Various lecture notes for 18300.

R. R. Rosales (MIT, Math. Dept., 2-337)

May 8, 2022

Abstract

Notes, both complete and/or incomplete, for MIT's 18.300 (Principles of Applied Mathematics). These notes will be updated from time to time. Check the date and version.

Contents

List of Figures

1 General Nonlinear First Order Equation

Here we consider the general nonlinear first order equation

$$\boldsymbol{H(\vec{p}, \vec{x}, \phi) = 0} \qquad (1.1)$$

for the (scalar) function $\boldsymbol{\phi = \phi(\vec{x})}$, where $\boldsymbol{\vec{p} = \nabla\phi}$, and $\boldsymbol{H}$ is a given scalar (smooth enough) function. This equation has the

characteristic form

$$\frac{\mathbf{d}}{\mathbf{dt}}\boldsymbol{x_n} = \boldsymbol{H_{p_n}}, \quad \frac{\mathbf{d}}{\mathbf{dt}}\boldsymbol{p_n} = \boldsymbol{-H_{x_n} - p_n\,H_\phi}, \quad \text{and} \quad \frac{\mathbf{d}}{\mathbf{dt}}\boldsymbol{\phi} = \boldsymbol{p_\ell\,H_{p_\ell}}, \qquad (1.2)$$

where (i) $1 \leq n \leq N$, (ii) subscripts indicate partial derivatives of H, (iii) we **use the summation convention** (add over repeated indexes), (iv) the curves defined by $\frac{\mathrm{d}}{\mathrm{d}t}x_n = H_{p_n}$ are called **bi-characteristics** (see remark 1.2), and (v) $\boldsymbol{t}$ is a parameter along the bi-characteristics (the reason we name the parameter $\boldsymbol{t}$ is indicated below). Note that **(1.2) is a complete system of $2N+1$ ode for $\vec{x}$, $\vec{p}$, and ϕ.** Note also that the equations yield

$$\frac{\mathbf{d}}{\mathbf{dt}}\boldsymbol{H} = \mathbf{0}. \qquad (1.3)$$

Remark 1.1 If $\boldsymbol{H}$ does not depend on $\boldsymbol{\phi}$ ($H_\phi = 0$), (1.2) has the standard Hamiltonian form of the equations for an isolated mechanical system in classical physics, with $\boldsymbol{H}$ the Hamiltonian (i.e.: Energy), and $\boldsymbol{\vec{p}}$ the vector of momentums. Then $\boldsymbol{t}$ **is time.** Further: the equation for ϕ along the bi-characteristics can be writen as $\frac{\mathbf{d}}{\mathbf{dt}}\boldsymbol{\phi} = \boldsymbol{\vec{p}} \cdot \frac{\mathbf{d}}{\mathbf{dt}}\boldsymbol{\vec{x}}$; i.e.: **$\boldsymbol{\phi}$ is the action in the energy surface $\boldsymbol{H = 0}$.** ♣

Proof of (1.2), ⇒. Assume that ϕ is a solution of (1.1), and define the curves $\frac{\mathrm{d}}{\mathrm{d}t}x_n = H_{p_n}$. Then, using the chain rule: [#1] $\frac{\mathrm{d}}{\mathrm{d}t}p_n = (p_n)_{x_\ell} H_{p_n} = (p_\ell)_{x_n} H_{p_n}$, since $\phi_{x_n\,x_\ell} = \phi_{x_\ell\,x_n}$. On the other hand, taking the x_n derivative of the equation leads to [#2] $H_{p_\ell}(p_\ell)_{x_n} + H_{x_n} + H_\phi\,p_n = 0$. Substituting [#2] into [#1] yields $\frac{\mathrm{d}}{\mathrm{d}t}p_n = -H_{x_n} - p_n\,H_\phi$. Finally, the last equation, $\frac{\mathrm{d}}{\mathrm{d}t}\phi = p_\ell\,H_{p_\ell}$, as well as (1.3), follow from the chain rule. **QED**.

Proof of (1.2), ⇐. Assume a function ϕ such that it satisfies (1.2), with the bi-characteristics starting (i.e.: $t = 0$) on a hyper-surface $\mathcal{S}$ where $H = 0$. Then, since (1.3) applies, ϕ solves (1.1).

Example 1.1 Empty equation. What happens if we start with an equation that has no solutions, for example take $\boldsymbol{H = \frac{1}{2}\,p^2 + 1}$. Then the characteristic equations are $\frac{\mathbf{d}}{\mathbf{d}t}\vec{\boldsymbol{x}} = \vec{\boldsymbol{p}}$, $\frac{\mathbf{d}}{\mathbf{d}t}\vec{\boldsymbol{p}} = \mathbf{0}$, and $\frac{\mathbf{d}}{\mathbf{d}t}\boldsymbol{\phi} = \boldsymbol{p^2}$. These are fine as ode, however: **they cannot provide a solution to the pde** because there is no surface $\mathcal{S}$ on which we can have $H = 0$. ♣

Remark 1.2 Bi-characteristics. The curves in (1.2) are called bi-characteristics — a special name because they are more special than a typical characteristic, as follows:

$$\text{\textit{Consider (1.1), and let us look for characteristics in the generic sense defined earlier: surfaces along which "weak singularities" can occur.}} \tag{1.4}$$

Thus we consider a curvilinear coordinate system $\{\eta_n\}$ such that the equation allows $\vec{p} = \nabla\phi$ to be discontinuous across the η_1 = constant surfaces. The relationship between $\vec{p}$ and $\vec{q}$ = "gradient of ϕ in the η coordinates" is (by the chain rule) $\boldsymbol{p_n = q_\ell\,\mathcal{N}_{\ell n}}$, where $\mathcal{N}_{\ell n} = (\eta_\ell)_{x_n}$. At a characteristic the equation should allow q_1 to be discontinuous, which leads to (chain rule) $0 = H_{q_1} = H_{p_n}(p_n)_{q_1} = H_{p_n}\mathcal{N}_{1n}$ — i.e.: the vectors $\{H_{p_n}\}$ and $\nabla\eta_1$ are orthogonal. Hence:

$$\text{\textit{A hyper-surface is characteristic if the vector } } \{H_{p_n}\} \text{\textit{ is tangential to the surface.}} \tag{1.5}$$

It follows that:

$$\text{\textit{The bi-characteristics are the intersections of all the possible characteristic hyper-surfaces.}} \tag{1.6}$$

One could say that the bi-characteristics are characteristics on steroids. ♣

Example 1.2 Eikonal equation/rays. The Eikonal equation[1] is

$$\boldsymbol{c^2\,(\nabla\phi)^2 = 1}, \tag{1.7}$$

where $\boldsymbol{c = c(\vec{x}) > 0}$ is the wave-front speed, and the wave-front at time t is given by the implicit equation $\boldsymbol{\phi = t}$. Here we can take

$$\boldsymbol{H = \frac{1}{2}(c^2 p^2 - 1)}, \tag{1.8}$$

where the normalization factor $\frac{1}{2}$ is so that "time" along the bi-characteristics agrees with the time along the fronts (i.e.: $\frac{\mathrm{d}}{\mathrm{d}t}\phi = 1$ below). The characteristic equations are then

$$\frac{\mathbf{d}}{\mathbf{d}t}\boldsymbol{x_n} = \boldsymbol{c^2\,p_n}, \quad \frac{\mathbf{d}}{\mathbf{d}t}\boldsymbol{p_n} = -\frac{\mathbf{1}}{\boldsymbol{c}}\,\boldsymbol{c_{x_n}}, \quad \text{and} \quad \frac{\mathbf{d}}{\mathbf{d}t}\boldsymbol{\phi} = \mathbf{1}. \tag{1.9}$$

Proof, using (1.2). The 1-st equation is obvious. The 2-nd and 3-rd follow from $\frac{\mathbf{d}}{\mathbf{d}t}\boldsymbol{p_n} = -\boldsymbol{c\,c_{x_n}\,p^2}$ and $\frac{\mathbf{d}}{\mathbf{d}t}\boldsymbol{\phi} = \boldsymbol{c^2p^2}$, upon use of $\boldsymbol{c^2p^2 = 1}$.

Note #1. In this case the bi-characteristics are also called *rays*.

Note #2. Since $c^2p^2 = 1$, the first equation above is the same as

$$\frac{\mathbf{d}}{\mathbf{d}t}\vec{\boldsymbol{x}} = \boldsymbol{c\,\hat{n}}, \tag{1.10}$$

where $\boldsymbol{\hat{n}}$ is the unit normal to the wave-front. Thus it is just a re-statement of the fact that the wave-front propagates normal to itself at speed c. The second equation is non-trivial, and governs how the normals to the front evolve as the front propagates. ♣

[1] For a wave-front propagating normal to itself at speed c.

1.1 Fermat's Principle

Eliminating $\vec{p}$ from (1.9) leads to

$$\frac{d}{dt}\left(\frac{1}{c^2}\frac{d}{dt}\vec{x}\right) = -\frac{1}{c}\nabla c. \quad (1.1.1)$$

However, from (1.10) we see that

$$ds = c\,dt, \quad (1.1.2)$$

where s is the **arc-length**. Thus we can write (see remarks 1.1.1 and 1.1.2)

$$\frac{d}{ds}\left(\frac{1}{c}\frac{d}{ds}\vec{x}\right) = \nabla\frac{1}{c}. \quad (1.1.3)$$

Fermat's principle states that: *The rays/bi-characteristics of the Eikonal equation are stationary points for the travel time*

$$T = \int_{\vec{x}_1}^{\vec{x}_2} dt = \int_{\vec{x}_1}^{\vec{x}_2} \frac{ds}{c}. \quad (1.1.4)$$

That is: consider all the path's connecting two fixed points $\vec{x}_1$ and $\vec{x}_2$. T is then a function on the set of paths, and the rays connecting $\vec{x}_1$ to $\vec{x}_2$ (there could be several) are the stationary values of T.

Note: the **rays are extrema, not minimums**, as often stated. See example 1.1.2

Proof. For an arbitrary path from $\vec{x}_1$ to $\vec{x}_2$, $\vec{x} = \vec{x}(t)$, we can write: $T = \int_{\vec{x}_1}^{\vec{x}_2} \frac{v}{c} dt$ — where $\vec{v} = \frac{d}{dt}\vec{x}$, $v = |\vec{v}|$, and $ds = v\,dt$. Hence, with $\mathcal{L} = \mathcal{L}(\vec{v}, \vec{x}) = v/c(\vec{x})$, the Euler-Lagrange variational equations,[2] $\frac{d}{dt}\mathcal{L}_{\vec{v}} = \mathcal{L}_{\vec{x}}$, applied to T yield $\frac{d}{dt}\left(\frac{1}{c\,v}\vec{v}\right) = v\,\nabla c^{-1}$. Since $ds = v\,dt$, this is the same as (1.1.3). **QED**

Remark 1.1.1 Compatibility of (1.1.3). Equation (1.1.3) is consistent with $|d\vec{x}/ds| = 1$. Specifically, if $|d\vec{x}/ds| = 1$ for any value of s, then $|d\vec{x}/ds| \equiv 1$.

Proof. Let $\vec{k} = d\vec{x}/ds$. Then expand (1.1.3) to $c^{-1}d\vec{k}/ds + \left(\vec{k}\cdot\nabla c^{-1}\right)\vec{k} = \nabla c^{-1}$, and dot multiply by $2\,c\,\vec{k}$ to obtain the equation $dk^2/ds = 2\,c\,(1-k^2)\,\vec{k}\cdot\nabla c^{-1}$. Hence, if $k^2 = 1$ at some point, $k^2 \equiv 1$. **QED**

Remark 1.1.2 Index of refraction. In the case of optics, if c_0 is the light speed in vacum, $n = c_0/c$ is the index of refraction. Thus (1.1.3) is equivalent to

$$\frac{d}{ds}\left(n\frac{d}{ds}\vec{x}\right) = \nabla n.$$

Example 1.1.1 Snell's law. yy

Example 1.1.2 Mirror laws and example where a ray is not a minimum. yy

---

The End.

[2] The Euler-Lagrange equations are derived elsewhere.