Mathematical Modeling
Class Notes

Leon Hostetler

June 25, 2019
Version 0.5

Contents

Preface iv

1 Introduction 1
1.1 The Grandfather Clock . . . 1

2 Proportions 4

3 Dimensional Analysis and Scaling 6
3.1 Checking Equation Consistency . . . 6
3.2 Natural Scales . . . 7
3.3 Nondimensionalization and Scale Models . . . 8
3.4 Buckingham π Theorem . . . 13

4 ODEs, Eigenvalues, Timescales, and Phase 18
4.1 Introduction . . . 18
4.2 Theoretical Background . . . 18
4.3 Pendulum Example . . . 20
4.4 Phase Diagrams . . . 22
4.5 Equilibrium and Stability . . . 28
4.6 Linearization . . . 29
4.7 Lanchester Combat Model . . . 33
4.8 Predator-prey Model . . . 36

5 Forced Behavior 40
5.1 Damped, Forced, Pendulum . . . 43

6 Spring Models 46
6.1 Spring Scale . . . 46
6.2 Double-ended Spring . . . 48
6.3 Spring Suspensions . . . 49

7 Heat Transfer 52
7.1 Double-pane Windows . . . 53
7.2 Insulated Wall . . . 55
7.3 Wine Cellar . . . 59

8 Electronics 61
8.1 Low-pass Filter . . . 61
8.2 High-pass Filter . . . 63
8.3 Bandpass Filter . . . 64
8.4 Electric Clock . . . 66

9 Growth and Decay 69
9.1 Filtration . . . 69

9.2 Logistic Growth 70
9.3 Competing Species 71
9.4 The SIR Epidemic Model 72

10 Traffic Flow **75**
10.1 Traffic Speed vs. Congestion 75
10.2 Traffic Flow on a Bridge 77
10.3 Whitham Traffic Flow Model 78
10.4 Solution to the Traffic Flow Equation 79
10.5 Example 79
10.6 Example 81
10.7 Shock waves and Expansion Fans 82
10.8 Example 84

Index **86**

Preface

About These Notes

My class notes can be found at www.leonhostetler.com/classnotes

Please bear in mind that these notes will contain errors. Any errors are certainly my own. If you find one, please email me at leonhostetler@gmail.com with the name of the class notes, the page on which the error is found, and the nature of the error.

Updates

Last Updated: June 25, 2019
Version 0.5: (Jun. 25, 2019) Updated layout.
Version 0.0: (May 16, 2017) First upload.

Chapter 1

Introduction

Hooke's law and Newton's law are example of mathematical models.

Mathematical modeling is just the process of representing the behavior of a system in terms of mathematical relationships in order to understand the system's behavior and to answer questions about its behavior.

In general, with mathematical modeling, we follow these steps:

1. Pose the problem/question. By posing the question well, we limit the scope of the model and make its development easier.
2. Create a simple conceptual model, reducing the system to its essential components and characteristics. For example, the complex apparatus of e pendulum clock can be conceptually represented as a swinging rod with a mass at the end.
3. Make simplifying assumptions. What factors can be ignored without significantly changing the accuracy of the model? Is a model which ignores friction a good enough model to answer the question with satisfactory precision?
4. Identify the quantities involved. Consider the question that you posed in the beginning. What are the parameters? What are the variables? What are the dependent and independent variables?
5. Derive a mathematical model by identifying the mathematical relationships between the quantities involved.
6. Simplify the mathematical model if necessary to approximate a solution.
7. Study the behavior of the solution of the equation.
8. Make design decisions or predictions about the system.

In general, we want the *simplest* model that answers the question posed. Use the minimal information needed to answer the question.

Tricks of the trade include

- Neglect small effects, e.g. friction, air resistance, variability of g, flexion of nearly rigid objects, etc.
- Assume the system is not affected by the environment, e.g. Coriolis effect
- Replace distributed quantities by lumped quantities. For example, if a distributed mass can be approximated as a point mass at its center of mass, do so.
- Assume linearity or linearize
- Assume parameters are constant when possible.

1.1 The Grandfather Clock

Question: What must be the length of the pendulum in a grandfather clock.

The key to understanding the question, is that the period of a grandfather clock depends on the length of the pendulum. In order to keep accurate time, the period of the pendulum much be 2 seconds, since that the clock ticks two seconds per swing of the pendulum.

We've posed the question, and now we need to create a simplified conceptual model. The essential component of the grandfather clock, as it relates to answering our question, is the pendulum itself. We can ignore the clock's case, and all the intricate machinery that converts the swings of the pendulum into the movement of the clock's minute and hour hands. We can even ignore the weights whose weight serves to drive the pendulum as well as the cables connecting the weights to the rest of the clock. So our simplified conceptual model of the grandfather clock is simply its pendulum with a pivot at one end and a mass at the other end.

Next, we make simplifying assumptions about our simplified conceptual model:

1. No friction at the pivot, and no air resistance. The time scale on which friction and air resistance operates is much larger than the time scale of a single swing of the pendulum. Therefore, we won't lose significant precision in our solution if we ignore these.
2. Rod is rigid. In reality, all objects are slightly flexible, but we can safely neglect this characteristic.
3. The rod is massless. Since the vast majority of the mass of the pendulum is concentrated at its end, the mass of the rod does not significantly alter the pendulum's moment of inertia, so we can neglect the rod's mass.

The next step is to identify the quantities involved. Since, we want to know the length of the pendulum, we want to treat it as the dependent variable. Other relevant variables and parameters include time, the mass at the end of the pendulum, and gravitational acceleration, which is the pendulum's driving force.

The next step is to identify the relationships. To do this, it is helpful to draw a diagram and label everything. By using the angle θ between the pendulum and the vertical, using a polar coordinate system centered on the mass at the end of the pendulum, and Newton's law, we find the following mathematical model for the pendulum

$$\ddot{\theta} = -\frac{g}{l}\sin\theta.$$

Our next step is to simplify as necessary until we can solve the problem. Because of the $\sin\theta$, the equation for the pendulum is a nonlinear ODE, which we are not able to solve. However, if we can linearize it (i.e. convert it into a linear ODE which we can solve) without losing too much precision, we are good. We can do this by assuming that $\theta << 1$ when measured in radians. This is because the Taylor expansion of sine is

$$\sin\theta = \theta - \frac{1}{3}\theta^3 + \cdots ,$$

and if $\theta << 1$ then $\theta^2/3$ and all higher order terms are extremely small. So we can discard the higher order terms and approximate $\sin\theta$ as

$$\sin\theta \approx \theta.$$

Considering the grandfather clock, it seems that we are justified in assuming $\theta << 1$. We guess that the length of the pendulum is approximately three feet and it swings approximately half a foot either side of straight down. A rough calculation gives us $\tan\theta_{max} \approx 1/6$ or $\theta_{max} \approx 0.165$. Calculating $\theta_{max}^3/3 \approx 0.00150$. The relative error when discarding the cubic term versus keeping it is a little less than 1% so our approximation is justified. Our simplified mathematical model is then

$$\ddot{\theta} = -\frac{g}{l}\theta.$$

The final step is to solve the problem. The linear ODE given above has the solution

$$\theta(t) = \theta_{max}\cos\left(\sqrt{\frac{g}{l}}t\right).$$

To solve for l, we need to specify t, and we know a convenient value. The period is $T = 2$s and the period of cosine is 2π. So

$$\sqrt{\frac{g}{l}} 2\,\mathrm{s} = 2\pi$$
$$l = \frac{g\ \mathrm{s}^2}{\pi^2} = 0.994\,\mathrm{m}.$$

So a grandfather clock's pendulum should be 0.994 meters long in order to keep time accurately.

Chapter 2

Proportions

Knowing that one quantity y is proportional to another quantity x

$$y \propto x,$$

allows us to use a constant of proportionality k to set up the equation

$$y = kx.$$

If one quantity is proportional to another, it tells us that the relationship is linear. In the example above, if we were to graph y versus x, we would get a straight line.

Example 2.0.1

A marathon runner is moving at a constant speed and the distance that she has traveled is proportional to the time that she has been running. If the runner has traveled 10 miles in 1 hour, what is the constant of proportionality?

Since $d \propto t$, we can say that $d = vt$ where v is the constant of proportionality. Plugging in our values and solving for v gives us $v = 10\,\text{mph}$. Notice that in this case, the constant of proportionality is her speed, that is, the rate at which she covers distance with respect to time.

Example 2.0.2

The gravitational force between two objects is inversely proportional to the square of the distance between them. If the gravitational force between them is 10 N when they are 1 m apart, what is the gravitational force between them when they are 2 m apart?

If we let F_g be the gravitational force between the two objects and r be the distance between them, then from what we are told,

$$F_g \propto \frac{1}{r^2},$$

which can also be written using a constant of proportionality k as

$$F_g = k\frac{1}{r^2}.$$

To answer the question, we start by solving for the constant of proportionality using the given information.

$$10\,\text{N} = k\frac{1}{(1\,\text{m})^2},$$

which gives us $k = 10\,\text{N}\cdot\text{m}^2$. Using this, we can find the gravitational force when they are 2m apart

$$F_g = 10\,\text{N}\cdot\text{m}^2\frac{1}{(2\,\text{m})^2} = \frac{5}{2}\text{N}.$$

If x goes up and w stays constant, what happens to y in the following?

$$y = xw$$
$$y = \frac{x}{w}.$$

Which of the following is closest to $1/10^2$?

$$\frac{1}{(10-1)^2} \quad \text{or} \quad \frac{1}{(10+1)^2}$$

Example 2.0.3

If $a \propto 1/b^2$, is it true that $\sqrt{a} \propto 1/b$?

If $a \propto 1/b^2$ then we know that $a = c^2/b^2$ for some constant c^2. Taking the square root of both sides gives us $\sqrt{a} = c/b$. Since c^2 is constant, c is constant, so $\sqrt{a} \propto 1/b$.

If $A \propto B$ *and* $A \propto C$, then $A \propto BC$.

To prove this, we start by noting that

$$A = \beta B = \gamma C,$$

where β and γ are the proportionality constants. But $A = f(B, C)$ is a function of B and C, so the constant of proportionality β that relates A and B is a function only of C. Likewise, γ is a function only of B, so

$$A = \beta(C)B = \gamma(B)C.$$

Rearranging the second equation gives us

$$\frac{B}{\gamma(B)} = \frac{C}{\beta(C)} = \lambda.$$

Since the left side is only a function of B, and the right side is only a function of C, they must both be equal to the same constant λ. Solving for $\gamma(B)$ gives us

$$\gamma(B) = \frac{B}{\lambda}.$$

Plugging this into $A = \gamma(B)C$ gives us

$$A = \frac{1}{\lambda}BC,$$

and since λ is just a constant,

$$A \propto BC.$$

Chapter 3

Dimensional Analysis and Scaling

All measurable quantities have units or dimensions. The fundamental units are

- length, denoted $\mathscr{L}$ (e.g. meters, feet, parsecs)
- time, denoted $\mathscr{T}$ (e.g. seconds, years)
- mass, denoted $\mathscr{M}$ (e.g. grams, kilograms)

From these fundamental units, we get derived quantities such as

- area, with units $\mathscr{L}^2$
- volume, with units $\mathscr{L}^3$
- velocity, with units $\frac{\mathscr{L}}{\mathscr{T}}$
- acceleration, with units $\frac{\mathscr{L}}{\mathscr{T}^2}$

and compound quantities such as

- force = mass × acceleration, with units $\frac{\mathscr{M}\mathscr{L}}{\mathscr{T}^2}$
- flow rate = volume / time, with units $\frac{\mathscr{L}^3}{\mathscr{T}}$
- weight = mass × gravitational acceleration, with units $\frac{\mathscr{M}\mathscr{L}}{\mathscr{T}^2}$
- mass flux = mass / (area × time), with units $\frac{\mathscr{M}}{\mathscr{L}^2\mathscr{T}}$

Keep in mind that dimensionless units are also frequently used in science. For example, an angle given in radians is a dimensionless number. It is defined as the arc length with units of length divided by the radial distance, which also has units of length. The units cancel, and we are left with a dimensionless number.

There are several things we use dimensional analysis for, and they are detailed in the sections that follow.

3.1 Checking Equation Consistency

All equations must be dimensionally consistent. That is, the units of one side of an equation must agree with the units on the other side.

For example, the approximate solution for a pendulum is

$$\frac{d^2\theta}{dt^2} = -\frac{g}{l}\theta,$$

where θ is the angle between the pendulum and the vertical, and l is the length of the pendulum. An analysis of the dimensions of the equation gives us

$$\frac{1}{\mathscr{T}^2} = \frac{\frac{\mathscr{L}}{\mathscr{T}^2}}{\mathscr{L}}.$$

Notice that the angle θ, measured in radians, is dimensionless. Simplifying the right side gives us

$$\frac{1}{\mathscr{T}^2} = \frac{1}{\mathscr{T}^2},$$

which shows that our equation makes sense from a dimensional standpoint. Both sides have units of one over time squared.

Checking dimensional consistency is the zeroth order test of your solution.

3.2 Natural Scales

Dimensional analysis can help us determine the natural scales of a problem.

For example, the period of a pendulum, which is independent of its mass, is a natural time scale to use when working with pendulum problems. It's faster and less messy when answering questions about a pendulum to use the time scale of periods rather than seconds.

The parameters of a mathematical model determine the natural scales. For example, the relevant parameters of a pendulum problem are length with units $\mathscr{L}$ and the gravitational acceleration with units $\mathscr{L}/\mathscr{T}^2$. The variables are time and the angle the pendulum makes with the vertical. To find the natural time scale for a pendulum problem, we note that there is only one way to combine the parameters length and gravitational acceleration to get something with units of time, and that is as

$$\sqrt{\frac{\mathscr{L}}{\frac{\mathscr{L}}{\mathscr{T}^2}}} = \mathscr{T}.$$

A natural time scale, then is a time proportional to the square root of the length divided by the gravitational acceleration. In this case, a convenient proportionality constant is 2π since that gives us the period of the pendulum

$$T = 2\pi\sqrt{\frac{l}{g}}.$$

Example 3.2.1

The heat equation is the partial differential equation

$$\frac{\partial u}{\partial t} = k\frac{\partial^2 u}{\partial x^2},$$

which describes the movement of heat. The function u gives the temperature, t is time, x is position, and k is the thermal diffusivity of the material. What are the units of k?

Since we know that the equation must be dimensionally consistent, we identify the units on both sides as

$$\frac{\text{DEG}}{\mathscr{T}} = k\frac{\text{DEG}}{\mathscr{L}}.$$

Solving for k tells us that k has the units

$$\frac{\mathscr{L}^2}{\mathscr{T}}.$$

For example, suppose a friend is using a metal skewer to roast a hot dog over a fire. The heat from the fire travels from the end of his skewer to his hand in the time that it takes his hot dog to be half roasted. How much longer of a skewer of the same metal will you need to fully roast your own hot dog before your hand is burned?

Earlier, we found that

$$k = \frac{\mathscr{L}^2}{\mathscr{T}},$$

that is, the proportionality constant from the heat equation has units of length-squared over time. Solving for time, we find that the time t for the heat from the fire to travel the length L of the skewer is proportional to the length-squared of the skewer divided by k

$$t = c\frac{L^2}{k}.$$

Note, c is just a dimensionless proportionality constant. You have a skewer of length L_2 such that the time for the heat to travel its length is $2t$, so

$$2t = c\frac{L_2^2}{k}.$$

Dividing these two equations by each other, allows us to get rid of the unknown constant of proportionality as well as the unknown constant k and solve for L_2 to get

$$L_2 = \sqrt{2}L.$$

So your skewer needs to be $\sqrt{2} \approx 1.41$ times the length of your friend's skewer. Notice that we were able to solve a heat problem without having to solve the partial differential heat equation.

3.3 Nondimensionalization and Scale Models

Nondimensionalization is the removal of units/dimensions from equations involving physical quantities by making a suitable substitution of variables. [1]We scale a dimensionful equation with respect to some reference values when converting it to a dimensionless equation.

Example 1

Recall the pendulum equation

$$\frac{d^2\theta}{dt^{*2}} = -\frac{g^*}{l^*}\theta,$$

where '*' indicates that the quantity is dimensionful (i.e. has units). To nondimensionalize this equation, we start by noting its dimensions. We know that t^* has units $\mathscr{T}$, l^* has units $\mathscr{L}$, and g^* has units $\mathscr{L}/\mathscr{T}^2$. Next, we define reference values T^* and L^* with which we can construct dimensionless versions of the above dimensionful parameters. At this point, it doesn't particularly matter exactly what the reference values are. Next, we define dimensionless parameters by dividing the dimensionful parameters by our reference values

$$t = \frac{t^*}{T^*}, \qquad l = \frac{l^*}{L^*}, \qquad g = \frac{g^*}{\frac{L^*}{T^{*2}}}.$$

Notice that t, l, and g are now dimensionless quantities. Next, we substitute these values into our original equation and simplify. To substitute for the second derivative, we need

[1] https://en.wikipedia.org/wiki/Nondimensionalization

to use the chain rule to write $\frac{d}{dt^*} = \frac{d}{dt}\frac{dt}{dt^*}$.

$$\begin{aligned}
\frac{d^2\theta}{dt^{*2}} &= -\frac{g^*}{l^*}\theta \\
\frac{d}{dt^*}\frac{d\theta}{dt^*} &= -\frac{g\frac{L^*}{T^{*2}}}{lL^*}\theta \\
\frac{d}{dt}\frac{dt}{dt^*}\frac{d\theta}{dt}\frac{dt}{dt^*} &= -\frac{g}{lT^{*2}}\theta \\
\frac{1}{T^{*2}}\frac{d}{dt}\frac{d\theta}{dt} &= -\frac{g}{lT^{*2}}\theta \\
\frac{d^2\theta}{dt^2} &= -\frac{g}{l}\theta.
\end{aligned}$$

Now we have a dimensionless equation of the same form as our original dimensionful equation.

Next, we want to define our reference time using a natural time scale for the system. In this case, the most natural time scale is the period of the pendulum,

$$T^* = \sqrt{\frac{l^*}{g^*}}.$$

Going back to the original pendulum equation

$$\frac{d^2\theta}{dt^{*2}} = -\frac{g^*}{l^*}\theta,$$

we can rewrite the left side as

$$\frac{1}{T^{*2}}\frac{d^2\theta}{dt^2} = -\frac{g^*}{l^*}\theta.$$

Substituting for T^*, we get

$$\frac{1}{\left(\sqrt{\frac{l^*}{g^*}}\right)^2}\frac{d^2\theta}{dt^2} = -\frac{g^*}{l^*}\theta,$$

which simplifies to

$$\frac{d^2\theta}{dt^2} = -\theta.$$

This shows us that if we scale an equation with respect to the natural scale, the dimensionful parameters disappear.

The solution of this differential equation is

$$\theta(t) = \theta_0 \cos t.$$

Notice that this has a period of $P = 2\pi$. To convert this value to a meaningful measure, with units of time, we just multiply it by our reference time

$$P^* = PT^* = 2\pi\sqrt{\frac{l^*}{g^*}}.$$

Example 2

Nondimensionalize the PDE

$$\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = 0.$$

When we're ready to nondimensionalize an equation, we need to note which variables have dimensions. We denote those variables with asterisks. In this case, all of them have dimensions, so

$$\frac{\partial u^*}{\partial t^*} + a^* \frac{\partial u^*}{\partial x^*} = 0.$$

The first step is to identify the independent variable(s), the dependent variable(s), the parameters, and all of their dimensions. In this case, our independent variable, the function which gives us the solution to the PDE, is u^*. To denote the units of a variable, we might put brackets around the variable, so $[u^*]$ indicates the units of u^*. Our independent variables are x^* which typically has units of length, denoted $[x^*] = \mathscr{L}$, and t^* with units $[t^*] = \mathscr{T}$. Our lone parameter in this PDE is a^*. We can deduce its units by the fact that the equation must be dimensionally consistent, and so each term must have the same units. The units are

$$\frac{[u^*]}{\mathscr{T}} + \frac{[a^*][u^*]}{\mathscr{L}} = 0.$$

We can cancel $[u^*]$ from both terms, and it's obvious that $[a^*] = \mathscr{L}/\mathscr{T}$.

Since a^* is a length divided by time the parameter must be a speed. This gives us a timescale for the PDE. Since we have only a single parameter with a time in it, this PDE has a single timescale.

The next step is to select scaling factors for the independent and dependent variables. We don't worry about scaling the parameters as that happens naturally as a result of scaling the variables. When scaling the variables, we generally prefer natural scales. In this case, we choose our scaling constants as L^* and U^*. A natural scale for a PDE is the length of the domain it is defined over. In this case, the PDE is defined on the domain $[0, L^*]$. A natural scale factor for u^* might be the average value of the initial conditions of u, which are typically given in the form $u(x, 0) = f(x)$. Since the parameter gives us a time, we define the scaling factor for time as L^*/a^*. Then,

$$\begin{aligned} u &= \frac{u^*}{U^*} \Leftrightarrow u^* = U^* u \\ x &= \frac{x^*}{L^*} \Leftrightarrow x^* = L^* x \\ t &= \frac{t^*}{\frac{L^*}{a^*}} \Leftrightarrow t^* = \frac{tL^*}{a^*}. \end{aligned}$$

Next, we make the change of variables and simplify.

$$\begin{aligned} \frac{\partial u^*}{\partial t^*} + a^* \frac{\partial u^*}{\partial x^*} &= 0 \\ \frac{\partial (U^* u)}{\partial \left(\frac{tL^*}{a^*}\right)} + a^* \frac{\partial (U^* u)}{\partial (L^* x)} &= 0 \\ \frac{U^* a^*}{L^*} \frac{\partial u}{\partial t} + \frac{U^* a^*}{L^*} \frac{\partial u}{\partial x} &= 0 \\ \frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} &= 0. \end{aligned}$$

We have now scaled the PDE, so that it is nondimensional. This PDE is now measuring in units of whatever a^* was.

Example 3

Nondimensionalize the **advection-diffusion equation**

$$\frac{\partial u^*}{\partial t^*} + a^* \frac{\partial u^*}{\partial x^*} = \nu^* \frac{\partial^2 u^*}{\partial {x^*}^2}.$$

We immediately notice that there are two parameters, a^* and ν^*, and to ensure that the equation is dimensionally consistent, their units must be

$$[a^*] = \frac{\mathscr{L}}{\mathscr{T}}, \qquad [\nu^*] = \frac{\mathscr{L}^2}{\mathscr{T}}.$$

This demonstrates that the advection-diffusion equation has two different time scales associated with it. One is a speed, since it is length divided by time, and the other is unclear what its physical meaning is. In general, the number of parameters in an ODE or PDE tells us how many different things are going on at the same time.

We can nondimensionalize the PDE with respect to either timescale. Typically, we nondimensionalize with respect to the timescale that dominates the behavior of whatever is being modeled by the PDE. In this case, we might not know, but we can nondimensionalize with respect to one timescale and then the other and then make observations about the behavior of the system by comparing the two.

If we nondimensionalize with respect to the speed, that is, the timescale associated with a^*, then

$$\begin{aligned} x &= \frac{x^*}{L^*} \Leftrightarrow x^* = L^* x \\ u &= \frac{u^*}{U^*} \Leftrightarrow u^* = U^* u \\ t &= \frac{t^*}{L^*/a^*} \Leftrightarrow t^* = \frac{L^*}{a^*} t. \end{aligned}$$

Making the change of variables and simplifying, we get

$$\begin{aligned} \frac{\partial u^*}{\partial t^*} + a^* \frac{\partial u^*}{\partial x^*} &= \nu^* \frac{\partial^2 u^*}{\partial x^{*2}} \\ \frac{\partial (U^* u)}{\partial \left(\frac{L^*}{a^*} t\right)} + a^* \frac{\partial (U^* u)}{\partial L^* x} &= \nu^* \frac{\partial^2 (U^* u)}{\partial L^* x^2} \\ \frac{a^*}{L^*} \frac{\partial u}{\partial t} + \frac{a^*}{L^*} \frac{\partial u}{\partial x} &= \frac{\nu^*}{L^{*2}} \frac{\partial^2 u}{\partial x^2} \\ \frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} &= \frac{\nu^*}{a^* L^*} \frac{\partial^2 u}{\partial x^2}. \end{aligned}$$

The PDE is now nondimensional, and the dimensionless constant $\nu^*/(a^* L^*)$ is the reciprocal of the **Peclet number** denoted Pe, so our nondimensional PDE can be written as

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = \frac{1}{\mathrm{Pe}} \frac{\partial^2 u}{\partial x^2}.$$

Notice that if $\mathrm{Pe} \to \infty$ then the PDE approximates

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0.$$

This is the one-dimensional **advection equation**. Consider a bottle of ink dumped into a river. If the river is flowing fast, then the most significant effect is that of the ink blot traveling downstream without spreading out much. This is **advection**. The advection timescale one can think of as the timescale over which the ink blot noticeably advects (i.e. moves) downstream. If the advection timescale is the most significant one, that is, if in some time period, the ink blot mostly advects, then our choice of timescale with which to nondimensionalize this problem was the right one.

The other thing going on at the same time, in addition to advection, is **diffusion**. This is the action in which the ink blot spreads out over time, growing larger. If the river is barely flowing at all, then in some period of time, the ink blot will most noticeably have diffused. If that is the case, then the most convenient nondimensionalization would have been with respect to the other variable. Doing so, we have that

$$t = \frac{t^*}{L^{*2}/\nu^*} \Leftrightarrow t^* = \frac{L^{*2}}{\nu^*}t.$$

The other substitutions remain the same. Then doing the change of variables and simplifying, we get the nondimensional PDE

$$\frac{\partial u}{\partial t} + \mathrm{Pe}\frac{\partial u}{\partial x} = \frac{\partial^2 u}{\partial x^2}.$$

Notice that if $\mathrm{Pe} \to 0$, then the PDE approximates

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}.$$

This is the one-dimensional **diffusion equation**.

To summarize, if the Peclet number is large, the ink blot most noticeably advects, and if it's small, the ink blot most noticeably diffuses. Returning to the Peclet number, we see that

$$\mathrm{Pe} = \frac{a^* L^*}{\nu^*} = \frac{a^*/L^*}{\nu^*/L^{*2}} = \frac{\tau_d}{\tau_a},$$

where τ_d is the diffusion timescale, and τ_a is the advection timescale. So the Peclet number, the dimensionless constant associated with our PDE, is the ratio of the PDE's two different timescales. We learned all of this without actually solving the PDE.

We can reach these conclusions in a more rigorous manner by returning to the original PDE. If $\nu = 0$, then the PDE becomes

$$\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = 0.$$

This PDE has a solution of the form

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

where f is any linear function. This can be verified by taking the partial derivatives and plugging it into the PDE. We want to know what this solution does as time advances. If we let $z = x - at$ then $u(x,t) = f(z)$. What this solution does as time advances, is that it retains the original shape of the $f(z)$, but moves to the right as time goes forward. In other words, at $t = 0$, the function $f(z)$ has some shape and is at some location. At $t = 1$, the function has the same shape but has been translated to the right. If $a = 1$, for example, the shape is translated one unit to the right for every unit of time that passes. So, as time passes, the solution to this PDE is a traveling wave moving to the right with speed a^*. So this PDE describes **advection**.

If instead of $\nu = 0$, we let $a = 0$, then the PDE becomes

$$\frac{\partial u}{\partial t} = \nu\frac{\partial^2 u}{\partial x^2}.$$

This PDE has solutions of the form

$$u(x,t) = \frac{1}{2t+1}e^{-\frac{x^2}{\nu(t+1)}}.$$

We want to know what happens as t changes. This is a Gaussian function and its graph is a bell curve. Notice that the first part, the fraction at the front, gives the amplitude. As time passes and t increases, the amplitude decreases since t is in the denominator. So as time passes, the bell curve is flattening. If we look at the exponential part of the solution, we see that the rate at which the bell curve is flattening is determined by ν. So ν gives the spreading rate. A wave that is flattening and spreading out like this, is **diffusion**.

Our original PDE contains both of these behaviors. That is, it describes a wave that is moving to the right (advection) with speed a^* while flattening and spreading out (diffusion) at a rate that depends on v^*/L^*.

3.4 Buckingham π Theorem

The laws of nature cannot depend on the units we use to measure things. Physical systems cannot depend on how we measure things because they don't know if we're using miles or meters, for example. As a consequence, any physical system must be expressible in terms of nondimensional parameters of the system.

This leads to the **Buckingham π theorem**: Given n variables or parameters and k fundamental units, there are $p = n - k$ dimensionless parameters $\pi_1, \pi_2, \dots, \pi_p$. The physical laws describing the system must then be expressible as a function of these dimensionless parameters $f(\pi_1, \pi_2, \dots, \pi_p) = 0$, or equivalently as $\pi_1 = g(\pi_2, \dots, \pi_p)$.

The significance of the Buckingham π theorem is that it allows us to determine how one variable is related to the others.

Note, π_i in this case are just the names of the parameters. They have nothing to do with the constant π.

Pendulum Example

How does the period T of a pendulum depend on the mass m, the length l, and the gravitational force g?

First, we figure out what the units are of all the variables:

$$[T] = \mathscr{T}, \qquad [m] = \mathscr{M}, \qquad [l] = \mathscr{L}, \qquad [g] = \frac{\mathscr{L}}{\mathscr{T}^2}.$$

So we have $n = 4$ variables and $k = 3$ fundamental units. So with this set of variables, there is $p = 1$ nondimensional parameter π we can construct as a combination of all the variables. To find that combination, we raise each variable to a power

$$\pi = T^a m^b l^c g^d.$$

Then we look at its units

$$[\pi] = \mathscr{T}^a \mathscr{M}^b \mathscr{L}^c \left(\frac{\mathscr{L}}{\mathscr{T}^2}\right)^d = \mathscr{T}^{a-2d} \mathscr{M}^b \mathscr{L}^{c+d}.$$

Since π is defined to be a nondimensional parameter, we know that all the exponents must be zero, that is

$$\begin{aligned} a - 2d &= 0 \\ b &= 0 \\ c + d &= 0. \end{aligned}$$

Now we pick one parameter arbitrarily as a free parameter. Choose $a = 1$ for convenience since it is the power of the variable that we are trying to find. Then solving for the others, we find that

$$a = 1, \qquad b = 0, \qquad c = -\frac{1}{2}, \qquad d = \frac{1}{2}.$$

So

$$\pi = T^1 m^0 l^{-\frac{1}{2}} g^{\frac{1}{2}} = T\sqrt{\frac{g}{l}}.$$

Checking the units of π, the parameter that describes the relationship between the four variables, shows us that it is indeed dimensionless.

This tells us that the equation of motion is of the form

$$f\left(T\sqrt{\frac{g}{l}}\right) = 0.$$

This is a root-finding problem. We know the function f is zero at the value $C = T\sqrt{g/l}$. Solving for T gives us

$$T = C\sqrt{\frac{g}{l}}.$$

The value of the constant C, which is actually 2π, cannot be found by dimensional analysis. It must be found by experimentation or by solving the actual ODE for the pendulum. However, the important thing is that we now know the mathematical relationship between the period of the pendulum and its length and the gravitational constant.

Projectile Example

If a projectile is launched straight up with initial speed v, ignoring air resistance, how does the maximum height of the projectile depend on the initial velocity.

We start by determining the units of the relevant variables height h and initial speed v.

$$[h] = \mathscr{L}, \qquad [v] = \frac{\mathscr{L}}{\mathscr{T}}.$$

We now have $n = 2$ variables and $k = 2$ fundamental units. Since $p = 0$, we do not have enough variables yet to use the Buckingham π theorem. We could add the mass of the projectile,

$$[m] = \mathscr{M},$$

however, we don't have $\mathscr{M}$ in another variable, so nothing will cancel powers of $\mathscr{M}$. Furthermore, since the unit of mass does not appear in the units of length, we can just ignore it from the outset. The key is to add the gravitational acceleration of the projectile

$$[g] = \frac{\mathscr{L}}{\mathscr{T}^2},$$

since it adds a third variable but no additional fundamental units. Now we have $n = 3$ variables and $k = 2$ fundamental units, so there is $p = n - k = 1$ nondimensional parameter that describes the system.

To find the combination of variables which gives that nondimensional parameter, we let

$$\pi = h^a v^b g^c,$$

with units

$$[\pi] = \mathscr{L}^a \left(\frac{\mathscr{L}}{\mathscr{T}}\right)^b \left(\frac{\mathscr{L}}{\mathscr{T}^2}\right)^c = \mathscr{L}^{a+b+c}\mathscr{T}^{-b-2c}.$$

So $a + b + c = 0$ and $-b - 2c = 0$. If we let $a = 1$ be the free parameter, then solving for the others gives us $b = -2$ and $c = 1$, so

$$\pi = \frac{hg}{v^2}.$$

This tells us that

$$f\left(\frac{hg}{v^2}\right) = 0,$$

which implies that

$$C = \frac{hg}{v^2}.$$

Solving for h tells us how the height of the projectile depends on the initial speed

$$h = C\frac{v^2}{g}.$$

Drag Example

How does the drag (i.e. air resistance) on a car depend on the car's size and speed?

The drag, D, being essentially a friction force, we know has the units

$$[D] = \mathscr{M}\frac{\mathscr{L}}{\mathscr{T}^2}.$$

What does it depend on? We know it depends on the size of the car and its speed. We can approximate the car as a sphere which has a single size parameter—its radius r. So we have that

$$[r] = \mathscr{L}, \qquad [v] = \frac{\mathscr{L}}{\mathscr{T}}.$$

So far we have three variables and three fundamental units. We need more independent variables in order to use the Buckingham π theorem. The drag force also depends on the viscosity of the air, μ, and the density of the air ρ. These have units

$$[\mu] = \frac{\mathscr{M}}{\mathscr{L}\mathscr{T}}, \qquad [\rho] = \frac{\mathscr{M}}{\mathscr{L}^3}.$$

Now we have $n = 5$ variables and $k = 3$ fundamental units, so these variables are related via $p = n - k = 2$ nondimensional parameters, π_1, and π_2.

We want to construct π_1 and π_2 by choosing combinations of the five variables. Both π_1 and π_2 need to have all the fundamental units. However, they can't both include D and μ since those have the same fundamental units, so the dimensions won't cancel. So we choose

$$\pi_1 = D^a v^b \rho^c r^d, \qquad \pi_2 = \rho^e v^f \mu^g r^h.$$

Looking at the units, we have

$$[\pi_1] = \left(\frac{\mathscr{M}\mathscr{L}}{\mathscr{T}^2}\right)^a \left(\frac{\mathscr{L}}{\mathscr{T}}\right)^b \left(\frac{\mathscr{M}}{\mathscr{L}^3}\right)^c \mathscr{L}^d = \mathscr{M}^{a+c}\mathscr{L}^{a+b-3c+d}\mathscr{T}^{-2a-b}$$

$$[\pi_2] = \left(\frac{\mathscr{M}}{\mathscr{L}^3}\right)^e \left(\frac{\mathscr{L}}{\mathscr{T}}\right)^f \left(\frac{\mathscr{M}}{\mathscr{L}\mathscr{T}}\right)^g \mathscr{L}^h = \mathscr{M}^{e+g}\mathscr{L}^{-3e+f-g+h}\mathscr{T}^{-f-g}.$$

Setting the exponents equal to zero and choosing $a = 1$ as the free parameter for π_1 and $e = 1$ as the free parameter for π_2 then solving for the others, we get

$$a = 1, \quad b = -2, \quad c = -1, \quad d = -2, \quad e = 1, \quad f = 1, \quad g = -1, \quad h = 1.$$

So our nondimensional parameters are

$$\pi_1 = \frac{D}{v^2\rho r^2}, \qquad \pi_2 = \frac{\rho v r}{\mu}.$$

Then by the Buckingham π theorem, we have that

$$f\left(\frac{D}{v^2\rho r^2}, \frac{\rho v r}{\mu}\right) = 0.$$

This can also be written as

$$\frac{D}{v^2\rho r^2} = g\left(\frac{\rho v r}{\mu}\right).$$

That is, the nondimensional parameter $D/(v^2\rho r^2)$ is some function g of the other nondimensional parameter $\rho v r/\mu$. The second nondimensional parameter has a name. It is the **Reynold's number** of the fluid through which the object is moving.

$$\mathrm{Re} = \frac{\rho v r}{\mu},$$

so we can write the relationship of the drag force D to the other variables as

$$D = \rho v^2 r^2 \cdot g(\mathrm{Re}).$$

We don't know what the function g is, so it must be found by experimentation or by solving the Navier-Stokes equation.

The experimental results tell us that $g \sim 1/\mathrm{Re}$ when Re is small and g approaches a constant value, when Re is large. This means there are two regimes to consider. When Re is small,

$$D = \rho v^2 r^2 \cdot C\frac{1}{\mathrm{Re}} = \rho v^2 r^2 \cdot C\frac{\mu}{\rho v r} = C\mu v r,$$

and when Re is large,

$$D = C\rho v^2 r^2.$$

When Re is small, which occurs when the fluid has a higher viscosity (e.g. water or molasses), the drag is called **Stoke's drag**. When Re is large which occurs when the fluid is air and we're dealing with meter-sized objects, the drag is called **form drag**.

So to summarize, we have two different models for how the drag force depends on the other variables. The one we should use for a given situation depends on the Reynold's number.

If the appropriate model for drag depends on the size of the object, how is it that scale models of airplanes tested in wind tunnels accurately depict the behavior of large airplanes? To ensure that a scale model accurately depicts the behavior of a real airplane, the parameters must be adjusted for the scale model in the wind tunnel so that the scale model in the wind tunnel and the real airplane in air have the same Reynold's number. To do that, you can increase the air speed in the wind tunnel or use a gas with lower viscosity.

A similar problem occurs with scale models used in older films. A flooding scene looks fake when filmed with water and scale buildings because the Reynold's number is not the same. To create a realistic flood using scaled-down buildings, the filmmakers need to use a different fluid.

Bomb Blast Example

What is the energy of an atomic bomb blast given that you have some snapshots of the blast depicting the radius of the compression wave at different times?

Energy has the units

$$[E] = \frac{\mathscr{M}\mathscr{L}^2}{\mathscr{T}^2}.$$

In addition to the radius R of the compression wave and the time t, we also need the density ρ of the air, since the speed at which the blast wave propagates obviously depends on the density of the fluid it is propagating through.

$$[R] = \mathscr{L}, \qquad [t] = \mathscr{T}, \qquad [\rho] = \frac{\mathscr{M}}{\mathscr{L}^3}.$$

We have $n = 4$ variables and $k = 3$ fundamental units, so there is $p = n - k = 1$ nondimensional parameter π describing the relationship between the energy of a bomb and the propagation of the blast wave.

To find the combination of variables which gives that nondimensional parameter, we let

$$\pi = E^a t^b \rho^c R^d,$$

with units

$$[\pi] = \left(\frac{\mathscr{M}\mathscr{L}^2}{\mathscr{T}^2}\right)^a \mathscr{T}^b \left(\frac{\mathscr{M}}{\mathscr{L}^3}\right)^c \mathscr{L}^d = \mathscr{M}^{a+c}\mathscr{L}^{2a-3c+d}\mathscr{T}^{-2a+b}.$$

If we let $a = 1$ be the free parameter, then

$$a = 1, \qquad b = 2, \qquad c = -1, \qquad d = -5,$$

then our nondimensional parameter is

$$\pi = \frac{Et^2}{\rho R^5}.$$

This tells us that

$$f\left(\frac{Et^2}{\rho R^5}\right) = 0,$$

which implies that

$$\frac{Et^2}{\rho R^5} = C.$$

Solving for E gives us

$$E = C\frac{\rho R^5}{t^2}.$$

To find C, we would have to perform a blast wave experiment with a small bomb.

Chapter 4

ODEs, Eigenvalues, Timescales, and Phase

4.1 Introduction

Consider radioactive decay. If q is the amount of radioactive substance and r is the decay rate, then the differential equation that models the system is

$$\frac{dq}{dt} = -rq.$$

By checking the units, we note that the units of r are $1/\tau$, so $1/r$ gives us a timescale for the behavior of this system.

We could easily solve this differential equation by separating variables and integrating, but notice that $e^{\lambda t}$ is an eigenfunction of d/dt. That is,

$$\frac{d}{dt}e^{\lambda t} = \lambda e^{\lambda t},$$

where λ is the **eigenvalue** and $e^{\lambda t}$ is the **eigenfunction**. If the amount of substance at time zero is given by $q(0) = q_0$, then in order for $q_0 e^{\lambda t}$ to be a solution of the differential equation, it has to be that $\lambda = -r$. So the solution is

$$q = q_0 e^{-rt}.$$

Exponents don't have units, so in order for the exponent on e to make sense, $1/\lambda = -1/r$ must have units of time. This tells us that $1/\lambda$ is proportional to the **characteristic time** or the **fundamental time** of the system and λ is proportional to the characteristic or fundamental frequency. In this sense, "eigenvalue" means **characteristic value**.

For time-dependent exponential decay, **half-life** is an arbitrary but useful timescale.

4.2 Theoretical Background

We can write any higher order ODE as a first order linear system of ODEs. For example, with

$$y''' + by'' + cy' + fy = 0,$$

we can make the substitutions

$$\begin{aligned} y_0 &= y \\ y_1 &= y' = y_0' \\ y_2 &= y'' = y_1', \end{aligned}$$

then

$$y_0' = y_1$$
$$y_1' = y_2$$
$$y_2' = -by_2 - cy_1 - fy_0.$$

Now we have three first order ODEs instead of one third order ODE. We can write this as

$$\begin{bmatrix} y_0' \\ y_1' \\ y_2' \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -f & -c & -b \end{bmatrix} \begin{bmatrix} y_0 \\ y_1 \\ y_2 \end{bmatrix},$$

or as

$$\vec{\boldsymbol{y}}' = A\vec{\boldsymbol{y}}.$$

The general solution is then of the form

$$\vec{\boldsymbol{y}} = \sum_{k=1}^{3} C_k \vec{\boldsymbol{z}}_k e^{\lambda_k t}.$$

For the more general case, we let $D = d/dt$ be the differential operator, then our ODE can be written as

$$\sum_{k=0}^{m} \alpha_k D^k y = f.$$

Let $y_k = D^k y = Dy_{k-1}$ for $k = 0, 1, 2, \ldots, m-1$, then we get the system of first order ODEs

$$Dy_{k-1} = y_k \qquad \text{for } k = 0, 1, 2, \ldots, m-1$$
$$Dy_m = \frac{1}{\alpha_m}\left(f - \sum_{k=0}^{m-1} \alpha_k y_k\right),$$

which can be written as the matrix equation

$$\vec{\boldsymbol{y}}' = f' - Ay,$$

with the general solution

$$\vec{\boldsymbol{y}} = \sum_{k=0}^{m} C_k \vec{\boldsymbol{z}}_k e^{\lambda_k t}.$$

The eigenvalues λ_k are the fundamental or **natural frequencies** of the physical system being modeled and the eigenvectors $\vec{\boldsymbol{z}}_k$ are the fundamental behaviors of the system. The expressions $\vec{\boldsymbol{z}}_k e^{\lambda_k t}$ are called the **fundamental modes**, **normal modes**, or **natural modes** of the system.

More Information: [1] [2] [3] [4]

[1] https://en.wikipedia.org/wiki/Normal_mode
[2] http://www.people.fas.harvard.edu/~djmorin/waves/normalmodes.pdf
[3] http://www-thphys.physics.ox.ac.uk/people/FrancescoHautmann/Cp4p/sl_nm_1.pdf
[4] http://www-thphys.physics.ox.ac.uk/people/FrancescoHautmann/Cp4p/nm-wv-notes12.pdf

4.3 Pendulum Example

Now consider again the pendulum equation

$$\ddot{\theta} + \beta\theta = 0.$$

The first step should always be to check the units. By dimensional analysis, we know that the units of β are $1/\tau^2$, so we have a timescale that is proportional to $1/\sqrt{\beta}$, which also means a frequency proportional to $\sqrt{\beta}$. We can use these facts to check our solution in the end.

To solve the ODE, we again use the eigenfunction of the derivative operator. Choose $\theta = \theta_0 e^{\lambda t}$. Plugging it and its second derivative into the ODE, we find that $\lambda = \pm i\sqrt{\beta}$. This tells us that there are two fundamental frequencies, $i\sqrt{\beta}$ and $-i\sqrt{\beta}$ both which have units of one over time.

We can also solve this as a first order system of equations. Notice that

$$\begin{aligned} \dot{\theta} &= v \\ \dot{v} &= \ddot{\theta} = -\beta\theta, \end{aligned}$$

where v is the angular speed of the pendulum. We can write this as the matrix equation

$$\begin{bmatrix} \dot{\theta} \\ \dot{v} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\beta & 0 \end{bmatrix} \begin{bmatrix} \theta \\ v \end{bmatrix}.$$

We can write this as

$$\dot{\vec{\Theta}} = A\vec{\mathbf{\Theta}},$$

where $\vec{\mathbf{\Theta}}$ is the vector containing θ and v and A is the 2×2 matrix. If we assume a solution of the form

$$\vec{\mathbf{\Theta}} = \vec{z} e^{\lambda t},$$

and substitute this into the matrix equation, we get

$$\begin{aligned} \dot{\vec{\Theta}} &= A\vec{\mathbf{\Theta}} \\ \lambda \vec{z} e^{\lambda t} &= A \vec{z} e^{\lambda t} \\ (\lambda \vec{z} - A\vec{z})\, e^{\lambda t} &= 0 \\ A\vec{z} &= \lambda \vec{z}. \end{aligned}$$

So λ must be an eigenvalue of the system and $\vec{z}$ must be an eigenvector. The general solution of the system is

$$\vec{\mathbf{\Theta}} = C_1 \vec{z}_1 e^{\lambda_1 t} + C_2 \vec{z}_2 e^{\lambda_2 t}.$$

The first term on the right shows the fundamental behavior of the system associated with the fundamental frequency λ_1, and the second term on the right shows the fundamental behavior of the system associated with the fundamental frequency λ_2. Since there are two λ, there are two timescales/frequencies, and the overall motion of the system is an arbitrary combination of the two orthogonal (i.e. independent) motions.

To find the eigenvalues, we solve the characteristic equation

$$\det(A - \lambda I) = 0,$$

for λ. In our case, we find that $\lambda = \pm i\omega$ if $\beta = \omega^2$. Then we find the eigenvectors $\vec{z}_k$ by plugging λ_k back into the matrix equation

$$(A - \lambda_k I)\vec{z}_k = 0.$$

In our case, we find that the eigenvector associated with $\lambda_1 = i\omega$ is

$$\vec{z}_1 = \begin{bmatrix} 1 \\ i\omega \end{bmatrix} z_1,$$

and the eigenvector associated with $\lambda_2 = -i\omega$ is

$$\vec{z}_2 = \begin{bmatrix} 1 \\ -i\omega \end{bmatrix} z_1.$$

For convenience, we let $z_1 = 1$, then our solution is

$$\vec{\mathbf{\Theta}} = \begin{bmatrix} \theta \\ v \end{bmatrix} = C_1 \begin{bmatrix} 1 \\ i\omega \end{bmatrix} e^{i\omega t} + C_2 \begin{bmatrix} 1 \\ -i\omega \end{bmatrix} e^{-i\omega t}.$$

We call the vector

$$\vec{\mathbf{\Theta}} = \begin{bmatrix} \theta \\ v \end{bmatrix},$$

the **state** or **phase** of the system. The state vector gives the position and speed of the system as a function of time. Our solution has two independent eigenvectors, so our physical system has two independent motions. What are they?

Breaking the vector equation into its two parts, we get

$$\begin{aligned} \theta(t) &= C_1 e^{i\omega t} + C_2 e^{-i\omega t} \\ v(t) &= C_1 i\omega e^{i\omega t} - C_2 i\omega e^{-i\omega t}. \end{aligned}$$

The constants C_1 and C_2 which are fixed by the initial conditions, must be such that θ and v are real, since the angle and the angular speed of the pendulum are always real. First, we consider only the $i\omega$ mode, which is

$$\begin{aligned} \theta(t) &= e^{i\omega t} \\ v(t) &= i\omega e^{i\omega t}. \end{aligned}$$

We can write i as $e^{i\frac{\pi}{2}}$ by plotting it on the complex plane, so we can write

$$\begin{aligned} \theta(t) &= e^{i\omega t} \\ v(t) &= \omega e^{i\omega t} e^{i\frac{\pi}{2}} = \omega e^{i(\omega t + \frac{\pi}{2})}. \end{aligned}$$

This tells us that the angle and speed of the pendulum are out of phase by $\pi/2 = 90°$. In other words, the imaginary unit i as a coefficient, serves to change the phase.

On the complex plane, $e^{i\omega t}$ traces out a counterclockwise circle of radius 1 as time passes. At $t = 0$, it is at angle 0. The expression $\omega e^{i(\omega t + \pi/2)}$ also traces out a circle, but it has radius ω and starts at angle $\pi/2$. If we plot the real parts of θ and v, we see that $\theta(t)$ traces out a cosine wave and $v(t)$ traces out a negative sine wave. That is,

$$\begin{aligned} \theta(t) &= \cos(\omega t) \\ v(t) &= \omega \cos\left(\omega t + \frac{\pi}{2}\right) = -\omega \sin \omega t. \end{aligned}$$

In summary, the eigenvectors give us the phase relations between the components of the state vector. The eigenvalues give us the frequency of each component/motion. For example, the eigenvector

$$\vec{z} = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

tells us that θ and v are in phase. The eigenvector

$$\vec{z} = \begin{bmatrix} 1 \\ -1 \end{bmatrix},$$

tells us that θ and v are 180° out of phase. The eigenvector,

$$\vec{z} = \begin{bmatrix} 1 \\ -i \end{bmatrix},$$

tells us that θ and v are 90° out of phase.

4.4 Phase Diagrams

We will be working with **dynamical systems** of the form

$$\begin{aligned} \dot{x} &= f(x, y) \\ \dot{y} &= g(x, y), \end{aligned}$$

where f and g could be nonlinear functions. If they are linear functions, then

$$\begin{aligned} f(x, y) &= a_1 x + b_1 y \\ g(x, y) &= a_2 x + b_2 y. \end{aligned}$$

The state (or phase) of the system is $(x(t), y(t))$. As a function of time, it describes a **trajectory** on the xy-plane. For example, at time t_0, the variables (x, y) are $(x(t_0), y(t_0))$ and they define a single point in the xy-plane. In the pendulum example, x and y are the angle and speed of the pendulum. At a later time, t_1, variables (x, y) are $(x(t_1), y(t_1))$ and define a second point in the xy-plane. Also, for all times between t_0 and t_1, there are other points. The set of points define a trajectory (including a direction) in the xy-plane. The graph of all possible trajectories is called the **phase diagram**.

Phase diagrams are often used to study the equilibria of dynamical systems and the behavior of those systems near equilibria. An **equilibrium point** or a **critical point** is where the two variables are not changing with respect to time. That is, $\dot{x} = \dot{y} = 0$. A point (x_{eq}, y_{eq}) is a critical point of $\dot{x} = f(x, y)$ and $\dot{y} = g(x, y)$ if

$$\begin{aligned} f(x_{eq}, y_{eq}) &= 0 \\ g(x_{eq}, y_{eq}) &= 0. \end{aligned}$$

Equilibrium points are important because most of the things we see in nature are in or near equilibrium. We are also interested in the **stability** of a system. That is, what is the behavior of the system near an equilibrium point?

An important feature of critical points on a phase diagram is that trajectories do not pass through critical points. They may be the endpoints of trajectories, but the trajectories do not pass on through. Phase diagrams are helpful because it is often easier to calculate one than to solve the original problem. And we can learn a lot about the problem by studying the phase diagram.

The phase diagram is really just the vector field

$$\vec{v}(x, y) = f(x, y)\, \hat{x} + g(x, y)\, \hat{y} = \dot{x}\hat{x} + \dot{y}\hat{y},$$

which gives the vectors tangent to the trajectories. The slope of a given trajectory at time t is given by

$$\frac{dy}{dx} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}} = \frac{g(x, y)}{f(x, y)},$$

which we can think of as the equation for a trajectory.

We will only be working with systems of two ODEs, but this covers a lot of mechanics. Newton's second law $F = ma$, for example, is a second order ODE that can be converted to two first order ODEs.

Pendulum Example

Recall the state of the pendulum was

$$\theta(t) = \theta_0 \cos(\omega t)$$
$$v(t) = \omega \theta_0 \sin(\omega t).$$

Notice that

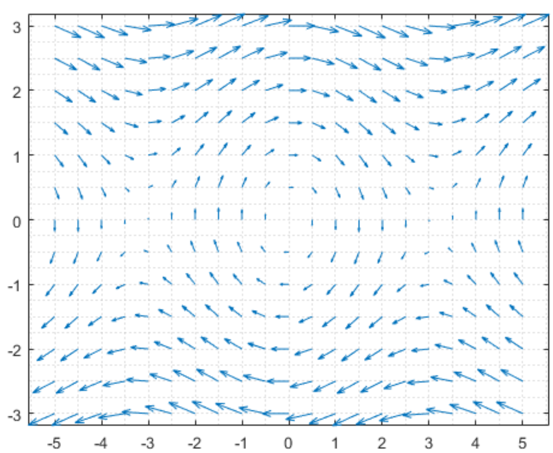
$$\theta^2 + \left(\frac{v}{\omega}\right)^2 = \theta_0^2 \cos^2(\omega t) + \theta_0^2 \sin^2(\omega t) = \theta_0^2,$$

describes a circle. We can plot $\frac{v}{\omega}$ versus θ to draw the state/phase of the pendulum, where the point $(\theta(t), v(t))$, which is on the circle described above, gives us the angle and angular speed of the pendulum at any time t.

Since the graph is a circle, and it relates θ and v, we can easily see the phase relation between θ and v by noting that when $\theta = 0$, $v =$max, and when $v = 0$, $\theta =$max. When $t = 0$, we get the point $(\theta_0, 0)$. From this graph, we can also plot the solution θ which is a cosine wave, by noting that at $t = 0$, θ is at max, then it goes through zero to negative max before going back up, and so on.

Since the pendulum's trajectory is a circle about $(0, 0)$, that point is a critical point.

The plot referred to above is a single possible trajectory of a pendulum. If the pendulum's energy is a little higher, its trajectory is a slightly larger circle, and if its energy is lower, its trajectory is a slightly smaller circle. The phase diagram shown below shows all possible trajectories of the pendulum.



Notice that for small energies, the trajectories are all circles, which correspond to a pendulum swinging. However, if the pendulum's energy is very large, the trajectories are no longer closed circles. This corresponds to pendulums with so much energy that they are rotating completely instead of swinging back and forth.

It can be plotted in MatLab with the following code:

```
[x,y] = meshgrid(-5:0.2:5,-3:0.2:3);
f = y;
g = -sin(x);
```

```
quiver(x,y,f,g)
grid minor
axis tight
```

Node Example

Suppose our system is described by

$$\begin{aligned}\dot{x} &= -x\\ \dot{y} &= -y.\end{aligned}$$

Recall that the critical points occur where $\dot{x} = \dot{y} = 0$. In this case, there is a critical point at $(x, y) = (0, 0)$. What is the behavior of the system if we start near the critical point?

We have that

$$\frac{g}{f} = \frac{\dot{y}}{\dot{x}} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}} = \frac{dy}{dx} = \frac{-y}{-x} = \frac{y}{x}.$$

Separating variables and integrating, we get

$$\begin{aligned}\frac{dy}{dx} &= \frac{y}{x}\\ \frac{1}{y}dy &= \frac{1}{x}dx\\ \int \frac{1}{y}\,dy &= \int \frac{1}{x}\,dx\\ \ln y &= \ln x + C\\ y &= Cx.\end{aligned}$$

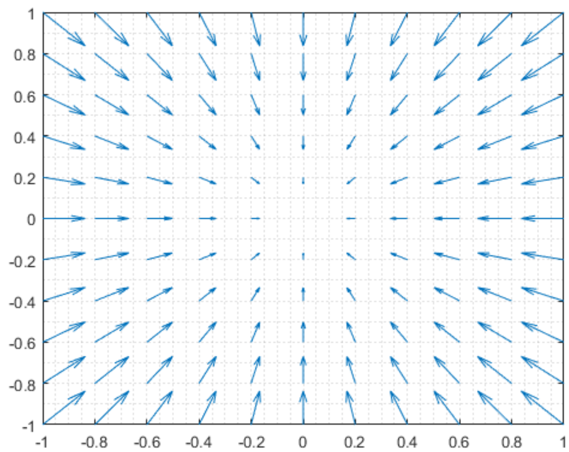
This tells us that on the phase diagram, the trajectories are straight lines through the origin. The arbitrary constant C depends on the initial condition.

To find the direction of the lines, we need to consider what happens as time passes. From the equations of the system, we know that if $x > 0$, then $\dot{x} < 0$, so the direction is to the left. If $x < 0$, then $\dot{x} > 0$, so the direction is to the right. So the phase diagram is a bunch of straight lines through the origin, and the direction arrows on those lines all point toward the origin.

To plot this phase diagram in MatLab, we use the code

```
[x,y] = meshgrid(-1:0.2:1,-1:0.2:1);
f = -x;
g = -y;
quiver(x,y,f,g)
grid minor
axis tight
```

which gives us the graph shown below.



What is the behavior of the system as a function of time? We know that all trajectories approach the critical point at $(0, 0)$, but how fast or how slow do they approach it? To find out, we just solve the ODE $\dot{x} = -x$ which has the solution $x = e^{-t}$. So the system approaches the critical point at an exponential decay rate.

There are four basic types of critical points. That is, there are four possible behaviors of a system near equilibrium. A **node** is a critical point in which all the trajectories go toward or away from it. That is, all trajectories start or end at the critical point. The example given above describes a node.

Center Example

The second type of critical point occurs when all trajectories are circles around the critical point. Such a critical point is called a **center**.

Consider the system described by the system of ODEs

$$\dot{x} = -y$$
$$\dot{y} = x.$$

Notice that there is again a critical point at $(0, 0)$. Taking the relevant derivative, we have that

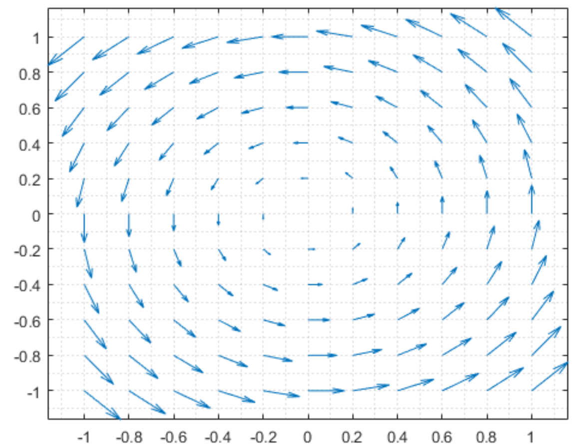
$$\frac{dy}{dx} = \frac{g}{f} = -\frac{x}{y}.$$

Separating variables and integrating, we find that

$$x^2 + y^2 = C,$$

that is, the trajectories are circles about the origin of varying radius. So $x(t)$ and $y(t)$ are oscillating in time.

When y is positive, $\dot{x}$ is negative, so the direction of motion is in the counterclockwise direction.



Saddle Example

The third type of critical point is called a **saddle**.

Consider the system described by the system of ODEs

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

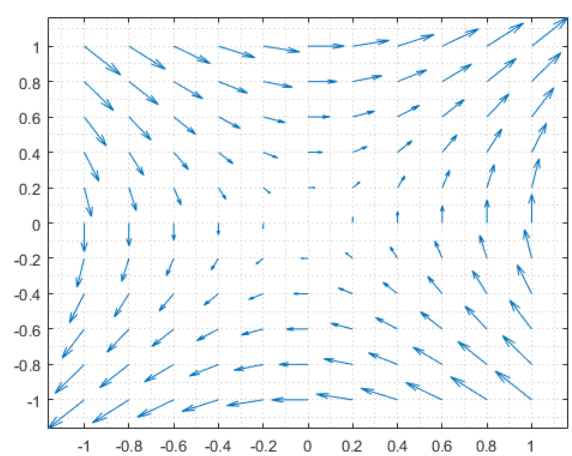
Notice that there is again a critical point at $(0, 0)$. Taking the relevant derivative, we have that

$$\frac{dy}{dx} = \frac{g}{f} = \frac{x}{y}.$$

Separating variables and integrating, we find that

$$-x^2 + y^2 = C,$$

that is, the trajectories are hyperbolas. If y is positive, then so is $\dot{x}$, and if y is negative, so is the change in x. This allows us to know the directions of the trajectories.



Spiral Example

The fourth and final kind of critical point is called a **spiral**
Consider the system described by the system of ODEs

$$\dot{x} = ax - y$$
$$\dot{y} = x + ay.$$

This is a linear system of ODEs, so there can be only one critical point, since there is only one solution to the system of equations. That critical point is at $(0, 0)$. Rearranging to get the relevant derivative, we get

$$\frac{dy}{dx} = \frac{\dot{y}}{\dot{x}} = \frac{x + ay}{ax - y}.$$

This is a nasty differential equation, but it becomes a lot easier if we convert it to polar coordinates using the transformation equations

$$x = r\cos\theta$$
$$y = r\sin\theta$$
$$r^2 = x^2 + y^2$$
$$\theta = \tan^{-1}\left(\frac{y}{x}\right).$$

Implicitly differentiating $r^2 = x^2 + y^2$ with respect to x gives us

$$r\frac{dr}{dx} = x + y\frac{dy}{dx}.$$

Implicitly differentiating $\theta = \tan^{-1}\left(\frac{y}{x}\right)$ with respect to x gives us

$$\begin{aligned} \frac{d\theta}{dx} &= \frac{1}{1 + \frac{y^2}{x^2}} \cdot \frac{x\frac{dy}{dx} - y}{x^2} \\ &= \frac{x\frac{dy}{dx} - y}{x^2 + y^2} \\ r^2\frac{d\theta}{dx} &= x\frac{dy}{dx} - y. \end{aligned}$$

Dividing the first result by the second result gives us

$$\frac{1}{r}\frac{dr}{d\theta} = \frac{x + y\frac{dy}{dx}}{x\frac{dy}{dx} - y}.$$

Substituting in $\frac{dy}{dx} = \frac{x+ay}{ax-y}$ and simplifying gives us

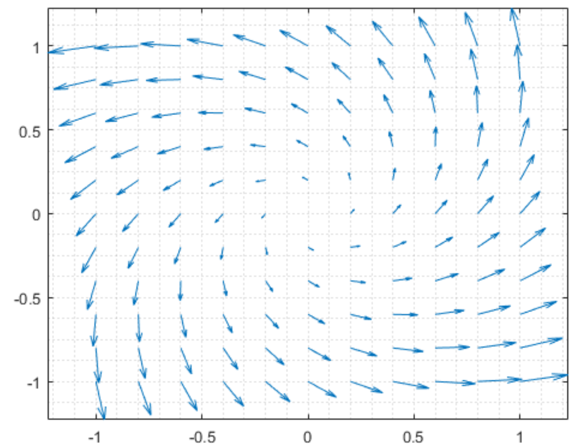
$$\frac{dr}{d\theta} = ar,$$

which has the solution

$$r = Ce^{a\theta}.$$

This solution is a spiral. If $a > 0$ then it is an increasing spiral, and if $a < 0$, it is a decreasing spiral.

A graph of the system for $a = 0.7$ is shown below.

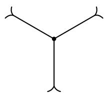
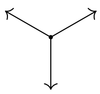
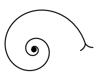
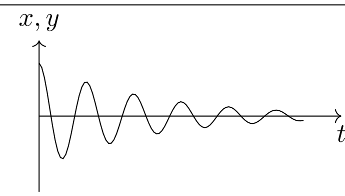
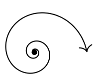
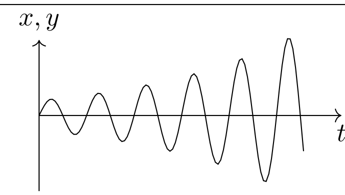
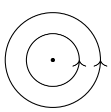
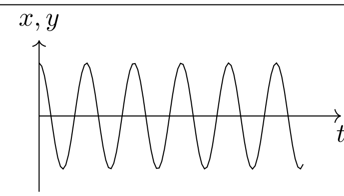
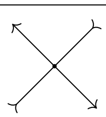


4.5 Equilibrium and Stability

The critical points of our system of ODEs correspond to equilibria of the system. If the state of the system remains in the neighborhood of an equilibrium, then the equilibrium is stable. More precisely, a critical point is globally stable if every trajectory inside some arbitrary circle remains inside the circle. Otherwise, the critical point is unstable.

In terms of the phase plane, if all the trajectories near a critical point stay near the point, then the critical point is stable. If all of the trajectories approach the critical point as time goes to infinity (e.g. a spiral on the phase plane), the critical point is asymptotically stable. If any path in the vicinity of the critical point moves away from the point, then the critical point is unstable. For example, a saddle on the phase plane is an unstable critical point. Although one specific trajectory is stable, not *all* paths stay in the vicinity of the critical point, so the saddle is unstable.

A summary of the topologies and classifications of the different critical points is shown in the table below.

Topology	Class	Stability	Time Behavior
	Node	Stable	Varies depending on direction of approach
	Node	Unstable	Varies depending on direction of approach
	Spiral	Stable	
	Spiral	Unstable	
	Center	Stable	
	Saddle	Unstable	Varies depending on direction of approach

4.6 Linearization

We want to be able to tell the topology and stability of a system without actually solving the differential equations. We don't have general techniques for solving nonlinear ODEs. In general, if our system of equations

$$\begin{aligned}\dot{x} &= f(x, y)\\ \dot{y} &= g(x, y),\end{aligned}$$

is nonlinear, then we cannot solve them. However, we need a way to identify the equilibria and stability using root-finding methods.

To find the stability of a nonlinear system, expand the nonlinear solution to the linear terms to find the linear stability. For small perturbations, this will be the same as the stability of the nonlinear system.

Let (x_0, y_0) be an equilibrium point. We want to approximate $f(x, y)$ and $g(x, y)$ for small perturbations about the equilibrium point. Doing a multivariable Taylor expansion of $f(x, y)$ and $g(x, y)$ about (x_0, y_0), gives us

$$
\begin{aligned}
f(x,y) \approx \quad & f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0) + \cdots \\
g(x,y) \approx \quad & g(x_0,y_0) + \frac{\partial g}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial g}{\partial y}(x_0,y_0)(y-y_0) + \cdots
\end{aligned}
$$

Since (x_0, y_0) is an equilibrium point, we know that $f(x_0, y_0) = g(x_0, y_0) = 0$. Ignoring the larger terms, we have as a first-order approximation (i.e. linear approximation), that

$$
\begin{aligned}
\dot{x} = f(x,y) &\approx \frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0) \\
\dot{y} = g(x,y) &\approx \frac{\partial g}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial g}{\partial y}(x_0,y_0)(y-y_0).
\end{aligned}
$$

We can write this as

$$
\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x}(x_0,y_0) & \frac{\partial f}{\partial y}(x_0,y_0) \\ \frac{\partial g}{\partial x}(x_0,y_0) & \frac{\partial g}{\partial y}(x_0,y_0) \end{bmatrix} \begin{bmatrix} x-x_0 \\ y-y_0 \end{bmatrix},
$$

or

$$
\dot{\vec{z}} = A\vec{z},
$$

where A is the **Jacobian matrix** of the system at the equilibrium point (x_0, y_0). We can denote A by

$$
A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.
$$

The solutions of this 2×2 matrix are of the form

$$
\vec{z} = c_1 \vec{z}_1 e^{\lambda_1 t} + c_2 \vec{z}_2 e^{\lambda_2 t},
$$

where $\vec{z}_i$ are the eigenvectors and λ_i are the eigenvalues. The eigenvalues tell us the topology and the stability of the critical points.

To find the eigenvalues, we calculate the characteristic equation, which is,

$$
\begin{aligned}
(a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} &= 0 \\
\lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{21}a_{22} &= 0 \\
\lambda^2 - \mathrm{Tr}(A)\,\lambda + \det A &= 0.
\end{aligned}
$$

So the eigenvalues depend on the trace and the determinant of A, which are easy to calculate for 2×2 matrices. If we define

$$
\begin{aligned}
p &= \mathrm{Tr}\ A \\
q &= \det A,
\end{aligned}
$$

then the characteristic equation becomes

$$
\lambda^2 - p\lambda + q = 0,
$$

and has solutions given by the quadratic formula

$$
\lambda = \frac{p \pm \sqrt{p^2 - 4q}}{2}.
$$

There are three possibilities for the eigenvalues:

Eigenvalues	Topology	p	$p^2 - 4q$
Both positive and real	Stable node	< 0	> 0
Both negative and real	Unstable node	> 0	> 0
Both real, one positive and one negative	Unstable saddle	> 0	> 0
Complex with $\Re(\lambda_i) = 0$	Stable center	0	< 0
Complex with $\Re(\lambda_i) < 0$	Stable spiral	< 0	< 0
Complex with $\Re(\lambda_i) > 0$	Unstable spiral	> 0	< 0

Table 4.1: Classifications of Equilibrium Points

- λ_1 and λ_2 are real and distinct if $p^2 - 4q > 0$
- λ_1 and λ_2 are real and repeated if $p^2 - 4q = 0$
- λ_1 and λ_2 are complex conjugates if $p^2 - 4q < 0$

Below is a graph of q versus p and the different regions delineated by the axis and the function $p^2 - 4q = 0$. For example, in the region where $p > 0$, $q > 0$, and $p^2 - 4q > 0$, the figure shows that critical points will be unstable nodes. Similarly, whenever $q < 0$, critical points will be saddles. When $q > 0$ and $p = 0$ the critical point is a center.

The graph gives the stability of the equilibrium points of a linear autonomous system of the form

$$\begin{aligned} \dot{x} &= Ax + By \\ \dot{y} &= Cx + Dy, \end{aligned}$$

with

$$\begin{aligned} p &= A + D \\ q &= AD - BC. \end{aligned}$$

As demonstrated above, this also works for the linearized form of non-linear systems. Then A, B, C, and D will be partial derivatives evaluated at the equilibrium points.

Example 4.6.1

Find the equilibria and stabilities of the system

$$\begin{aligned} \dot{x} &= -y \\ \dot{y} &= x. \end{aligned}$$

The only critical point is at $(0, 0)$.

We can write this as

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Computing the trace and the determinant of the coefficient matrix, we get

$$\begin{aligned} p &= 0 \\ q &= 1, \end{aligned}$$

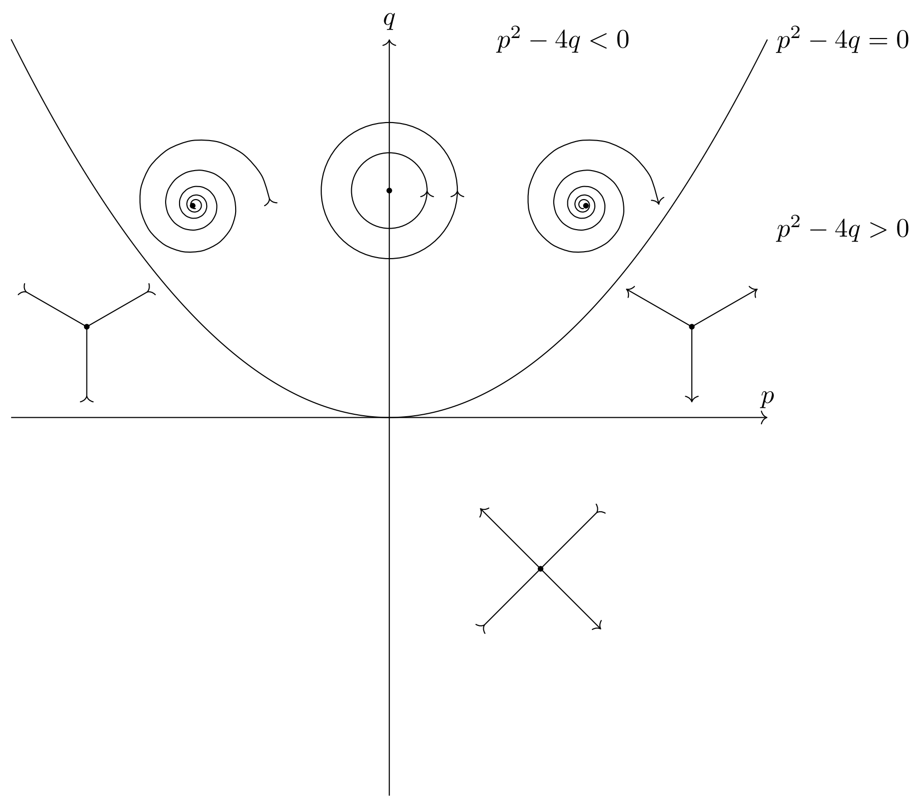


Figure 4.1: Classifications of Equilibrium Points in a Phase Plane

which implies that our critical point at $(0, 0)$ is a stable center.

Example 4.6.2

Find the equilibria and stabilities of the system

$$\dot{x} = ax - y$$
$$\dot{y} = x + ay.$$

The only critical point is at $(0, 0)$.

We can write this as

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} a & -1 \\ 1 & a \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Computing the trace and the determinant of the coefficient matrix, we get

$$p = 2a$$
$$q = a^2 + 1.$$

We have that $p^2 - 4q = -4 < 0$, so we are above the quadratic in the q versus p plane. So if $a > 0$ the critical point at $(0, 0)$ is an unstable spiral, and if $a < 0$ it is a stable spiral.

Example 4.6.3

Find the equilibria and stabilities of the system

$$\begin{aligned}\dot{x} &= x\\ \dot{y} &= -x + 2y.\end{aligned}$$

The only critical point is at $(0, 0)$.

We can write this as

$$\begin{bmatrix}\dot{x}\\ \dot{y}\end{bmatrix} = \begin{bmatrix}1 & 0\\ -1 & 2\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}.$$

Computing p and q, we get

$$\begin{aligned}p &= 3\\ q &= 2.\end{aligned}$$

We have that $p^2 - 4q = 1 > 0$, so we are below the quadratic in the q versus p plane. This implies that our critical point at $(0, 0)$ is an unstable node.

4.7 Lanchester Combat Model

Imagine that two armies A and B have engaged in battle. Let $A(t)$ and $B(t)$ be the strength (e.g. number of soldiers) of each army as a function of time. The strength of each army declines during the battle since some of its soldiers are getting killed. The rate of this decline is proportional to the strength of the opposing army. The larger the opposing army, the more soldiers will die per hour, for example. So we have that

$$\frac{dA}{dt} = -bB, \qquad \frac{dB}{dt} = -aA,$$

where b and a are the proportionality constants. We have included the negative signs since we know the rates are negative, and we want the constants to be positive. The proportionality constants denote the *effectiveness* of the armies. For example, b is the effectiveness of army B. The larger b is, the larger bB is and so the greater the rate of army A's decline.

Dividing the first equation by the second equation (justified by the chain rule and inverse function rule), we get

$$\frac{\frac{dA}{dt}}{\frac{dB}{dt}}\frac{dA}{dt} \cdot \frac{dt}{dB} = \frac{dA}{dB} = \frac{bB}{aA}.$$

we now have a separable differential equation that we can solve by integrating

$$\begin{aligned}\frac{dA}{dB} &= \frac{bB}{aA}\\ aA\,dA &= bB\,dB\\ \int aA\,dA &= \int bB\,dB\\ \frac{a}{2}A^2 &= \frac{b}{2}B^2 + C_1\\ aA^2 - bB^2 &= C.\end{aligned}$$

From the original system of ODEs, we note that there is a critical point at $(0, 0)$. Calculating the trace and determinant formed by the coefficient matrix from the equation

$$\begin{bmatrix} \dot{A} \\ \dot{B} \end{bmatrix} = \begin{bmatrix} 0 & -b \\ -a & 0 \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix},$$

we find that

$$p = 0$$
$$q = -ab.$$

This tells us that the critical point at $(0, 0)$ is a saddle. We can also see this in $aA^2 - bB^2 = C$, since this is the equation of a hyperbola centered at $(0, 0)$.

We can deduce the general appearance of the phase plane (the plot of A on the vertical axis and B on the horizontal axis) by examining the system of ODEs further. We know there is a saddle at $(0, 0)$. We only care about the first quadrant of the graph, since there cannot be negative troop strengths. When $B > 0$, $\dot{A} < 0$ and when $A > 0$, $\dot{B} < 0$, so all the arrows in the first quadrant point in the general direction of $(0, 0)$. If $B = 0$ then $\dot{A} = 0$, and if $A = 0$, $\dot{B} = 0$. This tells us that the trajectories are perpendicular to the axes where they cross them.

From the initial conditions, we know that the constant in $aA^2 - bB^2 = C$ is $C = aA_0^2 - bB_0^2$. We can therefore, rearrange this equation into the equation of a hyperbola as

$$\frac{A^2}{A_1^2} - \frac{B^2}{B_1}^2 = 1,$$

where

$$A_1 = \sqrt{\frac{aA_0^2 - bB_0^2}{a}}, \qquad B_1 = \sqrt{\frac{aA_0^2 - bB_0^2}{b}},$$

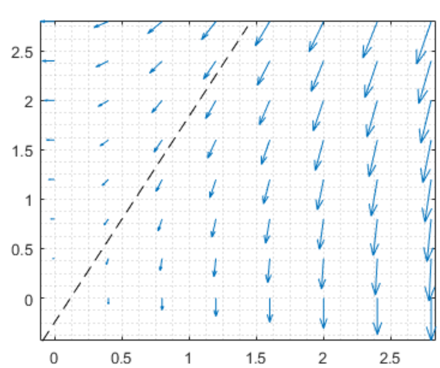
where A_0 and B_0 are the initial troop strengths. From the properties of hyperbolae, we obtain the slope of the asymptote as

$$\sqrt{\frac{b}{a}}.$$

So the equation of the asymptote is

$$A = \sqrt{\frac{b}{a}} B.$$

A plot of the phase plane (depicting the asymptote) is shown here for the case $b = 16$ and $a = 4$. We have A on the vertical axis and B on the horizontal axis.



The significance of the asymptote is that for initial points falling to the right of the asymptote, all trajectories end at the horizontal axis, and for initial points on the left of the asymptote, all trajectories end at the vertical asymptote. This means that if (B_0, A_0), which is the initial troop strengths, falls on the right side of the asymptote, then B will win since B will have a final nonzero value and A will be zero. Otherwise, A will win. In other words, A wins if

$$A_0 > \sqrt{\frac{b}{a}} B_0.$$

Otherwise, A loses if

$$A_0 < \sqrt{\frac{b}{a}} B_0.$$

To solve the time related questions, such as 'how long will the battle last?' we have to solve the system of ODEs to find $A(t)$ and $B(t)$.

We can write the system of ODEs represented by

$$\begin{bmatrix} \dot{A} \\ \dot{B} \end{bmatrix} = \begin{bmatrix} 0 & -b \\ -a & 0 \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix},$$

as $\dot{\vec{Q}} = R\vec{\boldsymbol{Q}}$ has a solution of the form

$$\vec{\boldsymbol{Q}} = C_1 \vec{\boldsymbol{z}}_1 e^{\lambda_1 t} + C_2 \vec{\boldsymbol{z}}_2 e^{\lambda_2 t},$$

where $\vec{\boldsymbol{z}}_i$ are the eigenvectors and λ_i are the eigenvalues. Computing the characteristic equation and solving for the eigenvalues, we find that

$$\lambda = \pm\sqrt{ab}.$$

Finding the eigenvectors, we find the general solution can be written as

$$\vec{\boldsymbol{Q}} = C_1 \begin{bmatrix} -\frac{\sqrt{ab}}{a} \\ 1 \end{bmatrix} e^{\sqrt{ab}t} + C_2 \begin{bmatrix} \frac{\sqrt{ab}}{a} \\ 1 \end{bmatrix} e^{-\sqrt{ab}t}.$$

So the troop strengths as functions of time are

$$\begin{aligned} A(t) &= -C_1 \sqrt{\frac{b}{a}} e^{\sqrt{ab}t} + C_2 \sqrt{\frac{b}{a}} e^{-\sqrt{ab}t} \\ B(t) &= C_1 e^{\sqrt{ab}t} + C_2 e^{-\sqrt{ab}t}. \end{aligned}$$

In MatLab, this solution can be obtained using the code

```
syms A(t) B(t) a b
eqn1 = diff(A) == -b*B;
eqn2 = diff(B) == -a*A;
[ASol(t) BSol(t)] = dsolve(eqn1, eqn2)
```

The initial conditions gives us

$$\begin{aligned} A(0) = A_0 &= -C_1 \sqrt{\frac{b}{a}} + C_2 \sqrt{\frac{b}{a}} \\ B(0) = B_0 &= C_1 + C_2, \end{aligned}$$

from which we find that

$$\begin{aligned} C_1 &= \frac{bB_0 - \sqrt{ab}A_0}{2b} \\ C_2 &= \frac{bB_0 + \sqrt{ab}A_0}{2b}. \end{aligned}$$

Substituting these into the general solution and simplifying gives us

$$A(t) = \frac{aA_0 - \sqrt{ab}B_0}{2a}e^{\sqrt{ab}t} + \frac{\sqrt{ab}B_0 + aA_0}{2a}e^{-\sqrt{ab}t}$$
$$B(t) = \frac{bB_0 - \sqrt{ab}A_0}{2b}e^{\sqrt{ab}t} + \frac{bB_0 + \sqrt{ab}A_0}{2b}e^{-\sqrt{ab}t}.$$

Example 4.7.1

$A_0 = 1200$ Romulan fighters and $B_0 = 1000$ Klingon fighters engage in battle. The effectiveness of both armies is $a = b = 0.5$ per hour. The Romulans are well on their way to victory, but when there are only 500 Klingon fighters remaining, the Romulans develop food poisoning, probably as a result of Klingon sabotage, and half of their remaining fighters are incapacitated. Which army will win? How long will the battle last? How many fighters will the remaining army have left?

The functions of time simplify to

$$A(t) = 100e^{0.5t} + 1100e^{-0.5t}$$
$$B(t) = -100e^{0.5t} + 1100e^{-0.5t}.$$

These functions are only valid before the Romulans develop food poisoning. To find the "half-life" of the Klingon army, we have to solve

$$500 = -100e^{0.5t_{1/2}} + 1100e^{-0.5t_{1/2}},$$

for $t_{1/2}$. A numerical approximation gives us $t_{1/2} = 1.005561$ hours. Plugging this into $A(t)$ tells is that there are 831 Romulan fighters left at that time. Half of them promptly fall sick.

Now we essentially have a new battle with the same values for a and b, but with $A_0 = 416$ and $B_0 = 500$. We solve again for the initial conditions, and we find that our new time equations are

$$A(t) = -42e^{0.5t} + 458e^{-0.5t}$$
$$B(t) = 42e^{0.5t} + 458e^{-0.5t}.$$

Clearly, the Klingons will win since $A(t)$ goes to zero.

Finding the zero of $A(t)$, we get

$$42e^{0.5t} = 458e^{-0.5t}$$
$$e^t = \frac{458}{42}$$
$$t = 2.39.$$

So in total, the fight lasts $2.39 + 1.00 = 3.39$ hours.

The number of Klingon fighters remaining is

$$B(2.39) = 42e^{0.5(2.39)} + 458e^{-0.5(2.39)} = 277.$$

4.8 Predator-prey Model

Consider a mountain valley containing rabbits and wolves. The rabbits eat planets, of which there are plenty, and the wolves eat rabbits. What is the system of ODEs that describes the populations of rabbits and wolves as a function of time? Is there an equilibrium population, where the population of both is stable?

The population of wolves and rabbits given by $W(t)$ and $R(t)$, we treat as continuous functions of time.

How does $R(t)$ vary with time? We know it depends on the population of rabbits via the reproduction rate and the population of wolves via the getting-eaten rate, so the rabbit growth rate is a function of the populations of both species

$$\dot{R} = f(R, W).$$

It makes sense that the rabbit growth rate (i.e. the time-rate of change of the rabbit population) is proportional to the population of rabbits. The more rabbits, the higher the growth rate because there are more rabbits reproducing. It also makes sense that the proportionality constant α is a function of the wolf population, because the more wolves, the slower the rabbit growth rate, so $\dot{R} = \alpha(W)R$. In the absence of any wolves, the simplest model of rabbit growth rate is a constant $\alpha(0) = a_1$. The presence of wolves reduces the rabbit growth rate and the simplest model is the linear one, so $\alpha(W) = a_1 - b_1 W$. So the simplest model of the rabbit growth rate is given by

$$\dot{R} = (a_1 - b_1 W)R.$$

Notice that our constants a_1 and b_1 are both positive. The sign has been made explicit in the equation for convenience.

How does $W(t)$ vary with time? We know it depends on the population of wolves via competition and the population of rabbits—their source of food, so the population of wolves is a function of the populations of both species

$$\dot{W} = g(R, W).$$

The simplest model for the growth rate of wolves is that it is proportional to the current population of wolves where the proportionality constant β depends on the population of rabbits. That is, $\dot{W} = \beta(R)W$. The more rabbits, the higher the growth rate of the wolf population, since there is more food, so a simple model is that the growth rate of the wolf population is proportional to the rabbit population. However, the more wolves, the slower their population growth rate since they are competing with each other for the rabbits. So the simplest model for the proportionality constant is $\beta(R) = -a_2 + b_2 R$. So the simplest model of the wolf population growth rate is given by

$$\dot{W} = (-a_2 + b_2 R)W.$$

We now have a pair of ODEs describing the population dynamics of wolves and rabbits

$$\begin{aligned} \dot{R} &= (a_1 - b_1 W)R \\ \dot{W} &= (-a_2 + b_2 R)W. \end{aligned}$$

We cannot solve this system because it is nonlinear. However, we can still find the equilibria and determine their stability.

To find the critical points, we determine where $\dot{R}$ and $\dot{W}$ are zero. We note that the critical points occur at

$$\begin{aligned} (R_0, W_0) &= (0, 0) \\ (R_1, W_1) &= \left(\frac{a_2}{b_2}, \frac{a_1}{b_1}\right). \end{aligned}$$

We know from the previous section, that the linear approximation of the stability at a critical point (R_i, W_i) is given by the Jacobian matrix

$$\begin{bmatrix} \frac{\partial \dot{R}}{\partial R}(R_i, W_i) & \frac{\partial \dot{R}}{\partial W}(R_i, W_i) \\ \frac{\partial \dot{W}}{\partial R}(R_i, W_i) & \frac{\partial \dot{W}}{\partial W}(R_i, W_i) \end{bmatrix}.$$

Evaluating the partial derivatives, the matrix becomes

$$\begin{bmatrix} a_1 - b_1 W_i & -b_1 R_i \\ b_2 W_i & -a_2 + b_2 R_i \end{bmatrix}.$$

Plugging in the critical point $(R_i, W_i) = (0, 0)$ gives us

$$\begin{bmatrix} a_1 & 0 \\ 0 & -a_2 \end{bmatrix}.$$

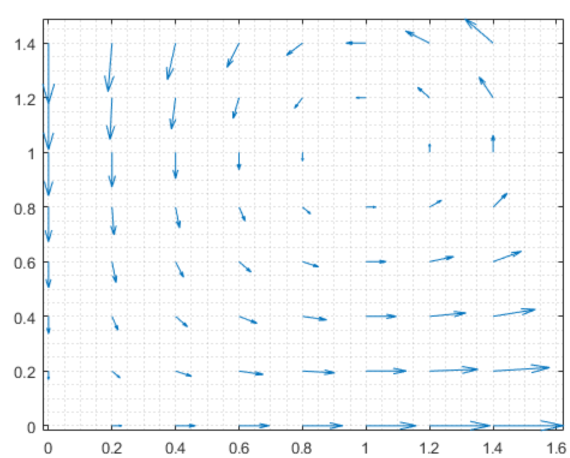
Evaluating the trace and determinant, we find that $p = a_1 - a_2$ and $q = -a_1 a_2$. Since a_1 and a_2 are positive constants, we have that $q < 0$. Checking the graph of q versus p given on an earlier page, we see that this corresponds to a saddle topology for the critical point $(0, 0)$. However, this is a rather uninteresting equilibrium point since it corresponds to all rabbits and wolves in our valley being dead.

Plugging the critical point $(R_i, W_i) = (a_2/b_2, a_1/b_1)$ into our matrix gives us

$$\begin{bmatrix} 0 & -\frac{a_2 b_1}{b_2} \\ \frac{a_1 b_2}{b_1} & 0 \end{bmatrix}.$$

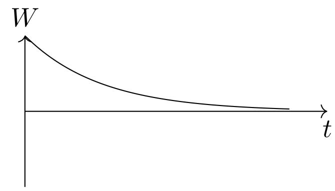
Taking the trace and determinant, we find that $p = 0$ and $q = a_1 a_2 > 0$ which corresponds to a center. This means there is a stable equilibrium at $(R_1, W_1) = (a_2/b_2, a_1/b_1)$.

If $a_1 = a_2 = b_1 = b_2 = 1$, then the phase plane ($W$ on the vertical axis and R on the horizontal axis) is given by the graph below. Note the saddle at $(0, 0)$ and the center at $(1, 1)$, the two equilibrium points of this system. With the noted values for the constants, this phase plane tells us that if the populations are $W = 1$ and $R = 1$, then there is stable equilibrium. This is obviously not realistic, but that's only because we chose unrealistic but easy to graph values for the constants.

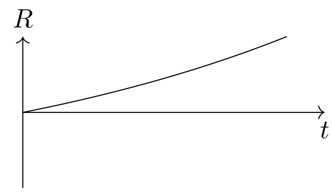


Notice also, that we only include the positive-positive quadrant of the phase plane. This is only because a negative number of wolves or rabbits does not make sense.

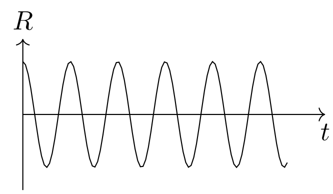
Notice that directly above the origin, the arrows are pointing toward the origin, meaning if you have wolves, but no rabbits, the population of wolves goes to zero. So the wolf population as a function of time looks like:



To the right of the origin, the arrows are pointing away from the origin, meaning if you have rabbits, but no wolves, the population of rabbits will increase forever. So the rabbit population as a function of time looks like



In the neighborhood of the stable center, both the wolf and rabbit populations oscillate. If you trace a trajectory about this critical point, you can see that as the population of rabbits decreases, so does the population of wolves, but then the population of rabbits increases, which then causes an increase in the population of wolves. Then the cycle repeats. So in the vicinity of that critical point, the rabbit population as a function of time looks like:



The wolf population in the vicinity of that critical point oscillates in the same manner, but it is out of phase with the rabbit population.

Chapter 5

Forced Behavior

So far, we have focused on the natural behavior of systems that can be modeled in differential equation form by

$$\left(\sum_{k=0}^{m} \alpha_k D^k \right) x = 0,$$

where $D = d/dt$ is a differential operator. Solutions to this differential equation are of the form $x(t) = Xe^{st}$, where X is the amplitude and s is the frequency. If we substitute this back into the differential equation, we get

$$\left(\sum_{k=0}^{m} \alpha_k D^k \right) Xe^{st} = 0$$
$$X \left(\sum_{k=0}^{m} \alpha_k s^k \right) e^{st} = 0.$$

Since e^{st} and the constant X are not zero, we can factor them out to get the **characteristic equation**

$$P(s) \equiv \sum_{k=0}^{m} \alpha_k s^k = 0.$$

$P(s)$ has roots s_k at $k = 0, 1, 2, \ldots, m$, so the general solution of the differential equation is

$$x(t) = \sum_{k=0}^{m} \beta_k e^{s_k t}, \tag{5.1}$$

where the coefficients β_k are determined by the initial conditions.

Equivalently, we can write the differential equation as a first order system of equations

$$\dot{\vec{x}} = A\vec{\boldsymbol{x}},$$

which has solutions

$$\vec{\boldsymbol{x}} = \sum_{k=0}^{m} B_k \vec{\boldsymbol{z}}_k e^{s_k t}, \tag{5.2}$$

where $\vec{\boldsymbol{z}}_k$ are the eigenvectors of the system and s_k are the eigenvalues of the system which satisfy the characteristic equation $\det(A - sI) = 0$. What's the difference between the two solutions Eq. (5.1) and (5.2)? They are the same, basically.

The solutions describe the natural or unforced behavior of the system being modeled by the differential equation. This is also called the **homogeneous solution**. The type of behavior is determined by the real and imaginary parts of s, as shown in Table (4.1) and Figure (4.1).

What if we add a forcing function? For example, instead of letting a mass on a spring oscillate on its own, what if we add a periodic forcing function by exerting a periodic force on the end of the spring opposite the mass? Forced behavior is very different from natural behavior and it often depends on the frequency of the forcing function. For example, hold a mass on a spring in your hand and let it oscillate at its natural frequency. If you now add a forcing function by periodically moving your hand up and down, the behavior of the mass on the spring will change. If you move your hand up and down at the right frequency, the amplitude of the periodic displacement of the mass will be much larger than the amplitude due to its natural oscillation. If you move your hand up and down very fast, however, the amplitude of the displacement of the mass will be much smaller.

The differential equation for a forced system is

$$\left(\sum_{k=0}^{m} \alpha_k D^k\right) x = f(t),$$

where $f(t)$ is the forcing function. The solution to this differential equation is the **forced solution** or **particular solution** as opposed to the homogeneous solution.

To "solve" this differential equation, assume $f(t) = Fe^{st}$, where F is the amplitude of the forcing, and s is the frequency of the forcing. The solutions are of the form $x = Xe^{st}$ if

$$\begin{aligned} \left(\sum_{k=0}^{m} \alpha_k D^k\right) Xe^{st} &= Fe^{st} \\ X\left(\sum_{k=0}^{m} \alpha_k s^k\right) &= F \\ X = \frac{F}{\sum_{k=0}^{m} \alpha_k s^k} &= \frac{1}{P(s)}F, \end{aligned}$$

where $P(s)$ is the characteristic polynomial. In other words, the behavior X is proportional to the amplitude F of the forcing function where the proportionality constant depends on the frequency s.

If s equals the natural frequency of the system then $P(s) = 0$, so $X = \infty$. This is the phenomenon of **resonance**.

The solution to the differential equation with forcing is

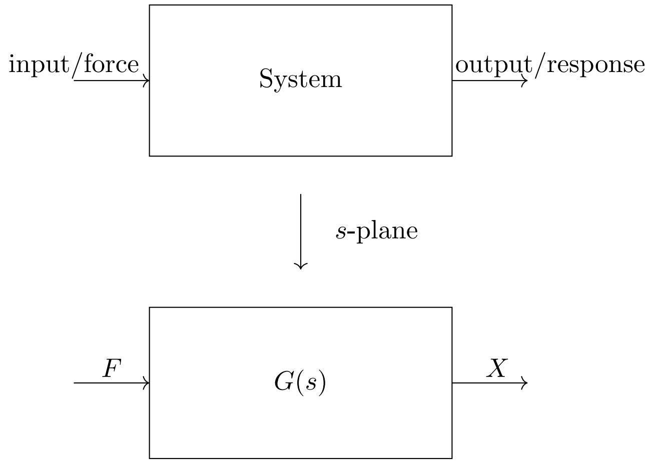
$$x(t) = Xe^{st} = \frac{1}{P(s)}Fe^{st} = G(s)Fe^{st},$$

where

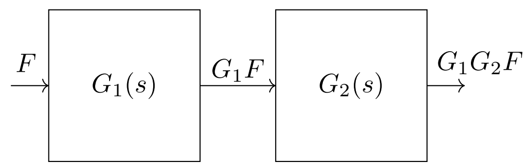
$$G(s) = \frac{1}{P(s)},$$

is the **transfer function**, which tells us how much of the input (i.e. forcing) is "transferred" to the output.

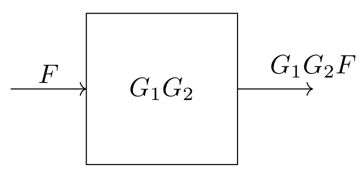
We can illustrate this visually with the figure below. That is, a system can be characterized by an input and an output. If we take the model into the s-plane, we see this as the amplitude of the input F related to the amplitude of the output X by the transfer function $G(s)$.



We can also chain many systems together creating complicated systems where the output of one subsystem forms the input of another subsystem. A computer, which consists, of many components and subsystems working together, is an example of such a complicated system.



We can also compact systems. For example, we can treat the two-part system shown above as a single system where the transfer function is the product of the two transfer functions in the two-part system.



If $G(s)$ is complex, which it will be for sinusoidal forcing, then the input (force) and the output (response) will have a phase difference. If G is complex, then we can write it in the exponential form

$$G = |G|e^{i\delta},$$

where $|G|$ is the real amplitude and δ is the phase angle. The phase angle can be calculated as

$$\delta = \tan^{-1}\left(\frac{\Im(G)}{\Re(G)}\right).$$

To formalize the relationship between the input and the repose of a system, we use the Laplace transform.

For a forced system, we ignore the natural/unforced behavior of the system since after a while this behavior decays to zero, and all that is left is the forced behavior.

We start with the ODE of the forced system

$$\left(\sum_{k=0}^{m} \alpha_k D^k\right) x(t) = f(t).$$

Next, we take the Laplace transform of both sides of the equation

$$\mathscr{L}\left\{\left(\sum_{k=0}^{m}\alpha_k D^k\right)x(t)\right\} = \mathscr{L}\left\{f(t)\right\}.$$

The Laplace transform is linear and it passes through the sum

$$\sum_{k=0}^{m}\alpha_k s^k \mathscr{L}\left\{x(t)\right\} = \mathscr{L}\left\{f(t)\right\}.$$

Notice that the some becomes the characteristic polynomial

$$P(s)\mathscr{L}\left\{x(t)\right\} = \mathscr{L}\left\{f(t)\right\},$$

so

$$\mathscr{L}\left\{x(t)\right\} = \frac{1}{P(s)}\mathscr{L}\left\{f(t)\right\}.$$

This gives us the general relationship between the input $f(t)$ and the output $x(t)$ via the transfer function

$$\mathscr{L}\left\{x(t)\right\} = G(s)\,\mathscr{L}\left\{f(t)\right\}.$$

In other words, the transfer function is the Laplace transform of the output divided by the Laplace transform of the input

$$G(s) = \frac{\mathscr{L}\left\{x(t)\right\}}{\mathscr{L}\left\{f(t)\right\}}.$$

Remember that $P(s)$ is the characteristic polynomial, and its roots are the eigenvalues or natural frequencies of the system. So if you know the transfer function, then the natural frequencies are the roots of

$$\frac{1}{G(s)} = 0.$$

5.1 Damped, Forced, Pendulum

A pendulum with friction at the pivot point is an example of a damped pendulum of length l and mass m. The damping force

$$\vec{\boldsymbol{F}}_{fric} = -c\vec{\boldsymbol{v}} = -cl\dot{\theta}\hat{\theta},$$

is proportional to the velocity and opposite in direction. The driving force is $f(t)\hat{\theta}$. The gravitational force is $m\vec{\boldsymbol{g}}$. The total force on the pendulum is

$$\vec{\boldsymbol{F}}_{tot} = (-mg\sin\theta - lc\dot{\theta})\hat{\theta} + f(t)\hat{\theta} = ml\ddot{\theta},$$

where the right side is $m\vec{\boldsymbol{a}}$ in angular form.

The ODE for the damped forced pendulum is then

$$\ddot{\theta} + \frac{c}{m}\dot{\theta} + \frac{g}{l}\sin\theta = f(t),$$

which, for small angles, can be linearly approximated as,

$$\ddot{\theta} + \frac{c}{m}\dot{\theta} + \frac{g}{l}\theta = f(t).$$

Assuming $\theta(t) = \Theta e^{st}$ and $f(t) = Fe^{st}$, then

$$s^2\Theta e^{st} + \frac{c}{m}s\Theta e^{st} + \frac{g}{l}\Theta e^{st} = Fe^{st}$$
$$\Theta = \frac{1}{s^2 + \frac{c}{m}s + \frac{g}{l}}F.$$

where F is the amplitude of the input/forcing function, and Θ is the amplitude of the response of the forced system. Writing this in terms of the transfer function, we get

$$\Theta = G(s)F.$$

where the transfer function is

$$G(s) = \frac{1}{s^2 + \frac{c}{m}s + \frac{g}{l}}.$$

Notice that the response of the system depends on the frequency s of the forcing function. If we choose $s = i\omega$, that is, the forcing function is sinusoidal, then

$$G(s) = \frac{1}{-\omega^2 + \frac{c}{m}i\omega + \frac{g}{l}}.$$

The actual solution of the ODE is

$$\theta(t) = \frac{F}{-\omega^2 + \frac{c}{m}i\omega + \frac{g}{l}}e^{i\omega t}.$$

Keep in mind that this is the "particular" or "forced" solution. Since we had a second order differential equation, there must be two solutions. The second solution is the "homogeneous" solution.

Recall that c comes from the friction force. If there is no friction, then

$$G(s) = \frac{1}{\frac{g}{l} - \omega^2}.$$

Recall that the natural frequency of a pendulum is $\omega_n = \sqrt{g/l}$. If we plug this in for ω in the equation, the denominator is zero. This tells us that if $\omega \to \sqrt{g/l}$, then $G \to \infty$. That is, if the frequency of the forcing function approaches the natural frequency of the pendulum, the response of the system becomes infinite. This is **resonance**. Notice that $G(s)$ is real in this case so $\Im(G) = 0$, so the phase angle is

$$\delta = \tan^{-1}\left(\frac{\Im(G)}{\Re(G)}\right) = 0.$$

Since there is no phase difference, this tells us that response is in phase with the forcing function.

In nature, we never see an infinite response for a finite input, because of friction. There is always friction in nature. For the transfer function with friction included

$$G(s) = \frac{1}{-\omega^2 + \frac{c}{m}i\omega + \frac{g}{l}},$$

there is no value of the forcing frequency ω that gives an infinite response G.

To find the amplitude and phase of the response, we put $G(s)$ into exponential form. To get the square of the amplitude, we can just multiply $G(s)$ by its complex conjugate

$$|G|^2 = \frac{1}{-\omega^2 + \frac{c}{m}i\omega + \frac{g}{l}} \cdot \frac{1}{-\omega^2 - \frac{c}{m}i\omega + \frac{g}{l}} = \frac{1}{\left(\frac{g}{l} - \omega^2\right)^2 + \frac{c^2\omega^2}{m^2}}.$$

Differentiating this with respect to ω allows us to find its maximum—its resonance frequency. To find the phase difference, we first rationalize $G(s)$ to separate it into its real and imaginary parts.

$$G(s) = \frac{1}{\left(\frac{g}{l} - \omega^2\right) + \frac{c}{m} i\omega} \cdot \frac{\left(\frac{g}{l} - \omega^2\right) - \frac{c}{m} i\omega}{\left(\frac{g}{l} - \omega^2\right) - \frac{c}{m} i\omega} = \frac{\frac{g}{l} - \omega^2}{\left(\frac{g}{l} - \omega^2\right)^2 + \frac{c^2\omega^2}{m^2}} - i \frac{\frac{c}{m}\omega}{\left(\frac{g}{l} - \omega^2\right)^2 + \frac{c^2\omega^2}{m^2}}.$$

To find the phase difference, we take the arctangent of the imaginary part of $G(s)$ divided by the real part

$$\delta = \tan^{-1}\left(\frac{-\frac{c}{m}\omega}{\frac{g}{l} - \omega^2}\right).$$

Notice that the phase difference is frequency dependent.

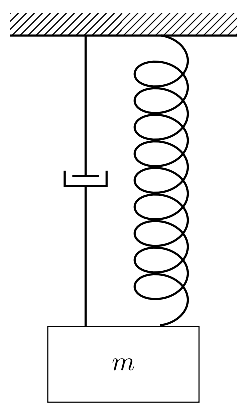
Chapter 6

Spring Models

6.1 Spring Scale

Consider a spring scale as is found in some fish scales, for example. The scale is essentially a damped spring upon which the object to be weighed is hung. The weight of the object extends the spring, and the amount of extension gives the weight of the object. The spring is damped because you don't want the object oscillating up and down for a long time—you want the spring to reach a new equilibrium as soon as possible.

The simplified conceptual model of a spring scale, is just a mass m hanging from a damped spring of length l which is attached to a support. In the diagram, the damping force is represented by what is called a "dashpot" which looks like a vertically-challenged butter churn.



Let the x direction be downward, and let x be the length of the spring with the mass attached, then the extension of the spring caused by attaching the mass is $x-l$. The total force on the mass is the sum of the spring force, the damping force, and the gravitational force

$$\vec{\boldsymbol{F}}_{net} = \vec{\boldsymbol{F}}_s + \vec{\boldsymbol{F}}_d + \vec{\boldsymbol{F}}_g.$$

where the individual forces are

$$\begin{aligned}
\vec{\boldsymbol{F}}_{net} &= m\ddot{x}\hat{x} \\
\vec{\boldsymbol{F}}_s &= -k(x-l)\hat{x} \\
\vec{\boldsymbol{F}}_d &= -b\dot{x}\hat{x} \\
\vec{\boldsymbol{F}}_g &= mg\hat{x}.
\end{aligned}$$

Putting it all together, we have that the x-component of the equation of motion, which is the only component, is

$$\ddot{x} + \frac{b}{m}\dot{x} + \frac{k}{m}(x - l) = g.$$

We are interested in the equilibrium state of the system, when $\ddot{x} = 0$ and $\dot{x} = 0$. In that case,

$$k(x - l) = mg = W,$$

where W is the weight of the mass on the spring. This is what we are trying to calculate with a spring scale. Recall that $x - l$ is how much the mass has stretched the spring.

We can write the equation of motion as

$$\ddot{x} + B\dot{x} + K\left(x - l - \frac{g}{K}\right) = 0,$$

where $B = b/m$ and $K = k/m$. We also want to know *how* equilibrium is approached, and to do that, we use the topologies from the phase plane, obtained from the Jacobian matrix. First, we change variables, so the equilibrium position of the spring (with the attached mass) is at the origin. To do that, let

$$z = x - \left(l + \frac{g}{K}\right),$$

then $z = 0$ is equivalent to $x - l = g/k$ which is equivalent to $k(x - l) = W$, and $\dot{z} = \dot{x}$ and $\ddot{z} = \ddot{x}$. Our new equation of motion is

$$\ddot{z} + B\dot{z} + Kz = 0.$$

To convert the ODE to a system of first order ODEs, we let

$$z_1 = z, \qquad z_2 = \dot{z}, \qquad \dot{z}_2 = \ddot{z},$$

which gives us

$$\begin{aligned} \dot{z}_1 &= z_2 \\ \dot{z}_2 &= -Bz_2 - Kz_1. \end{aligned}$$

We can write this as

$$\frac{d}{dt}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -K & -B \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}.$$

Computing the trace and determinant, we get $p = -B$ and $q = K$ and $p^2 - 4q = B^2 - 4K$. Consulting Figure (4.1), we see that if $B^2 - 4K > 0$ then the equilibrium is a stable node. That is, the spring very slowly reaches equilibrium. If $B^2 - 4K < 0$, then the equilibrium is a stable spiral. That is, the mass on the spring oscillates up and down for a bit before damping to a stop. An ideal spring scale will be constructed with B and K such that the spring reaches equilibrium as fast as possible and does not oscillate. This is the critically damped case when $B^2 - 4K = 0$.

$$\begin{aligned} B^2 - 4K &= 0 \\ \left(\frac{b}{m}\right)^2 - 4\frac{k}{m} &= 0. \end{aligned}$$

Since the spring constant k and the damping parameter b remain fixed for any given spring scale, the spring scale will be critically damped for only one optimal mass m_o.

$$\left(\frac{b}{m_o}\right)^2 - 4\frac{k}{m_o} = 0.$$

For a lighter mass m_l, the mass will slowly approach equilibrium since

$$\left(\frac{b}{m_l}\right)^2 - 4\frac{k}{m_l} > 0,$$

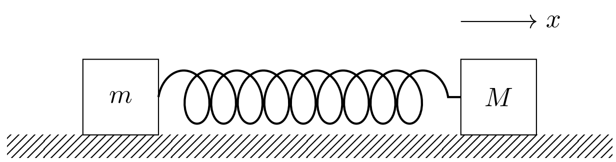
defines a stable node. On the other hand, if the mass being weighed m_h is heavier than the optimal mass, the mass will oscillate about the equilibrium point before settling down since

$$\left(\frac{b}{m_h}\right)^2 - 4\frac{k}{m_h} < 0,$$

defines a stable spiral.

6.2 Double-ended Spring

Consider the spring with a mass on each end shown in the diagram below.



Assume that the surface is rough (i.e. there is friction). How does the mass M move if we grab mass m and move it back and forth with our hand?

We start by finding the equation of motion for the system. To do that, identify the forces acting on mass M. The easiest force to identify is the friction force. If the motion of M is denoted by $x(t)$, then the friction force is opposite the velocity of M, so

$$F_{fric} = -b\dot{x},$$

where b is the damping parameter. The other force acting on M is the spring force. The spring force is given by Hooke's law $F_{sp} = -kd$, where d is the spring's displacement from equilibrium. If positive x is to the right, and the position of m is $x_m(t)$, then the distance between the two masses is $x(t) - x_m(t)$. However, we want to know how much the spring is stretched or compressed from its equilibrium length, which we will denote x_0. If the system is currently in equilibrium, then $x(t) - x_m(t) = x_0$ so $x(t) - x_m(t) - x_0 = 0$, and there is zero spring force acting on mass M. The spring force is then given by

$$F_{sp} = -k[x(t) - x_m(t) - x_0].$$

To confirm that this is the right equation, if we move m left one unit, then the distances between the blocks increases, so $x(t) - x_m(t) > x_0$, so $x(t) - x_m(t) - x_0 > 0$, so $F < 0$ due to the negative sign out front. In other words, the force on M will be to the left, as expected. Similarly, if we push m to the right then the distances between the blocks decreases, so $x(t) - x_m(t) < x_0$, so $x(t) - x_m(t) - x_0 < 0$, so $F > 0$ due to the negative sign out front. In other words, the force on M will be to the right, as expected. So this must be the right equation.

The total force on M is the sum of the spring and friction forces and by Newton's second law:

$$\begin{aligned} F_{tot} &= F_{fric} + F_{sp} \\ M\ddot{x} &= -b\dot{x} - k[x - x_m(t) - x_0]. \end{aligned}$$

Rearranging, we get the equation of motion

$$\ddot{x} + \frac{b}{M}\dot{x} + \frac{k}{M}(x - x_0) = \frac{k}{M}x_m(t).$$

If we redefine our coordinates such that $z = x - x_0$ and assume that mass m moves according to the forcing function $z_m(t) = Ae^{st}$, then our equation of motion becomes

$$\ddot{z} + \frac{b}{M}\dot{z} + \frac{k}{M}z = Ae^{st}.$$

Next, we identify which quantities in the ODE have units

$$\frac{d^2z^*}{dt^{2*}} + \frac{b^*}{M^*}\frac{dz^*}{dt^*} + \frac{k^*}{M^*}z^* = A^*e^{s^*t^*}.$$

Next, we nondimensionalize it with respect to the period of the oscillation, so we make the substitutions

$$z^* = z^*_{ref}z, \qquad t^* = \sqrt{\frac{M^*}{k^*}}t, \qquad A^* = \frac{z^*_{ref}k^*}{M^*}A, \qquad s^* = \sqrt{\frac{k^*}{M^*}}s,$$

and our ODE simplifies to

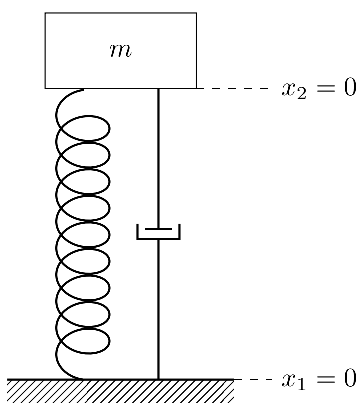
$$\frac{d^2z}{dt^2} + \varepsilon\frac{dz}{dt} + z = Ae^{st},$$

where

$$\varepsilon = \frac{b^*}{\sqrt{M^*k^*}}.$$

6.3 Spring Suspensions

Vehicle suspensions and seismic isolators in building foundations used to protect buildings during earthquakes can be modeled as a mass on a damped spring.



The goal of a suspension system is to minimize motion of the mass, whether that is a car or a building. In other words, for a given forcing function, which in the case of a car could be bumps in the road or in the case of a building could be an earthquake, we want the response of the system to be as close to zero as possible. For example, we want the building to remain still even while the ground is shaking, and we want the car to remain still even as it goes over bumps in the road.

In the diagram above, we have coordinate systems set such that the position of the mass is at $x_2 = 0$ and the position of the ground is at $x_1 = 0$ when the spring is in equilibrium such that the net spring force is zero. Since the gravitational force is accounted for by the equilibrium position of the spring, we can ignore it from now on. The two forces acting on the mass are the spring force and the damping force. The spring force is given by

$$F_s = -k(x_2 - x_1).$$

The damping force is given by

$$F_d = -b(\dot{x}_2 - \dot{x}_1).$$

Like any damping or friction force, it is proportional to the relative (vertical) speed of the mass and the ground. To make sure we got the sign right, consider the earthquake case. If the earthquake is over then $\dot{x}_1 = 0$. If the building is moving upward, then $\dot{x}_2 > 0$, which means the damping force is pointing down, hence the negative sign out front.

We want to know x_2 as a function of time given x_1 as a function of time, so

$$\begin{aligned} F_{net} &= F_s + F_d \\ m\ddot{x}_2 &= -k(x_2 - x_1) - b(\dot{x}_2 - \dot{x}_1). \end{aligned}$$

We are interested in minimizing the amplitude of the response, and to do that, we need to know the magnitude of the transfer function. If we separate the equation of motion so that the response is on the left and the input (i.e. forcing function) is on the right, we get

$$m\ddot{x}_2 + kx_2 + b\dot{x}_2 = kx_1 + \dot{x}_1.$$

If we let the initial conditions be zero since we don't care about them, then taking the Laplace transform of both sides gives us

$$\begin{aligned} \mathscr{L}\left\{m\ddot{x}_2 + kx_2 + b\dot{x}_2\right\} &= \mathscr{L}\left\{kx_1 + \dot{x}_1\right\} \\ m\mathscr{L}\left\{\ddot{x}_2\right\} + k\mathscr{L}\left\{x_2\right\} + b\mathscr{L}\left\{\dot{x}_2\right\} &= k\mathscr{L}\left\{x_1\right\} + \mathscr{L}\left\{\dot{x}_1\right\} \\ (ms^2 + k + bs)\mathscr{L}\left\{x_2\right\} &= (k + bs)\mathscr{L}\left\{x_1\right\}. \end{aligned}$$

Recall that

$$\text{transfer function} = \frac{\mathscr{L}\left\{\text{output}\right\}}{\mathscr{L}\left\{\text{input}\right\}},$$

so

$$G(s) = \frac{k + bs}{ms^2 + k + bs}.$$

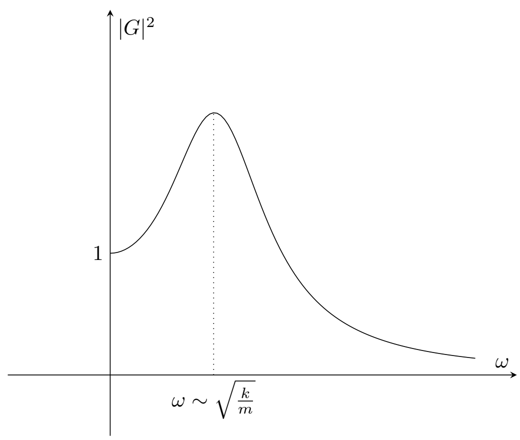
The value of s depends on the forcing function, that is, the earthquake or the bumps in the road. If we assume the forcing function is sinusoidal (e.g. evenly spaced, smooth bumps in the road or an earthquake such that the ground moves up and down in a sinusoidal manner), then $s = i\omega$, and

$$G(i\omega) = \frac{k + bi\omega}{-m\omega^2 + k + bi\omega}.$$

Multiplying G by its complex conjugate gives us the amplitude squared of the transfer function

$$|G|^2 = \frac{k^2 + \omega^2 b^2}{(k - m\omega^2)^2 + \omega^2 b^2}.$$

In order to minimize the response of the system, we want this value to be as close to zero as possible. If we graph $|G|^2$ versus the forcing frequency ω, we get a picture like the following.

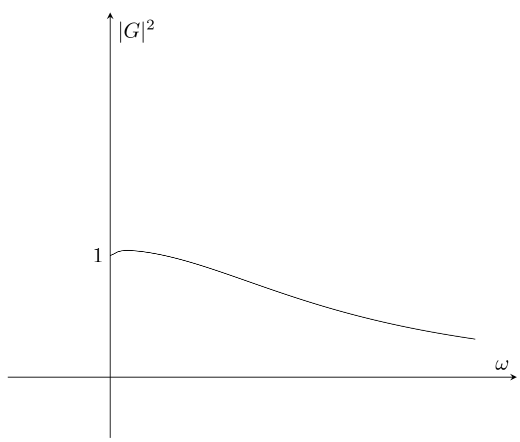


At the far left of the graph, the value of the transfer function is 1, which means the magnitude of the repose of the system is the same as the magnitude of the input. Near the natural frequency of the system, there is a resonance peak where the repose is much greater than the input. This is not at all what we want. We don't want a car to fly into the air when it encounters small bumps. To the right, the transfer function approaches zero, which is the regime in which the response of the system is much less than the input. This is where we want our suspension systems to be.

We may not be able to change the mass m or the forcing frequency ω (since that is determined by the earthquake or the speed at which bumps are encountered on the road), but we can adjust the spring constant k (by using a different spring) and the damping parameter b in order to get a more favorable graph.

By selecting k and b judiciously, we get a graph where the resonance peak is shifted far to the left and is essentially 1. So virtually any frequency results in a response that is less than the input.

For earthquakes, for example, there is some minimum frequency ω_{min} that causes damage. When designing a seismic isolation system, we would choose k and b such that on $[\omega_{min}, \infty)$, G is as small as possible. That is, we want G to be as small as possible over the range of shaking frequencies which can be caused by earthquakes and which cause damage to structures.



Chapter 7

Heat Transfer

Heat behaves like a fluid. It flows from high temperature regions to low temperature regions. We measure temperature and from that, we infer the heat. The amount of heat inferred from a temperature change ΔT in a mass m is proportional to both. That is,

$$\Delta Q = cm\Delta T,$$

where ΔQ is the heat change and c is the heat capacity of the material. We will be measure heat, Q, in calories. Heat flux is

$$F = \frac{\text{heat}}{\text{area} \cdot \text{time}}.$$

Fourier's law tells us that

$$F = -k\frac{\Delta T}{\Delta x},$$

where k is the **thermal conductivity** of the material. The heat flow per unit time through a surface with area A is then

$$AF = -Ak\frac{\Delta T}{\Delta x}.$$

Consider a house wall in the winter. The thickness of the wall is L, the outside temperature is T_o, and the inside temperature is T_i, so temperature difference divided by the thickness is

$$\frac{\Delta T}{\Delta x} = \frac{T_i - T_o}{L}.$$

Then the *heat loss* through the wall is

$$H = Ak\frac{T_i - T_o}{L},$$

or

$$H = AU(T_i - T_o),$$

where $U = k/l$ is the **coefficient of transmission**. Note that the coefficient of transmission depends on the properties of the material and the thickness of the material. The quantity $R = 1/U$ is the **thermal resistance** commonly called **R-value** when referring to insulation. The larger the thermal resistance, the lower the heat transfer for a given temperature difference. Windows are often noted with U-values. Lower is better.

Using conservation of energy, we can obtain the one-dimensional (e.g. in a wire) partial differential heat equation

$$\frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\left(\frac{k}{\rho c}\frac{\partial T}{\partial x}\right),$$

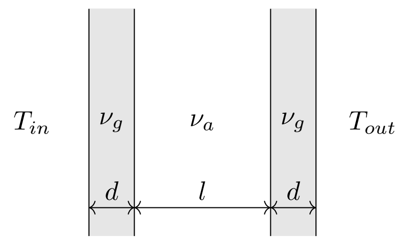
where ρ is the density of the material. If $k/(\rho c)$ is constant, the heat equation is typically written as

$$\frac{\partial T}{\partial t} = \nu \frac{\partial^2 T}{\partial x^2},$$

where $\nu = k/(\rho c)$ is the **thermal diffusivity**.

7.1 Double-pane Windows

Consider a double-paned glass window. The thermal diffusivity of the glass is ν_g, and the thermal diffusivity of the gas between the panes is ν_a. The inside temperature is T_{in} and the outside temperature is T_{out}. Develop a mathematical model for the heat transfer through the window.



The heat equation is

$$\frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\left(\nu \frac{\partial T}{\partial x}\right).$$

The thermal diffusivity of the window is not constant since there are different values for the glass panes and the gas between them, however, ν is piecewise constant.

The timescale associated with the outside temperature is about 1 hour. That is about the time it takes the outside temperature to change noticeably. We start by nondimensionalizing the heat equation to find what if any terms can be neglected. All of the quantities in the heat equation have dimensions

$$\frac{\partial T^*}{\partial t^*} = \frac{\partial}{\partial x^*}\left(\nu^* \frac{\partial T^*}{\partial x^*}\right).$$

which are

$$[T^*] = \text{Deg}, \qquad [t^*] = \mathscr{T}, \qquad [x^*] = \mathscr{L}, \qquad [\nu^*] = \mathscr{L}^2/\mathscr{T}.$$

We nondimensionalize it with respect to some reference temperature and time and d, the thickness of the glass

$$T = \frac{T^*}{T_{ref}}, \qquad t = \frac{t^*}{t_{ref}}, \qquad x = \frac{x^*}{d}.$$

Then our heat equation becomes

$$\frac{T_{ref}}{t_{ref}} \frac{\partial T}{\partial t} = \nu^* \frac{T_{ref}}{d^2} \frac{\partial^2 T}{\partial x^2}$$

$$\frac{\frac{d^2}{\nu^*}}{t_{ref}} \frac{\partial T}{\partial t} = \frac{\partial^2 T}{\partial x^2}.$$

Notice that the quantity

$$\frac{\frac{d^2}{\nu^*}}{t_{ref}} = \varepsilon$$

is the ratio of two timescales. If we let t_{ref} be the timescale over which the outside air changes, then ε gives the ratio of time it takes heat to go through the glass divided by the time it takes the outside temperature to change. For a glass window, $\nu \approx 3 \times 10^{-7}\,\mathrm{m}^2/\mathrm{s}$ and $d \approx 3 \times 10^{-3}\,\mathrm{m}$, and we know that $t_{ref} \approx 3600\mathrm{s}$, so the heat equation becomes

$$0.01 \cdot \frac{\partial T}{\partial t} = \frac{\partial^2 T}{\partial x^2}.$$

Since ε is very small, we can neglect $\frac{\partial T}{\partial t}$, and the heat equation simplifies to

$$\frac{\partial}{\partial x}\left(\nu \frac{\partial T}{\partial x}\right) = 0.$$

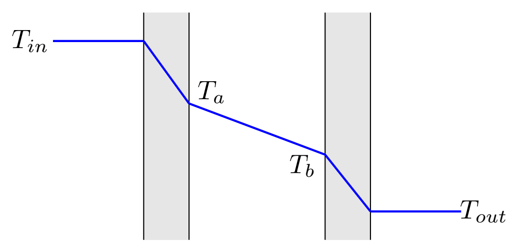
The quantity in parentheses is the flux F (with maybe a sign difference). The form of the differential equation implies that the quantity inside parentheses is constant. That is,

$$F = \nu \frac{\partial T}{\partial x} = \text{constant}.$$

This implies that the temperature T is a linear function with slope

$$\frac{\partial T}{\partial x} = \frac{F}{\nu}.$$

This means we can draw the solution without knowing the details. We know that the temperature as a function of position inside the glass is linear with slope F/ν_g and between the glass panes it is linear with slope F/ν_a.



By conservation of energy, and since we are at a steady state, we know that the heat flux through the first pane is the same as the heat flux through the air gap, which is the same as the heat flux through the second pane. From Fourier's law, then

$$F = \nu_g \frac{T_{in} - T_a}{d} = \nu_a \frac{T_a - T_b}{l} = \nu_g \frac{T_b - T_{out}}{d}.$$

Let $\beta = \nu_g l/(\nu_a d)$, then we get the system of equations

$$\begin{aligned} T_{in} - T_a &= T_b - T_{out} \\ \beta(T_{in} - T_a) &= T_a - T_b. \end{aligned}$$

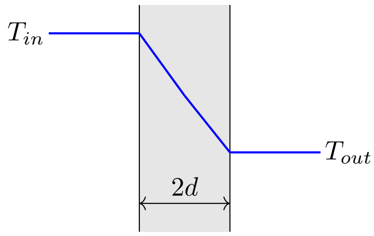
The inside and outside temperatures T_{in} and T_{out} are given, and solving the system of equations for T_a and T_b gives us

$$\begin{aligned} T_a &= \frac{T_{in}(1+\beta) + T_{out}}{2+\beta} \\ T_b &= \frac{T_{in} + (1+\beta)T_{out}}{2+\beta}. \end{aligned}$$

From Fourier's law we get

$$F = \frac{\nu_g}{d(2+\beta)}(T_{in} - T_{out}).$$

Compare this to the heat flux through a single-pane double thickness window. This means there is the same amount of glass for the heat to travel through as the double-pane windows but no air gap. That is, we want to know what the effectiveness of the air gap is.



Calculating the heat flux through the doubly thick single glass, we get

$$F_s = \nu_g \frac{T_{in} - T_{out}}{2d}.$$

Denoting the flux through the double pane window with F_d and the flux through the single pane window with F_s, the relative efficiency of the double pane window is

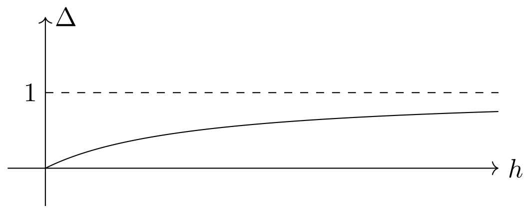
$$\Delta = \frac{F_s - F_d}{F_s} = \frac{\beta}{2 + \beta}.$$

Recall that $\beta = \nu_g l/(\nu_a d)$. If we let $R = \nu_g/\nu_a$ which is a function of the material and $h = l/d$ which is a function of how the materials are put together, then we can write the relative efficiency as

$$\Delta = \frac{hR}{2 + hR}.$$

A value of $\Delta = 0$ means the single pane window is as efficient as the double-pane window. This only occurs when the air gap in the double pane window is zero, so it is actually the same as a single pane, double thickness window. A value of $\Delta = 0.1$ means the heat flux through the double-pane window is 0.9 of the flux through the single pane window. A value of $\Delta = 0.5$ means the flux through the double-pane window is only half the flux through the single pane window.

Plotting Δ versus h with R fixed gives us the graph



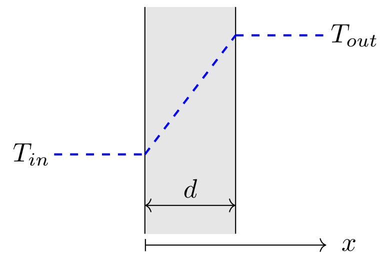
Notice that the graph tells us that there are diminishing returns when increasing the air gap between the windows. The wider the air gap, the more efficient, but to reach 99% efficiency, the gap would have to be very wide. This model is not valid for very large h because with a large gap between the panes, convection becomes significant. Our heat transfer model treats only conduction.

We can also increase the relative efficiency of the double-pane window by changing R, which can be done by using a different gas in the air gap. In practice, argon is often used to fill the gap since it has a low k, it is a safe noble gas, and its atoms are large so it's not as likely to leak away. A vacuum between the panes would be even better, but it's hard to maintain a vacuum, and the panes would have to be very thick.

7.2 Insulated Wall

Consider an 8-inch-thick insulated wall with

$$\nu = 0.01\,\mathrm{ft}^2/\mathrm{hr}.$$



The dashed lines in the diagram represent the mean temperature profile in each segment. The actual temperature profiles vary sinusoidally. The timescale of T_{out} is approximately 12 hours as it varies throughout the day. The temperature inside the wall will also vary sinusoidally, but to a lesser extent. If the insulated wall is perfect, T_{in} will vary much less than the outside temperature.

The timescale for the heat to move through the wall is

$$\frac{d^2}{\nu} = \frac{(8\,\text{in})^2}{0.01\,\text{ft}^2/\text{hr}} \approx 44\,\text{hr}.$$

This timescale is not negligible, so we must account for it in the heat equation. That is, we cannot set $\frac{\partial T}{\partial t}$ equal to zero.

Our heat equation is

$$T_t = \nu T_{xx}, \qquad 0 \leq x \leq d,$$

for the heat movement through the wall. The boundary conditions are

$$T(0,t) = T_{in}, \qquad T(d,t) = T_{out}.$$

To recast the heat equation in a more convenient form, we write

$$T_{out} = \bar{T}_{out} + \Delta T = T_{in} + \left(\bar{T}_{out} - T_{in}\right) + \Delta T,$$

where $\bar{T}_{out}$ is the average value of T_{out} and ΔT is the sinusoidally varying part of T_{out}. Next, we make the change of variables

$$\hat{T}(x,t) = T(x,t) - \left(T_{in} + \left(\bar{T}_{out} - T_{in}\right)\frac{x}{d}\right).$$

Now, our PDE is

$$\hat{T}_t = \nu \hat{T}_{xx},$$

and the boundary conditions become

$$\begin{aligned}
\hat{T}(0,t) &= T(0,t) - T_{in} \\
&= 0 \\
\hat{T}(d,t) &= T(d,t) - \left(T_{in} + \left(\bar{T}_{out} - T_{in}\right)\frac{d}{d}\right) \\
&= T_{out} - T_{in} - \left(\bar{T}_{out} - T_{in}\right) \\
&= T_{out} - \bar{T}_{out} \\
&= (\bar{T}_{out} + \Delta T) - \bar{T}_{out} \\
&= \Delta T.
\end{aligned}$$

Recall that T_{out} varies sinusoidally. What we have done with this change of variables is subtract away the mean value of T_{out} so that only the sinusoidally varying part remains. If we let

$$\begin{aligned}
\Delta T &= A e^{i\omega t} \\
\hat{T}(x,t) &= u(x) e^{i\omega t},
\end{aligned}$$

and substitute these into the PDE and the boundary conditions, we get the ordinary differential equation

$$\begin{aligned} i\omega u - \nu u_{xx} &= 0 \\ u(0) &= 0 \\ u(d) &= A. \end{aligned}$$

Dividing the differential equation by $-\nu$ gives us

$$u_{xx} - \frac{i\omega}{\nu} u = 0.$$

The Fourier transform converts PDEs to ODEs. Similarly, the Laplace transform converts ODEs to algebraic equations. Using the Laplace transform, gives us the characteristic equation

$$s^2 - \frac{i\omega}{\nu} = 0,$$

which tells us that

$$s = \pm\sqrt{i}\sqrt{\frac{\omega}{\nu}}.$$

After computing the square root of i, we get

$$s = \pm\frac{1+i}{\sqrt{2}}\sqrt{\frac{\omega}{\nu}},$$

so our general solution for $u(x)$ is

$$u(x) = C_1 e^{\sqrt{\frac{\omega}{2\nu}}(1+i)x} + C_2 e^{-\sqrt{\frac{\omega}{2\nu}}(1+i)x}.$$

From the boundary conditions, we get that $C_2 = -C_1$ and $u(d) = A$, so if we let

$$\alpha = \sqrt{\frac{\omega}{2\nu}},$$

our specific solution for $u(x)$ is

$$u(x) = \frac{A}{e^{\alpha(1+i)d} - e^{-\alpha(1+i)d}} \left(e^{\alpha(1+i)x} - e^{-\alpha(1+i)x} \right).$$

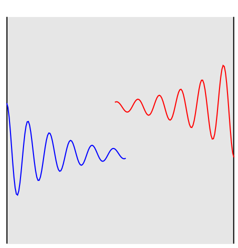
The first part is just a number, so to write it more compactly, we let

$$\beta = \frac{A}{e^{\alpha(1+i)d} - e^{-\alpha(1+i)d}}.$$

Recall that $\hat{T}(x,t) = u(x)e^{i\omega t}$. Plugging in $u(x)$, we get that the solution to the PDE is

$$\hat{T}(x,t) = \beta e^{i\omega t} \left(e^{\alpha(1+i)x} - e^{-\alpha(1+i)x} \right) = \beta \left(e^{\alpha x} e^{\alpha i \left(x + \frac{\omega}{\alpha} t\right)} - e^{-\alpha x} e^{-\alpha i \left(x - \frac{\omega}{\alpha} t\right)} \right).$$

This solution describes a wave of "heat" traveling left into the wall from the outside, and a wave of "cool" traveling right into the wall from the inside.



What is the heat flux at the inside of the wall? The flux is given by

$$\begin{aligned}\hat{F}(0,t) &= -k\frac{\partial \hat{T}}{\partial x}(0,t)\\ &= -2k\alpha\beta(1+i)e^{i\omega t}\\ &= -2k\sqrt{\frac{\omega}{2\nu}}\left(\frac{A}{e^{\alpha(1+i)d}-e^{-\alpha(1+i)d}}\right)(1+i)e^{i\omega t}\\ &= -2kA\sqrt{\frac{\omega}{2\nu}}\left(\frac{1+i}{e^{\alpha(1+i)d}-e^{-\alpha(1+i)d}}\right)e^{i\omega t}.\end{aligned}$$

Notice that the quantity in parentheses is a complex number. We can write this complex number in the form γe^{ψ} where γ is some real number and ψ is the phase shift. We can therefore, write the flux in the form

$$\hat{F}(0,t) = -2kA\sqrt{\frac{\omega}{2\nu}}\gamma e^{i(\omega t+\psi)},$$

If we let $e^{\alpha(1+i)d} - e^{-\alpha(1+i)d} = \mu + i\eta$, where μ and η are real numbers, then

$$\frac{1+i}{e^{\alpha(1+i)d}-e^{-\alpha(1+i)d}} = \frac{1+i}{\mu+i\eta} = \frac{\mu+\eta}{\mu^2+\eta^2} + i\frac{\mu-\eta}{\mu^2+\eta^2}.$$

To find the phase shift, we use

$$\psi = \tan^{-1}\left(\frac{\frac{\mu-\eta}{\mu^2+\eta^2}}{\frac{\mu+\eta}{\mu^2+\eta^2}}\right) = \tan^{-1}\left(\frac{\mu-\eta}{\mu+\eta}\right).$$

Recall that the outside temperature varies sinusoidally as

$$\Delta T = Ae^{i\omega t}.$$

We are now in position to calculate how long it takes the heat to travel through the wall by looking at the phase shift ψ in the flux through the inside of the wall. We have that $\nu = 0.01\,\text{ft}^2/\text{hr}$. The period of the sinusoidal variation of the outside temperature is

$$\frac{2\pi}{\omega} = 24\,\text{hr},$$

which gives us

$$\omega = 0.2618\,\text{hr}^{-1},$$

and so

$$\alpha = \sqrt{\frac{\omega}{2\nu}} = 3.6\,\text{ft}^{-1}.$$

Since $d = 8$ inches, this gives us

$$\alpha d = 2.4.$$

$$\mu + i\eta = e^{\alpha(1+i)d} - e^{-\alpha(1+i)d} = e^{2.4}e^{2.4i} - e^{-2.4}e^{-2.4i} = -8.06 + 7.51i,$$

and so

$$\mu = -8.06, \qquad \eta = 7.51.$$

So the phase difference between the outside temperature and the temperature of the inside of the wall is

$$\psi = \tan^{-1}\left(\frac{-8.06-7.51}{-8.06+7.51}\right),$$

which corresponds to the angle 268°. Since the period is 24 hours, the phase shift is 17.9 hours. So it effectively takes 17.9 hours for the outside heat to pass through the wall.

7.3 Wine Cellar

How deep does an underground wine cellar have to be placed in order for the temperature to remain nearly constant?

The driving force in this case is the sun heating the surface of the ground above the wine cellar. There are two timescales that we have to concern ourselves with. First, there is the yearly timescale. The surface temperature varies dramatically with the seasons. Second, there is the daily timescale. Near the middle of the day, the surface is often much cooler than at night. At the level of the wine cellar, the temperature oscillates with this two frequencies, the yearly variation being a much smaller frequency than the daily variation.

We will only consider the yearly timescale at first. This gives us the heat equation

$$T_t = \nu T_{xx},$$

for $0 < x < \infty$ since the depth of the ground is effectively infinity. We let x be the direction pointing down into the ground with $x = 0$ corresponding to the surface. The boundary conditions are

$$T(0,t) = Ae^{i\omega t}, \qquad |T| < M, \text{for } 0 < x < \infty.$$

That is, the temperature at the surface varies sinusoidally throughout the year and the temperature underground at any depth is bounded.

We assume a solution of the form

$$T = u(x)e^{i\omega t},$$

where $u(x)$ is the amplitude which varies with x—the depth underground. Substituting it into the PDE gives us an ordinary differential equation with boundary conditions

$$\begin{aligned} i\omega u - \nu u_{xx} &= 0 \\ u(0) &= A \\ u(\infty) &< \infty. \end{aligned}$$

The last boundary condition is just the statement that u is bounded for all values of x.

This differential equation of $u(x)$ has the general solution

$$u(x) = C_1 e^{\alpha(1+i)x} + C_1 e^{-\alpha(1+i)x},$$

where

$$\alpha = \sqrt{\frac{\omega}{2\nu}}.$$

The only way that $u(x)$ is bounded for all x, is if the first term is not there. This implies that $C_1 = 0$. The other boundary condition implies $C_2 = A$, so our solution is

$$u(x) = Ae^{-\alpha x}e^{-i\alpha x}.$$

Plugging this solution for $u(x)$ back into our solution T gives us

$$T(x,t) = Ae^{-\alpha x}e^{i(\omega t - \alpha x)}.$$

The imaginary unit in the exponential tells us that the temperature propagates into the ground as a sinusoidal wave. The $e^{-\alpha x}$ out front tells us that the amplitude of the wave is decreasing exponentially as it travels into the ground.

We want to place the wine cellar at a depth x such that the summer heat doesn't reach the cellar until about the time it starts getting cold up on the surface again. At that point, the door can be opened, quickly cooling the wine cellar down to the surface temperature. Notice the phase difference in $e^{i(\omega t - \alpha x)}$. We want to choose x such that the phase difference between the surface and the cellar is $\alpha x = \pi$, that is, a phase difference of half of a year.

For dry Earth,

$$\nu \approx 2 \times 10^{-3}\,\mathrm{cm}^2/\mathrm{s}.$$

The period of the sinusoidal temperature wave is one year or

$$\frac{2\pi}{\omega} = 3.15 \times 10^7\,\mathrm{s},$$

which gives us

$$\omega = 2 \times 10^{-7}\,1/\mathrm{s},$$

and

$$\alpha = 7 \times 10^{-3}\,1/\mathrm{cm}.$$

So the optimal depth of the wine cellar is

$$x = \frac{\pi}{\alpha} = 450\,\mathrm{cm} \approx 15\,\mathrm{ft}.$$

If the yearly temperature variation in the geographic location of the wine cellar is $A = 80^\circ$, then the temperature variation at 15 feet below the surface is

$$Ae^{-\pi} = 0.0432A \approx 3.5^\circ.$$

Now we need to look at the variation caused by the daily fluctuations in temperature. We did the yearly timescale, and now we need to do the daily timescale. The period is

$$\frac{2\pi}{\omega} = 24\,\mathrm{hr}\,,$$

which gives us

$$\omega = 7.27 \times 10^{-4}\,1/\mathrm{s},$$

and

$$\alpha = 4 \times 10^{-1}\,1\mathrm{cm},$$

so

$$\alpha x = 180.$$

So the temperature variation 15 feet below the ground due to the daily fluctuations at the surface is

$$Ae^{-180} \approx 0.$$

This tells us that the daily fluctuations in temperature can be neglected.

Chapter 8

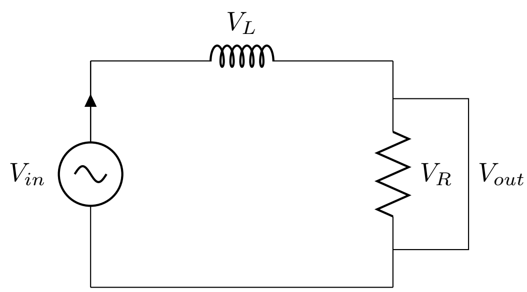
Electronics

Circuit diagrams are the simple conceptual models of electronics.

Review circuits from general physics.

8.1 Low-pass Filter

A low-pass filter is a piece of electronic hardware that filters out the high frequency, leaving only the lower frequencies.



Since the inductor and the resistor are in series, we have that

$$V_{in} = V_L + V_R.$$

From the inductor equation $V_L = L\frac{dI}{dt}$ and Ohm's law $V_R = IR$, we get

$$V_{in} = L\frac{dI}{dt} + IR,$$

which we can write as

$$V_{in}(t) = LI'(t) + RI(t).$$

This is a driven circuit, and we want to find the **voltage gain**

$$\frac{|V_{out}|}{|V_{in}|}.$$

Note that $V_{out} = V_R$. It is just the voltage difference measured across the resistor by a voltmeter.

Taking the Laplace transform of $V_{in}(t)$ gives us

$$\begin{aligned}\mathscr{L}\{V_{in}\} &= L\mathscr{L}\{I'\} + R\mathscr{L}\{I\} \\ &= L(-I(0) + s\mathscr{L}\{I\}) + R\mathscr{L}\{I\} \\ &= -LI(0) + (Ls + R)\mathscr{L}\{I\}.\end{aligned}$$

We assume because of the resistor, the transients have died out and $I(0) = 0$, so we get

$$\mathscr{L}\left\{V_{in}\right\} = (Ls + R)\mathscr{L}\left\{I\right\},$$

or

$$\mathscr{L}\left\{I\right\} = \frac{1}{Ls + R}\mathscr{L}\left\{V_{in}\right\}.$$

Taking the Laplace transform of Ohm's law gives us

$$\mathscr{L}\left\{I\right\} = \frac{1}{R}\mathscr{L}\left\{V_{out}\right\}.$$

Substituting this in gives us

$$\mathscr{L}\left\{V_{out}\right\} = \frac{R}{Ls + R}\mathscr{L}\left\{V_{in}\right\}.$$

This gives us our transfer function

$$G(s) = \frac{R}{Ls + R},$$

so the voltage gain is

$$\frac{|V_{out}|}{|V_{in}|} = |G(s)|.$$

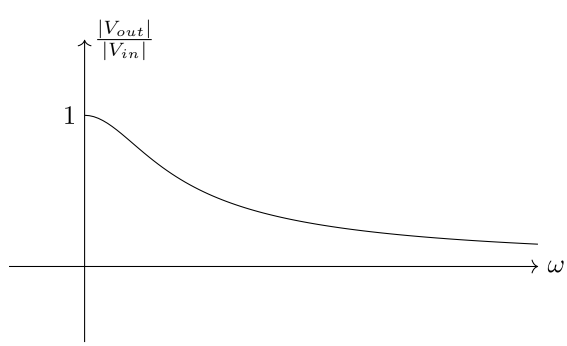
Assuming a sinusoidal output, then $s = i\omega$ and taking the magnitude of the complex quantity gives us

$$\frac{|V_{out}|}{|V_{in}|} = \sqrt{\frac{R^2}{R^2 + L^2\omega^2}} = \frac{1}{\sqrt{1 + \omega^2\left(\frac{L}{R}\right)^2}}.$$

The critical frequency is

$$\omega_c = \pm\frac{R}{L}.$$

Plotting the voltage gain as a function of ω gives us the following graph.

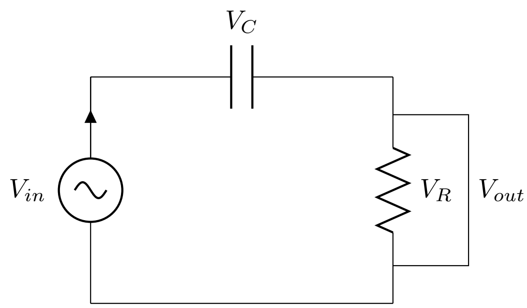


Since the system responds much more strongly to low input frequencies and gives essentially zero response for high input frequencies, this is a low-pass filter. In particular, notice that the output voltage is equal to the input voltage, that is, $|V_{out}|/|V_{in}| \approx 1$ only at very low input frequencies.

It makes sense that this RL circuit is a low-pass filter since an inductor resists changes in the current. At low frequencies (i.e. barely changing) the inductor has little effect, but at high frequencies (i.e. changing directions often) the inductor has a large effect.

8.2 High-pass Filter

A high-pass filter is an electronic devices that filters out the low frequencies and allows the high frequencies to pass through.



Since the components are in series, we have that

$$V_{in} = V_{out} + V_C.$$

The quantity V_{out} is just the voltage difference measured across the resistor, and by Ohm's law, that is $V_{out} = V_R = IR$. The voltage drop across the capacitor is

$$V_C = \int \frac{I}{C}\, dt,$$

so our voltage equation becomes

$$V_{in} = IR + \int \frac{I}{C}\, dt.$$

Differentiating with respect to time gives us

$$\frac{dV_{in}}{dt} = R\frac{dI}{dt} + \frac{I}{C}.$$

Taking the Laplace transform of both sides gives us

$$s\mathscr{L}\{V_{in}\} = sR\mathscr{L}\{I\} + \frac{1}{C}\mathscr{L}\{I\},$$

or

$$\mathscr{L}\{I\} = \frac{sC}{sRC+1}\mathscr{L}\{V_{in}\}.$$

From Ohm's law, we get

$$\mathscr{L}\{I\} = \frac{1}{R}\mathscr{L}\{V_{out}\}.$$

Substituting this in gives us

$$\mathscr{L}\{V_{out}\} = \frac{sRC}{sRC+1}\mathscr{L}\{V_{in}\}.$$

So our voltage gain is

$$\frac{|\mathscr{L}\{V_{out}\}|}{|\mathscr{L}\{V_{in}\}|} = |G(s)|.$$

where

$$G(s) = \frac{sRC}{sRC+1} = \frac{s}{s + \frac{1}{RC}}.$$

Assuming that $s = i\omega$, we get

$$|G(i\omega)| = \frac{\omega}{\sqrt{\omega^2 + \left(\frac{1}{RC}\right)^2}}.$$

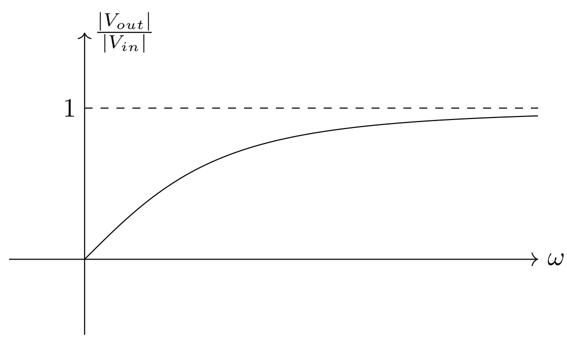
So our voltage gain is

$$\frac{|\mathscr{L}\{V_{out}\}|}{|\mathscr{L}\{V_{in}\}|} = \frac{\omega}{\sqrt{\omega^2 + \left(\frac{1}{RC}\right)^2}},$$

and the critical frequency is

$$\omega_c = \pm\frac{1}{RC}.$$

Plotting the voltage gain as a function of ω gives us the following graph.

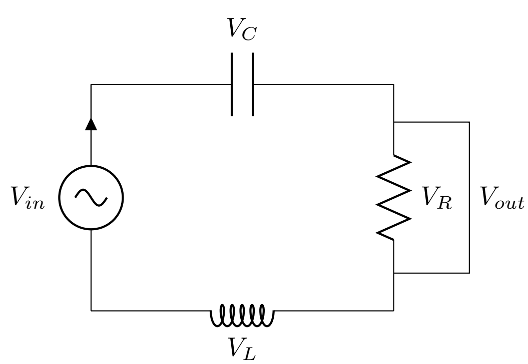


Since the system responds much more strongly to high input frequencies and gives essentially zero response for frequencies near zero, this is a high-pass filter.

It makes sense that this RC circuit is a high pass filter since the current in the circuit drops as the capacitor charges. When the capacitor is fully charged, the circuit is broken and there is no current at all. So the capacitor is more resistant to low frequencies than to the high frequencies.

8.3 Bandpass Filter

A band pass filter is an electronic device that passes a specific band of frequencies and filters out the lower and higher frequencies.



Since the elements are in series, we start with the voltage equation

$$V_{in} = V_C + V_{out} + V_L.$$

The voltage drop across the capacitor is

$$V_C = \int \frac{I}{C}\,dt.$$

The quantity V_{out} is just the voltage drop across the resistor, so by Ohm's law,

$$V_{out} = V_R = IR.$$

The voltage drop across the inductor is

$$V_L = L\frac{dI}{dt}.$$

So our voltage equation becomes

$$V_{in} = \int \frac{I}{C}\,dt + IR + L\frac{dI}{dt}.$$

Differentiating with respect to time gives us

$$\frac{dV_{in}}{dt} = \frac{I}{C} + R\frac{dI}{dt} + L\frac{d^2 I}{dt^2}.$$

Taking the Laplace transform gives us

$$s\mathscr{L}\left\{V_{in}\right\} = \frac{1}{C}\mathscr{L}\left\{I\right\} + Rs\mathscr{L}\left\{I\right\} + Ls^2\mathscr{L}\left\{I\right\} = \left(\frac{1}{C} + Rs + Ls^2\right)\mathscr{L}\left\{I\right\}.$$

From Ohm's law, we get

$$\mathscr{L}\left\{I\right\} = \frac{1}{R}\mathscr{L}\left\{V_{out}\right\}.$$

Substituting this in gives us

$$\mathscr{L}\left\{V_{in}\right\} = \left(\frac{1}{RCs} + 1 + \frac{Ls}{R}\right)\mathscr{L}\left\{V_{out}\right\}.$$

So our voltage gain is

$$\frac{|\mathscr{L}\left\{V_{out}\right\}|}{|\mathscr{L}\left\{V_{in}\right\}|} = |G(s)|.$$

where

$$G(s) = \frac{RCs}{CLs^2 + RCs + 1}.$$

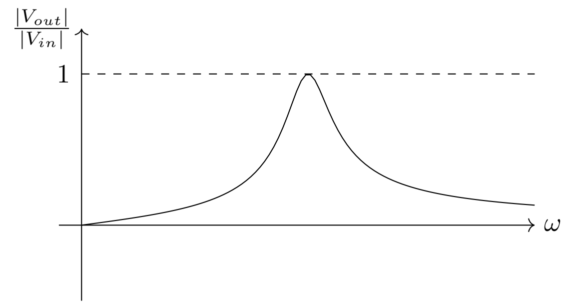
Assuming that $s = i\omega$, we get

$$|G(i\omega)| = \frac{RC\omega}{\sqrt{(1 - LC\omega^2)^2 + (RC\omega)^2}}.$$

So our voltage gain is

$$\frac{|\mathscr{L}\left\{V_{out}\right\}|}{|\mathscr{L}\left\{V_{in}\right\}|} = \frac{RC\omega}{\sqrt{(1 - LC\omega^2)^2 + (RC\omega)^2}}.$$

Plotting the voltage gain as a function of ω with the parameters $R = 1\,\Omega$, $C = 0.04\,\text{F}$, $L = 1\,\text{H}$, gives us the following graph:



Since the system responds much more strongly to a specific range of input frequencies and gives less response for lower frequency input and essentially zero response for lower and higher frequencies, this is a band-pass filter.

The location of the resonance peak is at

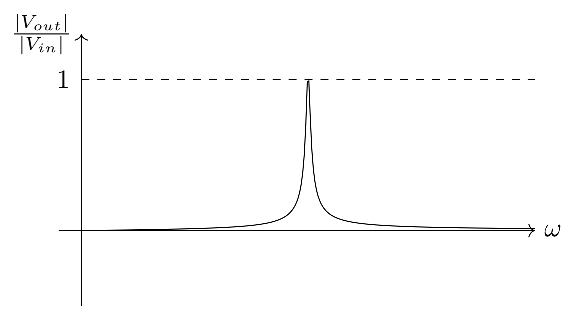
$$\omega_R = \frac{1}{\sqrt{LC}},$$

and can be found by noting that $|G|$ is maximized when $(1 - LC\omega^2)^2 = 0$.

The width of the resonance peak can be characterized by finding the full width at half maximum or FWHM. Since the maximum is 1, the FWHM is found by setting $|G|$ equal to $\frac{1}{2}$ and obtaining the two positive solutions, then taking their difference to get the width

$$\text{FWHM} = \sqrt{3}\frac{R}{L}.$$

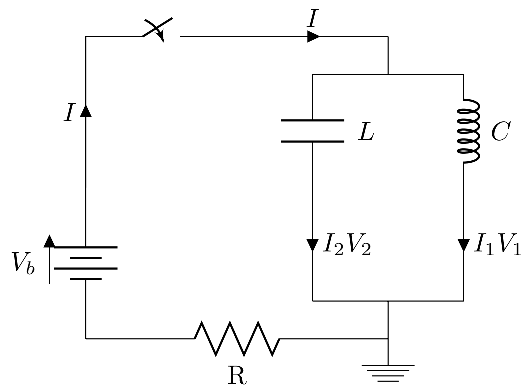
For example, by decreasing R in the circuit to $R = 0.1\,\Omega$, we obtain a much narrower resonance peak, which means our filter is much more effective at selecting a narrow range of frequencies to pass through.



8.4 Electric Clock

Of particular interest in the circuit diagram here is the LC loop, which is also called a **tank circuit** or **resonant circuit**. We start by closing the switch to charge the capacitor and get things moving in the tank circuit. Then we open the switch, but because of the capacitor and inductor in the tank circuit, the current in that part of the circuit oscillates even though nothing is happening in the rest of the circuit.

Recall that a charged capacitor will have a bunch of charges on one plate and a bunch of opposite charges on the other plate. This charge imbalance creates a voltage difference causing current to flow. The inductor opposes changes in current flow by storing energy in a magnetic field. So as the capacitor discharges, the inductor charges. Once the capacitor can not discharge any further, the inductor will continue to push current through the circuit since it resists changes in current flow. This results in charging the capacitor with the opposite polarity than what it began with. Then the process repeats with the current flowing in the opposite direction. The current in the tank circuit will continue to oscillate back and forth.



When we open the switch, the voltage across the capacitor and inductor will be the same and will be equal to the battery voltage

$$V_b = V_1 = V_2.$$

The current is given by

$$I = I_1 + I_2.$$

When we open the switch, however, I becomes 0, so

$$I_1 = -I_2.$$

The voltage across the capacitor is given by

$$V_1 = L\frac{dI_1}{dt}.$$

From the inductor equation, we have that

$$C\frac{dV_2}{dt} = I_2.$$

Since $V_1 = V_2$, we know that

$$\frac{dV_1}{dt} = \frac{dV_2}{dt}.$$

Substituting in known values, we get

$$\frac{d}{dt}\left(L\frac{dI_1}{dt}\right) = \frac{I_2}{C}.$$

Plugging in $I_1 = -I_2$ and simplifying, we get the differential equation

$$\frac{d^2I_1}{dt^2} + \frac{1}{LC}I_1 = 0.$$

The characteristic equation for this ODE is

$$s^2 + \frac{1}{LC} = 0.$$

This gives us the natural frequencies

$$s = \pm\frac{1}{\sqrt{LC}}i.$$

So the general solution is

$$I_1(t) = \alpha e^{i\omega t} + \beta e^{i\omega t},$$

where $\omega = 1/\sqrt{LC}$. This is a sinusoidal solution and the frequency is controlled by L and C. This is where the time-keeping part comes in. For example, if we wanted the current in the tank circuit to oscillate once per second, we would only have to choose a capacitor and an inductor with the appropriate values of L and C.

Chapter 9

Growth and Decay

The change in the amount of something is typically governed by the conservation law

$$\text{net change} = \text{amount added} - \text{amount lost},$$

which we can write as

$$\Delta\phi = \Delta\phi_A - \Delta\phi_L.$$

In many systems, for example, with population growth or nuclear decay the rate of change of the quantity ϕ is proportional to the quantity ϕ. That is

$$\frac{d\phi}{dt} \propto \phi.$$

The overall or net growth of ϕ depends on both the growth rate and the decay rate. Let $b > 0$ be the birth/growth rate then

$$\frac{d\phi_A}{dt} = b\phi.$$

Let $d > 0$ be the decay/death rate, so

$$\frac{d\phi_L}{dt} = -d\phi.$$

Overall then,

$$\Delta\phi = \frac{d\phi_A}{dt}\Delta t + \frac{d\phi_L}{dt}\Delta t = b\phi\Delta t - d\phi\Delta t.$$

As $\Delta t \to 0$,

$$\frac{d\phi}{dt} = (b-d)\phi = \mu\phi.$$

The solution to this differential equation is

$$\phi(t) = \phi_0 e^{\mu t},$$

where $\phi_0 = \phi(0)$. Notice that this is exponential growth or decay depending on the sign of μ. The timescale is given by $1/\mu$.

9.1 Filtration

Consider a polluted tank (or even a lake) of volume V. How long will it take to purify the tank to 5% of the original pollutant concentration if we have have some volume of outflow per unit time that goes through a filter? We will assume that the outflow is matched by an inflow of clear water so that the volume of the tank stays constant. We will also assume that the pollutants are well-mixed, that is, uniformly distributed throughout the tank.

The initial concentration of pollutants is

$$C = \frac{\text{mass of pollutant}}{\text{volume of tank}}.$$

The rate of the outflow, which is being filtered, is

$$r = \frac{\text{volume of outflow}}{\text{time}}.$$

By the law of conservation of mass, we know that the net change in pollutant is equal to the amount of pollutant added minus the amount of pollutant removed. The amount of pollutant in the tank at any time is the concentration of pollutant times the volume of the tank, CV. The net change in pollutant over time Δt is

$$\frac{d(CV)}{dt} = 0 - rC\Delta t.$$

The first term is 0 since no more pollutant is being added to the tank. The second term is the outflow rate times the concentration times the length of time, which gives the rate at which the pollutant is being removed from the tank. We know that the volume of the tank V is constant, so this simplifies to the differential equation

$$\frac{dC}{dt} = -\frac{r}{V}C,$$

which has the solution

$$C(t) = C_0 e^{-\frac{r}{V}t},$$

where $C_0 = C(0)$ is the initial concentration of pollutant. The associated timescale is r/V.

Our goal is to get to 5% of the original pollutant concentration, so we solve

$$\begin{aligned} 0.05C_0 &= C_0 e^{-\frac{r}{V}t} \\ \ln(0.05) &= -\frac{r}{V}t \\ t &= -\frac{V}{r}\ln(0.05) \approx \frac{3V}{r}. \end{aligned}$$

So it will takes us $t = 3V/r$ to reach the desired level of purity.

9.2 Logistic Growth

Logistic growth, or growth with finite resources, is a model relevant to a lot of different phenomena such as

- The population of bacteria in a petri dish where the available space is finite.
- The population of termites in an abandoned house where their food is finite.
- The products in a chemical reaction where the reactants are finite.

Here we develop the logistic growth model by considering a plant growing in a pot of pure nutrients (i.e. soil). The available nutrients in the pot are finite and we will assume that over time, all of the nutrients are converted into the plant. That is, the plant grows until there is no more soil remaining.

We let the total weight of the plant and the soil be a constant W, the dry weight of the plant be $w_p(t)$ and the weight of the soil be $w_s(t)$, then

$$w_s(t) = W - w_p(t).$$

Assume that the rate of change of the weight of the plant is proportional to the size of the plant and the amount of soil available. In other words, the rate of change of the weight of the plant is proportional to the weight of the plant where the proportionality constant depends on the weight of soil left. So our differential equation is

$$\frac{dw_p}{dt} = kw_p(t)w_s(t).$$

Substituting the first equation into the differential equation gives us

$$\frac{dw_p}{dt} = kw_p(t)[W - w_p(t)] = kWw_p - kw_p^2.$$

Separating variables and solving the differential equation gives us

$$\begin{aligned}
k\,dt &= \frac{1}{w_p(W - w_p)}dw_p \\
Wk\,dt &= \left(\frac{1}{w_p} + \frac{1}{W - w_p}\right) dw_p \\
Wk\,t &= \int \frac{1}{w_p}\,dw_p + \int \frac{1}{W - w_p}\,dw_p \\
Wk\,t &= \ln(w_p) - \ln(W - w_p) + C \\
Wk\,t &= \ln\left(\frac{w_p}{W - w_p}\right) + C \\
e^{Wkt} &= \frac{w_p}{W - w_p}e^C \\
w_p(t) &= \frac{We^{-C}e^{Wkt}}{1 + e^{-C}e^{Wkt}}.
\end{aligned}$$

In the first step, we expanded the term on the right using partial fraction decomposition.

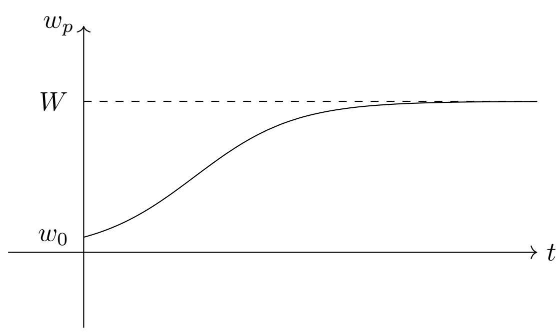
If we let $w_p(0) = w_0$, then we calculate that

$$e^{-C} = \frac{w_0}{W - w_0}.$$

Plugging this in and simplifying gives us the logistic growth model

$$w_p(t) = \frac{Ww_0}{w_0 + (W - w_0)e^{-Wkt}}.$$

Plotting the weight of the plant as a function of time gives us its logistic growth curve as shown in the graph below.



9.3 Competing Species

Refer to the predator-prey model that we did earlier in this chapter.

9.4 The SIR Epidemic Model

A simple but important mathematical model of disease spread is the **SIR model**.

In a population of N people, we have

$$S(t) = \text{number of susceptible people}$$
$$I(t) = \text{number of infected people}$$
$$R(t) = \text{number of removed people.}$$

We assume at any given time that all N people fall into one the three categories above, and that N stays constant. That is, there are no births or unrelated deaths during the epidemic. The susceptible people are all the non-infected people that have not had the disease. The removed people are all those who had the disease and died or recovered with immunity. That is, the removed people are not infectious and cannot become infected again.

Note that in the beginning (before the disease has occurred)

$$S(0) = N$$
$$I(0) = 0$$
$$R(0) = 0.$$

$R(t)$ depends on the number of infected people, and how quickly infected people die or recover. The rate of change of $R(t)$ is the rate at which infected people are removed. We can presume that the rate of change of R is proportional to the number of infected. That is,

$$\frac{dR}{dt} = \alpha I.$$

This makes sense when we consider the extreme $I = 0$. When there are no infected people around then the number of removed people does not change.

$S(t)$ depends on how many infected there are. The more infected people around, the faster that susceptible people will become infected. So we can presume that the rate of change of S is proportional to I. Also, we have to remember that S can only decrease since it starts at $S(0) = N$ and once infected or removed, individuals cannot become susceptible again. However, the rate at which susceptible people become infected is also proportional to the number of susceptible people around since the more susceptible people around, the more likely that infected people encounter susceptible people. So we have that

$$\frac{dS}{dt} = -\beta IS.$$

This makes sense when we consider the extremes $S = 0$ or $I = 0$. When there are zero infected people around or zero susceptible people around, then the number of susceptibles does not change.

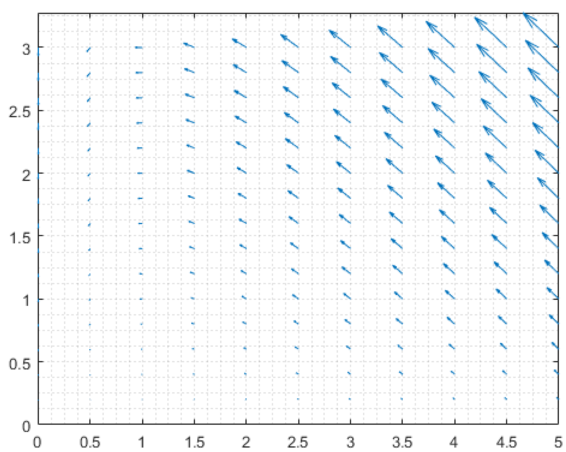
$I(t)$ depends on the number of susceptible people and how quickly susceptible people become infected people as well as how quickly those who are infected are removed. The net change in the number infected over some time period is the number of susceptible who become infected in that time period minus the number of infected who are removed in that time period

$$\frac{dI}{dt} = -\frac{dS}{dt} - \frac{dR}{dt} = \beta IS - \alpha I = I(\beta S - \alpha).$$

This equation can also be obtained by differentiating $N = S + I + R$ with respect to time. This result makes sense when we consider the extremes $S = 0$ and $I = 0$. If there are no susceptibles, then the rate of change of the infected depends only on the ones still infected. If there are no infecteds ($I = 0$), then the rate of change of infecteds is zero since nobody new can get infected if there are no infectious people around.

By examining the original differential equations for I and S, we can deduce the general behavior of the phase plane. First of all, we know that the phase plane will only contain the first quadrant since it does not make sense to have $I < 0$ or $S < 0$. We will plot I on the vertical axis and S on the horizontal axis. By the $\dot{I}$ ODE, we see that when $S = 0$, $\dot{I} < 0$, so along the vertical axis, the arrows will be pointing downward. When $I = 0$, we see from both ODEs that $\dot{S} = \dot{I} = 0$, so everywhere along the horizontal axis is stable (i.e. no arrows). When $I > 0$ and $S > 0$, we see that $\dot{S} < 0$, so all arrows will have a leftward component. By the $\dot{I}$ equation, we see that when $S = \alpha/\beta$, $\dot{I} = 0$, that is, the arrows are horizontal. When $S > \alpha/\beta$ then $\dot{I} > 0$, that is, the arrows have an upward component. When $S < \alpha/\beta$ then $\dot{I} < 0$, that is, the arrows have a downward component.

Below is a graph of the phase plane with $\alpha = \beta = 1$.



The Matlab code used was:

```
[x,y] = meshgrid(0:0.5:5,0:0.2:3);
a = 1;
b = 1;
f = -b.*x.*y;
g = y.*(b.*x-a);
quiver(x,y,f,g)
grid minor
axis tight
```

We can think of β as the infection parameter and α as the recovery/death parameter. High values mean these occur at a high rate.

Dividing the S equation by the I equation and simplifying gives us

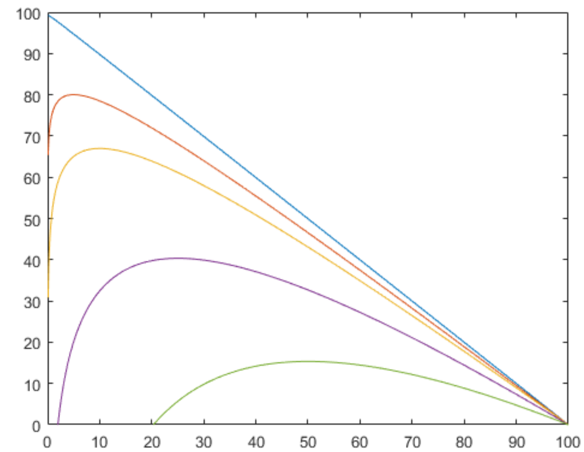
$$\frac{dS}{dI} = \frac{\beta S}{\alpha - \beta S}$$

$$dI = \frac{\alpha - \beta S}{\beta S} dS$$

$$I(S) - I_0 = \frac{\alpha}{\beta} \int_{S_0}^{S} \frac{1}{S'}\, dS' - \int_{S_0}^{S} dS'$$

$$I(S) = \frac{\alpha}{\beta} \ln\left(\frac{S}{S_0}\right) - S + I_0 + S_0.$$

Plotting some examples with different values of α/β gives us the following picture:



TheMatlab code used was:

```
N = 100;
x = 0:0.1:N;
p1 = 0.1.*log(x./N) - x + N;
p2 = 5.*log(x./N) - x + N;
p3 = 10.*log(x./N) - x + N;
p4 = 25.*log(x./N) - x + N;
p5 = 50.*log(x./N) - x + N;
plot(x,p1)
hold on
plot(x,p2)
hold on
plot(x,p3)
hold on
plot(x,p4)
hold on
plot(x,p5)
ylim([0,100]);
```

The top line corresponds to all the susceptible people getting infected before any of the infected are removed. Basically, everyone gets infected and nobody recovers or dies. The bottom line corresponds to the disease running out of steam before all the susceptible people were even infected. The middle lines correspond to essentially everyone getting infected at some point (since the line ends at $S = 0$), however, some of the infected are removed before others are infected. That is, not everyone is infected at once.

Chapter 10

Traffic Flow

10.1 Traffic Speed vs. Congestion

How is traffic speed related to how crowded the road is?

Here, we consider a one-lane road so that all the cars are moving single file. We denote the position of the front car by x_1, the next by x_2, and so on. The position of the ith car is then x_i. So we have the following relationships:

1. Assume the spacing $x_{i-1} - x_i$ is small enough that the car at x_{i-1} affects the car immediately behind it at x_i. That is, the highway is crowded enough that the behavior of the car in front affects the behavior of the car behind.
2. Braking
   - Spacing: How does the spacing between the cars affect the braking behavior of the cars? If the spacing is large then there is less braking, and if the spacing is small, there will be harder braking. We can say that the braking force is inverse proportional to the spacing.

     $$\text{braking} \propto \frac{1}{\text{spacing}}.$$

   - Speed difference: If the car in front decelerates slowly, the car behind brakes softly. If the car in front decelerates quickly, the car behind brakes hard. So the braking force is directly proportional to the difference in speeds

     $$\text{braking} \propto \text{difference in speed}.$$

     In other words, the braking force is proportional to the speed at which the spacing closes.

The simplest model of the braking force is

$$F = A\frac{\dot{x}_{i-1} - \dot{x}_i}{x_{i-1} - x_i}.$$

Note that $\dot{x}_i$ is the speed of the ith car and $\dot{x}_{i-1}$ is the speed of the car immediately ahead of it. So the numerator gives the difference in speed, and the denominator gives the spacing between the two cars. This model assumes that the braking force of the ith car directly proportional to the speed difference between it and the car in front and inversely proportional to the distance between them.

The braking force is the accelerating or decelerating force applied to the car—not the force that the driver applies to the brake pedal. It is given by

$$F_i = -m\ddot{x}_i,$$

where m is the mass of the ith car. Note that negative braking force is the same as accelerating.

However, we cannot just equate these two to get a model because it doesn't include reaction time. If we include reaction time, we get

$$\ddot{x}_i(t+\tau) = \lambda \frac{\dot{x}_{i-1}(t) - \dot{x}_i(t)}{x_{i-1}(t) - x_i(t)},$$

where τ is your reaction time and $\lambda = A/m$. This is known as a **delay differential equation**.

Notice that the right side is the result of differentiating a natural logarithm, so we can also write the equation as

$$\ddot{x}_i(t+\tau) = \lambda \frac{d}{dt} \ln(x_{i-1} - x_i).$$

We can integrate both sides of this equation with respect to time even though they are at different times.

$$\dot{x}_i(t+\tau) = \lambda \ln(x_{i-1} - x_i) + C.$$

Where the left side is now related to the speed of the ith car and the right side is related to the distance between it and the car in front of it. We cannot solve this differential equation for the spacing or the position.

We have that the speed of a car $\dot{x}_i$ depends on the traffic density which is related to $x_{i-1} - x_i$, the spacing between the cars. Let $\rho(x,t)$ be the traffic density which has units of number of cars per unit distance, then $v(\rho)$ is the speed of the traffic. That is, traffic speed is a function of how congested the road is.

Note that $v'(\rho) \leq 0$. That is, as the density increases, the traffic speed drops. Also, $v(0) = v_{max} =$ the speed limit. That is, when the density is zero (there are no other cars around), the traffic speed is at its maximum—the speed limit. In fact, $v(\rho) = v_{max}$ when $\rho \leq \rho_c$ where ρ_c is some critical density at which point the traffic speed begins to be affected. Finally, we know that $v(\rho_{max}) = 0$. That is, traffic stops moving when the density is at maximum. This corresponds to a complete traffic jam.

We will now make the simplifying assumptions that

1. All cars are moving at the same speed $\dot{x}_i = v(\rho)$
2. All cars have the same length L
3. All cars maintain the same distance d from each other

From second and third assumptions, we get $x_{i-1} - x_i = L + d$. Since the density of cars is the inverse of the spacing, we have that

$$\rho = \frac{1}{L+d}.$$

Our simplified ODE becomes

$$v(\rho) = \lambda \ln\left(\frac{1}{\rho}\right) + C, \qquad \text{for } \rho_c \leq \rho \leq \rho_{max}.$$

Since $v(\rho_{max}) = 0$, we get that $C = -\lambda \ln(1/\rho_{max})$. Since $v(\rho_c) = v_{max}$, we get that $\lambda = v_{max}/\ln(\rho_{max}/\rho_c)$, so our ODE simplifies to

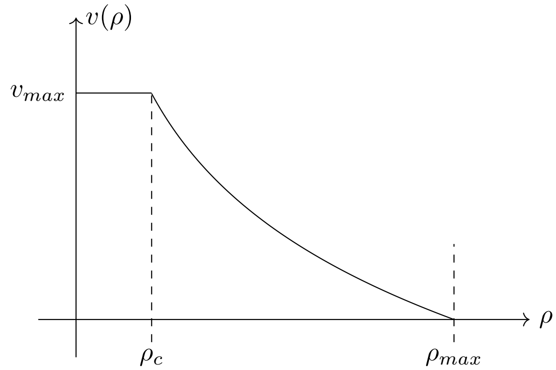
$$v(\rho) = v_{max} \frac{\ln(\rho_{max}/\rho)}{\ln(\rho_{max}/\rho_c}),$$

for $\rho_c \leq \rho \leq \rho_{max}$.

Overall, our function of traffic speed versus density is given by

$$v(\rho) = \begin{cases} v_{max}, & \rho \leq \rho_c \\ v_{max} \frac{\ln\left(\frac{\rho_{max}}{\rho}\right)}{\ln\left(\frac{\rho_{max}}{\rho_c}\right)}, & \rho_c \leq \rho \leq \rho_{max}. \end{cases}.$$

A plot of $v(\rho)$ is shown here:



10.2 Traffic Flow on a Bridge

How should you set the speed limit in a tunnel or on a bridge? This is equivalent to a road with no entrances or exits.

In this case, the cars are conserved. They satisfy a conservation law since every car that gets on the bridge must get off the bridge at some later time. Cars don't just appear in the middle of the bridge or vanish from them.

If one end of the bridge is at x_a and the other end is at x_b, then the total number of cars on the bridge can be obtained by integrating the density of the cars on the bridge over the length of the bridge

$$\int_{x_a}^{x_b} \rho(x,t)\,dx.$$

Since the cars on the bridge are conserved, the time rate of change of the number of cars on the bridge is equal to the number of cars coming onto the bridge per unit time (i.e. the flux in) minus the number of cars getting off the bridge per unit time (i.e. the flux out). So we have that

$$\frac{d}{dt}\int_{x_a}^{x_b} \rho(x,t)\,dx = F_{in} - F_{out}.$$

Since the flux is a function of the density, we can write

$$\frac{d}{dt}\int_{x_a}^{x_b} \rho(x,t)\,dx = F(\rho(x_a,t)) - F(\rho(x_b,t)).$$

However, by the fundamental theorem of calculus, we have that

$$\int_{x_a}^{x_b} F_x(\rho(x,t))\,dx = F(\rho(x_b,t)) - F(\rho(x_a,t)).$$

So we have that

$$\int_{x_a}^{x_b} \rho_t\,dx = -\int_{x_a}^{x_b} F_x(\rho)\,dx,$$

which we can write as

$$\int_{x_a}^{x_b} (\rho_t + F_x)\,dx = 0.$$

This implies that

$$\rho_t + F_x = 0.$$

This is now a PDE that describes traffic density.

Note that flux has units of number of cars per time, or alternatively,

$$[F] = \frac{\text{number of cars}}{\text{length}} \cdot \frac{\text{length}}{\text{time}}.$$

This is a density times a speed, so

$$F = \rho\, v(\rho).$$

By the chain rule,

$$\frac{\partial F(\rho)}{\partial x} = \frac{\partial F}{\partial \rho}\frac{\partial \rho}{\partial x},$$

so we can write our PDE as

$$\frac{\partial \rho}{\partial t} + \frac{\partial F}{\partial \rho}\frac{\partial \rho}{\partial x} = 0.$$

Since F is a function of ρ, the partial derivative $\frac{\partial F}{\partial \rho}$ is some unknown function c of ρ, so we can write the PDE as

$$\rho_t + c(\rho)\,\rho_x = 0.$$

This PDE is an **advection equation**. Its solution has the form

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

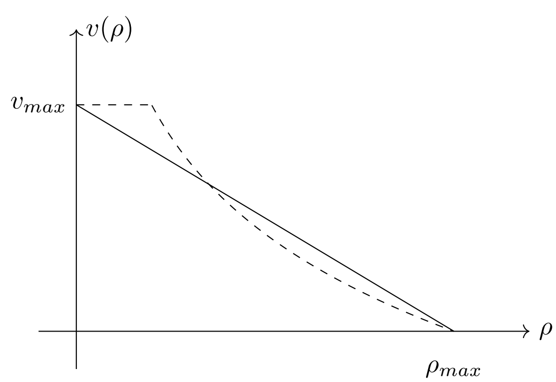
which describes density waves. The wave speed is given by

$$c(\rho) = \frac{\partial F}{\partial \rho} = \frac{\partial}{\partial \rho}(\rho v) = v + \rho v' \neq v.$$

Notice that in general, the speed that the density waves travel through traffic is not the same as the speed v of the cars.

10.3 Whitham Traffic Flow Model

A simple approximation to the traffic model that we derived in the first subsection of this section is given by Whitham. His model assumes that $v(\rho)$ is linear. That is, a plot of $v(\rho)$ with his model looks like:



The plot we derived earlier is given by the dotted curve. From here on (including the next subsections) we will work with Whitham's simpler model.

Whitham's plot is given by

$$v(\rho) = v_{max}\left(1 - \frac{\rho}{\rho_{max}}\right).$$

This implies that the flux is

$$F = \rho\left(1 - \frac{\rho}{\rho_{max}}\right),$$

which gives us the density wave speed

$$\frac{\partial F}{\partial \rho} = c(\rho) = 1 - \frac{2\rho}{\rho_{max}}.$$

This shows that the density wave can move forward and backward. If the traffic density is high then the density wave propagates backward.

10.4 Solution to the Traffic Flow Equation

Recall that our PDE is

$$\rho_t + c\rho_x = 0,$$

where c is a function of ρ. We will assume initial conditions of the form $\rho(x, 0) = \rho_0(x)$. We can solve this PDE using the **method of characteristics**.

Using the method of characteristics, our PDE becomes the system of ODEs

$$\dot{x} = c, \qquad \dot{t} = 1, \qquad \dot{\rho} = 0,$$

where the dot notation means a derivative with respect to the parameter s.

We parametrize the initial conditions curve with the new variable a, so that when $s = 0$, we get

$$x(0) = a, \qquad t(0) = 0, \qquad \rho(0) = \rho_0(a).$$

Solving the ODEs, using the initial conditions gives us

$$\begin{aligned} x(s) &= cs + a \\ t(s) &= s \\ \rho(s) &= \rho_0(a). \end{aligned}$$

To get the solution in the form of x and t, we invert these equations. The second one tells us that $s = t$, then the first one tells us that $a = x - ct$, so our solution in terms of x and t is

$$\rho(x, t) = \rho_0(x - c(\rho)t).$$

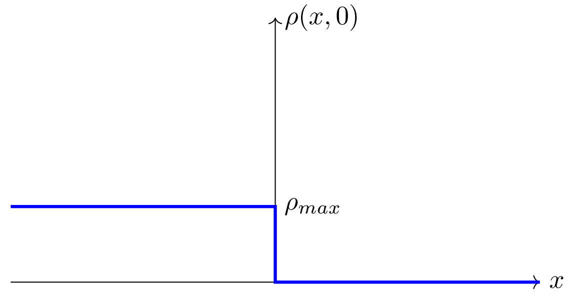
The equation $a = x - ct$, which can also be written as $x = ct + a$ allows us to plot the **characteristic curves** of the solutions.

10.5 Example

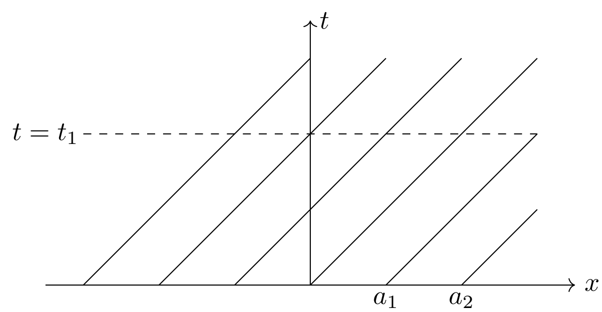
If c is a constant, that is, $c(\rho) = c$, then the solution to our traffic flow PDE is

$$\rho(x, t) = \rho_0(x - ct).$$

Consider a traffic light on a road on which the traffic moves from negative x toward positive x (i.e. from left to right on the graph below). If the light is red at $t = 0$, then the density to the left of the light is at ρ_{max} and the density to the right of the light is 0. This situation is depicted in the graph below.



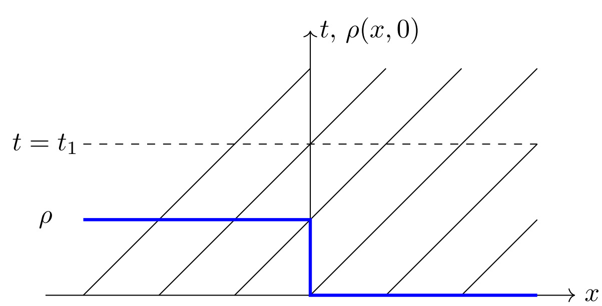
Recall that the relationship between x and t, is then $x = ct + a$, where c is a constant. We can use this to plot several characteristic curves shown below.



Notice that the slopes of the characteristic curves are $1/c$ since we have x on the horizontal axis and t on the vertical axis, and that we have a different characteristic curve for each value of a. Along any of the characteristic curves, the value of $\rho(x,t)$ is a constant, although this value is not necessarily the same for different curves.

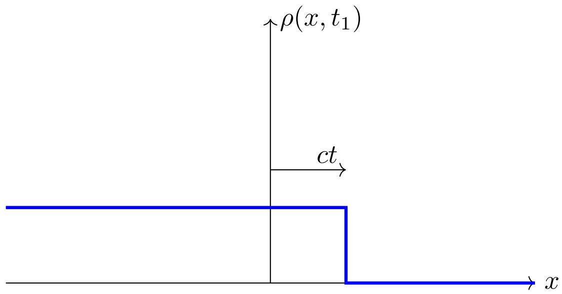
Now, we want to know what the traffic density is at some later time t_1 shortly after the light has turned green. To do this, we look at the dashed line $t = t_1$ in the plot above. Since the value of ρ is constant on any given characteristic curve, we can move along the line $t = t_1$, then at any intersection with a characteristic curve, we can follow that characteristic curve to the corresponding value of ρ. For example, if we look at the rightmost intersection of $t = t_1$ with a characteristic curve, we see that it intersects with the characteristic curve labeled a_1. Since we know that the value of ρ is constant along a characteristic curve, we trace this characteristic curve back to $t = 0$ where it has the positive x-value of a_1. But from the graph of ρ, we know that $\rho = 0$ for positive values of x at $t = 0$, so the value of ρ at $t = t_1$ is also zero. Doing this for all these intersections, gives us the plot of ρ versus x at $t = t_1$.

This process might be easier to understand if we overlay the $\rho(x,0)$ plot on top of the characteristic curves.



Plotting the new graph of ρ at $t = t_1$, obtained via the method described above, gives us the picture shown below. Notice that the pair of graphs for ρ at times 0 and t_1

shows that we have a density wave traveling to the right with speed c. At time t_1, the cars that were stopped at the light at $t = 0$ are now a distance of ct further to the right. This simple model assumes that when the light turns green the cars accelerate to their maximum speed instantaneously.



10.6 Example

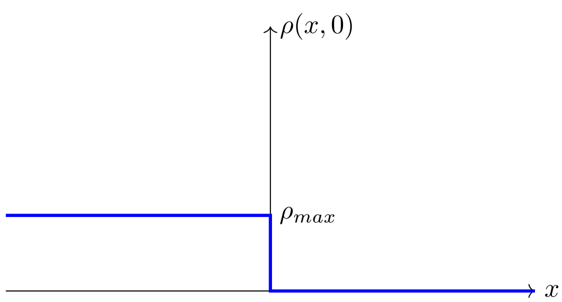
If $c(\rho) = \rho$, then our traffic flow PDE is

$$\rho_t + \rho\,\rho_x = 0.$$

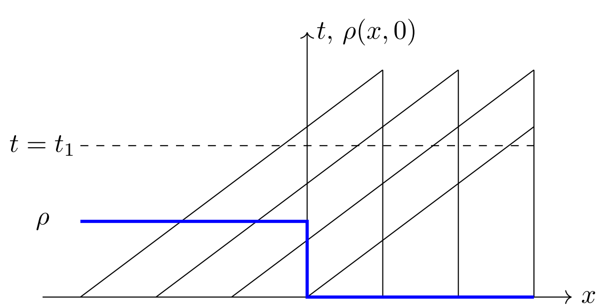
This is known as the inviscid **Burgers' equation**. Our solution is then of the form

$$\rho(x, t) = \rho_0(x - \rho t).$$

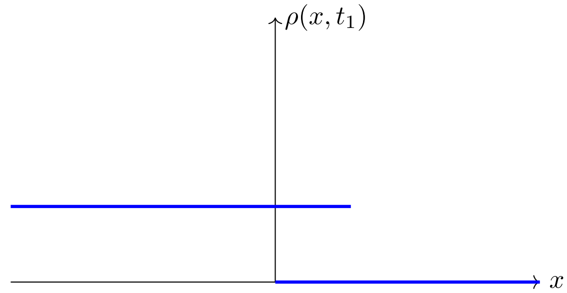
The relationship between x and t is now $x = \rho t + a$. Plugging this into the solution gives us $\rho = \rho_0(a)$, so $x = \rho_0(a)t + a$. If we consider again the traffic light situation, then ρ_0 is shown here:



To draw the characteristic curves, we use $x = \rho_0(a)t + a$, noting that they depend on the value of ρ_0 shown above. On the left of $x = 0$ we have that $\rho_0 = \rho_{max}$, so the characteristic curves are given by $x = \rho_{max}t + a$ which are lines with slope $1/\rho_{max}$. On the right of $x = 0$ we have that $\rho_0 = 0$, so the characteristic curves are given by $x = a$ which are vertical lines. We get a plot that looks like:

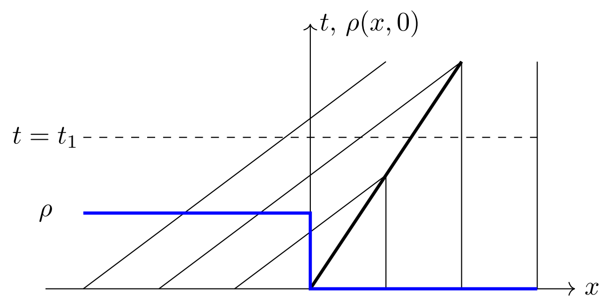


Plotting $\rho(x, t_1)$ using the same method described in the previous example now gives us a double valued function for some domain of x.

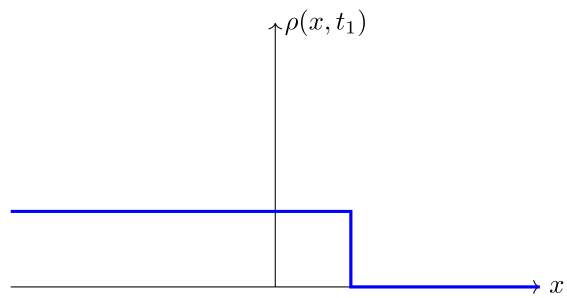


To get rid of the double-valued function, we introduce a **shock wave**. A shock wave is a propagating wave where the density and speed change abruptly.

We introduce a shock wave with slope $1/v_s$, where v_s is the speed of the shock wave, onto the characteristic plane with the thick line in the plot below.



Now, $\rho(x, t_1)$ is no longer double valued.

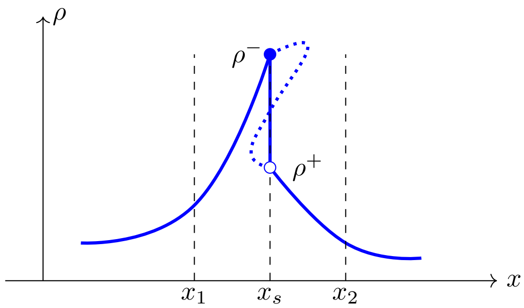


10.7 Shock waves and Expansion Fans

Whenever we have intersecting characteristic curves on our characteristic plane, we have to put in a shock wave so that our function remains single-valued.

In the case of traffic flow, we position the shock wave so that ρ, the traffic density, remains conserved.

We need to introduce a shock wave so that ρ is not multi-valued (thick dotted blue line), and is instead single-valued (solid blue line). The graph below shows an arbitrary shock wave. Notice that it is discontinuous—even the single-valued version. The value x_s is the location of the shock wave, but remember that the shock wave is moving in time, so this is more completely represented as a function of time, $x_s(t)$. The values ρ^- and ρ^+ are the values of ρ as it approaches the shock discontinuity from the left and right respectively.



Recall that the original form of the equation relating the density and the flux is

$$\frac{d}{dt}\int_{x_a}^{x_b} \rho\, dx + F(x_b) - F(x_a) = 0.$$

Picking a spot x_1 to the left and x_2 to the right of the discontinuity, we have that

$$\int_{x_1}^{x_2} \rho\, dx = \int_{x_1}^{x_s(t)} \rho\, dx + \int_{x_s(t)}^{x_2} \rho\, dx.$$

Differentiating this with respect to time gives us

$$\frac{d}{dt}\int_{x_1}^{x_2} \rho\, dx = (\rho^- - \rho^+)\frac{dx_s}{dt} + \int_{x_1}^{x_s} \rho_t\, dx + \int_{x_s}^{x_2} \rho_t\, dx.$$

Notice that $\rho^- - \rho^+$ is the height of the jump discontinuity. Remember that $\rho_t = -F_x$, so this simplifies to

$$\begin{aligned}\frac{d}{dt}\int_{x_1}^{x_2} \rho\, dx &= \frac{dx_s}{dt}(\rho^- - \rho^+) - \big(F(x_s^-) - F(x_1)\big) - \big(F(x_2) - F(x_s^+)\big)\\ &= \frac{dx_s}{dt}(\rho^- - \rho^+) + F(x_s^+) - F(x_s^-).\end{aligned}$$

So the speed of the shock wave is

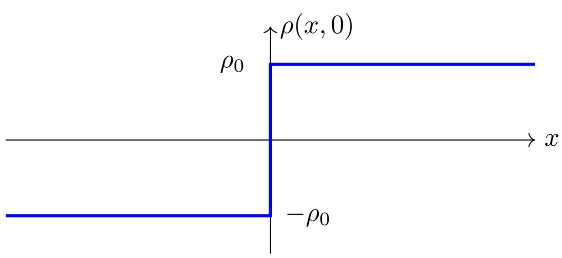
$$\frac{dx_s}{dt} = \frac{F(x_s^+) - F(x_s^-)}{\rho^+ - \rho^-}.$$

For example, if we have $c(\rho) = \rho$, and the traffic light case of the previous example, then the shock wave speed is

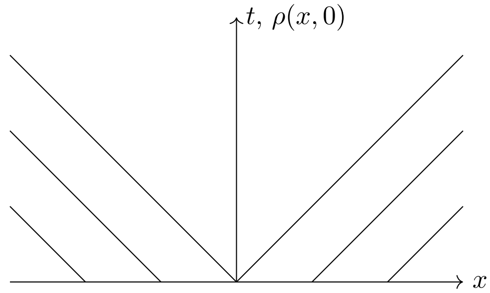
$$\frac{dx_s}{dt} = v_s = \frac{0 - \frac{1}{2}\rho_0^2}{\rho_0} = \frac{1}{2}\rho_0,$$

where $\rho_0 = \rho_{max}$. So the shock wave speed is half of the wave speed, and the slope of the shock wave on the characteristic plane is $1/v_s = 2/\rho_0$.

What if, instead of the red light traffic situation, we had $c = \rho$, but at $t = 0$, the density was $-\rho_0$ to the left of $x = 0$, and ρ_0 to the right? Then the graph of $\rho(x, 0)$ would look like:



Then our characteristic curves would look like:

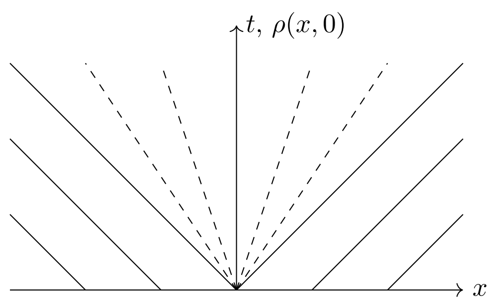


From the characteristic plane, we know that characteristic curves at the bottom left have slope $-1/\rho_0$, and on the bottom right they have slope $1/\rho_0$, but what about the empty V-shaped region in the middle? We know nothing about the characteristic curves in that region.

Recall that $x - ct = a$, so $c = (x-a)/t$. We claim that

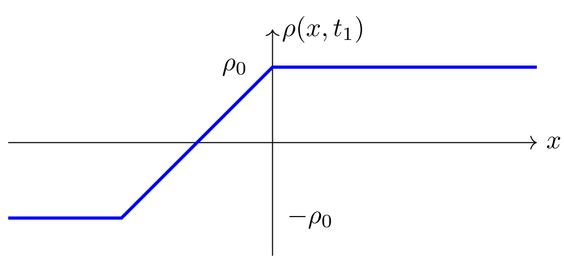
$$\rho = \alpha \frac{x-a}{t} + \beta,$$

for all α and β. This can be verified by plugging it into the PDE and simplifying. This tells us that for fixed t, ρ is linear in x. In other words, we get the following kind of graph, where the dashed lines show the previously missing characteristic curves.



These fan-like characteristics are called an **expansion fan**.

At some later time t_1, the density as a function of position looks like this:



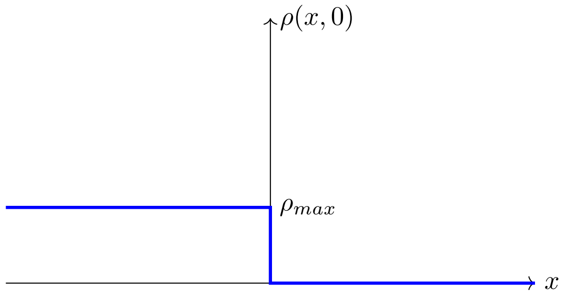
This shows an **expansion wave** moving to the left with speed ρ_0.

10.8 Example

From Whitham's traffic flow model, we have

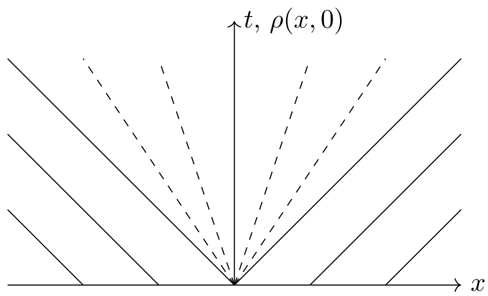
$$c(\rho) = v_{max}\left(1 - \frac{2\rho}{\rho_{max}}\right).$$

If you're sitting at the traffic light at $t = 0$, then the traffic density as a function of position is given by the following graph:

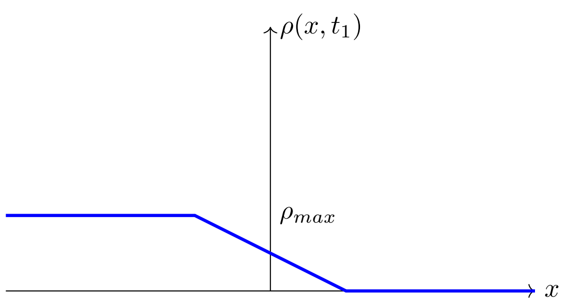


That is, when the light is red, the density of the cars on your side of the light is ρ_{max}. On the other side of the light, the density is 0. If you have some number of cars sitting in front of you, how soon after the light turns green can you start moving?

The characteristic curves look are shown in the following graph. Near the bottom left, the characteristic curves have slope $-1/v_{max}$. On the bottom right, the characteristic curves have slopes $1/v_{max}$. In the middle, there is an expansion fan.



At a later time t_1, shortly after the light turns green, the density as a function of position looks like the following graph:



The length of the sloped segment is $2v_{max}t_1$. This graph depicts a density wave propagating forward and an expansion wave propagating backward. Assuming that a car reaches v_{max} instantaneously, at time t_1 the car that was first in line is now $v_{max}t_1$ to the right of $x = 0$ (i.e. the light). As more and more of the cars near the light begin moving, the density just to left of the light starts dropping as is shown in the graph above. You begin to move when this expansion wave reaches you, and since the expansion wave is moving left from $x = 0$ at the same speed as the density wave is moving to the right of $x = 0$, we can conclude that you start to move when the first car from the other side of the light reaches your position. This very simple model does not account for the width of the intersection.

Index

Advection, 11, 78
Advection equation, 11, 12
Advection-diffusion equation, 10

Bandpass Filter, 64
Buckingham π theorem, 13
Burgers' equation, 81

Center, critical point, 25
Characteristic curve, 79
Characteristic equation, 40
Characteristic value, 18
Coefficient of transmission, 52
Critical point, 22

Delay differential equation, 76
Diffusion, 12
Diffusion equation, 12
Dynamical systems, 22

Equilibrium, 28
Equilibrium point, 22
Expansion fan, 82, 84
Expansion wave, 84

Forced solution, 41
Form drag, 16
Fourier's law, 52

High-pass Filter, 63
Homogeneous solution, 40

Lanchester combat model, 33
Linearization, 29
Logistic growth, 70
Low-pass filter, 61

Method of characteristics, 79

Node, 25
Nondimensionalization, 8

Particular solution, 41
Peclet number, 11
Phase diagram, 22
Predator-prey model, 36
Proportion, 4

R-value, 52
Reynold's number, 15

Saddle, critical point, 26
Shock wave, 82
Spiral, 27
Stability, 22, 28
Stoke's drag, 16

Tank circuit, 66
Thermal conductivity, 52
Thermal diffusivity, 53
Thermal resistance, 52
Traffic flow, 75
Transfer function, 41

Voltage gain, 61

Whitham traffic flow model, 84