Nonlinear First-Order PDE

Jonathan Ben-Artzi

Spring 2009

1 Introduction: Nonlinear First-Order PDE

Definition 1. Let $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and let $u : \bar{\Omega} \subseteq \mathbb{R}^n \to \mathbb{R}$. A first order partial differential equation for $u = u(x)$ is given by $F(Du, u, x) = 0$, where $F : \mathbb{R}^n \times \mathbb{R} \times \bar{\Omega} \to \mathbb{R}$ is a given function, and Du is the vector of partial derivatives of u.

Notation 2. We usually denote the entries of F as follows: $F = F(p, z, x)$. Thus $p \in \mathbb{R}^n, z \in \mathbb{R}, x \in \bar{\Omega}$.

Remark 3. The PDE $F(Du, u, x) = 0$ is usually accompanied by a boundary condition of the form $u = g$ on $\partial\Omega$. Such a problem is usually called a **boundary value problem**.

Example 4 (The eikonal equation)**.** The **eikonal equation**,

$$|Du| = 1, \tag{1.1}$$

introduced by Hamilton in 1827 is an approximation to the equations which govern the behavior of light traveling through varying materials. A solution, depending on parameters $\|a\| = 1, b \in \mathbb{R}$ is

$$u(x; a, b) = a \cdot x + b. \tag{1.2}$$

Example 5 (The Hamilton-Jacobi equation)**.** The (simple version of the) **Hamilton-Jacobi equation**

$$u_t + H(Du) = 0, \tag{1.3}$$

with $H : \mathbb{R}^n \to \mathbb{R}$ is an important equation from mechanics. $u = u(x, t) : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$. A solution, depending on parameters $a \in \mathbb{R}^n, b \in \mathbb{R}$ is

$$u(x, t; a, b) = a \cdot x - tH(a) + b \tag{1.4}$$

where $t \geq 0$.

Remark 6. For simplicity, in most of what follows, we restrict to $n = 2$. We call the two variables x, y. Thus, we reduce to the case

$$F(u_x, u_y, u, x, y) = 0. \tag{1.5}$$

In this case, the solution $u = u(x, y)$ is a surface in $\mathbb{R}^3$. The normal direction to the surface at each point is given by the vector $(u_x, u_y, -1)$. We will use this fact later, when we construct our solutions using *geometric* methods.

1.1 Linear Equations

The first subclass of nonlinear first-order PDE we consider are the *linear* equations, which have the form:

$$a(x,y)u_x + b(x,y)u_y = c_0(x,y)u + c_1(x,y). \tag{1.6}$$

Example 7. Consider the linear equation

$$u_x = c_0 u + c_1(x,y). \tag{1.7}$$

where c_0 is a constant, and $c_1(x,y)$ is a function of the two variables x and y. Since there is no differentiation with respect to y, the y variable may actually be treated as a parameter, and this equation reduces to an ODE.

To solve, we need some initial condition, for example

$$u(0,y) = y. \tag{1.8}$$

The solution is now simple:

$$u(x,y) = e^{c_0 x}\left(\int_0^x e^{-c_0\xi} c_1(\xi,y)\, d\xi + y\right). \tag{1.9}$$

What do we learn from this? We see that given an initial value along the y-axis, $u(0,y_0)$, we are able to extend to a solution $u(\cdot,y_0)$ parallel to the x-axis. This is true for any y_0.

Does this always work? We split into cases.

1. If for each y_0 we have one piece of information, say at $(0,y_0)$ as above, then we get a unique solution to the ODE along the line through y_0 that is parallel to the x-axis. Thus there is a unique solution to the PDE.

2. If the initial data is given along a different curve (i.e. not the y-axis necessarily) which has the same y value at two different x values, but with values that do not permit a solution to the ODE along that y value, then there is no solution to the PDE as well.

3. If we are given an "insufficient" amount of information (for example, if our initial data is given on a line parallel to the x-axis and *precisely overlaps* a solution to the ODE) then we have an infinite number of solutions to the PDE.

1.2 Quasi-Linear Equations

Another special case of nonlinear first-order PDE are the *quasi-linear* equations, where the nonlinearity appears only on the z (that is, u) variable of F. The general form of such an equation is

$$a(x,y,u)u_x + b(x,y,u)u_y = c(x,y,u). \tag{1.10}$$

2 The Method of Characteristics

The method of characteristics, developed by Hamilton in the 19th century, is essentially the method described above, only for more general examples: We want to construct the two-dimensional surface $u(x, y)$ in $\mathbb{R}^3$ that is a solution to our boundary value problem by translating the problem into infinitely many first order ODEs with initial data inherited from the data provided for the PDE.

2.1 Linear Equations

We return to the linear equation (1.6), and write the initial data in parametric form

$$\Gamma = (x_0(s), y_0(s), u_0(s)) \tag{2.1}$$

where $s \in I = (\alpha, \beta)$.

Γ is called the **initial curve**.

Now, we rewrite (1.6) as

$$(a, b, c_0 u + c_1) \cdot (u_x, u_y, -1) = 0. \tag{2.2}$$

As $(u_x, u_y, -1)$ is perpendicular to the solution surface, the vector $(a, b, c_0 u + c_1)$ must always lie in the tangent plane to the surface. Thus, the system of equations

$$\dot{x}(t) = a(x(t), y(t)) \tag{2.3}$$
$$\dot{y}(t) = b(x(t), y(t)) \tag{2.4}$$
$$\dot{u}(t) = c_0(x(t), y(t))u + c_1(x(t), y(t)), \tag{2.5}$$

defines a family of curves that lie in the solution surface. These equations are called the **characteristic equations** and the curves are called the **characteristics** of the PDE. This is an autonomous system, i.e. there's no dependence on the t variable (which makes sense, since there's no "dynamic" aspect to this problem).

Each characteristic must "start" at the initial curve, and thus we denote the characteristics by $(x(t, s), y(t, s), u(t, s))$ where the variable s tracks the starting point $\Gamma(s)$ on the initial curve:

$$x(0, s) = x_0(s),\ y(0, s) = y_0(s),\ u(0, s) = u_0(s) \tag{2.6}$$

To summarize: The parameter t measures how "far" we are from the initial curve. For $t = 0$ we are actually on it. The parameter s measures where we are along the initial curve. The resulting surface is called an **integral surface**.

Indeed, from calculus we know that (generally) two parameters describe a surface in $\mathbb{R}^3$.

2.2 General Nonlinear Equations

Let us derive the characteristic equations in *full generality*. We once again consider

$$F(Du, u, x) = 0 \tag{2.7}$$

where $x \in \Omega \subseteq \mathbb{R}^n$, with the boundary condition

$$u = g \ \text{ on } \Gamma \tag{2.8}$$

where $\Gamma \subseteq \partial\Omega$ and $g : \Gamma \to \mathbb{R}$ are given. We assume F, g to be smooth as well.

Suppose $\mathbf{x}(t) = (x^1(t), \ldots, x^n(t))$ is a parametric representation of a characteristic curve (which will be calculated later). We define two functions of the parameter t that track the values of u and Du along the characteristic:

$$z(t) := u(\mathbf{x}(t)), \tag{2.9}$$

and

$$\mathbf{p}(t) := Du(\mathbf{x}(t)), \tag{2.10}$$

so that $\mathbf{p}(t) = (p^1(t), \ldots, p^n(t))$ where

$$p^i(t) = u_{x_i}(\mathbf{x}(t)). \tag{2.11}$$

Differentiating (2.11), we have

$$\dot{p}^i(t) = \sum_{j=1}^{n} u_{x_i x_j}\left(\mathbf{x}(t)\right) \dot{x}^j(t). \tag{2.12}$$

Differentiating (2.7) with respect to x_i, we have

$$\sum_{j=1}^{n} \frac{\partial F}{\partial p_j}(Du, u, x) u_{x_j x_i} + \frac{\partial F}{\partial z}(Du, u, x) u_{x_i} + \frac{\partial F}{\partial x_i}(Du, u, x) = 0. \tag{2.13}$$

Evaluating (2.13) at $x = \mathbf{x}(t)$, we get (using (2.9), (2.10)) the identity

$$\sum_{j=1}^{n} \frac{\partial F}{\partial p_j}\left(\mathbf{p}(t), z(t), \mathbf{x}(t)\right) u_{x_i x_j}\left(\mathbf{x}(t)\right) + \frac{\partial F}{\partial z}\left(\mathbf{p}(t), z(t), \mathbf{x}(t)\right) p^i(t) + \frac{\partial F}{\partial x_i}\left(\mathbf{p}(t), z(t), \mathbf{x}(t)\right) = 0. \tag{2.14}$$

We also make the assumption

$$\dot{x}^j(t) = \frac{\partial F}{\partial p_j}\left(\mathbf{p}(t), z(t), \mathbf{x}(t)\right), \tag{2.15}$$

for $j = 1, \ldots, n$. Substituting (2.14) and (2.15) into (2.12) we have:

$$\dot{p}^i(t) = -\frac{\partial F}{\partial x_i}\left(\mathbf{p}(t), z(t), \mathbf{x}(t)\right) - \frac{\partial F}{\partial z}\left(\mathbf{p}(t), z(t), \mathbf{x}(t)\right) p^i(t) \tag{2.16}$$

for $i = 1, \ldots, n$. Last, we differentiate (2.9) to get

$$\dot{z}(t) = \sum_{j=1}^{n} \frac{\partial u}{\partial x_j}\left(\mathbf{x}(t)\right) \dot{x}^j(t) = \sum_{j=1}^{n} p^j(t) \frac{\partial F}{\partial p_j}\left(\mathbf{p}(t), z(t), \mathbf{x}(t)\right), \tag{2.17}$$

where to obtain the second equality we use (2.11) and (2.15).

Writing equations (2.15)-(2.17) in vector notation we have a system of $2n+1$ first-order ODEs called the **characteristic equations** of the PDE (2.7):

$$\dot{\mathbf{p}}(t) = -D_x F\left(\mathbf{p}(t), z(t), \mathbf{x}(t)\right) - D_z F\left(\mathbf{p}(t), z(t), \mathbf{x}(t)\right) \mathbf{p}(t) \tag{2.18}$$

$$\dot{z}(t) = D_p F\left(\mathbf{p}(t), z(t), \mathbf{x}(t)\right) \cdot \mathbf{p}(t) \tag{2.19}$$

$$\dot{\mathbf{x}}(t) = D_p F\left(\mathbf{p}(t), z(t), \mathbf{x}(t)\right). \tag{2.20}$$

Note that we *assumed* that (2.20) holds!

Remark 8. Notice that in the linear case we got only equations (2.19) and (2.20). It turns out that in linear (and quasi-linear) cases, the n-equations given in (2.18) are redundant.

3 Introduction to Variational Principles

In the previous section, we discovered the characteristic equations (2.18)-(2.20) which enable us to evolve initial data given to us along some $n-1$-dimensional surface into an n-dimensional solution surface. This procedure is *local* in the sense that we need to make sure that the ODEs can be solved with no conflicts or blow ups.

An alternative approach when considering physical systems, is to look at the entire (infinite dimensional) space of possible "paths" a system might take. The "true" path is usually a path that minimizes a certain quantity (by the *principle of least action*) and we will see that it satisfies certain equations.

A typical example is finding a geodesic between two points A and B on a manifold. In such a case, the quantity for which we seek a minimum is the *distance functional*, and the set on which we seek the minimum is the *set of all continuous (smooth) paths on the manifold, between A and B*.

3.1 Calculus of Variations

We start with a function

$$L : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R},$$

where $L = L(q, x)$ is called the **Lagrangian**. We now define the **action functional**

$$J[\mathbf{w}(\cdot)] := \int_0^t L\left(\dot{\mathbf{w}}(s), \mathbf{w}(s)\right)\, ds, \tag{3.1}$$

defined for functions $\mathbf{w}(\cdot) = (w^1(\cdot), \ldots, w^n(\cdot))$ belonging to the **admissible class**

$$\mathcal{A} := \{\text{smooth curves } \mathbf{w} \text{ with } \mathbf{w}(0) = y, \mathbf{w}(t) = x\}. \tag{3.2}$$

We seek a curve $\mathbf{x} \in \mathcal{A}$ that minimizes J over $\mathcal{A}$. The following theorem can be shown:

Theorem 9 (Euler-Lagrange equations)**.** The minimizer $\mathbf{x}(\cdot)$ solves the system of **Euler-Lagrange equations**

$$\frac{d}{ds}\left(\frac{\partial L}{\partial q}\left(\dot{\mathbf{x}}(s), \mathbf{x}(s)\right)\right) - \frac{\partial L}{\partial x}\left(\dot{\mathbf{x}}(s), \mathbf{x}(s)\right) = 0 \tag{3.3}$$

for each $0 \le s \le t$. These are n coupled second order PDEs.

4 Introduction to Hamilton-Jacobi Equations

4.1 Characteristics for the general Hamilton-Jacobi Equation

The general Hamilton-Jacobi equation is

$$G(Du, u_t, u, x, t) = u_t + H(Du, x) = 0. \tag{4.1}$$

The function H is called the **Hamiltonian**.

Denoting $q = (p, p_{n+1})$ and $y = (x, t)$, we have

$$G(q, z, y) = p_{n+1} + H(p, x). \tag{4.2}$$

Thus, we have the following partial derivatives:

$$\begin{aligned} D_q G &= (D_p H(p, x), 1) \\ D_y G &= (D_x H(p, x), 0) \\ D_z G &= 0. \end{aligned}$$

Plugging into (2.20), we get

$$\begin{cases} \dot{x}^i(s) = \frac{\partial H}{\partial p_i}\left(\mathbf{p}(s), \mathbf{x}(s)\right), \quad (i = 1, \dots, n) \\ \dot{x}^{n+1}(s) = 1 \end{cases} \tag{4.3}$$

so that the parameter s actually identifies with the time t. Equation (2.18) becomes

$$\begin{cases} \dot{p}^i(s) = -\frac{\partial H}{\partial x_i}\left(\mathbf{p}(s), \mathbf{x}(s)\right), \quad (i = 1, \dots, n) \\ \dot{p}^{n+1}(s) = 0 \end{cases} \tag{4.4}$$

and equation (2.19) becomes

$$\begin{aligned} \dot{z}(s) &= D_p H\left(\mathbf{p}(s), \mathbf{x}(s)\right) \cdot \mathbf{p}(s) + p^{n+1}(s) \\ &= D_p H\left(\mathbf{p}(s), \mathbf{x}(s)\right) \cdot \mathbf{p}(s) - H\left(\mathbf{p}(s), \mathbf{x}(s)\right) \end{aligned}$$

The first and third of the characteristic equations may be written in vector form

$$\dot{\mathbf{x}} = D_p H(\mathbf{p}, \mathbf{x}) \tag{4.5}$$

$$\dot{\mathbf{p}} = -D_x H(\mathbf{p}, \mathbf{x}) \tag{4.6}$$

and are called **Hamilton's equations**. (Here $\mathbf{x}(\cdot) = (x^1(\cdot), \dots, x^n(\cdot)), \mathbf{p}(\cdot) = (p^1(\cdot), \dots, p^n(\cdot))$.) The equation for z becomes trivial once we have Hamilton's equations.

4.2 Relating Hamilton's Equations with the Euler-Lagrange Equations

We begin by defining the **generalized momentum** corresponding to the position $\mathbf{x}(\cdot)$ and the velocity $\dot{\mathbf{x}}(\cdot)$ by

$$\mathbf{p}(s) := D_q L\left(\dot{\mathbf{x}}(s), \mathbf{x}(s)\right) \tag{4.7}$$

where $0 \leq s \leq t$. We make the following assumption:

> **"Convexity assumption": For all $x, p \in \mathbb{R}^n$ the equation $p = D_q L(q, x)$ can be uniquely solved for q as a smooth function of p and x, $q = \mathbf{q}(p, x)$.**

Definition 10 (Hamiltonian)**.** The **Hamiltonian** H associated with the Lagrangian L is

$$H(p, x) := p \cdot \mathbf{q}(p, x) - L(\mathbf{q}(p, x), x)$$

where $p, x \in \mathbb{R}^n$ and the function $\mathbf{q}(p, x)$ is defined implicitly by the assumption above.

The following theorem, whose proof relies on the Euler-Lagrange equations,

Theorem 11. The functions $\mathbf{x}(\cdot)$ and $\mathbf{p}(\cdot)$ satisfy Hamilton's equations

$$\dot{\mathbf{x}} = D_p H(\mathbf{p}, \mathbf{x}) \tag{4.8}$$

$$\dot{\mathbf{p}} = -D_x H(\mathbf{p}, \mathbf{x}) \tag{4.9}$$

for $0 \leq s \leq t$. Furthermore, the mapping

$$s \mapsto H(\mathbf{p}(s), \mathbf{x}(s))$$

is constant.

4.3 Motivating Example

We consider the Lagrangian

$$L(q, x) = \frac{1}{2} m |q|^2 - V(x)$$

where $m > 0$. Then we have

$$\begin{aligned} \frac{\partial L}{\partial q} &= mq \\ \frac{\partial L}{\partial x} &= -DV(x). \end{aligned}$$

Thus, the Euler-Lagrange equations (3.3) become

$$\begin{aligned} \frac{d}{ds}\left(\frac{\partial L}{\partial q}\left(\dot{\mathbf{x}}(s), \mathbf{x}(s)\right)\right) - \frac{\partial L}{\partial x}\left(\dot{\mathbf{x}}(s), \mathbf{x}(s)\right) &= \frac{d}{ds}\left(m\dot{\mathbf{x}}(s)\right) + DV(\mathbf{x}(s)) \\ &= m\ddot{\mathbf{x}}(s) - \mathbf{F}\left(\mathbf{x}(s)\right) = 0 \end{aligned}$$

where $\mathbf{F} := -DV$. This is precisely Newton's second law, describing the motion of a particle with mass m in a force field $\mathbf{F}$ generated by a potential V.

We check if the assumption above holds, namely, if the equation $p = D_q L(q, x)$ can be uniquely solved for q as a smooth function of p and x. We have

$$p = D_q L(q, x) = mq$$

So that

$$\mathbf{q}(p, x) = q = \frac{p}{m}.$$

Now, recalling Definition 10 (of the Hamiltonian H associated to L), we have

$$\begin{aligned} H(p, x) &= p \cdot \mathbf{q}(p, x) - L(\mathbf{q}(p, x), x) \\ &= p \cdot \frac{p}{m} - \frac{1}{2} m \left| \frac{p}{m} \right|^2 + V(x) \\ &= \frac{1}{2m} |p|^2 + V(x), \end{aligned}$$

which is the total energy (kinetic+potantial). By Theorem 11 the energy is constant for a solution.

4.4 The Initial Value Problem

The initial value problem for the Hamilton-Jacobi equation is the system

$$\begin{cases} u_t + H(Du) = 0 & \text{in } \mathbb{R}^n \times (0, \infty) \\ u = g & \text{on } \mathbb{R}^n \times \{t = 0\}, \end{cases} \tag{4.10}$$

where $u = u(x, t) : \mathbb{R}^n \times [0, \infty) \to \mathbb{R}$, $Du = D_x u = (u_{x_1}, \ldots, u_{x_n})$, and the Hamiltonian H and the initial function $g : \mathbb{R}^n \to \mathbb{R}$ are given.

5 Introduction to Conservation Laws

References

[1] Vladimir Arnold, *Mathematical Methods of Classical Mechanics*. Springer-Verlag, New York, 1978.

[2] Oliver Bühler, *A Brief Introduction to Classical, Statistical and Quantum Mechanics*. Courant Lecture Notes in Mathematics 13, New York, 2006.

[3] Lawrence C. Evans, *Partial Differential Equations*. AMS, Providence, RI, 1998.

[4] Yehuda Pinchover and Ya'akov Rubinstein, *Introduction to Partial Differential Equations (Hebrew)*. Technion, Haifa, Israel, Second Edition, 2003.