An Introduction to First-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, and the challenge it poses is finding the function that satisfies the relation at every point of a region, or, when given additional conditions, the unique function that does. First-order partial differential equations—those whose highest-order derivatives are of order one—are the natural entry point into classical solution theory, simple enough to be fully tractable and rich enough to contain the geometry that governs the whole subject.

What the course develops is how data prescribed on some surface in the domain of a first-order partial differential equation propagates along curves determined by the equation itself—curves along each of which the equation reduces to an ordinary differential equation. These are the characteristic curves, and the form they take, and the system needed to find them, depends on how deeply the solution is coupled into the structure of the equation. A first-order equation in which the coefficients of the derivatives depend only on position is a different object from one in which they depend on the solution itself, and both differ from an equation that depends nonlinearly on the entire gradient. Each of these levels demands its own version of the method, and understanding why is the same as understanding the classification.

The method of characteristics is developed across the full hierarchy of first-order equations—from linear through semilinear and quasilinear to fully nonlinear—treating each level not as a variation on the one before it but as a new setting in which a new structural feature of the method becomes necessary. The reader this course is written for has met ordinary differential equations and the calculus of several variables, and no prior knowledge of partial differential equations is assumed. The aim throughout is to equip the reader to take a first-order equation, identify its type, construct its characteristics, and carry the initial data along them to an explicit solution.

1. Origins and History of Partial Differential Equations

1.1. Functions of Several Variables and the Notion of a PDE

Content not available in this sample version; the complete version can be purchased on Leanpub at `https://leanpub.com/first-order-partial-differential-equations`.

1.2. D'Alembert and the Question of Data

Content not available in this sample version; the complete version can be purchased on Leanpub at `https://leanpub.com/first-order-partial-differential-equations`.

1.3. Cauchy, Characteristics, and the Answer

Content not available in this sample version; the complete version can be purchased on Leanpub at `https://leanpub.com/first-order-partial-differential-equations`.

2. General First-Order Equations and Their Classification

2.1. Partial Differential Equations and Their Classification

A partial differential equation of order k in n independent variables $x = (x_1, \ldots, x_n)$ is defined as follows.

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

$$F\big(x,\, u(x),\, Du(x),\, D^2u(x),\, \ldots,\, D^k u(x)\big) = 0,$$

where $u\colon \Omega \subset \mathbb{R}^n \to \mathbb{R}$ is the unknown function, $Du = (\partial_{x_1} u, \ldots, \partial_{x_n} u)$ is its gradient, and $D^j u$ 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. A first-order PDE in two variables x and t takes the form $F(x, t, u, u_x, u_t) = 0$. A *classical solution* of such an equation on an open set Ω is a function u that is continuously differentiable up to the order of the equation and satisfies the equation at every point of Ω. Throughout this course, all solutions are classical.

First-order PDEs fall into four types according to how deeply the solution u is coupled into the structure of the equation, and this classification determines both what the method of characteristics must do at each level and what can happen to solutions.

Definition 2.2. A first-order PDE in n variables is called *linear* if it has the form

$$\sum_{i=1}^{n} a_i(x)\, \frac{\partial u}{\partial x_i} + b(x)\, u = f(x),$$

where the coefficients a_i, b, and the right-hand side f depend only on the independent variables x.

In the linear case, the coefficients of the derivatives depend only on position, so the directions along which the equation propagates information are entirely fixed before any solution is sought, and the equation for u along each of those directions is linear. This is the most tractable level of the hierarchy.

Definition 2.3. A first-order PDE in n variables is called *semilinear* if it has the form

$$\sum_{i=1}^{n} a_i(x)\,\frac{\partial u}{\partial x_i} = f(x,u),$$

where the coefficients a_i of the derivatives depend only on x, but the right-hand side f may depend nonlinearly on u.

The semilinear case preserves the linear case's propagation directions: because the coefficients $a_i(x)$ still depend only on position, the geometry of how the equation carries information through the domain is unchanged. What changes is the behaviour of u along each direction, now governed by a nonlinear term, but the characteristics can still be determined independently of the solution.

Definition 2.4. A first-order PDE in n variables is called *quasilinear* if it has the form

$$\sum_{i=1}^{n} a_i(x,u)\,\frac{\partial u}{\partial x_i} = f(x,u),$$

where the coefficients a_i are allowed to depend on u as well as on x.

The quasilinear case introduces a genuine coupling between the solution and the structure of the equation. Because the coefficients $a_i(x,u)$ now depend on u, the directions along which the equation propagates information can no longer be determined before the solution is known, and the characteristics and the solution must be found together. This mutual dependence is the source of qualitatively new behaviour that the semilinear theory does not exhibit.

Definition 2.5. A first-order PDE in n variables is called *fully nonlinear* if it takes the general form

$$F(x,u,Du) = 0,$$

where F depends nonlinearly on the gradient Du itself.

The fully nonlinear case is the most general. Because F depends nonlinearly on the entire gradient Du, neither the propagation directions nor the solution can be extracted from the equation without simultaneously tracking how the gradient evolves. A new and larger space must be introduced to carry the method through, as will be seen in Section 6.

2.2. Hypersurfaces, Hyperplanes, and the Cauchy Problem

Content not available in this sample version; the complete version can be purchased on Leanpub at `https://leanpub.com/first-order-partial-differential-equations`.

3. Linear Equations and the Method of Characteristics

3.1. Why the Characteristics Are the Right Object to Study

The first-order linear PDE in n variables is

$$\sum_{i=1}^{n} a_i(x)\,\frac{\partial u}{\partial x_i} + b(x)\,u = f(x), \qquad x \in \Omega \subset \mathbb{R}^n, \tag{1}$$

with Cauchy data $u = u_0$ prescribed on a hypersurface Γ. The left side involves a sum of partial derivatives weighted by the coefficients $a_i(x)$, and that sum is precisely the directional derivative of u in the direction of the vector field $\mathbf{a}(x) = (a_1(x), \ldots, a_n(x))$. In the special case $b \equiv 0$, $f \equiv 0$, the equation therefore says

$$\mathbf{a}(x) \cdot \nabla u = 0,$$

meaning u does not change in the $\mathbf{a}$-direction. Along a curve in Ω whose tangent vector is always $\mathbf{a}$, the PDE reduces to $du/ds = 0$, where s is a parameter along the curve, and the solution is found by following these curves from Γ, where u_0 is known, into the domain.

These curves are defined by requiring their tangent vector to equal $\mathbf{a}(x)$ at every point,

$$\frac{dx_i}{ds} = a_i(x(s)), \qquad i = 1, \ldots, n, \tag{2}$$

carrying the characteristics forward through the domain from their starting points on Γ. This is an autonomous system of ordinary differential equations, and by the Picard–Lindelöf theorem there passes through every point of Ω exactly one curve satisfying (2). These are the *characteristic curves* of the equation, and the system (2) is the *characteristic system* for the x-variables.

For the general equation (1) with b and f not necessarily zero, u satisfies an ODE along each characteristic driven by b and f:

$$\frac{du}{ds} = f(x(s)) - b(x(s))\,u, \tag{3}$$

a linear first-order ODE for u along each characteristic $x(s)$. Because the x-equation (2) is independent of u in the linear case, the characteristics are determined first, and equation (3) then becomes a linear ODE in u alone, solved by an integrating factor along each characteristic with initial value $u_0(x_0)$ at the starting point $x_0 \in \Gamma$.

More generally, every equation in the characteristic system has the form of an ODE to be integrated from a known initial value, and the same two methods of integration are available for any such equation—not only for the u-equation but for any variable z—whether a spatial coordinate x_i, the solution value u, or, in the fully nonlinear case, a component p_i of the gradient—whose evolution along the characteristic is governed by an equation of the form

$$\frac{dz}{ds} = g(s), \qquad z(0) = z_0, \tag{4}$$

where g depends only on s and not on z itself. The first method integrates without limits, obtaining

$$z(s) = \int g(s)\, ds + C, \tag{5}$$

and then determines the constant of integration C by imposing $z(0) = z_0$. The second method integrates directly from the initial point, using the fundamental theorem of calculus to write

$$z(s) = z_0 + \int_0^s g(\sigma)\, d\sigma, \tag{6}$$

with the initial condition built into the lower limit of integration from the outset, so that no separate determination of a constant is needed. Both methods yield the same result, as the fundamental theorem guarantees; the second is more direct whenever g is integrable in closed form. When the ODE for a variable in the characteristic system does depend on z—as equation (3) does when $b \not\equiv 0$—the integrating factor remains the appropriate tool, and neither method above applies directly. The choice of procedure therefore depends on the structure of the ODE at hand, and recognising which case one is in is part of carrying the method through.

3.2. The Noncharacteristic Condition

The method requires that each characteristic actually cross Γ rather than lying within it. If a characteristic were to lie entirely inside Γ, the data prescribed along it would already be determined by the equation itself, leaving no freedom for u_0 and no means of propagating anything into the domain.

Definition 3.1. Let $\Gamma = \{\phi(x) = 0\}$ be a smooth hypersurface. The surface Γ is *noncharacteristic* for equation (1) at a point $x_0 \in \Gamma$ if

$$\sum_{i=1}^{n} a_i(x_0)\, \frac{\partial \phi}{\partial x_i}(x_0) \neq 0.$$

This condition is equivalently written as $\mathbf{a}(x_0) \cdot \nabla\phi(x_0) \neq 0$. If this quantity vanishes at x_0, then Γ is a *characteristic surface* for the equation at x_0.

For the flat surface $\Gamma = \{x_n = 0\}$, the noncharacteristic condition reduces simply to $a_n(x_0) \neq 0$, requiring that the coefficient of the derivative in the normal direction be nonzero. This condition governs the method of characteristics at every level of the classification hierarchy: in the semilinear and quasilinear cases its form is unchanged, with the propagation direction updated to reflect the dependence of a_i on u; in the fully nonlinear case it takes the form $F_{p_n} \neq 0$, stated explicitly in Section 6.

3.3. Worked Examples in the Linear Setting

Example 3.1. Consider the Cauchy problem

$$u_t + c\, u_x = 0, \qquad x \in \mathbb{R},\ t > 0$$

with initial data $u(x, 0) = \sin(x)$.

The initial surface is $\Gamma = \{t = 0\}$, described by $\phi(x, t) = t$, so $\nabla\phi = (0, 1)$. The coefficient vector is $\mathbf{a} = (c, 1)$, and the noncharacteristic condition requires

$$\mathbf{a} \cdot \nabla\phi = 1 \neq 0,$$

which holds everywhere on Γ, so the initial surface is noncharacteristic and the method applies.

Parametrise each characteristic by s, with starting point $x_0 \in \Gamma$. The characteristic system and its initial conditions are

$$\frac{dx}{ds} = c, \quad x(0) = x_0; \qquad \frac{dt}{ds} = 1, \quad t(0) = 0; \qquad \frac{du}{ds} = 0, \quad u(0) = u_0(x_0) = \sin(x_0).$$

Integrating the spatial equations gives $t = s$ and $x = x_0 + cs$. The equation for u gives $u = \sin(x_0)$, constant along each characteristic. Substituting $s = t$ and $x_0 = x - ct$ gives the solution

$$u(x, t) = \sin(x - ct).$$

Example 3.2. Consider the Cauchy problem

$$\frac{1}{x}\, u_x - \frac{1}{y}\, u_y = 4u, \qquad x > 0,\ y > 0$$

with data $u(x, x) = x^2$ prescribed on $\Gamma = \{y = x\}$.

The initial surface $\Gamma = \{y - x = 0\}$ is described by $\phi(x, y) = y - x$, so $\nabla\phi = (-1, 1)$. The coefficient vector is $\mathbf{a} = (1/x,\ -1/y)$, giving

$$\mathbf{a} \cdot \nabla\phi = \frac{1}{x}(-1) + \left(-\frac{1}{y}\right)(1) = -\frac{1}{x} - \frac{1}{y} = -\frac{x + y}{xy}.$$

At a point $(x, x) \in \Gamma$ this equals $-2/x \neq 0$ for $x > 0$, so the initial surface is noncharacteristic and the method applies.

Parametrise Γ by $x_0 > 0$, so the starting point on Γ is (x_0, x_0) with $u(0) = u_0(x_0) = x_0^2$. The characteristic system is

$$\frac{dx}{ds} = \frac{1}{x}, \quad x(0) = x_0; \qquad \frac{dy}{ds} = -\frac{1}{y}, \quad y(0) = x_0; \qquad \frac{du}{ds} = 4u, \quad u(0) = x_0^2.$$

Separating variables in the x-equation gives $x\,dx = ds$, so $x^2 = x_0^2 + 2s$. Likewise $y\,dy = -ds$ gives $y^2 = x_0^2 - 2s$. The u-equation integrates to $u = x_0^2 e^{4s}$.

Adding and subtracting the expressions for x^2 and y^2:

$$x^2 + y^2 = 2x_0^2 \implies x_0^2 = \frac{1}{2}(x^2 + y^2), \qquad x^2 - y^2 = 4s \implies s = \frac{1}{4}(x^2 - y^2).$$

Substituting into the expression for u gives the solution

$$u(x, y) = \frac{x^2 + y^2}{2}\, e^{x^2 - y^2}.$$

Example 3.3. Consider the Cauchy problem

$$y\,u_x - x\,u_y = 0, \qquad x > 0$$

with initial data $u(x, 0) = x^2$ prescribed on $\Gamma = \{y = 0,\, x > 0\}$.

The initial surface $\Gamma = \{y = 0\}$ is described by $\phi(x, y) = y$ and $\nabla\phi = (0, 1)$. The coefficient of u_y is $-x$, so

$$\mathbf{a} \cdot \nabla\phi = -x \neq 0 \quad \text{for } x > 0,$$

meaning the initial surface is noncharacteristic on $\{x > 0\}$ and the method applies.

With starting point $(x_0, 0)$ on Γ and $u(0) = u_0(x_0) = x_0^2$, the characteristic system is

$$\frac{dx}{ds} = y, \quad x(0) = x_0; \qquad \frac{dy}{ds} = -x, \quad y(0) = 0; \qquad \frac{du}{ds} = 0, \quad u(0) = x_0^2.$$

The x- and y-equations are coupled. Differentiating the first gives $d^2x/ds^2 = dy/ds = -x$, so $x'' + x = 0$. With $x(0) = x_0$ and $x'(0) = y(0) = 0$:

$$x(s) = x_0 \cos s, \qquad \text{and} \qquad y(s) = \frac{dx}{ds} \implies y(s) = -x_0 \sin s.$$

The u-equation gives $u = x_0^2$, constant along each characteristic. From $x^2 + y^2 = x_0^2 \cos^2 s + x_0^2 \sin^2 s = x_0^2$, the parameter x_0 is recovered as $x_0^2 = x^2 + y^2$, giving the solution

$$u(x, y) = x^2 + y^2.$$

Exercise 3.1. Consider the equation $x\,u_x + y\,u_y = u$ on $\{(x, y) : x > 0,\, y > 0\}$ with data $u(1, y) = y^2$ on the curve $\Gamma = \{x = 1\}$. Verify that Γ is noncharacteristic, write down the characteristic system, solve it explicitly, and express $u(x, y)$ in closed form.

4. Semilinear Equations and Fixed Characteristic Propagation

4.1. How the Semilinear Case Differs from the Linear

The semilinear first-order PDE is

$$\sum_{i=1}^{n} a_i(x)\,\frac{\partial u}{\partial x_i} = f(x,u), \tag{7}$$

where the nonlinearity enters only through the right-hand side $f(x,u)$. The coefficients $a_i(x)$ of the first-order derivatives remain functions of x alone, so the characteristic equation for the spatial position is exactly as in the linear case:

$$\frac{dx_i}{ds} = a_i(x(s)), \qquad i = 1,\ldots,n. \tag{8}$$

The characteristics in x-space can therefore be computed before the solution is known, and the noncharacteristic condition takes the same form as in the linear case, $\mathbf{a}(x_0)\cdot\nabla\phi(x_0) \neq 0$ on Γ. The geometry of information propagation is still fixed entirely by the equation's coefficients.

What changes is the ODE that u satisfies along each characteristic. In the linear case this ODE was linear in u; in the semilinear case it is

$$\frac{du}{ds} = f(x(s),u), \tag{9}$$

which may be nonlinear in u. Since $x(s)$ is already determined independently of u, equation (9) is an ODE for a single scalar function $u(s)$ with known coefficients, and the Picard–Lindelöf theorem still guarantees a unique local solution. The solution is found by integrating the system (8)–(9) from initial data u_0 on Γ along each characteristic.

4.2. Worked Examples of Semilinear Equations

Content not available in this sample version; the complete version can be purchased on Leanpub at `https://leanpub.com/first-order-partial-differential-equations`.

5. Quasilinear Equations and Coupled Characteristics

5.1. How the Quasilinear Case Differs from the Semilinear

The quasilinear first-order PDE is

$$\sum_{i=1}^{n} a_i(x,u)\,\frac{\partial u}{\partial x_i} = f(x,u), \tag{10}$$

where the key departure from the semilinear case is that the coefficients a_i are now allowed to depend on u as well as on x. Because the characteristic direction at each point depends on the value of u there, the characteristics and the solution must be determined simultaneously. The spatial position is governed by

$$\frac{dx_i}{ds} = a_i(x, u), \qquad i = 1, \dots, n, \tag{11}$$

in which the propagation direction depends on u and can no longer be determined before the solution is known. The solution value is propagated by

$$\frac{du}{ds} = f(x, u), \tag{12}$$

which feeds back into the direction equations through the a_i, making (11) and (12) a coupled system of $n + 1$ ODEs that must be solved together. The noncharacteristic condition takes the same form as before, with $\mathbf{a}(x_0, u_0)$ in place of $\mathbf{a}(x_0)$, requiring that $\mathbf{a}(x_0, u_0(x_0)) \cdot \nabla\phi(x_0) \neq 0$ on Γ and ensuring that the characteristics leave the initial surface transversally even as their directions are shaped by the solution they carry.

5.2. Worked Examples of Quasilinear Equations

Content not available in this sample version; the complete version can be purchased on Leanpub at `https://leanpub.com/first-order-partial-differential-equations`.

6. Fully Nonlinear Equations and Extended Characteristics

6.1. Why the Characteristics Must Be Extended into a Larger Space

Content not available in this sample version; the complete version can be purchased on Leanpub at `https://leanpub.com/first-order-partial-differential-equations`.

6.2. The Extended Characteristic System

Content not available in this sample version; the complete version can be purchased on Leanpub at `https://leanpub.com/first-order-partial-differential-equations`.

6.3. Worked Examples of Fully Nonlinear Equations

Content not available in this sample version; the complete version can be purchased on Leanpub at `https://leanpub.com/first-order-partial-differential-equations`.

7. Closing Remarks on the Method of Characteristics

Content not available in this sample version; the complete version can be purchased on Leanpub at `https://leanpub.com/first-order-partial-differential-equations`.