An Introduction to Second-Order Partial Differential Equations and the Method of Characteristics

A Foundation Course in Classical PDE Theory

Contents of the Course

Preface

A partial differential equation is an equation that an unknown function must satisfy involving its rates of change in several independent directions simultaneously. The challenge it poses is to find the function—or, given additional conditions, the unique function—that satisfies the relation at every point of a region. Second-order partial differential equations, in which the highest derivatives appearing are of order two, are the equations from which the mathematical description of physical reality is most extensively built: the vibration of a string, the diffusion of heat, the distribution of gravitational potential—all are governed by second-order equations, and each behaves in a manner entirely unlike the others.

The central question this course sets out to answer is what structural feature of a second-order equation determines the behaviour of its solutions and the correct way to find them. The answer emerges from the method of characteristics. In the first-order setting, the characteristics are curves along which the equation reduces to an ordinary differential equation, and finding the solution reduces to following those curves from where the data is known. For a second-order equation, the same question—along which curves does prescribed data fail to propagate into the domain?—leads to a condition that is quadratic rather than linear in the slope of the candidate curve. A quadratic has two real solutions, one, or none, and this trichotomy is the classification of second-order linear equations into three fundamental types, arising directly and inevitably from the attempt to apply the characteristic method at second order.

The course traces that development from beginning to end, introducing every term it uses, deriving every condition it applies, and working every example to an explicit solution. The reader it is written for has met ordinary differential equations and the calculus of several variables; no prior knowledge of partial differential equations is assumed. The aim throughout is to equip the reader to take a second-order linear PDE, determine its type, find its characteristics, reduce it to canonical form, and construct its solution.

1. Three Equations and the Problem They Posed

1.1. Partial Derivatives and the Notion of a PDE

A function of a single variable $f(x)$ has, at each point x, a single rate of change: the derivative $f'(x)$, which measures how f responds to a small change in x. When a function $u(x,t)$ depends on two independent variables—a spatial position x and a time t, for instance—it has at each point not one but two natural rates of change. The *partial derivative* $\partial u/\partial x$, written u_x, measures how u changes as x varies while t is held fixed; it is computed exactly as an ordinary derivative, treating t as a constant. The partial derivative $\partial u/\partial t$, written u_t, measures how u changes as t varies while x is held fixed. For a function like $u(x,t) = x^2 t + \cos(xt)$, for example,

$$u_x = 2xt - t\sin(xt), \qquad u_t = x^2 - x\sin(xt).$$

Second partial derivatives are computed by differentiating twice: $u_{xx} = \partial^2 u/\partial x^2$, $u_{tt} = \partial^2 u/\partial t^2$, and the *mixed* derivative $u_{xt} = \partial^2 u/\partial x\,\partial t$, obtained by differentiating first with respect to x and then with respect to t. For smooth functions, the order of differentiation in the mixed derivative does not matter: $u_{xt} = u_{tx}$.

A *partial differential equation* is an equation that an unknown function must satisfy involving one or more of its partial derivatives, imposing a constraint on how the function and its rates of change relate at every point of its domain, and the problem is to find all functions—or, given additional conditions, the unique function—that satisfy it. A *second-order* PDE is one in which the highest derivatives appearing are of order two. Three such equations stand at the historical and mathematical centre of this course, and understanding why they are central requires understanding where they came from.

1.2. Three Equations, Three Behaviours

The theory of second-order partial differential equations was not born from a single insight. It grew, by necessity, from three separate problems in mathematical physics that occupied the eighteenth and nineteenth centuries and for a long time were treated as three entirely separate theories. Each problem produced an equation with its own body of methods, its own known solutions, its own physical interpretation. The puzzle they eventually posed was not any individual mathematical difficulty but the cumulative puzzle of why three equations belonging formally to the same class behaved so entirely differently.

The first equation arrived from mechanics. In 1746, Jean le Rond d'Alembert set out to give a rigorous mathematical account of the motion of a vibrating string and showed that the transverse displacement $u(x,t)$ satisfies

$$u_{tt} = c^2 u_{xx}. \tag{1}$$

This is the one-dimensional wave equation. D'Alembert found that its general solution has the form $u(x,t) = f(x+ct) + g(x-ct)$ for two arbitrary functions f and g, expressing the motion as the superposition of two waveforms travelling in opposite directions at speed c. The solution was determined by prescribing the initial displacement and the initial velocity—two functions of x at time $t = 0$—and was then uniquely determined for all subsequent time.

The second equation arrived from the theory of heat. In his 1822 treatise, Joseph Fourier derived and studied the equation governing the diffusion of temperature in a conducting medium, which in one spatial dimension is

$$u_t = \kappa\, u_{xx}. \tag{2}$$

Fourier's solutions behaved quite differently from d'Alembert's. They did not propagate waveforms; they smoothed them. An initial temperature distribution, however rough, became smooth for all positive time. The solution was determined by a single initial condition—the temperature profile at $t = 0$—not two. And prescribing data at a later time and asking for the solution at an earlier time produced a problem that was, in a precise sense, catastrophically ill-determined.

The third equation arrived from the theory of gravitational and electrostatic potential. Pierre-Simon Laplace showed that the potential u in a region free of sources satisfies

$$u_{xx} + u_{yy} = 0. \tag{3}$$

The Laplace equation had no time variable. Its solutions were determined by their values on the *boundary* of the domain, not by any initial data in the interior. They were the smoothest possible functions on their domain, satisfying a maximum principle that forbade any interior point from being a strict maximum or minimum.

Three equations, three solution theories, three sets of intuitions, all three second-order and linear. What structural feature of each is responsible for the entirely different behaviour of its solutions? Answering this question is the work of this course, and the answer comes from the method of characteristics.

1.3. Cauchy's Question and the Road to Classification

Augustin-Louis Cauchy brought precision to the notion of an initial-value problem for PDEs through what is now called the Cauchy problem: given a PDE and data prescribed on a curve, find a solution in a neighbourhood of that curve. For second-order equations, this requires prescribing both the function value and its normal derivative on the initial curve, two pieces of data corresponding to the two free functions in the general solution. Cauchy also began the investigation of when this problem can fail. He observed that certain curves had the property that the equation, together with the Cauchy data on those curves, could not uniquely determine the highest-order derivatives of the solution; these were the characteristic curves.

It was Bernhard Riemann who, in the 1850s, gave this observation its full structural force. Riemann understood that the characteristics of a second-order equation were the geometric objects governing the propagation of information throughout the domain, in exact analogy with the characteristics of first-order equations. He observed that the number of real characteristic directions through a point was determined by an algebraic expression in the coefficients of the equation's highest-order terms, and that this number—two, one, or zero—was an invariant of the equation under smooth changes of coordinates. This was the first clear statement of what would become the classification theorem.

The method of characteristics applied to a second-order equation produces a quadratic condition, and a quadratic has exactly three qualitatively distinct cases. Those three cases are the three types, and the rest of the course is an account of what each case implies for the problem of finding solutions.

2. General Second-Order Equations and Their Classification

2.1. The General Second-Order Linear Equation

The historical equations of Section 1 are all instances of a single general concept, and the classification and the characteristic condition both rest on having that concept stated precisely.

Definition 2.1. A *partial differential equation* (PDE) of order k in n independent variables $x = (x_1, \dots, x_n)$ is an equation of the form

$$F\big(x,\, u(x),\, Du(x),\, D^2u(x),\, \dots,\, D^ku(x)\big) = 0,$$

where $u\colon \Omega \subset \mathbb{R}^n \to \mathbb{R}$ is the unknown function, $Du = (\partial_{x_1}u, \dots, \partial_{x_n}u)$ is its gradient, and D^ju denotes the collection of all partial derivatives of u of order exactly j. The *order* of the PDE is the order of the highest derivative actually appearing in F.

The definition is deliberately general, encompassing equations of any order, in any number of variables, linear or nonlinear. What singles out the setting of this course is the restriction to *second-order* equations in *two* independent variables, where the unknown function u depends on two variables and the equation involves u, its first partial derivatives u_x and u_y, and its three second partial derivatives u_{xx}, u_{xy}, and u_{yy}. A *classical solution* of such an equation on an open set $\Omega \subset \mathbb{R}^2$ is a function $u \in C^2(\Omega)$, twice continuously differentiable, that satisfies the equation at every point of Ω. Throughout this course, all solutions are classical.

The most general second-order linear PDE in two independent variables x and y is

$$A(x,y)\,u_{xx}+2B(x,y)\,u_{xy}+C(x,y)\,u_{yy}+D(x,y)\,u_x+E(x,y)\,u_y+F(x,y)\,u=G(x,y), \quad (4)$$

where A, B, C, D, E, F, and G are given smooth functions on an open domain $\Omega \subset \mathbb{R}^2$, and $u : \Omega \to \mathbb{R}$ is the unknown. The factor of 2 in front of the mixed-derivative term is a convention that ensures the matrix of second-order coefficients is symmetric, placing the classification within the natural framework of real symmetric matrices.

Definition 2.2. The *principal part* of equation (4) is its collection of second-order terms,

$$A\,u_{xx} + 2B\,u_{xy} + C\,u_{yy}.$$

The *coefficient matrix* of the principal part is

$$\mathcal{A}(x,y) = \begin{pmatrix} A(x,y) & B(x,y) \\ B(x,y) & C(x,y) \end{pmatrix}.$$

The lower-order terms in (4), those involving u_x, u_y, u, and G, affect the quantitative form of solutions but not the characteristic geometry of the equation. The characteristics, and therefore the classification, are determined entirely by the principal-part coefficients A, B, C, as will become clear in the derivation of the characteristic condition that follows.

2.2. Hypersurfaces, Hyperplanes, and the Cauchy Problem

Two geometric terms must be in place before the central problem can be stated precisely.

Definition 2.3. A *hypersurface* in $\mathbb{R}^n$ is a smooth surface of dimension $n-1$. In the plane $\mathbb{R}^2$, a hypersurface is a smooth curve. Formally, a hypersurface is a set of the form $\Gamma = \{(x,y) \in \Omega : \phi(x,y) = 0\}$ where ϕ is a smooth function with $\nabla\phi \neq 0$ on Γ.

Definition 2.4. A *hyperplane* in $\mathbb{R}^n$ is a flat hypersurface, one that can be written as $\{x : a \cdot x = c\}$ for some vector a and constant c. In the plane $\mathbb{R}^2$, a hyperplane is a straight line. The lines $\{t = 0\}$, $\{y = 0\}$, and $\{x = 0\}$ that serve as initial curves throughout this course are hyperplanes. The distinction between hypersurface and hyperplane is the distinction between a general smooth curve and its flat special case; the general theory applies to hypersurfaces, while the worked examples typically use hyperplanes for concreteness.

With these geometric terms in place, the Cauchy problem can be stated precisely.

Definition 2.5. The *Cauchy problem* for a second-order PDE consists of finding a function u satisfying the equation in a neighbourhood of a smooth curve Γ, subject to two conditions: the value $u\big|_{\Gamma} = u_0$ and the normal derivative $\partial_\nu u\big|_{\Gamma} = u_1$, where ∂_ν denotes differentiation in the direction normal to Γ. The curve Γ is the *initial curve*, u_0 the *initial value*, and u_1 the *initial normal derivative*. Together, u_0 and u_1 are the *Cauchy data*.

Two pieces of data are required because a second-order equation, when integrated, produces two free functions or constants, and the two pieces of Cauchy data are exactly what is needed to determine both. For the wave equation, for instance, the two free functions f and g in the general solution $f(x+ct)+g(x-ct)$ are pinned by specifying $u(x,0)=u_0(x)$ and $u_t(x,0)=u_1(x)$. On which curves Γ the Cauchy problem succeeds in determining the solution, and on which it fails, is the question the rest of this section answers.

2.3. The Characteristic Condition and Its Geometric Meaning

The characteristic question asks along which curves in Ω the Cauchy data fails to determine the second derivatives of u. A curve along which this failure occurs is a *characteristic curve* of the equation, and finding these curves requires analysing what the Cauchy problem demands of the equation.

Let $\Gamma = \{\phi(x,y)=0\}$ be a smooth curve, and suppose u and $\partial_\nu u$ are known on Γ. Knowing u on Γ means knowing all tangential derivatives of u along Γ. Both u_x and u_y are therefore known on Γ: their tangential components come from differentiating u_0 along the curve, and their normal component comes from $u_1 = \partial_\nu u$. What cannot be recovered from the Cauchy data alone is the second normal derivative u_{nn}.

Equation (4), evaluated on Γ, provides one linear relation among the three second-order quantities u_{xx}, u_{xy}, u_{yy}. Together with the two further relations obtained by differentiating the known first derivatives along Γ, this gives a 3×3 linear system for u_{xx}, u_{xy}, u_{yy}. The system fails to have a unique solution, meaning the Cauchy data fails to determine the second derivatives, precisely when its determinant vanishes. Computing this determinant and simplifying, the condition reduces to

$$A\,\phi_x^2 + 2B\,\phi_x\,\phi_y + C\,\phi_y^2 = 0. \tag{5}$$

This is the *characteristic condition.* A curve $\phi(x, y) = 0$ is a characteristic curve of (4) if and only if ϕ satisfies (5) at every point of the curve. Only the principal-part coefficients A, B, C enter the condition, because the lower-order terms contribute only known quantities to the linear system for the second derivatives and do not alter its coefficient matrix.

The complementary condition, that the initial curve is not characteristic, is what makes the Cauchy problem solvable.

Definition 2.6. Let $\Gamma = \{\phi(x, y) = 0\}$ be a smooth curve. The curve Γ is *noncharacteristic* for equation (4) at a point $(x_0, y_0) \in \Gamma$ if

$$A\phi_x^2 + 2B\phi_x\phi_y + C\phi_y^2 \neq 0 \quad \text{at } (x_0, y_0).$$

This condition is equivalently written as $\mathcal{A}(x_0, y_0)\, \nabla\phi(x_0, y_0) \cdot \nabla\phi(x_0, y_0) \neq 0$. If Γ is noncharacteristic at every one of its points, it is called a *noncharacteristic initial curve.*

The geometric content of this condition is that the characteristics are transverse to Γ, crossing the initial curve at a finite angle rather than running along it. When the noncharacteristic condition holds, the characteristics pass through Γ and carry information from the initial data into the domain.

When it fails—when Γ is itself a characteristic—the characteristics run along the initial curve, the data prescribed on Γ is redundant in one direction and insufficient in the other, and the Cauchy problem is either inconsistent or underdetermined. The noncharacteristic condition is the single geometric requirement on which the solvability of the Cauchy problem rests.

2.4. The Discriminant and the Classification

The characteristic condition (5) can be rewritten as a quadratic in the slope $\lambda = dy/dx$ of the characteristic curve. Assuming $\phi_y \neq 0$ and setting $\lambda = -\phi_x/\phi_y$ (the slope of the level curve $\phi = 0$), condition (5) becomes

$$A\,\lambda^2 - 2B\,\lambda + C = 0. \tag{6}$$

This is a quadratic in λ whose discriminant is

$$\Delta(x, y) = B(x, y)^2 - A(x, y)\,C(x, y). \tag{7}$$

The number of real solutions of (6) is determined entirely by the sign of Δ: two distinct real characteristic slopes when $\Delta > 0$, exactly one when $\Delta = 0$, and none when $\Delta < 0$. This trichotomy is the classification.

Definition 2.7. The equation (4) is called, at a point $(x_0, y_0) \in \Omega$: *hyperbolic* if $\Delta(x_0, y_0) > 0$, i.e., two distinct real characteristic directions; *parabolic* if $\Delta(x_0, y_0) = 0$, i.e., exactly one real characteristic direction; *elliptic* if $\Delta(x_0, y_0) < 0$, i.e., no real characteristic directions. If the type is the same at every point of Ω, the equation is *uniformly* of that type on Ω.

The names are borrowed from the classification of conic sections, which share the same discriminant. The connection to the coefficient matrix $\mathcal{A}$ is direct: $\det \mathcal{A} = AC - B^2 = -\Delta$. The equation is hyperbolic when $\mathcal{A}$ has eigenvalues of opposite sign (indefinite), parabolic when $\mathcal{A}$ is singular, and elliptic when $\mathcal{A}$ is definite.

The wave equation (1) has $A = 1$, $B = 0$, $C = -c^2$, giving $\Delta = c^2 > 0$: hyperbolic. The heat equation (2), viewed in the (x, t)-plane with $A = \kappa$, $B = 0$, $C = 0$ (no u_{tt} term), gives $\Delta = 0$: parabolic. The Laplace equation (3) has $A = 1$, $B = 0$, $C = 1$, giving $\Delta = -1 < 0$: elliptic. The three historical equations are the canonical representatives of the three types.

When $\Delta > 0$, the two real roots of (6),

$$\lambda_\pm = \frac{B \pm \sqrt{\Delta}}{A} \qquad (A \neq 0), \tag{8}$$

give two families of ODEs $dy/dx = \lambda_\pm$, each defining a family of characteristic curves filling the domain. When $\Delta = 0$, the two roots coincide and there is a single family. When $\Delta < 0$, there are no real roots and no real characteristic curves, so the equation has no preferred directions of propagation in the plane. These three cases are what the method of characteristics produces when applied to the general second-order equation, and they govern everything that follows.

3. Canonical Forms and How to Reach Them

3.1. Why Canonical Forms Are the Goal

Content not available in this sample version.

3.2. Reduction in the Hyperbolic, Parabolic, and Elliptic Cases

Content not available in this sample version.

3.3. Worked Examples of Reduction to Canonical Form

Content not available in this sample version.

4. Hyperbolic Equations and the Method of Characteristics

4.1. Two Characteristic Families and the Cauchy Problem

Content not available in this sample version.

4.2. Worked Examples of Solving Hyperbolic Equations

Content not available in this sample version.

5. Parabolic Equations and the Structure of Their Solutions

5.1. One Characteristic Family, One Free Coordinate

Content not available in this sample version.

5.2. Worked Examples of Solving Parabolic Equations

Content not available in this sample version.

6. Elliptic Equations and the Structure of Their Solutions

6.1. Complex Characteristics and the Laplacian Canonical Form

Content not available in this sample version.

6.2. Worked Examples of Solving Elliptic Equations

Content not available in this sample version.

7. What the Method of Characteristics Has Revealed

Content not available in this sample version.