and between equations (F.19) and (E.15); in other words, analogies between our formulation of elastodynamics and electrodynamics. In both cases we need the scalar and vector wave functions, as shown by equations (F.20) and (E.12) and by equations (F.19) and (E.15), respectively. Also, the left-hand sides of all four equations possess an identical form with c, $\sqrt{(\lambda + 2\mu)/\rho}$, $\sqrt{\mu/\rho}$ being the velocities with which, respectively, the electromagnetic wave, P wave, S wave propagate. The right hand-sides of equations (E.12) and (E.15) are zero, which means that we consider the case where there is no driving force, or source, generating the waves in an elastic medium. The right-hand sides of equations (F.19) and (F.20) indicate that the magnetic field is generated by the electric current while the electric field is generated by the electric charge. We could make the formal analogy even closer by considering sources in formulating equations (E.12) and (E.15). However, since the characteristic hypersurfaces, which are the focus of our study, are associated with the highest-order derivatives, the functions on the right-hand sides of equations (E.12) and (E.15) do not affect the characteristics. Unlike equations (E.12) and (E.15), equations (F.19) and (F.20) belong to a single system of equations. In the former case, the calibration equation is $\nabla \cdot A = 0$, while in the latter case, the analogous equation is $c^2 \nabla \cdot A = -\partial\phi/\partial t$, which relates the equations for A with the equation for ϕ.

G Oscillatory flow in incompressible fluid

Herein, we will formulate a differential equation that describes an oscillatory flow in an incompressible fluid. As we will see, depending on the coefficients, this equation is either an elliptic or a hyperbolic partial differential equation. The derivation in this appendix provides the physical motivation for the steady of characteristic hypersurfaces in Section ?.

G.1 Physical setup

Consider a spatial, so-called Eulerian, description with mass density given by

$$\rho = \rho_0 + \rho_1, \tag{G.1}$$

where ρ_0 corresponds to the mass density at equilibrium and ρ_1 to the Eulerian change of mass density due to a perturbation. Consider also the displacement vector given by

$$u = 0 + u$$

and the velocity vector given by

$$v = 0 + v, \tag{G.2}$$

where $v = \partial u/\partial t$. This notation for u and v emphasizes the fact that there is no displacement and no velocity at the state of equilibrium. Consider also the gravitational potential given by

$$U = U_0 + U_1, \tag{G.3}$$

with U_0 and U_1 corresponding to the equilibrium potential and to the change due to a perturbation, respectively. Finally, let us consider the stress tensor given by

$$\tilde{\mathcal{S}} = \tilde{\mathcal{S}}_0 + \tilde{\mathcal{S}}_1. \tag{G.4}$$

G.2 Conservation laws

Conservation of mass

Consider mass density given by?

$$\rho_1 = -\nabla \cdot \left[(\rho_0 + \rho_1)\, u \right].$$

Assuming that product $\rho_1 u$ is negligible, we can rewrite the above expression as

$$\rho_1 = -\nabla \cdot (\rho_0 u)\,. \tag{G.5}$$

Taking the derivative of equation (G.5) with respect to time and exchanging the order of spatial and temporal derivatives as well as using expression (G.2), we can write

$$\frac{\partial \rho_1}{\partial t} = -\nabla \cdot (\rho_0 v)\,, \tag{G.6}$$

which is the statement of conservation of energy.

Conservation of gravitational flux

Consider

$$\nabla^2 U = -4\pi G\rho, \tag{G.7}$$

where G is the gravitational constant.

Using equation (G.3) and in view of the linearity of the differential operator, we can write the left-hand side of equation (G.7) as

$$\nabla^2 U = \nabla^2 U_0 + \nabla^2 U_1.$$

Also, using equation (G.7), we can write

$$\nabla^2 U_0 = -4\pi G\rho_0$$

and

$$\nabla^2 U_1 = -4\pi G\rho_1$$

Invoking equation (G.5), we can rewrite the latter equation as

$$\nabla^2 U_1 = 4\pi G\nabla\cdot(\rho_0 u)\,. \tag{G.8}$$

Conservation of momentum

Consider a reference frame that is rotating with a steady angular velocity. We can write Newton's second law as

$$\rho\left[\frac{\partial v}{\partial t} + 2\Omega e_3 \times v + \Omega^2 e_3 \times (e_3 \times r)\right] = \nabla\cdot\tilde{S} + \rho\nabla U, \tag{G.9}$$

where Ω is the frequency of the angular rotation, r is the position vector from the origin on the rotation axis and e_3 is the unit vector along this axis. We note that the product of $-\rho$ and the second term on the left-hand side is the Coriolis force, and the product of $-\rho$ and the third term on the left-hand side is the centrifugal force. To write equation (G.9) more concisely, let us define

$$\mathcal{C} := \frac{1}{2}\Omega^2\left|e_3 \times r\right|^2$$

to write

$$\rho\left(\frac{\partial v}{\partial t} + 2\Omega e_3 \times v - \nabla\mathcal{C}\right) = \nabla\cdot\tilde{\mathcal{S}} + \rho\nabla U.$$

Reorganizing and defining the gravity potential as $W := U + \mathcal{C}$, we can write

$$\rho\left(\frac{\partial v}{\partial t} + 2\Omega e_3 \times v\right) = \nabla\cdot\tilde{\mathcal{S}} + \rho\nabla W.$$

Now, we wish to linearize this equation. First, we use equations (G.1) and (G.4) to replace ρ, U and $\tilde{\mathcal{S}}$, respectively, by the corresponding right-hand sides. Hence, we get

$$(\rho_0+\rho_1)\left(\frac{\partial v}{\partial t}+2\Omega e_3\times v\right)=\nabla\cdot\left(\tilde{\mathcal{S}}_0+\tilde{\mathcal{S}}_1\right)+(\rho_0+\rho_1)\nabla W.$$

Since $W:=U+\mathcal{C}$ and using equation (G.3), we can rewrite it as

$$\rho_0\left(\frac{\partial v}{\partial t}+2\Omega e_3\times v\right)+\rho_1\left(\frac{\partial v}{\partial t}+2\Omega e_3\times v\right)$$
$$=\nabla\cdot\tilde{\mathcal{S}}_0+\nabla\cdot\tilde{\mathcal{S}}_1+\rho_0\nabla\left(U_0+\mathcal{C}\right)+\rho_0\nabla U_1+\rho_1\nabla\left(U_0+\mathcal{C}\right)+\rho_1\nabla U_1.$$

Denoting

$$W_0:=U_0+\mathcal{C}, \tag{G.10}$$

we get

$$\rho_0\left(\frac{\partial v}{\partial t}+2\Omega e_3\times v\right)+\rho_1\left(\frac{\partial v}{\partial t}+2\Omega e_3\times v\right)$$
$$=\nabla\cdot\tilde{\mathcal{S}}_0+\nabla\cdot\tilde{\mathcal{S}}_1+\rho_0\nabla W_0+\rho_0\nabla U_1+\rho_1\nabla W_0+\rho_1\nabla U_1.$$

Ignoring the products and divergence of small terms as well as using the fact that?

$$0=\nabla\cdot\tilde{\mathcal{S}}_0+\rho_0\nabla W_0, \tag{G.11}$$

we obtain

$$\rho_0\left(\frac{\partial v}{\partial t}+2\Omega e_3\times v\right)=\nabla\cdot\tilde{\mathcal{S}}_1+\rho_0\nabla U_1+\rho_1\nabla W_0. \tag{G.12}$$

G.3 Hydrostatic equilibrium

Let us assume that

$$\tilde{\mathcal{S}}_0=-p_0 I,$$

which is the hydrostatic equilibrium. Herein, $\tilde{\mathcal{S}}_0$ is the stress tensor at the equilibrium, p_0 is the corresponding pressure and I stands for the unit 3×3 matrix.

The assumption of hydrostatic equilibrium has two important implication, which we describe below.

First, the surfaces of equal density coincide with the surfaces of equal gravitational potential. To see this, recall expression (G.11), namely,

$$\nabla\cdot\tilde{\mathcal{S}}_0+\rho_0\nabla W_0=0.$$

Denoting $g_0:=\nabla W_0$, we can write it as

$$\nabla\cdot\tilde{\mathcal{S}}_0=-\rho_0 g_0,$$

which implies that

$$\nabla p_0=\rho_0 g_0. \tag{G.13}$$

Examining equation (G.13) and in view of $g_0:=\nabla W_0$, we see that $\nabla p_0\parallel\nabla W_0$. This means that the surfaces of equal pressure coincide with the surfaces of equal gravitational potential.

Secondly, the surfaces of equal density coincide with the surfaces of equal gravitational potential. To see this, we take the curl of equation (G.13), which we rewrite using $g_0 := \nabla W_0$ to get

$$\nabla \times (\nabla p_0) = \nabla \times (\rho_0 \nabla W_0) .$$

Using the product rule and since the curl of the gradient vanishes identically, we get

$$0 = \nabla \rho_0 \times \nabla W_0. \tag{G.14}$$

Examining equation (G.14), we see that $\nabla \rho_0 \parallel \nabla W_0$. This means the surfaces of equal density coincide with the surfaces of equal gravitational potential.

G.4 Density stratification

In geophysical considerations we use the fact that the weight of the overlying material compresses matter in the Earth and causes the density to increase with depth. We can write

$$\nabla \rho_0 = (1 - \beta) \frac{\rho_0 g_0}{\alpha^2}, \tag{G.15}$$

where α is the sound speed and β is the departure from a strictly adiabatic density gradient or from uniform chemical composition. Let us also define

$$N^2 := \frac{1}{\rho_0} g_0 \cdot \left[\nabla \rho_0 - \frac{\rho_0 g_0}{\alpha^2} \right],$$

which, using expression (G.15), we can write as

$$N^2 = -\frac{\beta g_0 \cdot g_0}{\alpha^2}.$$

N is called the Brunt-Väisälä frequency.

We now consider a mass element that is displaced adiabatically by vector u to the point where the density of the neighbouring material is

$$\rho_0 (r + u) = \rho_0 (r) + u \cdot \nabla \rho_0.$$

At this point, the displaced mass element has the density given by

$$\rho_0 (r) + \left(\frac{\rho_0 g_0}{\alpha^2} \right) \cdot u.$$

Hence, the difference between the displaced-element mass density and density of its neighbourhood is

$$u \cdot \left(\frac{\rho_0 g_0}{\alpha^2} - \nabla \rho_0 \right), \tag{G.16}$$

which we can rewrite as

$$\beta \frac{\rho_0 u \cdot g_0}{\alpha^2}. \tag{G.17}$$

Examining expression (G.17), we conclude that if $\beta < 0$, the displaced mass element is buoyant and tends to move in direction $-u$. Also, if $\beta < 0$, N is real and — for small displacements

— the element of mass vibrates with frequency N. If $\beta > 0$, yet small, there is an exponential growth, which can be viewed as the onset of convection.

G.5 Equations of motion

Following Rayleigh hypothesis, we can write

$$\tilde{\mathcal{S}}_1 = \tilde{\tau} + u \cdot \nabla p_0 I, \tag{G.18}$$

where $\tilde{\tau}$ is the Cauchy stress tensor and the second term is the difference between stress on the mass element in the disturbed location and the stress that was experienced by this element in its equilibrium location. For a fluid,

$$\tilde{\tau} = (\lambda \nabla \cdot u)\, I, \tag{G.19}$$

with λ being the Lamé parameter that, herein, corresponds to incompressibility. Using expressions (G.13) and (G.19), we can rewrite expression (G.18) as

$$\tilde{\mathcal{S}}_1 = (\lambda \nabla \cdot u + \rho_0 u \cdot g_0)\, I.$$

Since the term in parentheses corresponds to the Eulerian pressure change that is associated with the disturbance, we can also write the above expression as

$$\tilde{\mathcal{S}}_1 = -p_1 I.$$

Equating the right-hand sides of the above two expressions, we can write this pressure as

$$p_1 = -\lambda \nabla \cdot u - \rho_0 u \cdot g_0. \tag{G.20}$$

Note that since in a fluid $\lambda = \rho_0 \alpha^2$ and since α is the adiabatic sound speed, we divide both sides of equation (G.20) by ρ_0 to write

$$\frac{p_1}{\rho_0} = -\alpha^2 \nabla \cdot u - u \cdot g_0.$$

We can view this equations as a statement of entropy conservation.

Using the product rule to write

$$\nabla \cdot (\rho_0 u) = \rho_0 \nabla \cdot u + u \cdot \nabla \rho_0$$

and recalling that in a fluid in equilibrium $\alpha^2 = \lambda / \rho_0$, we can rewrite expression (G.20) as

$$\begin{aligned} p_1 &= -\alpha^2 \left[\nabla \cdot (\rho_0 u) - u \cdot \nabla \rho_0\right] - \rho_0 g_0 \\ &= \alpha^2 \rho_1 + \alpha^2 u \cdot \left(\nabla \rho_0 - \frac{\rho_0 g_0}{\alpha^2}\right). \end{aligned}$$

Recognizing the dot product as expression (G.16), we can use expression (G.17) to write

$$p_1 = \alpha^2 \rho_1 - \beta \rho_0 u \cdot g_0.$$

Taking the derivative with respect to time and recalling expression (G.2), we obtain

$$\frac{\partial p_1}{\partial t} = \alpha^2 \frac{\partial \rho_1}{\partial t} - \beta \rho_0 v \cdot g_0. \tag{G.21}$$

Let us now examine equation (G.12). We can rewrite it as

$$\frac{\partial v}{\partial t} + 2\Omega e_3 \times v = -\frac{1}{\rho_0}\nabla \cdot \tilde{S}_1 + \nabla U_1 - \frac{1}{\rho_0} g_0 \nabla \cdot (\rho_0 u)\,. \tag{G.22}$$

Let us, for convenience, define

$$\chi := \frac{p_1}{\rho_0} - U_1. \tag{G.23}$$

Hence, immediately we can write $p_1/\rho_0 = \chi + U_1$ and, taking the gradient of this equation, get

$$\nabla\left(\frac{p_1}{\rho_0}\right) = \frac{1}{\rho_0}\nabla p_1 - \frac{p_1}{\rho_0^2}\nabla\rho_0 = \nabla\left(\chi + U_1\right).$$

Thus, we can write equation (G.22) as

$$\begin{aligned}
&\frac{\partial v}{\partial t} + 2\Omega e_3 \times v \\
&\qquad = -\nabla\chi + \left(\alpha^2 \nabla\cdot u + u\cdot g_0\right)\frac{\nabla\rho_0}{\rho_0} - \frac{g_0}{\rho_0}\left(\rho_0 \nabla\cdot u + u\cdot\nabla\rho_0\right) \\
&\qquad\qquad = -\nabla\chi + \nabla\cdot u\left(-g_0 + \alpha^2\frac{\nabla\rho_0}{\rho_0}\right) \\
&\qquad\qquad\qquad = -\nabla\chi + \nabla\cdot u\left[-g_0 + (1-\beta)\,g_0\right].
\end{aligned}$$

Consequently, we can write

$$\frac{\partial v}{\partial t} + 2\Omega e_3 \times v = -\nabla\chi - \beta g_0 \nabla\cdot u, \tag{G.24}$$

which together with equations (G.6), (G.8), (G.21) and (G.23) forms a complete system of equations of motion to describe the flow in a fluid.

G.6 Incompressible fluid

For our study of characteristic hypersurfaces that we discuss in Section ?, let us consider an oscillatory flow in an incompressible fluid.

Since there is no change in density in an incompressible fluid, $\partial\rho_1/\partial t = 0$ and, hence, equation (G.6) becomes

$$\nabla\cdot v = 0.$$

Also, if the fluid is incompressible, $\nabla\cdot u = 0$, and hence, equation (G.8) becomes

$$\nabla^2 U_1 = 0.$$

Using $\nabla\cdot u = 0$, for an incompressible fluid, equation (G.24) becomes

$$i\omega v + 2\Omega e_3 \times v = -\nabla\chi, \tag{G.25}$$

where we use $\partial/\partial t = i\omega$ with ω corresponding to the frequency of oscillation.

Taking the divergence of equation (G.25), we obtain

$$\nabla^2\chi = 2\Omega e_3 \cdot (\nabla \times v)\,.$$

Taking the curl of equation (G.25), we obtain

$$i\omega\nabla \times v = -2\Omega\nabla \times (e_3 \times v) = 2\Omega\,(e_3 \cdot \nabla)\,v = 2\Omega\frac{\partial v}{\partial x_3}.$$

Taking the dot product of e_3 with equation (G.25), we obtain

$$i\omega v_3 = -\frac{\partial \chi}{\partial x_3}.$$

Hence, writing

$$\frac{i\omega}{2\Omega}e_3 \cdot (\nabla \times v) = \frac{\partial v}{\partial x_3} = -\frac{1}{i\omega}\frac{\partial^2\chi}{\partial x_3^2},$$

we obtain

$$e_3 \cdot (\nabla \times \mathbf{v}) = \frac{2\Omega}{\omega^2}\frac{\partial^2\chi}{\partial x_3^2}.$$

Therefore,

$$\nabla^2\chi = 2\Omega e_3 \cdot (\nabla \times v) = \frac{4\Omega^2}{\omega^2}\frac{\partial^2\chi}{\partial x_3^2}. \tag{G.26}$$

We can rewrite equation (G.26) as

$$\left[\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \left(1 - \frac{4\Omega^2}{\omega^2}\right)\frac{\partial^2}{\partial x_3^2}\right]\chi = 0. \tag{G.27}$$

The differential operator in brackets is called the Poincaré (1885) operator. If $\omega > 2\Omega$, equation (G.27) is an elliptic partial differential equation. If $\omega < 2\Omega$, equation (G.27) is a hyperbolic partial differential equation. If $\omega = 2\Omega$, equation (G.27) reduces to a two-dimensional elliptic partial differential equation.

G.7 On boundary conditions

We assume a rigid boundary; namely, $\mathbf{v} \cdot \hat{\mathbf{n}} = 0$, where $\mathbf{n}$ is the unit outer normal vector. Also, we can write

$$i\omega\mathbf{v} = -\tilde{\Gamma} \cdot \nabla\chi, \tag{G.28}$$

where $\tilde{\Gamma}$ is a tensorial operator whose general form is

$$\tilde{\Gamma} = A\tilde{\mathbf{1}} + B\mathbf{e}_3\mathbf{e}_3 + C\mathbf{e}_3 \times \tilde{\mathbf{1}},$$

with A, B and C to be determined. Using expression (G.28), we can rewrite equation (G.25) as

$$-\tilde{\Gamma} \cdot \nabla\chi + \frac{2\Omega}{i\omega}\mathbf{e}_3 \times \left(-\tilde{\Gamma} \cdot \nabla\chi\right) = -\nabla\chi,$$

which leads to

$$\begin{cases} [c]cA + \frac{i}{\sigma}C = 1 \\ B - \frac{i}{\sigma}C = 0 \\ C - \frac{i}{\sigma}A = 0 \end{cases},$$

where $\sigma := \omega/2\Omega$. Therefore,

$$\tilde{\Gamma} = -\frac{1}{1-\sigma^2}\left(\sigma^2\tilde{\mathbf{1}} - \mathbf{e}_3\mathbf{e}_3 + i\sigma\mathbf{e}_3 \times \tilde{\mathbf{1}}\right).$$

Hence, the boundary condition is

$$\sigma^2\hat{\mathbf{n}} \cdot \nabla\chi - (\hat{\mathbf{n}} \cdot \mathbf{e}_3)(\mathbf{e}_3 \cdot \nabla\chi) + i\sigma(\hat{\mathbf{n}} \times \mathbf{e}_3) \cdot \nabla\chi = 0.$$

G.8 Characteristics

In cylindrical coordinates, (r, x_3, φ), equation (G.27) becomes

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\chi}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2\chi}{\partial\varphi^2} + \left(1 - \frac{4\Omega^2}{\omega^2}\right)\frac{\partial^2\chi}{\partial x_3^2} = 0.$$

If we consider the trial solution to be

$$\chi = \chi(r, x_3)\exp(im\varphi),$$

where $m \in \mathbb{Z}$, the characteristics are

$$x_3 \pm \frac{\sqrt{1-\sigma^2}}{\sigma}r = 0,$$

which represents cones about axis e_3 with the semi-apex angle given by $\arcsin\sigma$. We note that the characteristics are real if and only if $\omega < 2\Omega$.

H Transport equation on manifolds

H.1 Volume forms

Consider a general volume form in n dimensions:

$$f(x)|\mathrm{d}x_1 \wedge \mathrm{d}x_2 \wedge ... \wedge \mathrm{d}x_n|.$$

How does this change along a vector field X_H? To study this, we calculate the Lie derivative, as follows:

$$\mathcal{L}_{X_H}\left(f(x)|\mathrm{d}x_1 \wedge \mathrm{d}x_2 \wedge ... \wedge \mathrm{d}x_n|\right) = \left(\mathcal{L}_{X_H}f(x)\right)|\mathrm{d}x| + f(x)\left(\mathcal{L}_{X_H}|\mathrm{d}x|\right).$$

The first term in equation (**??**) is just the Lie derivative of function f along the curve times the wedge product. The second term is f multiplied by the Lie derivative of the wedge product; we need to know how $\mathcal{L}$ acts on $|\mathrm{d}x|$. Acting (L) on $|\mathrm{d}x|$, we obtain

$$\mathcal{L}_{X_H}|\mathrm{d}x| = \sum_i \mathrm{d}x_1 \wedge ... \wedge \mathrm{d}(\mathrm{d}x_i(X_H)) \wedge ... \wedge \mathrm{d}x_n$$

We need to know how the elements of the wedge product behave when acting on the vector field X_H. Consider the general form of vector field X_H:

$$X_H = \mathrm{d}x = \frac{\mathrm{d}x_i}{\mathrm{d}t}\frac{\partial}{\partial x_i} + \frac{\mathrm{d}p_i}{\mathrm{d}t}\frac{\partial}{\partial p_i}.$$

In view of Hamilton's equations for this system, namely, system (**??**), we write

$$X_H = \frac{\partial \sigma}{\partial p_i}\frac{\partial}{\partial x_i} - \frac{\partial \sigma}{\partial x_i}\frac{\partial}{\partial p_i}.$$

Now, we calculate the effect of element $\mathrm{d}x_i$ on vector field X_H:

$$\mathrm{d}x_i(X_H) = \frac{\partial \sigma}{\partial p_i}$$

The differential form of this element is then

$$\mathrm{d}(\mathrm{d}x_i(X_H)) = \mathrm{d}\left(\frac{\partial \sigma}{\partial p_i}\right) = \sum_j \left(\frac{\partial^2 \sigma}{\partial x_j \partial p_i}\mathrm{d}x_j + \frac{\partial^2 \sigma}{\partial p_j \partial p_i}\mathrm{d}p_j\right).$$

Hence, the effect of taking the Lie derivative of a wedge product is then

$$\mathcal{L}_{X_H}|\mathrm{d}x_1 \wedge ... \wedge \mathrm{d}x_n| = \sum_i \left(\frac{\partial^2 \sigma}{\partial x_j \partial p_i}|\mathrm{d}x_1 \wedge ... \wedge \mathrm{d}x_n| + \sum_j \left(\frac{\partial^2 \sigma}{\partial p_j \partial p_i}|\mathrm{d}x_1 \wedge ... \wedge \mathrm{d}p_j \wedge ... \wedge \mathrm{d}x_n|\right)\right).$$

Now, in view of the components of the slowness vector p, namely, $p_i = \frac{\partial \Psi}{\partial x_i}$, we write the differential form of $\mathrm{d}p_i$ as

$$\mathrm{d}p_j = \sum_k \frac{\partial^2 \Psi}{\partial x_k \partial x_j}.$$

We rewrite equation (**??**) as

$$\mathcal{L}_{X_H}|\mathrm{d}x_1 \wedge ... \wedge \mathrm{d}x_n| = \sum_i \left(\frac{\partial^2 \sigma}{\partial x_j \partial p_i}|\mathrm{d}x_1 \wedge ... \wedge \mathrm{d}x_n| + \sum_j \left(\frac{\partial^2 \sigma}{\partial p_j \partial p_i}\frac{\partial^2 \Psi}{\partial x_k \partial x_j}|\mathrm{d}x_1 \wedge ... \wedge \mathrm{d}x_n|\right)\right).$$

Our goal is to incorporate equation (**??**) into equation (**??**). To do so, we need a factor of 1/2 in front of the second term in equation (**??**).

H.2 Half-densities

Consider half-density ρ:

$$\rho = |\mathrm{d}x_1 \wedge ... \wedge \mathrm{d}x_n|^{\frac{1}{2}}.$$

Taking the square of ρ, we get

$$\rho^2 = |\mathrm{d}x_1 \wedge ... \wedge \mathrm{d}x_n|.$$

The Lie derivative of ρ^2 acts in the same way as taking the total derivative of x^2 with respect to x; in other words,

$$\mathcal{L}_{X_H}\rho^2 = 2\mathcal{L}_{X_H}\rho.$$

Rearranging terms, we obtain an expression for the Lie derivative of ρ, namely,

$$\mathcal{L}_{X_H}\rho = \frac{1}{2}\mathcal{L}_{X_H}\rho^2.$$

Now, consider the Lie derivative acting on the volume form ρA_0:

$$\mathcal{L}_{X_H}\rho A_0 = (X_H A_0)\rho + A_0\mathcal{L}_{X_H}\rho = (X_H A_0)\rho + \frac{A_0}{2}\mathcal{L}_{X_H}\rho^2.$$

H.3 Hamilton equations

Let the Hamiltonian of the system described by (**??**) be

$$\sigma(D)(\mathrm{d}\Psi).$$

Hamilton's equation of motion become:

$$\begin{aligned}\dot{x}_i &= \frac{\partial\sigma}{\partial p_i};\\ \dot{p}_i &= -\frac{\partial\sigma}{\partial x_i}.\end{aligned}$$

Define the following vector field along a ray:

$$X_H = \frac{\partial\sigma}{\partial p_i}\frac{\partial}{\partial x_i} - \frac{\partial\sigma}{\partial x_i}\frac{\partial}{\partial p_i}.$$

Since $\frac{\partial A}{\partial p} = 0$ (that is, amplitude is independent of direction), then we can say that along a ray parametrized by t,

$$\begin{aligned}\frac{\mathrm{d}A_0}{\mathrm{d}t} = \dot{x}_i\frac{\partial A_0}{\partial x_i} = \frac{\partial\sigma}{\partial p_i}\frac{\partial A_0}{\partial x_i} = 0\\ \rightarrow X_H(A_0) = 0.\end{aligned}$$

Returning to equation (**??**), we write it in terms of X_H:

$$\frac{1}{2}\sum_{i,j}\frac{\partial^2\sigma}{\partial p_i\partial p_j}\frac{\partial^2\Psi}{\partial x_i\partial x_j} + \chi A_0 + X_H A_0 = 0,$$

where the first term is zero,by definition of the eikonal equation, and we have let

$$\chi = \sum_{k\leq m}\sigma\left(D_{k-1}\right)\left(\mathrm{d}\psi\right).$$

H.4 Transport Equation

In view of equation (**??**), we can write the transport equation as

$$\mathcal{L}_{X_H}\rho A_0 + \left(\chi - \frac{1}{2}\sum_i\frac{\partial^2\sigma}{\partial x_i\partial p_j}\right)A_0\rho = 0.$$

Letting D^* be the adjoint to operator D, then (as shown in Gullemin and Sternberg), we can show that

$$\chi - \frac{1}{2}\sum_i \frac{\partial^2 \sigma}{\partial x_i \partial p_j} = \sigma(D - (-1)^m D^*),$$

we simplify equation (**??**) further:

$$\mathcal{L}_{X_H}\rho A_0 + \sigma(D - (-1)^m D^*)A_0\rho = 0.$$

With this result, we get a half-form that lives on the manifold that defines the surface of solution. Furthermore, we can find the solution to equation (**??**) that doesn't blow up when we approach a caustic. Of course, when we project onto the spatial plane, we will get solutions that are characteristic of the presence of a caustic, but that is natural for the coordinate system.

List of symbols

$e_1, \dots, e_n$ basis in $\mathbb{R}^n$
$x = (x_1, \dots, x_n)$ point in $\mathbb{R}^n$ represented by coordinates in a given basis
$p = (p_1, \dots, p_n)$ point in the dual space to $\mathbb{R}^n$
i, j, k, l summation indices
r, s parameters for parametrization of curves and surfaces
t time
$a, b, c, \dots$ auxiliary parameters or functions
$A, B, C, \dots, X, Y, Z$ tensors, matrices, vectors and auxiliary constants or functions
A, φ electromagnetic potentials
f, g, h functions
f' derivative of function f
$\Im$ imaginary part
$\Re$ real part
N normal vector
$X_1, \dots, X_n$ components of vector X in a given basis
D_X directional derivative in direction of X ; see expression (1.34) on page 14
$o(), O()$ see page 81
$\sim$ asymptotically equivalent to
α, β multiindices
ι imaginary unit: $\sqrt{-1}$
ν velocity
σ stress
ε strain
ρ density

References

Courant, R. and Hilbert, D. (1989). *Methods of mathematical physics*. Wiley.

Folland, G. B. (1995). *Introduction to Partial Differential Equations*. Princeton University Press, second edition.

Gel'fand, M., Graev, M. I., and Shapiro, Z. Y. (1969). Differential forms and integral geometry. *Functional Analysis and its Applications*, 3:101–114.

Guillemin, V. and Sternberg, S. (1977). *Geometric Asymptotics*. Number 14 in Mathematical Surveys. AMS.

Hörmander, L. (1983). *The Analysis of Linear Partial Differential Operators I-IV*. Grundlehren der matematischen Wissenschaften. Springer.

McOwen, R. C. (2003). *Partial differential equations: methods and applications*. Prentice Hall, 2nd edition.

Rogister, Y. and Slawinski, M. A. (2005). Analytic solution of raytracing equations for linearly inhomogeneous and elliptically anisotropic velocity model. *Geophysics*, 70:D37–D41.

Slawinski, M. A. (2003). *Seismic Waves and Rays in Elastic Media*, volume 34 of *Handbook of Geophysical Exploration*. Pergamon.

Spivak, M. (1999). *A comprehensive introduction to differential geometry*. Publish or perish.

Treves, F. (1980). *Introduction to Pseudodifferential and Fourier Integral Operators*. The University Series in Mathematics. Plenum.