2-Motivations From the Theory of ODE's
MATH 22C

1. Ordinary Differential Equations (ODE) and the Fundamental Role of the Derivative in the Sciences

Recall that a real valued function of a real variable $y = f(t)$ gives outputs y in terms of inputs t.

t=input=independent variable$\in \mathcal{R}$ (known)
y=output=dependent variable$\in \mathcal{R}$ (unknown)
f=name of function...

Examples: $y = t^2$, $y = e^t$, $y = \sin t$, $y = \ln t$, etc.

•**Defn:** An ODE is an equation involving an unknown function and its derivative.

•Recall: the derivative of a function $y = f(t)$ is defined as the rate of change of y with respect to t at a given value of t:

$$y' = f'(t) = \lim_{\Delta t \to 0} \frac{f(t + \Delta t) - f(t)}{\Delta t} = \frac{dy}{dt},$$

where dy can be interpreted as the *rise* and dt the *run* along the line tangent to the graph of f at the point t...so the fraction $\frac{dy}{dx} = \frac{rise}{run}$ makes sense as the slope.

•So why is the derivative so important? This is a deep philosophical question, but in a nutshell the derivative is important because *The fundamental laws of science come to us stated in terms of rates of change.* At the start one might think they could come to us as graphs, or as algebraic relations, or in some other form like pictures, but the experience of some three hundred years since the time of Newton is that they don't. They typically come to us stated in terms of rates of change.

–This is confirmed by many centuries of experience. But the more you think about it, the more subtle this point appears to be! One explanation for why fundamental laws use rates of change is that it is possible to give the relationships between how different things are changing at a given instant, but global relations are too complicated to find at the beginning. Taking this a bit deeper one might say that the derivative provides the *linearization* of the function f at a given point, (it is the slope of a *line*), and linear means the principle of superposition holds, so the relations between different rates of change can be added together from simpler relations...they don't interfere with each other...so the fundamental laws can be discovered by dissecting them into pieces and then putting them back together again at the level of rates of change.

–The acceleration in Newton's force law $F = ma$, reaction rates in chemistry, rates of continental drift in geology, rates at which plumes of chemicals move through the ground water in hydrology, rates of energy consumption by cells in biology. The law that defines the rate is usually the starting point for a quantitative understanding of any of these processes. To make a prediction, you have to "integrate up the rate law" to get a graph to compare with experiment. That is where the *integral* of calculus comes in. The Fundamental Theorem of Calculus relates the two:

–**Theorem:** (FTC) The derivative and the integral are related by the following fundamental relation:

$$\int_a^b f(t)dt = F(b) - F(a),$$

where $F(t)$ is an *anti-derivative* of $f(t)$ satisfying

$$F'(t) = f(t).$$

–The FTC tells us that the definite integral, the *area under the graph of f above the interval $[a,b]$*, defined in terms of a Riemann sum of approximating rectangular areas,

$$\int_a^b f(t)dt = \lim_{\Delta t \to 0} \sum_{k=1}^{N} f(x_k)\Delta x, \tag{1}$$

can be evaluated by finding an antiderivative and evaluating between the endpoints. Here, (1) is based on dividing $[a,b]$ into N equal spaces $a = x_0 < x_1 <, ..., < x_N = b$ of length $\Delta x = x_k - x_{k-1} = \frac{b-a}{N}$.

•The FTC can be viewed as a method for solving the simplest ODE's. In general, ODE's are too complicated to solve directly by the FTC alone. For a fundamental example, consider Newton's Force Law:

•Consider a rocket ship rising from the ground. To determine its height as a function of time, the starting point is Newton's force law $F = ma$. To get the formula for height vs time, you determine the mass $m(t)$ (it might change with time as it uses fuel), and the force $F(t) = F_r(t) - F_g(t) - F_f(t)$ (F_r equals the force of the rocket engine, F_g of gravity and F_f of friction—all can be determined separately as known functions—notice how you can dissect them and then put them back together again at the level of the rate of change!), and use these in $F = ma$ to solve for the acceleration,

$$a = \frac{d^2y}{dt^2} = \frac{F(t)}{m(t)} = G(t). \tag{2}$$

This is typical: The fundamental law stated in terms of rates of change leads to an ODE which must be solved to obtain the graph of the solution and the subsequent predictions. (You've already done this, so there is no point putting in explicit functions—the principle is all we are after here!) Because you can solve for a derivative in terms of known functions, the ODE (2) is of the simplest form, and can be solved directly by the FTC. That is, setting $y'(t) = v(t)$, $v'(t) = a(t)$ as the position and velocity, we use the Fundamental Theorem of Calculus and integrate once to obtain

$$v(t) = \int_{t_0}^{t} G(\xi)d\xi + v_0,$$

and then a second time to obtain

$$y(t) = \int_{t_0}^{t} v(\eta)d\eta + y_0 = \int_{t_0}^{t}\int_{t_0}^{\eta} G(\xi)d\xi d\eta + v_0(t - t_0) + y_0. \quad (3)$$

•**Conclude:** The equation (3) gives a formula for the height y as a function of time t for the rocket, that can be compared with experiment. In mathematics we would view this as an *exact* solution because it can be evaluated as a Riemann sum to arbitrary approximation. If the formula for $G(t)$ is simple enough, (often true if it can be expressed simply in terms of polynomials, sines, cosines, logs, or exponentials), then an antiderivative can be found, and the FTC can be applied to (3) to get a formula for the solution expressed in terms of the antiderivative. One strategy would be to approximate $G(t)$ in terms of polynomials, (every continuous function can be approximated arbitrarily well by polynomials), or in the case of Fourier analysis, in terms of sines and cosines, so that FTC can be applied directly to evaluate the associated approximate integrals in (3).

•Summary: The FTC provides a method for solving the simplest ODE's. But most of the important ODE's that arise cannot be so solved directly by integration because the unknown functions and their derivatives are entangled together in the equations and cannot be solved for in terms of known functions to integrate.

2. **Three Fundamental ODE's**

• An Ordinary Differential Equation is a in equation involving an unknown function and its derviatives.

– The *order* of an ODE is the highest order derivative appearing in the equation.

– The main distinction to make: LINEAR vs NONLINEAR

•Examples: The three most important ODE's in order of importance are: (We use $\dot{y}$ to denote the derivative $y' = dy/dt$ when t is the independent variable.)

(1) $\dot{y} = ky$ (The Exponential)

(2) $\ddot{y} + a^2 y = 0$ (The Harmonic Oscillator)

(3) $\dot{y} = y^2$ (Ricotti Equation)

---

(1) Recall the exponential function: $y = e^t$. This is *the only function whose derivative is itself:*

$$\frac{dy}{dt} = e^t.$$

• More generally: let $k \in \mathcal{R}$ be a constant (positive or negative). Then by the chain rule:

$$y = y_0 e^{kt} \quad \Rightarrow \quad \dot{y} = k y_0 e^{kt}$$

so that e^{kt} satisfies

$$\dot{y} = ky.$$

• This is the most important ODE. *If the rate of change of $y(t)$ is proportional to $y(t)$, then $y(t)$ grows or decays exponentially.*

• Said differently, the ODE e^{kt} satisfies tells us the most important property of the exponential: namely, e^{kt} *grows at a rate proportional to its size.*

• Population growth: The simplest model for population growth is that the population $P(t)$ should grow at a rate proportional to its size: So $P'(t) = kP(t)$, and we know $P(t) = P_0 e^{kt}$, where $P_0 = P(0)$ is the initial population.

• The simplest model for population growth is that the population $P(t)$ should grow at a rate proportional to its size: So $P'(t) = kP(t)$, so we know $P(t) = P_0 e^{kt}$.

• Now a first order ODE (involving smooth functions) has a unique solution satisfying one initial condition

$$\dot{y} = ky, \quad y(0) = y_0.$$

And we know that the function $y(t) = y_0 e^{kt}$ satisfies both:

$$\frac{dy}{dt} = y_0 \frac{d}{dt} e^{kt} = y_0 k e^{kt} = ky; \qquad y(0) = y_0 e^{k \cdot 0} = y_0.$$

So $y(t) = y_0 e^{kt}$ must be the unique solution to $\dot{y} = ky,\ y(0) = y_0$.

• Note that you cannot use the Fundamental Theorem of Calculus (FTC) directy to solve $\dot{y} = y$...i.e.,

$$y(t) = \int_0^t ky(\xi) d\xi + y_0,$$

fails because the unknown function ends up under the integral sign!

• The problem: you can only solve an ODE directly by the FTC if a known function is on the RHS and a single derivative is on the LHS, i.e., like

$$\frac{d^2y}{dt^2} = G(t),$$

for some known function $G(t)$. NOT: $\frac{d^2y}{dt^2} + y = G(t)$, or any other way the unknown and its devivatives get entangled in the equation.

• **Conclude:** Solving ODE's is hard! It requires being clever and there is no general formula like FTC for the solutions. Typically, for hard nonlinear ODE's, we give up even trying to get exact solutions, and settle for the qualitative behavior of solutions. Often, the qualitative information is more important than the exact formula anyway! The method of separation of variables is the most important method for solving linear and nonlinear scalar ODE's. It only works when t and y can be *separated*, but that's a lot of fundamental equations, both linear and nonlinear.

• Example: Solve $\dot{y} = ky$, $y(0) = y_0$ directly, without the formula. Solution: separation of variables:

$$\frac{y'}{y} = k \Rightarrow \int_0^t \frac{y'(\xi)}{y(\xi)} d\xi = \int_0^t k d\xi = kt$$

Substitute: $u(\xi) = y(\xi)$, $du = y' d\xi$:

$$\int_{u(0)}^{u(t)} \frac{du}{u} = \ln |u| \big]_{u(0)}^{u(t)} = \ln \frac{|u(t)|}{|u(0)|} = \ln \frac{y(t)}{y(0)}.$$

(Note that $y(t)$ cannot be zero, and hence cannot change sign!)

$$\ln \frac{y(t)}{y(0)} = kt,$$

and exponentiating both sides leads to

$$y(t) = y(0)e^{kt},$$

thereby deriving the answer we already knew!

---

(2) Consider now the Harmonic Oscillator $\ddot{y} + a^2 y = 0$. Recall the trig functions

$$\begin{aligned} y &= \sin t, & y &= \cos t \\ \dot{y} &= \cos t, & \dot{y} &= -\sin t \\ \ddot{y} &= -\sin t, & \ddot{y} &= -\cos t. \end{aligned}$$

Thus both $y = \sin t$ and $y = \cos t$ solve the ODE $\ddot{y} + y = 0$.

• More generally:

$$\begin{aligned} y &= \sin at, & y &= \cos at \\ \dot{y} &= a\cos at, & \dot{y} &= -\sin at \\ \ddot{y} &= -a^2 \sin t, & \ddot{y} &= -a^2 \cos t. \end{aligned} \tag{4}$$

Thus both $y = \sin at$ and $y = \cos at$ solve the general Harmonic Oscillator, the ODE

$$\ddot{y} + a^2 y = 0. \tag{2}$$

• Now the equation that sines and cosines satisfy, in some ways, tell us more about sines and cosines than their definitions as points on the unit circle. In fact, the equation they solve tells us the real reason why sines and cosines are fundamental to science. (It is not because of surveying and navigation, the reasons often given in introductory trigonometry texts!) The real reason is that everything in the universe is oscillating, and sufficiently small oscillations solve the Harmonic Oscillator, and hence must oscillate sinusoidally. That is a very powerful point, fundamental to science.

• Thus we have a general solution of $\ddot{y} + a^2 y = 0$ given by

$$y(t) = A \sin at + B \cos at.$$

Note that we can add two solutions together to get a general solution because the equation $\ddot{y} + a^2 y = 0$ is *linear* (i.e., the equation is a sum of terms each of which is no worse than a known function of t—in this case a constant—times y or a derivative of y) and *homogeneous* (zero is a solution because there is no function of t term like $g(t)$ sitting alone by itself, not multiplied by y or a derivative of y).

• Since a second order ODE (involving smooth functions) is determined by two initial conditions $y(0) = y_0$, $\dot{y}(0) = \dot{y}_0$, we can choose the constants A and B to meet any two initial conditions:

$$\begin{aligned} y_0 = y(0) &= A \sin(a \cdot 0) + B \cos(a \cdot 0) = B, \\ \dot{y}_0 = \dot{y}(0) &= Aa \cos(a \cdot 0) - Ba \sin(a \cdot 0) = aA, \end{aligned}$$

so solving for A and B in terms of the initial conditions gives

$$A = \frac{\dot{y}_0}{a}; \quad B = y_0.$$

The general solution of the Harmonic Oscillator $\ddot{y} + a^2 y = 0$ is thus:

$$y(t) = \frac{y_0'}{a} \sin at + y_0 \cos at.$$

• Note that all solutions of the Harmonic Oscillator really are exact sinusoidal oscillations. To see this, rewrite the general solution as

$$y(t)=\sqrt{A^2+B^2}\left(\frac{A}{\sqrt{A^2+B^2}}\sin at+\frac{B}{\sqrt{A^2+B^2}}\cos at\right),$$

and note that $\left(\frac{A}{\sqrt{A^2+B^2}},\frac{B}{\sqrt{A^2+B^2}}\right)=(\cos\theta,\sin\theta)$, because it is a unit vector, and hence points to the point on the unit circle θ radians around from the point $(1,0)$. It follows then that

$$\begin{aligned} y(t) &= \sqrt{A^2+B^2}\,(\cos\theta\sin at+\sin\theta\cos at) \\ &= \sqrt{A^2+B^2}\;2\sin\left(at+\theta\right). \end{aligned}$$

Conclude: The general solution of the harmonic oscillator is a true and exact *sinusoidal oscillation*! (We could just as well use cosines because $\cos\left(\frac{\pi}{2}-x\right)=\sin x$.)

• The equation $\ddot{y}+a^2y=0$ is a homogeneous, constant coefficient, second order, linear equation. In particular, *homogeneous* and *linear* means the principle of *superposition* holds. That is, sums of multiples of solutions are also solutions. That is, the solution space is a vector space: it is closed under addition and scalar multiplication of solutions. That means that you can put complicated solutions together from simpler ones, and pull complicated solutions apart into components. Nothing like this is true for nonlinear equations. For example adding two solutions is a solution:

$$\begin{aligned} y_1''(t)+a^2y_1(t)=0, \\ y_2''(t)+a^2y_2(t)=0, \end{aligned} \tag{5}$$

so adding gives

$$\left(y_1+y_2\right)''(t)+\left(y_1+y_2\right)(t)=0.$$

Note that if it were linear but not homogeneous, the solution space would not be a vector space, zero would not be a solution and superposition would not hold. For example,

$$\begin{array}{c} y_1''(t) + a^2 y_1(t) = g(t) \\ y_2''(t) + a^2 y_2(t) = g(t) \end{array} \tag{6}$$

so adding gives

$$(y_1 + y_2)''(t) + (y_1 + y_2)(t) = 2 \neq g(t).$$

But in this case, the difference between two solutions would solve the homogeneous equation. Thus for inhomogeneous linear equations, the solution space consists of any *particular* solution of the inhomogeneous equation, plus any solution in the vector space of homogeneous solutions.

• Just as the ODE $y' = ky$ tells us that exponential growth occurs whenever *the rate of change of y is proportional to y itself*, so also the ODE $\ddot{y} + a^2 y = 0$ tells us why sines and cosines are of fundamental importance to science. The main point: we now show that whenever a mass is subject to a restoring force, if the resulting oscillations are weak enough, then they evolve according to the Harmonic Oscillator $\ddot{y} + a^2 y = 0$. Thus, everything in the universe is vibrating, and every vibration, sufficiently weak, oscillates sinusoidally, because it solves the Harmonic Oscillator equation.

• To see this, assume a mass m is subject to a restoring force proportional to displacement:

$$Force = -ky.$$

For example, an ideal spring satisfies Hook's law: the force is *exactly* proportional to minus the displacement from equilibrium. So if we draw the y-axis as horizontal, and ask that $y = 0$ be the equilibrium position of a mass on a spring, then the force being $-y$ always points back toward the point of equilibrium. To solve for the position of the mass as a function of time, we start with the fundamental law, Newton's force law $F = ma$, stated in terms of rates of change:

$$m\frac{d^2y}{dt^2} = -ky,$$

or

$$\frac{d^2y}{dt^2} + \frac{k}{m}y = 0,$$

or

$$\ddot{y} + a^2 y = 0,$$

with $a^2 = k/m$.

• **Conclude:** Sinusoidal oscillations occur whenever the restoring force is proportional to displacement. We now show that whenever there is a restoring force, for weak enough oscillations, the restoring force is *always* approximately linear, and hence the mass will again *always* oscillate sinusoidally in the limit of weak vibrations. Well, that is just about everything that vibrates. Hello sinusoidal oscillations!

• So consider a general *restoring force*. By this we mean a force $F(y)$ that satisfies the three conditions:

(i) $F(0) = 0$.

(This says there should be a equilibrium position, which without loss of generality we can take as $y = 0$.)

(ii) $F(y) = F(0) + F'(0)y + Error, \quad |Error| \leq Const \cdot y^2.$

(This says the force should be smooth enough, i.e., have enough derivatives, so that Taylor's theorem with remainder should hold. Such an Error is said to be $O(y^2)$... pronounced *Big O of* y^2.)

(iii) $F'(0) = -k < 0$.

This says the force should be *restoring* in the sense that upon displacement, the force should point back toward the equilibrium position $y = 0$, so that the mass really oscillates around equilibrium.

It follows then, that for y sufficiently small, $y^2 << y$, and so $Error << 1$, and the force is well approximated by $F(y) = -ky$, just like the ideal spring satisfying Hooke's Law. So the oscillation looks like a solution of $\ddot{y} + a^2 y = 0$ with $a^2 = k$, which we have shown above is always sinusoidal.

Conclude: The equation $\ddot{y}+a^2y = 0$ satisfied by sines and cosines connects these basic functions of trigonometry to a fundamental principle of physics, the Harmonic Oscillator. Sines and cosines are of fundamental importance because everything in the universe is vibrating, and in the limit of weak vibrations, *all* vibrations oscillate sinusoidally. The ODE that sines and cosines satisfy give the deepest and most interesting property of them...

• Finally note that the second order Harmonic Oscillator $y''+a^2y = 0$ can be written as a first order system of equations, and that is how second order equations are generally studied. For example, setting $u = y$ and $v = y'$, we obtain

$$\begin{aligned} u'(t) &= y, \\ v'(t) &= y'' = -a^2 y = -a^2 u, \end{aligned}$$

so we can rewrite the $y'' + a^2 y = 0$ as a first order system in matrix form:

$$\begin{pmatrix} u \\ v \end{pmatrix}' = \begin{pmatrix} 0 & 1 \\ -a^2 & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}. \tag{7}$$

By the same procedure, an ODE of any order can be written as a first order system. And linear constant coefficient linear homongeneous sytems like the Harmonic Oscillator can be written in the matrix form

$$\mathbf{y} = \begin{pmatrix} y_1 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{pmatrix}' = \begin{pmatrix} a_{11} & . & . & . & a_{1n} \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ a_{n1} & . & . & . & a_{nn} \end{pmatrix} = A\mathbf{y}.$$

Eigenpairs (λ, R) that satisfy $AR = \lambda R$, (both real and imaginary), decouple a linear constant coefficient ODE in matrix form and lead to a basis for the vector space of solutions. Indeed, if $\mathbf{y} = \mathbf{R}e^{\lambda t}$, then

$$\mathbf{y}' = \mathbf{R}\lambda e^{\lambda t} = A\mathbf{R}e^{\lambda t} = A\mathbf{y},$$

satisfies the equation $\mathbf{y}' = A\mathbf{y}$, as claimed.

---

(2) Consider finally the third most important equation, the simplest nonlinear ODE $\dot{y} = y^2$, the so called Ricotti Equation. This is first order, homogeneous, and *nonlinear* because of the nonlinear y^2 on the right hand side. This is important because it shows how badly things can go wrong when an equation in *nonlinear*.

Note first that because the equation is nonlinear, superposition *does not hold*:

$$\begin{aligned} y_1'(t) &= y_1^2(t), \\ y_2'(t) &= y_2^2(t), \end{aligned} \tag{8}$$

so adding gives

$$(y_1 + y_2)'(t) = y_1^2(t) + y_2^2(t) \neq (y_1 + y_2)^2(t).$$

On the other hand, $y = 0$ is a constant equilibrium solution (called a *rest point* of the equations).

• Now most nonlinear equations cannot be solved in closed form, but once again we can solve this one by separation of variables:

$$\frac{y'}{y^2} = 1,$$

$$\int_0^t \frac{y'(\xi)}{y(\xi)} d\xi = \int_0^t d\xi = t,$$

so applying Leibniz's substitution principle with

$$u = y(\xi), \quad du = y'(\xi) d\xi$$

gives

$$\begin{aligned}\int_0^t \frac{y'(\xi)}{y(\xi)} d\xi \quad &= \textstyle\int_{u(0)}^{u(t)} \frac{du}{u^2} = -u^{-1}\big]_{u(0)}^{u(t)} = -\frac{1}{u(t)} + \frac{1}{u(0)} \\ &= -\textstyle\frac{1}{y(t)} + \frac{1}{y(0)}\end{aligned}$$

which upon solving for $y(t)$ gives the famous solution formula

$$y(t) = \frac{1}{\frac{1}{y_0} - t}.$$

• Conclude that since the first order ODE $\dot{y} = y^2$ should be determined by one initial condition $y(0) = y_0$, the unique solution must be given by the formula

$$y(t) = \frac{1}{\frac{1}{y_0} - t}.$$

In particular, from the formula you see that the denominator tends to zero as $t \to 1/y_0$, and so a solution does not exist for all t. In fact, starting at $t = 0$, a solution in general only exists for $0 \leq t < 1/y_0$, and this interval of existence gets smaller and smaller as we choose a larger and larger initial value y_0.

• **Conclude:** For linear ODE's almost nothing can go wrong: solutions always exist for all time and can be constructed by superposition. For nonlinear equations, terrible things can go wrong. In fact, solutions can *blow up* in an arbitrarily short amount of time, even when the nonlinearity is quadratic, the simplest possible. But this is not a *bad* thing, it is actually *wonderful*, because its tells us that interesting surprises are waiting to be discovered in nonlinear equations, and the mathematical surprises lead to surprising physical predictions! (Like shock-waves and black holes!)

3. The Harmonic Oscillator and the Waves they Propagate

—We have seen that the sinusoidal solutions of

$$\ddot{y} + a^2 y = 0$$

describe pretty much anything thing that vibrates—when a mass is in a restoring force,

$$\ddot{y} + a^2 y = 0$$

becomes the equation when the oscillations are sufficiently weak. Since everything in the universe is vibrating, the next question is, what is the equation that propagates the sinusoidal waves created by all these Harmonic Oscillators?

So consider sound waves generated by a Harmonic Oscillator at $x = 0$. (We want our picture on the horizontal x-axis, so call the dependent variable x instead of y). So assume an oscillator positioned at $x = 0$ along a horizontal x-axis, creates a sinusoidal variation in density around a constant density background $\bar{\rho}$. That is, assume

$$\rho(t) - \bar{\rho} = -\epsilon \sin(at)$$

is the density vibration created by the oscillator at $x = 0$ around the constant density ρ_0, and assume these variations in density propagate away to the right as sound waves moving at the speed of sound c. Now since density can be measured relative to any constant density, to make things simpler, lets assume $\rho(t)$ is the *difference* between the density and the ambient density $\bar{\rho}$, which means $\bar{\rho} = 0$, and our equation for the oscillating density is

$$\rho(t) = -\epsilon \sin at,$$

the general solution created by an harmonic oscillation. (We put $-\epsilon$ for convenience. This only changes the sign of the constant ϵ, which can have either sign, so there is no loss of generality. Since ϵ can have either sign, there is no loss of generality in also assuming $\alpha > 0$.) The amplitude ϵ is included to make the sinusoidal oscillations as small as we like.

Claim: A density profile $f(x)$ propagating to the right at speed c has the time-dependent profile

$$\rho(x,t) = f(x - ct).$$

To see this, ask for the speed of such a signal. That is, ask:

What is the speed of propagation of the constant values of f in the function $f(x - ct)$?

Answer: $f(x - ct)$ is constant when the input $x - ct$ is constant. Since the curve

$$x - ct = const$$

moves at speed

$$dx/dt = c,$$

it follows that

$$\rho = f(x - ct)$$

takes a constant value along each line

$$x - ct = const,$$

that is, along each line of speed c. That does it. It has to be. A density $\rho = f(x)$ at $t = 0$ propagating at speed c looks like

$$\rho(x,t) = f(x - ct).$$

So what is the sound wave $\rho(x,t)$ that propagates the harmonic oscillation of density

$$\rho(t) = -\epsilon \sin(at)$$

at the boundary $x = 0$ away to the right at speed c? We need the function $f(x - ct)$ so that $\rho(x,t) = f(x - ct)$ matches up with the oscillating density

$$\rho(t) = -\epsilon \sin(at)$$

at $x = 0$, which means

$$\rho(x,t) = +\epsilon \sin\left[\frac{a}{c}(x - ct)\right].$$

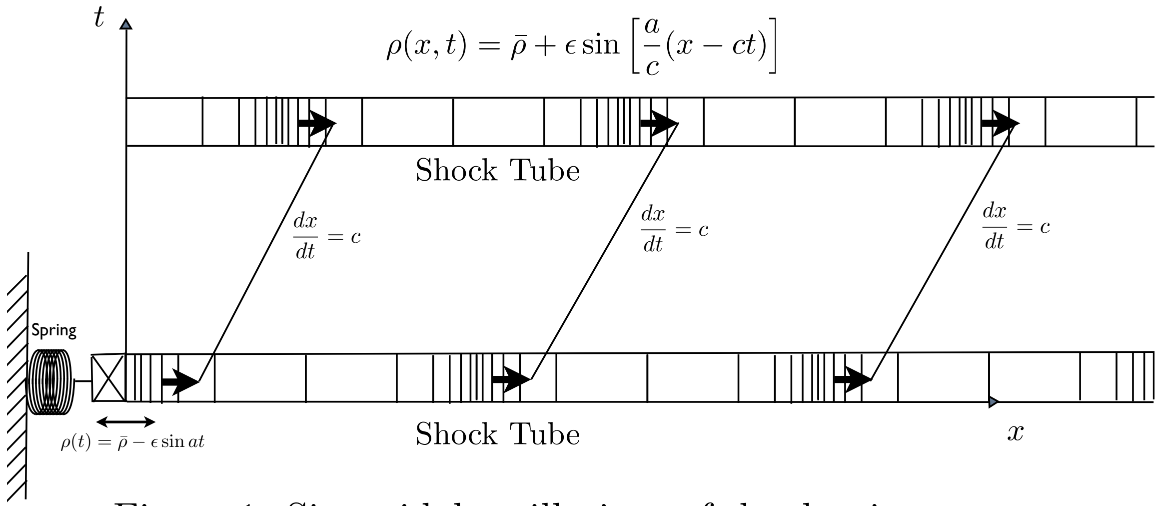

Figure 1: Sinusoidal oscillations of the density propagating at speed c by the transport equation $u_t - cu_x = 0$.

Conclude that

$$f(x - ct) = \epsilon \sin\left[\frac{a}{c}(x - ct)\right]$$

is the sound wave propagating to the right such that it matches up with the harmonic oscillation generating it at $x = 0$. The spatial profile of the wave at $t = 0$ is then

$$\rho(x) = \epsilon \sin \frac{a}{c} x.$$

Now the equation $\ddot{x} + a^2 x = 0$ for the Harmonic Oscillations is fundamental, so the equation that propagates these disturbances to the right at sound speed c must also be fundamental. So what is the simplest equation a right moving sound wave $\rho(x - ct)$ satisfies? Answer: The transport equation:

$$u_t + cu_x = 0.$$

Lets check:

$$u(x,t) = f(x - ct)$$

satisfies

$$u_t = \frac{\partial}{\partial t} u(x,t) = -cf'(x - ct),$$

and

$$u_x = \frac{\partial}{\partial x} u(x,t) = f'(x - ct),$$

so

$$u_t + cu_x = -cf'(x - ct) + cf'(x - ct) = 0.$$

Conclude: The fundamental equation that propagates signals to the right at speed c is the transport equation

$$u_t + cu_x = 0.$$

Solutions are of the form

$$u(x,t) = f(x - ct)$$

solve the transport equation and *are constant along lines of speed* c—hence represent sound waves moving to the right at speed c.

Now the transport equation

$$u_t + cu_x = 0$$

is, like the harmonic oscillator, *linear* and *homogeneous*, so the principle of superposition holds: Sums of multiples of solutions are also solutions. Since Harmonic Oscillators generate sinusoidal vibrations, it is natural in the theory of sound to decompose the right going waves into a sum of sinusoidal waves. This is called a Fourier decomposition of the sound wave, and keeps track of all the different frequencies created by harmonic oscillators with different $a's$.

It works like this: A function $f(x)$ can be decomposed into a sum of sine waves:

$$f(x) = \sum_{n=0}^{\infty} \epsilon_n \sin nx.$$

In this case, the $a's$ are restricted to $a = n$ because these are the frequencies that arise when a string is vibrated. As a result, the sound wave is

$$\rho(x,t) = f(x - ct) = \sum_{n=0}^{\infty} \epsilon_n \sin n(x - ct).$$

The main point to make here is that since the equation is linear, each term in the sum satisfies the equation, so the whole sum does too. (You have to worry about the infinite sum, but let's skip that detail for now.)

Conclude: A linear sound wave can carry all of the frequencies generated by the different Harmonic Oscillators, and the signal remains coherent: you can add them all up at one end, propagate them across the room, and decompose them into the same elements at the other end. Linear superposition is remarkable: That is why we can hear sounds clearly. Our ear simply breaks down the complicated signal into its elemental frequencies, and by this we can hear all the Harmonic Oscillators all over the room!

But all of this only works for weak signals where the *linear equations* apply. When the signal is strong, the evolution equation becomes *nonlinear*, and shock-waves and shock wave dissipation takes over!

- So consider the transport equation

$$u_t + cu_x = 0,$$

in the case of a variable sound speed c. If $c = c(x,t)$, then the equation is still a *linear* equation, but if $c = c(u)$

depends on the unknown solution u itself, (for example the Burgers equation is $u_t + uu_x = 0$ with $c(u) = u$), then the equation is a *nonlinear* equation. In either case, the equation still states that *signals (sound waves) propagate at speed c.* We can see this two ways:

First, to see it in a nutshell, write the equation as

$$u_t + cu_x = cu_x + u_t = \nabla_{(c,1)}u(x,t) = 0,$$

which says in words that *the solution $u(x,t)$ is constant in the direction $\overrightarrow{(x,t)} = \overrightarrow{(c,1)}$*, that is, along lines of speed

$$\frac{dx}{dt} = c. \tag{9}$$

If it is a linear equation, then $c = c(x,t)$, and (9) defines a non-autonomous scalar ODE of the form

$$\frac{dx}{dt} = c(x,t)$$

for the curves in the (x,t)-plane along which the solution u is constant. Such curves, called the *characteristics* of the PDE, are the one's with dx/dt=speed=inverse slope c at each (x,t) in the (x,t)-plane.

On the other hand, if $c = c(u)$ and the equation is *nonlinear*, then (9) says that the solution $u = u(x,t)$ *is constant along lines of speed* $c(u)$. That is: starting with the initial data $u(x,0)$ at $t = 0$, the sound wave starting at position x_0 at time $t = 0$ propagates the constant state $u_0 = u(x_0,0)$ as constant along the straight line of speed $dx/dt = c(u_0)$. **Conclude:** Even in the nonlinear case when the speed $c = c(u)$ depends on the unknown solution u, still, solutions propagate along curves of speed $c(u)$ in the (x,t)-plane.

A second way to see that solutions $u(x,t)$ of the transport equation (9) are constant along the curves $dx/dt = c$, is to ask for the curves $(x(t), t)$ along which the solution is constant. That is, ask that

$$\frac{d}{dt}u(x(t), t) = 0.$$

But using the chain rule this becomes

$$\frac{d}{dt}u(x(t), t) = u_x\dot{x}(t) + u_t = 0,$$

and so assuming $u(x, t)$ is a solution of (9) and comparing terms gives

$$\dot{x}(t) = c.$$

Conclude: Even in the case of a variable speed c, the solutions of the transport equation (9) are constant along the curves in the (x, t)-plane moving at speed c. These curves are called the *characteristic curves*, or just *characteristics* of the PDE (9).

—We now compare the difference between the *linear* transport equation (9), the case when the speed $c = c(x, t)$ depends on the independent variables (x, t), to the *nonlinear* transport equation, the case when $c = c(u)$ depends on the unknown solution u. The conclusion will be that nonlinearities can create *shock-waves*, but no such thing can happen when the equation is linear.

• Consider first the *nonlinear* case. To see what goes wrong, (really, it goes right, because it make things very interesting!), assume the simplest nonlinearity $c(u) = u$, in which case (9) is the so called *inviscid Burgers Equation*:

$$u_t + uu_x = 0. \tag{10}$$

Note first that this is obtained form the conservative form

$$u_t + \frac{1}{2}\left(u^2\right)_x = 0 \tag{11}$$

by differentiation. As we will see, the advantage of the conservative form, (a form in which all terms fall under a total derivative), is that we can use this formulation to extend the solutions beyond the point of shock wave formation. But first let's see how the shock-waves form in the first place.

The point is that equation (10) states that solutions $u(x,t)$ propagate as constant along curves of speed $dx/dt = u$, which are straight lines because, when u is constant, the speed is constant. But that means there is a consistency problem at a future time if at time $t = 0$,

$$u(x_1, 0) = u_L > u_R = u(x_2, 0)$$

with $x_1 < x_2$. That is, the bigger speed=initial state u_L is positioned initially at x_1 to the *left* of the smaller speed=initial state u_R, (the state u is the speed of the characteristic along which u is constant), at x_2. Thus the speed of the characteristic emanating from x_1 at $t = 0$ is u_L, the speed of the characteristic emanating from x_2 is u_R, so $u_L > u_R$ implies that the two characteristic will intersect at some future point $P = (x_*, t_*)$. So the solution coming from the left characteristic says $u = u_L$ at P, the solution coming in from the right characteristic says $u = u_R$ at P, and you can't have *one* solution taking *two* different values at the same point in spacetime–its *hello Houston, we have a problem!* What turns out to be so interesting is that this is exactly what happens in the compressible Euler equations, and we can continue the solution as a shock wave discontinuity starting at P. At that point the theory gets very interesting.

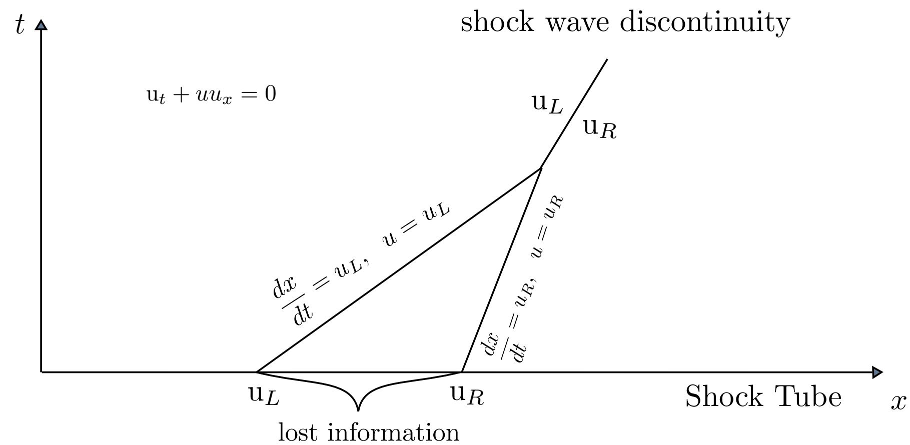

Figure 2: Nonlinear Burgers: "*u=const. along lines of speed u*" implies shock formation and loss of information

• Let's now compare the transport equation (9) in the nonlinear case $c = c(u)$ with the linear case $c = c(x,t)$. Recall that the nonlinear ODE $\dot{y} = y^2$ exibited *blowup in finite time*, but the linear equations all had solutions defined for all time. Our purpose now is to show that in the case of a linear transport equation we can never get shock-waves. **Conclude:** *shock-waves are a nonlinear phenomenon in PDE's.*

We now show that the non-existence of shock-waves for the linear transport equation is a consequence of the existence and unique theorem for ODE's, thereby motivating our next topic. This is because the transport equation $u_t + cu_x = 0$ says solutions $u(x,t)$ propagate as constant along the characteristic curves $dx/dt = c$ in the (x,t)-plane. Thus to rule out shock-waves, we only need to show that two characteristic curves satisfying

$$\frac{dx}{dt} = c(x,t)$$

can never cross in the (x,t)-plane when $c = c(x,t)$ is a function of the independent variables (x,t) alone. That is, the shock wave formed when two characteristics intersected at a point $P = (x_*, t_*)$, thereby determining one value u_L from the left and a different value u_R from the right...leaving a shock wave that jumps from u_L to u_R as the only way to continue the solution forward in time from the point P.

So consider the problem of showing that two solutions of the initial value problem

$$\dot{x} = c(x,t),$$

starting from two different initial values $x(0) = x_1$ and $x(0) = x_2$ never cross when $x_1 \neq x_2$. Now at first we might think things go wrong when $c = x^2$, because then the characteristic equation

$$\dot{x} = x^2,$$

is the Ricotti equation and hence solutions blow up in finite time. But really, just because the characteristic curve goes off to infinity in the (x,t)-plane at finite time, doesn't mean there is any real problem with the solution $u(x,t)$ being constant along that curve. Now we could make a physical condition that the speed $c(x,t)$ should remain bounded to rule out such things, and that would work. But let's be more systematic and show that two characteristics can *never* intersect in the (x,t)-plane when the equation is linear.

So to be more general, consider the linear first order equation

$$b(x,t)u_t + a(x,t)u_x = 0, \tag{12}$$

where $a(x,t)$ and $b(x,t)$ are assumed to be given and known, but otherwise arbitrary (smooth) functions of (x,t). (When $a = 1$, $b = c$ we're back to the special case above.) In this case, let us ask for the characteristic curves in the (x,t) plane along which a solution $u(x,t)$ is constant. Now a general curve in the (x,t)-plane can be parameterized by a real variable $\xi \in \mathcal{R}$, such a general curve being of form

$$(x(\xi), t(\xi)), \quad \xi \in \mathcal{R}.$$

Thus we ask when a solution $u(x,t)$ of (12) is constant along such a curve, i.e., when

$$\frac{d}{d\xi}u(x(\xi), t(\xi)) = u_x x'(\xi) + u_t t'(\xi) = 0.$$

The answer we get upon comparing the coefficients with (12) is:

$$\begin{pmatrix} x'(\xi) \\ t'(\xi) \end{pmatrix} = \begin{pmatrix} a(x,t) \\ b(x,t) \end{pmatrix}, \tag{13}$$

an *autonomous* nonlinear system of two equations in the plane, asking for a solution pair $(x(\xi), t(\xi)$ that solves (13). The ODE (13) is of the general form

$$\mathbf{y}' = f(\mathbf{y}), \tag{14}$$

where $\mathbf{y}$ denotes a vector in the (x,t)-plane

$$\mathbf{y} = \begin{pmatrix} x \\ t \end{pmatrix}$$

and

$$f(x,t) = \begin{pmatrix} a(x,t) \\ b(x,t) \end{pmatrix},$$

is the known vector valued function (or *vector field*) which defines the equations.

Now for the general problem, we go back to the notation of the general theory, in which case we take the *dependent* (unknown) variables $\mathbf{y} = (x,y)$, and in place of ξ we use

t instead of ξ to denote the *independent* variable. (For characteristic curves of a PDE $bu_t + au_x = 0$, the unknown curve is in the space-time plane (x,t)), so having used t for an unknown, we had to introduce a new variable $\xi \in \mathcal{R}$ as the parameter along the unknown curve $x(\xi), t(\xi)$...but in general applications, it's easier to think of an unknown curve in the (x,y)-plane and use t as the parameter, in which case our notation is that $(x(t), y(t))$ is the general unknown curve, and $f(\mathbf{y}) = (f_1(\mathbf{y}), f_2(\mathbf{y}))$ as the known vector field that defines the equations.) With this notation, (14) is *autonomous* because there is no explicit dependence on the independent variable t in the function $f(\mathbf{y})$. That is, it is not of the more general form $f(\mathbf{y}, t)$, the dependence of f on t being only through the unknown solution $\mathbf{y}(t)$.

Autonomous systems of two equations $\dot{\mathbf{y}} = f(\mathbf{y})$ are extremely important because one can graph the solutions explicitly in the (x,y)-plane. In this case, the (x,y) plane is called the *phase plane*, and the graph of solutions in the (x,y)-plane is called the *phase portrait* of the solutions. One reason this is such an important setting is because one can easily visualize the solutions, and the meaning of the equations. That is, let $f(\mathbf{y}) = \overrightarrow{(a(x,y), b(x,y))} = \vec{\mathbf{F}}(x,y)$, where $\vec{\mathbf{F}}$ is a vector valued function of (x,y) called a *vector field*. Now although the equations are nonlinear and so $\vec{\mathbf{F}}$ is unknown until the solution is know, the vector field $\vec{\mathbf{F}}$ *is* a known function of the position (x,y). This is because the equations are autonomous, and $\vec{\mathbf{F}}$ does not depend on the time t at which a solution takes the position (x,y). Thus the equation $\dot{\mathbf{y}} = \vec{\mathbf{F}}(x,y)$ asks for a curve $\mathbf{y} = (x(t), y(t))$ whose tangent vector $\dot{\mathbf{y}} = \vec{\mathbf{v}}$ is tangent to $\vec{\mathbf{F}}$ at each point of the curve. If you imagine $\vec{\mathbf{F}}$ as a *known* arrow drawn at each (x,y), then the solution asks for a curve that makes $\vec{\mathbf{F}}$ tangent to the curve, $\vec{\mathbf{v}} = \vec{\mathbf{F}}$, so the direction and length=speed

$\|\vec{\mathbf{v}}\|$ of $\vec{\mathbf{v}}$ agrees with $\vec{\mathbf{F}}$ at every point. This is visualized in Figure 3, and this is the right way to think about solutions of an autonomous system of ODE's.

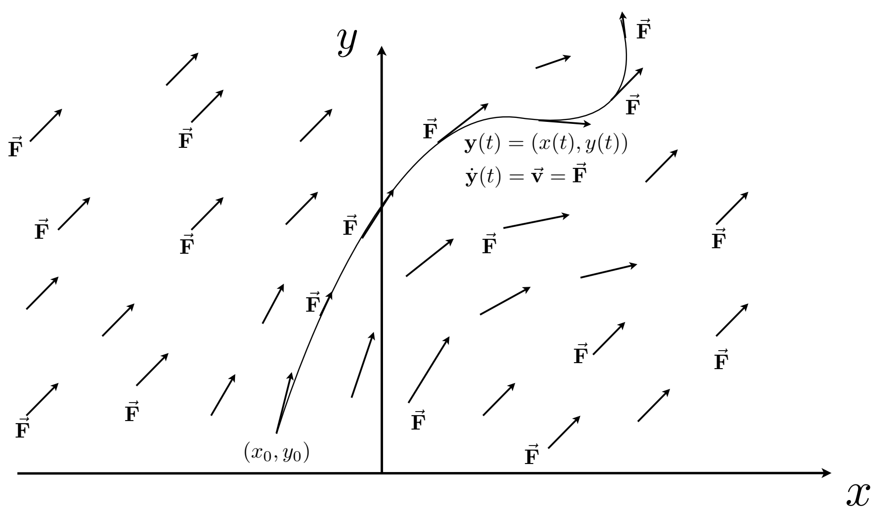

Figure 3

The reason this works so cleanly is mainly because there is always a unique solution curve $\mathbf{y}(t)$ passing through each point in the (x, y), and two different curves cannot intersect in the (x, y)-plane. This is what we now prove, and is exactly what we need to prove shock-waves never form in the general linear PDE $bu_t + au_x = 0$.

So here is the claim. Two solutions of an autonomous system $\dot{\mathbf{y}} = f(\mathbf{y})$ *cannot cross* in the (x, y)-plane, and this is a consequence of the existence and uniqueness theorem for ODE which we discuss in the next section. (This is arguably the most important direct consequence of the ODE theorem, but don't worry, for now just believe what the ODE theorem tells us: that two solutions of $\dot{\mathbf{y}} = f(\mathbf{y}, \mathrm{t})$ that agree at $t = t_0$, i.e. $\mathbf{y}_1(\mathrm{t}_0) = \mathbf{y_2}(\mathrm{t}_0)$, must, *as a logical consequence of our existence/uniqueness theorem*, agree for all t, i.e. $\mathbf{y}_1(\mathrm{t}) = \mathbf{y}_2(\mathrm{t})$. Again, for the general theorem, we go back to letting t be the independent variable, and $\mathbf{y} = (x, y)$.) To verify this, we first establish the following important property of solutions of *autonomous* systems:

• If $\mathbf{y}(t)$ is a solution of $\dot{\mathbf{y}} = f(\mathbf{y})$, then so is $\mathbf{y}(t - t_0)$.

That is, we can shift the time at which the initial condition is given. To see this just let $\tau = t - t_0$ and differentiate by the chain rule:

$$\frac{d}{dt}\mathbf{y}(t - t_0) = \frac{d}{d\tau}\mathbf{y}(\tau)\frac{d\tau}{dt} = f(\mathbf{y}(\tau)) \cdot 1 = f(\mathbf{y}(t - t_0)).$$

So $\mathbf{y}(t - t_0)$ is also a solution as claimed. (Note that this would *not* be true for non-autonomous systems of form $\dot{\mathbf{y}} = f(\mathbf{y}, t)$, i.e., when f depends explicitly on the independent variable t!)

• So assume that we have two solutions $\mathbf{y}_1(t)$, $\mathbf{y}_2(t)$ of the equation

$$\dot{\mathbf{y}}(t) = f(\mathbf{y}(t)), \tag{15}$$

and assume they cross, (which could happen at *two different times* t_1, t_2), at a point $\mathbf{y}_0 = (x_0, y_0)$ in the (x, y)-plane, so that

$$\mathbf{y}_1(t_1) = \mathbf{y}_0 = \mathbf{y}_2(t_2), \tag{16}$$

for two times t_1 and t_2, (see Figure 4).

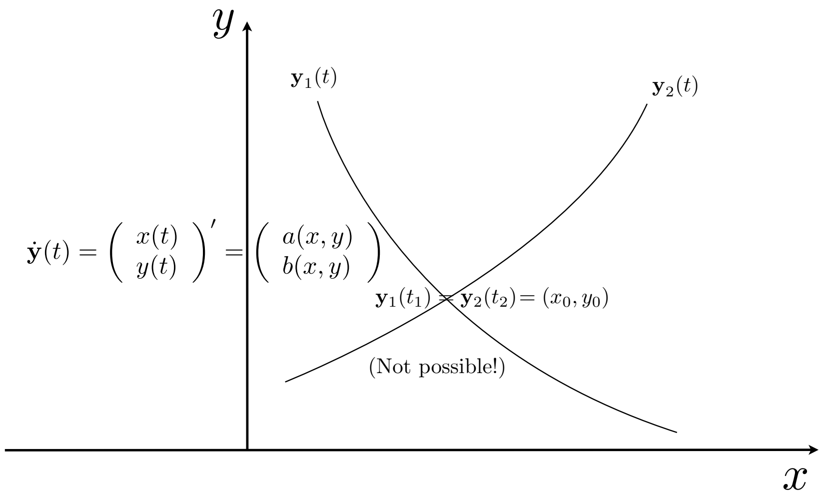

Figure 4

We show that either one solution is a re-parameterization of the other, (i.e., $\mathbf{y}_2(t) = \mathbf{y}_1(t + t_0)$ for constant $t_0 = t_2 - t_1$, so both $y_1(t)$ and $y_2(t)$ trace out the same curve in the (x, y)-plane), or else the initial value problem

$$\dot{\mathbf{y}} = f(\mathbf{y}), \tag{17}$$

$$\mathbf{y}(t_0) = \mathbf{y}_0, \tag{18}$$

has two distinct solutions, (in violation the existence uniqueness theorem of ODE's—next topic). For this, consider the solution $\bar{\mathbf{y}}_2(t) = \mathbf{y}_2(t + (t_2 - t_1))$. But by (16),

$$\bar{\mathbf{y}}_2(t_0) = \mathbf{y}_2(t_1) = \mathbf{y}_0.$$

Thus both $\bar{\mathbf{y}}_2(t)$ and $\mathbf{y}_1(t)$ solve the same equation $\dot{\mathbf{y}}(t) = f(\mathbf{y}(t))$ with the same initial conditions. But the existence and uniqueness theorem tells us there is only one such solution, so they must be the same solution, i.e., $\bar{\mathbf{y}}_2(t) = \mathbf{y}_1(t)$, (at least for t close to t_1, which is all we need!).

Conclude: Either $\bar{\mathbf{y}}_2(t) \equiv \mathbf{y}_2(t + t_0) = \mathbf{y}_1(t)$, and hence $\mathbf{y}_2(t)$ is just a a re-parameterization of $\mathbf{y}_1(t)$, or else the initial value problem (17) has two different solutions. Since the latter choice is ruled out, they must be the same solution, in which case two *different* solution curves cannot intersect at the same point in the (x, y) plane. As an application it follows that two characteristics along which solutions of $a(x,t)u_t + b(x,t)u_x = 0$ are constant cannot cross in the (x, t) plane. The case $b = 1$, $a = c(x, t)$ is the case $u_t + c(x, t)u_x = 0$ gives $t = \xi$, and so the characteristics

$$\dot{y} = c(x, t),$$

the curves moving at speed c in the (x, t)-plane, also can never intersect in the (x, t)-plane.

We have established: shock-waves cannot form in solutions of the linear equation $a(x, t)u_t + b(x, t)u_x = 0$.

Example: Describe the solution of the initial value problem

$$u_t + xu_x = 0, u(x, 0) = x.$$

Solution: Since its a linear PDE, we know we can solve it by characteristics and shock-waves will not form because the characteristics will never cross in the (x, t)-plane. The solutions are constant along curves $x(t)$ of speed $dx/dt = c(x, t) = x$, which has solution $x(t) = x_0e^t$.. Thus the solution starting at x_0 at $t = 0$ will continue constant along the curve $x = x_0e^t$.. These curves are graphed in Figure 5.

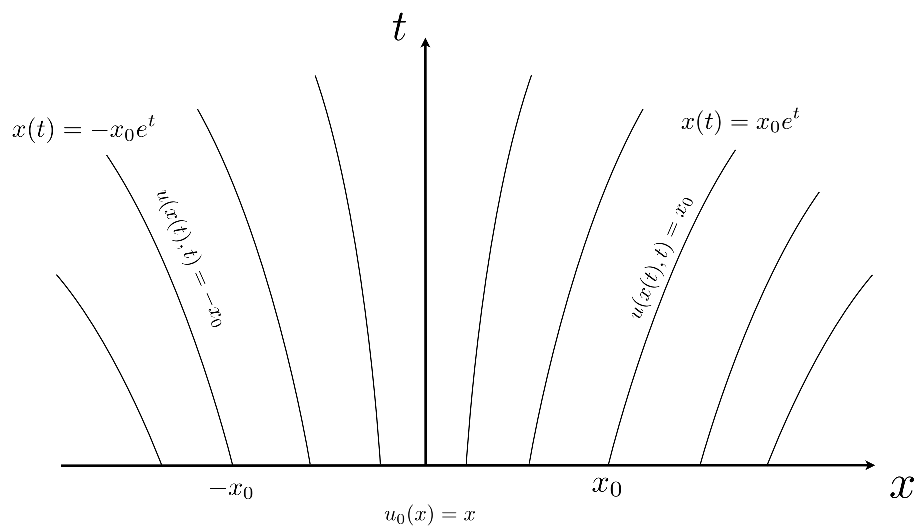

Figure 5

Example: Describe the solution of the initial value problem

$$u_t + x^2u_x = 0, u(x, 0) = x.$$

Solution: Since its a linear PDE, we know we can solve it by characteristics and shock-waves will not form because the characteristics will never cross in the (x, t)-plane. The solutions are constant along curves $x(t)$ of speed $dx/dt = c(x, t) = x^2$, which is our *nonlinear* ODE which blows up in finite time, and has solution satisfying $x(0) = x_0$ given by

$$x(t) = \frac{1}{\frac{1}{x_0} - t} = \frac{x_0}{1 - x_0t}$$

Thus the solution starting at x_0 at $t = 0$ will continue constant along the curve $x = \frac{x_0}{1-x_0t}$. These curves are graphed in Figure 6.

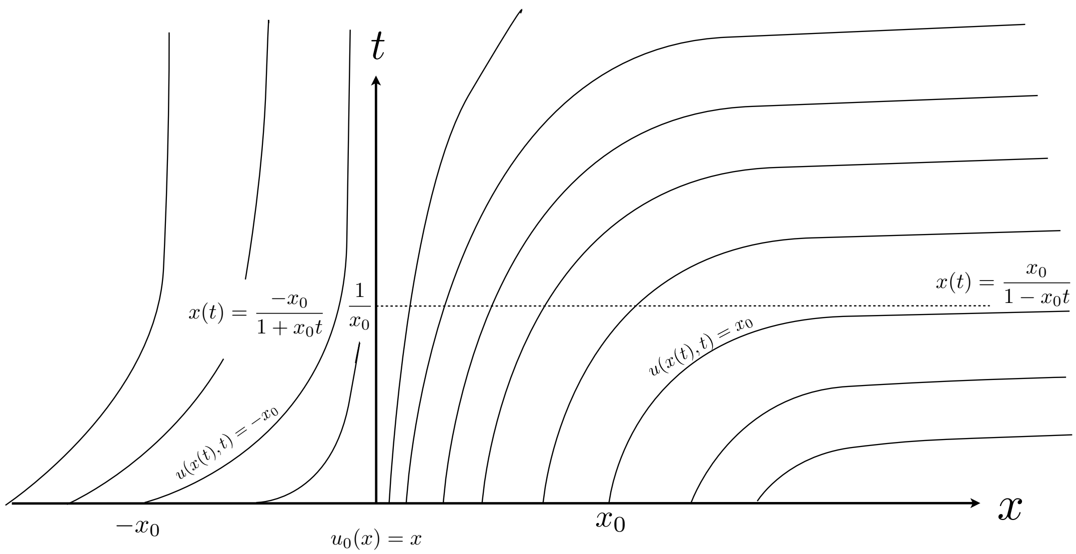

Figure 6

Note that the fact that $x(t) \to \infty$ as $t \to 1/x_0$ doesn't really matter to the linear PDE because it just says the lines of constant u tend to infinity at a finite value of t. The issue for characteristics is whether the solution can be defined until either $x \to \infty$ or $t \to \infty$; i.e., as long as the solutions don't cross, and run out to the end of the (x,t)-plane, the solution of the linear PDE is defined everywhere by assigning it by its initial value along each characteristic curve. We say the curves *foliate* the space-time, meaning they cover the plane, and never cross. The only way things can go wrong with the initial value problem when the x-axis $t = 0$ is itself a characteristic curve. Since this is a curve of infinite speed, it will never happen as long as $b = 1$ and the speed $a(x,t) = c(x,t) \neq \infty$. However it would happen if $b(x,t) = 0$) Thus the initial value problem always has a solution for all time as long as $b(x,t) \neq 0$, so the initial data is *non-characteristic*, that is, not assigned along a characteristic curve.

Conclude: The fact that nonlinear ODE's can blow up in finite time, is not an impediment to solutions of linear PDE's existing for all time by the method of characteristics!

Boundary Value Problems: Just as the characteristics tell us when you can assign initial values at $t = 0$, (say $u(x, 0) = u_0(x)$) and obtain a unique solution, they give us the more general principle that you can assign values of u on any curve that crosses the characteristics just once, because the solution of $au_t + bu_x = 0$ must take the assigned constant value at one point to all the other points along that characteristic curve. This principle also of characteristics also tells us which boundary value problems can be solved, and which ones cannot be solved because they are inconsistent. For example, in both Figures 5 and 6, the t-axis is a characteristic curve, and so one cannot assign values of u freely at the boundary $x = 0$ in either Example 1 or 2. On the other hand, since $c(x, t) > 0$ whenever $x > 0$, we can assign values of u arbitrarily along the boundary $x = x_0$, $t > 0$, together with initial data along $t = 0$ for $x \geq x_0$. The characteristics in Figures 5 and 6 show that this is consistent in both cases because the values of u assigned by these initial boundary values can propagate consistently as constant along the characteristics that cross the boundary. On the other hand, if you try to assign values freely along the boundary $x = -x_0 < 0$, $t > 0$ together with initial data along the x-axis for $x \leq -x_0$ in either Figure 5 or 6, the problem is ill-posed because the characteristics that pass through $x < -x_0$ $t = 0$ also hit the boundary $x = -x_0$, $t > 0$, making it impossible for u to be constant along characteristics unless the boundary data is determined by the initial data. Without the knowledge of the characteristics, we can't tell ahead of time which boundary value problems make sense, and which do not make sense!

Conclude: The characteristics tell us which boundary value problems make sense, and which do not make sense.

The purpose of the next section is to make precise the fundamental existence and uniqueness theorem for the initial value problem (17) of ODE's, (the main theorem we applied above to prove that shock-waves can't form in linear PDE's). It turns out that having a continuous vector field is not enough, but we require more smoothness than this, namely, *Lipschitz Continuity* is required. As a result, we have shown that the characteristics curves (13), the curves along which a solution $u(x,t)$ of the *linear* equation $b(x,t)u_t + a(x,t)u_x = 0$ is constant, can never cross in the (x,t)-plane, so long as $a(x,t)$ and $b(x,t)$ are Lipschitz continuous functions of (x,t).

Summary: PDE's can be reduced to ODE's via their characteristics, and shock-waves are a purely *nonlinear* phenomenon!