

## VII. Aspects of Partial Differential Equations, 275-320

---

DOI: [10.3792/euclid/9781429799911-7](https://doi.org/10.3792/euclid/9781429799911-7)

from

*Advanced Real Analysis*  
*Digital Second Edition*  
*Corrected version 2017*

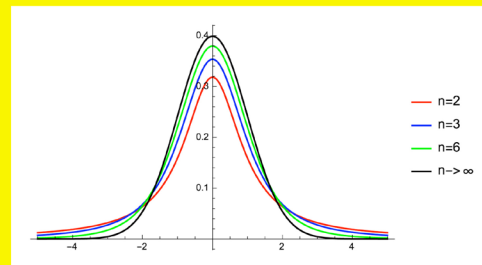
Anthony W. Knapp

Full Book DOI: [10.3792/euclid/9781429799911](https://doi.org/10.3792/euclid/9781429799911)

ISBN: 978-1-4297-9991-1

**ADVANCED  
REAL ANALYSIS**  
Digital Second Edition  
Corrected version 2017

Anthony W. Knapp



Anthony W. Knapp  
81 Upper Sheep Pasture Road  
East Setauket, N.Y. 11733–1729, U.S.A.  
Email to: [aknapp@math.stonybrook.edu](mailto:aknapp@math.stonybrook.edu)  
Homepage: [www.math.stonybrook.edu/~aknapp](http://www.math.stonybrook.edu/~aknapp)

Title: Advanced Real Analysis  
Cover: Normal distribution as a limit of Gosset's  $t$  distribution; see page 421.

Mathematics Subject Classification (2010): 46–01, 42–01, 43–01, 35–01, 34–01, 47–01, 58–01, 60A99, 60F05, 28C10, 42C40, 65T60.

First Edition, ISBN-13 978-0-8176-4382-9

©2007 Anthony W. Knapp  
Published by Birkhäuser Boston

Digital Second Edition, not to be sold, no ISBN  
©2016 Anthony W. Knapp, corrected version issued in 2017  
Published by the Author

All rights reserved. This file is a digital second edition of the above named book. The text, images, and other data contained in this file, which is in portable document format (PDF), are proprietary to the author, and the author retains all rights, including copyright, in them. The use in this file of trade names, trademarks, service marks, and similar items, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

All rights to print media for the first edition of this book have been licensed to Birkhäuser Boston, c/o Springer Science+Business Media Inc., 233 Spring Street, New York, NY 10013, USA, and this organization and its successor licensees may have certain rights concerning print media for the digital second edition. The author has retained all rights worldwide concerning digital media for both the first edition and the digital second edition.

The file is made available for limited noncommercial use for purposes of education, scholarship, and research, and for these purposes only, or for fair use as understood in the United States copyright law. Users may freely download this file for their own use and may store it, post it online, and transmit it digitally for purposes of education, scholarship, and research. They may not convert it from PDF to any other format (e.g., EPUB), they may not edit it, and they may not do reverse engineering with it. In transmitting the file to others or posting it online, users must charge no fee, nor may they include the file in any collection of files for which a fee is charged. Any exception to these rules requires written permission from the author.

Except as provided by fair use provisions of the United States copyright law, no extracts or quotations from this file may be used that do not consist of whole pages unless permission has been granted by the author (and by Birkhäuser Boston if appropriate).

The permission granted for use of the whole file and the prohibition against charging fees extend to any partial file that contains only whole pages from this file, except that the copyright notice on this page must be included in any partial file that does not consist exclusively of the front cover page. Such a partial file shall not be included in any derivative work unless permission has been granted by the author (and by Birkhäuser Boston if appropriate).

Inquiries concerning print copies of either edition should be directed to Springer Science+Business Media Inc.

## CHAPTER VII

### Aspects of Partial Differential Equations

**Abstract.** This chapter provides an introduction to partial differential equations, particularly linear ones, beyond the material on separation of variables in Chapter I.

Sections 1–2 give an overview. Section 1 addresses the question of how many side conditions to impose in order to get local existence and uniqueness of solutions at the same time. The Cauchy–Kovalevskaya Theorem is stated precisely for first-order systems in standard form and for single equations of order greater than one. When the system or single equation is linear with constant coefficients and entire holomorphic data, the local holomorphic solutions extend to global holomorphic solutions. Section 2 comments on some tools that are used in the subject, particularly for linear equations, and it gives some definitions and establishes notation.

Section 3 establishes the basic theorem that a constant-coefficient linear partial differential equation  $Lu = f$  has local solutions, the technique being multiple Fourier series.

Section 4 proves a maximum principle for solutions of second-order linear elliptic equations  $Lu = 0$  with continuous real-valued coefficients under the assumption that  $L(1) = 0$ .

Section 5 proves that any linear elliptic equation  $Lu = f$  with constant coefficients has a “parametrix,” and it shows how to deduce from the existence of the parametrix the fact that the solutions  $u$  are as regular as the data  $f$ . The section also deduces a global existence theorem when  $f$  is compactly supported; this result uses the existence of the parametrix and the constant-coefficient version of the Cauchy–Kovalevskaya Theorem.

Section 6 gives a brief introduction to pseudodifferential operators, concentrating on what is needed to obtain a parametrix for any linear elliptic equation with smooth variable coefficients.

#### 1. Introduction via Cauchy Data

The subject of partial differential equations is a huge and diverse one, and a short introduction necessarily requires choices. The subject has its origins in physics and nowadays has applications that include physics, differential geometry, algebraic geometry, and probability theory. A small amount of complex-variable theory will be extremely helpful, and this will be taken as known for this chapter. We shall ultimately concentrate on single equations, as opposed to systems, and on partial differential equations that are linear. After the first two sections the topics of this chapter will largely be ones that can be approached through a combination of functional analysis and Fourier analysis.

Let us for now use subscript notation for partial derivatives, as in Section I.1. A **system of  $p$  partial differential equations** in  $N$  variables for the unknown functions  $u^{(1)}, \dots, u^{(m)}$  consists of  $p$  expressions

$$F_k(u^{(1)}, \dots, u^{(m)}, u_{x_1}^{(1)}, \dots, u_{x_1}^{(m)}, \dots, u_{x_N}^{(1)}, \dots, u_{x_N}^{(m)}, u_{x_1 x_1}^{(1)}, \dots, u_{x_1 x_1}^{(m)}, \dots) = 0,$$

$1 \leq k \leq p$ , in an open set of  $\mathbb{R}^N$ ; it is assumed that the partial derivatives that appear as variables have bounded order. When  $p = 1$ , we speak of simply a **partial differential equation**. The highest order of a partial derivative that appears is the **order** of the equation or system. We might expect that it would be helpful if the number  $p$  of equations in a system equals the number  $m$  of unknown functions, but one does not insist on this condition as a matter of definition. A system in which the number  $p$  of equations equals the number  $m$  of unknown functions is said to be “determined,” but nothing is to be read into this terminology without a theorem. We shall work only with determined systems. The equation or system is **linear homogeneous** if each  $F_k$  is a linear function of its variables. It is **linear** if each  $F_k$  is the sum of a linear function and a function of the  $N$  domain variables that is taken as known.

The classical equations that we would like to include in a more general theory are the three studied in Section I.2 in connection with the method of separation of variables—the heat equation, the Laplace equation, and the wave equation—and one other, namely the Cauchy–Riemann equations. With  $\Delta$  denoting the Laplacian  $\Delta u = u_{x_1 x_1} + \dots + u_{x_N x_N}$ , the first three of these equations in  $N$  space variables are

$$u_t = \Delta u, \quad \Delta u = 0, \quad \text{and} \quad u_{tt} = \Delta u.$$

The **Cauchy–Riemann equations** are ordinarily written as a system

$$u_x = v_y, \quad u_y = -v_x,$$

but they can be written also as a single equation if we think of  $u$  and  $v$  as real and write  $f = u + iv$ . Then the system is equivalent to the single equation

$$\frac{\partial f}{\partial \bar{z}} = 0 \quad \text{or} \quad f_{\bar{z}} = 0, \quad \text{where} \quad \frac{\partial}{\partial \bar{z}} = \frac{\partial}{\partial x} + i \frac{\partial}{\partial y}.$$

Guided in part by the theory of ordinary differential equations of Chapter IV in *Basic*, we shall be interested in existence-uniqueness questions for our equation or system, both local and global, and in qualitative properties of solutions, such as regularity, the propagation of singularities, and any special features. For a particular equation or system we might be interested in any of the following three problems:

- (i) to find one or more particular solutions,
- (ii) to find all solutions,
- (iii) to find those solutions meeting some initial or boundary conditions.

Problems of the third type as known as **boundary-value problems** or **initial-value problems**.<sup>1</sup> The method of separation of variables in Section I.2 is particularly adapted to solving this kind of problem in special situations.

For ordinary differential equations and systems these three problems are closely related, as we saw in the course of investigating existence and uniqueness in Chapter IV of *Basic*. For partial differential equations they turn out to be comparatively distinct. We can, however, use the kind of setup with first-order systems of ordinary differential equations to get an idea how much flexibility there is for the solutions to the system. Let us treat one of the variables  $x$  as distinguished<sup>2</sup> and suppose, in analogy with what happened in the case of ordinary differential equations, that the system consists of an expression for the derivative with respect to  $x$  of each of the unknown functions in terms of the variables, the unknown functions, and the other first partial derivatives of the functions. Writing down general formulas involves complicated notation that may obscure the simple things that happen; thus let us suppose concretely that the independent variables are  $x, y$  and that the unknown functions are  $u, v$ . The system is then to be

$$\begin{aligned}u_x &= F(x, y, u, v, u_y, v_y), \\v_x &= G(x, y, u, v, u_y, v_y).\end{aligned}$$

With  $x$  still regarded as special, let us suppose that  $u$  and  $v$  are known when  $x = 0$ , i.e., that

$$\begin{aligned}u(0, y) &= f(y), \\v(0, y) &= g(y).\end{aligned}$$

The real-variable approach of Chapter IV of *Basic* is not very transparent for this situation; an approach via power series looks much easier to apply. Thus we assume whatever smoothness is necessary, and we look for formal power series solutions in  $x, y$ . The question is then whether we can determine all the partial derivatives of all orders of  $u$  and  $v$  at a point like  $(0, 0)$ . It is enough to see that the system and the initial conditions determine  $\frac{\partial^k u}{\partial x^k}(0, y)$  and  $\frac{\partial^k v}{\partial x^k}(0, y)$  for all  $k \geq 0$ . For  $k = 0$ , the initial conditions give the values. For  $k = 1$ , we substitute  $x = 0$  into the system itself and get values, provided we know values of all the variables at  $(0, y)$ . The values of  $u$  and  $v$  come from  $k = 0$ , and the values of  $u_y$  and  $v_y$

---

<sup>1</sup>The distinction between these terms has nothing to do with the mathematics and instead is a question of whether all variables are regarded as space variables or one variable is to be interpreted as a time variable.

<sup>2</sup>It is natural to think of this variable as representing time and to say that the differential equation and any conditions imposed at a particular value of this variable constitute an initial-value problem.

come from differentiating those expressions with respect to  $y$ . For  $k = 2$ , we differentiate each equation of the system with respect to  $x$  and then put  $x = 0$ . For each equation we get a sum of partial derivatives of  $F$ , evaluated as before, times the partial of each variable with respect to  $x$ . For the latter we need expressions for  $u_x, v_x, u_{xy}$ , and  $v_{xy}$ ; we have them since we know  $u_x(0, y)$  and  $v_x(0, y)$  from the step  $k = 1$ . This handles  $k = 2$ . For higher  $k$ , we can proceed inductively by continuing to differentiate the given system, but let us skip the details. The result is that the initial values of  $u(0, y)$  and  $v(0, y)$  are enough to determine unique formal power-series solutions satisfying those initial values.

Next, under the hypothesis that  $F, G, f$ , and  $g$  are holomorphic functions of their variables near an initial point, one can prove convergence of the resulting two-variable power series near  $(0, 0)$ . This fact persists when the number of equations and the number of unknown functions are increased but remain equal, and when the domain variables are arbitrary in number. The theorem is as follows.

**Theorem 7.1** (Cauchy–Kovalevskaya Theorem, first form). Let a system of  $p$  partial differential equations with  $p$  unknown functions  $u^{(1)}, \dots, u^{(p)}$  and  $N$  variables  $x_1, \dots, x_N$  of the form

$$\begin{aligned} u_{x_1}^{(1)} &= F_1(u^{(1)}, \dots, u^{(p)}, u_{x_2}^{(1)}, \dots, u_{x_2}^{(p)}, \dots, u_{x_N}^{(1)}, \dots, u_{x_N}^{(p)}), \\ &\vdots \\ u_{x_1}^{(p)} &= F_p(u^{(1)}, \dots, u^{(p)}, u_{x_2}^{(1)}, \dots, u_{x_2}^{(p)}, \dots, u_{x_N}^{(1)}, \dots, u_{x_N}^{(p)}), \end{aligned} \quad (*)$$

be given, subject to the initial conditions

$$\begin{aligned} u^{(1)}(0, x_2, \dots, x_N) &= f_1(x_2, \dots, x_N), \\ &\vdots \\ u^{(p)}(0, x_2, \dots, x_N) &= f_p(x_2, \dots, x_N). \end{aligned} \quad (**)$$

Suppose that  $f_1, \dots, f_p$  are holomorphic in a neighborhood in  $\mathbb{C}^{N-1}$  of the point  $(x_2, \dots, x_N) = (x_2^0, \dots, x_N^0)$  and that  $F_1, \dots, F_p$  are holomorphic in a neighborhood in  $\mathbb{C}^{Np}$  of the value of the argument  $u^{(1)}, \dots, u_{x_N}^{(p)}$  of the  $F_j$ 's that corresponds to  $(0, x_2^0, \dots, x_N^0)$ . Then there exists a neighborhood of  $(x_1, x_2, \dots, x_N) = (0, x_2^0, \dots, x_N^0)$  in  $\mathbb{C}^N$  in which the system (\*) has a holomorphic solution satisfying the initial conditions (\*\*). Moreover, on any connected subneighborhood of  $(0, x_2^0, \dots, x_N^0)$ , there is no other holomorphic solution satisfying the initial conditions.

We omit the proof since we shall use the theorem in this generality only as a guide for how much in the way of initial conditions needs to be imposed to expect uniqueness without compromising existence. Initial conditions of the form (\*\*) for a system of equations (\*) are called **Cauchy data**.

We shall, however, make use of a special case of Theorem 7.1, where a better conclusion is available.

**Theorem 7.2.** In the Cauchy–Kovalevskaya system of Theorem 7.1, suppose that the functions  $F_k$  in the system (\*) are of the form

$$F_k(u^{(1)}, \dots, u^{(p)}, u_{x_2}^{(1)}, \dots, u_{x_2}^{(p)}, \dots, u_{x_N}^{(1)}, \dots, u_{x_N}^{(p)}) \\ = \sum_{i=1}^p a_i u^{(i)} + \sum_{i=1}^p \sum_{j=2}^N c_{ij} u_{x_j}^{(i)} + h_k(x_1, \dots, x_N)$$

with the  $a_i$  and  $c_{ij}$  constant and with each  $h_j$  a given entire holomorphic function on  $\mathbb{C}^N$ . Suppose further that the functions  $f_j(x_2, \dots, x_N)$  in the initial conditions (\*\*) are entire holomorphic functions on  $\mathbb{C}^N$ . Then the system (\*) has an entire holomorphic solution satisfying the initial conditions (\*\*).

This theorem is proved in Problems 6–9 at the end of the chapter without making use of Theorem 7.1. We shall use it in proving Theorem 7.4 below, which in turn will be applied in Section 5.

Since our interest is really in single equations and we want to allow order  $> 1$ , we can ask whether we can carry over to partial differential equations the familiar device for ordinary differential equations of introducing new unknown functions to change a higher-order equation to a first-order system.

Recall with an ordinary differential equation of order  $n$  for an unknown function  $y(t)$  when the equation is  $y^{(n)} = F(t, y, y', \dots, y^{(n-1)})$ : we can introduce unknown functions  $y_1, \dots, y_n$  satisfying  $y_1 = y, y_2 = y', \dots, y_n = y^{(n-1)}$ , and we obtain an equivalent first-order system  $y_1' = y_2, \dots, y_{n-1}' = y_n, y_n' = F(t, y_1, y_2, \dots, y_n)$ . Values for  $y, y', \dots, y^{(n-1)}$  at  $t = t_0$  correspond to values at  $t = t_0$  for  $y_1, y_2, \dots, y_n$  and give us equivalent initial-value problems.

For a single higher-order partial differential equation of order  $m$  in which the  $m^{\text{th}}$  derivative of the unknown function with respect to one of the variables  $x$  is equal to a function of everything else, the same kind of procedure changes a suitable initial-value problem into an initial-value problem for a first-order system as above. But if we ignore the initial values, the solutions of the single equation need not match the solutions of the system. Let us see what happens for a single second-order equation in two variables  $x, y$  for an unknown function  $u$  under the assumption that we have solved for  $u_{xx}$ . Thus consider the equation

$$u_{xx} = F(x, y, u, u_x, u_y, u_{xy}, u_{yy})$$

with initial data

$$\begin{aligned}u(0, y) &= f(y), \\u_x(0, y) &= g(y).\end{aligned}$$

This is another instance in which the initial data are known as **Cauchy data**: the equation has order  $m$ , and we are given the values of  $u$  and its derivatives through order  $m - 1$  with respect to  $x$  at the points of the domain where  $x = 0$ . For this example, introduce variables  $u, p, q, r, s, t$  equal, respectively, to  $u, u_x, u_y, u_{xx}, u_{xy}, u_{yy}$ . With these interpretations of the variables, the given equation becomes  $r = F(x, y, u, p, q, s, t)$ , and we differentiate this identity to make it more convenient to use. Then  $u$  yields a solution of a system of six first-order equations, namely

$$\begin{aligned}u_x &= p, \\p_x &= r, \\q_x &= p_y, \\r_x &= F_x + pF_u + rF_p + sF_q + r_yF_s + s_yF_t, \\s_x &= r_y, \\t_x &= s_y.\end{aligned}$$

The choice here of  $q_x = p_y$  rather than  $q_x = s$  is important; we will not be able to invert the initial-value problem without it. The initial data will be values of  $u, p, q, r, s, t$  at  $(0, y)$ , and we can read off what we must use from the above values of  $u(0, y)$  and  $u_x(0, y)$ , namely

$$\begin{aligned}u(0, y) &= f(y), \\p(0, y) &= g(y), \\q(0, y) &= f'(y), \\r(0, y) &= F(0, y, f(y), g(y), f'(y), g'(y), f''(y)), \\s(0, y) &= g'(y), \\t(0, y) &= f''(y).\end{aligned}$$

If  $u$  satisfies the initial-value problem for the single equation, then the definitions of  $u, p, q, r, s, t$  give us a solution of the initial-value problem for the system.

Let us show that a solution  $u, p, q, r, s, t$  of the initial-value problem for the system has to make  $u$  be a solution of the initial-value problem for the single equation. What needs to be shown is that  $u_y = q$ ,  $u_{xy} = s$ , and  $u_{yy} = t$ . We use the same kind of argument with all three.



For  $u_y = q$ , we see from the system that  $(u_y)_x = (u_x)_y = p_y = q_x$ , so that  $(u_y - q)_x = 0$ . Therefore  $u_y(x, y) - q(x, y) = h(y)$  for some function  $h$ . Setting  $x = 0$  gives  $h(y) = u_y(0, y) - q(0, y) = f'(y) - f'(y) = 0$ . Thus  $h(y) = 0$ , and we obtain  $u_y = q$ .

Similarly for  $u_{xy} = s$ , we start from  $u_{xxy} = p_{xy} = r_y = s_x$ , so that  $(u_{xy} - s)_x = 0$ . Therefore  $u_{xy}(x, y) - s(x, y) = k(y)$  for some function  $k$ . Setting  $x = 0$  gives  $k(y) = u_{xy}(0, y) - s(0, y) = p_y(0, y) - s(0, y) = g'(y) - g'(y) = 0$ . Thus  $k(y) = 0$ , and we obtain  $u_{xy} = s$ .

Finally for  $u_{yy} = t$ , we start from  $u_{xyy} = (u_{xy})_y = s_y = t_x$ , so that  $(u_{yy} - t)_x = 0$ . Therefore  $u_{yy}(x, y) - t(x, y) = l(y)$  for some function  $l$ . Setting  $x = 0$  gives  $l(y) = u_{yy}(0, y) - t(0, y) = f''(y) - f''(y) = 0$ . Thus  $l(y) = 0$ , and we obtain  $u_{yy} = t$ .

The conclusion is that the given second-order equation with two initial conditions is equivalent to the system of six first-order equations with six initial conditions. In other words the Cauchy data for the single equation lead to Cauchy data for an equivalent first-order system. It turns out that if a single equation of order  $m$  has one unknown function and is written as solved for the  $m^{\text{th}}$  derivative of one of the variables  $x$ , and if the given Cauchy data consist of the values at  $x = x_0$  of the unknown function and its derivatives through order  $m - 1$ , then the equation can always be converted in this way into an equivalent first-order system with given Cauchy data. The steps of the reduction to Theorem 7.1 are carried out in Problems 10–11 at the end of the chapter. The result is as follows.

**Theorem 7.3** (Cauchy–Kovalevskaya Theorem, second form). Let a single partial differential equation of order  $m$  in the variables  $(x, y) = (x, y_1, \dots, y_{N-1})$  of the form

$$D_x^m u = F(x, y; u; \text{all } D_x^k D_y^\alpha u \text{ with } k < m \text{ and } k + |\alpha| \leq m) \quad (*)$$

be given, subject to the initial conditions

$$D_x^i u(0, y) = f^{(i)}(y) \quad \text{for } 0 \leq i < m. \quad (**)$$

Here  $\alpha$  is assumed to be a multi-index  $\alpha = (\alpha_1, \dots, \alpha_{N-1})$  corresponding to the  $y$  variables. Suppose that  $f^{(0)}, \dots, f^{(m-1)}$  are holomorphic in a neighborhood in  $\mathbb{C}^{N-1}$  of the point  $(y_1, \dots, y_{N-1}) = (y_1^0, \dots, y_{N-1}^0)$  and that  $F$  is holomorphic in a neighborhood of the value of its argument corresponding to  $x = 0$  and  $(y_1, \dots, y_{N-1}) = (y_1^0, \dots, y_{N-1}^0)$ . Then there exists a neighborhood of  $(x, y_1, \dots, y_{N-1}) = (0, y_1^0, \dots, y_{N-1}^0)$  in  $\mathbb{C}^N$  in which the system  $(*)$  has a holomorphic solution satisfying the initial conditions  $(**)$ . Moreover, on any connected subneighborhood of  $(0, y_1^0, \dots, y_{N-1}^0)$ , there is no other holomorphic solution satisfying the initial conditions.

In the special case that  $F$  is the sum of a known entire holomorphic function and a linear combination with constant coefficients of  $x$ ,  $y$ , and the various  $D_x^k D_y^\alpha u$ , the steps that reduce Theorem 7.3 to Theorem 7.1 perform a reduction to Theorem 7.2. We therefore obtain a better conclusion under these hypotheses, as follows.

**Theorem 7.4.** Let a single partial differential equation of order  $m$  in the variables  $(x, y) = (x, y_1, \dots, y_{N-1})$  of the form

$$D_x^m u = ax + b_1 y_1 + \dots + b_{N-1} y_{N-1} + \sum_{\substack{0 \leq k < m \\ k + |\alpha| \leq m}} c_{k,\alpha} D_x^k D_y^\alpha u + h(x, y_1, \dots, y_{N-1}) \quad (*)$$

be given, subject to the initial conditions

$$D_x^i u(0, y) = f^{(i)}(y) \quad \text{for } 0 \leq i < m. \quad (**)$$

Suppose that  $f^{(0)}, \dots, f^{(m-1)}$  are entire holomorphic on  $\mathbb{C}^{N-1}$  and that  $h$  is entire holomorphic on  $\mathbb{C}^N$ . Then the equation (\*) has an entire holomorphic solution satisfying the initial conditions (\*\*).

The steps in the reduction of this theorem to Theorem 7.2 are indicated for  $N = 2$  in Problem 11 at the end of the chapter, and the steps for general  $N$  are similar. We shall make use of Theorem 7.4 to prove the existence of certain “fundamental solutions” in Section 5.

As we said, in this reduction from an initial-value problem for a single equation to an initial-value problem for a first-order system, the equation without initial values is *not* always equivalent to the system without initial values. A simple example will suffice. In the second-order setup as above, let the given equation be  $u_{xx} = -u_{yy} + 4$ . That is, let  $F(x, y, u, u_x, u_y, u_{xy}, u_{yy}) = -u_{yy} + 4$ . This equation has  $u = x^2 + y^2$  as a solution, for example. If we introduce variables  $u, p, q, r, s, t$  as above, we find that  $F(x, y, u, p, q, s, t) = -t + 4$ , and we obtain the system

$$\begin{aligned} u_x &= p, \\ p_x &= r, \\ q_x &= p_y, \\ r_x &= F_x + pF_u + rF_p + sF_q + r_y F_s + s_y F_t = -s_y, \\ s_x &= r_y, \\ t_x &= s_y. \end{aligned}$$

If we put

$$u = x^2, \quad p = 2x, \quad q = s = 0, \quad r = t = 2,$$

we find that this tuple  $(u, p, q, r, s, t)$  solves the system. But  $u = x^2$  is not a solution of  $u_{xx} = -u_{yy} + 4$ .

There is a still more general Cauchy–Kovalevskaya Theorem than anything we have considered, still involving local holomorphic systems, data, and solutions. It amounts to whatever one can get by combining the Implicit Function Theorem, the technique of reduction of order via an increase in the number of equations, and Theorem 7.1. We omit the precise statement. The word “noncharacteristic” is used to describe situations in which the Implicit Function Theorem applies for this purpose.

Cauchy data are not the only kinds of initial data that one might consider. In fact, none of the examples with separation of variables in Section I.2 used Cauchy data. A typical example from that section is the Dirichlet problem for the Laplacian in the unit disk. The equation can be written as  $u_{xx} = -u_{yy}$ , and Cauchy data would consist of values of  $u(x_0, y)$  and  $u_x(x_0, y)$ . This amounts to two functions on a piece of a line in the plane, and one could handle two functions of a suitable curve in the plane after applying the Implicit Function Theorem. By contrast, the Dirichlet problem requires just a single function on the unit circle for a unique solution. A more apt comparison is to think of a Sturm–Liouville problem as being an ordinary-differential-equations analog of the Dirichlet problem. A particular Sturm–Liouville problem to compare with the Dirichlet problem for the disk is the equation  $u_{xx} = 0$  with boundary conditions  $u(0) = u(\pi) = 0$ . The region is a ball in 1-dimensional space, and the function is specified on the boundary; the function is uniquely determined without specifying the derivative on the boundary. However, if the equation is changed to  $u_{xx} = -\lambda u$  for some positive constant  $\lambda$ , then there is a nonunique solution when  $\lambda$  is the square of a nonzero integer.

## 2. Orientation

After this essay on what is appropriate for existence and uniqueness, let us turn to some other aspects of partial differential equations and systems. A few principles and observations will influence what we do in the upcoming sections of this chapter.

*The subjects of linear systems and nonlinear systems of partial differential equations cannot be completely separated.*

For example let  $a(x, y)$  and  $b(x, y)$  be given functions on an open set in  $\mathbb{R}^2$ , and consider the single linear equation

$$a(x, y)u_x + b(x, y)u_y = 0$$

for an unknown function  $u(x, y)$ . If we look for curves  $c(t) = (x(t), y(t))$  along which such a function  $u(x, y)$  is constant, the condition on  $c$  is that  $(\frac{d}{dt})u(x(t), y(t)) = 0$ , hence that

$$x'(t)u_x(x(t), y(t)) + y'(t)u_y(x(t), y(t)) = 0.$$

One way for this equation to be satisfied is that  $c(t) = (x(t), y(t))$  satisfy the system

$$\begin{aligned}x'(t) &= a(x, y), \\y'(t) &= b(x, y),\end{aligned}$$

of two ordinary differential equations. This system is nonlinear, and the condition for  $c(t)$  to solve it is that  $c(t)$  be an integral curve. Thus  $u$  is a solution if it is constant along each integral curve. If we introduce two parameters, one varying along an integral curve and the other indexing a family of integral curves, then we obtain solutions by letting  $u$  be any function of the second parameter. Under reasonable assumptions, these solutions turn out to be the only solutions locally, and thus the solution of a certain linear partial differential equation reduces to solving a nonlinear system in fewer variables. Despite this circumstance the partial differential equations of interest to us will be the linear ones.

*As we have seen, there is a distinction between the reduction of a partial differential equation to a first-order system of Cauchy type and the reduction of a Cauchy problem for the equation to the corresponding Cauchy problem for the first-order system.*

One consequence is that finding a several-parameter set of solutions of a partial differential equation may not be very helpful in solving a specific boundary-value problem about the equation. With an eye on the wave equation, let us take as an example a homogeneous linear equation with constant coefficients. Let  $P : \mathbb{R}^{N+1} \rightarrow \mathbb{C}$  be a polynomial such as  $P(x_0, x_1, \dots, x_N) = x_0^2 - x_1^2 - \dots - x_N^2$  in the case of the wave equation,  $x_0$  being the time variable. We write the equation in our notation with  $D$  as

$$P(D)u = 0,$$

understanding as usual that  $\partial/\partial x_j$  is to be substituted in  $P$  everywhere that  $x_j$  appears. If  $a$  is any  $(N + 1)$ -tuple, then  $(\partial/\partial x_j)e^{a \cdot x} = a_j e^{a \cdot x}$ . Consequently  $P(D)e^{a \cdot x} = P(a)e^{a \cdot x}$ , and  $e^{a \cdot x}$  solves the equation  $P(D)u = 0$  whenever  $P(a) = 0$ . Concretely with the wave equation, let  $\alpha$  be a real number, let  $\beta = (\beta_1, \dots, \beta_N)$  be in  $\mathbb{R}^N$ , and write  $x = (t, x')$ . Then  $e^{\alpha t - \beta \cdot x'}$  solves the wave equation whenever  $\alpha^2 = |\beta|^2$ . Apart from the one constraint  $\alpha^2 = |\beta|^2$ , we obtain an  $N$ -parameter family of solutions of the wave equation. But this family of solutions is not of any obvious help in solving boundary-value problems such as those encountered in Section I.2. We shall discuss this example further shortly.

*Global problems involving linear partial differential equations with constant coefficients lend themselves to use of the Fourier transform.*

The reason is that the Fourier transform carries differentiation into multiplication by a function. Specifically under suitable conditions on  $f$ , the relevant formula is  $\mathcal{F}\left(\frac{\partial f}{\partial x_j}\right)(\xi) = 2\pi i \xi_j (\mathcal{F}f)(\xi)$  if we use  $\xi$  for the Fourier transform variable.

Thus, at least on a formal level, to find a solution of an inhomogeneous equation  $P(D)u = f$ , we can take the Fourier transform of both sides, obtaining  $P(2\pi i \xi)(\mathcal{F}u)(\xi) = (\mathcal{F}f)(\xi)$ . Then we divide by  $P(2\pi i \xi)$  and take the inverse Fourier transform. In Section III.1 we carried out the steps of this process for the equation  $(1 - \Delta)u = f$  when  $f$  is in the Schwartz space. In this case the polynomial is  $1 + 4\pi^2 |\xi|^2$ , and we found that there is a solution  $u$  in the Schwartz space.

In practice the function  $P(2\pi i \xi)$  may be zero in some places, and then we have to check what happens with the division. There will also be a matter of ensuring that the inverse Fourier transform is well defined where we want it to be.

In Section 3 we shall use multiple Fourier series to see that a linear equation  $P(D)u = f$  with constant coefficients and with  $f$  in  $C_{\text{com}}^\infty(\mathbb{R}^N)$  always has a solution in a neighborhood of a point. It is of interest also to know what happens when  $f$  is replaced by a function with fewer derivatives or even by a distribution of compact support. This matter is addressed in Problem 5 at the end of the chapter.

*For a linear partial differential equation of order  $m$ , the terms with differentiations of total order  $m$  are especially important. Moreover, a linear equation with variable coefficients can sometimes be studied near a point  $x_0$  of the domain by applying a “freezing principle.”*

We explain the notion of a freezing principle in a moment. We shall now make use of the notation of Chapter V for linear differential operators  $L$ , often writing an equation under study as  $Lu = f$  with  $f$  known and  $u$  unknown. Here  $L$  is given by

$$L = P(x, D) = \sum_{|\alpha| \leq m} a_\alpha(x) D^\alpha$$

for some  $m$ , or we can write

$$L = P(x, D_x) = \sum_{|\alpha| \leq m} a_\alpha(x) D_x^\alpha$$

if the variable  $x$  of differentiation needs emphasis. It is customary to assume that  $m$  is the order of  $L$ , in which case some  $a_\alpha(x)$  with  $|\alpha| = m$  is not identically zero.

The domain is to be an open set in real Euclidean space, usually  $\mathbb{R}^N$ ; thus  $x$  varies in that open set, and the multi-index  $\alpha$  is an  $N$ -tuple of nonnegative integers.

The idea of a freezing principle is that the behavior of solutions of  $P(x, D)u = f$  near  $x = x_0$  can sometimes be studied by considering solutions of the equation  $(P(x_0, D_x)u)(x) = f(x)$  and making estimates for how much effect the variability of  $x$  might have. For equations that are “elliptic” in a sense that we define shortly, the classical approach to the equations via something called “Gårding’s inequality” used this idea and worked well. We shall indicate a more recent approach via “pseudodifferential operators” in Section 6 and will omit any discussion of details concerning Gårding’s inequality in our development. The freezing principle is somewhat concealed within the mechanism of pseudodifferential operators, but it is at least visible in the notation that is used for such operators.

As far as theorems for nonelliptic operators are concerned, the idea of a freezing principle is meaningful but has its limitations. We have noted that linear differential equations with constant coefficients are at least locally solvable, a result that will be proved in Section 3. But the same is not always true for equations with variable coefficients. In 1957 Hans Lewy gave an example in  $\mathbb{R}^3$  involving the linear differential operator

$$P(x, D) = -(D_1 + iD_2) + 2i(x_1 + ix_2)D_3.$$

For a certain function  $f$  of class  $C^\infty$  that is nowhere real analytic, the equation  $P(x, D)u = f$  admits no solution in any nonempty open set. By contrast, if  $f$  is holomorphic, the Cauchy–Kovalevskaya Theorem (Theorem 7.3) ensures the existence of local solutions.

In the linear differential operator  $P(x, D_x) = \sum_{|\alpha| \leq m} a_\alpha(x) D_x^\alpha$ , the terms of highest order are of special interest; we group them and give them their own name:

$$P_m(x, D_x) = \sum_{|\alpha|=m} a_\alpha(x) D_x^\alpha.$$

In line with the freezing principle, when one takes a Fourier transform, one does not apply the Fourier transform to the coefficients of  $L$ , only to the various  $D_x^\alpha$ ’s. Recalling that  $D_x^\alpha$  goes into multiplication by  $(2\pi i)^{|\alpha|} \xi^\alpha$  under the Fourier transform, we introduce the expressions<sup>3</sup>

---

<sup>3</sup>The Fourier transform variable  $\xi$  lies in the dual space of  $\mathbb{R}^N$ . To take maximum advantage of this fact in more advanced treatments, one wants to identify  $\mathbb{R}^N$  with the tangent space at  $x$  to the domain open set. Then  $\xi$  is to be regarded as a member of the dual of the tangent space of  $x$ , and to some extent, the formalism makes sense on smooth manifolds. We elaborate on these remarks in Chapter VIII.

$$P(x, 2\pi i\xi) = \sum_{|\alpha| \leq m} a_\alpha(x) (2\pi i\xi)^\alpha$$

and

$$P_m(x, 2\pi i\xi) = \sum_{|\alpha|=m} a_\alpha(x) (2\pi i\xi)^\alpha.$$

These are called the **symbol** and the **principal symbol** of  $L$ , respectively.

EXAMPLES. The Laplacian, the wave operator, and the heat operator have order  $m = 2$ , while the Cauchy–Riemann operator has  $m = 1$ . In all these cases except the heat operator, the symbol and the principal symbol coincide. The operators written with the notation  $D$  are

$$\begin{aligned} \Delta = \Delta_x &= D_1^2 + \cdots + D_N^2 && \text{in } \mathbb{R}^N && \text{(Laplacian),} \\ \frac{\partial}{\partial \bar{z}} &= D_1 + iD_2 && && \text{(Cauchy–Riemann operator),} \\ \square &= D_0^2 - \Delta_x && \text{in } \mathbb{R}^{N+1} && \text{(wave operator),} \\ &D_0 - \Delta_x && \text{in } \mathbb{R}^{N+1} && \text{(heat operator).} \end{aligned}$$

The principal symbols  $P_m(x, 2\pi i\xi)$  in each case are independent of  $x$  and are as follows:

$$\begin{aligned} -4\pi^2(\xi_1^2 + \cdots + \xi_N^2) &&& \text{(Laplacian),} \\ 2\pi i\xi_1 - 2\pi\xi_2 &&& \text{(Cauchy–Riemann operator),} \\ -4\pi^2\xi_0^2 + 4\pi^2(\xi_1^2 + \cdots + \xi_N^2) &&& \text{(wave operator),} \\ 4\pi^2(\xi_1^2 + \cdots + \xi_N^2) &&& \text{(heat operator).} \end{aligned}$$

*Complex analysis inevitably plays an important role in the study of partial differential equations.*

We already saw that complex analysis is useful in addressing the Cauchy problem. The Lewy example shows that complex analysis has to play a role in drawing a distinction between linear equations with constant coefficients, where we always have local existence of solutions, and linear equations with variable coefficients, where local existence can fail if the inhomogeneous term of the equation is merely  $C^\infty$ . Actually, the complex analysis that enters the local existence theorem in Section 3 for linear equations with constant coefficients is rather primitive and can be absorbed into facts about polynomials in several variables. Complex analysis enters in a more serious way for more advanced theorems about partial differential equations, but we shall not pursue theorems that go in this direction beyond one application in Section 5 of Theorem 7.4.

*Linear partial differential equations can exhibit behavior of kinds not seen in ordinary differential equations.*

The operator  $L$  on an open set in  $\mathbb{R}^N$  is said to be **elliptic** at  $x$  if  $P_m(x, 2\pi i\xi) = 0$  for  $\xi \in \mathbb{R}^N$  only when  $\xi = 0$ . The operator  $L$  is **elliptic** if it is elliptic at every point  $x$  of its domain. The Laplacian and the Cauchy–Riemann operator are elliptic, but the wave operator and the heat operator are not. A linear ordinary differential operator with nonvanishing coefficient for the highest-order derivative is automatically elliptic. We shall be especially interested in elliptic operators, which are relatively easy to handle.

In Section I.2 we considered the **Dirichlet problem** for the unit disk in  $\mathbb{R}^2$ , namely the problem of finding a function  $u$  satisfying  $\Delta u = 0$  in the interior and taking prescribed values on the boundary. The problem was solved by the Poisson integral formula. No matter how rough the function on the boundary was, the solution  $u$  in the interior was a smooth function. Theorem 3.16 extended this conclusion of smoothness, showing that solutions of  $\Delta u = 0$  in any open set of  $\mathbb{R}^N$  are automatically  $C^\infty$ . This behavior is typical of solutions of linear elliptic differential equations with smooth coefficients.

Other partial differential equations can behave quite differently. Consider the wave equation  $((\frac{\partial}{\partial t})^2 - \Delta_x)u = 0$  with  $x \in \mathbb{R}^n$ . We have seen that  $u(t, x) = e^{\alpha t - \beta \cdot x}$  is a solution if  $\alpha$  is a number and  $\beta$  is a vector with  $\alpha^2 = |\beta|^2$ . But actually the exponential function is not important here. If  $f$  is any  $C^2$  function of one variable, then  $f(\alpha t - \beta \cdot x)$  is a solution as long as  $\alpha^2 = |\beta|^2$  is satisfied: in fact,  $((\frac{\partial}{\partial t})^2 - \Delta_x)f(\alpha t - \beta \cdot x) = f''(\alpha t - \beta \cdot x)(\alpha^2 - |\beta|^2)$ . Such a solution represents an undistorted progressing wave; the roughness of the wave is maintained as time progresses. Again, this kind of behavior is not exhibited by elliptic equations.

In the special case that  $L$  is of order 2 with real coefficients and a point  $x_0$  is specified, we can make a linear change of variables in  $\xi$  to bring the order-two terms of the operator into a certain standard form at  $x_0$  that makes the question of ellipticity transparent. This change of variables amounts to replacing the standard basis  $e_1, \dots, e_N$  used for determining the first partial derivatives  $D_1, \dots, D_N$  by a new basis  $e'_1, \dots, e'_N$  and the corresponding first partial derivatives  $D'_1, \dots, D'_N$ . The result is as follows.

**Proposition 7.5.** If  $L = P(x, D)$  is of order 2 and has real coefficients in an open set of  $\mathbb{R}^N$  and if a point  $x_0$  is specified, then there exists a nonsingular  $N$ -by- $N$  real matrix  $M = [M_{ij}]$  such that the definition  $D'_j = \sum_k M_{jk} D_k$  exhibits  $L$  at  $x_0$  as of the form  $\kappa_1 D_1'^2 + \dots + \kappa_N D_N'^2$  with each  $\kappa_j$  equal to  $+1$ ,  $-1$ , or  $0$ . The principal symbol of  $L$  at  $x_0$  is then  $-4\pi^2 \sum_j \kappa_j \xi_j'^2$ , where  $\xi_j' = \sum_k M_{jk} \xi_k$ .

REMARKS. We see immediately that  $L$  is elliptic at  $x_0$  if and only if all  $\kappa_j$  are  $+1$  or all are  $-1$ . This is the situation with the Laplacian. In Section 4 we



shall prove a maximum principle for certain elliptic operators of order 2 with real coefficients, generalizing the corresponding result for the Laplacian given in Corollary 3.20. If one  $\kappa_j$  is  $+1$  and the others are  $-1$ , or if one is  $-1$  and the others are  $+1$ , the operator is said to be **hyperbolic** at  $x_0$ ; this is the situation with the wave operator. Much is known about hyperbolic operators of this kind and about generalizations of them, but the study of such operators remains a continuing subject of investigation.

**Lemma 7.6** (Principal Axis Theorem). If  $B$  is a real symmetric matrix, then there exist a nonsingular real matrix  $M$  and a diagonal matrix  $C$  whose diagonal entries are each  $+1$ ,  $-1$ , or  $0$  such that  $B = M^{\text{tr}} C M$ .

PROOF. By the finite-dimensional Spectral Theorem for self-adjoint operators, choose an orthogonal matrix  $P$  such that  $P B P^{-1}$  is some real diagonal matrix  $E$ . Any real number is the product of a square and one of  $+1$ ,  $-1$ , and  $0$ , and thus  $E = Q C Q$  with  $C$  as in the lemma and with  $Q = Q^{\text{tr}}$  diagonal and nonsingular. Since  $P$  is orthogonal,  $P^{-1} = P^{\text{tr}}$ , and therefore  $B = P^{\text{tr}} Q^{\text{tr}} C Q P$ . This proves the lemma with  $M = Q P$ .  $\square$

PROOF OF PROPOSITION 7.5. Let the principal symbol be

$$P_2(x, 2\pi i \xi) = \sum_{|\alpha|=2} a_\alpha(x) (2\pi i \xi)^\alpha = -4\pi^2 \sum_{|\alpha|=2} a_\alpha(x) \xi^\alpha.$$

We rewrite this in matrix notation, viewing  $\xi = (\xi_1, \dots, \xi_N)$  as a column vector and converting  $\{a_\alpha(x)\}$  into a matrix by defining

$$\begin{aligned} b_{jj}(x) &= a_\alpha(x) && \text{if } \alpha \text{ is } 2 \text{ in the } j^{\text{th}} \text{ entry and } 0 \text{ elsewhere,} \\ b_{jk}(x) &= \frac{1}{2} a_\alpha(x) && \text{if } \alpha \text{ is } 1 \text{ in the } j^{\text{th}} \text{ and } k^{\text{th}} \text{ entries and } 0 \text{ elsewhere.} \end{aligned}$$

Then  $B(x) = [b_{jk}(x)]$  is a symmetric matrix, and

$$P_2(x, 2\pi i \xi) = -4\pi^2 \sum_{j,k} b_{jk}(x) \xi_j \xi_k = -4\pi^2 \xi^{\text{tr}} B(x) \xi.$$

We apply the lemma to the real symmetric matrix  $B = B(x_0)$  to obtain  $B(x_0) = M^{\text{tr}} C(x_0) M$  with  $M$  nonsingular and with  $C(x_0)$  diagonal of the form in the lemma. Define  $C(x)$  by  $B(x) = M^{\text{tr}} C(x) M$ , write  $C(x) = [c_{jk}(x)]$  and  $M = [m_{jk}]$ , and put  $\xi' = M \xi$ . Then  $P_2(x, 2\pi i \xi) = -4\pi^2 \xi^{\text{tr}} B(x) \xi = -4\pi^2 \xi'^{\text{tr}} (M^{\text{tr}} C(x) M) \xi = -4\pi^2 \xi'^{\text{tr}} C(x) \xi'$ . If we set  $D'_j = \sum_k M_{jk} D_k$ , then the algebraic manipulations for the order-two part of  $L$  are the same as with the principal symbol and show that the order-two part of the operator is given by  $P_2(x, D) = \sum_{j,k} b_{jk}(x) D_j D_k = \sum_{j,k} c_{jk}(x) D'_j D'_k$ . The matrix  $C(x_0)$  is diagonal with diagonal entries  $+1$ ,  $-1$ , and  $0$ , and the proposition follows.  $\square$

*Ways are needed for making routine the passage via the Fourier transform between differentiations and multiplications by polynomials.*

We are going to be using the Fourier transform to transform any linear equation  $Lu = f$ , at least in the constant-coefficient case, into a problem involving division by a polynomial and inversion of a Fourier transform. It is inconvenient to check repeatedly the technical conditions in Proposition 8.1 of *Basic* that relate differentiations and multiplications by polynomials. Weak derivatives and Sobolev spaces as discussed in Chapter III, and distributions as discussed in Chapter V, all help us handle easily the passage via the Fourier transform between differentiations and multiplications by polynomials.

*“Fundamental solutions” are useful for obtaining all solutions of a linear partial differential equation, especially for constant-coefficient equations. In the case of an elliptic equation, a substitute for a fundamental solution that is easier to find is a “parametrix,” which at least reveals qualitative properties of solutions.*

In Section I.3 we encountered Green’s functions in connection with Sturm–Liouville theory. The operator  $L$  under study in that section was a second-order ordinary differential operator, and a Green’s function was the kernel of an integral operator  $T_1$  that we used. To understand symbolically what was happening there, let us take  $r = 1$  in Section I.3, and then the operator  $T$ , which is the same as the operator  $T_1$  for  $r = 1$  in that section, sets up a one-one correspondence between a class of functions  $u$  and a class of functions  $f$ , the relationship being that  $u = Tf$  and  $Lu = f$ . In other words  $T$  was a two-sided inverse of  $L$ . The operator  $T$  was of the form  $Tf(x) = \int_a^b G(x, y)f(y) dy$ . If we think symbolically of taking  $f$  to be a point mass  $\delta_{x_0}$  at  $x_0$ , then we find that  $T(\delta_{x_0})(x) = G(x, x_0)$ , and the relationship is to be  $L(G(\cdot, x_0)) = \delta_{x_0}$ . In other words the Green’s function at  $x_0$  is a **fundamental solution**  $u$  of the equation  $Lu = f$  in the sense that application of  $L$  to it yields a point mass at  $x_0$ .

These matters can easily be made rigorous with distributions of the kind introduced in Chapter V. In the case that  $L$  has constant coefficients, the notion of a fundamental solution is especially useful because the operator  $L$  commutes with translations. If a certain  $u$  produces  $Lu = \delta_0$ , then translation of that  $u$  by some  $x_0$  produces a solution of  $Lu = \delta_{x_0}$ . In short, one obtains a fundamental solution for each point by finding it just for one point, and all solutions may be regarded as the sum of a weighted average of fundamental solutions at the various points plus a solution of  $Lu = 0$ . In practice we can carry out this process of weighted average by means of convolution of distributions. Corollary 5.23 carried out the details for the Laplacian in  $\mathbb{R}^N$ , once Theorem 5.22 had identified a fundamental solution at 0.

In the case of the Laplacian in all of  $\mathbb{R}^N$ , Theorem 5.22 showed that a fundamental solution at 0 is a multiple of  $|x|^{-(N-2)}$  if  $N > 2$ . But fundamental solutions

are at best inconvenient to obtain for other equations, and a certain amount of the qualitative information they yield, at least in the elliptic case, can be obtained more easily from a “parametrix,” which is a kind of approximate fundamental solution. To illustrate matters, consider the inhomogeneous version  $\Delta u = f$  of the Laplace equation, which is known as **Poisson’s equation**. Suppose that  $f$  is in  $C_{\text{com}}^{\infty}(\mathbb{R}^N)$  and we seek information about a possible solution  $u$ . We shall use the Fourier transform, and therefore  $u$  had better be a function or distribution whose Fourier transform is well defined. But let us leave aside the question of what kind of function  $u$  is, going ahead with the computation. If we take the Fourier transform of both sides, we are led to ask whether the following inverse Fourier transform is meaningful:

$$-4\pi^2 \int_{\mathbb{R}^N} e^{2\pi i x \cdot \xi} |\xi|^{-2} \widehat{f}(\xi) d\xi.$$

Here  $\widehat{f}(\xi)$  is in the Schwartz space, but the singularity of  $|\xi|^{-2}$  at the origin does not put  $|\xi|^{-2} \widehat{f}(\xi)$  into any evident space of Fourier transforms. To compensate, we use Proposition 3.5f to introduce a function  $\chi \in C_{\text{com}}^{\infty}(\mathbb{R}^N)$  that is identically 0 near the origin and is identically 1 away from the origin. Then  $\chi(\xi)|\xi|^{-2} \widehat{f}(\xi)$  has no singularity and is in fact in the Schwartz space. It thus makes sense to define

$$Qf(x) = -4\pi^2 \int_{\mathbb{R}^N} e^{2\pi i x \cdot \xi} \chi(\xi) |\xi|^{-2} \widehat{f}(\xi) d\xi,$$

where  $Qf(x)$  is the Schwartz function with

$$\widehat{Qf}(\xi) = -4\pi^2 \chi(\xi) |\xi|^{-2} \widehat{f}(\xi).$$

Since  $\Delta f$  is in  $C_{\text{com}}^{\infty}(\mathbb{R}^N)$  and  $Qf$  is a Schwartz function,  $Q\Delta f$  and  $\Delta Qf$  are Schwartz functions. Applying the Fourier transform operator  $\mathcal{F}$ , as it is defined on the Schwartz space, we calculate that

$$\mathcal{F}(Q\Delta f) = \chi \widehat{f} = \mathcal{F}(\Delta Qf).$$

Hence  $\mathcal{F}(Q\Delta f - f) = \mathcal{F}(\Delta Qf - f) = (\chi - 1) \widehat{f}$ .

The function  $\chi - 1$  on the right side is in  $C_{\text{com}}^{\infty}(\mathbb{R}^N)$ , and it is therefore the Fourier transform of some Schwartz function  $K$ . Since  $\mathcal{F}$  carries convolutions into products, we have  $\widehat{Kf} = \widehat{K} * \widehat{f}$ , and consequently

$$Q\Delta = \Delta Q = 1 + (\text{convolution by } K).$$

The operator of convolution by  $K$  is called a “smoothing operator” because, as follows from the development of Chapter V, it carries arbitrary distributions of

compact support into smooth functions. The operator  $Q$  that gives a two-sided inverse for  $\Delta$  except for the smoothing term is called a **parametrix** for  $\Delta$ .

The parametrix does not solve our equation for us, but it does supply useful information. As we shall see in Section 5, a parametrix will enable us to see that whenever  $u$  is a distribution solution of  $\Delta u = f$  on an open set  $U$ , with  $f$  an arbitrary distribution on  $U$ , then  $u$  is smooth wherever  $f$  is smooth. In particular, any distribution solution of  $\Delta u = 0$  is a smooth function. The argument will apply to any elliptic linear partial differential equation with constant coefficients. A first application of the method of pseudodifferential operators in Section 6 shows that the same conclusion is valid for any elliptic linear partial differential equation with smooth variable coefficients.

### 3. Local Solvability in the Constant-Coefficient Case

We come to the local existence of solutions to linear partial differential equations with constant coefficients.

**Theorem 7.7.** Let  $U$  be an open set in  $\mathbb{R}^N$  containing 0, and let  $f$  be in  $C^\infty(U)$ . If  $P(D)$  is a linear differential operator with constant coefficients and with order  $\geq 1$ , then the equation  $P(D)u = f$  has a smooth solution in a neighborhood of 0.

The proof will use multiple Fourier series as in Section III.7. Apart from that, all that we need will be some manipulations with polynomials in several variables and an integration. As in Section III.7, let us write  $\mathbb{Z}^N$  for the set of all integer  $N$ -tuples and  $[-\pi, \pi]^N$  for the region of integration defining the Fourier series.

We shall give the idea of the proof, state a lemma, prove the theorem from the lemma, and then return to the proof of the lemma. The idea of the proof of Theorem 7.7 is as follows: We begin by multiplying  $f$  by a smooth function that is identically 1 near the origin and is identically 0 off some small ball containing the origin (existence of the smooth function by Proposition 3.5f), so that  $f$  is smooth of compact support, the support lying well inside  $[-\pi, \pi]^N$ . If we regard  $f$  as extended periodically to a smooth function, we can write  $f(x) = \sum_{k \in \mathbb{Z}^N} d_k e^{ik \cdot x}$  by Proposition 3.30e. Let the unknown function  $u$  be given by  $u(x) = \sum_{k \in \mathbb{Z}^N} c_k e^{ik \cdot x}$ . Then  $P(D)u(x)$  is given by

$$P(D)u(x) = \sum_{k \in \mathbb{Z}^N} c_k P(ik) e^{ik \cdot x},$$

and thus we want to take  $c_k P(ik) = d_k$ . We are done if  $\frac{d_k}{P(ik)}$  decreases faster than any  $|k|^{-n}$ , by Proposition 3.30c and our computations. So we would like to prove that

$$|P(ik)|^{-1} \leq C(1 + |k|^2)^M \quad \text{for all } k \in \mathbb{Z}^N$$

and for some constants  $C$  and  $M$ , and then we would be done. Unfortunately this is not necessarily true; the polynomial  $P(x) = |x|^2$  is a counterexample. What is true is the statement in the following lemma, and we can readily adjust the above idea to prove the theorem from this lemma.

**Lemma 7.8.** If  $R(x)$  is any complex-valued polynomial not identically 0 on  $\mathbb{R}^N$ , then there exist  $\alpha \in \mathbb{R}^N$  and constants  $C$  and  $M$  such that

$$|R(k + \alpha)|^{-1} \leq C(1 + |k|^2)^M \quad \text{for all } k \in \mathbb{Z}^N.$$

PROOF OF THEOREM 7.7. Apply the lemma to  $R(x) = P(ix)$ . Because of the preliminary step of multiplying  $f$  by something, we are assuming that  $f$  is smooth and has support near 0. Instead of extending  $f$  to be periodic, as suggested in the discussion before the lemma, we extend the function  $f(x)e^{-i\alpha \cdot x}$  to be smooth and periodic. Thus write

$$f(x)e^{-i\alpha \cdot x} = \sum_{k \in \mathbb{Z}^N} d_k e^{ik \cdot x},$$

and put  $c_k = \frac{d_k}{R(k + \alpha)}$ . Since the  $|d_k|$  decrease faster than  $|k|^{-n}$  for any  $n$ , Lemma 7.8 and Proposition 3.30c together show that  $\sum_{k \in \mathbb{Z}^N} c_k e^{ik \cdot x}$  is smooth and periodic. Define

$$u(x) = e^{i\alpha \cdot x} \sum_{k \in \mathbb{Z}^N} c_k e^{ik \cdot x} = \sum_{k \in \mathbb{Z}^N} c_k e^{i(k + \alpha) \cdot x}.$$

This function is smooth but maybe is not periodic. Application of  $P(D)$  gives

$$\begin{aligned} P(D)u(x) &= \sum_{k \in \mathbb{Z}^N} c_k P(i(k + \alpha)) e^{i(k + \alpha) \cdot x} \\ &= e^{i\alpha \cdot x} \sum_{k \in \mathbb{Z}^N} \frac{d_k}{R(k + \alpha)} P(i(k + \alpha)) e^{ik \cdot x} \\ &= e^{i\alpha \cdot x} \sum_{k \in \mathbb{Z}^N} d_k e^{ik \cdot x} = e^{i\alpha \cdot x} (f(x)e^{-i\alpha \cdot x}) = f(x), \end{aligned}$$

and hence  $u$  solves the equation for the original  $f$  in a neighborhood of the origin.  $\square$

The proof of Lemma 7.8 requires two lemmas of its own.

**Lemma 7.9.** For each positive integer  $m$  and positive number  $\delta < \frac{1}{m}$ , there exists a constant  $C$  such that

$$\int_{-1}^1 |x - c_1|^{-\delta} \cdots |x - c_m|^{-\delta} dx \leq C$$

for any  $m$  complex numbers  $c_1, \dots, c_m$ .

PROOF. For  $1 \leq j \leq m$ , let  $E_j$  be the subset of  $[-1, 1]$  where  $|x - c_j|^{-\delta}$  is the largest of the  $m$  factors in the integrand. The integral in question is then

$$\begin{aligned} &\leq \sum_{j=1}^m \int_{E_j} |x - c_1|^{-\delta} \cdots |x - c_m|^{-\delta} dx \\ &\leq \sum_{j=1}^m \int_{E_j} |x - c_j|^{-m\delta} dx \leq \sum_{j=1}^m \int_{-1}^1 |x - c_j|^{-m\delta} dx \\ &\leq \sum_{j=1}^m \int_{-1}^1 |x - \operatorname{Re} c_j|^{-m\delta} dx \leq m \sup_{r \in \mathbb{R}} \int_{-1}^1 |x - r|^{-m\delta} dx. \end{aligned}$$

On the right side the integrand decreases pointwise with  $|r|$  when  $|r| \geq 1$ , and hence the expression is equal to

$$\begin{aligned} &m \sup_{-1 \leq r \leq 1} \int_{-1}^1 |x - r|^{-m\delta} dx \\ &= m \sup_{-1 \leq r \leq 1} \left( \int_{-1}^r (r - x)^{-m\delta} dx + \int_r^1 (x - r)^{-m\delta} dx \right) \\ &= m(1 - m\delta)^{-1} \sup_{-1 \leq r \leq 1} \left( (1 + r)^{1-m\delta} + (1 - r)^{1-m\delta} \right) \\ &\leq 2^{2-m\delta} m(1 - m\delta)^{-1}. \quad \square \end{aligned}$$

**Lemma 7.10.** If  $R(x)$  is any complex-valued polynomial on  $\mathbb{R}^N$  of degree  $m > 0$ , then  $|R(x)|^{-\delta}$  is locally integrable whenever  $\delta < \frac{1}{m}$ .

PROOF. We first treat the special case that  $x_1^m$  has coefficient 1 in  $R(x)$  and that integrability on the cube  $[-1, 1]^N$  is to be checked. Write  $x'$  for  $(x_2, \dots, x_N)$ , so that  $x = (x_1, x')$ . Then  $R(x) = x_1^m + \sum_{j=0}^{m-1} x_1^j p_j(x')$ , where each  $p_j$  is a polynomial. For fixed  $x'$ ,  $R(x_1, x')$  is a monic polynomial of degree  $m$  in  $x_1$  and factors as  $(x_1 - c_1) \cdots (x_1 - c_m)$  for some complex numbers  $c_1, \dots, c_m$  depending on  $x'$ . Applying Lemma 7.9, we see that  $\int_{-1}^1 |R(x_1, x')|^{-\delta} dx_1 \leq C$ . Integration in the remaining  $N - 1$  variables therefore gives  $\int_{[-1, 1]^N} |R(x)|^{-\delta} dx \leq 2^{N-1} C$ .

Turning to the general case, suppose that  $R(x)$  and a point  $x_0$  are given. We want to see that  $F(x) = R(x + x_0)$  has the property that  $|F(x)|^{-\delta}$  is integrable on some neighborhood of the origin in  $\mathbb{R}^N$ . The function  $F$  is still a polynomial of degree  $m$ . Let  $F_m$  be the sum of all its terms of total degree  $m$ . This cannot be identically 0 on the unit sphere since it is a nonzero homogeneous function,<sup>4</sup> and thus  $F_m(v_1) \neq 0$  for some unit vector  $v_1$ . Extend  $\{v_1\}$  to an orthonormal basis of  $\mathbb{R}^N$ , and define  $G(y_1, \dots, y_N) = F_m(y_1 v_1 + \cdots + y_N v_N)$ . The function  $G$  is a polynomial of degree  $m$  whose coefficient of  $y_1^m$  is  $F_m(v_1)$  and hence is not 0, and it is obtained by applying an orthogonal transformation to the variables of  $F$ . Therefore  $|G|^{-\delta}$  and  $|F|^{-\delta}$  have the same integral over a ball centered at the origin. The special case shows that  $|G|^{-\delta}$  is integrable over some such ball, and hence so is  $|F|^{-\delta}$ .  $\square$

<sup>4</sup>A function  $F_m$  of several variables is **homogeneous of degree  $m$**  if  $F_m(rx) = r^m F_m(x)$  for all  $r > 0$  and all  $x \neq 0$ .

PROOF OF LEMMA 7.8. Let  $R$  have degree  $m$ , which we may assume is positive without loss of generality. The function  $S(x) = |x|^{2m} R\left(\frac{x}{|x|^2}\right)$  is then a polynomial of degree  $\leq 2m$ , and Lemma 7.10 shows that any number  $\delta$  with  $\delta < \frac{1}{2m}$  has the property that  $|R|^{-\delta}$  and  $|S|^{-\delta}$  are integrable for  $|x| \leq 1$ . Using spherical coordinates and making the change of variables  $r \mapsto 1/r$  in the radial direction, we see that

$$\begin{aligned} \int_{|x| \geq 1} |R(x)|^{-\delta} |x|^{-2N} dx &= \int_{r=1}^{\infty} \int_{\omega \in S^{N-1}} |R(r\omega)|^{-\delta} r^{-2N} d\omega r^{N-1} dr \\ &= \int_{r=0}^1 \int_{\omega \in S^{N-1}} |R(r^{-1}\omega)|^{-\delta} d\omega r^{N-1} dr \\ &= \int_{|x| \leq 1} |R(x/|x|^2)|^{-\delta} dx \\ &= \int_{|x| \leq 1} |S(x)|^{-\delta} |x|^{2m\delta} dx \\ &\leq \int_{|x| \leq 1} |S(x)|^{-\delta} dx. \end{aligned}$$

The right side is finite. Since  $(1 + |x|^2)^{-N} \leq 1 + |x|^{-2N}$ , we see that

$$\int_{\mathbb{R}^N} |R(x)|^{-\delta} (1 + |x|^2)^{-N} dx < \infty.$$

Define  $E = \{\alpha \in \mathbb{R}^N \mid 0 \leq \alpha_j < 1 \text{ for all } j\}$ . By complete additivity, we can rewrite the above finiteness condition as

$$\int_{\alpha \in E} \left[ \sum_{k \in \mathbb{Z}^N} |R(k + \alpha)|^{-\delta} (1 + |k + \alpha|^2)^{-N} \right] d\alpha < \infty.$$

Every pair  $(l, \beta)$  with  $l \in \mathbb{Z}$  and  $\beta \in [0, 1)$  has  $(l + \beta)^2 \leq 2(1 + l^2)$ . Summing  $N$  such inequalities gives  $|k + \alpha|^2 \leq 2N + 2|k|^2 \leq 2N(1 + |k|^2)$ . Thus we obtain  $1 + |k + \alpha|^2 \leq 3N(1 + |k|^2)$ ,  $(1 + |k + \alpha|^2)^{-N} \geq (3N)^{-N} (1 + |k|^2)^{-N}$ , and

$$\int_{\alpha \in E} \left[ \sum_{k \in \mathbb{Z}^N} |R(k + \alpha)|^{-\delta} (1 + |k|^2)^{-N} \right] d\alpha < \infty.$$

Therefore  $\sum_{k \in \mathbb{Z}^N} |R(k + \alpha)|^{-\delta} (1 + |k|^2)^{-N}$  is finite almost everywhere  $[d\alpha]$ . Fix an  $\alpha$  for which the sum is finite. If

$$\sum_{k \in \mathbb{Z}^N} |R(k + \alpha)|^{-\delta} (1 + |k|^2)^{-N} = K < \infty,$$

then  $|R(k + \alpha)|^{-\delta} (1 + |k|^2)^{-N} \leq K$  for all  $k \in \mathbb{Z}^N$  and hence  $|R(k + \alpha)|^{-1} \leq K^{1/\delta} (1 + |k|^2)^{N/\delta}$  for all  $k \in \mathbb{Z}^N$ . This proves Lemma 7.8.  $\square$

#### 4. Maximum Principle in the Elliptic Second-Order Case

In this section we work with a second-order linear homogeneous elliptic equation  $Lu = 0$  with continuous real-valued coefficients in a bounded connected open subset  $U$  of  $\mathbb{R}^N$ . It will be assumed that only derivatives of  $u$ , and not  $u$  itself, appear in the equation; in other words we assume that  $L(1) = 0$ . The conclusion will be that a real-valued  $C^2$  solution  $u$  cannot have an absolute maximum or an absolute minimum inside  $U$  without being constant. This result was proved already in Corollary 3.20 for the special case that  $L$  is the Laplacian  $\Delta$ .

Let us use notation for  $L$  of the kind in Proposition 7.5 and its proof. Then  $L$  is of the form

$$Lu = \sum_{i,j} b_{ij}(x) D_i D_j u + \sum_k c_k(x) D_k u$$

with the matrix  $[b_{ij}(x)]$  real-valued and symmetric. Ellipticity of  $L$  at  $x$  means that  $\sum_{i,j} b_{ij}(x) \xi_i \xi_j \neq 0$  for  $\xi \neq 0$ . Thus  $|\sum_{i,j} b_{ij}(x) \xi_i \xi_j|$  has a positive minimum value  $\mu(x)$  on the compact set where  $|\xi| = 1$ . By homogeneity of  $|\sum_{i,j} b_{ij}(x) \xi_i \xi_j|$  and  $|\xi|^2$ , we conclude that

$$\left| \sum_{i,j} b_{ij}(x) \xi_i \xi_j \right| \geq \mu(x) |\xi|^2$$

for some  $\mu(x) > 0$  and all  $\xi$ . The positive number  $\mu(x)$  is called the **modulus of ellipticity** of  $L$  at  $x$ .

**EXAMPLE.** Let  $L$  be the sum of the Laplacian and first-order terms, i.e.,  $Lu = \Delta u + \sum_k c_k(x) D_k u$ . Suppose that  $u$  is a real-valued  $C^2$  function on  $U$  and that  $u$  attains a local maximum at  $x_0$  in  $U$ . By calculus,  $D_i u(x_0) = 0$  for each  $i$  and  $D_i^2 u(x_0) \leq 0$ , so that  $Lu(x_0) \leq 0$ . Therefore if we know that  $Lu(x)$  is  $> 0$  everywhere in  $U$ , then  $u$  can have no local maximum in  $U$ . To obtain a maximum principle, we want to relax two conditions and still get the same conclusion. One is that we want to allow more general second-order terms in  $L$ , and the other is that we want to get a conclusion from knowing only that  $Lu(x)$  is  $\geq 0$  everywhere. The first step is carried out in Lemma 7.11 below, and the second step will be derived from the first essentially by perturbing the situation in a subtle way.

**Lemma 7.11.** Let  $L = \sum_{i,j} b_{ij}(x) D_i D_j + \sum_k c_k(x) D_k$ , with  $[b_{ij}(x)]$  symmetric, be a second-order linear elliptic operator with real-valued coefficients in an open subset  $U$  of  $\mathbb{R}^N$  such that for every  $x$  in  $U$ , there is a number  $\mu(x) > 0$  such that  $\sum_{i,j} b_{ij}(x) \xi_i \xi_j \geq \mu(x) |\xi|^2$  for all  $\xi \in \mathbb{R}^N$ . If  $u$  is a real-valued  $C^2$  function on  $U$  such that  $Lu > 0$  everywhere in  $U$ , then  $u$  has no local maximum in  $U$ .



PROOF. Suppose that  $u$  has a local maximum at  $x_0$ . Applying Proposition 7.5, we can find a nonsingular matrix  $M$  such that the definition  $D'_i = \sum_j M_{ij} D_j$  makes the second-order terms of  $L$  at  $x_0$  take the form  $\kappa_1 D_1'^2 + \cdots + \kappa_N D_N'^2$  with each  $\kappa_i$  equal to  $+1$ ,  $-1$ , or  $0$ . Examining the hypotheses of the lemma, we see that all  $\kappa_i$  must be  $+1$ . Hence the change of basis at  $x_0$  via  $M$  converts the second-order terms of  $L$  into the form  $D_1'^2 + \cdots + D_N'^2$ . The argument in the example above is applicable at  $x_0$ , and the lemma follows.  $\square$

**Theorem 7.12** (Hopf maximum principle). Let

$$L = \sum_{i,j} b_{ij}(x) D_i D_j + \sum_k c_k(x) D_k,$$

with  $[b_{ij}(x)]$  symmetric, be a second-order linear elliptic operator with real-valued continuous coefficients in a connected open subset  $U$  of  $\mathbb{R}^N$ . If  $u$  is a real-valued  $C^2$  function on  $U$  such that  $Lu = 0$  everywhere in  $U$ , then  $u$  cannot attain its maximum or minimum values in  $U$  without being constant.

PROOF. First we normalize matters suitably. We have  $|\sum_{i,j} b_{ij}(x) \xi_i \xi_j| \geq \mu(x) |\xi|^2$  with  $\mu(x) > 0$  at every point. By continuity of the coefficients and connectedness of  $U$ , the expression within the absolute value signs on the left side is everywhere positive or everywhere negative. Possibly replacing  $L$  by  $-L$ , we shall assume that it is everywhere positive:

$$\sum_{i,j} b_{ij}(x) \xi_i \xi_j \geq \mu(x) |\xi|^2 \quad \text{for all } x \in U.$$

Because of the continuity of the coefficients of  $L$ , the coefficient functions are bounded on any compact subset of  $U$  and the function  $\mu(x)$  is bounded below by a positive constant on any such compact set. Since  $u$  can always be replaced by  $-u$ , a result about absolute maxima is equivalent to a result about absolute minima. Thus we may suppose that  $u$  attains its absolute maximum value  $M$  at some  $x_1$  in  $U$ , and we are to prove that  $u$  is constant in  $U$ . Arguing by contradiction, suppose that  $x_0$  is a point in  $U$  with  $u(x_0) < M$ .

The idea of the proof is to use  $x_0$  and  $x_1$  to produce an open ball  $B$  with  $B^{\text{cl}} \subseteq U$  and a point  $s$  in the boundary  $\partial B$  of  $B$  such that  $u(s) = M$  and  $u(x) < M$  for all  $x$  in  $B^{\text{cl}} - \{s\}$ . See Figure 7.1. For a suitably small open ball  $B_1$  centered at  $s$ , we then produce a  $C^2$  function  $w$  on  $\mathbb{R}^N$  such that  $Lw > 0$  in  $B_1$  and  $w$  attains a local maximum at the center  $s$  of  $B_1$ . The existence of  $w$  contradicts Lemma 7.11, and thus the configuration with  $x_0$  and  $x_1$  could not have occurred.

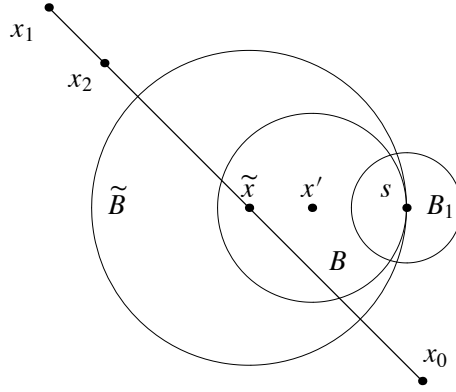


FIGURE 7.1. Construction in the proof of the Hopf maximum principle.

Since  $U$  is a connected open set in  $\mathbb{R}^N$ , it is pathwise connected. Let  $p : [0, 1] \rightarrow U$  be a path with  $p(0) = x_0$  and  $p(1) = x_1$ . Let  $\tau$  be the first value of  $t$  such that  $u(p(t)) = M$ ; necessarily  $0 < \tau \leq 1$ . Define  $x_2 = p(\tau)$ . Choose  $d > 0$  such that  $B(d; p(t))^{\text{cl}} \subseteq U$  for  $0 \leq t \leq \tau$ , and then fix a point  $\tilde{x} = p(t)$  with  $0 \leq t < \tau$  and with  $|\tilde{x} - x_2| < \frac{1}{2}d$ . By definition of  $d$ ,  $B(d; \tilde{x})^{\text{cl}} \subseteq U$ . Let  $\tilde{B}$  be the largest open ball contained in  $U$ , centered at  $\tilde{x}$ , and having  $u(x) < M$  for  $x \in \tilde{B}$ . Since  $u(x_2) = M$  and  $|\tilde{x} - x_2| < \frac{1}{2}d$ ,  $\tilde{B}$  has radius  $< \frac{1}{2}d$ . Thus  $\tilde{B}^{\text{cl}} \subseteq B(d; \tilde{x})^{\text{cl}} \subseteq U$ . The construction of  $\tilde{B}$  and the continuity of  $u$  force some point  $s$  of the boundary  $\partial\tilde{B}$  to have  $u(s) = M$ . Let  $B$  be any open ball properly contained in  $\tilde{B}$  and internally tangent to  $\tilde{B}$  at  $s$ . Then  $B^{\text{cl}} \subseteq \tilde{B} \cup \{s\}$ , and hence  $u(x) < M$  everywhere on  $B^{\text{cl}}$  except at  $s$ , where  $u(s) = M$ . Write  $B = B(R; x')$ .

To construct  $B_1$ , fix  $R_1 > 0$  with  $R_1 < \frac{1}{2}R$ , and let  $B_1 = B(R_1; s)$ . If  $x$  is in  $B_1^{\text{cl}}$ , then  $|x - \tilde{x}| \leq |x - s| + |s - \tilde{x}| \leq R_1 + \frac{1}{2}d < \frac{1}{2}R + \frac{1}{2}d \leq d$ , and hence  $B_1^{\text{cl}} \subseteq B(d; \tilde{x})^{\text{cl}} \subseteq U$ . Since  $B^{\text{cl}}$  and  $B_1^{\text{cl}}$  are compact subsets of  $U$ , the coefficients of  $L$  are bounded on  $B^{\text{cl}} \cup B_1^{\text{cl}}$ , and the ellipticity modulus is bounded below by a positive number. Let us say that

$$|b_{ij}(x)| \leq \beta, \quad |c_k(x)| \leq \gamma, \quad \mu(x) \geq \mu > 0 \quad \text{for } x \in B^{\text{cl}} \cup B_1^{\text{cl}}.$$

The next step is to construct an auxiliary function  $z(x)$  on  $\mathbb{R}^N$  to be used in the definition of  $w(x)$ . Let  $\alpha$  be a (large) positive number to be specified, and set

$$z(x) = e^{-\alpha|x-x'|^2} - e^{-\alpha R^2}.$$

The function  $z(x)$  is  $> 0$  on  $B$ , is  $0$  on  $\partial B$ , and is  $< 0$  off  $B^{\text{cl}}$ . Let us see that we can choose  $\alpha$  large enough to make  $L(z)(x) > 0$  for  $x$  in  $B_1$ . Performing the

differentiations explicitly, we obtain

$$\begin{aligned} L(z)(x) &= 2\alpha e^{-\alpha|x-x'|^2} \left( 2\alpha \sum_{i,j} b_{ij}(x)(x_i - x'_i)(x_j - x'_j) \right. \\ &\quad \left. - \sum_k (b_{kk}(x) - c_k(x)(x_k - x'_k)) \right) \\ &\geq 2\alpha e^{-\alpha|x-x'|^2} (2\alpha\mu|x-x'|^2 - (\beta + \gamma|x-x'|)). \end{aligned}$$

All points  $x$  in  $B_1$  have  $\frac{1}{2}R < |x - x'| < \frac{3}{2}R$  and therefore satisfy

$$L(z)(x) \geq 2\alpha e^{-\alpha|x-x'|^2} (2\alpha\mu\frac{1}{4}R^2 - (\beta + \frac{3}{2}\gamma R)).$$

Consequently we can choose  $\alpha$  large enough so that  $L(z)(x) > 0$  for  $x$  in  $B_1$ . Fix  $\alpha$  with this property.

Let  $\epsilon > 0$  be a (small) positive number to be specified, and define

$$w = u + \epsilon z.$$

For  $x$  in  $B_1$ , we have  $Lw = Lu + \epsilon Lz > 0$ . Also,

$$w(s) = u(s) + \epsilon z(s) = u(s) = M \quad \text{since } s \text{ is in } \partial B.$$

Let us see that we can choose  $\epsilon$  to make  $w(x) < M$  everywhere on  $\partial B_1$ . We consider  $\partial B_1$  in two pieces. Let  $C_0 = \partial B_1 \cap B^{\text{cl}}$ . Since  $C_0$  is a subset of  $B^{\text{cl}} - \{s\}$ ,  $u(x) < M$  at every point of  $C_0$ . By compactness of  $C_0$  and continuity of  $u$ , we must therefore have  $u(x) \leq M - \delta$  on  $C_0$  for some  $\delta > 0$ . Since the function  $z(x)$  is everywhere  $\leq 1 - e^{-\alpha R^2}$ , any  $x$  in  $C_0$  must have

$$w(x) = u(x) + \epsilon z(x) \leq M - \delta + \epsilon(1 - e^{-\alpha R^2}).$$

By taking  $\epsilon$  small enough, we can arrange that  $w(x) \leq M - \frac{1}{2}\delta$  on  $C_0$ . Fix such an  $\epsilon$ . The remaining part of  $\partial B_1$  is  $\partial B_1 - C_0$ . Each  $x$  in this set has

$$w(x) = u(x) + \epsilon z(x) \leq M + \epsilon z(x) < M.$$

Thus  $w(x) < M$  everywhere on  $\partial B_1$ , as asserted.

Since  $w(s) = M$  and  $w(x) < M$  everywhere on  $\partial B_1$ ,  $w$  attains its maximum in  $B_1^{\text{cl}}$  somewhere in the open set  $B_1$ . Since  $Lw > 0$  on  $B_1$ , we obtain a contradiction to Lemma 7.11. This completes the proof.  $\square$

### 5. Parametrixes for Elliptic Equations with Constant Coefficients

In this section we use distribution theory to derive some results about an elliptic equation  $P(D)u = f$  with constant coefficients. Initially we work on  $\mathbb{R}^N$ , yet in the end we will be able to work on any nonempty open set. We think of  $f$  as known and  $u$  as unknown. But we allow  $f$  to vary, so that we can see the effect on  $u$  of changing  $f$ . It will be important to be able to allow solutions that are not smooth functions, and thus  $u$  will be allowed to be some kind of distribution.

We begin by obtaining a parametrix, which at first will be a tempered distribution that approximately inverts  $P(D)$  on  $\mathcal{S}'(\mathbb{R}^N)$ . More specifically it inverts  $P(D)$  on  $\mathcal{S}'(\mathbb{R}^N)$  up to an error term given by an operator equal to convolution with a Schwartz function.

At this point we can use the version Theorem 7.4 of the Cauchy–Kovalevskaya Theorem to obtain a **fundamental solution**, i.e., a member  $u$  of  $\mathcal{D}'(\mathbb{R}^N)$  such that  $P(D)u = \delta$ . This step is carried out in Corollary 7.15 below. Convolution of  $P(D)u = \delta$  with a member  $f$  of  $\mathcal{E}'(\mathbb{R}^N)$  shows that Corollary 7.15 implies a global existence theorem: any elliptic equation  $P(D)u = f$  with  $f$  in  $\mathcal{E}'(\mathbb{R}^N)$  has a solution in  $\mathcal{D}'(\mathbb{R}^N)$ .

But it is not necessary, for purposes of examining regularity of solutions, to have an existence theorem. The next step is to modify the parametrix to have compact support. Once that has been done, the parametrix will invert  $P(D)$  on  $\mathcal{D}'(\mathbb{R}^N)$ , up to a smoothing term, and we will deduce a regularity theorem about solutions saying that the singular support of  $u$  is contained in the singular support of  $f$ . In particular, solutions of  $P(D)u = 0$  on  $\mathbb{R}^N$  are smooth. Finally we localize this result to see that the inclusion of singular supports persists even when the equation  $P(D)u = f$  is being considered only on an open set  $U$ .

The starting point for our investigation is the following lemma.

**Lemma 7.13.** If  $P(D)$  is an elliptic operator with constant coefficients, then the set of zeros of  $P(2\pi i\xi)$  in  $\mathbb{R}^N$  is compact.

REMARK. The polynomial  $P(2\pi i\xi)$  is the symbol of  $P(D)$ , as defined in Section 2. The important fact about the symbol is that the Fourier transform satisfies  $\mathcal{F}(P(D)T) = P(2\pi i\xi)\mathcal{F}(T)$ , which follows immediately from the formula  $\mathcal{F}(D^\alpha T) = (2\pi i)^{|\alpha|}\xi^\alpha \mathcal{F}(T)$ . This fact accounts for our studying the particular polynomial  $P(2\pi i\xi)$ .

PROOF. Let  $P$  have order  $m$ , and let  $Z$  be the set of zeros of  $P(2\pi i\xi)$  in  $\mathbb{R}^N$ . Since  $P(D)$  is elliptic, the principal symbol  $P_m(2\pi i\xi)$  is nowhere 0 on the unit sphere of  $\mathbb{R}^N$ . By compactness of the sphere,  $|P_m(2\pi i\xi)| \geq c > 0$  there, for some constant  $c$ . Taking into account the homogeneity of  $P_m$ , we see that  $|P_m(2\pi i\xi)| \geq c|\xi|^m$  for all  $\xi$  in  $\mathbb{R}^N$ . If we write  $P(2\pi i\xi) = P_m(2\pi i\xi) + Q(2\pi i\xi)$ , then

$Q(2\pi i\xi)| \leq C|\xi|^{m-1}$  for  $|\xi| \geq 1$  and for some constant  $C$ . If  $\xi$  is in  $Z$  and  $|\xi| \geq 1$ , then we have  $c|\xi|^m \leq P_m(2\pi i\xi) = |Q(2\pi i\xi)| \leq C|\xi|^{m-1}$ , and we conclude that  $|\xi| \leq C/c$ . This proves the lemma.  $\square$

Fix an elliptic operator  $P(D)$ , and choose  $R > 0$  by the lemma such that all the zeros in  $\mathbb{R}^N$  of  $P(2\pi i\xi)$  lie in the closed ball of radius  $R$  centered at the origin. Fix numbers  $R'$  and  $R''$  with  $R' > R'' > R$ . Let  $\chi$  be a smooth function on  $\mathbb{R}^N$  with values in  $[0, 1]$  such that  $\chi(\xi)$  is 0 when  $|\xi| \leq R''$  and is 1 when  $|\xi| \geq R'$ . The formal computation is as follows: if we define  $v$  in terms of  $f$  by

$$v(x) = \int_{\mathbb{R}^N} e^{2\pi i x \cdot \xi} \frac{\mathcal{F}(f)(\xi)}{P(2\pi i\xi)} \chi(\xi) d\xi,$$

then Fourier inversion gives

$$\begin{aligned} (P(D)v)(x) &= \int_{\mathbb{R}^N} e^{2\pi i x \cdot \xi} \mathcal{F}(f)(\xi) \chi(\xi) d\xi \\ &= f(x) + \int_{\mathbb{R}^N} e^{2\pi i x \cdot \xi} (\chi(\xi) - 1) \mathcal{F}(f)(\xi) d\xi, \end{aligned}$$

and the second term on the right side will be seen to be a smoothing term. Let us now state a precise result and use properties of distributions to make this computation rigorous.

**Theorem 7.14.** Let  $P(D)$  be an elliptic operator on  $\mathbb{R}^N$  with constant coefficients. Then there exist a distribution  $k \in \mathcal{S}'(\mathbb{R}^N)$  and a Schwartz function  $h \in \mathcal{F}^{-1}(C_{\text{com}}^\infty(\mathbb{R}^N))$  such that

$$P(D)k = \delta + T_h,$$

as an equality in  $\mathcal{S}'(\mathbb{R}^N)$ . Here  $\delta$  is the Dirac distribution  $\langle \delta, \varphi \rangle = \varphi(0)$ . Consequently whenever  $f$  is in  $\mathcal{E}'(\mathbb{R}^N)$ , then the distribution  $v = k * f$  is tempered and satisfies  $P(D)v = f + (h * f)$ .

**REMARKS.** The convolution operator  $f \mapsto k * f$  is called a **parametrix** for  $P(D)$  on  $\mathcal{E}'(\mathbb{R}^N)$ . More precisely it is a right parametrix, and a left parametrix can be defined similarly. The operator  $f \mapsto h * f$  is called a **smoothing operator** because  $h * f$  is in  $C^\infty(\mathbb{R}^N)$  whenever  $f$  is in  $\mathcal{E}'(\mathbb{R}^N)$ . To see the smoothing property, we observe that  $h$ , as a Schwartz function, is identified with a tempered distribution when we pass to  $T_h$ . Theorem 5.21 shows that  $T_h * f$  is a tempered distribution with Fourier transform  $\mathcal{F}(h)\mathcal{F}(f)$ . Both factors  $\mathcal{F}(h)$  and  $\mathcal{F}(f)$  are smooth functions, and  $\mathcal{F}(h)$  has compact support. Therefore  $\mathcal{F}(h * f)$  is smooth of compact support, and  $h * f$  is a Schwartz function.

PROOF. The function  $\sigma(\xi) = \chi(\xi)/P(2\pi i\xi)$  is smooth and is bounded on  $\mathbb{R}^N$  because, in the notation used in the proof of Lemma 7.13,  $|P(2\pi i\xi)| \geq |P_m(2\pi i\xi)| - |Q(2\pi i\xi)| \geq (c|\xi| - C)|\xi|^{m-1}$  and because  $(c|\xi| - C)|\xi|^{m-1} \geq 1$  as soon as  $|\xi|$  is large enough. Since  $\sigma$  is bounded, integration of the product of  $\sigma$  and any Schwartz function is meaningful, and  $T_\sigma$  is therefore in  $\mathcal{S}'(\mathbb{R}^N)$ . Define  $k = \mathcal{F}^{-1}(T_\sigma)$ . This is in  $\mathcal{S}'(\mathbb{R}^N)$  and has  $\mathcal{F}(k) = T_\sigma$ . Define  $h = \mathcal{F}^{-1}(\chi - 1)$ . Since  $\chi - 1$  is in  $C_{\text{com}}^\infty(\mathbb{R}^N)$ ,  $h$  is in  $\mathcal{S}(\mathbb{R}^N)$ .

Now let  $f$  in  $\mathcal{E}'(\mathbb{R}^N)$  be given, and define  $v = k * f$ . Theorem 5.21 shows that  $v$  is in  $\mathcal{S}'(\mathbb{R}^N)$  and that  $\mathcal{F}(v) = \mathcal{F}(k)\mathcal{F}(f) = \sigma\mathcal{F}(f)$ . Then

$$\begin{aligned} \mathcal{F}(P(D)v) &= P(2\pi i\xi)\mathcal{F}(v) = P(2\pi i\xi)\sigma(\xi)\mathcal{F}(f) \\ &= \chi(\xi)\mathcal{F}(f) = \mathcal{F}(f) + (\chi(\xi) - 1)\mathcal{F}(f) = \mathcal{F}(f) + \mathcal{F}(h)\mathcal{F}(f). \end{aligned}$$

Taking the inverse Fourier transform of both sides yields  $P(D)v = f + h * f$  as asserted. For the special case  $f = \delta$ , we have  $v = k * \delta = k$ , and then  $P(D)k = \delta + T_h$ . This completes the proof.  $\square$

The function  $h$  is the inverse Fourier transform of a member of  $C_{\text{com}}^\infty(\mathbb{R}^N)$ , specifically  $h(x) = \int_{\mathbb{R}^N} e^{2\pi i x \cdot \xi} (\chi(\xi) - 1) d\xi$ . Since the integration is really taking place on a compact set, we see that we can replace  $x$  by a complex variable  $z$  and obtain a holomorphic function in all of  $\mathbb{C}^N$ . In other words,  $h$  extends to a holomorphic function on  $\mathbb{C}^N$ . If we single out any variable, say  $x_1$ , then the ellipticity of  $P(D)$  implies that  $D_{x_1}^m$  has nonzero coefficient in  $P(D)$ , and  $P(D)w = h$  is therefore an equation to which the global Cauchy–Kovalevskaya Theorem applies in the form of Theorem 7.4. The theorem says that the equation  $P(D)w = h$ , in the presence of globally holomorphic Cauchy data, has not just a local holomorphic solution but a global holomorphic one. Therefore  $P(D)w = h$  has an entire holomorphic solution  $w$ . Let us regard  $w$  and  $h$  as yielding distributions  $T_w$  and  $T_h$  on  $C_{\text{com}}^\infty(\mathbb{R}^N)$ , so that the equation reads  $P(D)T_w = T_h$ . Subtracting this from  $P(D)k = \delta + T_h$  yields  $P(D)(k - T_w) = \delta$ . In summary we have the following corollary.

**Corollary 7.15.** If  $P(D)$  is an elliptic operator on  $\mathbb{R}^N$  with constant coefficients, then there exists  $e$  in  $\mathcal{D}'(\mathbb{R}^N)$  with  $P(D)e = \delta$ .

The distribution  $e$  is called a **fundamental solution** for  $P(D)$  in  $\mathcal{D}'(\mathbb{R}^N)$ . A consequence of the existence of  $e$  is that  $P(D)u = f$  has a solution  $u$  in  $\mathcal{D}'(\mathbb{R}^N)$  for each  $f$  in  $\mathcal{E}'(\mathbb{R}^N)$ . This represents an improvement in the conclusion (fundamental solution vs. parametrix) of Theorem 7.14.

Think of Corollary 7.15 as being an existence theorem. We now turn to a discussion of the regularity of solutions. For this we do not need the existence result, and thus we shall proceed without making further use of Corollary 7.15.

**Proposition 7.16.** Let  $P(D)$  be an elliptic operator on  $\mathbb{R}^N$  with constant coefficients. Then the tempered distribution  $k = \mathcal{F}^{-1}(T_\sigma)$ , where  $\sigma(\xi) = \chi(\xi)/P(2\pi i\xi)$ , is a smooth function on  $\mathbb{R}^N - \{0\}$ . Therefore, for any neighborhood of 0, the elliptic operator  $P(D)$  has a parametrix  $k_0 \in \mathcal{E}'(\mathbb{R}^N)$  with compact support in that neighborhood. In particular, there is a smooth function  $h_1$  with support in that neighborhood such that whenever  $f$  is in  $\mathcal{E}'(\mathbb{R}^N)$ , then the distribution  $v = k_0 * f$  is in  $\mathcal{E}'(\mathbb{R}^N)$  and satisfies  $P(D)v = f + (h_1 * f)$ .

SKETCH OF PROOF. One checks that

$$D^\beta(\xi^\alpha k) = (2\pi i)^{|\beta|}(-2\pi i)^{-|\alpha|}\mathcal{F}^{-1}(T_{\xi^\beta D^\alpha \sigma}).$$

Here  $\xi^\beta D^\alpha \sigma$  is a  $C^\infty$  function, and we are interested in its integrability. It is enough to consider what happens for  $|\xi| \geq R'$ , where  $\sigma(\xi) = 1/P(2\pi i\xi)$ . The function  $1/P(2\pi i\xi)$  is bounded above by a multiple of  $|\xi|^{-m}$ , and an inductive argument on the order of the derivative shows that  $|\xi^\beta D^\alpha \sigma| \leq C|\xi|^{|\beta|-|\alpha|-m}$  for  $|\xi| \geq R'$ , for a constant  $C$  independent of  $\xi$ .

Take  $\beta = 0$ . If  $|\alpha|$  is large enough, we see that  $D^\alpha \sigma$  is in  $L^1(\mathbb{R}^N)$ . Then  $\mathcal{F}^{-1}(D^\alpha \sigma) = (2\pi i)^{|\alpha|}\xi^\alpha k$  is given by the usual integral formula for  $\mathcal{F}$ , but with  $e^{-2\pi i x \cdot \xi}$  replaced by  $e^{2\pi i x \cdot \xi}$ . Therefore  $\xi^\alpha k$  is a bounded continuous function when  $|\alpha|$  is large enough. Applying this observation to  $(\sum_{j=1}^n |\xi_j|^{2l})k$  for large enough  $l$ , we find that  $k$  is a continuous function on  $\mathbb{R}^N - \{0\}$ .

Next take  $|\beta| = 1$  and increase  $l$  by 1, writing  $\alpha'$  for the new  $\alpha$ . Then  $\xi^\beta D^{\alpha'} \sigma$  is integrable, and it follows<sup>5</sup> that  $\xi^{\alpha'} k$  has a pointwise partial derivative of type  $\beta$  and is continuous. Thus the same thing is true of  $k$  on  $\mathbb{R}^N - \{0\}$ .

Iterating this argument by adding 1 to one of the entries of  $\beta$  to obtain  $\beta'$ , we find for each  $\beta$  that we consider, that the functions  $D^\beta(\sum_{j=1}^n |\xi_j|^{2l})k$  and  $D^{\beta'}(\sum_{j=1}^n |\xi_j|^{2l'})k$  are integrable for  $l'$  sufficiently large, and we deduce that  $D^\beta k$  has all first partial derivatives continuous. Since  $\beta'$  is arbitrary,  $k$  equals a smooth function on  $\mathbb{R}^N - \{0\}$ .

To finish the argument, let  $k$  and  $h$  be as in Theorem 7.14, and let  $\psi$  in  $C_{\text{com}}^\infty(\mathbb{R}^N)$  be identically 1 near 0 and have support in whatever neighborhood of 0 has been specified. If we write  $k = \psi k + (1 - \psi)k$ , then  $k_0 = \psi k$  has support in that same neighborhood, and  $T = (1 - \psi)k$  is of the form  $T_{h_0}$  for some smooth function  $h_0$ , by what we have shown. Substituting  $k = k_0 + T_{h_0}$  into  $P(D)k = \delta + T_h$ , we find that  $P(D)k_0 = \delta + T_h - T_{P(D)h_0}$ . The function  $h_1 = h - P(D)h_0$  is smooth, and it must have compact support since  $P(D)k_0$  and  $\delta$  have compact support.  $\square$

**Corollary 7.17.** If  $u$  is in  $\mathcal{D}'(\mathbb{R}^N)$  and  $P(D)$  is elliptic, then  $\text{sing supp } u \subseteq \text{sing supp } P(D)u$ , where “sing supp” denotes singular support.

<sup>5</sup>The precise result to use is Proposition 8.1f of *Basic*.

REMARK. At first glance it might seem that the rough spots of  $P(D)u$  are surely at least as bad as the rough spots of  $u$  for any  $D$ . But consider a function on  $\mathbb{R}^2$  of the form  $u(x, y) = g(y)$  and apply  $P(D) = \partial/\partial x$ . The result is 0, and thus  $\text{sing supp } u$  can properly contain  $\text{sing supp } P(D)u$  for  $P(D) = \partial/\partial x$ . The corollary says that this kind of thing does not happen if  $P(D)$  is elliptic.

PROOF. Let  $E = (\text{sing supp } P(D)u)^c$ . By definition the restriction of  $P(D)u$  to  $C_{\text{com}}^\infty(E)$  is of the form  $T_\psi$  with  $\psi$  in  $C^\infty(E)$ . Let  $U$  be any nonempty open set with  $U^{\text{cl}}$  compact and with  $U^{\text{cl}} \subseteq E$ . It is enough to exhibit a smooth function  $\eta$  equal to  $u$  on  $U$ . Choose an open set  $V$  with  $V^{\text{cl}}$  compact such that  $U^{\text{cl}} \subseteq V \subseteq V^{\text{cl}} \subseteq E$ . Multiply  $\psi$  by a smooth function of compact support in  $E$  that equals 1 on  $V^{\text{cl}}$ , obtaining a function  $\psi_0 \in C_{\text{com}}^\infty(E)$  such that  $\psi_0 = \psi$  on  $V$ .

Choose an open neighborhood  $W$  of 0 such that  $W = -W$  and such that the set of sums  $U^{\text{cl}} + W^{\text{cl}}$  is contained in  $V$ . Applying Proposition 7.16, we can write  $P(D)k_0 = \delta + h'$  with  $k_0 \in \mathcal{E}'(\mathbb{R}^N)$  and  $h' \in C_{\text{com}}^\infty(\mathbb{R}^N)$ . The proposition allows us to insist that the support of  $k_0^\vee$  be contained in  $W$ . Then also  $h'$  has support contained in  $W$ .

We are to produce  $\eta \in C^\infty(U)$  with  $\langle T_\eta, \varphi \rangle = \langle u, \varphi \rangle$  for all  $\varphi \in C_{\text{com}}^\infty(U)$ . Our choice of  $W$  forces  $k_0^\vee * \varphi$  to have support in  $V$ . Hence

$$\langle k_0 * P(D)u, \varphi \rangle = \langle P(D)u, k_0^\vee * \varphi \rangle = \langle T_\psi, k_0^\vee * \varphi \rangle = \langle T_{\psi_0}, k_0^\vee * \varphi \rangle = \langle k_0 * \psi_0, \varphi \rangle.$$

On the other hand, application of Corollary 5.14 gives

$$\langle k_0 * P(D)u, \varphi \rangle = \langle P(D)k_0 * u, \varphi \rangle = \langle (\delta + h') * u, \varphi \rangle = \langle u, \varphi \rangle + \langle h' * u, \varphi \rangle.$$

Combining the two computations, we see that  $\langle u, \varphi \rangle = \langle k_0 * \psi_0 - h' * u, \varphi \rangle$ , and the proof is complete if we take  $\eta$  to be  $k_0 * \psi_0 - h' * u$ .  $\square$

The final step is to localize the result of Corollary 7.17.

**Corollary 7.18.** If  $P(D)$  is elliptic with constant coefficients, if  $U$  is nonempty and open in  $\mathbb{R}^N$ , and if  $u$  and  $f$  are members of  $\mathcal{D}'(U)$  with  $P(D)u = f$ , then  $\text{sing supp } u \subseteq \text{sing supp } f$ . Consequently if  $f$  is a smooth function on  $U$ , then so is  $u$ .

REMARKS. For the Laplacian this result gives something beyond the results in Chapter III: Part of the statement is that *any* distribution solution  $u$  of  $\Delta u = 0$  on an open set  $U$  equals a smooth function on  $U$ . Previously the best result of this kind that we had was Corollary 3.17, which says that any distribution solution equal to a  $C^2$  function is a smooth function.



PROOF. It is enough to prove that  $E \cap \text{sing supp } u \subseteq E \cap \text{sing supp } f$  for each open set  $E$  with  $E^{\text{cl}}$  compact and  $E^{\text{cl}} \subseteq U$ . Choose  $\psi$  in  $C_{\text{com}}^\infty(U)$  with  $\psi$  equal to 1 on  $E^{\text{cl}}$ . The equality  $\langle \psi u, \varphi \rangle = \langle u, \psi \varphi \rangle = \langle u, \varphi \rangle$  for all  $\varphi \in C_{\text{com}}^\infty(E)$  shows that  $E \cap \text{sing supp } u = E \cap \text{sing supp } \psi u$ . Regard  $\psi u$  as in  $\mathcal{E}'(\mathbb{R}^N)$ , and define  $g = P(D)(\psi u)$ . Both  $\psi u$  and  $g$  are in  $\mathcal{E}'(\mathbb{R}^N)$ , and every  $\varphi \in C_{\text{com}}^\infty(E)$  satisfies

$$\begin{aligned} \langle g, \varphi \rangle &= \langle P(D)(\psi u), \varphi \rangle = \langle \psi u, P(D)^{\text{tr}} \varphi \rangle \\ &= \langle u, P(D)^{\text{tr}} \varphi \rangle = \langle P(D)u, \varphi \rangle = \langle f, \varphi \rangle. \end{aligned}$$

Hence  $E \cap \text{sing supp } g = E \cap \text{sing supp } f$ . Application of Corollary 7.17 therefore gives

$$E \cap \text{sing supp } u = E \cap \text{sing supp } \psi u \subseteq E \cap \text{sing supp } g = E \cap \text{sing supp } f,$$

and the result follows.  $\square$

## 6. Method of Pseudodifferential Operators

Linear elliptic equations with variable coefficients were already well understood by the end of the 1950s. The methods to analyze them combined compactness arguments for operators between Banach spaces with the use of Sobolev spaces and similar spaces of functions. Those methods were of limited utility for other kinds of linear partial equations, but some isolated methods had been developed to handle certain cases of special interest. In the 1960s a general theory of pseudodifferential operators was introduced to include all these methods under a single umbrella, and it and its generalizations are now a standard device for studying linear partial differential equations. They provide a tool for taking advantage of point-by-point knowledge of the zero locus of the principal symbol.

As with distributions, pseudodifferential operators make certain kinds of calculations quite natural, and many verifications lie behind their use. We shall omit most of this detail and concentrate on some of the ideas behind extending the theory of the previous section to variable-coefficient operators.

We start with a nonempty open subset  $U$  of  $\mathbb{R}^N$  and a linear differential operator  $P(x, D) = \sum_{|\alpha| \leq m} a_\alpha(x) D^\alpha$  whose coefficients  $a_\alpha(x)$  are in  $C^\infty(U)$ . If  $u$  is in  $C_{\text{com}}^\infty(U)$ , we can regard  $u$  as in  $C_{\text{com}}^\infty(\mathbb{R}^N)$ . The function  $u$  is then a Schwartz function, and the Fourier inversion formula holds:

$$u(x) = \int_{\mathbb{R}^N} e^{2\pi i x \cdot \xi} \widehat{u}(\xi) d\xi,$$

where  $\widehat{u}$  is the Fourier transform  $\widehat{u}(\xi) = \int_{\mathbb{R}^N} e^{-2\pi i x \cdot \xi} u(x) dx$ . Applying  $P$  gives

$$\begin{aligned} P(x, D)u(x) &= \sum_{|\alpha| \leq m} a_\alpha(x) (2\pi i)^{|\alpha|} \int_{\mathbb{R}^N} e^{2\pi i x \cdot \xi} \xi^\alpha \widehat{u}(\xi) d\xi \\ &= \int_{\mathbb{R}^N} e^{2\pi i x \cdot \xi} \left( \sum_{|\alpha| \leq m} a_\alpha(x) (2\pi i)^{|\alpha|} \xi^\alpha \right) \widehat{u}(\xi) d\xi = \int_{\mathbb{R}^N} e^{2\pi i x \cdot \xi} P(x, 2\pi i \xi) \widehat{u}(\xi) d\xi, \end{aligned}$$

where  $P(x, 2\pi i \xi)$  is the symbol. The basic idea of the theory is to enlarge the class of allowable symbols, thereby enlarging the class of operators under study, at least enough to include the parametrices and related operators of the previous section. The enlarged class will be the class of pseudodifferential operators.

In the constant-coefficient case, in which  $P(x, 2\pi i \xi)$  reduces to  $P(2\pi i \xi)$ , what we did in essence was to introduce an operator of the above kind, at first with  $1/P(2\pi i \xi)$  in the integrand in place of  $P(2\pi i \xi)$  but then with  $\chi(\xi)/P(2\pi i \xi)$  instead of  $1/P(2\pi i \xi)$  in the integrand in order to eliminate the singularities. When we composed the two operators, the result was the sum of the identity and a smoothing operator.

In the variable-coefficient case, the operator we use has to be more complicated. Suppose that we want  $P(x, D)G = 1 + \text{smoothing}$ , with  $G$  given by the same kind of formula as  $P(x, D)$  but with its symbol  $g(x, \xi)$  in some wider class. If the equation in question is  $P(x, D)u = f$ , then our computation above shows that we want to work with  $P(x, D) \left( \int_{\mathbb{R}^N} e^{2\pi i x \cdot \xi} g(x, \xi) \widehat{f}(\xi) d\xi \right)$ . The effect of putting  $P(x, D)$  under the integral sign is not achieved by including  $P(x, 2\pi i \xi)$  in the integrand, because the product  $e^{2\pi i x \cdot \xi} g(x, \xi)$  is being differentiated. A brief formal computation shows that  $D^\alpha (e^{2\pi i x \cdot \xi} g(x, \xi)) = e^{2\pi i x \cdot \xi} ((D_x + 2\pi i \xi)^\alpha g(x, \xi))$ , where the subscript  $x$  is included on  $D_x$  to emphasize that the differentiation is with respect to  $x$ . Thus we want  $P(x, D_x + 2\pi i \xi)g(x, \xi)$  to be close to identically 1, differing by the symbol of a "smoothing operator." We cannot simply divide by  $P(x, D_x + 2\pi i \xi)$  because of the presence of the  $D_x$ . What we can do is expand in terms of degrees of homogeneity in  $\xi$  and sort everything out. When degrees of homogeneity are counted,  $\xi^\alpha$  has degree  $|\alpha|$  while  $D_x$  has degree 0. Expansion of  $P$  gives

$$P(x, D_x + 2\pi i \xi) = P_m(x, 2\pi i \xi) + \sum_{j=0}^{m-1} p_j(x, \xi, D_x),$$

where  $P_m$  is the principal symbol and  $p_j$  is homogeneous in  $\xi$  of degree  $j$ . No  $D_x$  is present in  $P_m$  because degree  $m$  in  $\xi$  can occur only from terms  $(2\pi i \xi)^\alpha$  in  $(D_x + 2\pi i \xi)^\alpha$ . Since the constant function of  $\xi$  has homogeneity degree 0 and

since degrees of homogeneity add, let us look for an expansion of  $g(x, \xi)$  in the form

$$g(x, \xi) = \sum_{j=0}^{\infty} g_j(x, \xi),$$

with  $g_j$  homogeneous in  $\xi$  of degree  $-m - j$ . Expanding the product

$$(P_m(x, 2\pi i\xi) + \sum_{k=0}^{m-1} p_k(x, \xi, D_x)) \left( \sum_{j=0}^{\infty} g_j(x, \xi) \right) = 1$$

and collecting terms by degree of homogeneity, we read off equations

$$\begin{aligned} P_m(x, 2\pi i\xi)g_0(x, \xi) &= 1, \\ P_m(x, 2\pi i\xi)g_1(x, \xi) + p_{m-1}(x, \xi, D_x)g_0(x, \xi) &= 0, \\ P_m(x, 2\pi i\xi)g_2(x, \xi) + p_{m-1}(x, \xi, D_x)g_1(x, \xi) + p_{m-2}(x, \xi, D_x)g_0(x, \xi) &= 0, \end{aligned}$$

and so on. Dividing each equation by  $P_m(x, 2\pi i\xi)$ , we obtain recursive formulas for the  $g_j(x, \xi)$ 's, except for the problem that  $P_m(x, 2\pi i\xi)$  vanishes for  $\xi = 0$ . To handle this vanishing, we again have to introduce a function like  $\chi(\xi)$  by which to multiply  $g_j$ , and it turns out that in order to produce convergence,  $\chi$  has to be allowed to depend on  $j$ . After the  $g_j$ 's have been adjusted, we need to assemble an adjusted  $g$  from them and form a **right parametrix**, namely the pseudodifferential operator  $G$  corresponding to symbol  $g(x, \xi)$  such that  $P(x, D)G = 1 + R$ , where  $R$  is a "smoothing operator."

To make all this at all precise, we need to be more specific about a class of symbols, about the definition of the corresponding pseudodifferential operators, about the recognition of "smoothing operators," and about the assembly of the symbol from the sequence of homogeneous terms.

Fix a nonempty open set  $U$  in  $\mathbb{R}^N$ , and fix a real number  $m$ , not necessarily an integer. The symbol class known as  $S_{1,0}^m(U)$  and called the class of **standard symbols** of order  $m$  consists of the set of all functions  $g$  in  $C^\infty(U \times \mathbb{R}^N)$  such that for each compact set  $K \subseteq U$  and each pair of multi-indices  $\alpha$  and  $\beta$ , there exists a constant  $C_{K,\alpha,\beta}$  with<sup>6</sup>

$$|D_\xi^\alpha D_x^\beta g(x, \xi)| \leq C_{K,\alpha,\beta} (1 + |\xi|)^{m-|\alpha|} \quad \text{for } x \in K, \xi \in \mathbb{R}^N.$$

Then  $D_\xi^\alpha D_x^\beta g$  will be a symbol in the class  $S_{1,0}^{m-|\alpha|}(U)$ . Let  $S_{1,0}^{-\infty}(U)$  be the intersection of all  $S_{1,0}^{-n}(U)$  for  $n \geq 0$ .

<sup>6</sup>The symbol class  $S_{1,0}^m(U)$  is not the historically first class of symbols to have been studied, but it has come to be the usual one. Classes  $S_{\rho,\delta}^m(U)$  occur frequently as well, but we shall not discuss them.

EXAMPLES.

(1) If  $P(x, D) = \sum_{|\alpha| \leq m} a_\alpha D^\alpha$  with all  $a_\alpha$  in  $C^\infty(U)$ , then its symbol  $P(x, 2\pi i\xi) = \sum_{|\alpha| \leq m} a_\alpha(x)(2\pi i)^{|\alpha|} \xi^\alpha$  is in  $S_{1,0}^m(U)$ .

(2) If  $P(x, D)$  in Example 1 is elliptic, then the parametrix  $g(x, \xi)$  that we construct will be in  $S_{1,0}^{-m}(U)$ .

(3) With  $P$  and  $g$  formed as in Examples 1 and 2, the error term  $r(x, \xi)$  such that  $P(x, D_x + 2\pi i\xi)g(x, \xi) = 1 + r(x, \xi)$  will be in  $S_{1,0}^{-\infty}(U)$ . The corresponding pseudodifferential operator will be a “smoothing operator” in a sense to be defined below.

To a standard symbol  $g$ , we associate a **pseudodifferential operator**  $G = G(x, D)$  first on smooth functions and then on distributions.<sup>7</sup> The associated  $G : C_{\text{com}}^\infty(U) \rightarrow C^\infty(U)$  for a symbol  $g \in S_{1,0}^m(U)$  is given by

$$(G\varphi)(x) = \int_{\mathbb{R}^N} e^{2\pi i x \cdot \xi} g(x, \xi) \widehat{\varphi}(\xi) d\xi \quad \text{for } \varphi \in C_{\text{com}}^\infty(U), x \in U.$$

One readily checks that  $G\varphi$  is indeed in  $C^\infty(U)$  and that  $G : C_{\text{com}}^\infty(U) \rightarrow C^\infty(U)$  is continuous. The associated  $G : \mathcal{E}'(U) \rightarrow \mathcal{D}'(U)$  is given by<sup>8</sup>

$$\langle Gf, \varphi \rangle = \int_{\mathbb{R}^N} \left[ \int_U e^{2\pi i x \cdot \xi} g(x, \xi) \varphi(x) dx \right] \mathcal{F}(f)(\xi) d\xi \quad \text{for } f \in \mathcal{E}'(U).$$

(Recall that  $\mathcal{F}(f)$  is a smooth function, according to Theorem 5.20.) One readily checks that  $\langle Gf, \varphi \rangle$  is well defined, that  $Gf$  is in  $\mathcal{D}'(U)$ , and that when  $f = T_\psi$  for some  $\psi \in C_{\text{com}}^\infty(U)$ , then  $G(T_\psi) = T_{G\psi}$ .

The error term in constructing a parametrix is ultimately handled by the following fact: if  $g$  is a symbol in  $S_{1,0}^{-\infty}(U)$ , then  $G$  carries  $\mathcal{E}'(U)$  into  $C^\infty(U)$ . For this reason the pseudodifferential operators with symbol in  $S_{1,0}^{-\infty}(U)$  are called **smoothing operators**.

With the definitions made, let us return to the construction of a right parametrix for the elliptic differential operator  $P(x, D)$ . Let us write  $p_m(x, \xi, D_x)$  for the principal symbol  $P_m(x, 2\pi i\xi)$  in order to make the notation uniform. The

<sup>7</sup>Pseudodifferential operators can be used with other domains, such as Sobolev spaces, in order to obtain additional quantitative information. But we shall not pursue such lines of investigation here. Further comments about this matter occur in Section VIII.8.

<sup>8</sup>Our standard procedure for defining operations on distributions has consistently been to define the operation on smooth functions, to exhibit an explicit formula for the transpose operator on smooth functions and observe that the transpose is continuous, and to use the transpose operator to define the operator on distributions. This procedure avoids the introduction of topologies on spaces of distributions. In the present discussion of the operation of a pseudodifferential operator on distributions, we defer the introduction of transpose to Section VIII.6.

recursive computation given above produces expressions  $g_j(x, \xi)$  for  $j \geq 0$  such that

$$\left(\sum_{k=0}^m p_k(x, \xi, D_x)\right)\left(\sum_{j=0}^{\infty} g_j(x, \xi)\right) = 1$$

in a formal sense. The actual  $g_j(x, \xi)$ 's are not standard symbols because the formula for  $g_j(x, \xi)$  involves division by  $(p_m(x, \xi))^{j+1}$  and because  $p_m(x, \xi)$  vanishes at  $\xi = 0$ . However, the product  $\chi_j(\xi)g_j(x, \xi)$  is a standard symbol if  $\chi_j$  is a smooth function identically 0 near  $\xi = 0$  and identically 1 off some compact set. Thus we attempt to form the sum

$$g(x, \xi) = \sum_{j=0}^{\infty} \chi_j(\xi)g_j(x, \xi)$$

and use it as parametrix. Again we encounter a problem: we find that convergence is not automatic. More care is needed. What works is to define  $\chi_j(\xi) = \chi(R_j^{-1}|\xi|)$ , where  $\chi : \mathbb{R} \rightarrow [0, 1]$  is a smooth function that is 0 for  $|t| \leq \frac{1}{2}$  and is 1 for  $|t| \geq 1$ . One shows that positive numbers  $R_j$  tending to infinity can be constructed so that the partial sums in the series for  $g(x, \xi)$  converge in  $C^\infty(U \times \mathbb{R}^N)$  and the result is in the symbol class  $S_{1,0}^{-m}(U)$ . Let  $G$  be the pseudodifferential operator corresponding to  $g(x, \xi)$ .

A little computation shows that

$$P(x, D_x + \xi)g(x, \xi) = 1 + r(x, \xi),$$

where 
$$r(x, \xi) = -1 + \chi_0(\xi) - \sum_{j=1}^{\infty} r_j(x, \xi)$$

and 
$$r_j(x, \xi) = \sum_{k=1}^{\min\{j,m\}} [\chi_{j-k}(\xi) - \chi_j(\xi)]p_{m-k}(x, \xi, D_x)g_{j-k}(x, \xi).$$

The function  $r_j(x, \xi)$  is in  $C^\infty(U \times \mathbb{R}^N)$  and vanishes for  $|\xi| > R_j$ . This fact, the identities already established, and the construction of the numbers  $R_j$  allow one to see that  $\sum_{j=n+1}^{\infty} r_j(x, \xi)$  is in  $S_{1,0}^{-n}(U)$ . Since the remaining finite number of terms of  $r(x, \xi)$  have compact support in  $\xi$ , they too are in  $S_{1,0}^{-n}(U)$  and then so is  $r(x, \xi)$ . Since  $n$  is arbitrary,  $r(x, \xi)$  is in  $S_{1,0}^{-\infty}(U)$ . Hence the corresponding pseudodifferential operator is a smoothing operator. Consequently we obtain, as an identity on  $C_{\text{com}}^\infty(U)$  or on  $\mathcal{E}'(U)$ ,

$$P(x, D)G = 1 + R$$

with  $R$  a smoothing operator. Therefore  $G$  is a right parametrix for  $P(x, D)$ .

From the existence of a *right* parametrix, it can be shown that  $P(x, D)u = f$  is locally solvable.<sup>9</sup> If we could obtain a **left parametrix**, i.e., a pseudodifferential operator  $H$  with  $HP(x, D) = 1 + S$  for a smoothing operator  $S$ , then it would follow that singular supports satisfy

$$\text{sing supp } u = \text{sing supp } f \quad \text{whenever } f \text{ is in } \mathcal{E}'(U) \text{ and } P(x, D)u = f.$$

Inclusion in one direction follows from the local nature of  $P(x, D)$  in its action on  $u$ :  $\text{sing supp } f = \text{sing supp } P(x, D)u \subseteq \text{sing supp } u$ . Inclusion in the reverse direction uses the “pseudolocal” property of any pseudodifferential operator and of  $H$  in particular, namely that  $\text{sing supp } Hf \subseteq \text{sing supp } f$ . It goes as follows:

$$\begin{aligned} \text{sing supp } u &= \text{sing supp } (1 + S)u = \text{sing supp } HP(x, D)u \\ &= \text{sing supp } Hf \subseteq \text{sing supp } f. \end{aligned}$$

In particular, if  $f$  is in  $C_{\text{com}}^\infty(U)$ , then  $u$  is in  $C^\infty(U)$ . Constructing a left parametrix  $H$  with the techniques discussed so far is, however, more difficult than constructing the right parametrix  $G$  because we cannot so readily determine the symbol of  $HP(x, D)$  for a general pseudodifferential operator  $H$ .

Let us again work with the general theory, taking  $g$  to be in  $S_{1,0}^m(U)$  and denoting the corresponding pseudodifferential operator  $G : C_{\text{com}}^\infty(U) \rightarrow C^\infty(U)$  by

$$(G\varphi)(x) = \int_{\mathbb{R}^N} e^{2\pi i x \cdot \xi} g(x, \xi) \widehat{\varphi}(\xi) d\xi \quad \text{for } \varphi \in C_{\text{com}}^\infty(U).$$

The distribution  $T_{G\varphi}$ , which we write more simply as  $G\varphi$ , acts on a function  $\psi$  in  $C_{\text{com}}^\infty(U)$  by

$$\begin{aligned} \langle G\varphi, \psi \rangle &= \int_{\mathbb{R}^N} \int_U e^{2\pi i x \cdot \xi} g(x, \xi) \psi(x) \widehat{\varphi}(\xi) dx d\xi \\ &= \int_{\mathbb{R}^N} \int_U \int_U e^{2\pi i (x-y) \cdot \xi} g(x, \xi) \psi(x) \varphi(y) dy dx d\xi. \end{aligned}$$

If we think of  $\psi(x)\varphi(y)$  as a particular kind of function  $w(x, y)$  in  $C_{\text{com}}^\infty(U \times U)$ , then we can extend the above formula to define a linear functional  $\mathcal{G}$  on all of  $C_{\text{com}}^\infty(U \times U)$  by

$$\langle \mathcal{G}, w \rangle = \int_{\mathbb{R}^N} \left[ \int_{U \times U} e^{2\pi i (x-y) \cdot \xi} g(x, \xi) w(x, y) dx dy \right] d\xi.$$

It is readily verified that  $\mathcal{G}$  is continuous on  $C_{\text{com}}^\infty(U \times U)$  and hence lies in  $\mathcal{D}'(U \times U)$ . The expression written formally as

$$\mathcal{G}(x, y) = \int_{\mathbb{R}^N} e^{2\pi i (x-y) \cdot \xi} g(x, \xi) d\xi$$

is called the **distribution kernel** of the pseudodifferential operator  $G$ . This expression is not to be regarded as a function but as a distribution that is evaluated by the formula for  $\langle \mathcal{G}, w \rangle$  above.

The first serious general fact in the theory is as follows.

<sup>9</sup>More detail about this matter is included in Section VIII.8.

**Theorem 7.19.** If  $G$  is a pseudodifferential operator on an open set  $U$  in  $\mathbb{R}^N$ , then the distribution kernel  $\mathcal{G}$  of  $G$  is a smooth function off the diagonal of  $U \times U$ , and  $G$  is **pseudolocal** in the sense that

$$\text{sing supp } Gf \subseteq \text{sing supp } f \quad \text{for all } f \in \mathcal{E}'(U).$$

We give only a few comments about the proof, omitting all details. The first conclusion of the theorem is proved by using the known decrease of the derivatives of  $g(x, \xi)$ . For example, to see that  $\mathcal{G}$  is given by a continuous function, one uses the decrease of  $D_\xi^\alpha g(x, \xi)$  in the  $\xi$  variable to exhibit  $(x - y)^\alpha \mathcal{G}$ , for  $|\alpha| > m + N$ , as equal to a multiple of the continuous function  $\int_{\mathbb{R}^N} e^{2\pi i(x-y)\cdot\xi} D_\xi^\alpha g(x, \xi) d\xi$ . The second conclusion of the theorem, the pseudolocal property, can be derived as a consequence by using an approximate-identity argument.

To establish a general theory of pseudodifferential operators, the next step is to come to grips with the composition of two pseudodifferential operators. If we have two pseudodifferential operators  $G$  and  $H$  on the open set  $U$ , then each maps  $C_{\text{com}}^\infty(U)$  into  $C^\infty(U)$ , and their composition  $G \circ H$  need not be defined. But the composition is sometimes defined, as in the case that  $H$  is a differential operator and in the case that  $H$  is replaced by  $\psi(x)H$ , where  $\psi$  is a fixed member of  $C_{\text{com}}^\infty(U)$ . Thus let us for the moment ignore this problem concerning the image of  $H$  and make a formal calculation of the symbol of the composition anyway. Say that  $G = G(x, D)$  and  $H = H(x, D)$  are defined by the symbols  $g(x, \xi)$  and  $h(x, \xi)$ . Substituting from the definition of  $H(x, D)\varphi(x)$  and allowing any interchanges of limits that present themselves, we have

$$\begin{aligned} G(x, D)H(x, D)\varphi(x) &= G(x, D) \int_{\mathbb{R}^N} e^{2\pi i x \cdot \xi} h(x, \xi) \widehat{\varphi}(\xi) d\xi \\ &= \int_{\mathbb{R}^N} G(x, D_x) [e^{2\pi i x \cdot \xi} h(x, \xi)] \widehat{\varphi}(\xi) d\xi \\ &= \int_{\mathbb{R}^N} e^{2\pi i x \cdot \xi} (e^{-2\pi i x \cdot \xi} G(x, D_x) [e^{2\pi i x \cdot \xi} h(x, \xi)]) \widehat{\varphi}(\xi) d\xi. \end{aligned}$$

This formula suggests that the composition  $J = G \circ H$  ought to be a pseudodifferential operator with symbol

$$\begin{aligned} j(x, \xi) &= e^{-2\pi i x \cdot \xi} G(x, D_x) [e^{2\pi i x \cdot \xi} h(x, \xi)] \\ &= e^{-2\pi i x \cdot \xi} \int_{\mathbb{R}^N} e^{2\pi i x \cdot \eta} g(x, \eta) [e^{2\pi i x \cdot \xi} h(x, \xi)] \widehat{(\cdot)}(\eta) d\eta. \end{aligned}$$

Let us suppose that the Fourier transform of  $h(x, \xi)$  in the first variable is meaningful, as it is when  $h(\cdot, \xi)$  has compact support. Write  $\widehat{h}(\cdot, \xi)$  for this Fourier transform. Then the above expression is equal to

$$\int_{\mathbb{R}^N} e^{2\pi i x \cdot (\eta - \xi)} g(x, \eta) \widehat{h}(\eta - \xi, \xi) d\eta = \int_{\mathbb{R}^N} e^{2\pi i x \cdot \eta} g(x, \eta + \xi) \widehat{h}(\eta, \xi) d\eta.$$

If we form the infinite Taylor series expansion of  $g(x, \eta + \xi)$  about  $\eta = 0$  and assume that it converges, we have

$$g(x, \eta + \xi) = \sum_{\alpha} \frac{1}{\alpha!} D_{\xi}^{\alpha} g(x, \xi) \eta^{\alpha}.$$

Substituting and interchanging sum and integral, we can hope to get

$$\begin{aligned} j(x, \xi) &= \sum_{\alpha} \frac{1}{\alpha!} \int_{\mathbb{R}^N} e^{2\pi i x \cdot \eta} D_{\xi}^{\alpha} g(x, \xi) \eta^{\alpha} \widehat{h}(\eta, \xi) d\eta \\ &= \sum_{\alpha} \frac{(2\pi i)^{-|\alpha|}}{\alpha!} D_{\xi}^{\alpha} g(x, \xi) \int_{\mathbb{R}^N} e^{2\pi i x \cdot \eta} (D_x^{\alpha} h)^{\widehat{}}(\eta, \xi) d\eta. \end{aligned}$$

In view of the Fourier inversion formula, we might therefore expect to obtain

$$j(x, \xi) = \sum_{\alpha} \frac{(2\pi i)^{-|\alpha|}}{\alpha!} D_{\xi}^{\alpha} g(x, \xi) D_x^{\alpha} h(x, \xi).$$

We shall see that such a formula is meaningful, but in an asymptotic sense and not as an equality.

This discussion suggests four mathematical questions that we want to address:

- (i) If we are given a possibly divergent infinite series of symbols as on the right side of the formula for  $j(x, \xi)$  above, how can we extract a genuine symbol to represent the sum of the series?
- (ii) Put  $G(x, D_x + \xi)\varphi(x) = \int_{\mathbb{R}^N} e^{2\pi i x \cdot \eta} g(x, \eta + \xi) \widehat{\varphi}(\eta) d\eta$ . In what sense of  $\sim$  is it true that  $G(x, D_x + \xi)\varphi(x) \sim \sum_{\alpha} \frac{(2\pi i)^{-|\alpha|}}{\alpha!} D_{\xi}^{\alpha} g(x, \xi) D_x^{\alpha} \varphi(x)$ ?
- (iii) How can we handle the matter of compact support?
- (iv) How can we show, under suitable hypotheses that take (iii) into account, that  $j(x, \xi)$  is given by  $G(x, D_x + \xi)(h(x, \xi))$  and therefore that we obtain a formula from (ii) for  $j(x, \xi)$  involving  $\sim$ ?

The path that we shall follow is direct but not optimal. In Section VIII.6 we shall take note of an approach that is tidier and faster, but insufficiently motivated by the present considerations.

Question (i) is fully addressed by the following theorem.

**Theorem 7.20.** Suppose that  $\{m_j\}_{j \geq 0}$  is a sequence in  $\mathbb{R}$  decreasing to  $-\infty$ , and suppose for  $j \geq 0$  that  $g_j(x, \xi)$  is a symbol in  $S_{1,0}^{m_j}(U)$ . Then there exists a symbol  $g(x, \xi)$  in  $S_{1,0}^{m_0}(U)$  such that for all  $n \geq 0$ ,

$$g(x, \xi) - \sum_{j=0}^{n-1} g_j(x, \xi) \quad \text{is in } S_{1,0}^{m_n}(U).$$

The theorem is proved in the same way that we constructed a right parametrix for an elliptic differential operator earlier in this section. We can now give a



precise meaning to  $\sim$  in terms of a notion of an asymptotic series. If  $\{m_j\}_{j \geq 0}$  is a sequence in  $\mathbb{R}$  decreasing to  $-\infty$ , if  $g(x, \xi)$  is a symbol in  $S_{1,0}^{m_0}(U)$ , and if  $g_j(x, \xi)$  is a symbol in  $S_{1,0}^{m_j}(U)$  for each  $j \geq 0$ , then we write

$$g(x, \xi) \sim \sum_{j=0}^{\infty} g_j(x, \xi)$$

if for all  $n \geq 0$ ,

$$g(x, \xi) - \sum_{j=0}^{n-1} g_j(x, \xi) \text{ is in } S_{1,0}^{m_n}(U).$$

If the given sequence  $\{m_j\}_{j \geq 0}$  is a finite sequence ending with  $m_r$ , we can extend it to an infinite sequence with  $g_j(x, \xi) = 0$  for  $j > r$ , and in this case the definition of  $\sim$  is to be interpreted to mean that  $g(x, \xi) - \sum_{j=0}^r g_j(x, \xi)$  is the symbol of a smoothing operator.

For (ii), we have just attached a meaning to  $\sim$ . We define  $G(x, D_x + \xi)\varphi(x) = \int_{\mathbb{R}^N} e^{2\pi i x \cdot \eta} g(x, \eta + \xi) \widehat{\varphi}(\eta) d\eta$ . The precise statement that is proved to yield the asymptotic expansion of (ii) is the following.

**Proposition 7.21.** Let  $U$  be open in  $\mathbb{R}^N$ , fix  $g$  in  $S_{1,0}^m(U)$ , and let  $K$  be a compact subset of  $U$ . Then for any nonnegative integers  $M$  and  $R$  such that  $R > m + N$ , there exists a constant  $C$  such that

$$\begin{aligned} & \left| G(x, D_x + \xi)\varphi(x) - \sum_{|\alpha| < n} \frac{(2\pi i)^{-|\alpha|}}{\alpha!} D_\xi^\alpha g(x, \xi) D_x^\alpha \varphi(x) \right| \\ & \leq C \left\{ (1 + |\xi|^m) \int_{|\xi + \eta| \leq |\xi|/2} |\widehat{\varphi}(\eta)| d\eta \right. \\ & \quad \left. + \sum_{|\alpha|=N} |\xi|^{m-R} \sup_y [|D^\alpha \varphi(y)| (1 + |\xi||x - y|)^{-M}] \right\} \end{aligned}$$

for all  $\varphi$  in  $C_K^\infty$ , all  $x$  in  $K$ , and all  $\xi$  with  $|\xi| \geq 1$ .

We shall not make further explicit use of this proposition. The proof of the result is long, and we omit any discussion of it.

We turn to questions (iii) and (iv). Question (iii) is addressed by a definition and some remarks concerning it, and question (iv) is addressed by the theorem that comes after those remarks. Continuing with our pseudodifferential operator  $G$  on the open set  $U$ , we say that  $G$  is **properly supported** if the subset  $\text{support}(\mathcal{G})$  of  $U \times U$  has compact intersection with  $K \times U$  and with  $U \times K$  for every compact subset  $K$  of  $U$ . See Figure 7.2.

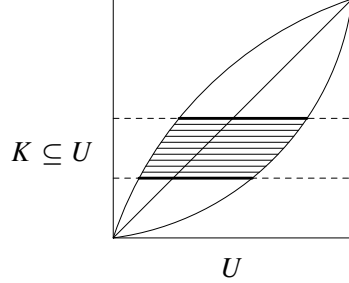


FIGURE 7.2. Nature of the support of the distribution kernel of a properly supported pseudodifferential operator. The open set  $U$  in this case is an open interval, and the oval-shaped region represents  $\text{support}(\mathcal{G})$ . The shaded region is an example of a set  $(U \times K) \cap \text{support}(\mathcal{G})$ .

Suppose that  $G$  is properly supported,  $K$  is compact in  $U$ , and  $\varphi$  is in  $C_{\text{com}}^{\infty}(U)$  with support contained in  $K$ . Introduce projections  $p_1(x, y) = x$  and  $p_2(x, y) = y$ . Define  $L = p_1((U \times K) \cap \text{support}(\mathcal{G}))$ ; the set  $L$  is compact since  $G$  is properly supported and since the continuous image of a compact set is compact. Let us see that  $G\varphi$  has support contained in  $L$ . To do so, we write  $\psi \otimes \varphi$  for the function  $(x, y) \mapsto \psi(x)\varphi(y)$ , and then we have

$$\langle G\varphi, \psi \rangle = \int_{\mathbb{R}^N} \int_U \int_U e^{2\pi i(x-y)\cdot \xi} g(x, \xi) \psi(x) \varphi(y) dy dx d\xi = \langle \mathcal{G}, \psi \otimes \varphi \rangle.$$

If  $\psi$  is in  $C_{\text{com}}^{\infty}(L^c \cap U)$ , then  $F = p_1^{-1}(\text{support } \psi) \cap p_2^{-1}(\text{support } \varphi)$  is the compact support of  $\psi \otimes \varphi$ , and

$$F \cap \text{support}(\mathcal{G}) \subseteq p_1^{-1}(L^c) \cap (U \times K) \cap \text{support}(\mathcal{G}) = p_1^{-1}(L^c) \cap p_1^{-1}(L) = \emptyset.$$

Thus  $\langle \mathcal{G}, \psi \otimes \varphi \rangle = 0$ ,  $\langle G\varphi, \psi \rangle = 0$ , and  $G\varphi$  is supported in  $L$ .

Thus the properly supported pseudodifferential operator  $G$  carries  $C_{\text{com}}^{\infty}(U)$  into itself, and Lemma 5.2 shows that it does so continuously. Then  $G$  is continuous also as a mapping of the dense vector subspace  $C_{\text{com}}^{\infty}(U)$  of  $C^{\infty}(U)$  into  $C^{\infty}(U)$ . Because of the completeness of  $C^{\infty}(U)$ ,  $G$  extends to a continuous map of  $C^{\infty}(U)$  into itself.

Similarly one checks that any properly supported pseudodifferential operator carries  $\mathcal{E}'(U)$  into  $\mathcal{E}'(U)$ . Therefore the composition  $G \circ H$  of two pseudodifferential operators, whether regarded as acting on  $C_{\text{com}}^{\infty}(U)$  or as acting on  $\mathcal{E}'(U)$ , is well defined if  $H$  is properly supported.

**Theorem 7.22.** Let  $U$  be an open subset of  $\mathbb{R}^N$ .

(a) If  $G$  is a pseudodifferential operator on  $U$ , then there exists a properly supported pseudodifferential operator  $G^\#$  on  $U$  such that  $G - G^\#$  is in  $S_{1,0}^{-\infty}(U)$ , hence such that  $G - G^\#$  is a smoothing operator.

(b) If  $G$  and  $H$  are properly supported pseudodifferential operators on  $U$  with symbols  $g$  in  $S_{1,0}^m(U)$  and  $h$  in  $S_{1,0}^{m'}(U)$ , then  $G \circ H$  is a properly supported pseudodifferential operator with symbol  $j$  in  $S_{1,0}^{m+m'}(U)$ , and

$$j(x, \xi) \sim \sum_{\alpha} \frac{(2\pi i)^{-|\alpha|}}{\alpha!} D_{\xi}^{\alpha} g(x, \xi) D_x^{\alpha} h(x, \xi).$$

All that is needed from (b) in many cases is the following weaker statement.

**Corollary 7.23.** Let  $U$  be an open subset of  $\mathbb{R}^N$ . If  $G$  and  $H$  are properly supported pseudodifferential operators on  $U$  with symbols  $g$  in  $S_{1,0}^m(U)$  and  $h$  in  $S_{1,0}^{m'}(U)$ , then  $G \circ H$  is a properly supported pseudodifferential operator whose symbol  $j(x, \xi)$  is in  $S_{1,0}^{m+m'}(U)$  and has the property that

$$j(x, \xi) - g(x, \xi)h(x, \xi)$$

is a symbol in  $S_{1,0}^{m+m'-1}(U)$ .

This is enough of the general theory so that we can see how to prove a theorem with consequences beyond the subject of pseudodifferential operators. A pseudodifferential operator  $G$  on  $U$  with symbol  $g(x, \xi)$  in  $S_{0,1}^m(U)$  is said to be **elliptic of order  $m$**  if for each compact subset  $K$  of  $U$ , there are constants  $C_K$  and  $M_K$  such that

$$|g(x, \xi)| \geq C_K (1 + |\xi|)^m \quad \text{for } x \in K \text{ and } |\xi| \geq M_K.$$

In particular, an elliptic differential operator of order  $m$  satisfies this condition. A (two-sided) **parametrix**  $H$  for a properly supported pseudodifferential operator  $G$  with symbol  $g \in S_{1,0}^m(U)$  is a properly supported pseudodifferential operator  $H$  of order  $-m$  such that  $H \circ G = 1 + \text{smoothing}$  and  $G \circ H = 1 + \text{smoothing}$ .

**Theorem 7.24.** If  $G$  is a properly supported elliptic pseudodifferential operator of order  $m$ , then  $G$  has a parametrix  $H$ .

REMARKS. We saw in Theorem 7.19 that  $\text{sing supp } Gf \subseteq \text{sing supp } f$  for  $f$  in  $\mathcal{E}'(U)$ . The same argument as with the left parametrix before that theorem shows now from the parametrix of Theorem 7.24 that  $\text{sing supp } Gf \supseteq \text{sing supp } f$  and therefore that  $\text{sing supp } Gf = \text{sing supp } f$  for  $f$  in  $\mathcal{E}'(U)$ . In particular, solutions of elliptic equations are smooth wherever the given data are smooth.

PARTIAL PROOF. Let  $\rho : U \times \mathbb{R}^n \rightarrow [0, 1]$  be a smooth function with the properties that

- (i)  $\rho$  equals 1 in a neighborhood of each point  $(x, \xi)$  where  $g(x, \xi) = 0$ ,
- (ii) for each compact subset  $K$  of  $U$ , there is a constant  $T_K$  such that  $\rho(x, \xi) = 0$  for  $x$  in  $K$  and  $|\xi| \geq T_K$ .

We omit the verification that  $\rho$  exists and is the symbol of a smoothing operator. Put

$$h_0(x, \xi) = (1 - \rho(x, \xi))g(x, \xi)^{-1}.$$

This is a smooth function by (i), and we omit the step of checking that  $h_0$  is in  $S_{1,0}^{-m}(U)$ . Let  $H_0$  be the pseudodifferential operator with symbol  $h_0$ . Apply Theorem 7.22a to find a properly supported  $H_0^\#$  whose symbol  $h_0^\#$  has  $h_0^\# \sim h_0$ . We write  $h_0^\# = h_0 + r_0$  with  $r_0$  in  $S_{1,0}^{-\infty}(U)$ .

Corollary 7.23 shows that  $H_0^\#G$  is a well-defined properly supported operator whose symbol  $j_0(x, \xi)$  is in  $S_{1,0}^0(U)$  and has the property that  $j_0 - h_0^\#g$  is in  $S_{1,0}^{-1}(U)$ . Since

$$j_0 - h_0^\#g = j_0 - (h_0 + r_0)g = j_0 - [(1 - \rho)g^{-1} + r_0]g = j_0 - 1 + \rho - r_0g$$

and since  $\rho$  and  $r_0g$  are the symbols of smoothing operators,  $j_0 - 1$  must be in  $S_{1,0}^{-1}(U)$ . Therefore  $H_0^\#G = 1 + R$  for a pseudodifferential operator  $R$  whose symbol  $r$  is in  $S_{1,0}^{-1}(U)$ .

The equality  $H_0^\#G = 1 + R$  shows that  $R$  is properly supported. By Corollary 7.23,  $R^k$  is a properly supported pseudodifferential operator for all integers  $k \geq 1$ , and its symbol  $r_k$  is in  $S_{1,0}^{-k}(U)$ . We form the asymptotic series

$$1 - r_1 + r_2 - r_3 + \cdots$$

and use Theorems 7.20 and 7.22a to obtain a properly supported pseudodifferential operator  $E$  whose symbol is in  $S_{1,0}^0(U)$  and has

$$e \sim 1 - r_1 + r_2 - r_3 + \cdots. \quad (*)$$

For any integer  $n \geq 1$ , we have

$$\begin{aligned} & (1 - R + R^2 - R^3 + \cdots \pm R^{n-1})H_0^\#G \\ &= (1 - R + R^2 - R^3 + \cdots \pm R^{n-1})(1 + R) = 1 \mp R^n. \quad (**) \end{aligned}$$

Because of (\*),  $E - (1 - R + R^2 - R^3 + \cdots \pm R^{n-1})$  has symbol in  $S_{1,0}^{-n}(U)$ . Since the symbol  $j_0$  of  $H_0^\#G$  is in  $S_{1,0}^0(U)$ , the product

$$(E - (1 - R + R^2 - R^3 + \cdots \pm R^{n-1}))H_0^\#G \quad \text{has symbol in } S_{1,0}^{-n}(U).$$

Also, (\*\*) implies that

$$(1 - R + R^2 - R^3 + \cdots \pm R^{n-1})H_0^\#G - 1 = \mp R^n \quad \text{has symbol in } S_{1,0}^{-n}(U).$$

Adding shows that

$$EH_0^\#G - 1 \quad \text{has symbol in } S_{1,0}^{-n}(U).$$

Since  $n$  is arbitrary,  $EH_0^\#G - 1$  is a smoothing operator. Thus  $H = EH_0^\#$  is a left parametrix for  $G$ .

In similar fashion we can use the assumption “properly supported” to obtain a right parametrix  $\tilde{H}$  for  $G$ . We omit the details. The operators  $H$  and  $\tilde{H}$  give us equations

$$HG = 1 + S \quad \text{and} \quad G\tilde{H} = 1 + \tilde{S}$$

for suitable properly supported smoothing operators  $S$  and  $\tilde{S}$ . Computing the product  $HG\tilde{H}$  in two ways shows that

$$HG\tilde{H} = (1 + S)\tilde{H} = \tilde{H} + S\tilde{H} = \tilde{H} + \text{smoothing}$$

$$\text{and} \quad HG\tilde{H} = H(1 + \tilde{S}) = H + H\tilde{S} = H + \text{smoothing}.$$

Hence  $H = \tilde{H} + S_0$  with  $S_0$  properly supported smoothing. Consequently

$$GH = G\tilde{H} + GS_0 = 1 + \tilde{S} + GS_0 = 1 + \text{smoothing},$$

and the left parametrix  $H$  is also a right parametrix.  $\square$

**BIBLIOGRAPHICAL REMARKS.** The proof of Theorem 7.7 is adapted from Taylor’s *Pseudodifferential Operators*, and the proof of Theorem 7.12 is taken from the book by Bers, John, and Schechter. The approach to pseudodifferential operators used in Section 6 is now considered outdated, and a more streamlined approach requiring additional motivation appears in Section VIII.6.

## 7. Problems

1. Suppose that  $P(x, D) = \sum_{|\alpha| \leq m} a_\alpha(x) D^\alpha$  with each  $a_\alpha$  in  $C^\infty(\Omega)$ . Prove that if  $P(x, D)u = 0$  for all functions  $u \in C^m(\Omega)$ , then all the coefficients  $a_\alpha$  are 0.
2. **(Harmonic measure)** Let  $\Omega$  be a bounded nonempty connected open subset of  $\mathbb{R}^N$ , let  $\partial\Omega$  be its boundary  $\partial\Omega = \Omega^{\text{cl}} - \Omega$ , and let  $L$  be an elliptic linear differential operator on  $\Omega$  of the form  $L(u) = \sum_{i,j} b_{ij}(x) D_i D_j u + \sum_k c_k(x) D_k u$  with real-valued coefficients of class  $C^2$  such that  $b_{ij}(x) = b_{ji}(x)$  for all  $i$  and  $j$ . Let  $S$  be the vector subspace of real-valued continuous functions  $u$  on  $\Omega^{\text{cl}}$  such that  $Lu(x) = 0$  for all  $x \in \Omega$ . Prove for each point  $p$  in  $\Omega$  that there exists a Borel measure  $\mu_p$  on  $\partial\Omega$  with  $\mu_p(\partial\Omega) = 1$  such that  $u(p) = \int_{\partial\Omega} u(x) d\mu_p(x)$  for all  $u$  in  $S$ .

3. This problem identifies a fundamental solution of the Cauchy–Riemann operator in  $\mathbb{R}^2$ . It makes use of Green’s Theorem, which relates line integrals in  $\mathbb{R}^2$  with double integrals, for an annulus centered at the origin.
- For  $\varphi$  in  $C_{\text{com}}^\infty(\mathbb{R}^2)$ , let  $P(x, y) = \frac{x\varphi(x, y)}{x^2+y^2}$  and  $Q(x, y) = \frac{y\varphi(x, y)}{x^2+y^2}$ . Prove that  $\lim_{\varepsilon \downarrow 0} \oint_{|(x, y)|=\varepsilon} (P dx + Q dy) = 0$ .
  - With  $P$  and  $Q$  as in (a), verify that  $\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = \frac{y\varphi_x - x\varphi_y}{x^2+y^2}$ .
  - Conclude from (a) and (b) that  $\iint_{\mathbb{R}^2} \frac{y\varphi_x - x\varphi_y}{x^2+y^2} dx dy = 0$ .
  - Repeat (a) with  $P(x, y) = -\frac{y\varphi(x, y)}{x^2+y^2}$  and  $Q(x, y) = \frac{x\varphi(x, y)}{x^2+y^2}$ , showing that  $\lim_{\varepsilon \downarrow 0} \oint_{|(x, y)|=\varepsilon} (P dx + Q dy) = 2\pi\varphi(0, 0)$  if the line integral is taken counterclockwise around the circle.
  - With  $P$  and  $Q$  as in (d), verify that  $\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = \frac{x\varphi_x + y\varphi_y}{x^2+y^2}$ .
  - Conclude from (d) and (e) that  $\iint_{\mathbb{R}^2} \frac{x\varphi_x + y\varphi_y}{x^2+y^2} = -2\pi\varphi(0, 0)$ .
  - Conclude from (c) and (f) that  $\frac{1}{2\pi} \iint_{\mathbb{R}^2} \frac{1}{z} \frac{\partial \varphi}{\partial \bar{z}} dx dy = -\varphi(0, 0)$ .
  - Let  $T$  be the locally integrable function  $1/(2\pi z)$ , regarded as a member of  $\mathcal{D}'(\mathbb{R}^2)$ . Prove that  $\frac{\partial}{\partial \bar{z}}(T) = \delta$ .
4. On  $\mathbb{R}^1$ , the **Heaviside distribution**  $H$  is the distribution given by the **Heaviside function**  $H(x)$  equal to 1 for  $x \geq 0$  and to 0 for  $x < 0$ .
- Prove that  $D_x H = \delta$ , so that  $H$  is a fundamental solution for the elliptic operator  $D_x$  on  $\mathbb{R}^1$ .
  - Show that the function  $f(x) = \max\{x, 0\}$  on  $\Omega = (-1, 1)$  has the Heaviside function as weak derivative on  $\Omega$  and that  $f$  is in  $L_1^p(\Omega)$  for every  $p$  with  $1 \leq p < \infty$ .
  - Does the restriction of the Heaviside function to  $\Omega = (-1, 1)$  have a weak derivative on  $\Omega$ ? Why or why not?
  - Show that the distribution  $H \times \delta$  on  $\mathbb{R}^2$  given by  $\langle H \times \delta, \varphi \rangle = \int_0^\infty \varphi(x, 0) dx$  for  $\varphi \in C_{\text{com}}^\infty(\mathbb{R}^2)$  is a fundamental solution of the operator  $D_x$  on  $\mathbb{R}^2$ .
  - Find the support and the singular support of the distribution  $H$  on  $\mathbb{R}^1$  and of the distribution  $H \times \delta$  on  $\mathbb{R}^2$ .
5. Let  $U$  be an open set in  $\mathbb{R}^N$  containing 0, let  $f$  be in  $\mathcal{E}'(U)$ , and let  $P(D)$  be a linear differential operator with constant coefficients and with order  $\geq 1$ . By taking into account the theory of periodic distributions in Problems 12–13 of Chapter V and by suitably adapting the proof that Lemma 7.8 implies Theorem 7.7, prove that the equation  $P(D)u = f$  has a distribution solution in some neighborhood of 0.

Problems 6–9 prove the global version of the Cauchy–Kovalevskaya Theorem given as Theorem 7.2 for the linear constant-coefficient case. The result is an ingredient used in deriving Corollary 7.15 from Theorem 7.14. For the statement the domain variables are  $t$  and  $x$  with  $x = (x_1, \dots, x_N)$ , and the unknown functions are the  $p$

components of a function  $u(t, x)$  with values in  $\mathbb{C}^p$ . Write  $D_t$  for  $\partial/\partial t$  and  $D_j$  for  $\partial/\partial x_j$ . The Cauchy problem in question is

$$D_t u = \sum_{j=1}^N A_j D_j u + Bu + F(t, x),$$

$$u(0, x) = g(x),$$

where  $A_j$  and  $B$  are  $p$ -by- $p$  matrices of complex constants,  $F$  is an entire holomorphic function from  $\mathbb{C}^{N+1}$  to  $\mathbb{C}^p$ , and  $g$  is an entire holomorphic function from  $\mathbb{C}^N$  to  $\mathbb{C}^p$ . The conclusion is that the unique formal power-series solution of the Cauchy problem converges and defines an entire holomorphic function from  $\mathbb{C}^{N+1}$  to  $\mathbb{C}^p$  that solves the problem. For a vector  $v = (v_1, \dots, v_p)$  in  $\mathbb{C}^p$ , let  $\|v\|_\infty = \max\{|v_1|, \dots, |v_p|\}$ .

6. Let  $\alpha$  denote a multi-index  $\alpha = (\alpha_1, \dots, \alpha_N)$  of integers  $\geq 0$ . Prove that

$$|\alpha|! \leq (|\alpha|)!, \text{ that } \sum_{|\alpha|=l} \frac{1}{\alpha!} = \frac{N^l}{l!}, \text{ and that } \sum_{l=0}^{\infty} \binom{q+l}{l} z^l = (1-z)^{-q-1} \text{ if } |z| < 1.$$

7. Show that iterated substitution into the system  $D_t u = \sum_{j=1}^N A_j D_j u + Bu + F$  leads to an expression for  $D_t^m u$  as the sum of two kinds of terms: For one kind, there are  $2^m$  terms of the form  $\sum T_1 \cdots T_m D_x^\alpha u$  with each  $T_i$  equal to an  $A_{j_i}$  or to  $B$ , with  $D^\alpha$  equal to the product of the  $D_{j_i}$  for which  $T_i = A_{j_i}$ , and with the sum taken over  $j_i$  from 1 to  $N$ . For the other kind, there are  $\sum_{s=0}^{m-1} 2^s = 2^m - 1$  terms with something operating on  $F$ , the terms corresponding to  $s$  being the ones  $\sum T_1 \cdots T_s D_x^\alpha D_t^{m-1-s} F$  with each  $T_i$ , the  $D^\alpha$ , and the sum all as above.

8. (a) How does one compute  $D_x^\beta D_t^m u(0, 0)$  from the expression in the previous problem?

(b) Why is it enough to prove, for any given  $r > 0$ , that the values  $D_x^\beta D_t^m u(0, 0)$  satisfy  $\sum_{m \geq 0} \sum_{\beta} (\beta! m!)^{-1} \|D_x^\beta D_t^m u(0, 0)\|_\infty r^{|\beta|+m} < \infty$ ?

9. Choose a constant  $M \geq 1$  with  $\|Bv\|_\infty \leq M\|v\|_\infty$  and  $\|A_j v\|_\infty \leq M\|v\|_\infty$  for all  $j$ . Let  $R$  be a positive number to be specified. Choose  $C = C(R)$  such that  $\sum_{m \geq 0} \sum_{\beta} (\beta! m!)^{-1} \|D_t^m D_x^\beta F(0, 0)\|_\infty R^{|\beta|+m}$  and  $\sum_{\beta} (\beta!)^{-1} \|D_x^\beta g(0)\|_\infty R^{|\beta|}$  are both  $\leq C$ .

(a) Among the  $2^m$  terms of the first kind in Problem 7, show that each one for which  $k$  of the  $m$  factors  $T_1, \dots, T_m$  are  $B$  is  $\leq M^m N^{m-k} C R^{-(m-k)} (m-k)!$ , so that the sum of the contributions from the terms of the first kind to  $\|D_t^m u(0, 0)\|_\infty$  is  $\leq \sum_{k=0}^m \binom{m}{k} M^m N^{m-k} C R^{-(m-k)} (m-k)!$ .

(b) Taking into account the result of Problem 8a, adjust the estimate in part (a) of the present problem to bound the sum of the contributions from the terms of the first kind to  $\|D_t^m D_x^\beta u(0, 0)\|_\infty$ .

(c) Summing over  $m \geq 0, l \geq 0$ , and  $\beta$  with  $|\beta| = l$  the estimate in part (b) and using the formulas in Problem 6, show that the contribution of the terms of the first kind to the series in Problem 8b is finite if  $R$  is chosen large enough so that  $Nr/R \leq \frac{1}{2}$  and  $2MrN/R < 1$ .

- (d) For the  $2^m - 1$  terms of the second kind in Problem 7, replace  $T_1 \cdots T_s$  by  $T_1 \cdots T_{m-1}$ , treating the missing factors as the identity  $I$ , each such factor accompanying a differentiation  $D_t$ . If there are  $k$  factors of  $B$ , show that the term is  $\leq M^{m-1}(N+1)^{m-1-k}CR^{-(m-1-k)}(m-1-k)!$ . Arguing in a fashion similar to the previous parts to this problem, show that consequently the contribution of the terms of the second kind to the series in Problem 8b is finite if  $R$  is chosen large enough so that  $Nr/R \leq \frac{1}{2}$  and  $2Mr(N+1)/R < 1$ .

Problems 10–12 concern the reduction to a first-order system of the Cauchy problem for a single  $m^{\text{th}}$ -order partial differential equation that has been solved for  $D_x^m u$ . They generalize the discussion of a second-order equation in two variables that appeared in Section 1 and reduce Theorems 7.3 and 7.4 to Theorems 7.1 and 7.2, respectively. In two variables  $(x, y)$ , the equation is

$$D_x^m u = F(x, y; u; D_x u, D_y u; D_x^2 u, \dots; D_x^{m-1} D_y u, \dots, D_y^m u),$$

and the Cauchy data are

$$D_x^i u(0, y) = f^{(i)}(y) \quad \text{for } 0 \leq i < m.$$

10. In the case of two variables  $(x, y)$ , introduce variables  $u^{i,j}$  for  $i + j \leq m$ . Show that the given Cauchy problem is equivalent to the following Cauchy problem for a first-order system

$$\begin{aligned} D_x u^{i,j+1} &= D_y u^{i+1,j} && \text{for } i + j + 1 \leq m, \\ D_x u^{i,0} &= u^{i+1,0} && \text{for } 0 \leq i < m, \\ D_x u^{m,0} &= F_x + u^{1,0} F_{u^{0,0}} + u^{2,0} F_{u^{1,0}} + (D_y u^{1,0}) F_{u^{0,1}} + \cdots + (D_y u^{1,m-1}) F_{u^{0,m}} \end{aligned}$$

with Cauchy data

$$\begin{aligned} u^{i,j}(0, y) &= D_y^j f^{(i)}(y) && \text{for } i + j \leq m, (i, j) \neq (m, 0), \\ u^{m,0}(0, y) &= F(0, y; f^{(0)}(y); f^{(1)}(y), D_y f^{(0)}(y); \dots, D_y^m f^{(0)}(y)). \end{aligned}$$

11. What changes to the setup and argument in Problem 10 are needed to handle more variables, say  $(x, y_1, \dots, y_{N-1})$ ?
12. Back in the situation of two variables  $(x, y)$  as in Problem 10, suppose that  $F$  is a linear combination, with constant coefficients, of  $u, D_x u, D_y u, \dots, D_y^m u$ , plus an entire holomorphic function of  $(x, y)$ , and suppose that  $f^{(0)}, \dots, f^{(m-1)}$  are entire holomorphic functions of  $y$ . Prove that the reduction to first order as in Problem 10 leads to a Cauchy problem for a first-order system of the type in Problems 6–9. Conclude that the Cauchy problem for the given  $m^{\text{th}}$ -order equation in the situation of constant coefficients has an entire holomorphic solution.