Intro to PDEs

Recall that an Ordinary Differential Equation (ODE) is an equation that relates a function and its derivatives. For example, if y is a function of x, $y = y(x)$ (remark on notation) (so y is a dependent variable and x an independent variable), the equation:

$$\frac{d^2y}{dx^2} + y = 0 \qquad (*)$$

is a ODE. It can also be written as $y'' + y = 0$. In an ODE, the <u>function is the unknown</u>, so the question to be answered in this example is: can we find a <u>function</u> y that <u>satisfies</u> $(*)$ (ie, find y such that when y is plugged into $*$) an equality is obtained)?

In this case, the answer is yes. e.g. $y(x) = \cos x$, then $\frac{d^2y}{dx^2} + y = -\cos x + \cos x = 0$ ✓. But also $\tilde{y}(x) = \sin x$: $\frac{d^2\tilde{y}}{dx^2} + \tilde{y} = -\sin x + \sin x = 0$

Compare to an algebraic equation, e.g, $x^2 - 1 = 0$: can we find a number x such that the equation is satisfied.

Notice that, as for algebraic equations, ODEs can admit multiple (more than one) solutions.

If our function is a function of more than one variable, e.g. $z = z(x,y)$, then we need to consider partial derivatives: $\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}$.

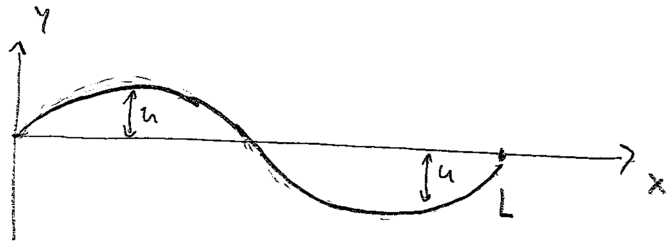
A Partial Differential Equation (PDE) is an equation that relates a function (the unknown) and its partial derivatives.

Ex:
$$\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} + z = 0 \quad (**)$$
In this example, the question to be answered is: can we find a function z (of x and y) that satisfies $(**)$? A function $z = z(x,y)$ (remark on notation) that satisfies $(**)$ is called a solution to $(**)$.

Remark: It's always easy to verify if a given function is a solution to a PDE: just plug it in and see if equality is satisfied.

Let's see other examples of PDEs.

Consider a string of length L with both ends attached. Suppose the string can vibrate with an amplitude u. (no motion in the horizontal direction)



u changes over space (x) and time (t), so u is a function of t and x:

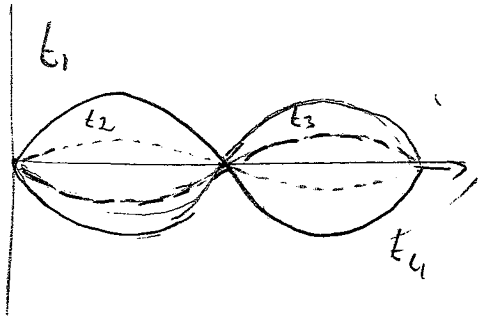
$$u = u(t,x).$$

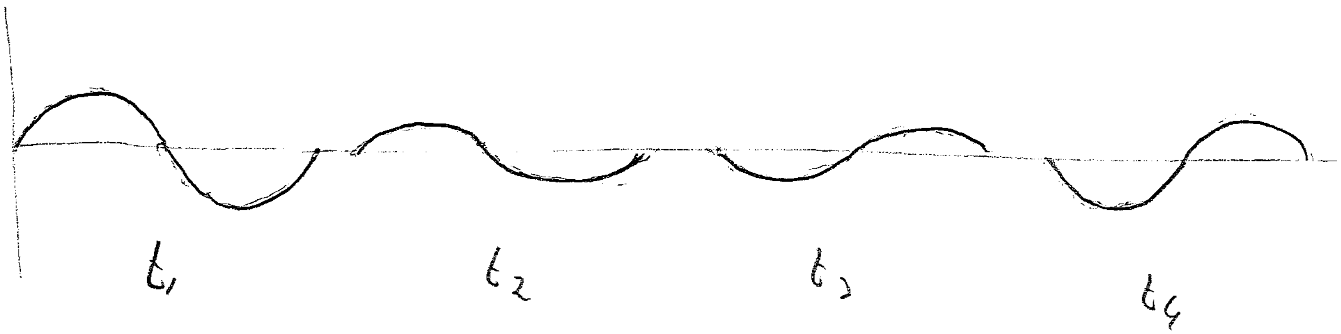
The <u>dynamics</u> (behavior over space and time) of the amplitude u is described by the <u>wave equation</u>:

$$\frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = 0$$

Above, c is a constant that depends on the nature of the string. If the string has constant density ρ (mass/length) and constant tension T (force), then $c = \sqrt{\frac{T}{\rho}}$. Notice that $c = \sqrt{\frac{N}{kg/m}} = \sqrt{\frac{N\,m}{kg}} = \sqrt{\frac{kg \cdot m/s^2 \cdot m}{kg}} = \sqrt{\frac{m^2}{s^2}} = \frac{m}{s}$ = units of velocity

We will see later on that c corresponds exactly to the speed of waves propagating along the string.





imagine water waves moving on the ocean surface.

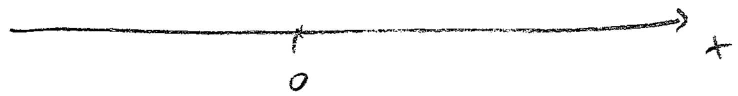
Notice that in this example, a solution u must satisfy:

$$u(t,0) = 0 \quad \text{and} \quad u(t,L) = 0$$

These are extra conditions that must be added to the PDE, and they are called boundary conditions (B.C.). A PDE + B.C is known as a boundary value problem

$$\begin{cases} \dfrac{\partial^2 u}{\partial t^2} - c^2 \dfrac{\partial^2 u}{\partial x^2} = 0 \\ u(t,0) = 0 = u(t,L) \end{cases}$$

As another example, consider an infinitely long rod (idealization), which we identify with the x-axis:



Let u denote the temperature on the rod. u can vary over space and time, so

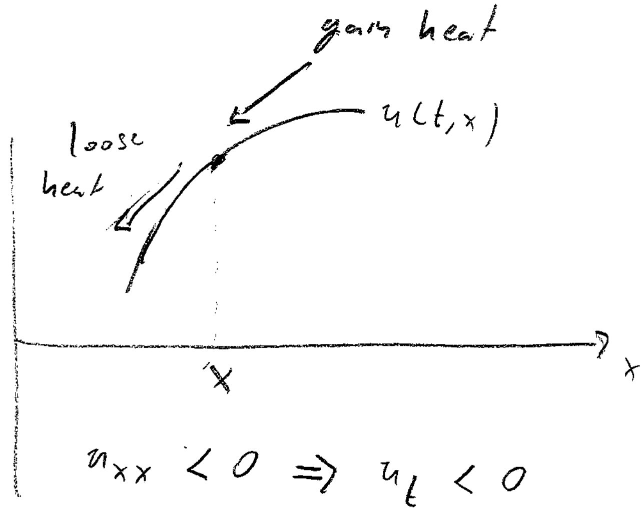
$u = u(t, x)$.

The temperature dynamics is described by the <u>heat equation</u>

$$u_t - k u_{xx} = 0 \quad , \quad \text{notation } u_t = \frac{\partial u}{\partial t} , \quad u_{xx} = \frac{\partial^2 u}{\partial x^2}$$

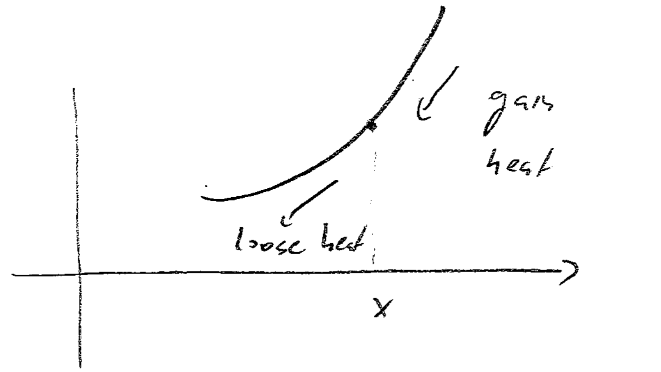
where $k > 0$ is a constant depending on the nature of the rod known as thermal diffusivity.

We can interpret this equation physically upon imagining the graph of u for a fixed time t



$u_{xx} < 0 \Rightarrow u_t < 0$

temperature at x will (instanteneously) decrease



$u_{xx} > 0 \Rightarrow u_t > 0$

temperature at x will (instanteneously) increase.

Suppose that u is a solution to the heat equation, and suppose that $t=0$ corresponds to "now". The temperature at later times, $u(t,x)$, will depend on the temperature at $t=0$. Thus we cannot completely calculate $u(t,x)$ without knowing $u(0,x)$. Mathematically, this means that the solution u contains <u>undetermined functions</u> that must be determined with the knowledge of $u(0,x)$ (compare to ODEs). Hence, to completely find u we must also be given

ODEs $\downarrow$ undetermined <u>constants</u>

$$u(0,x) = u_0(x)$$

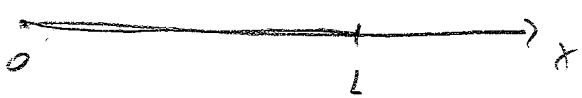
where u_0 is a known function of x only, called an <u>initial condition</u> (I.C.)

Thus, the complete problem to find u is

$$\begin{cases} u_t - k u_{xx} = 0 \\ u(0,x) = u_0(x) \end{cases}$$

A PDE + I.C. is know as an <u>initial value problem</u>

Suppose now that instead of an infinite rod we consider a rod of length L:



Assume that the temperature of the rod at the endpoints is kept constant by some external heat sources, so that $u(t,0) = T_1$, $u(t,L) = T_2$ where T_1 and T_2 are known numbers. These are boundary conditions, and in this case the problem of completely determining u becomes

$$\begin{cases} u_t - k u_{xx} = 0 \\ u(0,x) = u_0(x) \\ u(t,0) = T_1, \ u(t,L) = T_2 \end{cases}$$

A PDE + I.C. + B.C. is known as an initial-boundary value problem.

Notice that we can choose different boundary conditions. For instance, suppose that the rod endpoints are insulated, so that no heat can be transferred across the endpoints. In this case the rate of change of the temperature across the endpoints has to be zero:

$$u_x(t,0) = 0 \quad \text{and} \quad u_x(t,L) = 0$$

and the initial-boundary value problem is
$$\begin{cases} u_t - k u_{xx} = 0 \\ u(0,x) = u_0(x) \\ u_x(t,0) = 0 = u_x(t,L) \end{cases}$$

Going back to the wave equation, we notice that to completely determine u we also need to be given the initial configuration of the string, which in this case is not only $u(0,x)$ but also $\partial_t u(0,x)$ (by Newton's law, to determine the motion of an object, we need to provide initial position and initial velocity).

The complete initial-boundary value problem for the wave equation is

$$\begin{cases} \dfrac{\partial^2 u}{\partial t} - c^2 \dfrac{\partial^2 u}{\partial x^2} = 0 \\ u(0,x) = u_0(x), \quad \dfrac{\partial u}{\partial t}(0,x) = u_1(x) \\ u(t,0) = 0 = u(t,L) \end{cases}$$

where u_0 and u_1 are known functions (of x only)

Let's see another example

Suppose we want to study the electric and magnetic field in a region of space. Let E and B denote the electric and magnetic fields, respectively. These are vectors:

$$E = (E^1, E^2, E^3) \quad \text{and} \quad B = (B^1, B^2, B^3)$$

(If we used arrows for vectors $\vec{E} = E^1 \vec{i} + E^2 \vec{j} + E^3 \vec{k}$)

(Notice that we do not employ any special notation for vectors, ie, $\vec{E}$, $\underline{\underline{E}}$, etc)

E and B depend on time, which we denote by t, and space, which we denote by (x,y,z): So E and B are vector valued functions or vector functions: or vector fields:

$$E = E(t,x,y,z) \quad \text{and} \quad B = B(t,x,y,z)$$

(ie, $E^i = E^i(t,x,y,z)$; $B^i = B^i(t,x,y,z)$; Remark on notation)

Suppose that there exists a distribution of charge $\rho = \rho(t,x,y,z)$ and and electric current $J = (J^1, J^2, J^3)$, $J = J(t,x,y,z)$ in space

The behaviour of E and B in the presence of ρ and J is described by Maxwell's equations,

$$\nabla \cdot E = \frac{\rho}{\varepsilon_0}$$ (Gauss law: the electric flux leaving a volume is proportional to the charge inside)

$$\nabla \cdot B = 0$$ (There are no magnetic monopoles; compare to Gauss law)

$$\frac{\partial B}{\partial t} = -\nabla \times E$$ (Faraday's law: voltage induced in a closed circuit is proportional to the rate of change of the magnetic flux it encloses - connection via Stoke's theorem)

$$\overbrace{\frac{\partial E}{\partial t}}^{\text{displacement current}} = \frac{1}{\mu_0 \varepsilon_0} \nabla \times B - \frac{1}{\varepsilon_0} J$$ (Ampère's law: the magnetic field induced around a closed loop is proportional to the electric current plus displacement current it encloses)

where: ε_0 = permittivity of free space = $8.85 \cdot 10^{-12}$ F/m (farads per meter)

μ_0 = permeability of free space = $4\pi \cdot 10^{-7}$ N/A^2 (newtons per Ampère square)

$\nabla \cdot$ = divergence = div, $\nabla \cdot (\sigma^1, \sigma^2, \sigma^3) = \frac{\partial \sigma^1}{\partial x} + \frac{\partial \sigma^2}{\partial y} + \frac{\partial \sigma^3}{\partial z}$

$\nabla \times$ = curl or rotational = curl, $\nabla \times (\sigma^1, \sigma^2, \sigma^3) = \det \begin{bmatrix} \vec{i} & \vec{j} & \vec{k} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ \sigma^1 & \sigma^2 & \sigma^3 \end{bmatrix}$

$$= \left(\frac{\partial \sigma^3}{\partial y} - \frac{\partial \sigma^2}{\partial z}, \frac{\partial \sigma^1}{\partial z} - \frac{\partial \sigma^3}{\partial x}, \frac{\partial \sigma^2}{\partial x} - \frac{\partial \sigma^1}{\partial y} \right)$$

The first two equations can be written as

$$\frac{\partial E^1}{\partial x} + \frac{\partial E^2}{\partial y} + \frac{\partial E^3}{\partial z} = \frac{\rho}{\varepsilon_0} \quad \text{and} \quad \frac{\partial B^1}{\partial x} + \frac{\partial B^2}{\partial y} + \frac{\partial B^3}{\partial z} = 0$$

The last two equations are vector equations and each corresponds to three scalar equations:

$$\frac{\partial B^1}{\partial t} = -\left(\frac{\partial E^3}{\partial y} - \frac{\partial E^2}{\partial z}\right), \quad \frac{\partial B^2}{\partial t} = -\left(\frac{\partial E^1}{\partial z} - \frac{\partial E^3}{\partial x}\right), \quad \frac{\partial B^3}{\partial t} = -\left(\frac{\partial E^2}{\partial x} - \frac{\partial E^1}{\partial y}\right)$$

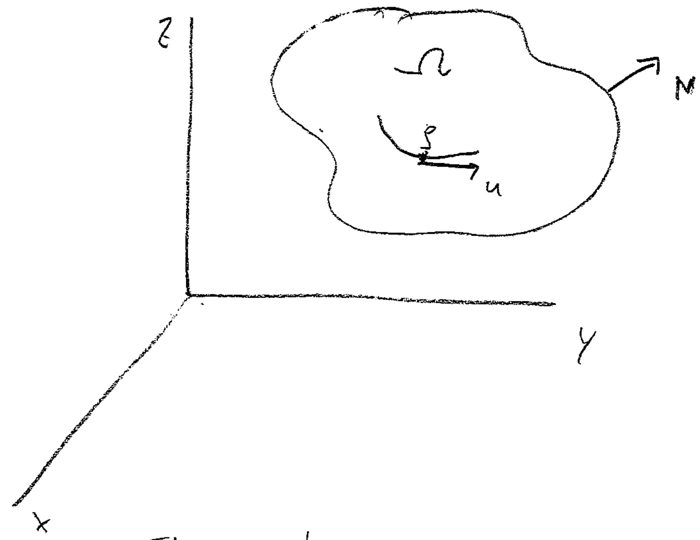
(recall $\frac{\partial}{\partial t}(v^1, v^2, v^3) = \left(\frac{\partial v^1}{\partial t}, \frac{\partial v^2}{\partial t}, \frac{\partial v^3}{\partial t}\right)$; e.g. curves in multivariable calculus)

$$\frac{\partial E^1}{\partial t} = \frac{1}{\mu_0 \varepsilon_0}\left(\frac{\partial B^3}{\partial y} - \frac{\partial B^2}{\partial z}\right) - \frac{1}{\varepsilon_0} J^1, \quad \frac{\partial E^2}{\partial t} = \frac{1}{\mu_0 \varepsilon_0}\left(\frac{\partial B^1}{\partial z} - \frac{\partial B^3}{\partial x}\right) - \frac{1}{\varepsilon_0} J^2, \quad \frac{\partial E^3}{\partial t} = \frac{1}{\mu_0 \varepsilon_0}\left(\frac{\partial B^2}{\partial x} - \frac{\partial B^1}{\partial y}\right) - \frac{1}{\varepsilon_0} J^3$$

Suppose we know ρ and J. Then, to study E and B, we would like to solve the above equations, finding the vector fields E and B.

Each of Maxwell's equations are Partial Differential Equations (PDEs), and together they form a system of PDEs

Another example: consider an $\overbrace{\text{inviscid}}^{\text{no viscosity}}$ fluid moving in a region Ω of space (e.g., water, viscous: honey)



Let u denote the fluid's velocity and ρ its density. Both quantities vary over space and time, u is a vector and ρ a scalar. Thus

$$u = u(t,x,y,z), \quad u = (u^1, u^2, u^3)$$

$$\rho = \rho(t,x,y,z).$$

The dynamics of the fluid is described by the Euler equations,

$$\begin{cases} \partial_t u + (u\cdot\nabla) u = -\frac{1}{\rho}\nabla p \\ \partial_t \rho + \operatorname{div}(\rho u) = 0 \\ \left.\begin{array}{l} u(0,x,y,z) = u_0(x,y,z) \\ \rho(0,x,y,z) = \rho_0(x,y,z) \end{array}\right\} \text{I.C.} \\ u\cdot N = 0 \text{ on } \underbrace{\partial\Omega}_{\text{boundary of } \Omega} \ \} \text{ B.C.} \end{cases}$$

Here, p is a function describing the pressure of the fluid. It depends on ρ and its specific form depends on the nature of the fluid, e.g. $p = A\rho^r$, A, r = constant.

$(u\cdot\nabla)u$ means the following

$$u\cdot\nabla = (u^1,u^2,u^3)\cdot\left(i\frac{\partial}{\partial x} + j\frac{\partial}{\partial y} + k\frac{\partial}{\partial y}\right)$$

$$= u^1\frac{\partial}{\partial x} + u^2\frac{\partial}{\partial y} + u^3\frac{\partial}{\partial z}$$

It is convenient to denote $(x, y, z) = (x^1, x^2, x^3)$ (remark on notation). Then

$$u \cdot \nabla = \sum_{i=1}^{3} u^i \frac{\partial}{\partial x^i} \quad \text{so that} \quad (u \cdot \nabla) u = \sum_{i=1}^{3} u^i \frac{\partial}{\partial x^i} (u^1, u^2, u^3)$$

$$= \left(\sum_{i=1}^{3} u^i \frac{\partial u^1}{\partial x^i}, \sum_{i=1}^{3} u^i \frac{\partial u^2}{\partial x^i}, \sum_{i=1}^{3} u^i \frac{\partial u^3}{\partial x^i} \right)$$

The first equation is a vector equation corresponding to three scalar equations

$$\partial_t (u^1, u^2, u^3) + \sum_{i=1}^{3} u^i \frac{\partial}{\partial x^i} (u^1, u^2, u^3) = -\frac{1}{\rho} \left(\frac{\partial p}{\partial x^1}, \frac{\partial p}{\partial x^2}, \frac{\partial p}{\partial x^3} \right)$$

so

$$\frac{\partial u^j}{\partial t} + \sum_{i=1}^{3} u^i \frac{\partial u^j}{\partial x^i} = -\frac{1}{\rho} \frac{\partial p}{\partial x^j}, \quad j = 1, 2, 3.$$

The unknowns in Euler's equations are u and ρ, and we have a system of PDEs.

<u>Notation</u> $\frac{\partial}{\partial x^i} = \partial_i$. We also adopt the <u>sum convention</u>: in a sum (like $\sum_{i=1}^{3} u^i \frac{\partial}{\partial x^i}$) whenever an index appears twice (i in $\sum_{i=1}^{3} u^i \frac{\partial}{\partial x^i}$), the summation symbol is omitted and the sum is implicitly understood. So: $u^i \frac{\partial u^j}{\partial x^i} = \sum_{i=1}^{3} u^i \frac{\partial u^j}{\partial x^i}$ or $u^i \partial_i u^j = \sum_{i=1}^{3} u^i \partial_i u^j$

and $\partial_t u^j + u^i \partial_i u^j = -\frac{1}{\rho} \partial_j p$

Another example is the Schrödinger equation:

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \Delta \Psi + V \Psi$$

which is the fundamental equation of quantum mechanics. Here, the unknown function is $\Psi = \Psi(t, x, y, z)$, called the wave function, $i = \sqrt{-1}$, $V = V(x,y,z)$, is the potential function (potential energy) which is a known function depending on the system we want to study (e.g. atom, molecule), m is the mass of the particle being described (e.g. electron) and $\hbar$ is a constant called Planck's constant.

$$\hbar = 1.05 \cdot 10^{-34} \text{ J}\cdot\text{s}$$

$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ (so $\Delta \Psi = \frac{\partial^2 \Psi}{\partial x^2} + \frac{\partial^2 \Psi}{\partial y^2} + \frac{\partial^2 \Psi}{\partial z^2}$) is called the Laplacian operator, also written as ∇^2. Notice that Ψ is in general a complex function. The interpretation of Ψ is that $\iiint_R |\Psi(t,x,y,z)|^2 \, dx\, dy\, dz$ is the probability of finding the particle (e.g. electron) in the region R of space.

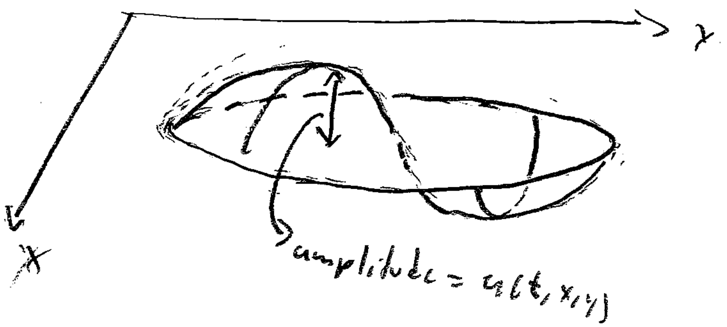
The wave and heat equations can be generalized to two and three dimensions:

$$\frac{\partial^2 u}{\partial t^2} - c^2 \Delta u = 0 \quad \text{and} \quad \frac{\partial u}{\partial t} - k \Delta u = 0$$

respectively, where

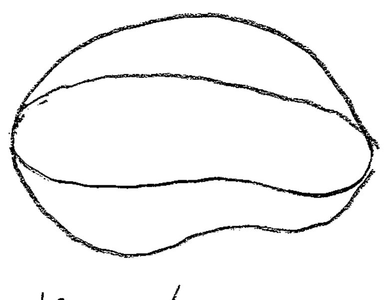
$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ in 2 dimensions and $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ in 3 dimensions.

Physical examples are a vibrating membrane (wave, 2d) and the temperature or a solid (heat, 3d)



wave 2d

(comment: 2d, (x,y), the z-direction is u



temperature u(t,x,y,z)

The equation

$$\Delta u = 0$$

is called Laplace's equation; it corresponds to stationary solutions of the wave and heat equations.

A more complicated example is the following.

The dynamics of gravity are described by Einstein's equations of general relativity:

$$-\frac{1}{2} g^{\mu\nu} \partial_{\mu\nu} g_{\alpha\beta} - \frac{1}{2} g^{\mu\nu} \partial_{\alpha\beta} g_{\mu\nu} + \frac{1}{2} g^{\mu\nu} \partial_{\alpha\mu} g_{\beta\nu} + \frac{1}{2} g^{\mu\nu} \partial_{\beta\mu} g_{\alpha\nu} + F_{\alpha\beta} = T_{\alpha\beta}, \quad (*)$$

where: the indices α, β, μ, ν vary from 0 to 3; we have four independent variables x^0, x^1, x^2, x^3, with $x^0 = t$, $(x^1, x^2, x^3) = (x, y, z)$. For each pair of indices, e.g. $\mu\nu$, $\partial_{\mu\nu} = \frac{\partial^2}{\partial x^\mu \partial x^\nu}$ (e.g., $\partial_{02} = \frac{\partial^2}{\partial x^0 \partial x^2} = \frac{\partial^2}{\partial t \partial y}$). The sum convention is adopted, e.g. $g^{\mu\nu} \partial_{\alpha\mu} g_{\beta\nu} = \sum_{\mu=0}^{3} \sum_{\nu=0}^{3} g^{\mu\nu} \partial_{\alpha\mu} g_{\beta\nu}$.

The functions $g_{\alpha\beta}$, $\alpha, \beta = 0, \dots, 3$, correspond to the "gravitational field". Each $g_{\alpha\beta}$ is a function of (x^0, x^1, x^2, x^3), $g_{\alpha\beta} = g_{\alpha\beta}(x^0, x^1, x^2, x^3)$, and they are the unknown functions to be found.

Notice that there are 16 $g_{\alpha\beta}$'s, and that for each pair α, β, we have a different equation (*), thus there are also 16 equations and Einstein's equations are a system of PDEs.

The functions $g^{\alpha\beta}$, $\alpha, \beta = 0, .., 3$, are the entries of the matrix of the $g_{\alpha\beta}$'s.

$F_{\alpha\beta}$ is an expression involving first and zero order derivatives of the functions $g_{\mu\nu}$.

$T_{\alpha\beta}$ is a known matrix ($\alpha, \beta = 0, .., 3$) called the stress-energy tensor. It contains information about the distribution of matter and energy.

Formal definition of PDEs

Let us first define PDEs for a function of two variables.

Def. A first order PDE for a function $u = u(x,y)$ (the unknown) is an equation that can be written as

$$F(x, y, u(x,y), u_x(x,y), u_y(x,y)) = 0$$

or in simpler form

$$F(x, y, u, u_x, u_y) = 0 \quad (*)$$

for some function F defined on a subset of $\mathbb{R}^5$, i.e., $F : D \to \mathbb{R}$, where $D \subseteq \mathbb{R}^5$. A solution to a PDE is a function u that satisfies the equality $(*)$.

Ex: Consider $u_x + u_y = 0$. This can be written as $F(x,y,u,u_x,u_y) = 0$ where $F(p_1, p_2, p_3, p_4, p_5) = p_4 + p_5$.

Ex: Consider $(u_x)^2 + x u_x + y u_y = 0$. This can be written as $F(x,y,u,u_x,u_y) = 0$ where $F(p_1, p_2, p_3, p_4, p_5) = p_4^2 + p_1 p_4 + p_2 p_5$

Remarks on terminology and notation

Recall that $\mathbb{R}^5 = \mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}$ and more generally $\mathbb{R}^n = \underbrace{\mathbb{R}\times\cdots\times\mathbb{R}}_{n \text{ times}}$.

$D \subseteq \mathbb{R}^5$ means that D is a subset of $\mathbb{R}^5$ (possibly $D=\mathbb{R}^5$), and $F: D \to \mathbb{R}$ means that F is a function with domain D and co-domain $\mathbb{R}$ (thus, for each quintuple $(p_1, p_2, p_3, p_4, p_5) \underset{\text{belongs}}{\in} D$, $F(p_1, p_2, p_3, p_4, p_5)$ is a real number, i.e., $F(p_1, p_2, p_3, p_4, p_5) \in \mathbb{R}$).

Notice that when we write a function F for a given (first order) PDE, it is important (i.e, p_1 matches x, p_2 matches y, p_3 matches u, etc.

Above we defined first order PDEs. First order meant that the highest derivative appearing is of order one.

Def. A <u>second order PDE</u> for the a function $u=u(x,y)$ (the unknown) is an equation that can be written as

for most functions of interest, $u_{xy} = u_{yx}$, so we could have omitted u_{yx}, in which case $D \subseteq \mathbb{R}^8$, but it is convenient to write all derivatives.

$$F(x, y, u, u_x, u_y, u_{xx}, \overbrace{u_{xy}, u_{yx}}, u_{yy}) = 0$$

for some function $F: D \to \mathbb{R}$, where $D \subseteq \mathbb{R}^9$

Ex: $u_{tt} - c^2 u_{xx} = 0$ can be written as $F(p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8, p_9)$ where $F(p_1, \dots, p_9) = p_9 - c^2 p_6$ (where we denote y by t and $u = u(x,t)$)

Ex: $u_t - k u_{xx} = 0$ can be written as $F(p_1, \dots, p_8, p_9)$

$F(p_1, \dots, p_8) = p_5 - k p_6$.

We can also define first and second order PDEs for functions of more variables. For example, a first order PDE for a function $u = u(x,y,z)$ is

$$F(x, y, z, u, u_x, u_y, u_z) = 0$$

while a second order PDE for $u(x,y,z)$ is

$$F(x, y, z, u, u_x, u_y, u_z, u_{xx}, u_{xy}, u_{xz}, u_{yy}, u_{yz}, u_{zz}) = 0$$

We will now give the general definition of a PDE for a function of an arbitrary number of variables and of arbitrary order.

Some notations, terminology, and conventions

We will denote $\mathbb{R}^n = \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{n\text{-times}}$. Thus an element $x \in \mathbb{R}^n$ is an ordered n-tuple $x = (x^1, x^2, \ldots, x^n)$. Notice that we denote the components of x by superscripts (but sometimes subscripts are also used, $x = (x_1, \ldots, x_n)$). We think of elements of $\mathbb{R}^n$ as vectors so that the usual vector operations (addition, multiplication by scalars, etc) are defined, e.g.

$$x + y = (x^1, x^2, \ldots, x^n) + (y^1, y^2, \ldots, y^n) = (x^1+y^1, x^2+y^2, \ldots, x^n+y^n)$$

We will not employ any special notation for vectors (such as $\vec{x}$ etc).

When $n=2$ or $n=3$, we sometimes use (x,y) and (x,y,z) rather than (x^1, x^2) and (x^1, x^2, x^3), respectively. The value of n is sometimes called the dimension of space ($\mathbb{R}^2$ is two-dimensional, $\mathbb{R}^3$ is three-dimensional).

In many scenarios, we will take x to be an independent variable, so that calculus operations such as differentiation, div, curl (when $n=3$) are defined, e.g.

$\operatorname{div} x = \frac{\partial x^1}{\partial x^1} + \cdots + \frac{\partial x^n}{\partial x^n} = n$. The components x^i, $i=1,\ldots,n$ are also called coordinates and for them calculus operations also apply, e.g., $\frac{\partial x^1}{\partial x^2} = 0$, $\frac{\partial x^3}{\partial x^3} = 1$.

Def. A PDE of order m for a function u of n variables, is an equation in u

$$F\Big(x_1, x_2, \dots, x_n, u, \frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}, \dots, \frac{\partial u}{\partial x_n}, \frac{\partial^2 u}{\partial x_1^2}, \frac{\partial^2 u}{\partial x_1 \partial x_2}, \dots, \frac{\partial^2 u}{\partial x_1 \partial x_n}, \frac{\partial^2 u}{\partial x_2^2}, \dots,$$

$$\frac{\partial^2 u}{\partial x_n^2}, \text{ third order derivatives}, \dots, \frac{\partial^m u}{\partial x_1^m}, \frac{\partial^m u}{\partial x_1 \underbrace{\partial x_2 \cdots \partial x_2}_{m-1 \text{ times}}}, \dots, \frac{\partial^m u}{\partial x_n^m}\Big) = 0$$

where $F : D \to \mathbb{R}$ with

$$D \subseteq \underbrace{\mathbb{R}^n}_{\substack{\text{variables}\\ x_1, \dots, x_n}} \times \underbrace{\mathbb{R}}_{\substack{\text{function}\\ u}} \times \underbrace{\mathbb{R}^n}_{\substack{1^{st} \text{ partial}\\ \text{derivatives}}} \times \underbrace{\mathbb{R}^{n^2}}_{\substack{2^{nd} \text{ partial}\\ \text{derivatives}}} \times \underbrace{\mathbb{R}^{n^3}}_{\substack{3^{rd} \text{ partial}\\ \text{derivatives}}} \times \dots \times \underbrace{\mathbb{R}^{n^{m-1}}}_{\substack{\text{partial derivatives}\\ \text{of order } m-1}} \times \underbrace{\mathbb{R}^{n^m}}_{\substack{\text{partial derivatives}\\ \text{of order } m.}}$$

Notation We denote by $D^k u$ the set of all k^{th} order $\overset{\text{Remark on terminology}}{(\text{partial})}$ derivatives of u.

Thus the above can be written as $F(x, u, Du, D^2u, \dots, D^{m-1}u, D^m u) = f$

Def. A solution for a PDE $F(x, u, \dots, D^m u) = 0$ is a function that satisfies the equality

Terminology. In $F(x, u, \dots, D^m u) = 0$ we also refer to $u = u(x^1, \dots, x^n)$ as the dependent variable and to $x = (x^1, \dots, x^n)$ as the independent variable.

Linear vs non-linear PDEs

Given a PDE $F(x, u, Du, \dots, D^m u) = 0$ we can always write

$$F(x, u, Du, \dots, D^m u) = F_H(x, u, Du, \dots, D^m u) + f(x)$$

where every term in F_H contains u or its derivatives, and f is a function of x only.

EX $u_x^2 + u_y + xy = 0$, $F(p_1, p_2, p_3, p_4, p_5) = p_4^2 + p_4 + p_1 p_2$

$F_H(p_1, p_2, p_3, p_4, p_5) = p_4^2 + p_5$, $f(p_1, p_2) = p_1 p_2$, or

$F_H(u, u_x, u_y) = u_x^2 + u_y$, $f(x, y) = xy$

EX: $x u_x + y u_y + u - xy = 0$, $F(p_1, p_2, p_3, p_4, p_5) = p_1 p_4 + p_2 p_5 + p_3 - p_1 p_2$

so $F_H(p_1, p_2, p_3, p_4, p_5) = p_1 p_4 + p_2 p_5 + p_3$ and $f(p_1, p_2) = -p_1 p_2$

or $F_H(x, y, u, u_x, u_y) = x u_x + y u_y + u$, and $f(x, y) = -xy$.

Def We call $F_H(x,u,Du,\dots,D^m u)$ the <u>homogeneous part</u> of the PDE $F(x,u,Du,\dots,D^m u)=0$. The PDE is called <u>homogeneous</u> if $F=F_H$ (so $f=0$) and non-homogeneous otherwise.

Ex: $xu_x+yu_y+u=0$ is homogeneous, $u_x^2+u_y=x$ is non-homogeneous.

Def. A PDE $F(x,u,Du,\dots,D^m u)=0$ is called <u>linear</u> if F_H can be written as

$$F_H(x,u,Du,\dots,D^m u)=F_0(x,u)+F_1(x,Du)+\dots+F_m(x,D^m u)$$

where each F_i is linear on its second argument, i.e., for any functions u and v and any constants a and b:

$$F_0(x,au+bv)=aF_0(x,u)+bF_0(x,v),$$
$$F_1(x,aDu+bDv)=aF_1(x,Du)+bF_1(x,Dv)$$
$$\vdots$$
$$F_m(x,aD^m u+bD^m v)=aF_m(x,D^m u)+bF_m(x,D^m v).$$

The PDE is called <u>non-linear</u> otherwise.

<u>Remark on notation</u>: The above means, for fixed k, $F_k(x,D^k u)=$ sum of terms linear on each partial derivative of order k, e.g. $F_1(x,Du)=F_{11}(x,\frac{\partial u}{\partial x_1})+F_{12}(x,\frac{\partial u}{\partial x_2})+\dots+F_{1n}(x,\frac{\partial u}{\partial x_n})$

<u>Remark</u> Notice that it follows that

$$F_H(x,au+bv,aDu+bDv,\dots,aD^m u+bD^m v)=aF(x,u,\dots,D^m u)+bF(x,v,\dots,D^m v)$$

if the PDE is linear.

Ex: The PDE $x u_x + y u_y + u - xy = 0$ is a (non-homogeneous) linear PDE

$$F_H(x, u, u_x, u_y) = x^2 u_x + y^2 u_y + u$$

$$F_H(x, \underbrace{au+bv}_{\substack{\text{replace}\\ u \text{ by}\\ au+bv}}, \underbrace{au_x+bv_x}_{\substack{\text{replace}\\ u_x \text{ by}\\ au_x+bv_x}}, \underbrace{au_y+bv_y}_{\substack{\text{replace}\\ u_y \text{ by}\\ au_y+bv_y}}) = \overbrace{x^2(au_x+bv_x)}^{F_x(x,y,au_x+bv_x)} + \overbrace{y^2(au_y+bv_y)}^{F_y(x,y,au_y+bv_y)} + \overbrace{au+bv}^{F_0(x,y,au+bv)}$$

$$= a x^2 u_x + b x^2 v_x + a y^2 u_y + b y^2 v_y + au + bv$$

$$= a(x^2 u_x + y^2 u_y + u) + b(x^2 v_x + y^2 v_y + v)$$

$$= a F_H(x, u, u_x, u_y) + b F_H(x, v, v_x, v_y)$$

Ex: The PDE $u_x^2 + u_y = 0$ is a (homogeneous) non-linear PDE

$$F_H(x, u, u_x, u_y) = u_x^2 + u_y$$

$$F_H(x, au+bv, au_x+bv_x, au_y+bv_y) = (au_x+bv_x)^2 + au_y+bv_y = a^2u_x^2 + b^2v_x^2 + ab\,u_x u_y + au_y + bv_y$$

$$\neq a(u_x^2 + u_y) + b(v_x^2 + v_y)$$

There is another way of thinking of linear PDEs, as follows. Consider F_H as an *operator*. An operator is an operation that assigns to a given function another function. For instance, in the example $F_H(x,u,Du) = xu_x + yu_y + u$, F_H assigns to the function u the new function $xu_x + yu_y + u$. Similarly, in the example $F_H(x,u,Du) = u_x^2 + u_y$, F_H assigns to u the new function $u_x^2 + u_y$

Remark: Notice that differentiation and integration are also operators. For instance, $\frac{\partial}{\partial x^1}$ assigns to a given u the new function $\frac{\partial u}{\partial x^1}$. Similarly $\int dx$ assigns to a given $f(x)$ the new function $\int f(x)\,dx$.

Remark. If you think about, an operator is also a function (comment)

Thinking of F_H as an operator, we will write

$F_H(x,u,\ldots,D^m u) = Pu$, where we think of P as an operator and of Pu as P applied to u (the same way we think of $\frac{\partial}{\partial x^1}$ and of $\frac{\partial u}{\partial x^1}$)

We call P the operator associated with the PDE

Then, $F(x, u, \dots, D^m u) = 0$ is a linear equation if and only if P is a linear operator, i.e.,

$$P(au + bv) = aPu + bPv$$ for any two functions u, v and constants a, b such that both sides are well defined.

(Exercise)

<u>Remark</u> Notice that the concept of an operator is more general than that of the homogeneous part of a PDE, and we can talk about operators without having a PDE (e.g., $\frac{\partial}{\partial x}$).

<u>Terminology</u> We sometimes use the term $\underbrace{\text{differential operator}}_{}$ (mention the order of P) to emphasize that a certain operator is constructed out of derivatives. (as will be most operators in this course).

When an operator comes from a PDE ($P = F_H$), we call it the <u>operator associated</u> with the PDE.

A PDE $F(x, u, \dots, D^m u) = 0$ can equivalently be written as

$$Pu = f,$$

where f is a known function (equal to $-f$ of before)

Prop Let $u_1, \dots, u_k$ be solutions to the homogeneous linear PDE

$$Pu = 0$$

Then, the function $v = c_1 u_1 + c_2 u_2 + \dots + c_k u_k$ is also a solution, where $c_1, \dots, c_k$ are constants.

Proof We need to show that $Pv = 0$. Compute

$$Pv = P(c_1 u_1 + c_2 u_2 + \dots + c_k u_k) \underset{\text{by linearity}}{=} c_1 \underbrace{Pu_1}_{=0} + c_2 \underbrace{Pu_2}_{=0} + \dots + c_k \underbrace{Pu_k}_{=0} = 0 + \dots + 0 = 0.$$

$\boxtimes$

Notice that the result is not true for non-homogeneous linear PDEs.

Prop A first order PDE $Pu = f$ for $u = u(x,y)$ is linear if and only if it can be written as

$$a u_x + b u_y + c u = f$$

where $a, b,$ and c are known functions of x and y

proof) First, assume that the PDE is written as stated:

$a u_x + b u_y + c u = f$, and let's show that it is linear. In this case, $Pu = a u_x + b u_y + c u$, so

$$P(c_1 u + c_2 v) = a \frac{\partial}{\partial x}(c_1 u + c_2 v) + b \frac{\partial}{\partial y}(c_1 u + c_2 v) + c(c_1 u + c_2 v) = a Pu + b Pv.$$

Reciprocally, assume that the PDE is linear. We want to show that it can be written as stated.

By linearity $$F_H(x,y,u,u_x,u_y) = \underbrace{F_x(x,y,u_x) + F_y(x,y,u_y)}_{F_1(x,y,Du)} + F_0(x,y,u) \quad (F_x, F_y = \text{not derivatives})$$

Since $F_x(x,y,u_x)$ is a linear function of u_x, we must have $F_x(x,y,u_x) = a(x,y) u_x$ and similarly $F_y(x,y) = b(x,y) u_y$, $F_0(x,y,u) = c(x,y) u$, for some functions a, b, c.

Furthermore, since this holds for any function u, if we choose $u(x,y) = 1$, $F_H(x,y,1,0,0) = c(x,y)$.

Then, choosing $u(x,y) = x$:

$$F_H(x,y,x,1,0) = a(x,y) + c(x,y) \cdot x \implies a(x,y) = F_H(x,y,x,1,0) - \overbrace{c(x,y)}^{\text{already known}} x$$

and similarly $b(x,y) = F_H(x,y,y,0,1) - c(x,y) y$.

$\square$

We can use similar ideas to give the precise form of P for equations of higher order and of more variables. For example, a linear second order PDE for $u = u(x,y)$ can be written as

$$a^{xx} u_{xx} + a^{yy} u_{xy} + a^{xy} u_{xy} + b^x u_x + b^y u_y + cu = f$$

where $a^{xx}, \ldots, c$ are functions of x and y.

A linear first order PDE for $u = u(x^1, \ldots, x^n)$ can be written as

$$a^1 \frac{\partial u}{\partial x^1} + a^2 \frac{\partial u}{\partial x^2} + \cdots + a^n \frac{\partial u}{\partial x^n} + bu = f$$

where $a^1, \ldots, b$ are functions of $(x^1, \ldots, x^n)$.

A linear second order PDE for $u = u(x^1, \ldots, x^n)$ can be written as

$$\sum_{i,j=1}^{n} a^{ij} \partial^2_{ij} u + \sum_{i=1}^{n} b^i \partial_i u + cu = f$$

where $\partial^2_{ij} = \frac{\partial}{\partial x^i \partial x^j}$, $\partial_i = \frac{\partial}{\partial x^i}$, and the coefficients are functions of $(x^1, \ldots, x^n)$

<u>Notation</u> For coordinates $(x^1, \ldots, x^n)$, we set

$$\partial_i = \frac{\partial}{\partial x^i}, \quad \partial^2_{ij} = \partial_{ij} = \frac{\partial}{\partial x^i \partial x^j}, \quad \partial^3_{ijk} = \partial_{ijk} = \frac{\partial}{\partial x^i \partial x^j \partial x^k}$$

We also sometimes write ∂_{x^i}, $\partial_{x^i x^j}$ etc.

We also adopt the sum convention: when there is a sum over indices $i, j, k, \ldots$ etc ranging from 1 to n, and indices appear repeated in the sum, we omit the summation symbol and the sum is implicitly understood. Thus,

$$\sum_{i,j=1}^{n} a^{ij} \partial_{ij} = a^{ij} \partial_{ij}, \quad \sum_{i=1}^{n} b^i \partial_i = b^i \partial_i \quad \text{etc.}$$

Thus, a linear equation cannot have any non-linear function of u or its derivatives: $u \cdot u_x$, $\cos(u)$, u_y^2, etc. cannot be present.

Notice that the coordinates can have non-linear functions, eg

$$x^2 u_x + u = 0 \quad \text{is a linear eq.}$$

First order PDEs

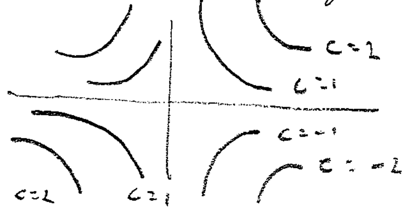
We will develop a general method for solving first order PDEs. In order to understand the intuition behind the method, let us work out some simple examples first.

Ex: Consider $a u_x + b u_y = 0$, where a and b are constants not both zero

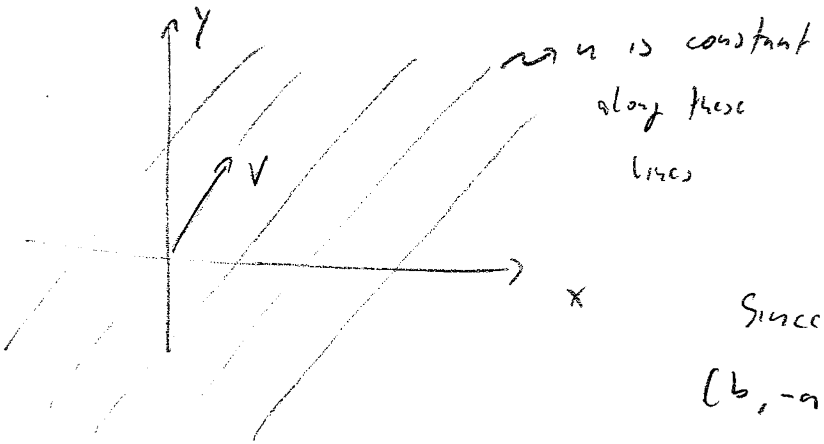
Notice that $a u_x + b u_y$ is the directional derivative of u in the direction of the vector $v = (a,b)$: $D_v u = \nabla u \cdot v = \left(\frac{\partial u}{\partial x} i + \frac{\partial u}{\partial y} j \right) \cdot (a i + b j) = a u_x + b u_y$.

Thus, $a u_x + b u_y = 0$ means that ∇u is orthogonal to v for any (x,y) where u is defined. But ∇u is orthogonal to the level curves of u: recall that the level curves of $g = g(x,y)$ are the sets $\Gamma_c = \{ (x,y) \in \mathbb{R}^2 \mid g(x,y) = c \}$. So, for each $c \in \mathbb{R}$, Γ_c is a curve on the plane along which g is constant. For example, if $g(x,y) = xy$, then $xy = c \Rightarrow y = \frac{c}{x}$

or $x = 0$
$y = 0$



Therefore, we know that u must be constant in the direction of V:



The line through (x_0, y_0) in the direction of V is $(r(t) = tv + r_0)$

$$(x, y) = t(a,b) + (x_0, y_0), \quad t \in \mathbb{R}$$

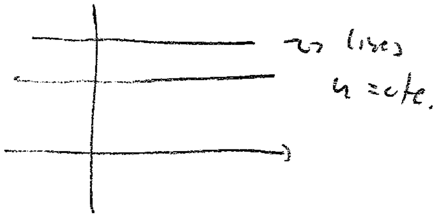
Since $(b, -a)$ is orthogonal to V, we have

$$(b,-a)\cdot(x,y) = (b,-a)\cdot(x_0,y_0) \Rightarrow \boxed{bx - ay = \text{constant}}$$

These are the lines in the direction of V (another way of seeing this: slope $= \frac{b}{a}$, $y = \frac{b}{a}x + \text{intercept}$). Therefore $u(x,y)$ depends only on $bx-ay$, i.e., $u(x,y)$ is a function of $bx-ay$, i.e. $\boxed{u(x,y) = f(bx-ay)}$, where f is any function of one-variable. For instance, take $f(w) = w^2$, so $u(x,y) = (bx-ay)^2$ solves the equation. Verify $u_x(x,y) = 2(bx-ay)\cdot b$, $u_y(x,y) = 2(bx-ay)(-a)$, so

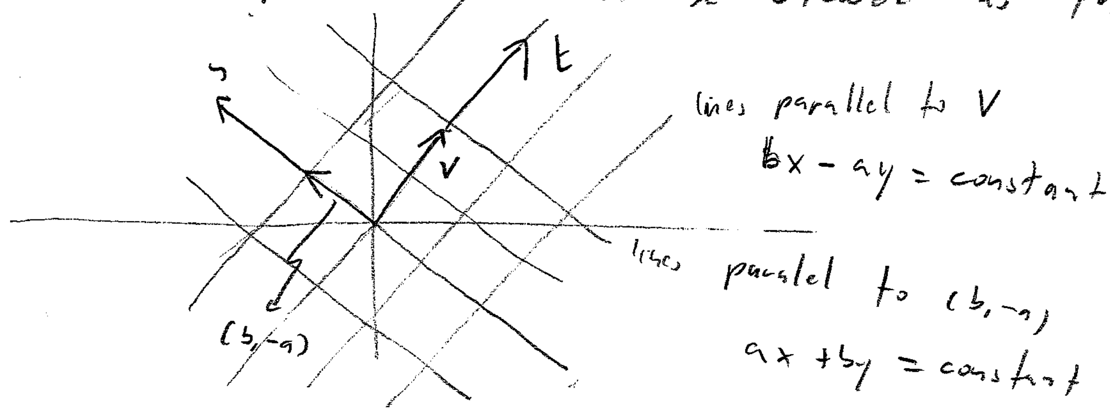
$$a u_x(x,y) + b u_y(x,y) = 2ab(bx-ay) - 2ab(bx-ay) = 0.$$

In the particular case when $b=0$, we have $a u_x = 0$. This means that u does not depend on x, so $u(x,y) = f(y)$ and u is constant along the lines $y = cte$.



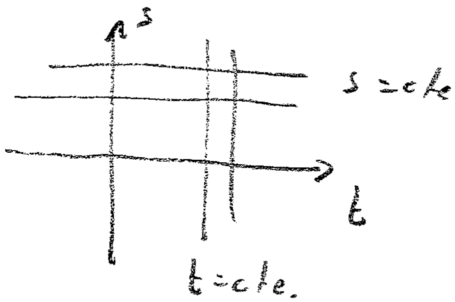
Let's see that the previous case can be viewed as this one.

Idea:



let $s = bx - ay$, so the lines parallel to V correspond to $s = cte$

$t = ax + by$, so the lines parallel to $(b,-a)$ correspond to $t = cte$



Let $\sigma(t,s) = u(x,y)$. (σ is u written in the variables (t,s)). Then

$$\frac{\partial u}{\partial x} = \frac{\partial \sigma}{\partial x} = \frac{\partial \sigma}{\partial t}\frac{\partial t}{\partial x} + \frac{\partial \sigma}{\partial s}\frac{\partial s}{\partial x} = \frac{\partial \sigma}{\partial t} a + \frac{\partial \sigma}{\partial s} b$$

$$\frac{\partial u}{\partial y} = \frac{\partial \sigma}{\partial y} = \frac{\partial \sigma}{\partial t}\frac{\partial t}{\partial y} + \frac{\partial \sigma}{\partial s}\frac{\partial s}{\partial y} = \frac{\partial \sigma}{\partial t} b - \frac{\partial \sigma}{\partial s} a$$

So $$0 = a u_x + b u_y = a^2 \frac{\partial \sigma}{\partial t} + ab\frac{\partial \sigma}{\partial s} + b^2 \frac{\partial \sigma}{\partial t} - ab\frac{\partial \sigma}{\partial s} = (a^2 + b^2)\frac{\partial \sigma}{\partial t} \overset{a^2+b^2 \neq 0}{\Rightarrow} \sigma = f(s)$$

So $u(x,y) = \sigma(t,\sigma) = f(s) = f(bx - ay)$.

Ex: Solve $4u_x - 3u_y = 0$ with initial condition $u(0,y) = y^3$.

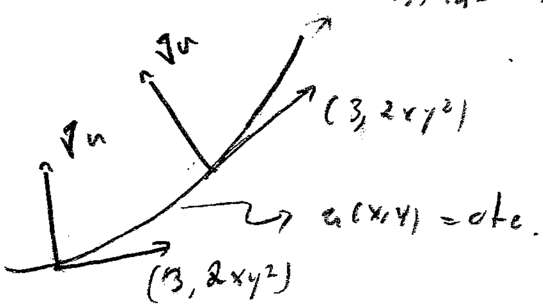
$u(x,y) = f(-3x - 4y)$. But $u(0,y) = f(-4y) = y^3$. Let $w = -4y$, so $y = -\frac{w}{4}$, and $f(-4y) = f(w) = y^3 = -\frac{w^3}{64}$, $f(w) = -\frac{w^3}{64}$ and $u(x,y) = \frac{(3x+4y)^3}{64}$.

Ex: Let us now investigate an equation where the coefficients are not constants. Consider (the linear PDE)

$3u_x + 2xy^2 u_y = 0$, which we can write as $\nabla u \cdot (3, 2xy^2) = 0$.

Thus, for each point (x,y), $\nabla u(x,y)$ is orthogonal to the vector $(3, 2xy^2)$.

Thus, u is constant along the curves that have $(3, 2xy^2)$ as tangent vector.



Hence, we need to find such curves. If the curve is written $y = y(x)$, then the slope of the tangent vector is $\frac{dy}{dx}$, which has to equal $\frac{2xy^2}{3}$:

$\frac{dy}{dx} = \frac{2xy^2}{3}$ $\Rightarrow$ equation for the curves along which u is constant.

$$\frac{dy}{y^2} = \frac{2}{3}x\,dx, \quad -\frac{1}{y^2} = \frac{1}{3}x^2 + C', \quad \text{so} \quad \frac{1}{3}x^2 + \frac{1}{y} = \text{constant} \quad (y \neq 0)$$

Therefore $u(x,y) = f\left(\frac{1}{3}x^2 + \frac{1}{y}\right)$.

Let's verify that this indeed solves the equation:

$$u_x(x,y) = f'\left(\frac{1}{3}x^2 + \frac{1}{y}\right) \cdot \frac{2}{3}x, \quad u_y(x,y) = f'\left(\frac{1}{3}x^2 + \frac{1}{y}\right)\left(-\frac{1}{y^2}\right)$$

$$3u_x + 2xy^2 u_y = 2x f'\left(\frac{1}{3}x^2 + \frac{1}{y}\right) + 2xy^2 f'\left(\frac{1}{3}x^2 + \frac{1}{y}\right)\left(-\frac{1}{y^2}\right) = 2x f'(\quad) - 2x f'(\quad) = 0.$$

<u>Upshot:</u> Given $a(x,y)u_x + b(x,y)u_y = 0$, the solution u will be given by $u(x,y) = f(A(x,y))$, where f is arbitrary and $A(x,y)$ is found by solving $\frac{dy}{dx} = \frac{b(x,y)}{a(x,y)}$ (so $A(x,y) = $ cste are solutions to $\frac{dy}{dx} = \frac{b}{a}$).

If further conditions are given, we can then determine f.

Notice that in the above examples we solve the PDE by solving certain ODEs. We will now generalize this idea for arbitrary first order PDEs.

The method of characteristics: This is a method for solving 1^{st} order PDEs upon transforming the PDE into a system of ODEs.

Consider the equation

$$a(x,y,u)\,u_x + b(x,y,u)\,u_y = c(x,y,u). \quad (*)$$

Ex:

$$x^2 u\, u_x + \cos xy\, e^u u_y = \frac{u^2}{1+x^2+y^2}$$

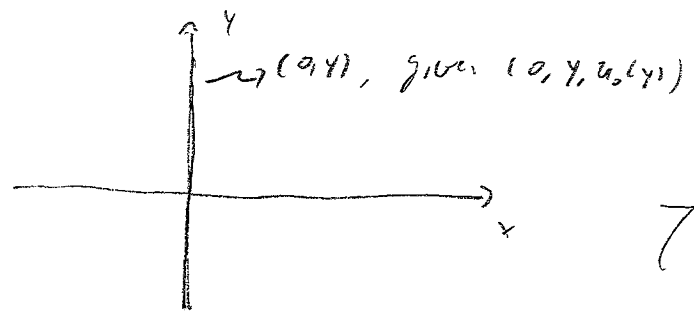
Equation $(*)$ is in general non-linear, but it is linear with respect to the derivatives of u, i.e.

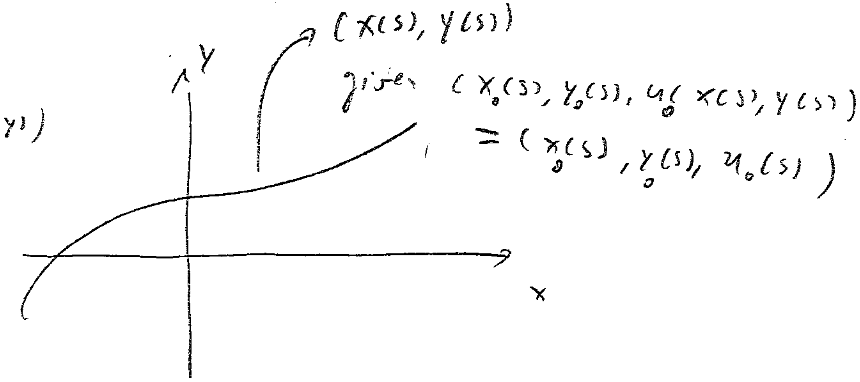
$$a(x,y,c_1u+c_2v)(c_1u+c_2v)_x = c_1 a(x,y,c_1u+c_2v)\,u_x + c_2\, a(x,y,c_1u+c_2v)\,v_x,$$

and similarly for b. Thus, we sometimes call $(*)$ a quasi-linear equation. When a, b and c do not depend on u, then $(*)$ is a linear equation.

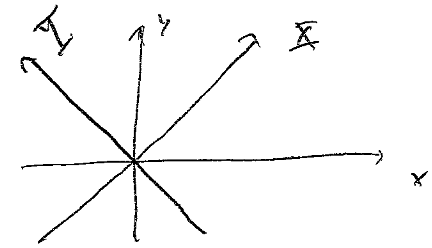
Consider first the linear case:

$a(x,y)\,u_x + b(x,y)\,u_y = c_0(x,y)\,u + f(x,y)$ (**). Assume we are given an initial condition for (**). The initial condition is typically of the form $u(0,y) = u_0(y)$, but it is convenient to consider that the initial condition is given along a curve on the xy-plane:





E.g. changing coordinates, $(0,y)$ becomes a curve $(\bar{X}(s), \bar{Y}(s))$:



So we write the initial condition as $\Gamma(s) = (x_0(s), y_0(s), u_0(s))$.

Write (**) as

$$(a, b, c_0 u + f)\cdot(u_x, u_y, -1) = 0$$

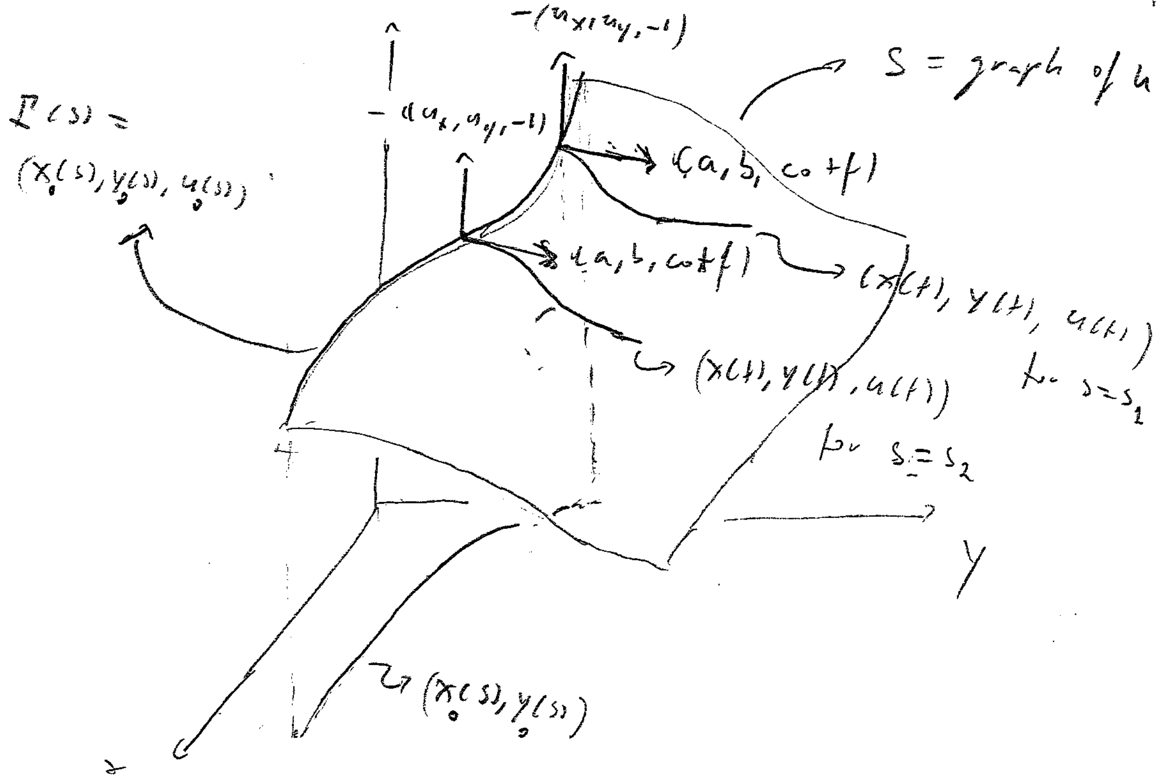
Recall (from multivariable calculus) that the graph of u can be viewed as a parametrized surface (or parametric surface) S: $r(x,y) = (x, y, u(x,y))$ and that $(u_x, u_y, -1)$ is normal to the surface S.

This means that $(a, b, cu+f)$ is tangent to S (for each (x,y))

Therefore, the equations:

$$\frac{dx}{dt}(t) = a(x(t), y(t)), \quad \frac{dy}{dt}(t) = b(x(t), y(t)), \quad \frac{du}{dt}(t) = c(x(t), y(t))\,u(t) + f(x(t), y(t)) \qquad (***)$$

defines curves $(x(t), y(t), u(t))$ that lie on S. The idea for solving the PDE is to solve $(***)$ with initial condition starting on $\Gamma(s)$



($(u_x, u_y, -1)$ points down, so we drew $-(u_x, u_y, -1)$). Notice that for each s we have a different curve, so we write

$x = x(t,s)$, $y = y(t,s)$

$u = u(t,s)$.

Notice that for each s parametrizing the initial conditions we have a different curve $(x(t), y(t), u(t))$, thus we write:

$$x = x(t,s), \quad y = y(t,s), \quad u = u(t,s)$$

Thus, we seek to solve

$$(G) \begin{cases} \dot{x}(t,s) = a(x(t,s), y(t,s)) \\ \dot{y}(t,s) = b(x(t,s), y(t,s)) \\ \dot{u}(t,s) = c(x(t,s), y(t,s))\, u(t,s) + f(x(t,s), y(t,s)) \end{cases}$$

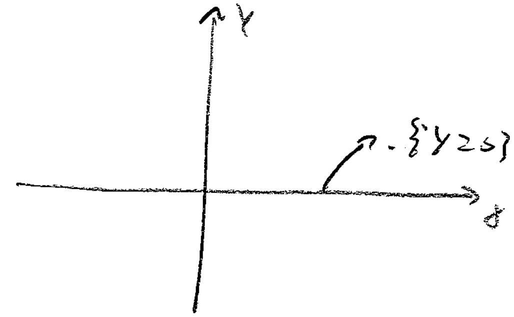
with initial condition

$$x(0,s) = x_0(s), \quad y(0,s) = y_0(0,s), \quad u(0,s) = u_0(s),$$

where $\bullet$ means derivative with respect to t. Notice that the parameter s only enters in the initial condition, but we need to consider it when solving (G) (see ex. below)

The eq. (G) are called <u>characteristic equations</u> and its solutions <u>characteristic curves</u>

Ex: Solve $u_x + u_y = 2$ with $u(x,0) = x^2$



The characteristic eq. are $(a=1, b=1)$

$$\dot{x} = 1, \quad \dot{y} = 1, \quad \dot{u} = 2 .$$

We need to parametrize the initial condition. We can choose $(x,0) = (s,0)$, so $x(0,s) = s$, $y(0,s) = 0$, $u(0,s) = s^2$. Solving the eq:

$$x(t,s) = t + f_1(s), \quad y(t,s) = t + f_2(s), \quad u(t,s) = 2t + f_3(s).$$

Using the initial conditions $x(0,s) = s = f_1(s)$, $y(0,s) = 0 = f_2(s)$, $u(0,s) = s^2 = f_3(s)$, thus

$$x(t,s) = t+s, \quad y(t,s) = t, \quad u(t,s) = 2t+s^2.$$

$r(t,s) = (x(t,s), y(t,s), u(t,s))$ is a parametric surface corresponding to the graph of u, so it is one form of representing the solution. However, we would like to have $u = u(x,y)$, so we put $y = t$, $s = x - t = x - y$, so that

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

Now let's consider the quasilinear case:

$$a(x,y,u)\,u_x + b(x,y,u)\,u_y = c(x,y,u)$$

which we can write as

$$(a,b,c)\cdot(u_x,u_y,-1) = 0$$

Thus, as in the linear case, (a,b,c) is tangent to the parametrized surface S given by the graph of (although now the vector (a,b,c) depends not only on (x,y) but also on $u(x,y)$)

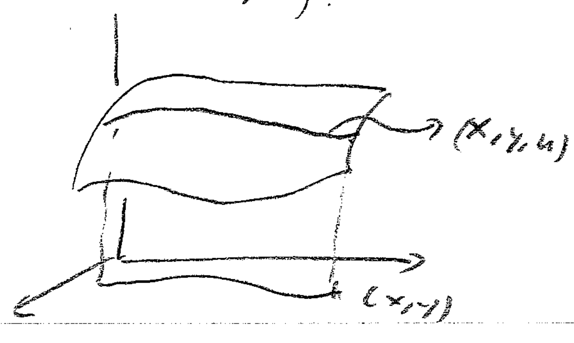
Thus we can again write ODE for the curves on S tangent to (a,b,c):

$$\begin{cases} \dot{x}(t,s) = a(x(t,s), y(t,s), u(t,s)) \\ \dot{y}(t,s) = b(x(t,s), y(t,s), u(t,s)) \\ \dot{u}(t,s) = c(x(t,s), y(t,s), u(t,s)) \end{cases}$$

These are again called characteristic eq. and its solutions called characteristic curves. Sometimes we also call characteristic curves their projection onto the xy-plane, i.e. $(x(t,s), y(t,s))$:

with

$$x(0,s) = x_0(s),\quad y(0,s) = y_0(s),\quad u(0,s) = u_0(s)$$



The main difference between the linear and quasi-linear cases is that in the linear case the first two eq. are independent of the third.

Ex: Solve

$$u u_x - u u_y = u^2 + (x+y)^2 \quad \text{with} \quad u(x,0) \equiv 1$$

The characteristic eq. are $\begin{cases} \dot{x} = u \\ \dot{y} = -u \\ \dot{u} = u^2 + (x+y)^2 \end{cases}$ with $x(0,s) = s$, $y(0,s) = 0$, $u(0,s) = 1$

Adding the first two eq. $(x+y)^\circ = 0$, $x+y = f(s)$, but $x(0,s) + y(0,s) = s + 0 = s = f(s)$.

Plugging into the u-eq: $\dot{u} = u^2 + s^2$ or $\dfrac{du}{u^2+s^2} = dt \Rightarrow \frac{1}{s}\tan^{-1}\left(\frac{u}{s}\right) = t + g(s)$. Use

$u(0,s) = 1$: $\frac{1}{s}\tan^{-1}\left(\frac{1}{s}\right) = g(s)$. Thus $u = s\tan\left(st + \tan^{-1}\left(\frac{1}{s}\right)\right)$.

Next, we want $u = u(x,y)$. We know $s = x+y$. Need: $t = t(x,y)$. Plug u into $\dot{x}$-eq:

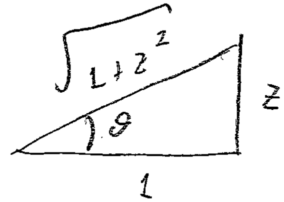
$\dot{x} = s\tan\left(st + \tan^{-1}\left(\frac{1}{s}\right)\right) \Rightarrow x = s\int \tan\big(\underbrace{st + \tan^{-1}\left(\tfrac{1}{s}\right)}_{v,\ dv = s\,dt}\big)\,dt = \int \tan v = -\log|\cos v| + h(s)$,

$x = -\log\left|\cos\left(st + \tan^{-1}\left(\frac{1}{s}\right)\right)\right| + h(s)$. Plug in $t=0$: $s = -\log\left|\cos\left(\tan^{-1}\left(\frac{1}{s}\right)\right)\right| + h(s)$, so

$x = -\log\left|\cos\left(st + \tan^{-1}\left(\frac{1}{s}\right)\right)\right| + \log\left|\cos\left(\tan^{-1}\left(\frac{1}{s}\right)\right)\right| + \underbrace{s}_{=x+y}$.

$y = \log\left|\dfrac{\cos\left(st + \tan^{-1}\left(\frac{1}{s}\right)\right)}{\cos\tan^{-1}\left(\frac{1}{s}\right)}\right|$, thus $e^y = \dfrac{\cos\left(st + \tan^{-1}\left(\frac{1}{s}\right)\right)}{\cos\tan^{-1}\left(\frac{1}{s}\right)}$

Recall that $\cos \tan^{-1} z = \dfrac{1}{\sqrt{1+z^2}}$



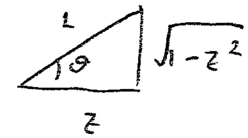
$\tan\theta = z$, $\tan^{-1} z = \theta$

$\cos\tan^{-1}\theta = \cos\theta = \dfrac{1}{\sqrt{1+z^2}}$

thus $\cos\tan^{-1}\left(\frac{1}{s}\right) = \dfrac{1}{\sqrt{1+\left(\frac{1}{s}\right)^2}} = \dfrac{s}{\sqrt{1+s^2}}$, $\dfrac{s e^y}{\sqrt{1+s^2}} = \cos\left(st + \tan^{-1}\left(\frac{1}{s}\right)\right)$. Take $\cos^{-1}$

$st = \cos^{-1}\left(\dfrac{s e^y}{\sqrt{1+s^2}}\right) - \tan^{-1}\left(\dfrac{1}{s}\right)$. Plugging into the formula for u (notice that the $\tan^{-1}\left(\frac{1}{s}\right)$ terms cancel out)

$u = s \tan \cos^{-1}\left(\dfrac{s e^y}{\sqrt{1+s^2}}\right)$. But $\tan\cos^{-1} z = \dfrac{\sqrt{1-z^2}}{z}$



$$u = s\,\frac{\sqrt{1-\left(\frac{s e^y}{\sqrt{1+s^2}}\right)^2}}{\frac{s e^y}{\sqrt{1+s^2}}} = \frac{\sqrt{1-\frac{s^2 e^{2y}}{1+s^2}}}{e^y/\sqrt{1+s^2}} = \frac{\sqrt{1+s^2-s^2e^{2y}}}{e^y} = \frac{\sqrt{1+(x+y)^2-(x+y)^2e^{2y}}}{e^y}$$

Notice that $u(x,0) = 1$.

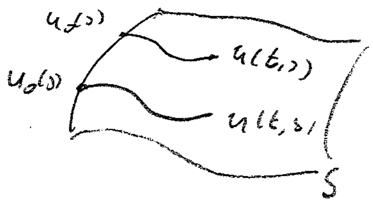
We can verify by direct computation that $u(x,y)$ satisfies the equation (verified with Mathematica).

Idea of the method of characteristics

in some cases u does not change at all $\Rightarrow$ u cte along the characteristic

1) Find curves (characteristic curves) along which we can $\overbrace{\text{compute how } u \text{ changes}}$

2) Since we know u at the initial point of those curves (initial condition), we can thus compute u at "any" point on the characteristic curve, thus forming the surface $S = \text{graph}(u)$



Notice how it was important to interpret the equation geometrically to derive the method.

Terminology The problem $\begin{cases} a(x,y,u)\,u_x + b(x,y,u)\,u_y = c(x,y,u) & (*) \\ x(0,s) = x_0(s), \quad y(0,s) = y_0(s), \quad u(0,s) = u_0(s) \end{cases}$ is called the Cauchy problem for the quasilinear eq $(*)$

The surface $r(s,t) \equiv (x(t,s), y(t,s), u(t,s))$ is called an integral surface (since the characteristic curves are integral curves).

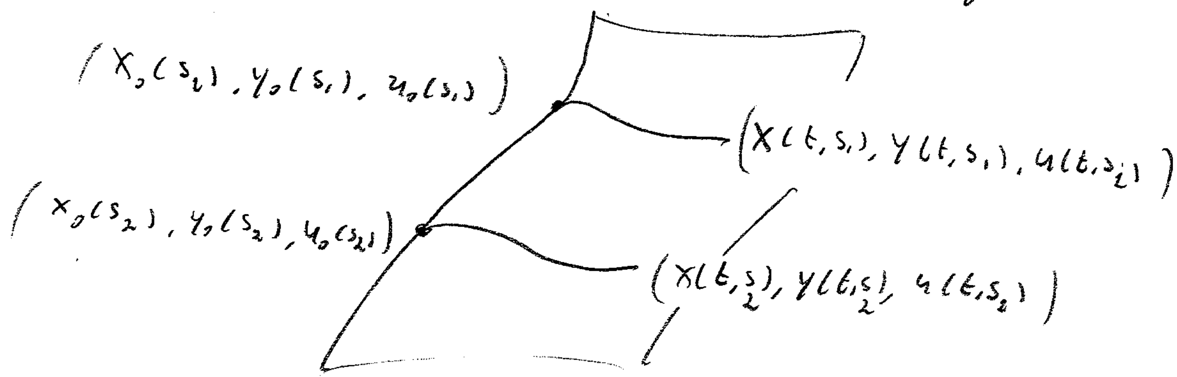
Remark. Notice that the characteristic equations can be non-linear even if the PDE is linear (recall ex $3u_x + 2xy^2 u_y = 0$)

Terminology: characteristics = characteristic curves.

$\Gamma(s)$ sometimes used for its projection as well. $\Leftarrow$ characteristics (x,y,u) vs projected characteristics (x,y) (= characteristics)

What could go wrong?

Intersecting characteristic: In the method of characteristics, solutions are constructed by propagating the initial conditions along the characteristic curves:



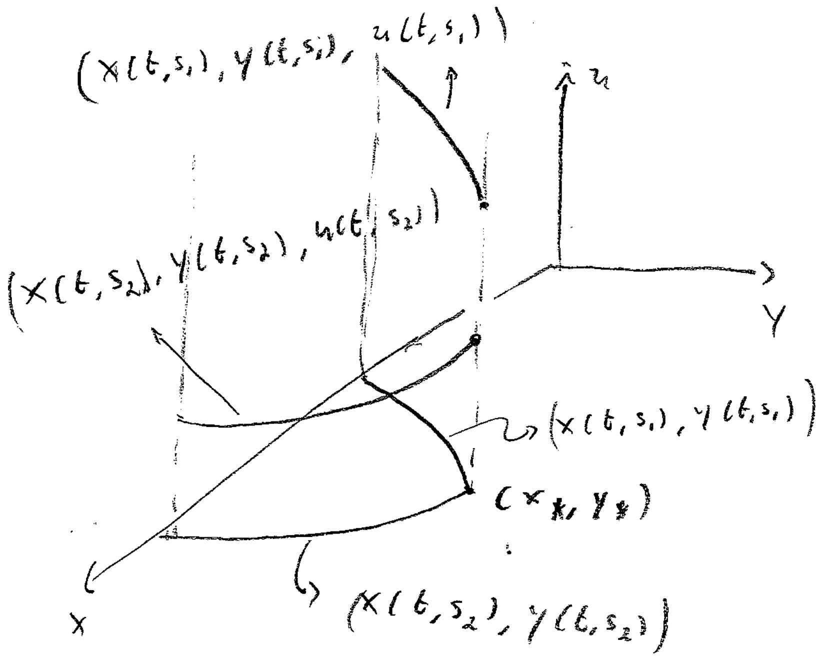
But what if for some t_*, projected characteristics starting at s_1 and s_2 ($s_1 \neq s_2$) intersect?

Then $X(t_*,s_1) = X(t_*,s_2) = x_*$

$Y(t_*,s_1) = Y(t_*,s_2) = y_*$

but which u do we take for the solution at (x_*,y_*)? $u(t_*,s_1)$ or $u(t_*,s_2)$?

Notice that there is no reason to expect $u(t_*,s_1) = u(t_*,s_2)$



(very important ex)

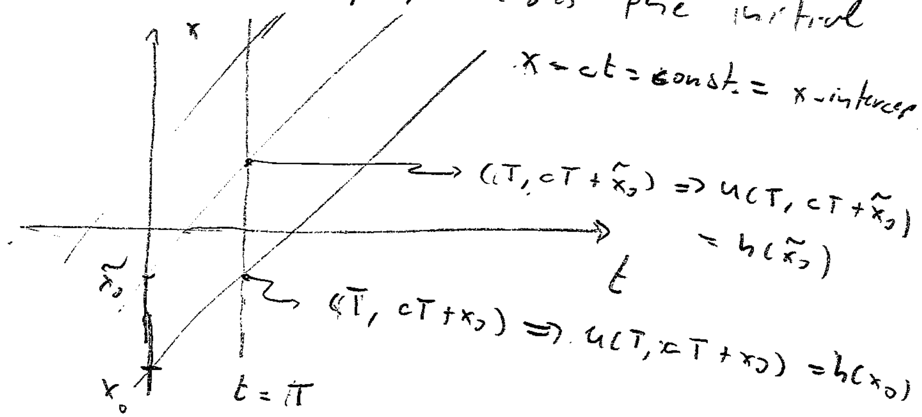
Ex: Consider $u_t + u u_x = 0$ with $u(x,0) = h(x)$ (so h is a known function)

Before analyzing this equation, consider the simpler version $u_t + c u_x = 0$, $c = cte.$

The characteristic eq. are
$$\begin{cases} \dot{x} = c & x(0,s) = s \\ \dot{t} = 1 & t(0,s) = 0 \\ \dot{u} = 0 & u(0,s) = h(s) \end{cases}$$
where $\dot{} = \frac{\partial}{\partial \tau}$ (since we have label t instead of y)

Then, $(x, t, u) = (s + c\tau, \tau, h(s))$. Eliminating $s = x - c\tau = x - ct$, we find

$u(t,x) = h(x - ct)$ This tells us that the solution simply moves the initial value of u with speed c along the x-axis:



Notice that in this case the (projected) characteristics are all parallel, thus they never intersect.

Back to $u_t + u u_x = 0$, we have

$$\begin{cases} \dot{x} = u \longrightarrow \dot{x} = h(s) \Rightarrow x = \tau h(s) + s \\ \dot{t} = 1 \Rightarrow t = \tau \\ \dot{u} = 0 \Rightarrow u = h(s) \end{cases}$$

Since $h(s) = u(\tau, s)$,

$x = \tau u(\tau, s) + s$

$x(0,s)$, $t(0,s) = 0$

$u(0,s) = h(s)$

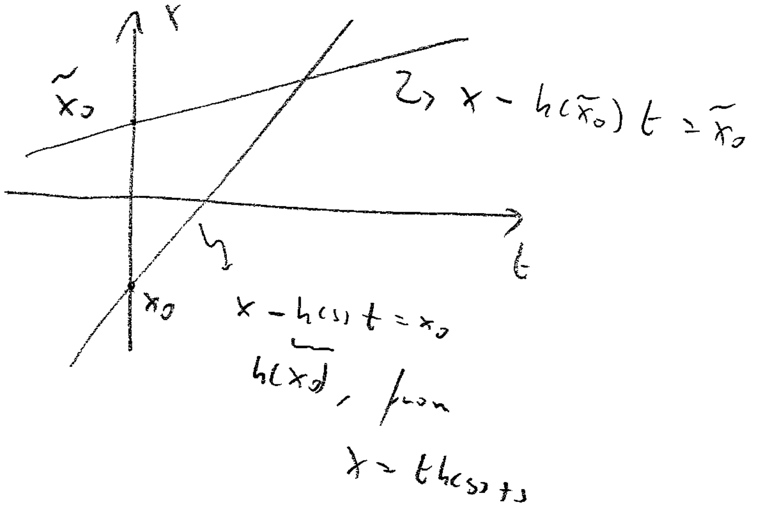
Solving for s as before, $s = x - \tau u = x - tu$, so

$u(t,x) = h(s) = h(x - tu)$

The formula $u(t,x) = h(x - tu(t,x))$ gives u implicitly, but we can double check that it is a solution. Indeed, with $t = \tau$ and writing x in terms of $(\tau, s) = (t,s)$ we have $u(t, x(t,s)) = h(s)$. (where we used $x = \tau h(s) + s = tu + s$)

Taking ∂_t: $u_t(t, x(t,s)) + u_x(t, x(t,s)) \underbrace{\dot{x}(t,s)}_{= u(t, x(t,s))} = 0$

In order to analyze the solutions, let's go back to the characteristic curves: $x(\tau, s) = \tau h(s) + s$; $t(\tau, s) = \tau$, so the (projected) characteristics are $x = t h(s) + s$. For each s, the characteristic is a straight line with slope (speed) $h(s)$. Hence, in general, different characteristics have different slopes, so they will intersect.



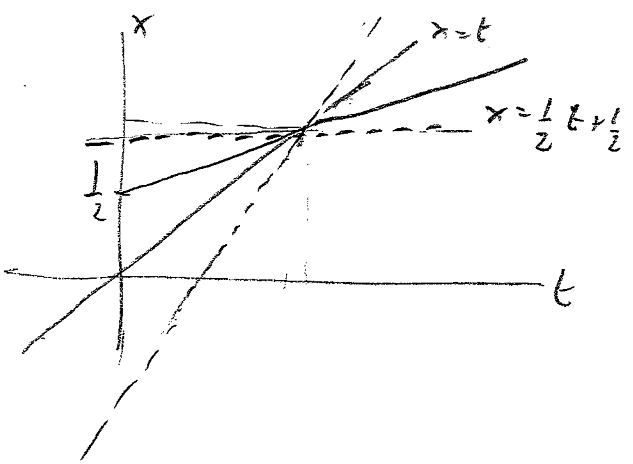
We thus expect something wrong to happen. Indeed, differentiating $u(t,x) = h(x - tu(t,x))$ w.r.t. x:

$$u_x(t,x) = h'(x - tu(t,x))(1 - t u_x(t,x)). \quad \text{Or}$$

$$u_x(t,x) = \frac{h'(x - tu(t,x))}{t h'(x - tu(t,x)) + 1}$$

- Or, writing $x - tu = s$ we have $u_x = \dfrac{h'(s)}{1 + t h'(s)}$. Thus, for each s (corresponding to a point on the x-axis, parametrization of the initial condition) there exists a $t_* = t_*(s)$, namely, $t_* = -\dfrac{1}{h'(s)}$, for which $u_x = \infty$. At this moment u is no longer a solution to the PDE. When $u_x = \infty$, we say that the solution blows-up.

To see that the blow-up is indeed due to the intersection of characteristics, let us consider the example $h(s) = 1 - s$. Take the characteristics corresponding to $s = 0$ and $s = \frac{1}{2}$: $x = h(s)t + s \Rightarrow x = t$, $x = \frac{1}{2}t + \frac{1}{2}$.



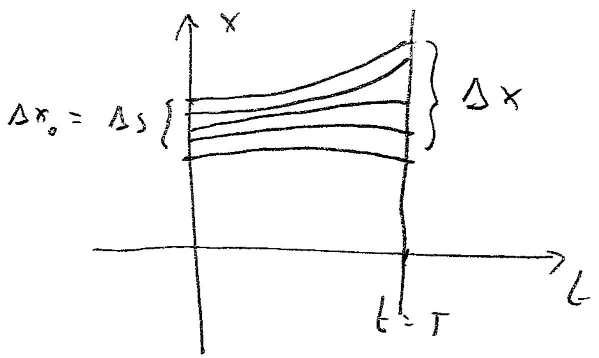
They intersect at $t = 1$. But $h'(s) = -1$, so $t_* = -\dfrac{1}{h'(s)} = 1$, and for $t = 1$ we get $u_x = \infty$.

In fact, all characteristics intersect at $x = 1$:

$$x = x_0 + t\, h(x_0) = x_0 + t(1 - x_0) \Rightarrow x(1) = 1 \text{ for any } x_0$$

How can we see that the blow-up is indeed due to intersecting characteristics (as opposed to, say, a coincidence at $t=1$ in the previous example)?

Consider the characteristics $x = t\,h(s) + s$ emanating from an interval of length Δs at $t=0$, and reaching a fixed time $t=T$.



The characteristics leaving $t=0$ reach $t=T$, forming an interval of length Δx at $t=T$. Δx need not to be equal to Δs, but $\Delta x > 0$ unless the interval at $t=T$ is collapsed to a point.

We have $\Delta x \simeq \frac{dx}{ds}\Delta s$ or, infinitesimally, $dx = \frac{dx}{ds}ds$.

When characteristics intersect, $\Delta x \to 0$, ie, $dx = 0$, or $\frac{dx}{ds} = 0$. But computing:

$$\frac{dx}{ds} = t\,h'(s) + 1 = 0 \Rightarrow t = -\frac{1}{h'(s)}, \text{ as before.}$$

The regions where $u_x = \infty$ in Burger's equations are called shock waves, and they correspond to several phenomena of interest

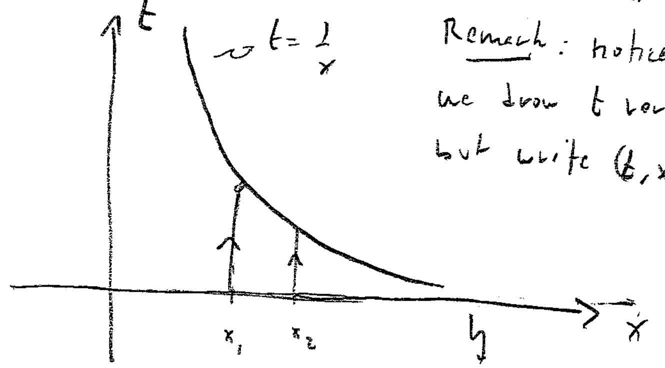
traffic flow, glacier waves, airplanes breaking the sound barrier.

The eq. $u_t + u u_x = 0$ is called Burger's eq (compare to Euler's)

__Solutions may exist only locally__ When a solution $u=u(x,y)$ exists for all (x,y) (or for all (x,y) in a pre-determined range; say, if y represents time we may be interested only on $y \geq 0$), we say that the solution exists __globally__. When a solution exists only on a neighborhood of $\Gamma(s)$ we say that it exists __locally__.

__EX__: Consider
$$\begin{cases} \frac{2}{\pi}\frac{1}{x}\frac{1}{1+u^2}u_t = 1 \\ u(0,x) = 0, \quad x>0 \end{cases}$$
. The characteristic eq. are
$$\begin{cases} \dot{x} = 0 \\ \dot{t} = \frac{2}{\pi}\frac{1}{x}\frac{1}{1+u^2} \\ \dot{u} = 1 \end{cases} \quad \text{with} \quad \begin{array}{l} x(0,s) = s \\ t(0,s) = 0 \\ u(0,s) = 0 \end{array}$$
$$\dot{} = \frac{\partial}{\partial\tau}$$

we find $x = s$, $u = \tau$. Then, plugging into the eq. for t: $\dot{t} = \frac{2}{\pi}\frac{1}{s}\frac{1}{1+\tau^2}$, which has solution $t = \frac{2}{\pi}\frac{1}{s}\tan^{-1}(\tau)$, after using $t(0,s)=0$. Thus, for each $s=x$, t exists only up to $\frac{1}{x}$, since $\tan^{-1}(\tau) \to \frac{2}{\pi}$ as $\tau \to \infty$. To see what's happening, use $\tau = u$ and solve for u to get $u(t,x) = \tan\left(\frac{\pi x t}{2}\right)$, so u blows-up along $xt=1$.



$\Gamma(s) = (t(s), x(s))$

Remark: notice that we draw t vertically but write (t,x)

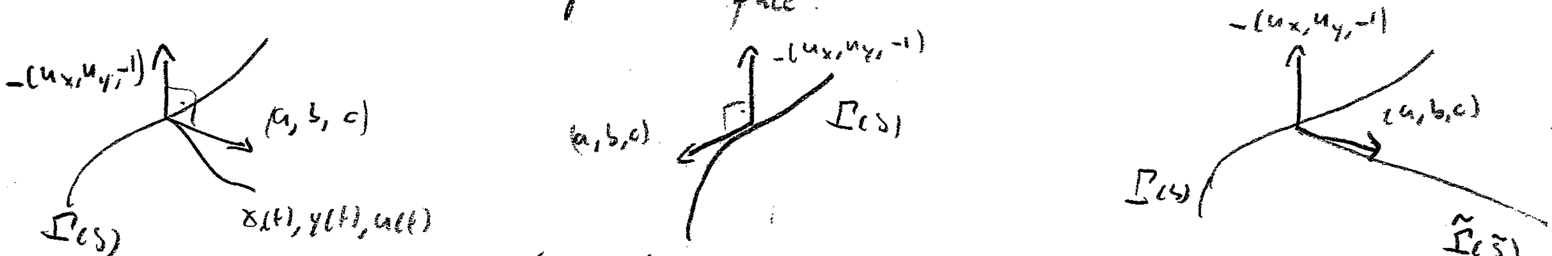
Notice that here the characteristics don't intersect, but they "end" on $t = 1/x$
(with t = time, we would like to have global existence)

<u>Inversion $(t,s) \mapsto (x,y)$</u> In the examples we worked out we could always go back to the original variables (x,y), i.e., we could write $t=t(x,y)$, $s=s(x,y)$. But this may not always be the case. Recall that the map $(x,y) = (x(t,s), y(t,s))$ is invertible (so $t=t(x,y)$ and $s=s(x,y)$ are well defined) if the Jacobian $J = \dfrac{\partial(x,y)}{\partial(t,s)}$

$$= \det\begin{bmatrix} \frac{\partial x}{\partial t} & \frac{\partial x}{\partial s} \\ \frac{\partial y}{\partial t} & \frac{\partial y}{\partial s} \end{bmatrix}$$

is not zero, $J \neq 0$.

One thing to notice is that while the dependence of the characteristics on t comes from the PDE, the dependence on s comes from the initial condition $\Gamma(s)$. Since the PDE and the initial condition are independent of each other, for a given PDE there always exist initial conditions for which the Jacobian vanishes. These are cases where the characteristic curves don't go "inside" the integral surface:



(a,b,c) doesn't point to the "inside"

(a,b,c) points inside for $\Gamma(s)$, but not for $\tilde{\Gamma}(\tilde{s})$

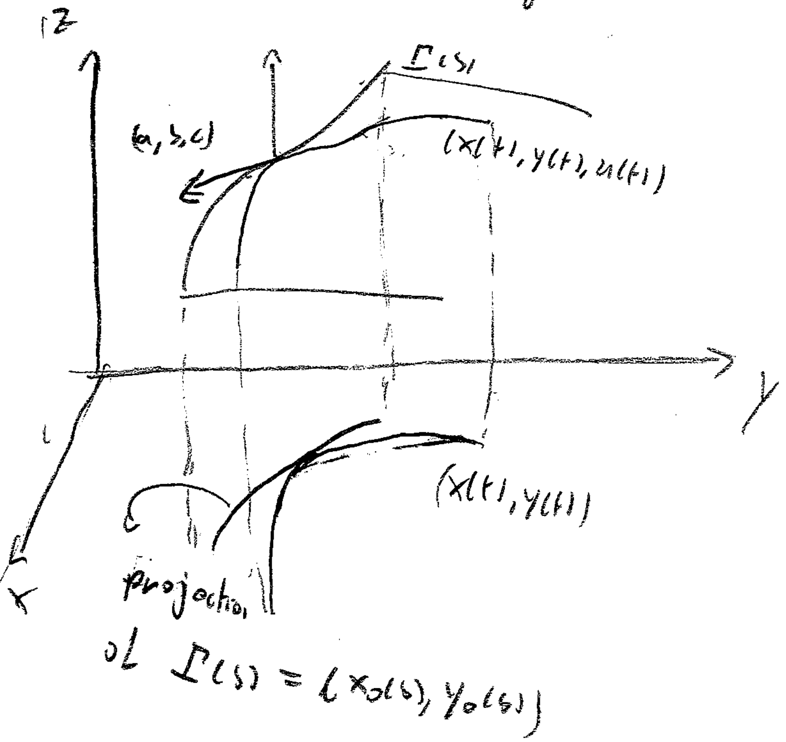
Computing J and evaluating at $t=0$

$J(t,s) = x_t(t,s)\,y_s(t,s) - x_s(t,s)\,y_t(t,s)$ so with $x(0,s) = x_0(s)$, $y(0,s) = y_0(s)$ and

$J(0,s) = a(x_0,y_0)\dfrac{\partial y_0}{\partial s} - b(x_0,y_0)\dfrac{\partial x_0}{\partial s}$ $x_t = a \quad y_t = b$

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

Thus $J(0,s)=0$ if and only if the vectors (a,b) and $\left(\dfrac{\partial x_0}{\partial s}, \dfrac{\partial y_0}{\partial s}\right)$ are linearly dependent. $J(0,s)=0$ if the (projected) characteristic curve $(x(t,s), y(t,s))$ at $t=0$ and $\Gamma(s)$ are tangent



In general, in order to guarantee a unique solution to a first order PDE with given I.C. we need $J(0,s) \neq 0$ (for all s)

$J(0,s)\neq 0$ is called the <u>transversality condition</u>, when it is satisfied $\Gamma(s)$ is said to be <u>non-characteristic</u> and characteristic otherwise.

(I'll give exercise)

Def. Consider

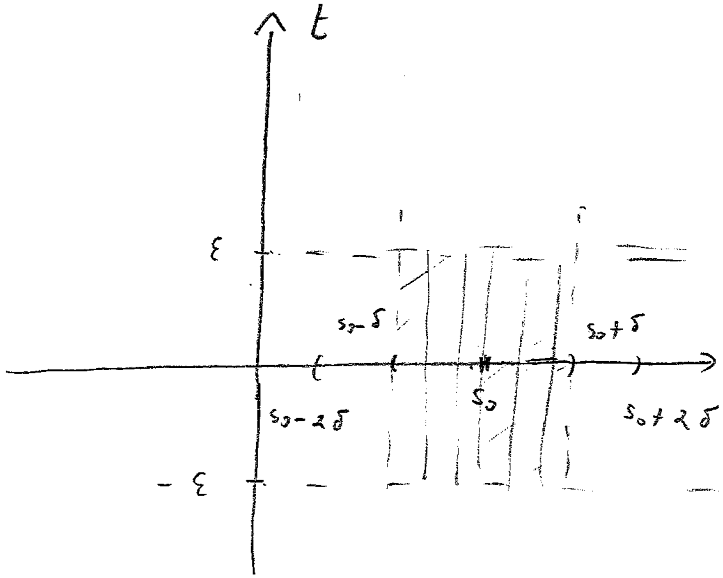
$$(*)\begin{cases} a(x,y,u)\,u_x + b(x,y,u)\,u_y = c(x,y,u) & \text{with initial condition} \\ \Gamma(s) = (x_0(s), y_0(s), u_0(s)) \end{cases}$$

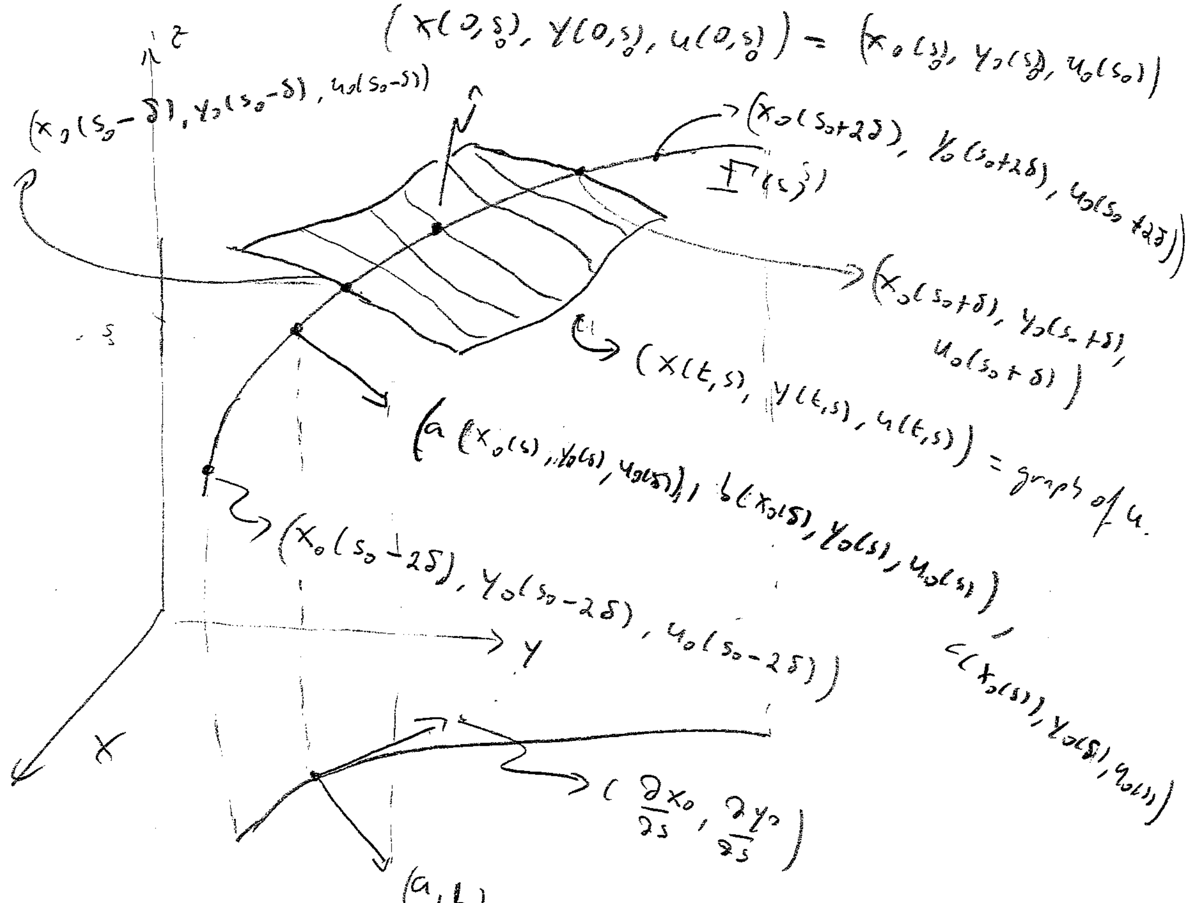
We say that the equation and the initial condition satisfy the <u>transversality condition</u> (or that $\Gamma(s)$ is non-characteristic for the PDE) at a point s on Γ if

$$J(0,s) = \underbrace{x_t(0,s)\,y_s(0,s) - y_t(0,s)\,x_s(0,s)}_{} \neq 0$$

$$= \det\begin{bmatrix} \frac{\partial x}{\partial t} & \frac{\partial x}{\partial s} \\ \frac{\partial y}{\partial t} & \frac{\partial y}{\partial s} \end{bmatrix}\Bigg|_{t=0} = \det\begin{bmatrix} a(x_0,y_0) & \frac{\partial x_0}{\partial s} \\ b(x_0,y_0) & \frac{\partial y_0}{\partial s} \end{bmatrix}, \quad x_0 = x_0(s),\ y_0 = y_0(s)$$

<u>Theorem</u> Consider the Cauchy problem $(*)$, and assume that a, b, and c are <u>smooth functions</u> (i.e., all their partial derivatives of any order exist). Assume that the transversality condition holds for each s in an interval $(s_0 - 2\delta, s_0 + 2\delta)$ of the initial curve $\Gamma(s)$. Then, there exists an $\varepsilon > 0$ such that the Cauchy problem $(*)$ has a unique solution in the neighborhood $(t,s) \in (-\varepsilon, \varepsilon) \times (s_0 - \delta, s_0 + \delta)$ of the initial curve. If the transversality condition fails for an interval of s values, then $(*)$ has no solution or infinitely many solutions.





Remarks : • Notice that we are given an initial condition, so the statement about infinitely many solutions is not because of "undetermined functions".

• Notice that we can only guarantee a solution on the smaller interval $(s_0-\delta, s_0+\delta)$.

• Existence on $(-\varepsilon, \varepsilon)$ means that the integral surface exists on "both sides" of $\Gamma(s)$ (positive and negative "time") proof: By the existence and uniqueness theorem for ODEs, for each $s \in (s_0-2\delta, s_0+2\delta)$, there exists a characteristic curve $(x(t,s), y(t,s), u(t,s))$ starting at $(x_0(s), y_0(s), u_0(s))$ and existing on tan interva $(-\varepsilon_*(s), \varepsilon_*(s))$. Considering s on the small interval $(s_0-\delta, s_0+\delta)$ we can guarantee that there exists a $\varepsilon > 0$ such that $(-\varepsilon, \varepsilon) \subseteq (-\varepsilon_*(s), \varepsilon_*(s))$ for all $s \in (s_0-\delta, s_0+\delta)$. The transversality condition implies that the family of characteristics $(x(t,s), y(t,s), u(t,s))$, $(t,s) \in (-\varepsilon, \varepsilon) \times (s_0-\delta, s_0+\delta)$ gives a parametrization of a smooth surface.

Next, we show that the constructed surface gives a solution to the PDE. Since $J(0,s) \neq 0$, by continuity and making ε smaller if necessary, we have that $J(t,s) \neq 0$ for $(t,s) \in (-\varepsilon, \varepsilon) \times (s_0-\delta, s_0+\delta)$. Thus, we can invert $x(t,s)$, $y(t,s)$ and set

$$u(x,y) = u(t(x,y), s(x,y))$$

Compute: $\frac{\partial u}{\partial x} = \frac{\partial u}{\partial t}\frac{\partial t}{\partial x} + \frac{\partial u}{\partial s}\frac{\partial s}{\partial x}$; $\frac{\partial u}{\partial y} = \frac{\partial u}{\partial t}\frac{\partial t}{\partial y} + \frac{\partial u}{\partial s}\frac{\partial s}{\partial y}$

But $$1 = \frac{\partial t}{\partial t} = \frac{\partial}{\partial t}(t(x,y)) = \frac{\partial t}{\partial x}\underbrace{\frac{\partial x}{\partial t}}_{=a} + \frac{\partial t}{\partial y}\underbrace{\frac{\partial y}{\partial t}}_{=b} = a\frac{\partial t}{\partial x} + b\frac{\partial t}{\partial y}$$

$$0 = \frac{\partial s}{\partial t} = \frac{\partial}{\partial t}(s(x,y)) = \frac{\partial s}{\partial x}\overbrace{\frac{\partial x}{\partial t}} + \frac{\partial s}{\partial y}\overbrace{\frac{\partial y}{\partial t}} = a\frac{\partial s}{\partial t} + b\frac{\partial s}{\partial y}$$

Thus, $a u_x + b u_y = a(u_t t_x + u_s s_x) + b(u_t t_y + u_s s_y) = u_t\underbrace{(a t_x + b t_y)}_{=1} + u_s\underbrace{(a s_x + b s_y)}_{=0}$

$a u_x + b u_y = u_t$: But $u_t = c$ by the characteristic equations, thus

$a u_x + b u_y = c$, showing existence.

We now move to prove uniqueness. Suppose we are given a solution $f = f(x,y)$, and let $(x(t,s), y(t,s), u(t,s))$ be a characteristic curve. We must have

$f(x(0,s), y(0,s), u(0,s)) = u(0,s)$ since by assumption f solves (*).

Set

$$F(t,s) := u(t,s) - f(x(t,s), y(t,s)).$$

Compute

$$F_t(t,s) = u_t(t,s) - f_x(x,y)\,\dot{x}(t,s) - f_y(x,y)\,\dot{y}(t,s)$$

$$= c(x,y,u) - f_x(x,y)\,a(x,y,u) - f_y(x,y)\,b(x,y,u) \quad (\text{use } u = F+f)$$

$$F_t(t,s) = c(x,y,F+f) - f_x(x,y)\,a(x,y,F+f) - f_y(x,y)\,b(x,y,F+f) \quad (**)$$

By the initial condition, $F(0,s) = 0$. But $F(t,s)=0$ is a solution to the ODE (**). [plug $F=0$: $af_x+bf_y=c$ which is satisfied since f is a solution]

By the uniqueness result for ODEs, we conclude that

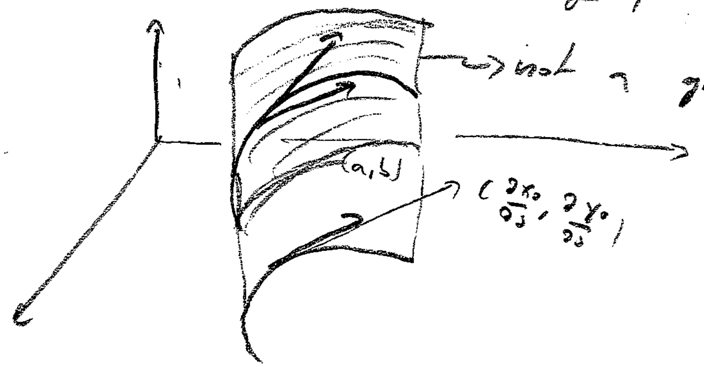
$$F(t,s) = u(t,s) - f(x(t,s), y(t,s)) = 0$$

and since this holds for any $s \in (s_0-\delta, s_0+\delta)$, we conclude that $f=u$, showing uniqueness.

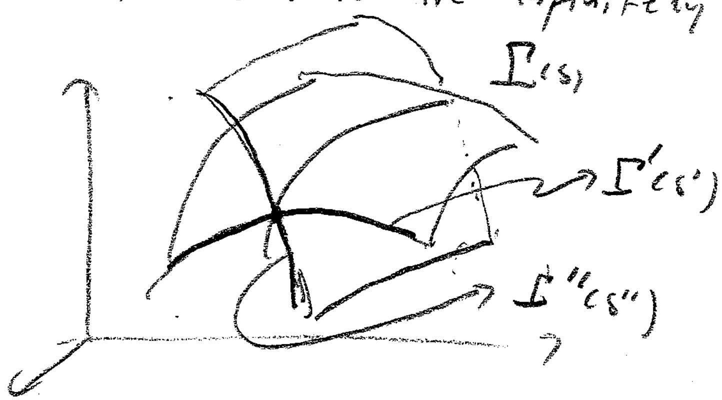
Suppose now that the transversality condition does not hold on an interval of s values, so $\left(\frac{\partial x_0}{\partial s}, \frac{\partial y_0}{\partial s}\right)$ and $\left(a(x_0(s), y_0(s)),\ b(x_0(s), y_0(s))\right)$ are linearly dependent.

Then there are two possibilities: (i) a characteristic in the direction (a,b) starting at (x_0,y_0) agrees with $\Gamma(s)$, or (ii) it doesn't.

In case (ii), the characteristic and $\Gamma(s)$ have the same projection but don't agree, so the characteristic cannot belong to an integral surface, hence there is no solution.:



In case (i), we can construct a solution as follows. Pick $(x_0, y_0, u_0) \in \Gamma(s)$, let $\Gamma'(s')$ be a curve through (x_0, y_0, u_0) transversal to $\Gamma(s)$. By the existence part of the theorem, we can construct a parametrized surface S', which gives a solution to (*) (S' will contain $\Gamma(s)$ since now $\Gamma(s)$ is a characteristic). This works for any choice of transversal $\Gamma'(s')$, thus there are infinitely many solutions.



$\boxtimes$

Remarks

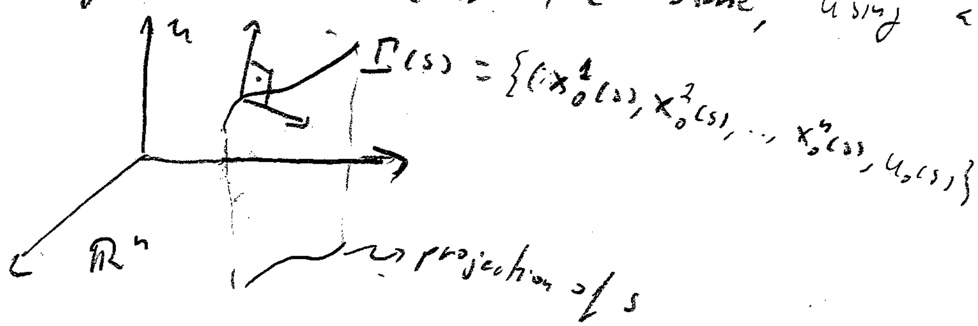
- The method of characteristics can be adapted to deal with "fully non-linear equations", i.e., non-linear equations that are not quasi-linear.

Ex: $u_x^2 + u_y^2 = 1$

$u_x u_x = a(x,y,u,u_x)\, u_x$

- The method can also be applied for functions of n variables $u = u(x^1, \dots, x^n)$.

The basic geometric idea is the same, using a non-transversality condition



Our previous experience with first order PDEs shows that we need to be more precise about the domain of definition of a PDE and its solution.

<u>Def.</u> Let $\mathcal{U} \subseteq \mathbb{R}^n$ be a <u>domain</u> in $\mathbb{R}^n$ (i.e., an open and connected set). A PDE of order m for a function $u = u(x^1, \dots, x^n)$ in $\mathcal{U}$ is an equation

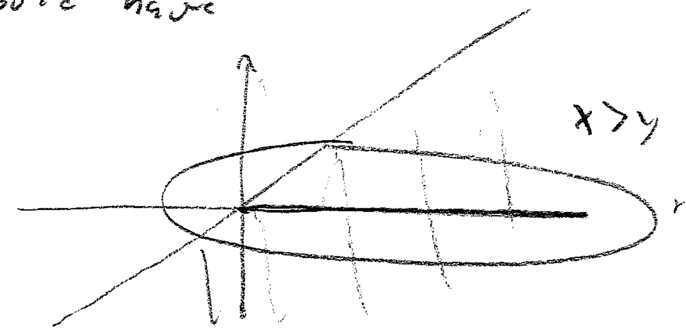
$$F(x, u, Du, \dots, D^m u) = 0. \quad (*)$$

A <u>solution</u> to the PDE is a function defined on $\mathcal{U}$ which satisfies $(*)$. We sometimes use the terminology <u>global solution</u> for a function that satisfies $(*)$ on all of $\mathcal{U}$, and <u>local solution</u> when a function satisfies $(*)$ only in part of $\mathcal{U}$.

<u>Ex:</u> Solve $u_x + u_y = 0$ in $\mathbb{R}^2$. We verify that $u(x,y) = \cos\sqrt{x-y}$ solves the equation. This solution is local because it is defined only on $\{(x,y) \mid x > y\}$ (notice that we cannot include $x = y$ because u_x, u_y blow-up there). On the other hand, if we take $\mathcal{U} = \{(x,y) \mid x > y\}$ then this solution is global in $\mathcal{U}$.

Remark The distinction global vs local solutions becomes important when we have an I.C. For example, we could have

$$\begin{cases} u_x + u_y = 0 & \text{in } \mathbb{R}^2 \\ u(x,0) = \cos\sqrt{x}, & x > 0 \end{cases}$$



The function $u(x,y) = \cos\sqrt{x-y}$ is thus a local solution as it exists only in $\{(x,y) \mid x > y\}$, i.e., in a neighborhood of the initial region $\{y=0\}$.

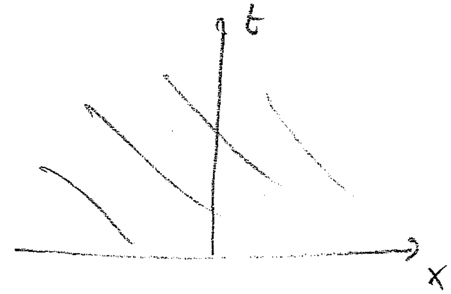
Notation Let u be a function defined in a domain $\mathcal{U} \subseteq \mathbb{R}^n$. We write

$$u \in C^k(\mathcal{U}) \quad \text{or simply} \quad u \in C^k \text{ when } \mathcal{U} \text{ is understood}$$

to mean that all derivatives of order k of u exist and are continuous in $\mathcal{U}$. If all derivatives of u exist, we write $u \in C^\infty$ and say that u is smooth. (so when we write $u \in C^k$, we can possibly have $k = \infty$).

The wave equation

We will study the one-dimensional wave equation:



$$u_{tt} - c^2 u_{xx} = 0 \quad \text{in} \quad (x,t) \in (-\infty,\infty) \times (0,\infty)$$

Consider the change of variables $\alpha = x + ct$, $\beta = x - ct$ and set $v(\alpha,\beta) = u(x,t)$, i.e., $u(x,t) = v(\alpha(x,t), \beta(x,t))$. Then

$$u_t = v_\alpha \frac{\partial \alpha}{\partial t} + v_\beta \frac{\partial \beta}{\partial t} = c\, v_\alpha - c\, v_\beta ,$$

$$u_{tt} = c\, \partial_t (v_\alpha - v_\beta) = c\left(v_{\alpha\alpha} \frac{\partial \alpha}{\partial t} + v_{\alpha\beta} \frac{\partial \beta}{\partial t} - v_{\beta\alpha} \frac{\partial \alpha}{\partial t} - v_{\beta\beta} \frac{\partial \beta}{\partial t} \right)$$

$$= c\left(v_{\alpha\alpha} c + v_{\alpha\beta}(-c) - v_{\beta\alpha} c - v_{\beta\beta}(-c) \right) = c^2 (v_{\alpha\alpha} - 2 v_{\alpha\beta} + v_{\beta\beta})$$

And

$$u_x = v_\alpha \frac{\partial \alpha}{\partial x} + v_\beta \frac{\partial \beta}{\partial x} = v_\alpha + v_\beta$$

$$u_{xx} = \partial_x (v_\alpha + v_\beta) = v_{\alpha\alpha} \frac{\partial \alpha}{\partial x} + v_{\alpha\beta} \frac{\partial \beta}{\partial x} + v_{\beta\alpha} \frac{\partial \alpha}{\partial x} + v_{\beta\beta} \frac{\partial \beta}{\partial x} = v_{\alpha\alpha} + 2 v_{\alpha\beta} + v_{\beta\beta}$$

Hence

$$0 = u_{tt} - c^2 u_{xx} = c^2 (v_{\alpha\alpha} - 2 v_{\alpha\beta} + v_{\beta\beta}) - c^2 (v_{\alpha\alpha} + 2 v_{\alpha\beta} + v_{\beta\beta})$$

$$= -c^2 4 v_{\alpha\beta} \implies v_{\alpha\beta} = 0 \quad \text{since } c \neq 0$$

Write this as $\partial_\beta(\partial_\alpha v) = 0$. Thus, $\partial_\alpha v$ is independent of β, hence it is a function only of α: $\partial_\alpha v(\alpha,\beta) = f(\alpha)$, for some function f.

Integrating in α: $v(\alpha,\beta) = \int f(\alpha)\,d\alpha + G(\beta)$ where G is a function only of β. But $\int f(\alpha)\,d\alpha = F(\alpha)$ = function only of α. Thus

$$v(\alpha,\beta) = F(\alpha) + G(\beta).$$

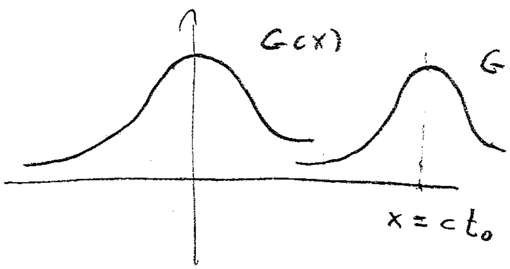
Now recall that $\alpha = x+ct$, $\beta = x-ct$, and $u(x,t) = v(\alpha,\beta)$, so

$$\boxed{u(x,t) = F(x+ct) + G(x-ct)} \qquad (*)$$

This shows the following:

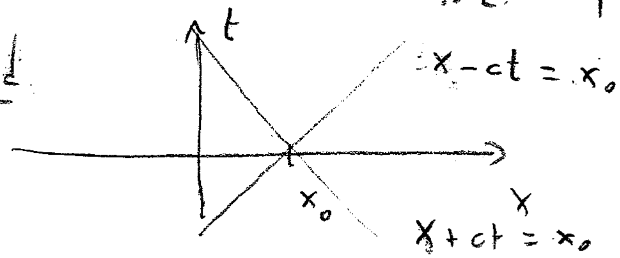
Prop If u is a solution of the one-dimensional wave equation, then there exist two functions $F, G \in C^2$ such that $(*)$ holds. Conversely, given any two functions $F, G \in C^2$, formula $(*)$ defines a solution to the one dimensional wave equation.

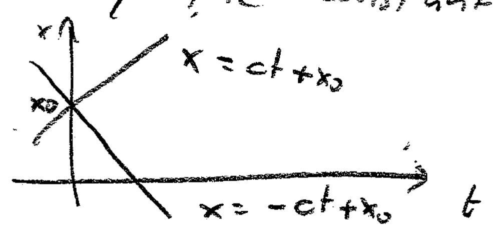
For any fixed $t_0 > 0$, the graph of $G(x - ct_0)$ is the same as that of $G(x)$, except that it is shifted to the right by the distance ct_0. Thus, as t moves, the graph of $G(x-ct)$ moves to the right. Since the distance travelled by the graph after time t is ct, the graph is moving at speed c.



Similarly, the graph of $F(x+ct)$ moves to the left at speed c.

Since $F(x+ct)$ and $G(x-ct)$ are solutions to the wave equation, we interpret them as waves moving (propagating) at speed c. G is called a <u>forward</u> wave and F a <u>backward</u> wave. Formula (*) above says that any solution to the wave equation in one dimension is a $\underbrace{\text{sum}}_{\text{"superposition"}}$ of a forward and a backward wave. That's why the constant c is the <u>wave speed</u>.





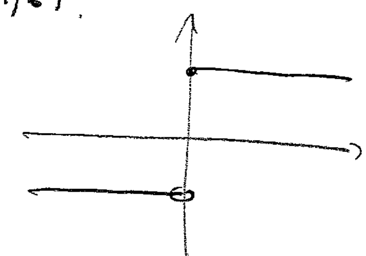
We derived (*) assuming that F and G are C^2. But given formula (*), we can ask if it is valid to consider functions that are not C^2. For instance, take $F(x) = G(x) = |x|$. Then (*) would give

$$u(x,t) = |x+ct| + |x-ct|$$

Since $|x|$ is C^2 except at $x=0$, we see that the above formula fails to give a solution to the wave equation only when $x = \pm ct$, so it seems that $|x+ct| + |x-ct|$ is "almost" a solution. This motivates the following:

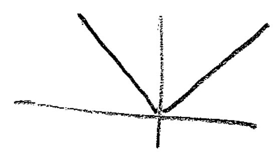
<u>Def</u>. A piecewise C^k function $f : (a,b) \to \mathbb{R}$ is a function that is C^k excep at a countable number of isolated points $\{x_i\}_{i=1}^{\infty} \subset (a,b)$.

<u>Ex</u>: $f(x) = \begin{cases} 1, & x \geq 0 \\ -1, & x < 0 \end{cases}$ is a piecewise smooth (C^∞) function

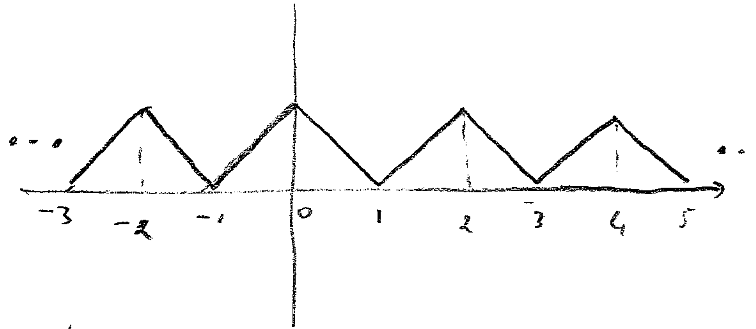


In this example $\{x_i\}_{i=1}^{\infty} = 1$ point $= 0$.

<u>Ex</u>: $f(x) = |x|$ is a piecewise smooth function

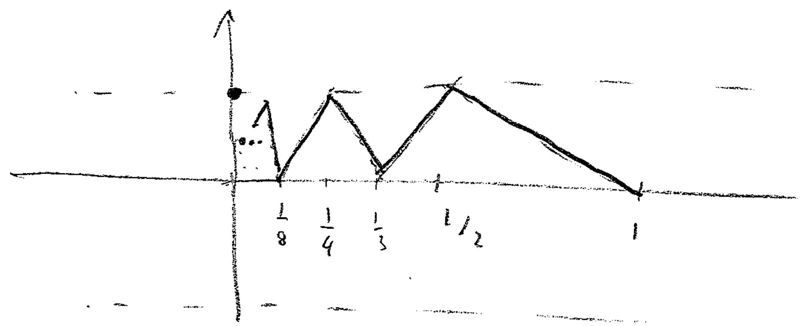


Ex: The function



is a piecewise smooth function. In this case $\{x_i\}_{i=1}^{\infty}$ = integers.

Ex: The function



is not piecewise smooth because in this case $\{x_i\}_{i=1}^{\infty} = \{\frac{1}{n}\}_{n=1}^{\infty} \cup \{0\}$, and thus $\{0\}$ is not isolated.

Remarks: • The points $\{x_i\}_{i=1}^{\infty}$ where a piecewise C^k function fails to be C^k are sometimes called the singularities of the function.

• If $f:(a,b) \to \mathbb{R}$ is piecewise C^k and (a,b) is finite, then the number of singularities must be finite. But there can be infinitely many singularities if (a,b) is infinite (ex, $(a,b) = (-\infty,\infty)$, see above example).

Def. Let F and G be picewise C^2 functions. We call the function

$$u(x,t) = F(x+ct) + G(x-ct)$$

a generalized solution of the wave equation (aka a weak solution of the wave equation).

A generalized solution satisfies the wave equation except at the singularities.

A solution that contains no singularities (i.e., a solution in the "old" sense) is called a classical solution

Generalized solutions are important, e.g. in the study of shocks. The terminology generalized or weak solution may carry different meanings depending on the context.